"From Problems to Numbers and Back"

Lecture Notes to 'A Discipline-neutral Introduction to Mathematical Modeling'

Kees van Overveld, Emiel van Berkum, and Tijn Borghuis[1]

version: July 2014

Contents

1 No model without a Purpose **11**

1.1 Models that Everybody Knows 11
1.2 Various Kinds of Modeling Purposes 13
1.2.1 Purposes of Modeling in the Context of Research 13
1.2.2 Purposes of Modeling in the Context of Design 16
1.3 Modeling Approaches 18
1.3.1 Material - Immaterial 18
1.3.2 Static - Dynamic 21
1.3.3 Continuous - Discrete 21
1.3.4 Numerical - Symbolic 23
1.3.5 Geometric - Non-geometric 23
1.3.6 Deterministic - Stochastic 25
1.3.7 Calculating - Reasoning 27
1.3.8 Black Box - Glass Box 28
1.4 Stages in The Modeling Process 29
1.4.1 Definition Stage (1) 29
1.4.2 Conceptualization Stage (2) 30
1.4.3 Formalization Stage (3) 31
1.4.4 Execution Stage (4) 32
1.4.5 Conclusion Stage (5) 33
1.4.6 Learn from what you have done: Evaluation for the Modeler 34
1.5 Reflections: Iterating the Modeling Process 35
1.5.1 Reflection after the Definition Stage (1') 35
1.5.2 Reflection after the Conceptualization Stage (2') 35
1.5.3 Reflection after the Formalization Stage (3') 36
1.5.4 Reflection after the Execution Stage (4') 36
1.5.5 Reflection after the Conclusion Stage (5') 37
1.6 Example 39
1.7 Mathematical Tools: Functions of Two Variables 44

1.7.1 Functions of Several Variables 44
1.7.2 Graphs and Contour Plots 46
1.7.3 Partial Derivatives 47
1.7.4 Tangent Planes 49
1.8 Summary 50
1.9 Learning goals 51
1.9.1 Knowledge 51
1.9.2 Skills 51
1.9.3 Attitude 51
1.10 Questions 52
1.11 Exercises 54

2 The Art of Omitting **57**
2.1 The Conceptual Model 57
2.2 Concepts and Entities 58
2.3 Properties 59
2.4 Relations 62
2.5 Constructing a Conceptual Model 63
2.6 Quantities 67
2.6.1 Types of Quantities 68
2.6.2 Operations on Quantities: Ordering 71
2.7 Units, Scales and Dimensions 73
2.7.1 Counting is Easier than Measuring 73
2.7.2 Units and Dimensions 77
2.7.3 Dimensions and Formulas 78
2.8 Mathematics: Functions of two Variables continued 79
2.8.1 Extrema of Functions of several Variables 80
Local Extrema and Critical Points 80
Global Extrema 82
2.8.2 Constrained Optimization 83
2.9 Summary 86
2.10 Learning goals 87
2.10.1 Knowledge 87
2.10.2 Skills 88
2.10.3 Attitude 88
2.11 Questions 88
2.12 Exercises 90

3 Time for Change **93**

3.1 Change needs Time . . . 93
3.2 Introduction to Processes . . . 94
3.2.1 States and State Charts . . . 94
State Space . . . 96
State Space Reduction: Symmetry . . . 98
State Space Reduction: Projection . . . 99
Exposed and Hidden Properties . . . 100
Hierarchy and Orthogonality . . . 101
3.2.2 Applying State Charts . . . 101
3.3 Time and State Transitions . . . 103
3.3.1 Partially Ordered Time . . . 103
3.3.2 Totally Ordered Time . . . 106
Time Lapses . . . 106
Total Ordering, Causality . . . 106
Causality and Functions . . . 107
Evolution of Dynamical Systems . . . 108
3.3.3 Totally Ordered Time; Equal Intervals . . . 109
Equal Intervals: Periodicity . . . 110
Recursion . . . 111
Time Lapses with Equal Length; Sampling . . . 112
3.3.4 Totally Ordered Time; Equal, Infinitesimal Intervals: Continuous Time . . . 117
Evolution of Dynamical Systems and Differential Equations . . . 118
3.4 Mathematical Preliminaries - More about Differential Equations . . . 120
3.4.1 Direction Fields . . . 120
3.4.2 Euler's method . . . 121
3.4.3 Equilibrium Solutions and Stability . . . 124
Material for the exam . . . 125
Exercises . . . 125
3.5 Summary . . . 126
3.6 Learning goals . . . 126
3.6.1 Knowledge . . . 126
3.6.2 Skills . . . 127
3.6.3 Attitude . . . 127
3.7 Questions . . . 127
3.8 Exercises . . . 130

4 The Function of Functions **133**
4.1 What is a Formal Model? . . . 133
4.2 A Practical Route to Formal Models . . . 135

4.2.1 Naming . . . 135
4.2.2 Structure . . . 137
Chain of Dependencies . . . 137
To-do-list . . . 141
4.2.3 Forming Mathematical Expressions . . . 142
Dimensional Analysis . . . 142
Relation Wizard and Function Selector . . . 142
Two models is better than one model . . . 143
Wisdom of the Crowds . . . 143
4.3 Examples . . . 143
4.3.1 Example 1: the Detergent Problem . . . 144
Estimating Accuracy . . . 147
4.3.2 Example 2: The Chimney Sweepers Problem . . . 149
The Consequence of Purpose for the Applicability of a Model . . . 149
The Benefits of a Second Model . . . 153
4.3.3 Example 3: the Peanut Butter Problem . . . 153
The Relation Wizard . . . 154
Function Selector . . . 157
Glass Box Models, Black Box Models . . . 160
Functions of One or More Arguments . . . 162
4.4 Mathematical preliminaries . . . 164
4.5 Summary . . . 164
4.6 Learning goals . . . 165
4.6.1 Knowledge . . . 165
4.6.2 Skills . . . 165
4.6.3 Attitude . . . 166
4.7 Questions . . . 166
4.8 Exercises . . . 168

5 Roles of Quantities in a Functional Model 171
5.1 Functional Models . . . 171
5.2 The 4-Categories Approach to Building Functional Models . . . 173
5.2.1 The Structure of a Functional Model . . . 173
Before Developing a Model: look out for Trivial Cases . . . 175
Case 1: maximize s, minimize W . . . 175
Case 2: minimize W, minimize t . . . 176
Case 3: minimize t, maximize s . . . 177
Analysis . . . 177
5.2.2 Input of the Functional Model: Category I . . . 179

No Dependency-by-Anticipation 180
5.2.3 Output of the Functional Model: Category II 181
Using Cat.-II Quantities: Requirements, Desires or Wishes 182
5.2.4 Limitations from Context: Category III 184
5.2.5 Constructing the Functional Model: Category IV 185
5.3 Executing a Functional Model 187
5.3.1 Optimisation in a Functional Model 187
Lumping and Penalties 188
Searching Cat.-I Space 190
5.3.2 An Alternative to Lumping: Dominance 191
Trade-offs and the Pareto Front 192
Approximating the Pareto Front: Genetic Programming 193
Pareto: Post Processing 197
Last Resorts 198
What about the Optimal Cycling Trip? 199
5.4 Mathematical preliminaries 200
5.5 Summary 200
5.6 Learning goals 200
5.6.1 Knowledge 200
5.6.2 Skills 201
5.6.3 Attitude 201
5.7 Questions 201
5.8 Exercises 204

6 Models and Confidence 207
6.1 What do we mean by 'Confidence' in a Model? 207
6.1.1 Verification and Validation, Accuracy and Precision 209
Verification: is model X correct? 209
Validation: is X the correct model? 211
Verification and Validation: Weak and Strong, Assumptions and Results . 212
6.1.2 Distributions to Indicate Uncertainty 216
Distributions and the Usefulness of a Model Outcome 217
6.2 Distance and Similarity 219
6.3 Confidence in Black Box Models 220
6.3.1 Features from Data Sets and from Distributions 220
Mean 221
Standard Deviation 222
Correlation 223
6.3.2 Example of the Value of a Black Box Model: Linear Least Squares 224

6.3.3 Validating a Black Box Model . . . 225
Residual Error . . . 225
Distinctiveness . . . 226
Common Sense . . . 227
6.4 Confidence in Glass Box Models . . . 228
6.4.1 Structural Validity Assessment . . . 230
6.4.2 Quantitative Validity Assessment . . . 232
Sensitivity Analysis . . . 232
The Contribution of Model Errors to Uncertainty; Model Improvement . . . 235
6.5 Mathematical preliminaries . . . 236
6.6 Summary . . . 236
6.7 Learning goals . . . 237
6.7.1 Knowledge . . . 237
6.7.2 Skills . . . 237
6.7.3 Attitude . . . 237
6.8 Questions . . . 238
6.9 Exercises . . . 240

7 A Working Model - and then? 243
7.1 The Need for Interpretation . . . 243
7.2 Criteria for Modeling . . . 244
7.2.1 Definition Stage, Inside, Qualitative: Genericity . . . 246
7.2.2 Definition Stage, Inside, Quantitative: Scalability . . . 247
7.2.3 Definition Stage, Outside, Qualitative: Specialization . . . 249
7.2.4 Definition Stage, Outside, Quantitative: Audience . . . 253
7.2.5 Conclusion Stage, Inside, Qualitative: Convincingness . . . 254
7.2.6 Conclusion Stage, Inside, Quantitative: Distinctiveness . . . 257
7.2.7 Conclusion Stage, Outside, Qualitative: Surprise . . . 261
7.2.8 Conclusion Stage, Outside, Quantitative: Impact . . . 262
7.3 Criteria for Modeling and Purposes . . . 263
7.4 Mathematical preliminaries . . . 267
7.5 Summary . . . 267
7.6 Learning goals . . . 268
7.6.1 Knowledge . . . 268
7.6.2 Skills . . . 268
7.6.3 Attitude . . . 268
7.7 Questions . . . 269
7.8 Exercises . . . 270
Notes . . . 275

Index **295**

Chapter 1

No model without a Purpose

'Perhaps reality is just a better model'

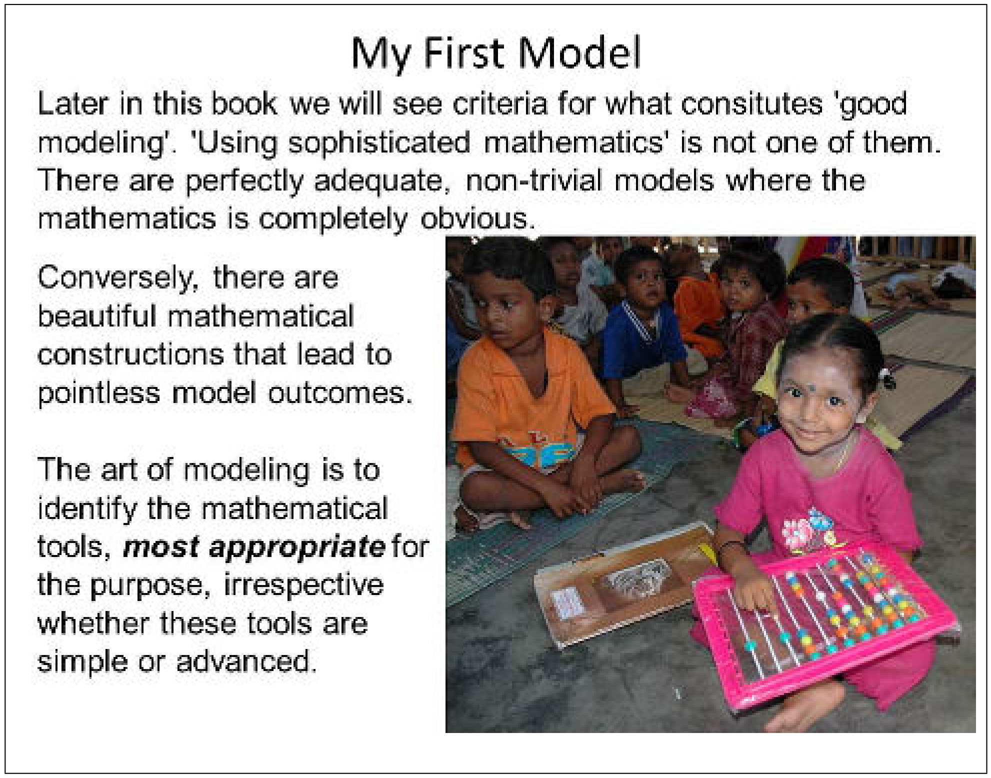
T zero minus two days. After months of planning and preparation, the metamorphosis is complete. Acres of wasteland have been magically transformed into the venue of the next mega pop festival. Five huge stages have been erected, catering and sanitary facilities, sufficient to serve a thirsty and sweaty crowd, are in place; PR machinery has been running at full steam; nearly one hundred thousand tickets have been sold; and an impressive line up of hip rock stars has been booked. And then ... the e-mail arrives. It does not contain the threat of a terrorist attack, neither the announcement of the main attraction's withdrawal. A short paragraph merely contains the result of a computer predicting air movements in the atmosphere. Something about wind speed, a sharp drop of air pressure... Although the sky still looks perfectly innocent, tables with numbers foretell a 95% chance of a major thunderstorm with dangerously strong gusts of wind to pass over the festival terrain, exactly during the opening night. There is no discussion. The director cancels the festival. Sponsors loose their money, ticket holders receive a refund accompanied by a frustrating little e-mail apology, and two or three management bureaus go bankrupt.

1.1 Models that Everybody Knows

The little story above, although fictitious, gives a hint about the potential impact of using MODELS. Let us analyse what is happening.

People look to the sky to know if it will be raining soon. Dark clouds means: rain expected. The computer program, causing the cancelation of the pop festival in the above story, seems to do something else. It uses the equivalent of 'looking at the sky': it takes recent, accurate measurements of temperatures, humidities and air pressures as input, but it does *not* draw its conclusions on mere inspection of the current state. Instead, it runs calculations in an attempt to predict the future.

Running calculations to predict the future in itself is not magic. Suppose you want to know if you can afford to buy, say a 25 Euro book, and have enough money left to buy food later today. You do a number of things:



- *Collect data* : do a 'measurement' to assess the amount of money in your wallet. This measurement could be more or less accurate (do you also count the copper coins?); sometimes you know that a measurement can have a systematic error (if you ignore the copper coins, only counting bank notes and silver coins, you are certain that your measurement is an under estimate);
- *Find out the value of one or more known constants* : in this case, the price of the book: 25 Euro; a meal: at least 5 Euro;
- *Perform a mathematical operation* [1]: in this case, you subtract the known, exact value (25 Euro) from the 'measured' value (the amount of money in your wallet); the result is some number - typically with some uncertainty (in this case, the uncertainty comes from ignoring the copper coins);
- *Interpret the outcome of this* MATHEMATICAL OPERATION (=subtraction, producing a number) in terms of your initial problem (='will I have enough money left for buying food'). This interpretation yields an answer such as 'yes', 'no', 'probably', 'unlikely' - or it can even prompt you to do a new calculation or a new measurement.

All the above steps are trivial - in fact, few people will call this 'applying a MODEL'. It is immediately clear how the future, regarding your options for buying food, will look like. 'Predicting the future' can be as simple as merely doing a single subtraction. It can also be as complicated as running a weather simulator. In both cases, however, we perform the same kinds of steps.

Some models are set up to predict future events, such as weather models and our simple wallet model. Prediction, however, is by far not the only purpose of models. In the next section we will investigate the various purposes models can have.

[1] The photograph of the little girl doing arithmetic on a toy abacus was taken from http://commons.wikimedia.org/wiki/Abacus#mediaviewer/File:Indian_pre-school_girl_in_pink_shirt_plays_with_abacus.jpg

1.2 Various Kinds of Modeling Purposes

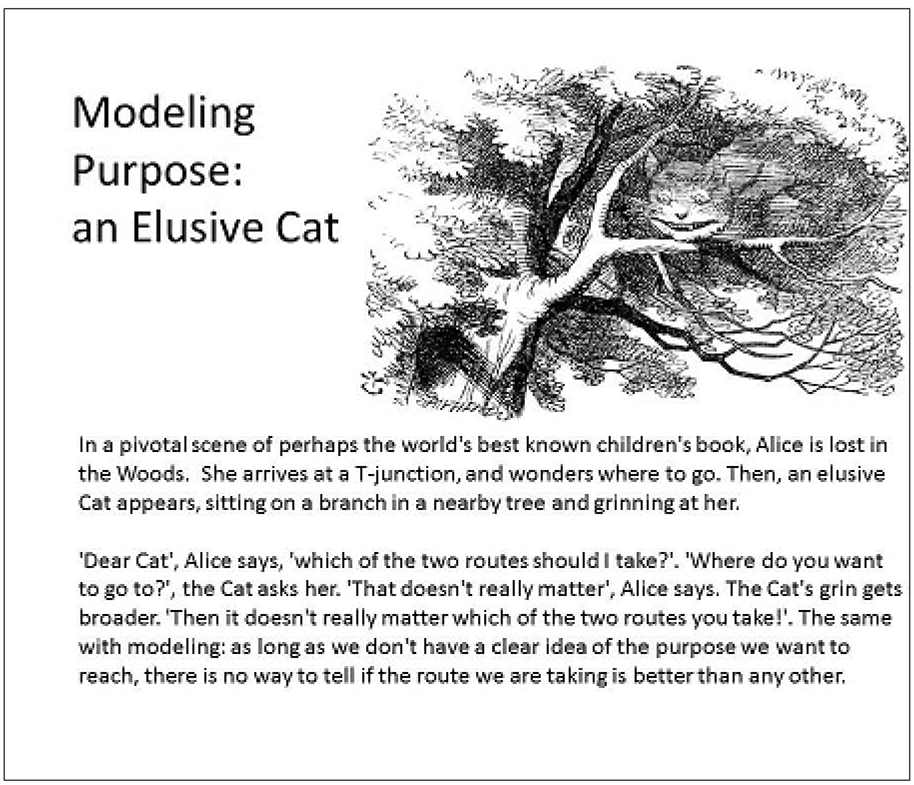
It is important to realize, before developing or using a model what the model should do for you: what is it that the modeler *wants* from the model? This is what we call the MODEL'S PURPOSE. A model, just like medicine, should only be used once the purpose is perfectly clear. Few people would use medicine just to see what will happen: similarly, it is pointless to use a model without clear need. Taking a medicine is a means to an end, namely, to become healthy. Similarly, deploying a model is a means to an end.

For this reason, it is impossible to say that 'a model is good' or 'a model is bad': a model can only be good or bad *with respect to some given purpose*[2].

Purposes can belong to the context of SCIENTIFIC RESEARCH or to the context TECHNOLOGIC DESIGN.

1.2.1 Purposes of Modeling in the Context of Research

Let us consider the example of our Solar System. We might want to know something about the Solar System: the Solar System could be our object of research. We can study the Solar System in various ways. For instance:



1. *A planets is seen as the visual manifestations of an ancient deity.* This explains why Mars is red (the god of war, associating to fire and blood), why Mercury runs so fast (the messenger of the gods), and Saturn so slowly (slowness being one of the archetypical attributes of the god Saturn).
2. *A planet is seen as a celestial bodies* , its trajectory therefore is a system of circles centered round the earth, according to Medieval belief. This is the Ptolemaic view, after Ptolmy (approximately AD 90-168) ▷1 . This model predicts the location of planets on the night sky at arbitrary dates to rather good accuracy - sufficient, for instance, to do astrological predictions, or navigation at sea using the stars.
3. *A planet is seen as a point moving round the sun in ellipses.* Johannes Keppler (1571-1630), by observing and analyzing the accurate measurements of Tycho

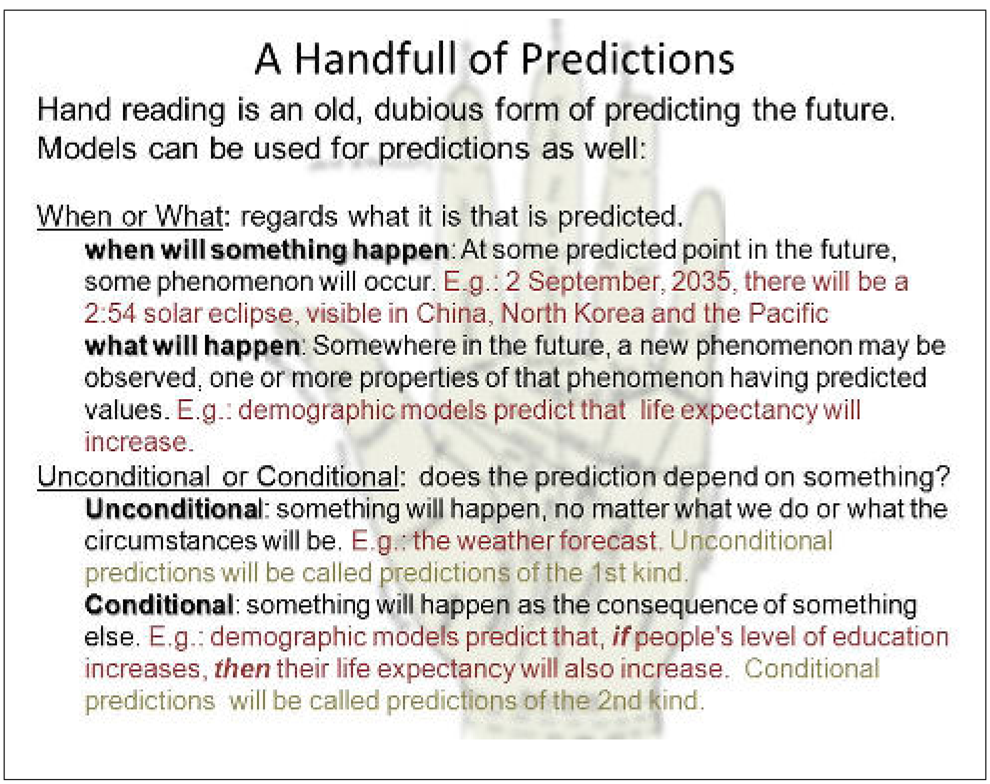
[2] Image 'Alice in Wonderland' from http://commons.wikimedia.org/wiki/John_Tenniel#mediaviewer/File:Alice_par_John_Tenniel_24.png

Brahe (1546-1601), found that there was a relation between the planet's distance to the sun and its speed: the so-called 'equal areas-law' ▷2 . This model was used to summarize the EMPIRICAL findings of earlier astronomers.

4. *A planet is seen as a point masses, moving in the sun's gravity field.* Newton (1642-1727) postulated a model for the interaction between any two bodies with mass. This interaction is a force, experienced by the two bodies. It is the force that pulls the ripe apple from the tree, the unfortunate child from its bicycle, and the cast stone back to the ground. It is called gravity, and Newton postulated that it not only works between earthly objects and our home planet, but also in the cosmos between stars and planets. The merits of Newton's model are, that it explains Keppler's laws, and that it predicts, for instance, whether an approaching meteorite will collide with the earth.

5. *Planets are seen as multiple objects* , mutually interacting through gravity. Newton lacked the mathematical tools to find the orbits of three or more objects (for instance: earth - sun - moon) as a result of their mutual gravity. A general CLOSED-FORM SOLUTION of this so-called three body problem is impossible to find, but high-accuracy numerical approximations are often used instead ▷3 .

Let us investigate, from the above 5 cases, what purposes models for scientific research can have.



- *ancient deities:* to EXPLAIN X is answering a question of the form 'Why is X the case?'. People have beliefs about a system, e.g., a religious system, containing gods and their attributes. Further, there is a set of observations, *viz.* lights in the night sky at varying locations, having various colors. By assuming correspondences between the two, some 'why' questions could be answered. 'Why is that 'star' red?' Answer: 'because it is Mars, the god of war'. Whether such answers are satisfying is a matter of agreement. Nowadays, in the Academic world such answers would no longer be accepted ▷4 .

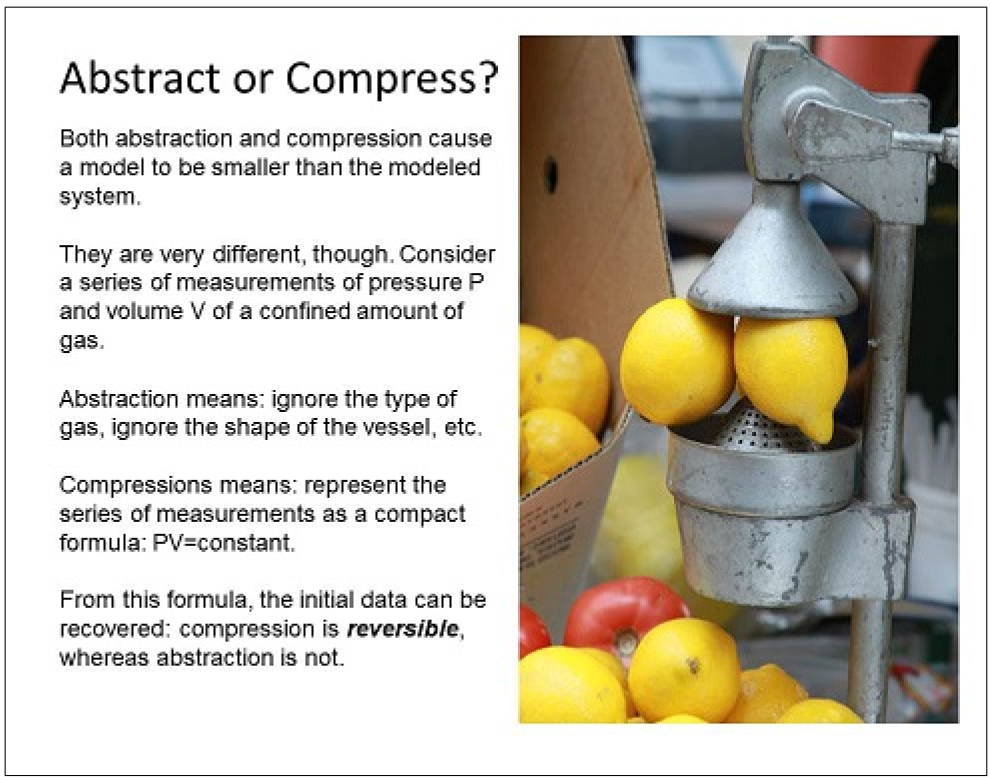
An explanation assumes the presence of some accepted THEORY in which the explanation should make sense ▷5 .

- *Ptolmaic view:* to PREDICT[3] X is answering a question of the form 'When will X happen?'

[3] The figure of the hand used for divination (predicting the future) is taken from http://commons.wikimedia.

or 'What will happen to X?' or 'When will X happen if ...' or 'What will happen to X if ...'. The four types of prediction are explained in the Scheme 'A Handful of Predictions'. Whether an explanation is acceptable is SUBJECTIVE and depending on who the explanation is intended for. A *prediction*, on the other hand, can be OBJECTIVELY verified.

- *Keppler's laws:* to COMPRESS[4] X is answering a question of the form 'How can X be written down in more compact form?'. Tycho Brahe's results comprised lengthy, numerical tables. The three concise Keppler's laws fit on a single Web-page, and contain the same information, in the sense that all of Brahe's data can be reconstructed with Keppler's formulas. In itself, the compression does not add to the understanding: Keppler does not give an answer as to *why* planets closer to the sun move faster. In that respect, compression differs from explanation. Good compression, however, can aid analysis - which in turn can lead to explanation or even prediction. Newton said that he was 'standing on the shoulders of giants': if Keppler would not have compressed Tycho Brahe's numbers into a form that was easy to apprehend, Newton might have failed to find his own results.
- *Two-body gravity:* to ABSTRACT X is answering a question of the form 'What is essential in X?'; to UNIFY X and Y is answering a question of the form 'What do X and Y have in common?'. Before Newton, it was believed that motions of falling and thrown objects on the earth and the trajectories of planets followed different rules. After Keppler's contribution, there was a useful, compact model for planetary motion, although no explanation. Newton's work added two ingredients: the compression became an explanation: now, there was an answer to the question 'why do planets, close to the sun, move faster?' ▷6 . His second contribution was to extend the scope of his model for planetary motion to the sublunary domain. The mechanism that pulls the apple from the tree is the same that pulls the moon towards the earth. This is *unification*: providing a representation that allows explanations and predictions in two domains that initially were thought to be unrelated. Unification is a special case of *abstraction*: leaving out details in the hope that the remaining essentials allow for *generic*



org/wiki/Category:Chiromancy#mediaviewer/File:Chart_of_the_Hand.png

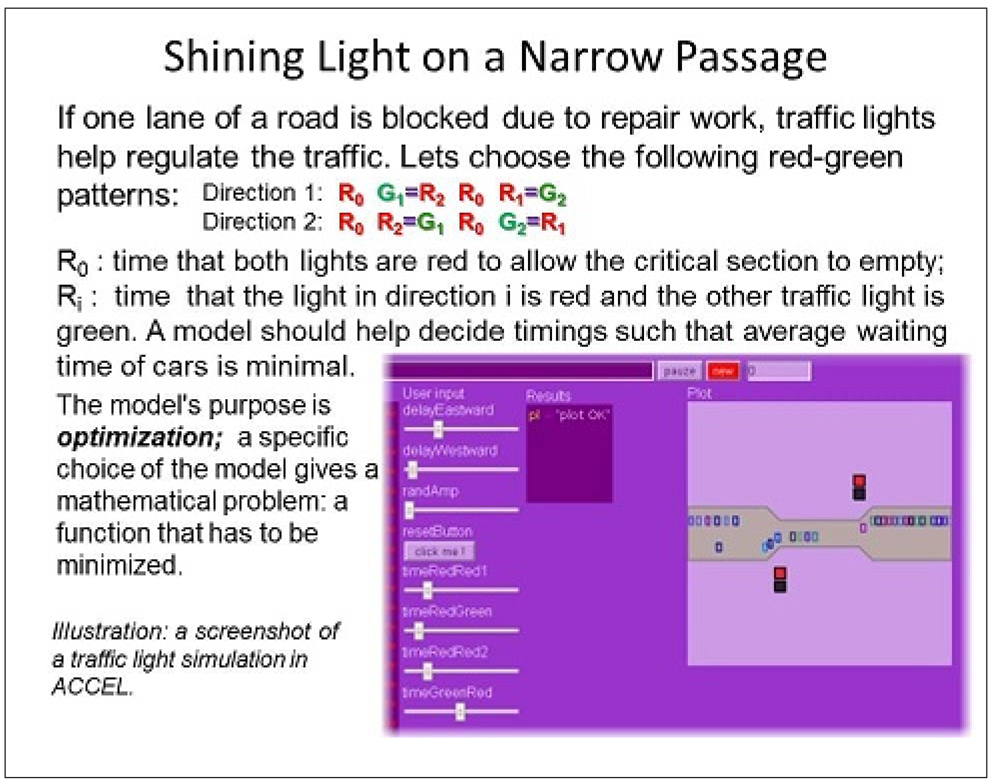
[4] The image of the lemon squeezer is taken from http://upload.wikimedia.org/wikipedia/commons/2/2e/Lemon_squeezer.jpg?uselang=nl

explanation, applying to two (or more) domains.

- *Multi-body gravity:* to ANALYSE X is answering a question of the form 'What is the relation between X and some Y?'. With today's super computers, numerical experiments are being performed to find out if we should worry about future cosmological disasters. These models are used for *predictions*; ideally, however, we would like to use them for the purpose of *analysis*. Analysis is a broader category of purpose: its aim is to DEDUCE properties of a system under study, not studying merely numerical outcomes, but rather the *laws* and *regularities* that underlay these. An analysis contributes to the *understanding* of governing mechanisms or principles; hence, on the basis of an analysis it is often possible to do predictions. Indeed, to do plausible predictions some form of analysis is typically required; not all forms of analysis, however, lead to predictions. One may analyse a single, unique phenomenon ('what was the course of events that led to the 1912 Titanic disaster') , where such analysis will not give rise to any predictions at all.

1.2.2 Purposes of Modeling in the Context of Design

The Solar System can be studied; perhaps it can be partially understood, but we cannot intervene with it. This contrasts with ARTIFACTS such as machines, tools or organizations. These are the result of human decision making: there is an intention, for instance: creating VALUE, and this intention should be fulfilled by doing adequate things. The examples in Section 1.1 fall in this category: the values to be created in the festival case are entertainment, safety for the public and profit for the sponsors and organizers. The values to be created in the wallet-example are the ability to buy food and, if possible, the possession of a particular book. The decisions are, respectively, were whether or not the festival should be canceled because of weather conditions, and whether or not the book should be bought.



There are different ways to do design. Some design processes comprise the following stages; for every stage, we indicate the associated modeling purpose.

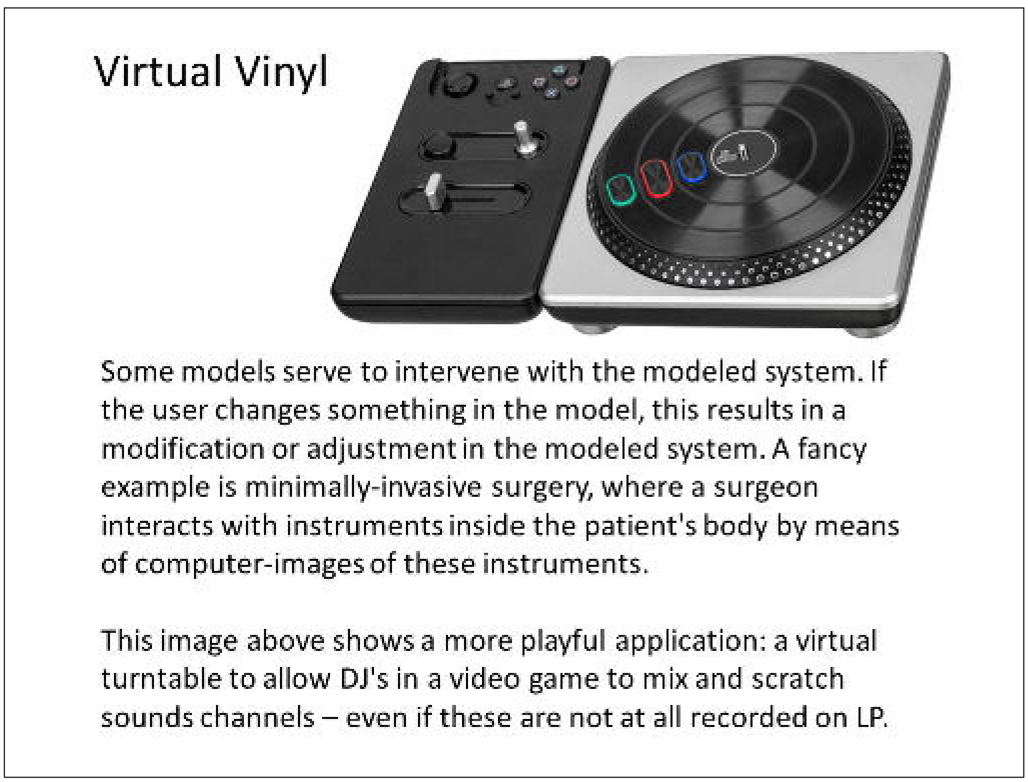
1. *Context formation:* The designer needs to understand the context of the artifact-to-be-designed (we will abbreviate this with ATBD). A model is sometimes used to represent this context. The

purpose of this model is *data compression*: a large and heterogenous amount of information should be condensed into a structured and accessible format.

2. *Requirement elicitation:* this activity may result in a user requirement document. This document describes the ATBD as the user will perceive it, and as such it is a model. The purpose of this model is to COMMUNICATE ▷[7] the intentions of the user to the designer. To *Communicate* X is answering a question of the form 'How do I make sure person P knows X?'. A second purpose, related to requirement elicitation, is to SPECIFY. Specifying X is answering the question 'Which requirements should X (which does not yet exist) fulfill?'.

3. *Option generation:* the elementary activity in a design process is taking a decision. A decision can have two or more outcomes. An example of two outcomes is: 'do intervention X or omit it' (a yes-no decision). With more than two outcomes, we distinguish CLOSED and OPEN decisions. *Closed* means that the list of possible outcomes is fully known beforehand. *Open* means that the designer has to invent the alternatives. A model may be used to aid in suggesting options in an open decision. This is called EXPLORATION. To explore X is answering the question 'What are the various possibilities for X?' ▷[8] .

4. *Selecting the most appropriate option:* once a number of possible options is found ▷[9] , the designer should select one. Often this relates to some form of optimality. The model involved is OPTIMIZATION. To optimize X is answering the question 'What should we choose for some Y such that X is as large or as small as possible?'. A more general purpose is to DECIDE X, that is answering the question: 'What do I choose for X such that some condition is fulfilled?'.



5. *Verification:* once the ATBD is realized, it may be necessary to verify that it indeed satisfies the initial list of requirements. Such models intend to VERIFY X, that is to answer the question 'Is it true that X holds?'.

6. *Steering and control:* Some artifacts, when realized, need no further models to operate. Many modern artifacts, however, employ models for their operation. Think of user interfaces for machines or software[5]. This user interface presents certain options for intervention (sliders,

[5] The image of the virtual turntable is taken from http://upload.wikimedia.org/wikipedia/commons/7/71/

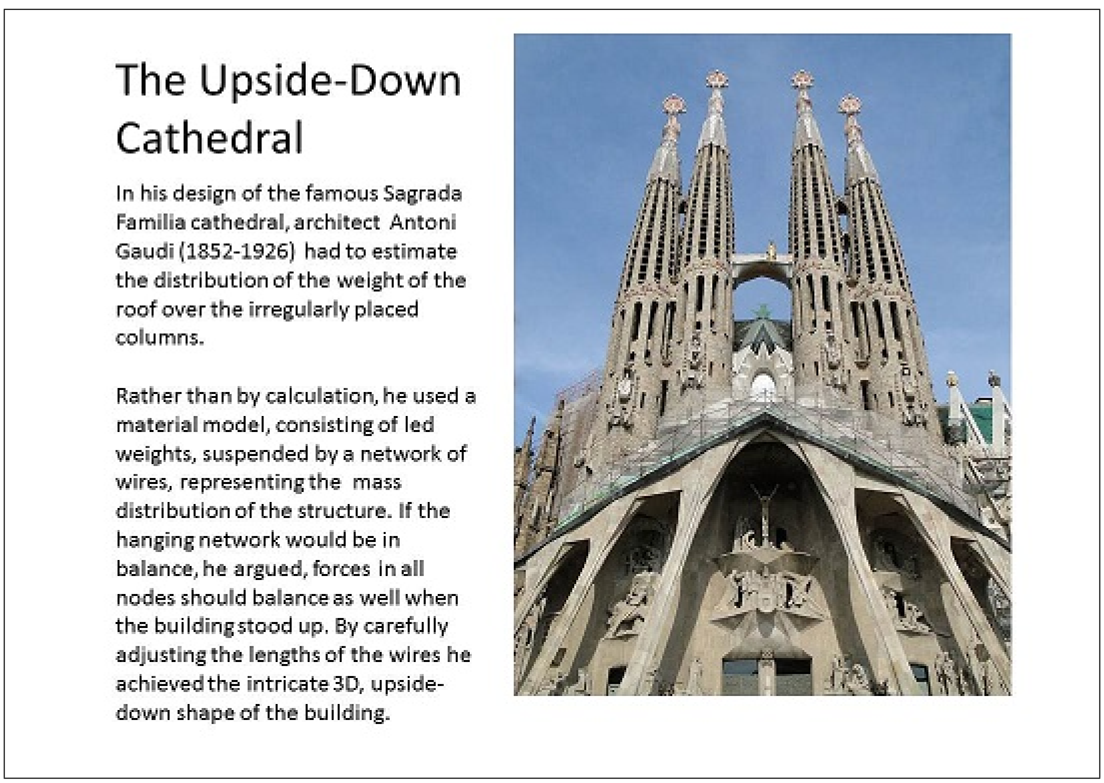
buttons, ...). Conversely, the state of the artifact may be represented by dials, numbers on a display etc. They allow to STEER or *control* the modelled system, that is: affecting the status quo in the modeled system. To steer X is answering the question 'What interventions should take place in X such certain conditions hold?'. Think of the screen of an air traffic controller. This contains moving dots, each dot representing an aircraft. The operating controller may use a cursor to select or interact with one of the dots in order to instruct the respective pilot. The display, and its accompanying software form a model with the purpose to control or steer aircraft. This is an example where the steering or controlling model has a human being in the loop. Other examples exist where the steering or controlling is fully automated, e.g. a thermostat.

7. *Training:* Some models, such as flight simulators, driving simulators or simulators for delicate equipment serve to TRAIN prospect operators. Training to use X is answering the question: 'How could a prospect user learn to operate X without actual dealing with X?'. This includes the domain of medical or surgical training where the modeled system is (part of) a patient.

In Table 1.1 we summarize the various purposes for modeling.

1.3 Modeling Approaches

To meet their purposes, MODELS come in various kinds. We identify a number of DIMENSIONS that can be used to categorize models.



1.3.1 Material - Immaterial

To estimate[6] the forces on a vessel (say, a tanker, a freight ship or a ferry boat) due to water friction, a scaled-down copy of the vessel is sometimes built, and towed in a water tank. Measuring the water

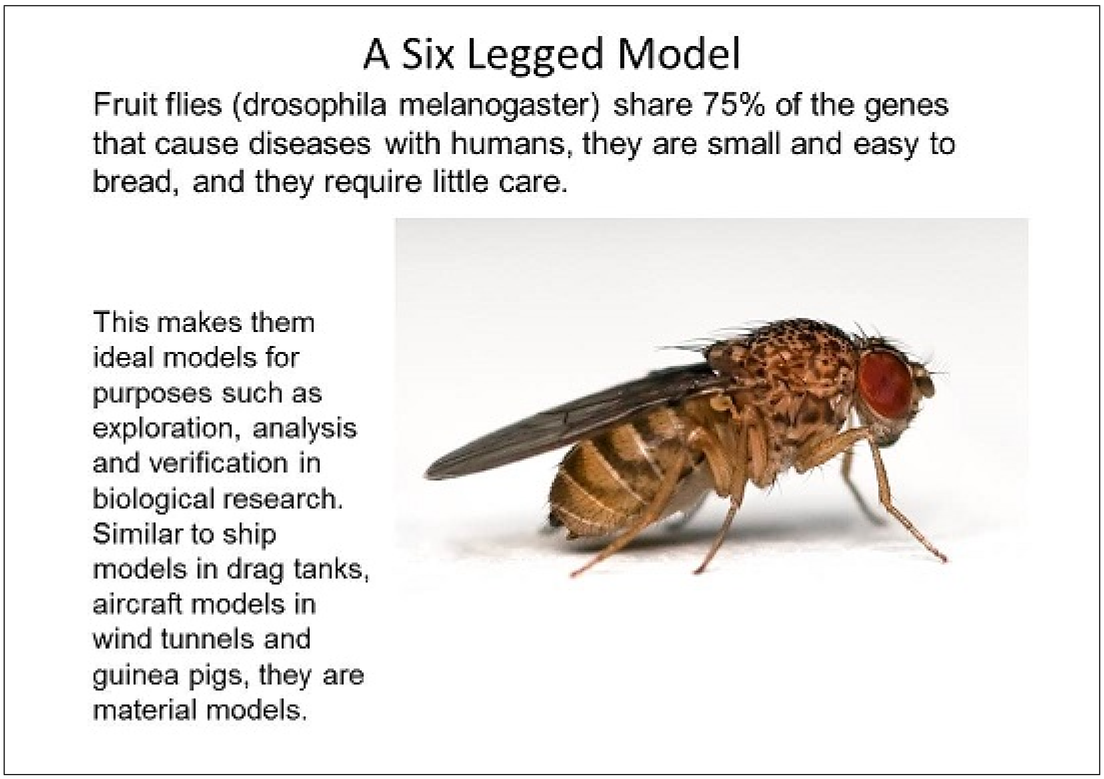
DJ-Hero-PS3-Turntable.jpg?uselang=nl

[6]The image of the Sagrada Famila cathedral was taken from http://commons.wikimedia.org/wiki/Sagrada_Familia#mediaviewer/File:Sagrada_Familia_03.jpg

Purpose	Typically found in Research (R) or Design (D)	Relevant questions
explanation	R	Who should be convinced? How non-obvious is the phenomenon to be explained?
prediction 1 ('unconditional')	R & D	What do we assume to stay the same until the prediction is to be fulfilled? How accurate should the time be foretold?
prediction 2 ('conditional')	R & D	What do we assume to stay the same until the prediction is to be fulfilled? What is the condition, and how can it vary?
compression	R	Is the data set sufficiently coherent to expect something meaningful (no outliers, no different phenomena in one data set, ...)? In what form whould the compression represent the data (formula, graph, ...)?
abstraction	R	Abstraction means: leaving out details. Are the details to leave out sufficiently irrelevant? Are the statements (predictions) about the abstraction still sufficiently concrete to be relevant for the initial phenomena?
unification	R	Why, and to what extent, should the phenomena to be unified, be similar? Are the statements (predictions) of the unified model still sufficiently concrete to be relevant for the initial phenomena?
analysis	R	What do we hope to learn from the result of the analysis?
verification	R & D	Is the route to verification independent from the route that led to the construction (e.g., verifying is a computer program is correct by just reading the code is not very helpful)? If no counterexample is found, what does this prove?
communication	R & D	To whom is the communication directed? What does the receiving party know? What does the receiving party need to know?
exploration	D	Is the collection of alternatives to be produced sufficiently broad (varied, complete, ...) ?
decision	D	Is the set of alternatives known? Is it a closed or open set?
optimization	D	How optimal does it *need* to be? Can it be assessed if the solution is or is not optimal? How much effort can be saved by going for sub-optimal?
specification	D	Which details can be left unspecified? Can the specification be realized?
steering, control	D	What sorts of perturbations should the controller be able to accommodate?
train	D	What sorts of scenario's should the trainee be exposed to?

Table 1.1: Purposes for models

drag, and scaling up the measured drag, thereby taking the dimensions of the real ship into account, is a routine approach to obtain quantitative results that may be more reliable than solving approximate equations. In another application domain: to predict the effect of new medication on the human metabolism, guinea pigs may be injected with the experimental drug, their physiological reactions being interpreted in terms of risks to the human body.



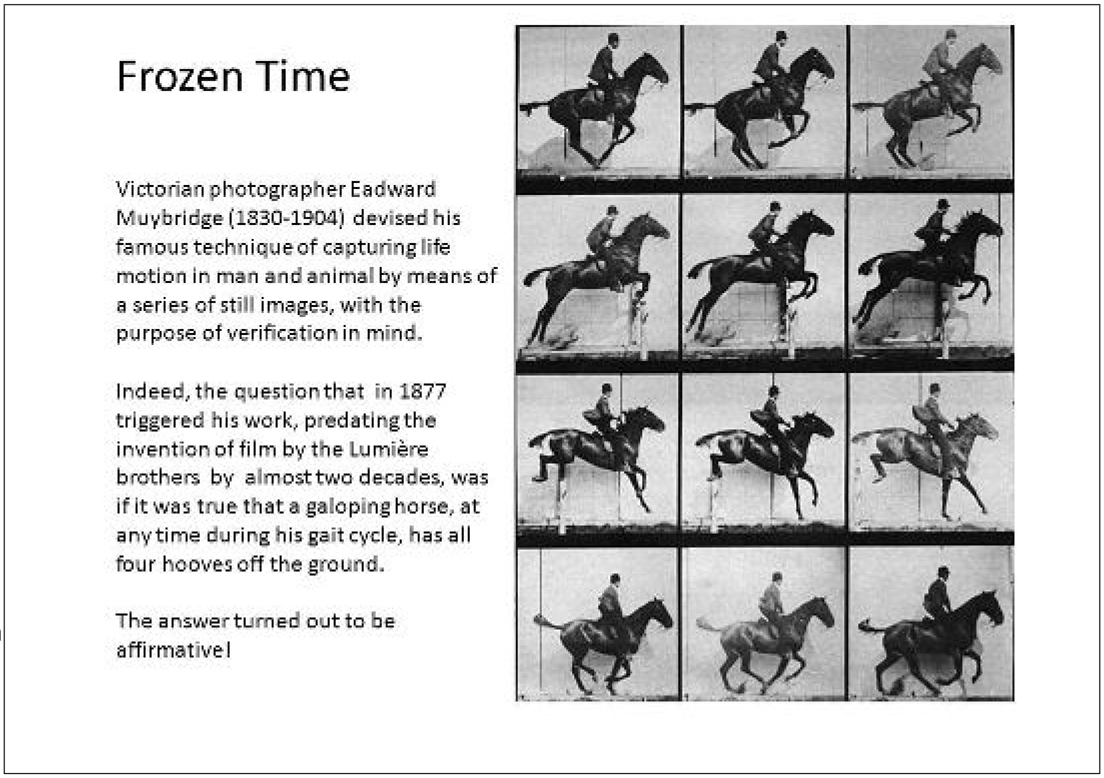
Scaled-down vessels and guinea pigs[7] can act as part of a model. They need a non-material context, however, in order to fulfill a purpose. The measured forces on the scaled-down vessel need to be converted into forces acting on a (hypothetical) full-scale vessel using dedicated mathematical transformations. The guinea pig experiment with the needs to be repeated a number of times in order to be statistically significant, and both doses and symptoms need to be translated to the case of human anatomy in order to mean something. Still, for the sake of brevity, models involving material objects are sometimes called 'material models'. The object occurring in material models can be both a natural object (like a guinea pig) or an artifact (like a scaled-down vessel).

MATERIAL MODELS contrast with IMMATERIAL MODELS.

Immaterial models contain no other material objects than the carriers of information (e.g., paper and ink, or computer memory). Immaterial models are, for instance, mathematical models, consisting of equations and functions; logical models, consisting of facts and rules to connect facts, or software models, consisting of computer instructions, for instance to drive a computer simulation. These three are commonly called FORMAL MODELS or formal systems ▷10 , see also Section 1.4.3 and Chapters 4 and 5.

[7]The photograph of a fruit fly was taken from http://commons.wikimedia.org/wiki/Drosophila#mediaviewer/File:Drosophila_repleta_lateral.jpg

1.3.2 Static - Dynamic



Some problems require time to be taken into account[8], and some don't. To know how strong an electric lamp needs to be to illuminate a segment of motor way at night one can safely ignore the PROCESS of illuminating (that is, light waves traveling through space, reflecting, etc.). It is allowed to talk about STATIONARY quantities, such as light intensity, electric power, reflection coefficients, etc. On the other hand, in a model that predicts how long something takes (say, cooking a 7 kg turkey in a 250 °C oven), it is obvious that the quantities occurring in the model will depend on time. Models that involve time are called DYNAMIC; the opposite is called STATIC.

Many quantities depend on time, but time does not depend on anything else ▷11 . Also, time-dependent quantities may depend on earlier values of the same or other quantities (*memory*), but not on a *future* value of any other quantity. Chapter 3 is entirely devoted to dynamic models.

1.3.3 Continuous - Discrete

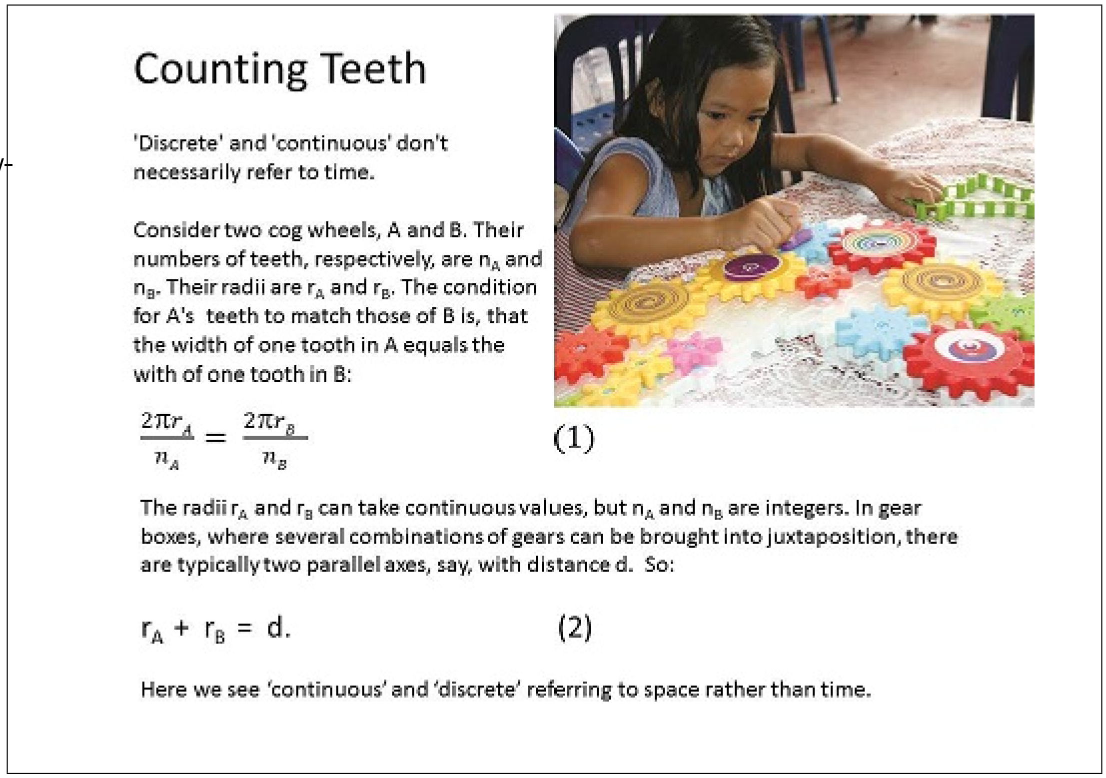
All models involve quantities. Quantities often correspond to sensorial impressions. We see a raindrop gliding down on the window pane, and the vertical position h_r of this raindrop assumes *all* values between the top, h_T, of the window and the window sill, h_{WS}. We see that h_r passes through *all possible values* in between h_T and h_{WS}. It doesn't skip any location. Even if we would follow the raindrop with a magnifying glass, we would see h_r to occupy *all* values between h_T and h_{WS}. The property corresponding to 'not even skipping the smallest hole' is called CONTINUITY. A quantity, like h_r, that can assume a range of values without even the smallest hole is called CONTINUOUS.

Not all quantities are continuous. If I am in a room, I can count the people in the room.

[8] The Muybridge image was taken from http://upload.wikimedia.org/wikipedia/commons/b/bd/Muybridge_horse_jumping.jpg?uselang=nl

COUNTING means: making a correspondence between entities in a set and numbers 1,2,3, $\cdots$, assuming that for every entity, it is clear whether or not it belongs to the set. So: there is nobody standing in the doorway, no people showing on photographs, or other dubious cases. The highest number en*count*ered is the number of people in the room. This will be an INTEGER number: 0, 1, 2, 3, $\cdots$, but never 4.7 or 5.3. Quantities that can only occur as integer numbers are called DISCRETE.

Many problems involve continuous quantities[9]. To do something meaningful with continuous quantities, however, we may have to resort to SAMPLING. Sampling approximates a continuous quantity by a discrete quantity. This discrete quantity occurs in steps that are so small that in practice no problems occur when working with the discrete quantity instead of the continuous one.



The Lumière brothers, when they invented motion picture in 1895, introduced sampling of moving scenes. A money system, using a smallest coin as unit, samples wealth.

Sampling is admitted if the behavior of the HYPOTHETICAL continuous quantity can be RECONSTRUCTED from a set of discrete samples. To find the height y at time t of a football from two images taken at times t_0 and t_1 in a film, showing the ball at heights y_0 and y_1, respectively, we may do LINEAR INTERPOLATION, that is calculate:

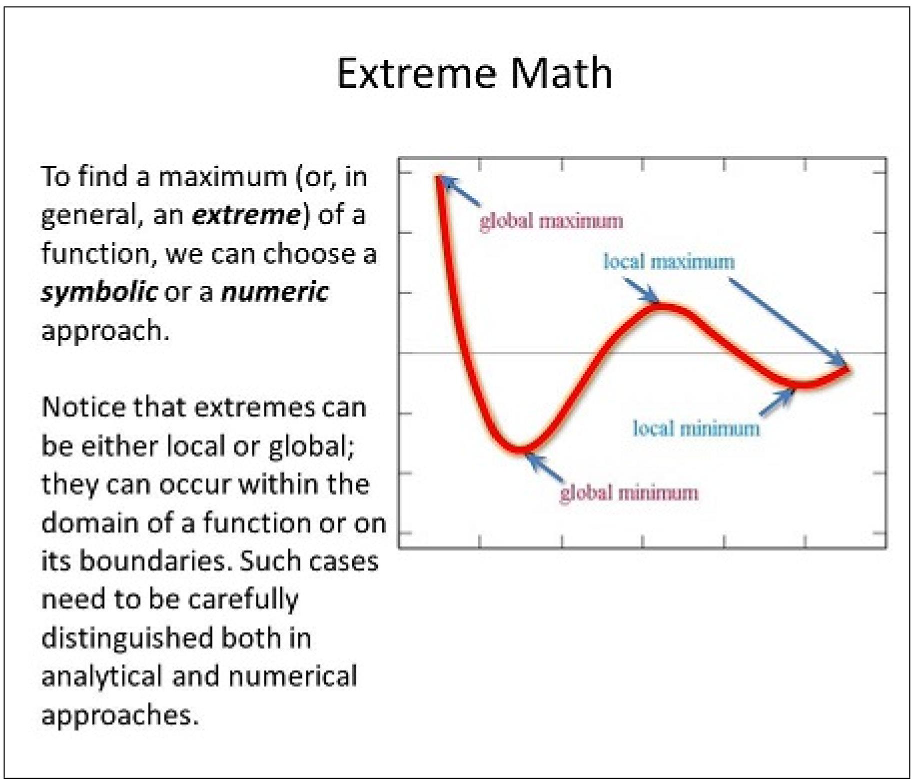
$$y = y_0 + \frac{t - t_0}{t_1 - t_0}(y_1 - y_0). \tag{1.1}$$

This is a simple example of reconstruction, where a continuous quantity (y) is reconstructed from a set of discrete samples of this quantity (y_0,y_1). This occurs in many modeling contexts ▷12

[9]The image of the toy gears is taken from http://upload.wikimedia.org/wikipedia/commons/1/18/Gear_toy_and_young_girl.jpg?uselang=nl

1.3.4 Numerical - Symbolic

To illustrate the difference between numerical models and symbolic models, we look at an example. We want to know, for some function $y = f(x)$, for which value for x between x_0 and x_1, y reaches its maximum. We can reason as follows. We solve $\frac{d}{dx}f(x) = 0$ for x, and record all solutions x_i between x_0 and x_1. These correspond to so-called STATIONARY POINTS of f: points where it has a horizontal tangent. We only record x_i where f has a local maximum. Then, we evaluate $y = f(x_i)$ in all x_i-values; we also evaluate $y = f(x_0)$ and $y = f(x_1)$, and the largest of all these y-s is the answer.



This procedure uses mathematical ANALYSIS. That is: it performs operations on symbols and formulas, e.g. by differentiation, rather than numbers. It applies to any differentiable function f - but in many cases we won't be able to solve $\frac{d}{dx}f(x) = 0$. For instance, if f is a high degree polynomial, or a complicated transcendental function .

An alternative procedure is that we program a computer to evaluate $y = f(x)$ in, say, 100000 EQUIDISTANT values for x between x_0 and x_1, and store these y-values. The y-value which is the largest is a likely approximation of the absolute maximum. That is: we apply a simple ITERATIVE ALGORITHM.

Most computer programs ▷[13] operate on numbers rather than non-numerical symbols. This approach is therefore called NUMERICAL modeling.

The two strategies both frequently occur in modeling. They have a number of striking differences, though (see Table 1.2) ▷[14] .

1.3.5 Geometric - Non-geometric

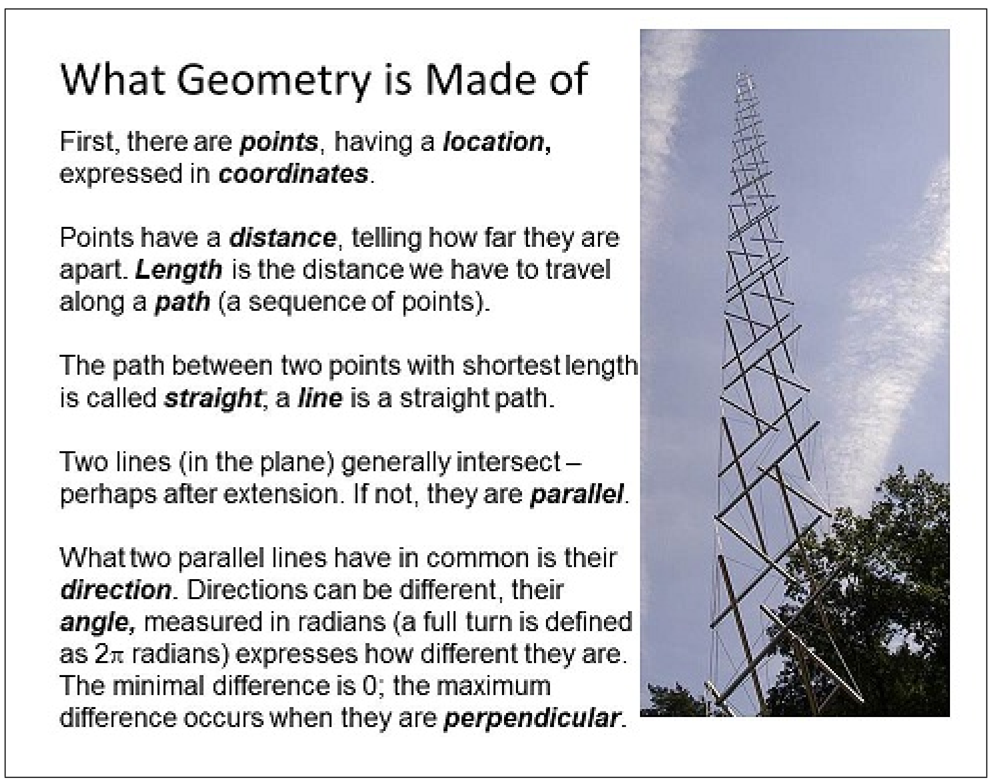
According to our sensorial experience, three independent quantities relate to space, for instance horizontal distance, vertical distance and depth distance, measured from some given reference point and using a set of reference directions. Quantities used to distinguish spatial locations are usually called COORDINATES. Many interesting problems are governed by fewer than three spatial coordinates. For a train on a track, for instance, only one spatial coordinate may be relevant,

Feature	Symbolic	Numerical
precision	irrelevant, but if the answer needs to be a number, we need numerical approximations at the end of the calculation.	a point of concern: numerical calculations typically introduce round-off errors that can make the outcome fully unreliable.
generality	limited. Many mathematical results exist, but in order to actually calculate something one is typically restricted to few simple cases. For instance: linear functions, low-degree polynomials, or trigonometric functions.	generic. Numerical operation is not restricted to closed form 'simple cases'.
ease of use	proficient use of mathematical analysis requires abstract thinking and precision, plus knowledge of textbook mathematical results.	many standard, reasonably robust methods are available, for instance in libraries (Matlab, Excel and others).

Table 1.2: Differences between symbolic and numerical modeling

as the train can not move sideways nor up and down. For a game of chess, two coordinates are enough. From the latter example, we learn that spatial coordinates can be either continuous or discrete.

Apparently, there are some fundamental concepts (DISTANCE ▷15 , STRAIGHT, LENGTH, LINE, COPLANAR, PARALLEL, PERPENDICULAR, DIRECTION and ANGLE), related to our perception of space, that have been given a more precise mathematical definition[10]. They have become MATHEMATICAL OBJECTS ▷16 . The area of mathematics that studies the relations between the sorts of objects mentioned above is called GEOMETRY.



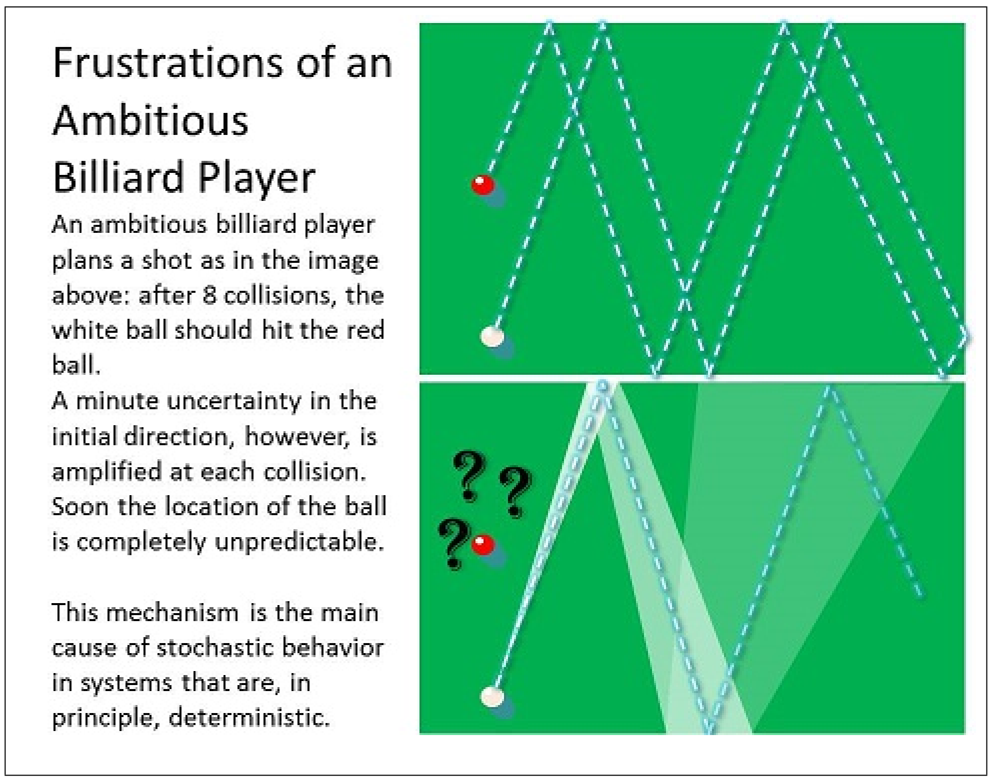
Models hinging on space-related quantities are called *geometric* models. Often, space-related quantities are properties of material objects. For instance, in the design of an artificial heart-valve, both the shape of the heart and the spatial patterns of the blood flow need to be accurately represented. Similar in the design of machine parts (3D geometry) or urban planning (2D).

[10]The image of the needle tower, by Kenneth Snelson, is taken from http://commons.wikimedia.org/wiki/File:KrollerMuller_ParkSculpture4.jpg?uselang=nl

1.3.6 Deterministic - Stochastic

This chapter started with two examples of models: a weather model that predicts a thunder storm, and the wallet model that predicts whether or not I can afford buying a meal tonight. The wallet model makes a certain statement ('if you buy a 25 Euro book, it is certain that you will have enough money left to spend 5 Euro's for a meal'), whereas the weather model does a probabilistic prediction ('there is a 95% chance that ...').

To understand this difference between so-called DETERMINISTIC and STOCHASTIC models, we study an example from physics.



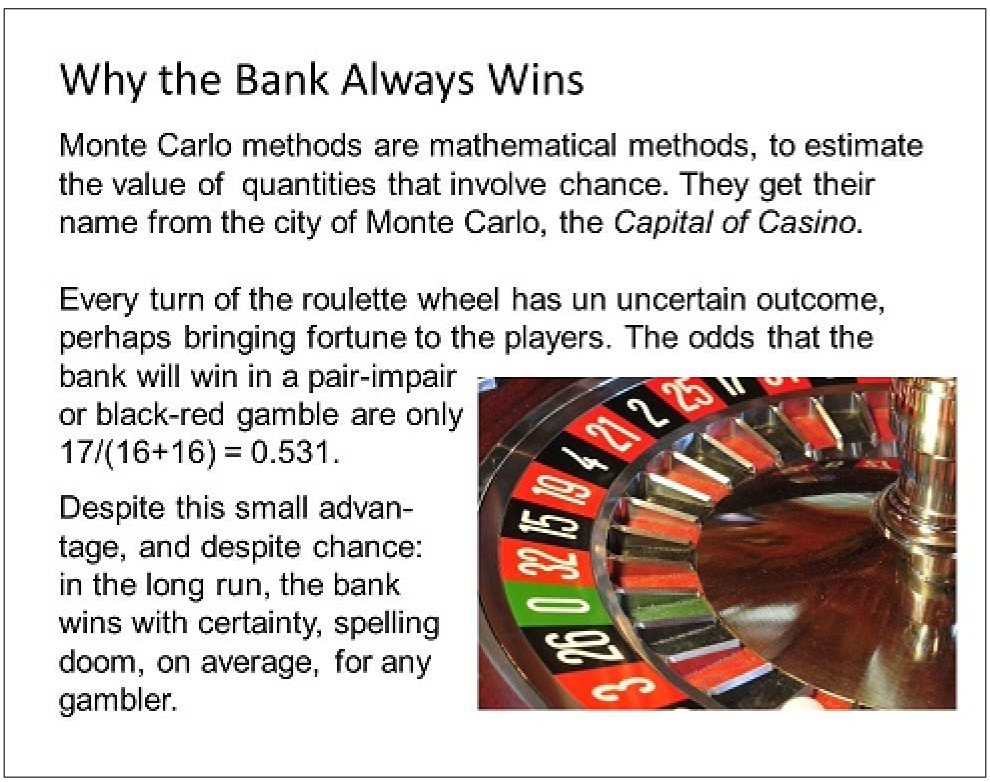
Let us consider a gas, consisting of atoms in a closed container. We view these atoms as miniature billiard balls. Just as billiard balls, atoms can collide with each other, and they can collide with the walls of the container. First, suppose that there are very few atoms. Just as a proficient billiard player can determine what routes the balls will follow by careful observation and control of the initial configuration, the trajectories of these few atoms could, in principle, be fully known from their initial configuration as well. The system is behaving in a deterministic manner: the outcome leaves no room for uncertainty if the initial configuration is known with 100% accuracy ▷[17] . In practice, however, small uncertainties are unavoidable. Still, a trained billiard player can perform a shot resulting in two or sometimes even more planned collisions with each of the balls (caramboles) with a large chance of success. This is increasingly difficult if the number of balls increases, or if the time period during which the balls roll increases. Unavoidably, uncertainty creeps in, and soon the pattern of collisions becomes completely unpredictable or RANDOM. Compare this to rolling a die: there, also, a large amount of collisions, each time causing a change in the position and orientation of the die, makes the outcome unpredictable.

In the case of the gas, every collision of an atom with the wall of the container gives a small transfer of momentum to the container, and, similar to the randomness of collisions with sufficiently many billiard balls, these events of momentum transfers appear in an unpredictable sequence.

The predictability decreases if we increase the number of atoms. As we say, we make the ENSEMBLE bigger. The time sequence of impacts on the walls is increasingly random, but at the same time the *percentual* fluctuations in the number of impacts per second will vary less and less with increasing ensemble size [▷18]. Similar, the percentual fluctuations in transferred momentum per second will get smaller.

By the time the amount of atoms approaches a couple of million [▷19] - which is still an extremely tiny puff of gas - the fluctuations in the transferred momentum are so small that we perceive the resultant effect of the momentum transfers as a force that is *constant* in time: it is the *pressure, excerted by the gas* on the walls of the container. So, despite the inherently stochastic nature of the collection of swarming atoms, there is a property of the system *as a whole* that is not at all stochastic. This is an example of a BULK QUANTITY of the ensemble. We have LUMPED the many properties of the individual atoms (their locations and speeds) together, to find a new, so-called EMERGENT [▷20] quantity that is not at all stochastic: the *pressure*. Indeed, the pressure P of n moles of a gas in a volume V at temperature T behaves in a deterministic way, as described for instance by the gas law $PV = nRT$ (R being the so-called gas constant).

In this example we see how AVERAGING-OUT of the individual properties of a sufficient number of elementary entities in a stochastic ensemble may lead to determinstic behavior of bulk-properties[11]. Most deterministic laws of nature present themselves in this way: they are deterministic, because they relate bulk-properties, applying to ensembles of sufficient size. Also outside physics, emergent behavior and bulk-properties occur: think of demographic phenomena, economy, traffic, communication and many other fields.



There are modeling strategies based on this idea of ensmebles. In such a model, sometimes called MONTE CARLO MODEL, the purpose is to obtain emergent behavior out of a sufficiently large repetition of a simulated experiment, where individual variations are assumed to be random and UNCORRELATED.

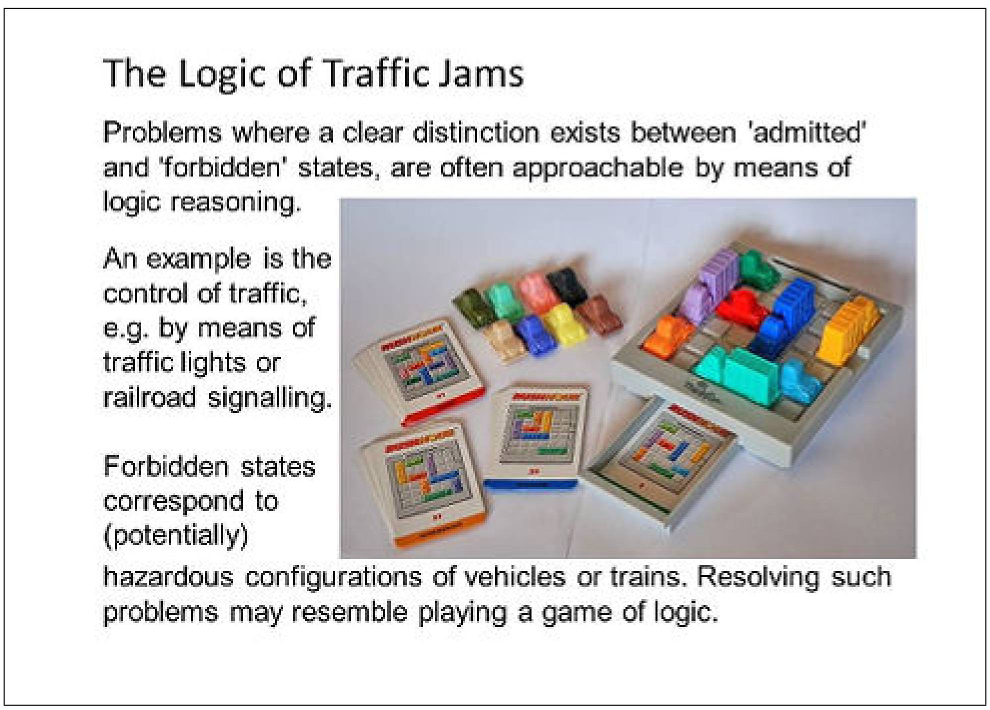
For example, we may want to find out if eating apples is healthy. We take two equal sized groups of people, one group of apple-eaters and the other of people who don't eat apples. Next we ask everybody for the numbers of times (s)he visited a physician over the last three years. Obviously, the numbers vary from individual to individual, but if apples are healthy we expect that, apart from the random variations,

[11] The roulette table image comes from http://commons.wikimedia.org/wiki/Roulette#mediaviewer/File:13-02-27-spielbank-wiesbaden-by-RalfR-064.jpg

there is a significant difference between the averages in the two groups.

Lumping omits details in individual entities in models of stochastic systems. Also in deterministic models, details may be deliberately omitted. For example, consider the traffic lights problem discussed in Section 1.2.2 (figure 'Shining Light on a Narrow Passage'; go to this link for an interactive demonstration). Here the flow of traffic has to be modeled to solve the problem. Typically, random components are present in the flow of traffic and a stochastic model lies at hand. A deterministic model, however, might be useful to get some knowledge concerning the problem and might be good enough for the given purpose. One might assume, for instance, that the cars arrive in a deterministic pattern with a constant inter arrival time. So the variability in the traffic flow is not accounted for in this model and detailed knowledge is omitted by deliberately ignoring the randomness in the modeled system.

1.3.7 Calculating - Reasoning



Many models involve *calculations*, either with numbers or with symbols. There is a different class of FORMAL models, however, where answers do not come from operating upon numbers[12].

For instance: consider a software system to control signals and switches (Dutch: 'wissel') of a railroad junction. As we will discuss in more detail in 3.2.1, the combination of all states of the signals and the states of the switches can easily amount to many billions of possible combinations. Many of these are illegal: for instance, if the signals on both incoming branches of a switch show green, we can foresee a collision if two trains approach this switch at the same time. It is impossible, though, to verify all states by hand to see which are admitted. So it is impossible to verify if the software leads to safe train traffic by just inspecting the computer program.

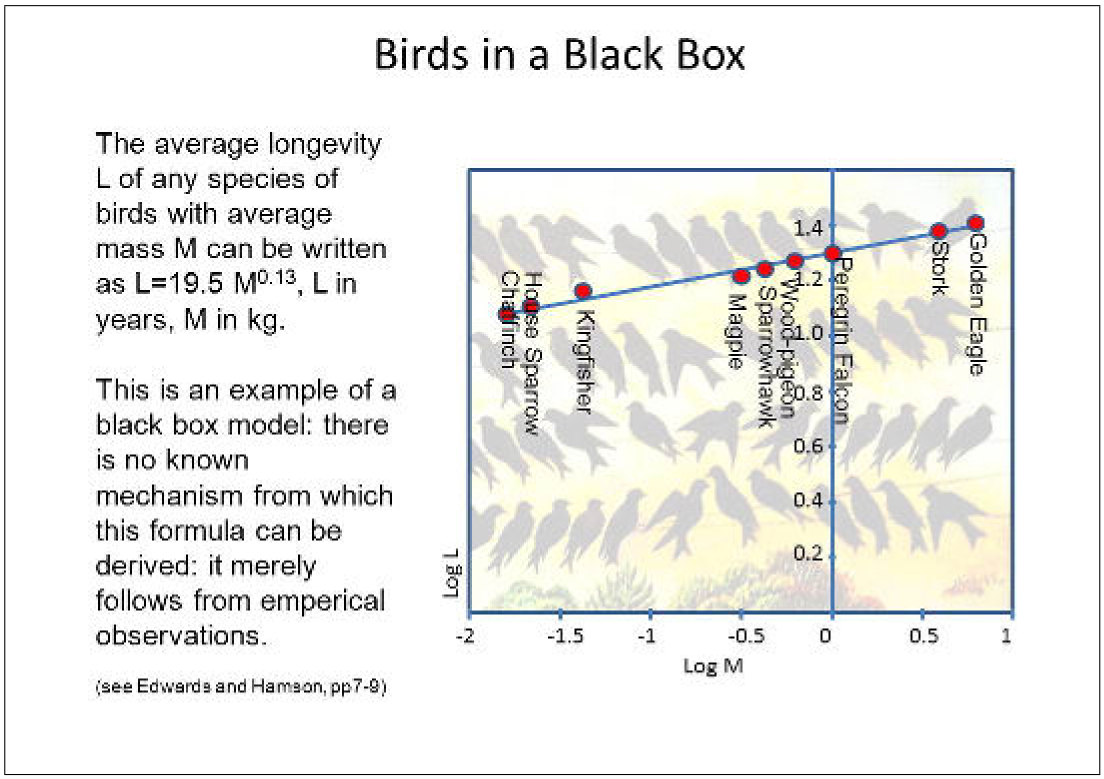
Instead we make a model of the junction-control software in terms of logical expressions. By combining these expressions, using another computer program, we may be able to prove that indeed the software system is safe, or we may be able to spot flawed configurations. Logical expressions, in some sense, resemble arithmetic expressions. There are also quantities, and the values of some quantities depend on others. But instead of addition, multiplication, etc., the

[12] The image of the rush hour game is taken from http://commons.wikimedia.org/wiki/Category:Logic_puzzles#mediaviewer/File:Rush_Hour_sliding_block_puzzle.jpg

logic expressions consist of such operations as ' AND', ' OR', ' IMPLIES', etc.; the values of the quantities are 'TRUE' or 'FALSE'. Using a logical model we can REASON rather than CALCULATE.

Other examples include the retrieval of information from DATABASES, or the application of logical rules in an EXPERT SYSTEM or KNOWLEDGE BASE.

1.3.8 Black Box - Glass Box



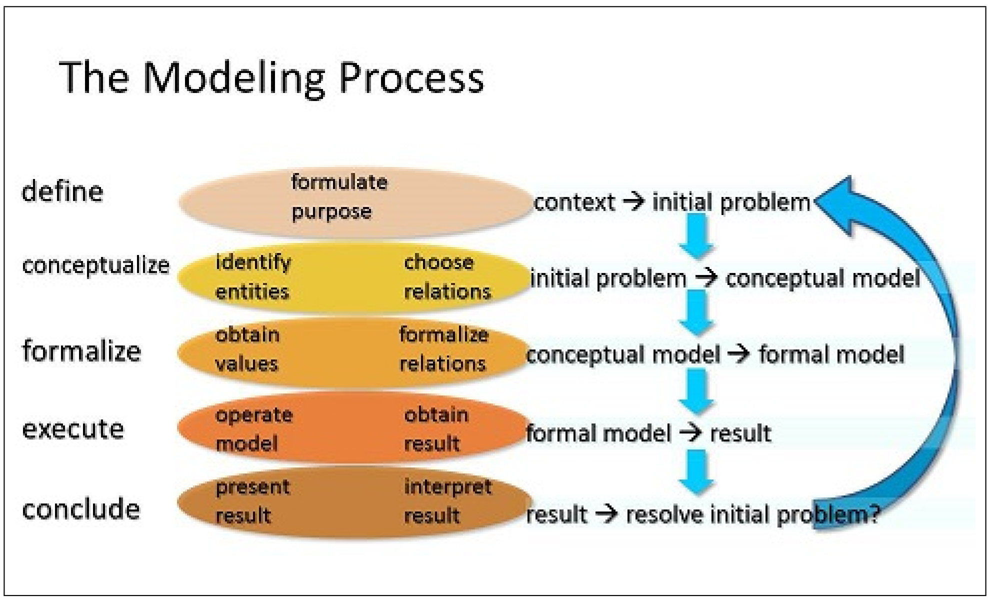
For a number of species of birds, we assess their longevity and their average mass[13]. We have no A PRIORI idea if these two are related. Plotting the data in a graph may help our intuition. Such a plot is a model with the purpose of *compression*: it may suggest us how to proceed. If mass and longevity would be unrelated, the data points would be scattered all over the place. The plot seems to suggest, however, a smooth dependency. Plotting the data on a log-log graph, we find that the behavior is reasonably well described as $\log_{10}(L) = 0.13 \log_{10}(M) + 1.29$, for L=longevity in years and M=mass in kg. We have made a form of *compression* here: we start with a data set with several dozens of numbers, and we bring this information back to one formula with two FITTED quantities.

This formula says nothing about any species of birds *not* occurring in the initial data. Since we have no explanation of why longevity should be related to mass, we cannot make any claims about new species of birds. But we might want to check if the data for other kinds of birds happen to be close to the straight line in the log-log plot ▷21 . Then the compressing power of our formula improves: we can compress more data with the same accuracy with a single formula. Next, even though we have no explanatory model, we may be tempted to start doing predictions. Compare this to the Keppler approach to planet motions, or the relationship between longevity and mass in birds from the Figure 'Birds in a Black Box'.

[13]The oil painting of birds on wires is taken from http://commons.wikimedia.org/wiki/File:Birds_on_a_Wire.jpg?uselang=nl

This type of modeling is called BLACK BOX modeling. The idea of a black box is, that the inside is invisible. We cannot see 'inside' the biological mechanisms that cause a certain species of bird to attain a certain mass, or that causes them to reach a certain age. Still, the fact that the data points are not random, but allow compression, suggests that there *is* a connection between these two mechanisms. Perhaps both are caused by some, hidden, third mechanism. The stronger the compression, the stronger the suggestion that there is a deeper mechanism ▹22 hidden in the system. A black box model, however, does not attempt to reveal these mechanisms.

On the other hand, a transparent box or GLASS BOX offers an unrestricted view on the internal workings inside the box. Glass box modeling starts from POSTULATING the way the we think things behave.



For instance, an engineer investigating the behavior of tall buildings during an earthquake may postulate that the buildings behave as an upside-down pendulum. Starting from this postulate, (s)he sets up a model for such a pendulum. So: the internal mechanisms of a glass box may be completely 'wrong' (a tall building *is* no pendulum), but they are used none the less to obtain predictions or other statements about the MODELED SYSTEM.

The term WHITE BOX is sometimes used instead of glass box.

A combination of a black box model and a white (glass) box model for the same system is sometimes called a GREY BOX model.

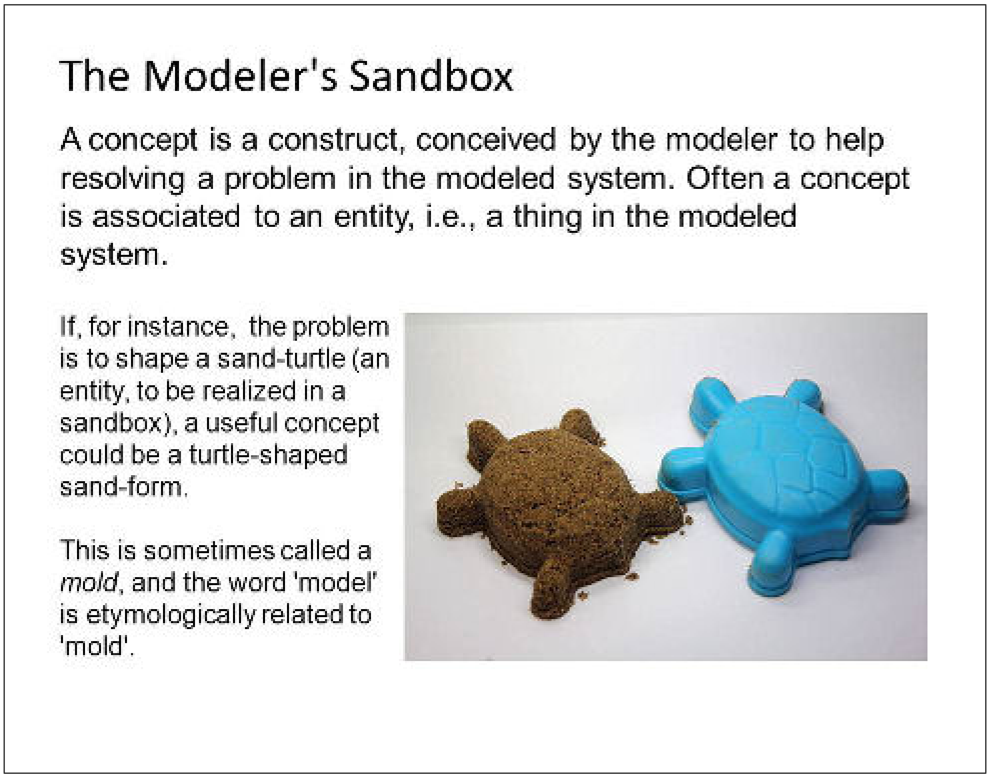
1.4 Stages in The Modeling Process

The MODELING process assumes a number of different stages, as depicted in the diagram 'The Modeling Process' ▹23 . The stages in the modeling process are called **definition**, **conceptualization**, **formalization**, **execution**, and **conclusion**. Further, after the completion of a modeling process, the modeller may do an **evaluation**. Below the stages are elaborated.

1.4.1 Definition Stage (1)

Any modeling process starts with the problem DEFINITION. The overview of purposes in Table 1.1 provides some help here, although models often should serve several purposes at once. It also may happen that, during the modeling process, the purpose changes.

1.4.2 Conceptualization Stage (2)



During the CONCEPTUALIZATION stage, the CONCEPTUAL MODEL is constructed. The conceptual model, to be elaborated in Chapter 2, is a collection of entities[14], their properties, and relations between them, but these are not yet in mathematical form. So there are no equations yet, and no mathematical derivations: these constitute the formal model, to be elaborated in Chapters 4 and 5.

- *Conceptualization (2.1): Identify Entities and Properties*

The things that occur in the problem domain could just be called 'things', but we prefer the term ENTITIES. To some of the entities, occurring in the model domain, we associate concepts in the conceptual model. The concepts in the conceptual model stand for those entities of the modeled system (i.e., the problem domain) that we want to take into consideration. The set of entities to be considered must be large enough to suit the purpose. For instance, it should contain elements so that we can assess if the problem is solved. Also the set of entities should not contain too much as this makes the model unnecessarily complex.

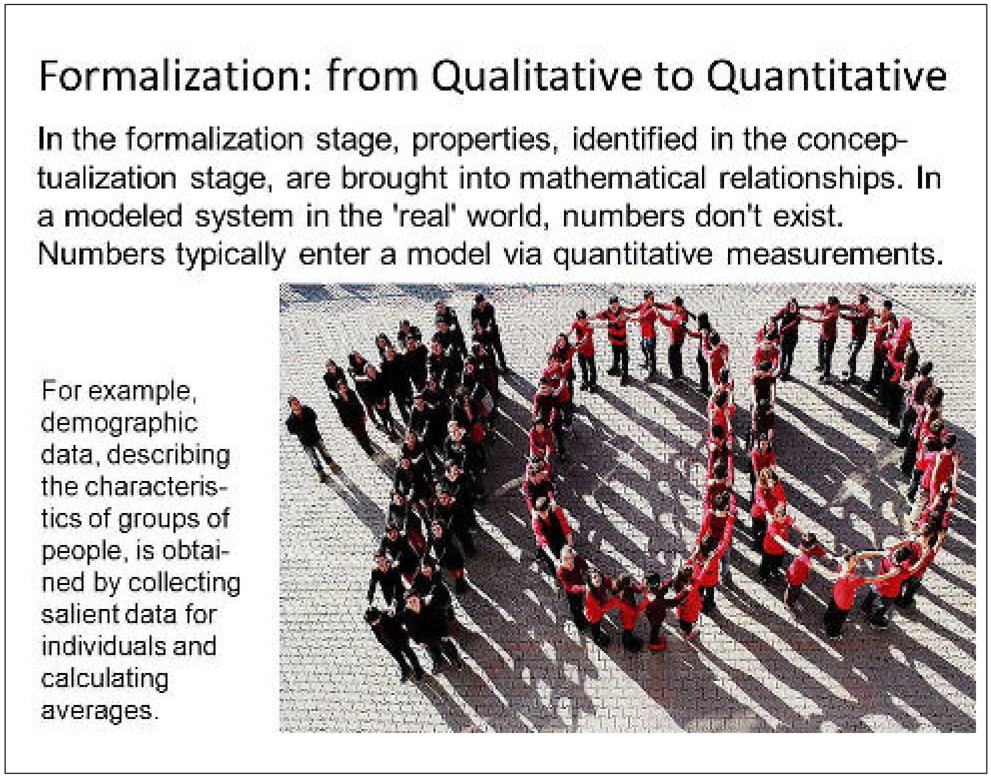
- *Conceptualization (2.2): Identify Relations between Entities and Properties*

If entities in a model are isolated, the model won't *do* anything. We need relations as well. Relations are often depicted in a schematic drawing. In such a so-called GRAPH the entities are boxes or circles. The general name is NODES. The relations between the entities are ARCS or ARROWS. The *meaning* of such arcs or arrows should be made explicit, and used consistently. We call this LABELING. This way to denote a set of entities is called and ENTITY-RELATION GRAPH.

[14]The image of the turtle mold was taken from http://upload.wikimedia.org/wikipedia/commons/7/70/Sandfoermchen-3.jpg?uselang=nl

1.4.3 Formalization Stage (3)

To 'do' mathematics (calculations) or logic (reasoning), we need quantities to compute or argue with. Many of them will represent properties of the entities from the conceptual model.



The FORMALIZATION stage transforms the conceptual model, consisting of concepts, properties and qualitative relations, to a FORMAL MODEL consisting of quantities and quantitative relations connecting them.

For quantities to occur in computations, they have to have values, and they have to take part in formal relations. An example of a formal relation, also called FORMULA is '1+1=2'. The symbols '1', '2', '+' and '=' have formally defined meanings. ' FORMAL' means: that which is defined in a logically consistent system, and does not require human interpretation in order to be operated. Arithmetic is the best known example of a formal system, but also the rules to play games such as tic-tac-toe, go or chess are formal systems.

In the formalization stage, we transform the conceptual model to one that is expressed in a FORMAL form, that is: one that no longer relies on human interpretation. There is a number of different ways values can be BOUND to quantities. One very common one, is that values come in as measured data.

Measured data[15] either can be RAW ▷24 data, or it can be processed data. Processing data often amounts to grouping a collection of numbers (=measured values) together, and summarize the information in these measured values into few numbers. The latter is called AGGREGATED data. Averaging, such as in Section 1.3.8, is an example of aggregation.

- *Formalization (3.1): Obtain Values for Quantities*

This stage is about the collection of data that serves as input for the model: either raw data or processed data, such as aggregated data.

- *Formalization (3.2): Formalize Relations between Quantities*

If the model is a glass box model, it contains knowledge about the mechanisms inside the modeled

[15] The image of the celebrating crowd forming the number '100' is taken from http://commons.wikimedia.org/wiki/Category:Human_formations#mediaviewer/File:100_years.jpg

system. For a black box model, these relations come in the form of HYPOTHESES, for instance that there is a linear relation between log(longevity) and log(mass) in birds, as in the example from Section 1.3.8. A hypothesis is a postulated proposition or relation, that is assumed to be provisionally true, but that will be subject to testing.

Recipes to Reckon

Formal operations are those manipulations to symbols (including numbers) that can take place without interpretation.

For instance, the equality

$$(a+b)^2 = a^2 + b^2 + 2ab$$

holds irrespective of the values or the meanings of a and b.

Parallel to collecting, discovering or introducing relations, the modeler should collect, discover or introduce ASSUMPTIONS. Assumptions should be documented; they can assist later to assess or limit the plausibility of the model's result, and they can inspire to do MODEL REFINEMENT. Model refinement means: one or more iterations of the modeling process in order to obtain a model that better fulfills the purpose.

1.4.4 Execution Stage (4)

Once we have a formal model[16] that we feel confident with, we can start *using* the model for its purpose. This is called the EXECUTION of the model. In many cases, this involves some form of calculating or reasoning, e.g. to solve equations, to search for an optimal solution, or to perform some algorithm.

- *Execution (4.1): Do Operations with the Model*

The OPERATIONS with a model should comply with its intentions. These operations should lead to fulfillment of the purpose. For instance, if the purpose is 'compression', the model should produce a compact representation of data; if the purpose is 'verification', the model should produce a result 'TRUE' or 'FALSE', etc..

For some of the purposes from Table 1.1, such as 'exploration' and 'communication', there may be no need for formal operations. Think again of our black box model regarding birds' masses and longevity. If the sole purpose of the model were to communicate empirical findings, the data from a table plotted on a log-log scale could be an adequate result.

But for formal models, in general, there is at least some formal operation in the modeling process.

In some formal models this amounts to mathematical handwork. Say, deriving expressions, applying

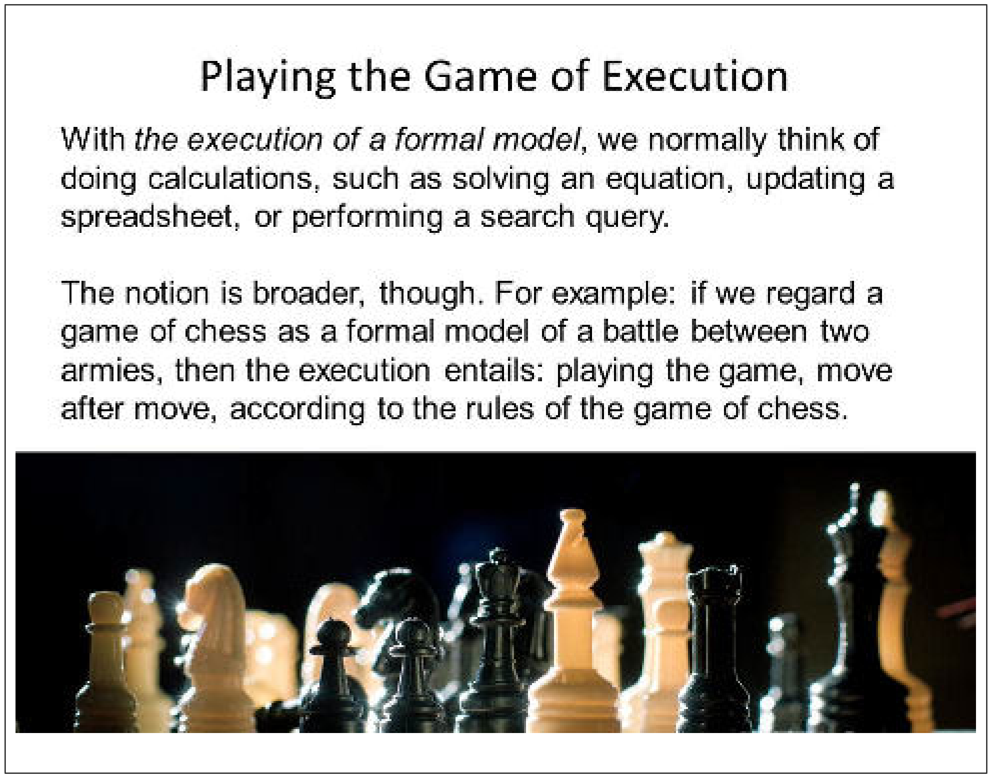
[16] The woodcut with medieval calculus-masters was taken from http://upload.wikimedia.org/wikipedia/commons/3/38/Gregor_Reisch%2C_Margarita_Philosophica%2C_1508_%281230x1615%29.png?uselang=nl

transformations, doing CALCULUS, etc.. Most formal models, however, involve a computer; in that case, *operating the model* amounts to running a computer program. The relations between quantities, established in stage (6), then lead to the statements instructing the computer to perform calculations or INFERENCES for obtaining the desired result.

• *Execution (4.2): Obtain a Result*

Most purposes from Table 1.1 cause a model to produce a result in the form of some mathematical object[17]: the value of a quantity, a set of numbers, a graph, etc.

Apart from obtaining the sought mathematical object, the modeler should always strive for obtaining insight as well. After having completed the operations with the model, looking at the obtained result, the modeler's first question should be: 'so what?'. This is elaborated further in the following stages.



1.4.5 Conclusion Stage (5)

The CONCLUSION stage comes after the execution stage. The *results* obtained from the execution are mathematical objects: numbers, graphs, and perhaps more advanced things. The purpose of the model, however, was not stated in mathematical terms. Therefore, there is always the need for a translation back to the problem domain. This translation involves PRESENTATION and INTERPRETATION of the result.

• *Conclusion (5.1): Present the Result*

The purpose of the model relates to a problem, and therefore some PROBLEM OWNER, STAKE HOLDERS, and a PROBLEM CONTEXT. For none of these, in general, the model outcome will be appealing, useful or even comprehensible.

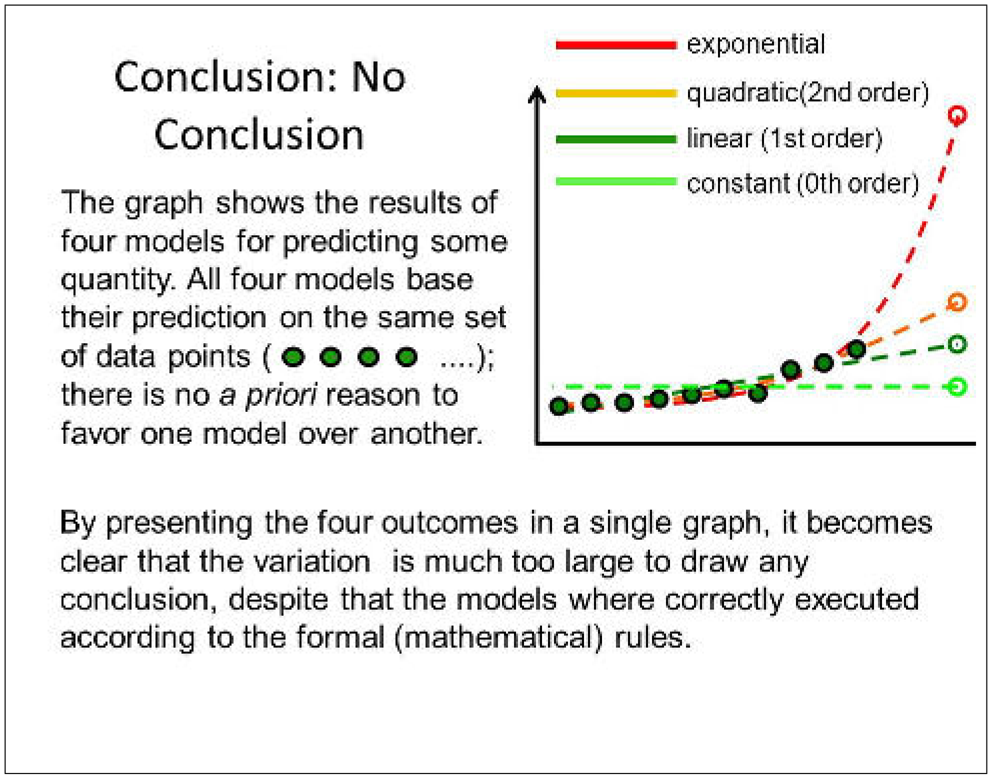
Take for instance a model for predicting the weather. The problem owner is, say, a meteorological institute that sells advise regarding weather conditions; stake holders are people who want to know about the weather. The problem context, among other things contains people's tendency to complain about bad weather, but also to complain about good weather if it was predicted wrongly,

[17]The photograph of the chessmen is taken from `http://upload.wikimedia.org/wikipedia/commons/0/08/Chessmen_in_backlight.jpg?uselang=nl`

and to have a short memory with respect to the quality of earlier predictions.

The model's outcome, obviously, is a table with thousands of numbers, say, representing predicted temperatures as a function of time and place. This table is next to useless for clients such as festival organizers, and it does not fit with the problem context. Therefore, over the last decades, there has been a considerable development in the presentation of weather maps - nowadays including computer animation and even simulated 3D effects.

- *Conclusion (5.2): Interpret the Result*



A well-thought of presentation[18] is crucial for the impact, convincingness and overall success of a modeling effort. But it may not be sufficient to fulfill the initial purpose.

Consider the following example. A numerical model is used to calculate the concentration of some medicine in the blood in dependence of several metabolical conditions; from this concentration, the model does predictions about effectiveness and side effects of the medication. The numerical outcomes are presented using professional graphics. For patients, and even for physicians, however, this presentation has little value. Indeed: the initial purpose is: when should a patient take his pills? Therefore, the presentation should be interpreted, leading to answers such as: 'it is best to take two pills just after breakfast, and a third one later in the day if the fever returns', or something similar. The step from graphs and tables to this form of recommendation is called *interpretation*. It involves non-trivial skills, and it may require to consult domain specialists who have a profound feeling for the problem context.

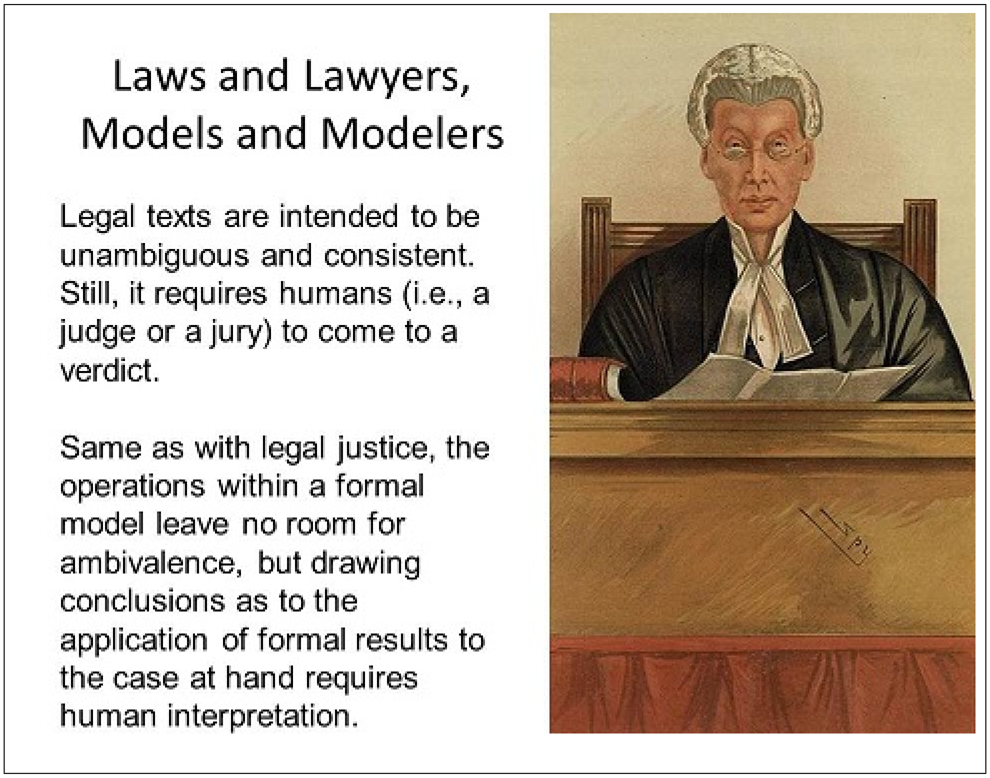
1.4.6 Learn from what you have done: Evaluation for the Modeler

If a valid model leads to plausible results, and the interpretation of these results show that the right problem has been addressed, there is much to be proud of. There is, however, always room for improvement. Even if the current problem instance does not allow or require this improvement: for the modeler proper, pondering on this 'room for improvement' may be advisable. There will

[18]The image of the graph was based on http://commons.wikimedia.org/wiki/Category:Extrapolation#mediaviewer/File:Fstaals_extrapolatie_8.jpg

be next problems requiring next models, and a secondary purpose, hidden in every problem, is the modeler's desire to keep learning.

Therefore a last stage, devoted to REFLECTION should conclude every modeling exercise[19]. Chapter 7 is devoted to systematic techniques for reflecting on the entire modeling process.



1.5 Reflections: Iterating the Modeling Process

1.5.1 Reflection after the Definition Stage (1')

The modeling process is an ITERATIVE process. Iterations may occur at any stage. A change at the level of the identification of the purpose, however, can have drastic consequences since every subsequent step may be affected.

After the problem owner communicates his version to the modeler, the modeler will reformulate the problem in terms of purpose and context. Before going further, it is essential that the modeler goes back to the problem owner to verify if this new, redefined version of the problem is still in accordance with the problem owner's intentions.

1.5.2 Reflection after the Conceptualization Stage (2')

Identifying the right set of entities is almost impossible to do the first time right. Revisions to the conceptual model are common in each modeling process.

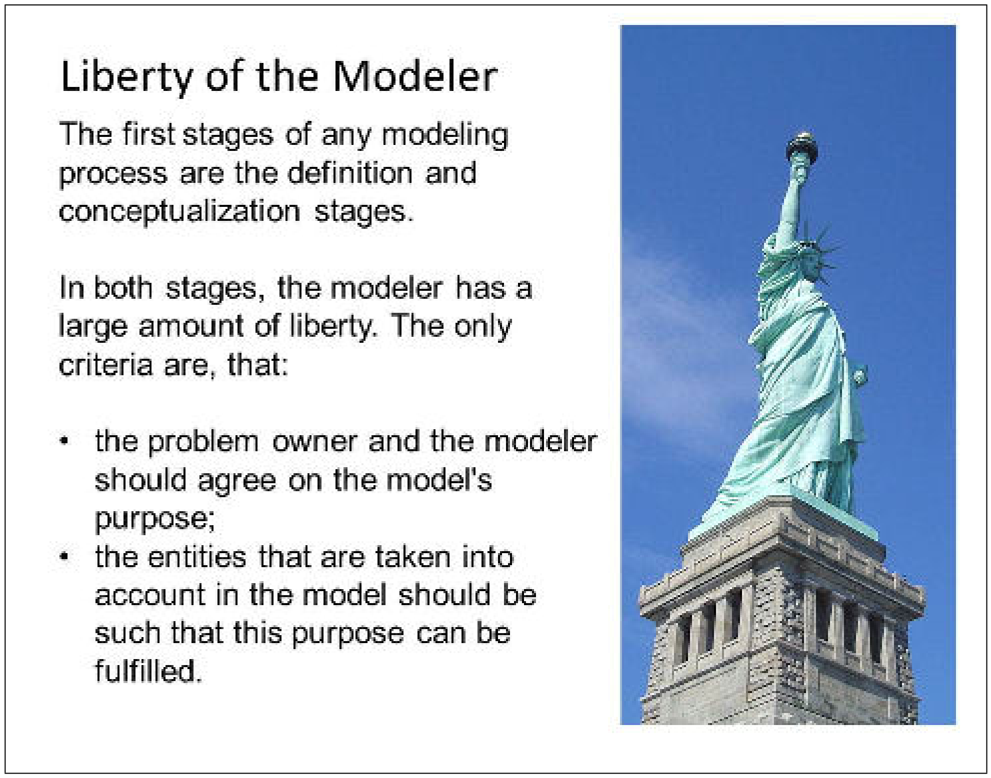
At the end of the conceptualization stage, the model cannot yet *do* anything. It cannot yet compute anything, as there are no formulas, no computed numbers, and no computer script. We have gained a fair amount of insight, though. We should check if the conceptual model reflects our intuition about the things that matter. Didn't we forget anything? Is there not too much detail to start with? Also, we should see if the relations that we identified are correct and necessary

[19] The image of a judge was taken from http://upload.wikimedia.org/wikipedia/commons/5/5b/Chitty_JW_Vanity_Fair_1885-03-28.jpg?uselang=nl

- and if not, we should correct this.

1.5.3 Reflection after the Formalization Stage (3')

During formalization[20], there will come a point that a first formal version of the model is ready. This is a good moment for a reflection. The modeler should seek arguments to support the conclusion that the formulas are correct, or that they are at least good enough for the model's purpose.



This is also called the *validation* and *verification* of the model. Validation means: checking that the model produces output that is valuable for the model's purpose, for instance: that it gives sufficiently small uncertainties. This typically involves: running the model on input data sets for which the empirical outputs are known, and see if the model reproduces these known outputs, prior to executing the model for the actual purpose at hand. Such known outputs are called GROUND TRUTH: if the model produces output that *differs* from ground truth data, we are certain that the model is *wrong*.

Verification amounts to checking the logical and mathematical consistency of the formulas. Consistency is a necessary, but not a sufficient condition for an adequate formalization.

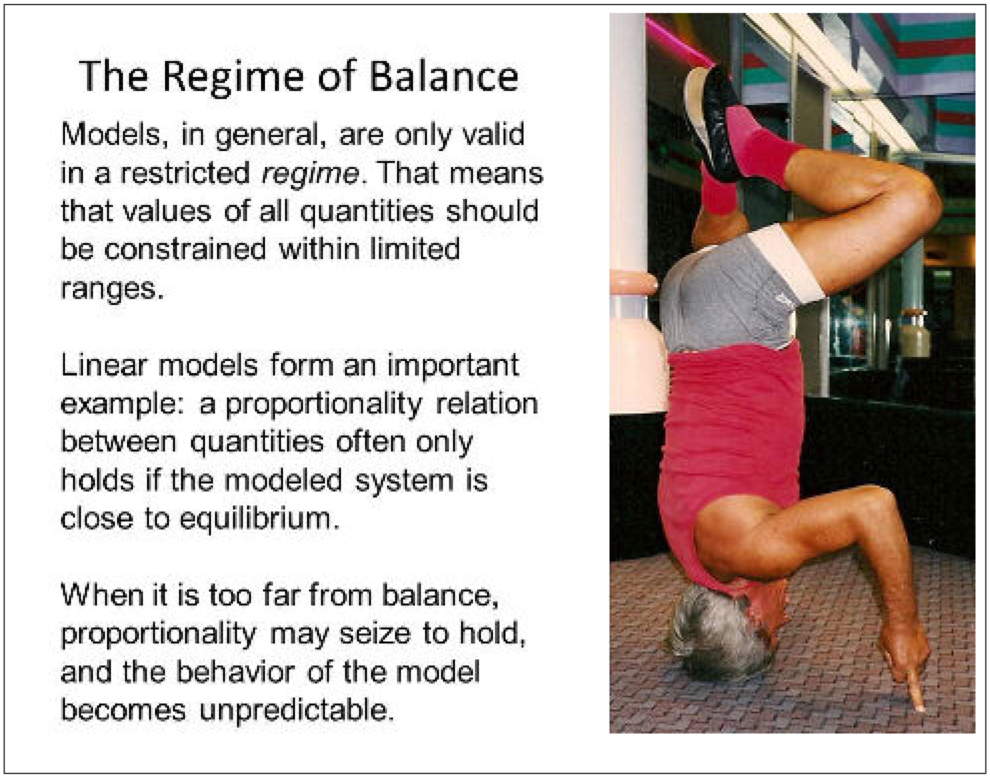
Similar, successful validation against ground truth data is necessary but not sufficient. It is always possible that the tests with ground data miss a peculiar case. Indeed, any form of NON-EXHAUSTIVE testing resembles sampling few oranges from a full batch: even if the sampled ones are sweet, there is no guarantee that there are no sour oranges in the batch ▷[25] . In Chapter 6 we learn some techniques for validation and verification.

1.5.4 Reflection after the Execution Stage (4')

The reflection on the outcome of the execution of a model should verify if these numerical outcomes fall in the REGIMES that were assumed in the various parts of the calculations.

[20] The image of the statue of liberty is taken from http://commons.wikimedia.org/wiki/Statue_of_liberty#mediaviewer/File:Majestic_Liberty.jpg

A regime is a range of values for the quantities in a model such that the model behaves similar, or a range of values for the quantities in a model such that the same set of assumptions hold.



In physical systems, for instance, there is very often a distinction between the linear regime and non-linear regimes. In the linear regime the system is near a given rest state or EQUILIBRIUM state. Think of a spring that is gently pulled by some external force. For small force the spring's elongation is proportional to the force. If the force ceases to be, the spring will return to its rest shape. For larger forces this needs not to be the case. Now suppose that, during the formalization, we have derived formulas that are only valid in the linear regime, whereas, after the execution we found that a particular spring is elongated beyond this regime: in that case, the model outcome cannot be trusted[21].

In general: conclusions obtained in one regime cannot be carried over to another regime.

1.5.5 Reflection after the Conclusion Stage (5')

The initial problem was not stated in mathematical terms, and therefore the outcome of the execution of the formal model had to be translated back into non-mathematical terms. In this last reflection stage, we focus at the question: 'did we solve the initial problem', including 'did we do a proper presentation of the results and is our interpretation of the results adequate'? This reflection inevitably requires interviewing the problem owner: (s)he is the only one with enough contextual knowledge to assess if the initial problem indeed was solved.

We summarize the modeling process in Table 1.3.

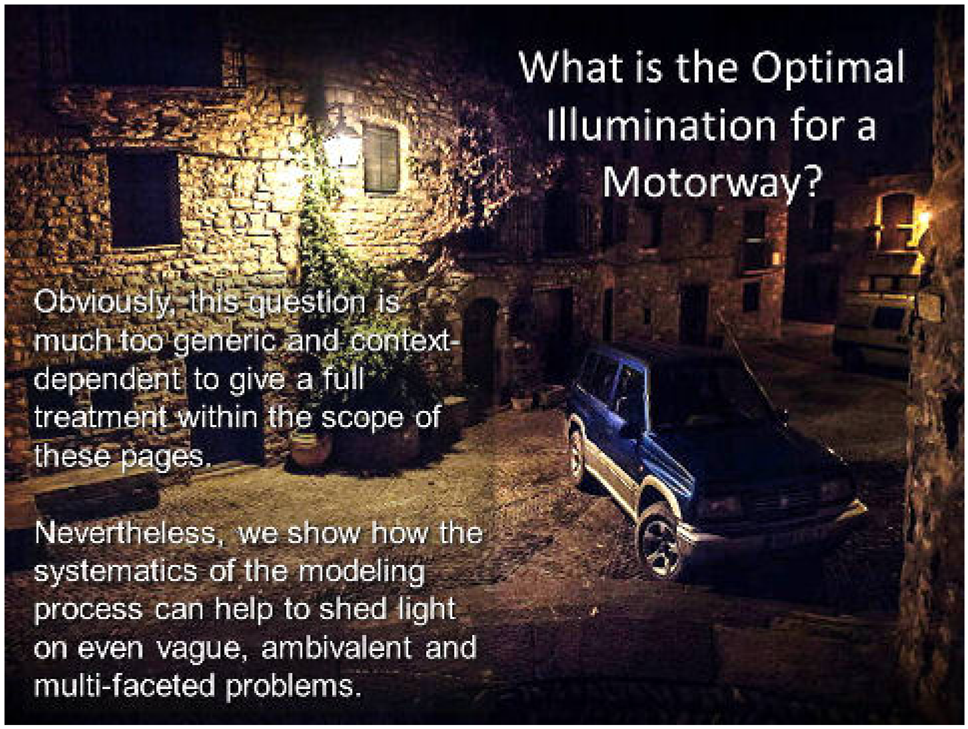
[21]The image of the equilibrist is taken from http://commons.wikimedia.org/wiki/Category:Balance#mediaviewer/File:UPSTREAM_FITNESS-5.jpg

Stage	What to do	Reflection
1. Definition	**(1.1)Formulate the Model's Purpose** • who is the problem owner? • who are the stake holders? • what is the problem context? • what purpose(s) do we have to deal with?	**(1')Assess the Plausibility of the Problem Definition** • does the problem owner recognize the redefined version of the problem? • under what conditions can the problem be considered to be solved?
2. Conceptualization	**(2.1)Identify Entities and Properties** • what are the most important entities and properties? • what are their relations? **(2.2)Identify Relations between Entities and Properties** • which entities occur in relations? • which properties occur in relations? • what is the meaning of these relations?	**(2')Assess the Plausibility of the Conceptual Model** • do we include the crucial entities and properties? • do we include the crucial relations? • is the conceptual model sufficiently simple?
3. Formalization	**(3.1)Obtain Values for Quantities** • glass box: can we propose plausible mechanisms? • black box: do we have raw data or processed data? • do we understand the assumptions that underly these data? **(3.2)Formalize Relations between Quantities** • what causal mechanisms should each relation express? • for every relation, what assumptions underly this relation? • how to express the mechanism we want to express in mathematics, logic or computer language?	**(3')Assess the Plausibility of the Formal Model** • is there ground truth data to validate the model? • can 'special cases' help test the model? • are there independent models to test our model with?
4. Execution	**(4.1)Do Operations with the Model** • what sort of operations do we do? • how do we do these operations? **(4.2)Obtain a result** • in what form does the result arrive? • when do we have sufficient results?	**(4')Assess the Plausibility of the Result** • do the results comply with assumptions? • are they valid for the purpose? • do we need to refine the model?
5. Conclusion	**(5.1)Present the Result** • what presentation styles exist for this type of result? • what presentation is adequate, given the problem owner? • does the presentation capture the essence of our result? **(5.2)Interpret the Result** • to whom should the interpretation be meaningful? • what does this person need to know? • is the interpretation valid? • does this interpretation raise any further questions?	**(5')Assess the Plausibility of the Answer to the Initial Purpose** • has the initial purpose been fulfilled? • does the model outcome further contribute to the initial problem?
Evaluate		**Learn from what you have done** • what did go really well? • how can we consolidate this? • what did not go that well? • how can we improve this?

Table 1.3: Overview of the modeling process

1.6 Example

In this section we show how the modeling process as described in Section 1.4 could be executed in practice. We consider the case of illuminating a segment of public motor way with street lamps[22]. Notice: the various stages of the modeling process are explained in more detail in forthcoming chapters. In each stage we will come back to the streetlamp example to show how the introduced techniques apply there. The elaboration in the the present section is therefore no more than a first, brief, introduction.



(1.1) Formulate the Model's Purpose

Some examples of purposes for a street lamp model could be:

- *verify:* could LED lamps be used for illuminating the road segment?
- *decide:* should we have yes or no adaptive illumination?
- *optimize:* what is the best height for lanterns for this road segment?
- *steer or control:* for real time operation of the adaptive switch on/off strategy.
- *analyze:* how do the benefits of adaptive illumination depend on the traffic flow?
- ... and many others.

We choose as initial purpose 'optimization': in terms of costs, what would be the optimal way to illuminate the road?

(2.1) Identify Entities and Properties

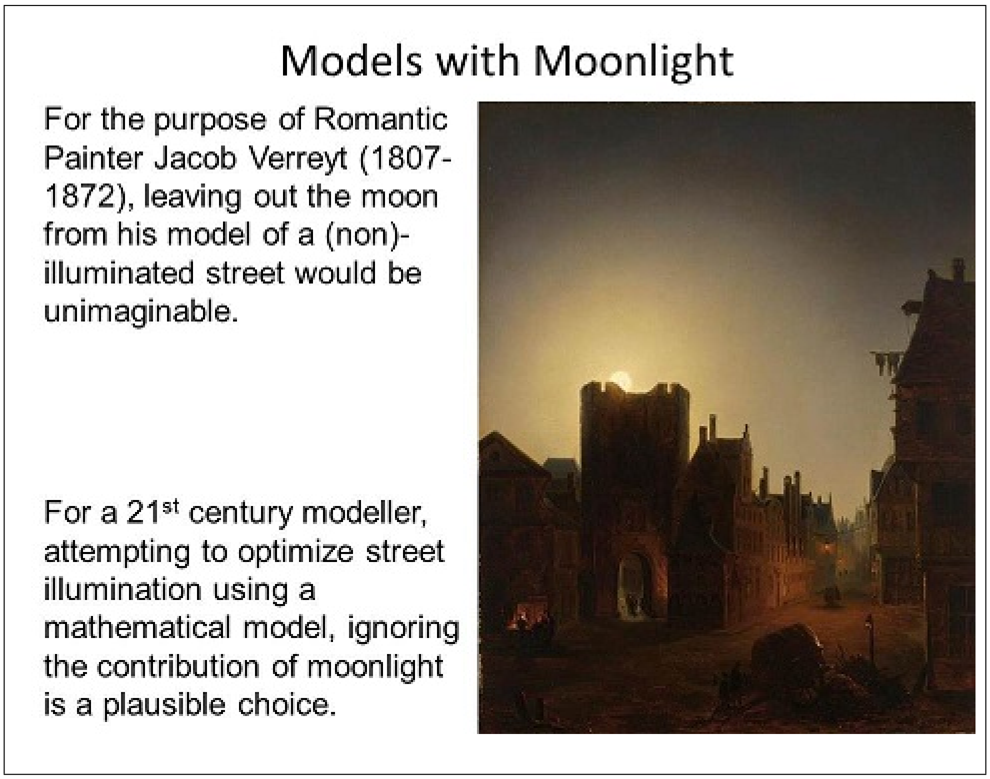
The problem owner is the provincial road agency; stake holders are car drivers, people living nearby (think of light pollution!), street workers, energy producers, insurance companies and perhaps many more.

In this stage we must decide what the SCOPE of the model will be: for instance, in case a detailed prediction of the long term financial exploitation is required, maintenance and replacement should be modeled as well. This scope gives us a first hint as to what entities should be considered, and

[22] The illustration is taken from http://commons.wikimedia.org/wiki/Category:Darkness#mediaviewer/File:4x4_(9351322825).jpg

therefore: which concepts should be present in the model.

The crucial question that should be answered is: 'with how little money can we safely illuminate the motorway?', and perhaps: 'could we obtain additional savings if street lamps are temporarily switched off during the absence of cars?' Entities involved therefore include motor traffic, drivers, street lamps, energy, and perhaps the equipment to implement adaptivity.



Next we look at the properties of these concepts that should be accounted for[23].

- *motor traffic* : The concept 'motor traffic' is complex. In order to answer the purpose for the model, however, we can capture the essential features in few properties, for instance the average time between the passage of two subsequent cars, and perhaps the average speed of the cars.
- *drivers* : The concept 'drivers' has a number of properties: some have a physical origin (what is the minimum light intensity needed to see a road marking at some given distance?; what is the maximum light intensity to prevent blinding?). Others come from psychological mechanisms: what is the minimal distance a driver needs to see an illuminated road segment in front of him in order to drive safely?
- *street lamps* : The concept 'street lamps' can be captured by the distance between subsequent lamps, their height, and their illumination power.

A more detailed elaboration is given in Section 2.5.

(2.2) Identify Relations between Entities and Properties

Relations among concepts should be made explicit. E.g., the intensity of the street lamps has a relation to the energy that is consumed, and the energy has a relation to the costs of the illumination system. Further, there is a relation between the intensity of the street lamps and the reflection on the road as perceived by the drivers - and this perceived illumination level has a relation to the maximum and minimum values we should reckon with.

(2') Assess the Plausibility of the Conceptual Model

In Chapter 2 we learn that the conceptual model is drawn in the form of a graph. Then there are

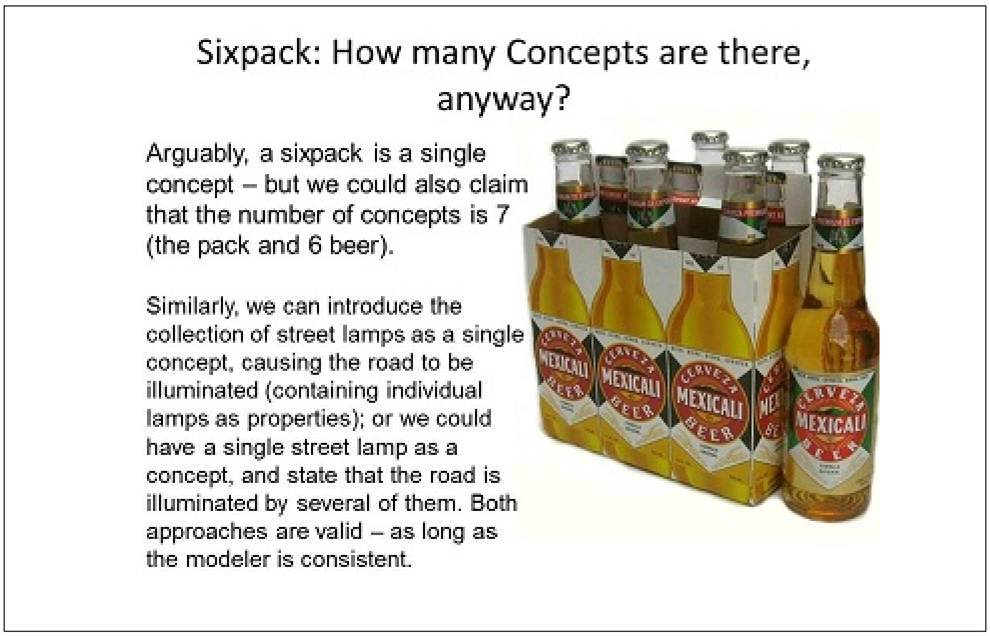
[23]The reproduction of the Jacob Verreyt painting was taken from http://commons.wikimedia.org/wiki/File:Moonlit_streetscene-Jacob_Verreyt.jpg

some early SANITY CHECKS we could perform. For instance, concepts or properties that don't engage in any relation should be taken out of the conceptual model. Also, if there are insufficient concepts or properties to express the initial purpose, the conceptual model cannot be complete.

(3.1) Obtain Values for Quantities

For the various quantities, there are different sources for their values, and hence different procedures to obtain them. Some examples:

- *motor traffic, drivers* : properties related to traffic[24] and drivers may need data gathering and aggregation. To asses the opinions of drivers about the new lighting scheme, for instance, may require a market survey including an experimental set up using a driving simulator.
- *street lamps* : properties of the street lamps are, within bounds, free for a designer to choose, so these could be used to find an optimal configuration.



(3.2) Formalize Relations between Quantities

Relations occur in many different forms, such as:

- Quantities have dimensions (energy, distance, energy / distance, time, price / distance, etc.). As we will see in Section 2.7, studying these dimensions will give help in constructing the right mathematical relations.
- Some relations will have the form of dependencies: the light intensity on a given point at the surface of the road depends on the height, the distance and the power of the lanterns. In Chapter 4, we will see how a careful, step-by-step derivation of such dependencies leads to a plausible and transparent formal model.
- Other relations are constraints: the minimum intensity in any point should not be less than a certain value, and similar for the maximum intensity. In Chapter 5 we will learn how constraints can be represented mathematically, and how they can be resolved in simple cases.

(3') Assess the Plausibility of the Formal Model

Some behaviors of the traffic - street lamps system are intuitive, and the model should reproduce these. For instance:

- if lamps are taller, the light distribution underneath will fluctuate less;
- with a given power per lantern, energy consumption should be proportional to the number of lanterns;

[24]The photograph of the six pack was taken from http://commons.wikimedia.org/wiki/File:Mexicali_Beer_6_Pack.jpg

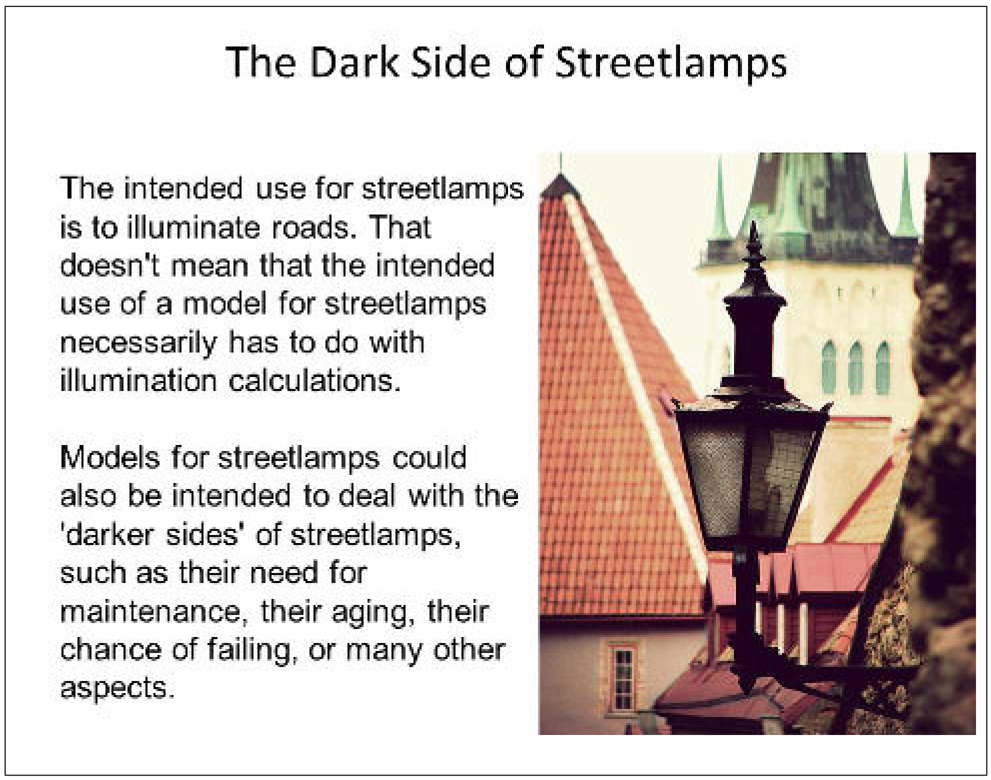
- local extrema in the illumination are likely to occur immediately below street lamps, and in the middle between adjacent street lamps (provided that these are modeled as point sources);
- if the lower limit for the intensity, needed for drivers to see road markings, decreases, energy consumption should decrease.

In general, assessing the plausibility of a formal model is called its *validation*; in Chapter 6 we learn which techniques can be used.

(4.1) Do Operations with the Model

Required operations depend on the purpose[25]. Some examples:

- *conditional prediction* : given a certain traffic density and given a configuration of street lamps, what will be the power consumption for adaptive street lighting;
- *optimization* : assuming non-adaptive illumination only, what will be the optimal configuration of street lamps (in terms of height, distance and power) such that energy costs are minimal;
- *optimization* : given a certain traffic density, what will then be the optimal configuration of street lamps such that power consumption is minimal for adaptive street lighting;
- *analysis* : given a street lamp configuration and adaptive illumination scheme, how do energy savings depend on traffic distributions?
- *simulation* : given a street lamp configuration and a scheme for adaptive lighting, would drivers appreciate this scheme?



(4.2) Obtain a Result Depending on the operation, results could be:

- *conditional prediction* : an average amount of power consumption;
- *optimization* (for non-adaptive schemes): height, distance and power per street lamp for an optimal configuration;
- *optimization* (for adaptive schemes, for a given average traffic density): height, distance and power per street lamp for an optimal configuration;

[25] The photograph of a worn-out streetlamp in Talinn was taken from http://commons.wikimedia.org/wiki/File:Talinn-Aare_Piiroja_-Vintage_series.jpg?uselang=nl

- *analysis* : a table with the energy savings listed for a number of average traffic densities (cars/hour);
- *simulation* : a collection of data, obtained from interviewing drivers that have made a test run in the simulator, regarding their opinions on adaptive lighting.

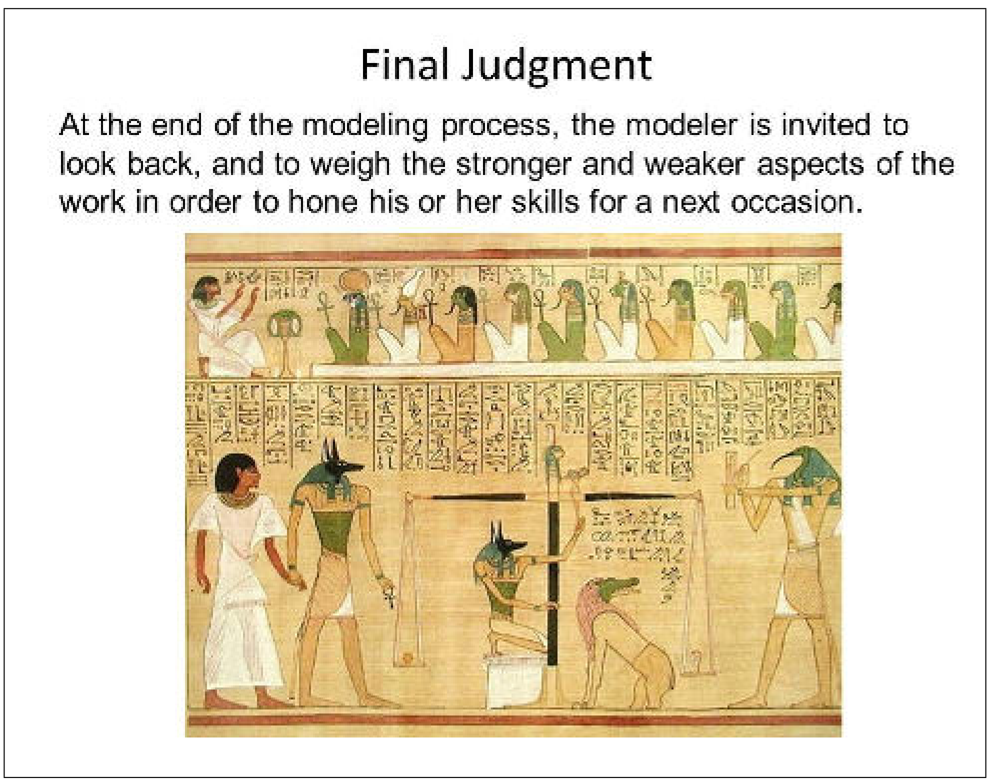
(4') Assess the Plausibility of the Result

As an example we consider optimization: the power consumption should be calculated for values of height, distance between street lamps and power per lamp for a range of values near the optimal values to verify if the optimum is *stable*. That is: if a small change in one of the quantities would cause a dramatic change of the calculated power consumption, we should mistrust the relevance of the outcome (see further Chapter 6).

(5.1) Present the Result

In order to convince the problem owner and other stake holders, a mere presentation of the data (height, distance, power) is certainly insufficient. A presentation might consist of graphs showing the energy consumption as a function of each of these quantities.

(5.2) Interpret the Result



An interpretation should attempt to account for (some of) the approximations and assumptions used in the model. It should also try to provide some intuition behind the model outcome: 'in hindsight, the result could be understood because ...'. For instance, it might try to explain why the optimal height of the street lamps in an adaptive lighting scenario is less than that of a standard, non adaptive lighting scenario.

(5') Assess the Plausibility of the Answer to the Initial Problem

The only way to do this is by actual conversation with the original problem owner.

Learn from what you have done [26]

A straightforward approach to modeling the illumination distribution over a road surface due to

[26] The image of the Egyptian, jackal-headed god Anubis, weighting the heart of the deceased (individual clad in white, left) and compare its weight to that of a feather to come to a final assessment 'sufficient' or 'insufficient', was taken from http://commons.wikimedia.org/wiki/Category:Papyrus_of_Hunefer#mediaviewer/File:BD_Hunefer_cropped_1.jpg

street lamps could take both width and length of the road into account, that is, consider it a 2D problem. Perhaps this is not necessary, though. Maybe the variation of the light intensity perpendicular to the road axis can be ignored in comparison to the axial direction. That is: we could interpret the outcome of the model as ground truth data, and approximate these with a simpler 1-D model. If the accuracy of the approximation is sufficient, we might have done our optimization calculations much more efficiently. This is good to know for a next problem involving road illumination.

1.7 Mathematical Tools: Functions of Two Variables

In this chapter models have been discussed. Models in which, amongst other things, quantities play a role. Moreover, relations between quantities have to be defined in the modeling process. Many relations are expressed as functions. Sometimes one has to maximize or minimize a function. For all the above things we use the analysis of real functions. Elementary calculus, however, does not elaborate on functions of several variables. Yet, in many models functional relations involving multiple quantities play a role.

Resistance is Functional

Perhaps the best known formula in physics is Ohm's law, V=IR. Many people call this an equation – but that assumes that there is a quantity with unknown value that should be obtained.

In many applications, it is more natural to regard the relation V = IR as a function of two variables. This can be done in three ways:

$V = f_V(I,R) = I \times R$

$I = f_I(R,V) = V / R$

$R = f_R(V,I) = V / I$

An example is the problem, introduced in Figure 'Shining Light on a Narrow Passage', where the length of the red and green period of traffic lights has to be set. The function that gives the average waiting time depends on multiple quantities, for three of which a value has to be chosen. Therefore we introduce the essentials of functions of several variables here.

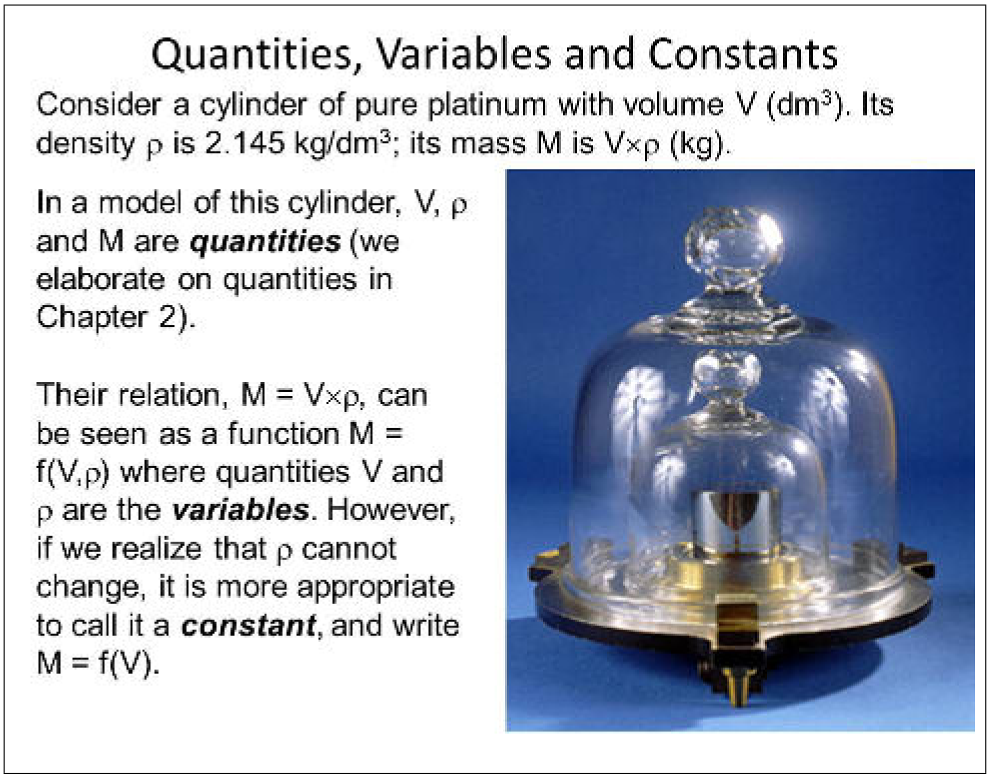
1.7.1 Functions of Several Variables

In the core course Calculus functions $f(x)$ of one variable x have been discussed. The domain and range of these functions where subsets of the real numbers. Now we need functions of several variables[27]. First we discuss functions of two variables.

[27] The portrait of Georg (not: Gerog!) Ohm is taken from http://commons.wikimedia.org/wiki/File:Gerog_Ohm.jpg?uselang=nl

A FUNCTION OF TWO VARIABLES is a rule that assigns a real number $f(x, y)$ to each ordered pair of real numbers (x, y) in the domain $D \subset \mathbb{R}^2$ of the function. As for functions of one variable, the DOMAIN CONVENTION is that the domain of a function of two variables is the largest set of pairs (x, y) for which the function $f(x, y)$ can be evaluated, unless the domain is explicitly given by a smaller set, due to constraints coming from the modeled system. Analogously one can define a function of n variables.

Example 1: The function $f(x, y) = \sqrt{x^2 + y^2}$ gives the distance of a point to the origin.



Example 2: In Section 1.2.2 we mentioned the traffic lights problem. Here the length of the red and green period of traffic lights have to be set. The following pattern for the traffic lights can be chosen.

direction I	$R_0\ G_1\ R_0\ R_1$
direction II	$R_0\ R_2\ R_0\ G_2$

A purpose might be to minimize the average waiting time for a car. One needs a model for the traffic flow to estimate the average waiting time for a car. Suppose that f_i is the flow of traffic in direction i in number of cars per minute ($i = 1, 2$) and f_0 is the number of cars per minute that can pass the part of the road with one lane. Furthermore, let R_0 is the time both traffic lights are red simultaneously and R_i is the time the traffic light in direction i is red and the other traffic light is green ($i = 1, 2$; $R_1 = G_2$; $R_2 = G_1$).

For the model we assume that the cars arrive in a deterministic pattern with a constant inter arrival time (see also the discussion in Section 1.2.2). Furthermore we assume that the variables above have such values that there is no queue anymore at the moment that the traffic light turns red again. It can be shown that in this case the average waiting time F is

$$F = \frac{f_0}{2(f_1 + f_2)} \cdot \frac{\frac{f_1}{f_0 - f_1}(2R_0 + R_1)^2 + \frac{f_2}{f_0 - f_2}(2R_0 + R_2)^2}{2R_0 + R_1 + R_2}. \tag{1.2}$$

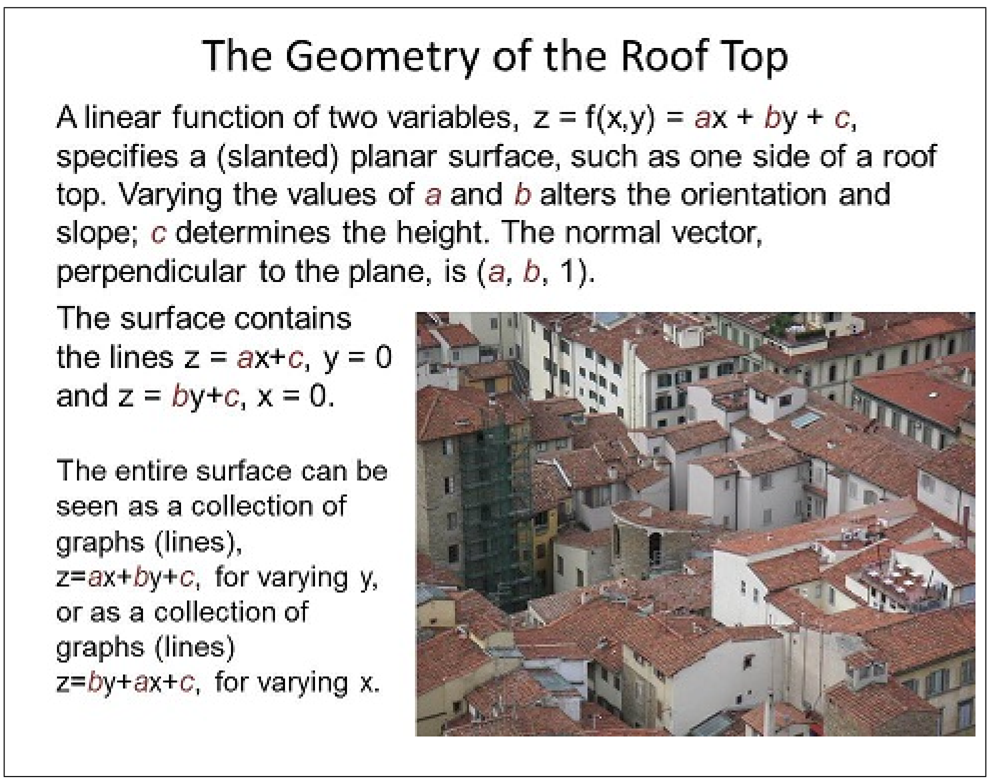
The derivation of this formula is give in Appendix **??**.

This function can be seen as a function of 6 VARIABLES[28]. The frequencies f_0, f_1 and f_2,

[28]The photograph of the standard kilogram was taken from http://commons.wikimedia.org/wiki/Category:Kilogram#mediaviewer/File:National_prototype_kilogram_K20_replica.jpg

however, are given or can be measured for the given road. The values of these quantities won't change; f_0, f_1 and f_2 are considered to be CONSTANTS. As follows: suppose that for the specific road 800 cars arrive per hour in one direction and 300 cars in the other direction. So, $f_1 = 40/3$ cars per minute and $f_2 = 5$ cars per minute. Furthermore, we assume that 40 cars can pass the blocked part of the road in one minute in one direction if the traffic light is green ($f_0 = 40$ cars per minute). With f_0, f_1, and f_2 being constant, F has been reduced to a function of three variables (R_0, R_1 and R_2, in minutes). If we further suppose that the length of the blocked part of the road is 500 m and the speed of the cars is 20 km/hour, then R_0 should be at least 1.5 minutes. If we choose $R_0 = 1.5$ (=another constant), then F reduces further to a function of only two variables. The average waiting time (in minutes) is then equal to

$$F = \frac{6}{77} \cdot \frac{7(3+R_1)^2 + 2(3+R_2)^2}{3+R_1+R_2}. \tag{1.3}$$



The function of Expression 1.3 is a function of two variables (R_1 and R_2). This function should be minimized. If one deals with a function of one variable, the derivative (if possible) is taken in order to find extremal values. However, now we have two variables. To find extremal values of functions of several variables one needs so-called *partial derivatives*. These will be introduced in Section 1.7.3.

1.7.2 Graphs and Contour Plots

We are familiar with drawing a graph of a function of one variable. As we know this is a two-dimensional figure. Drawing the graph of a function of two variables results in a three dimensional construction. Using smart projections one can visualize such a construction as an (two-dimensional) image. In the elementary course Calculus we have seen this for some simple functions. A plane[29] in three-dimensional space can be seen as a function of two variables. Consider for example the function $f(x,y) = (-4x - 5y + 32)/6$. If we write $z = (-4x - 5y + 32)/6$, then we recognize the plane
$4x + 5y + 6z = 32$ with normal vector $(4, 5, 6)$. Here we use $()$ to denote vectors. In Section

[29] The fotograph of rooftops was taken from http://www.rgbstock.nl/photo/n7LahDA/Rooftops+5

?? notational issues are discussed further. Go to this link to experiment with graphs for linear functions of two variables.

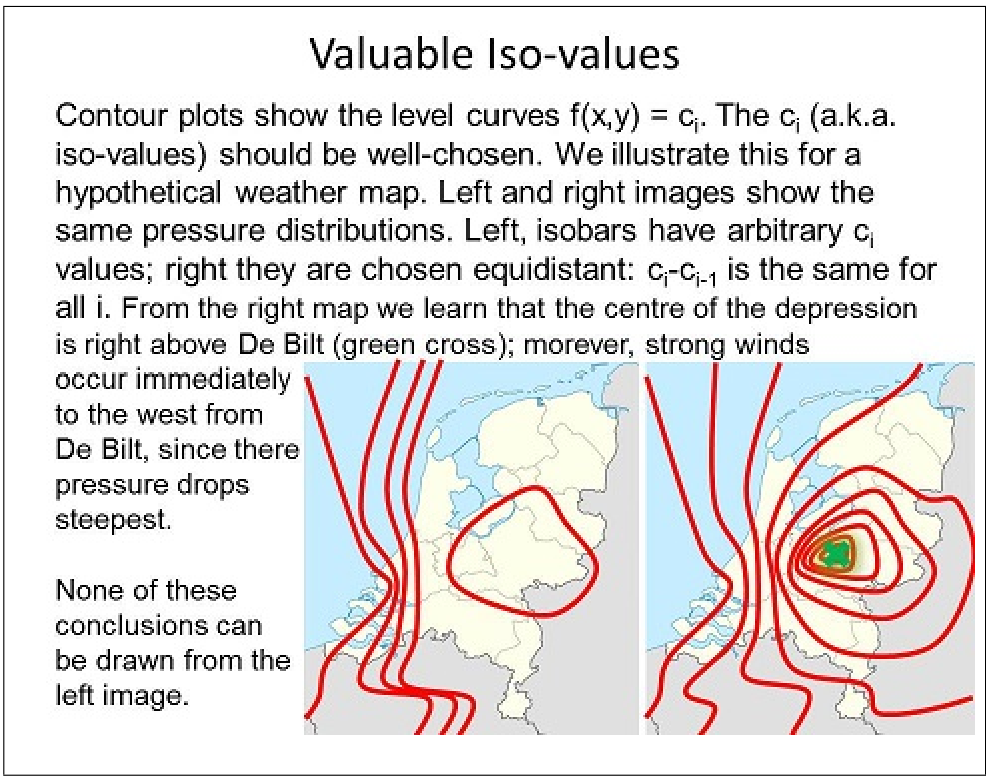
A graph of the function $f(x,y) = \sqrt{x^2+y^2}$ of Example 1 in Section 1.7.1 can be found at this link ▷26

A graph of the function given in Expression 1.2 of Example 2 in Section 1.7.1 is found at this link ▷27 ; the values of the constants f_0, f_1, f_2, and R_0 can be adjusted at will.

One can imagine that the visualization of a graph of a function of more than two variables becomes more or less impossible.

Another way to visualize a function of two variables in a two-dimensional picture is a CONTOUR PLOT[30]. A contour plot is a collection of level curves. A LEVEL CURVE of the function $f(x,y)$ is the two-dimensional curve defined by $f(x,y) = c$, for some constant c. The level curves of the function $f(x,y) = \sqrt{x^2+y^2}$ are circles. If we would add a constant, say z_{shift} to f, the level curves would still be circles, although for different values of c. By going to this link, one can experiment by varying values for c, in combination with values for z_{shift}. Notice that, for any given function f, level curves don't necessarily exist for all values of c.

The contour plot of the the function F of Expression 1.2 can be found by going to this link. Here, the values for the constants f_0, f_1, f_2, and R_0 can again be adjusted.



1.7.3 Partial Derivatives

Recall that for a function $f(x)$ of one variable the *derivative function* $f'(x)$ is defined as

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h},$$

for values of x for which the limit exists. At a point $x = a$ the value $f'(a)$ is the instantaneous rate of change of the function $f(x)$ with respect to x.

For a function of two variables, *partial derivatives* can be defined.

[30]The map of the Netherlands was taken from http://commons.wikimedia.org/wiki/File:Netherlands_location_map.svg?uselang=nl

The PARTIAL DERIVATIVE $\frac{\partial f}{\partial x}$ of $f(x,y)$ with respect to x is defined by

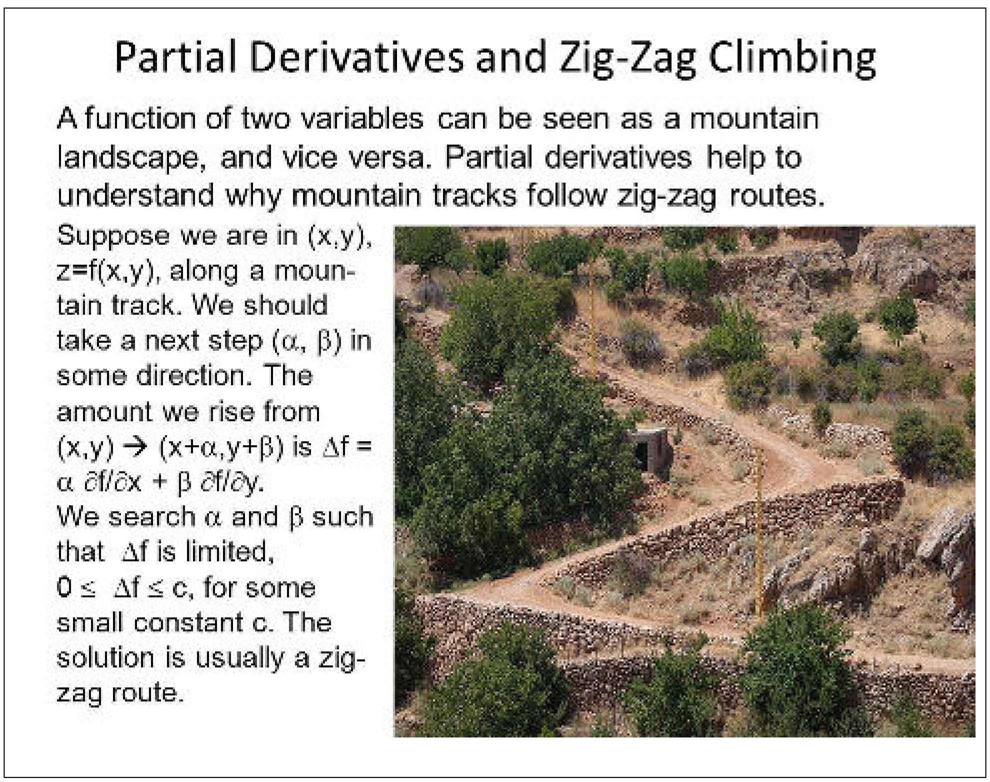
$$\frac{\partial f}{\partial x}(x,y) = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}, \tag{1.4}$$

for values of (x,y) for which the limit exists.
The *partial derivative* $\frac{\partial f}{\partial y}$ *of* $f(x,y)$ *with respect to* y is defined by

$$\frac{\partial f}{\partial y}(x,y) = \lim_{h\to 0} \frac{f(x,y+h) - f(x,y)}{h}, \tag{1.5}$$

for values of (x,y) for which the limit exists.
Several notations are used for the partial derivative with respect to x. We have

$$\frac{\partial f}{\partial x}(x,y) = \frac{\partial}{\partial x} f(x,y) = D_x f(x,y) = f_x(x,y) = \frac{\partial f}{\partial x} = \frac{\partial z}{\partial x} = \frac{\partial z}{\partial x}(x,y). \tag{1.6}$$



As for functions of one variable a partial derivative gives the instantaneous rate of change. The instantaneous rate of change in the direction of x at the point (a,b) is given by $\frac{\partial f}{\partial x}(a,b)$. It is just as easy (or difficult)[31] to compute partial derivatives as to compute derivatives of functions of one variable. If one computes the partial derivative of the function $f(x,y)$ with respect to x, just compute the ordinary derivative with respect to x, while treating y as a constant.

Example 1: The partial derivative $\frac{\partial f}{\partial x}$ of the function $f(x,y) = \sqrt{x^2+y^2}$ is equal to $\frac{1}{2} \cdot \frac{1}{\sqrt{x^2+y^2}} \cdot 2x$.

Example 2: The partial derivative of the function given in Expression 1.3 with respect to R_1 is equal to

$$\frac{\partial F}{R_1} = \frac{6}{77} \cdot \frac{14(3+R_1)(3+R_1+R_2) - 7(3+R_1)^2 - 2(3+R_2)^2}{(3+R_1+R_2)^2}. \tag{1.7}$$

[31] The photograph of a Lebanese mountain track was taken from http://commons.wikimedia.org/wiki/File:Zig-zag_lebanese_mountain_road.jpg?uselang=nl

1.7.4 Tangent Planes

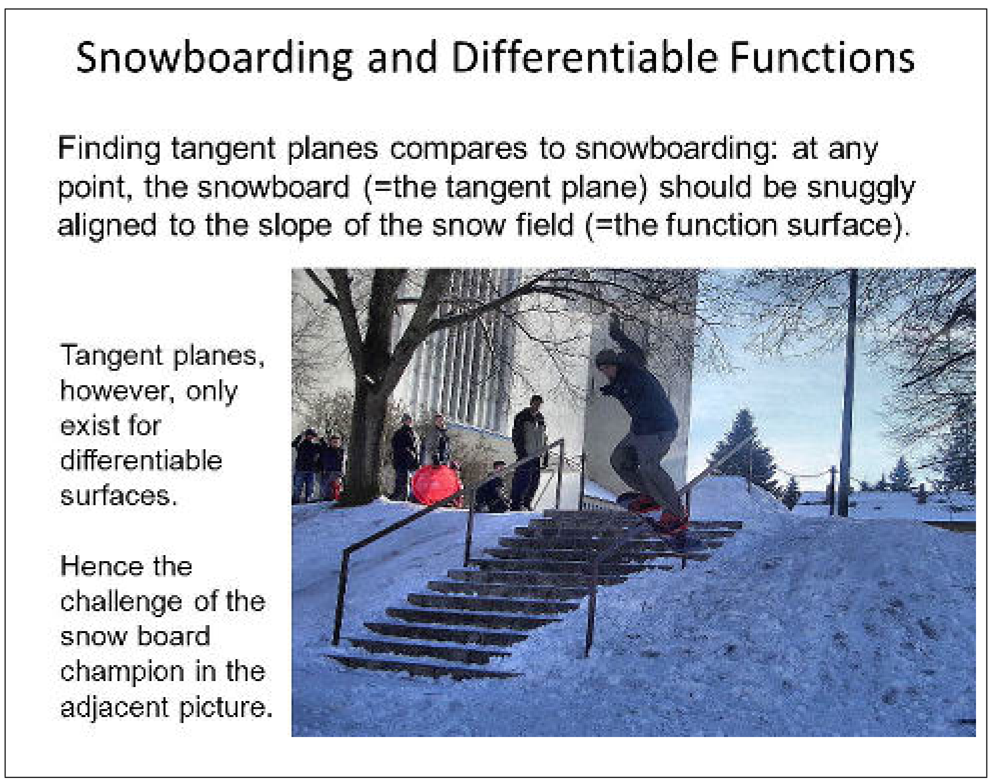
For a function of one variable the *tangent line* to the graph $y = f(x)$ at $x = a$ is given by the equation

$$f'(a)(x-a) - (y - f(a) = 0. \tag{1.8}$$

Using this the *linear approximation* of the function $f(x)$ can be defined. The linear approximation of $f(x)$ at $x = a$ is the function

$$L(x) = f'(a)(x-a) + f(a). \tag{1.9}$$

Likewise, there exists a linear approximation to a function of two variables[32] $f(x, y)$ at a point (a, b). Suppose that $f(x, y)$ has continuous first partial derivatives at (a, b). The equation of the TANGENT PLANE to the graph $z = f(x, y)$ at (a, b) is given by $f_x(a, b)(x - a) + f_y(a, b)(y - b) - (z - f(a, b)) = 0$.



We conclude that a normal vector to the tangent plane is equal to $(f_x(a, b), f_y(a, b), -1)$.

By analogy with Expression 1.9, the LINEAR APPROXIMATION $L(x, y)$ of the function $f(x, y)$ at the point (a, b) is defined as $L(x, y) = f_x(a, b)(x - a) + f_y(a, b)(y - b) + f(a, b)$.

Example 1: The tangent plane to $z = f(x, y)$ with $f(x, y) = \sqrt{x^2 + y^2}$ at $(4, 3)$ is equal to $\frac{4}{5}(x-4) + \frac{3}{5}(y-3) - (z-5) = 0$.

This can be written as $4x + 3y - 5z = 0$.

Example 2: The tangent plane to $z = F(R_1, R_2)$, where F is the function of Expression 1.3 at the point $(-\frac{5}{3}, \frac{5}{3})$ is equal to $z = \frac{16}{11}$.
Note that this is a horizontal plane. Both partial derivatives are equal to zero.

To experiment with tangent planes and normal vectors, click here.

From now on, when we encounter a relation between 3 quantities, where one quantity depends on the other two, we can describe this as a function of two variables, and we have seen ways to visualise

[32]The photograph of the snowboard trick is taken from http://commons.wikimedia.org/wiki/Snowboard#mediaviewer/File:Backside_Boardslide.jpg

the functional dependency. In Section 2.8 we learn how we can do more sophisticated things with functions of multiple variables, such as finding extremes, which is useful for optimization models.

1.8 Summary

- A model can only be meaningful with a clearly defined *purpose*;
  - purposes come from *research* (aim: to produce knowledge or understanding) or *design* (aim: to create or add value)
  - purposes are: *explanation*, *prediction* (two cases!), *compression*, *abstraction*, *unification*, *communication*, *documentation*, *analysis*, *verification*, *exploration*, *decision*, *optimization*, *specification*, *realization*, *steering and control*. See Table 1.1.
- Modeling approaches can be distinguished on a number of *dimensions*:
  - *static - dynamic*: does time play a role?
  - *continuous - sampled - discrete*: does the modeled system involve 'counting' or 'measuring'?
  - *numeric - symbolic*: do results follow from operations on numbers or expressions?
  - *geometric - non-geometric*: do features from 2D or 3D space play a role?
  - *deterministic - stochastic*: does probability play a role?
  - *calculating - reasoning*: does the purpose rely on numbers or on propositions?
  - *black box - glass box*: does the model start from data or from mechanisms?
- Modeling is a process involving 5 stages, each stage consisting of two activities (=two subsequent blocks in Table 1.3) and a reflection:
  - *define*: establish the purpose
  - *conceptualize*: devise a representation of the modeled system in terms of concepts, properties and relations
  - *formalize*: devise a representation of the conceptual relations in terms of mathematical expressions
  - *execute*: perform the appropriate operations (often involves running a computer program)
  - *conclude*: devise an adequate presentation and interpretion
- Functions of several variables;
  - A *function of two variables* $f(x, y)$ is a rule that assigns a real number $f(x, y)$ to each ordered pair of real numbers (x, y) in the domain $D \subset \mathbb{R}^2$ of the function
  - A graph of a function of two variables results in a three dimensional figure
  - A *level curve* of the function $f(x, y)$ is the two-dimensional curve defined by $f(x, y) = c$, for some constant c
  - A *contour plot* is the collection of a number of level curves

- The *partial derivative* of the function $f(x, y)$ with respect to x is the ordinary derivative with respect to x, while treating y as a constant
- The equation of the *tangent plane* to $z = f(x, y)$ at (a, b) can be found by use of the partial derivatives
- The *linear approximation* $L(x, y)$ of the function $f(x, y)$ at the point (a, b) can be found by use of the partial derivatives

1.9 Learning goals

1.9.1 Knowledge

You should be able to name at least 6 purposes for models; for all purposes, introduced in Section 1.2, you should know their differences and main characteristics. You should be able to name all dimensions introduced in Section 1.3. You should be able to name all stages in the modeling process as introduced in Section 1.4, and for each stage, you should know which activities of the modeling process take place in that stage. You should comprehend the meaning of all the terms introduced in this chapter, as they appear in the index. You should possess a working knowledge of the analysis of real functions of several variables as explained in Section 1.7 and of the relevant sections in the calculus book of either *Adams* or *Smith & Minton*:

Adams: §12.1 until 'Using maple graphics', §12.3 until 'Distance from a point to a surface', §12.6 until Definition 5 (so only Linear Approximations).

Smith & Minton: §12.1 until 'Density plots', §12.3 until 'Higher-order partial derivatives', §12.4 until 'Increments and differentials'.

1.9.2 Skills

In this section, with 'problem' we mean: a problem that does not require domain-specific knowledge exceeding your present knowledge.

For a model, developed in the context of a given problem, given with sufficient detail, you should be able to determine its purpose(s). For a given problem domain, you should be able to identify several possible purposes that models could have. For a model, given with sufficient detail, you should be able to determine each of its dimensions as introduced in Section 1.3. For a given problem domain with a given model with given purpose, you should be able to suggest a global direction for a modification to the model to satisfy an alternative purpose. For a given problem, you should be able to set up a proposal for approach according to the modeling process.

For a given function of two variables, you should be able to find the domain and range, to draw graphs and contour plots and be able to interpret the plots. You should be able to compute the derivatives of a function of two variables, compute the tangent plane and the linear approximation of the function at a given point.

1.9.3 Attitude

When confronted with a problem that might benefit from a formal approach, you should consider to use a model. When approaching a problem by using a model, you should have the attitude to first formulate a purpose. When devising a model, you should consider various of the modeling dimensions from Section 1.3 before you choose a definitive route. When approaching a problem by means of a model, you should be inclined to follow the modeling process as explained in Section 1.4. When dealing with quantitative dependencies among quantities, you should consider applying the calculus of functions of several variables.

1.10 Questions

1. What is the meaning of 'reliable' in models?
2. Without using a model, how do people predict the weather? Are you sure that there is no model involved?
3. In your own words, explain the difference between the various types of predictions we discuss.
4. In your own words, explain the difference between documentation and communication.
5. In purpose 'steering or control' we talk about 'a human in the loop'. What does 'loop' mean here?
6. Explain in your own words the difference between purposes 'abstraction' and 'unification'.
7. Explain in your own words the difference between purposes 'specification' and 'realization'.
8. Explain in your own words what 'continuity' means.
9. In your own words, give the meaning of 'sampling'.
10. What is the relation between 'sampling' and 'reconstruction'.
11. When is 'absolute certainty' difficult to attain in numeric models, when is it no problem?
12. In your own words, explain what an angle is.
13. We say that 'angle is a special kind of distance'. What do we mean by that?
14. In your own words, what is 'emergent'?
15. We say 'logical expressions [...] resemble arithmetic expressions'. What do we mean by that?
16. In your own words, explain the difference between black box and glass box.
17. What is the difference between conceptualization and formalization?
18. Reflection on a modeling stage asks for plausibility, not for correctness. Why?
19. Give some similarities and differences between 'assumption' and 'hypothesis'.
20. What are the benefits of a log-log scale?
21. Both in step 5 (obtaining values for quantities) and in step 9 (obtaining a result) in Table 1.3, we obtain values. What is the difference between these steps?

22. In the example of the modeling proces in Section 1.6, why should it be that the optimal height of the street lamps in an adaptive lighting scenario is less than that of a standard, non adaptive lighting scenario?

23. In the example of the modeling proces in Section 1.6, why should it be that for a given traffic distribution, savings will be less if street lamps are further apart?

24. In the example of the modeling proces in Section 1.6, why could it be that the optimal height of the street lamps in an adaptive lighting scenario is different from that of a standard, non adaptive lighting scenario?

25. Why is the lower left most cell in Table 1.3 empty?

1.11 Exercises

1. Think of an example, analogous to the 'can-we-afford-to-buy-this-book-advisor model' where a prediction on the basis of mathematics at first sight seems fully reliable; next analyse under what assumptions the model holds, and analyse a scenario where the model still could go wrong.

2. (*) Consider Keppler's work based on Tycho Brahe's results, and Newton's work based on Keppler's work. In both cases you could say that X's model explains Y's results.

   (a) For X being (Keppler, Newton) and Y being (Brahe, Keppler), formulate what explains what;

   (b) Give an argument why these two forms of explanation are similar;

   (c) Give an argument in what respect the two forms of explanation differ.

3. We discussed 5 models for the solar system to illustrate various purposes for models, used in scientific research. Give another example of a system for which you give at least 3 very different models, and discuss the purposes of these models.

4. We claim that the purpose 'explanation' is to a large extent a social construct: it depends on the willingness of some community whether or not an explanation is acceptable. Give some examples of explanations of phenomena that are acceptable in one community, but not acceptable in another community.

5. Give an example where an explanation can be objectively shown to be wrong.

6. Give an example of a modeling situation with purpose 'explanation' without the possibility of 'prediction', and one with 'prediction' without the possibility of 'explanation'.

7. We discussed various purposes for models; Table 1.1 gives a summary. Not all purposes are independent, in the sense that fulfilling one purpose sometimes needs an other purpose to be fulfilled as well. Find examples of pairs of purposes where one purpose implies another purpose.

8. We explain linear *inter*polation in the text.

   (a) What is linear *extra*polation? (If necessary, look up the answer - but (re-)formulate it in your own terms).

   (b) Analogous to Expression 1.1, give a formula for linear extrapolation.

   (c) Think of an example where extrapolation is the answer to a modeling purpose.

   (d) Generally, extrapolation is thought to be riskier than interpolation. Give a reason.

9. The discussion of the black box model for birds mass and longevity did not start from a problem. Try to think of a problem where the subsequent models (inspiration model: plotting a graph; compression model: fitting the data points with a formula; prediction model: using the formula to say something about the expected longevity of a new bird species with given mass) could help to provide a solution.

10. Consider the example with bird's data from Section 1.3.8.

(a) Suppose that there are two sub species of blackbirds that accidentally were not distinguished when compiling the table with average masses and average longevities. What could be a worst case consequence for the model?

(b) Draw a general conclusion from this example with respect to the validity of averaging.

11. In 1.4.5, we state that a weather model where the output consists of numbers is unsuitable for the problem owner, for the stakeholders, and for the problem context. Give examples of each three, and explain why a purely numerical output is insufficient for each of the three.

12. We make an inventory of some 15 purposes in Table 1.1. Consider one domain (examples of 'domain' in this context are the weather, the national economy, public transport, a treatment for some disease, transport phenomena, ...), and give for each purpose a plausible problem. You are not supposed to *solve* these problems, of course!

13. Apart from the purposes from Table 1.1, we discuss 7 dimensions that can help distinguishing modeling approaches. One of these has three values (discrete - sampled - continuous); all others have 2. All in all this gives $2^6 \times 3 \times 15 = 2880$ combinations. Find some (at least 3) examples of combinations of modeling purpose and modeling approach that are unlikely, and explain why these are unlikely.

14. In assignment 13, we propose 2880 combinations of modeling purposes and modeling approaches. Answer the following questions:

    (a) Pick one combination and give a casus (=a domain and a problem in this domain) for which this combination is fitting. The casus should be so concrete that we can verify that indeed this combination of purpose and each of the dimensions applies.

    (b) Change one of the dimensions and do the same.

    (c) Change the purpose and do the same.

15. In secondary school (in Holland: 'VWO') final exams physics, there are often assignments where 'something' has to be calculated. Find two of such assignments in different chapters of physics (chapters in physics are, for instance, mechanics, light, atomic physics, heat, electricity and magnetism, ...); for each answer the following questions:

    (a) Assume that this assignment is a model to help solving a problem. Propose a problem for which this assignment could be (part of) the model.

    (b) Using the purposes in Table 1.1, identify the purpose of this model.

    (c) Using the dimensions, characterise the approach.

16. Our modeling process contains both a conceptualization, a formalization and an execution stage. Think of a problem that is approached by means of a model, containing a conceptual model, where the formalization stage is skipped.

17. In Section 1.4.5 we give an example of an interpretation. Find another example yourself.

Exercises concerning Section 1.7

Adams: §12.1: 4, 6, 12, 14, 20, 23, 27, 28, 45 (? ACCEL ?), 48 (? ACCEL ?); §12.3: 2, 7, 13, 16; §12.6: 1, 3.

Smith and Minton: §12.1: 3, 4, 7(a), 7(b), 16, 21 (? ACCEL ?), 27, 35 (?ACCEL?), 53, 54; §12.3: 1, 3, 19; §12.4: 1, 3, 7, 8.

Chapter 2

The Art of Omitting

'Making models is like playing golf: the holes cause the excitement'

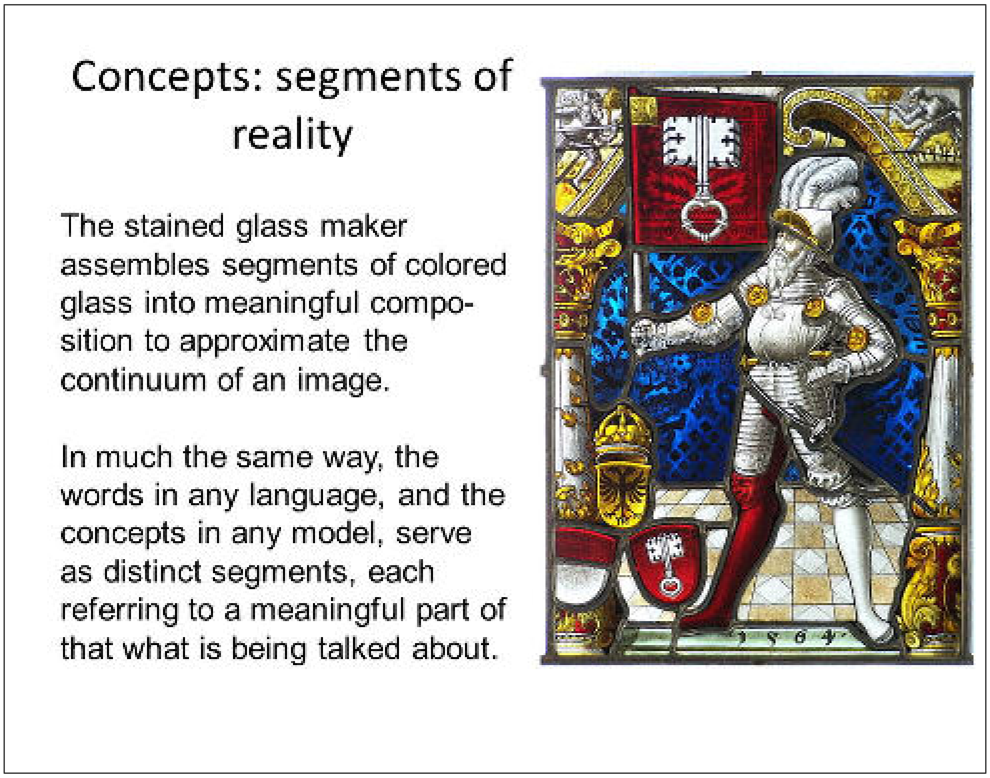
When the Spanish troops, headed by notorious conquistador Herman Cortez, invaded the Mid America's in the 16th century, natives were shocked. Not only did they suffer devastating losses because of the military superiority of the invaders, they also had to overcome their fear being confronted with mythical creatures from hell: horrendous monsters, galloping on four legs and using two further extremities to shoot arrows. Had they only realized that these creatures were not single entities, but rather human soldiers on horseback, as some found out after a horseman fell from his saddle, their panic might have been less. They were not to blame ▷[28], of course: horse does not occur as an indigenous animal in America. What is really interesting, is that the Indians failed to see rider and animal as two separate entities.

2.1 The Conceptual Model

The ability to see entities as separate and DISJOINT, to some extent, is inborn. Young children, before the age of language, focuss their attention to distinguish parts of the environment, taking subsequent samples rather than having their gaze move in a smooth and featureless fashion. One second here, the next second there. This is a first form of SEGMENTATION. When words are learned, this takes on a next stage. Words invite to seeing the world to consist of isolated objects. 'Nose' refers to a different part of the face than 'mouth'. These distinctions are meaningful: a nose has different purpose than a mouth (you should not put soup in your nose), and words help to differentiate. There are no separate words for the left part of the nose and the right part of the nose, which makes sense as long as there is no practical purpose for such discrimination. Apparently, language categorizes the things in the world into chunks in a meaningful way.

2.2 Concepts and Entities

Modelers create their own world. A world that is initially empty, and that becomes inhabited by the CONCEPTS [▷29], introduced, one by one, by the modeler.



The word 'concept'[1] derives from 'to conceive', that is: 'to imagine', 'to form as a mental image or as a thought'. A concept is a placeholder for an ENTITY in the system that is being modeled.

Concepts point to, or *refer* to entities in the modeled system. Once a concept enters the model, it receives a name. Names are important [▷30]: a name endows an entity with INDIVIDUAL IDENTITY. By means of a name, it can be distinguished from all other concepts [▷31]. A modeler should give conscious thought to naming every new concept added to the model.

Entities are only relevant for a modeler if they CORRESPOND to a concept in the model. Such concept is said to REPRESENT the entity it corresponds to. A concept represents the entity it was introduced for, in much the same way as a flag represents a country, or a letter in a Western alphabet represents a vocal sound.

Naming a concept not necessarily requires deep thinking. We may borrow names from everyday language. The concept that is to represent a lantern will be called `lantern`, rather than `X` or `Pineapple` [▷32]. Concept's names resemble the answer given to a child learning language when it asks 'how do you call *that*?', while pointing at something.

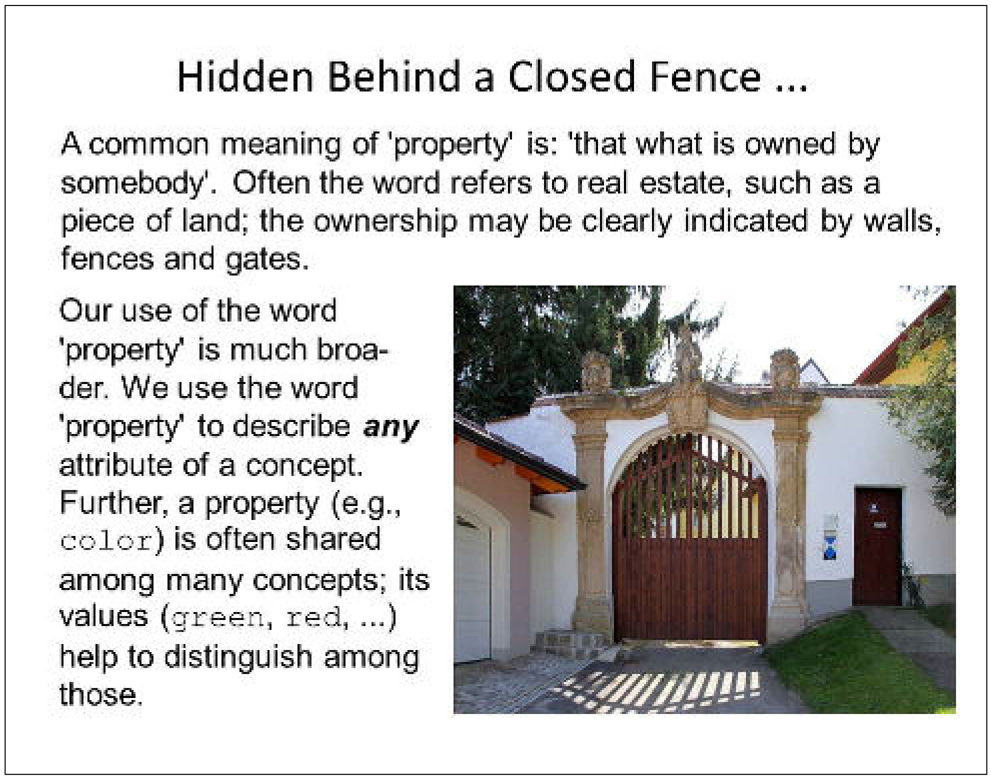
The `lantern`-concept as it occurs in a model allows the modeler to reason about lanterns, but it *is* not a lantern: never will it shine, and no dog will ever urinate against it. Most likely it will be a little rectangle on a piece of paper, together with other rectangles called `traffic`, `road`, etc., forming the conceptual model of the road illumination system.

[1]The image of the stained glass window was taken from http://commons.wikimedia.org/wiki/File:Muzeum_Su%C5%82kowskich_-_Zabytkowy_Witra%C5%BC.jpg?uselang=nl

2.3 Properties

An entity in the real world, may surprise us: 'I didn't know that this lantern was rusty' or 'I did not even think of the possibility that it *could be* rusty'. In the first case we realized that 'rustiness' is a property[2] of lanterns, but we we hadn't assessed if this particular lantern was or was not rusty. In the second case, we did not even consider the property 'rustiness', until we saw brown spots, and realized that these were patches of rust.

Such a discovery is impossible for concepts in a conceptual model. All the properties of a concept are explicitly known, as they result from a definition by the modeler. Their value can also be the result of a definition, or it can result from calculating or inferring, using relations with other properties.



But: what, actually, *is* a property?

First, a PROPERTY is a means to distinguish concepts from each other. When segmenting the world into concepts, we need to say in what respect two concepts are different. In the example in the introduction of this chapter: one of the ways to distinguish horse and soldier is by the property `numberOfLegs`. A horse has four, and a soldier has two, and therefore they can be distinguished. Obviously there may be more distinchuishing properties, for instance `hasTail`, `canTalk`, or `getsSalary`. For a horse the values of these properties are, respectively `true, false, false`, whereas for a soldier they are `false, true, true`. In all situations, when we have two concepts `C1` and `C2` that are different, there must be at least one property that takes on different values for `C1` and `C2`.

It follows that a PROPERTY carries part of the information in a concept. It is an aspect of a concept. A concept's properties together carry all information in the concept.

A PROPERTY always comes in the form of a *name* and a *set of values*.

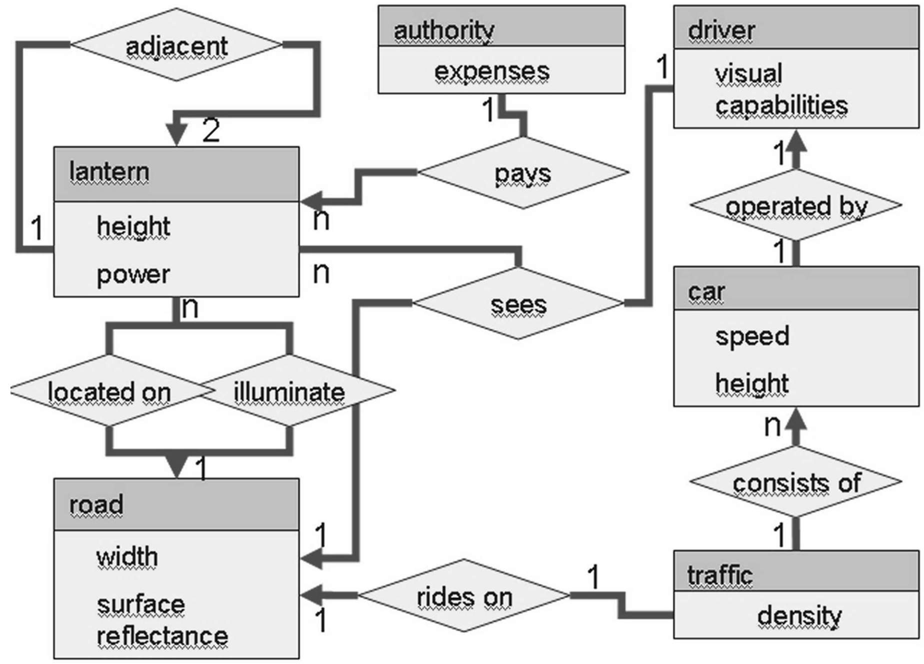
The name of a property is used to refer to the property. So we can talk about the property `color` with name `'color'`.

The set of values property can have is called the TYPE of that property. The type of a property

[2]The image of the gate is taken from http://commons.wikimedia.org/wiki/File:K%C3%B6nigstetten_-_G%C3%B6ttweiger_Herrenhof,_Portal.JPG?uselang=nl

can be limited to just a single element. A set with one element is called a SINGLETON. If a property's type is a singleton, the element of the type is called 'the VALUE' of the property.

The property `color`, could, dependening on the purpose of the model, have a singleton value such as {`green`}, or a set {`green`, `red`}, or a RANGE, {`light Green ... darkGreen`} - assuming that, for some color, we kan assess if it is 'between' light green and dark green. We might also define the type `colors` to represent all colors that can be distinguished, e.g. by a human being or on a computer screen. In a *range* all elements are known when only two extreme values and a notion of ordering is given. For instance, the *range* of all integers between 3 and 6 is the *set* {3,4,5,6}; the ordering is '$<$': $3 < 4$, $4 < 5$, $5 < 6$, and 4 and 5 are the only two integers x with $3 < x < 6$. Ranges are denoted with three dots between the lower and upper element; so {`lightGreen ... darkGreen`} is indeed a range of colors.

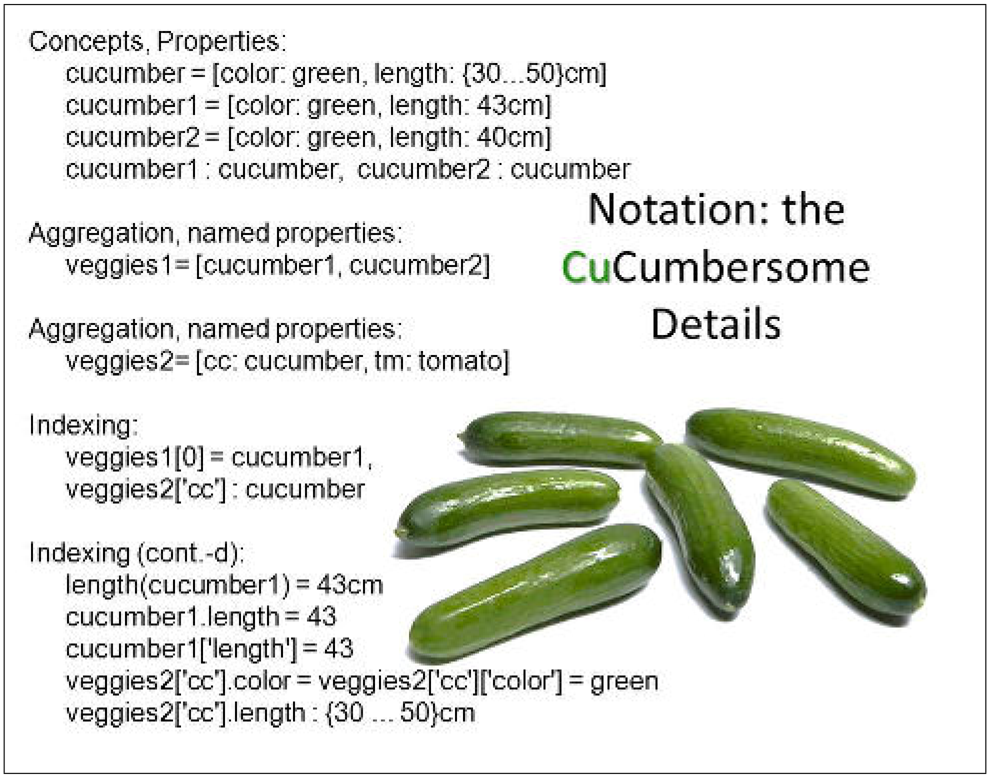


Conceptual model in the form of an Entity Relation Graph

Figure 2.1: Part of a model for the road illumination problem. Concepts are denoted as rectangular blocks. Every concept has a name (top) and perhaps some properties (below the name). Relations are directed arcs, or arrows. Relations have names, written in the diamond shape label. The arity of a relation is indicated by both ends of each arc. Most relations connect two concepts. The relation 'sees', however, connects three concepts: indeed, 'seeing' involves (multiple) lanterns, a driver and a road. (Illustration source: Kees van Overveld)

In summary: a property is a chunk of information about a concept. All information about a concept is captured in its properties. We say: a CONCEPT is defined as a bundle of properties.

An example of the idea of a 'bundle of properties' is a vector. In mathematics we write e.g. $(4, 2, -3)$ or $\langle 4, 2, -3 \rangle$ or $\begin{pmatrix} 4 \\ 2 \\ -3 \end{pmatrix}$, all meaning that the numbers 4, 2 and -3 are not to be considered as loose values, but instead belong together - for example because they refer to a location in 3D space. We say that the values 4, 2, -3 are AGGREGATED. The differences in notations occur because the development of mathematical notation took place over several centuries with contributions of numerous authors. This is not problematic, as long as mathematical expressions are only used by humans, and as long as notation in a single document is consistent. Since some decades, however, mathematical notation is also used to program computers, and mathematical expressions are to be communicated over the Internet. Then the need of standardization is more urgent. So if we will denote aggregations (such as vectors), to be used in the context of automated processing, we need to do so in a uniform manner. Our vector then becomes `[4,2,-3]`.



In this example[3] of an aggregation, we don't give explicit names for the properties. Sometimes this is acceptable. By convention, when denoting a 3D location by a vector, the first number denotes the horizontal coordinate, the second number the vertical coordinate, and the third number is the depth. In computer context, we can refer to one of the members of an aggregation without named properties by setting `p[0]` to refer to the first element of vector `p`; `p[1]` to refer to the second element, et cetera. The number between [and], used to single out one element from an aggregation, is called the INDEX.

In many cases, however, we want property names to be given explicitly. Then we write, again for the same vector, `[x:4,y:2,z:-3]`. Here, `x` is the name of a property with value 4, etc. Using named properties, we don't have to be careful with the order of the elements: `[x:4,y:2,z:-3]` =`[z:-3,x:4,y:2]`, whereas `[4,2,-3]`$\neq$`[-3,4,2]`. If an aggregation is given with named properties, we refer to the value of a property with name `x` as `p['x']` (notice that we need quotes here to signify that `x` is the name of a property, and not an index. In the expression `[x:4,y:2,z:-3]` quotes surrounding `x`, `y`, or `z` are not necessary, because the only thing that can occur before ':' is a name: the name of a concept or the name of a property).

Our notation allows values to be numbers, but other types are also admitted. We may even write down an aggregation such as `M=[[1,0,0],[0,1,0],[0,0,1]]`. This is a vector with three (unnamed) elements, being `[1,0,0]`, `[0,1,0]`, and `[0,0,1]`. In other words, the elements of `M` are themselves vectors, and `M` is a vector of vectors - in other words, a matrix. In a mathematical context we would write `M` as $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

We use square brackets, [and], to denote aggregation instead of parentheses, (and). Apart from convention, this has also a conceptual reason. Paretheses are used to denote FUNCTION application. For a function $f : x \to x^2$, the expression $f(3)$ denotes the function f, applied to

[3]The image of cucumbers was taken from http://www.rgbstock.com/photo/mqyBYfE/cucumber

the argument value 3, yielding 9.

Functions don't necessarily need to be mathematical calculations (as in the example: taking the square). Any recipe that associates, to an element from a set called this recipe's DOMAIN, a single unique element from a set called this recipe's RANGE, is a function. For instance, the property `color` takes as domain the set of all concepts that have a color. `tomato` and `canary` are part of the domain of `color`; e.g. `water`, `wind`, and `democracy` are not. Therefore, `color` can be called a function, and we can formally write `color(tomato)=red` (notice the use of parentheses instead of square brackets). Here, `tomato` is the ARGUMENT and `red` is the RETURN VALUE of `color`. Although many computer languages may not support this, we can view *any* property `P` of a concept `C` as a function, where `C` is an element of its domain, and `P(C)` is a value in its range. If `C` is an aggregation with named properties, we could just as well write `C['P']` for `P(C)` ▷33 .

2.4 Relations

In natural language[4], we distinguish substantives ('nose', 'lantern', 'dog'), verbs ('sneeze', 'shine', 'bark') and PREPOSITIONS. Prepositions *connect* substantives. They often correspond to RELATIONS. In particular relations in space or *spatial relations*. 'The dog is near the lantern' could be replaced by `near(dog,lantern)` without much loss of information ▷34 .

Apart from prepositions, *verbs* also often connect substantives. The verb 'sees', above, was an example; other examples are: `produces(machine, sausages)` for 'this machine *produces* sausages', `wait(passenger,train)` for 'the passenger *waits for* the train', `likes(john,marshMellows)` for 'John *likes* marsh mellows', etc. For this reason, relations are also often expressed by verbs.

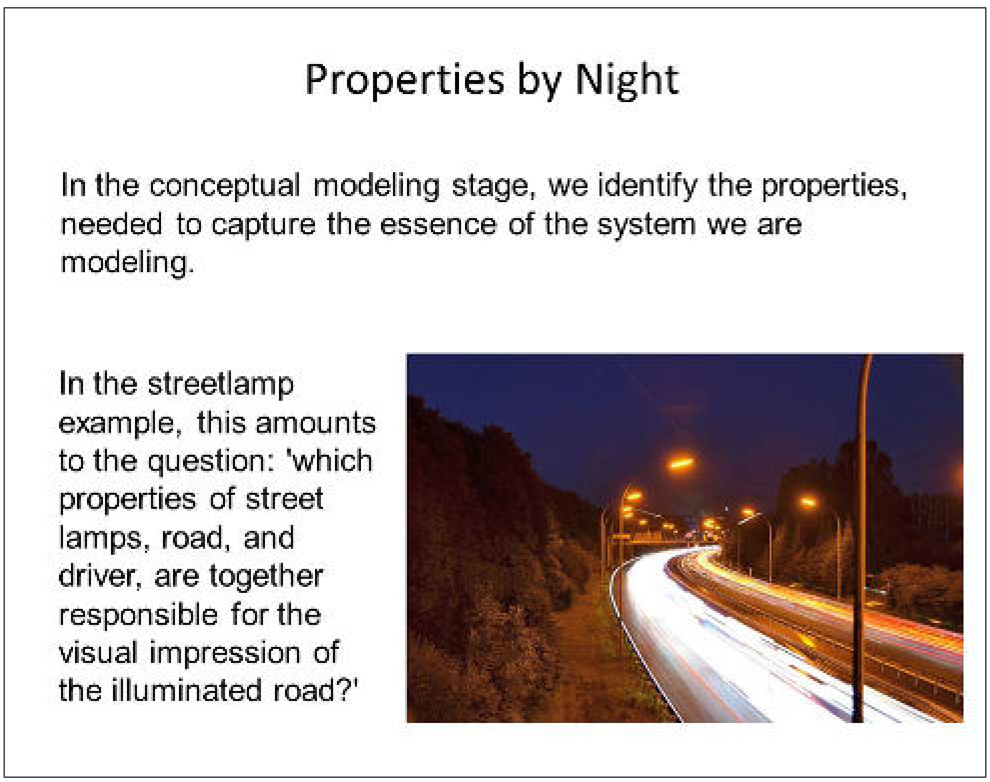
If there is more than one concept in our model, it is necessary to say something about the relations in the model (which concept is related to which concept?), and to what these relations mean. There is a significant difference between the meaning of 'dog near lantern' and 'dog above lantern',

[4] The image of French fries comes from http://commons.wikimedia.org/wiki/File:Pommes_frites_med_fritessaus.jpg?uselang=nl

which again is different from 'lantern above dog'.

Just as properties, relations can also be seen as functions. In Section **??**, we have seen that a function can have multiple arguments. An example of a relation, denoted as a function would be `smarter(John,Peter)`. This function yields the value `true` if `John` is smarter than `Peter`, and `false` otherwise. Its domain is the set of pairs of humans; its range is the set {`true`,`false`}. Again, as with properties, we see that a relation-seen-as-a-function does not necessarily mean that there is some mathematical computation involved: it could be that the outcome of the function evaluation `smarter(x,y)` amounts to looking up some information on `x` and `y`.

2.5 Constructing a Conceptual Model



Construction of a conceptual model[5] is described with a 4-step process. We follow the street lamp example from Section 1.6.

1: establish concepts. First identify the entities for which we need corresponding concepts. We write down things that come to mind when we think of road illumination. Say, `lantern`, `road`, `moon`, `car`, `tree` ▷35.

After brief reflection, we want to skip `moon`, because our system should also work on moonless nights, and we skip `tree` because trees complicate things and should be omitted for a first iteration. Next we add two more concepts, `driver` and `traffic`, because without either of these, illuminating roads is pointless.

A first inventory of concepts should at least contain enough concepts to be able to formulate the problem.

Often the main challenge of this step is not to include too many concepts. Conceptual models often are unnecessarily complex because they contain concepts contributing little to the purpose of the model, but obscuring its working. Hence the title of this chapter: the art of making a good conceptual model is the art of omitting the unnecessary.

[5]The image of the illuminated motorway was taken from http://commons.wikimedia.org/wiki/File:Motorway_(7858495690).jpg?uselang=nl

2: establish properties. For each of the concepts we ask 'what do we need to know of this concept?', in other words: 'which properties do we need?'. For the concepts found in the example in step 1, this may yield the following:

lantern: `height`, `power`;

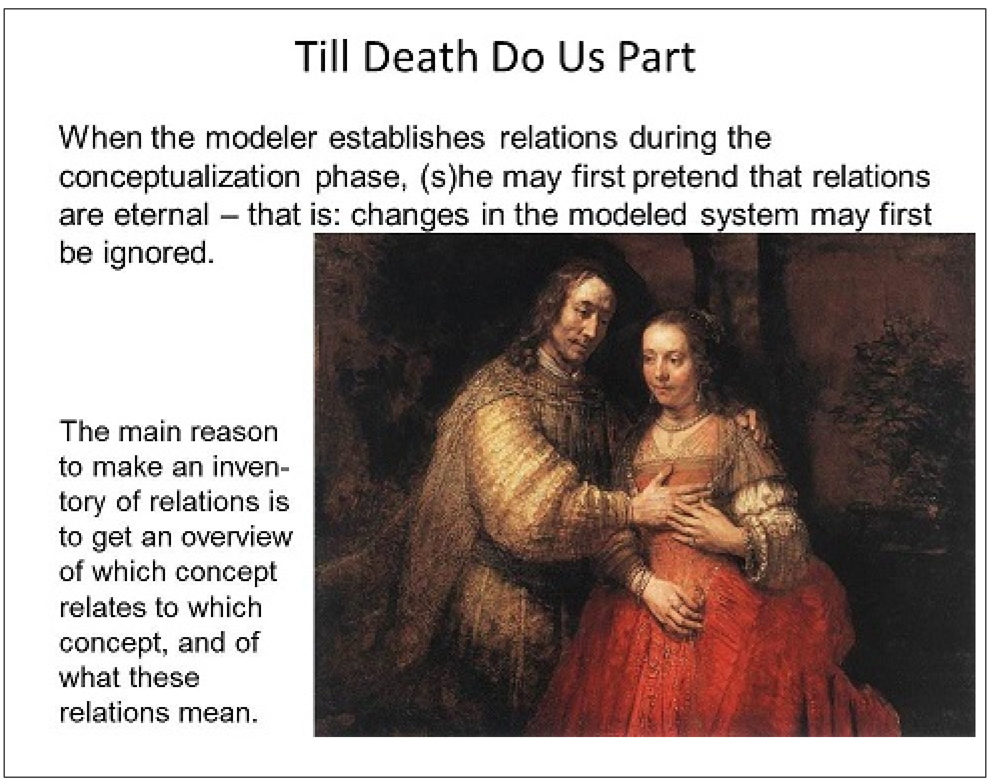
road: `width`, `reflectivity`;

car: `speed`, `height`;

driver: `visualCapabilities`;

traffic: `density`.

Some of these properties are clearly needed (like the height and power of the lantern). Others may be discarded after a moment of reflection. E.g, the height of the car: a truck driver sits 2 meters above the road surface, and a motorist perhaps not even 1 meter, but they experience the illumination conditions not very different. Yet others may require additional work: the visual capabilities of a driver cannot be represented by just a number.



3: establish types of the properties. Every property has a type, determining its set of values[6]. Here, we think of values for the following properties (we use the so-called dot notation, to be explained in detail in Section **??**. The expression `a.b` means 'property `b` of concept `a`').

lantern.height: {5.0 ... 25.0}m - this is the range of heights for lamp posts; the actual value could correspond to an optimum sought for, in case the purpose of the model is 'optimization'. The unit 'm' (meter) is the notation for the type of `lantern.height` signifies that this property is an amount of meters.

lantern.power: {100, 2000}W - these are the powers of LED lamps and gas discharge lamps, respectively. This may come in if the purpose of the model is to aid in what-if analysis or decision support: in this case, the consequence of taking LED lamps or gas discharge lamps.

road.width: {14.40}m - this is a singleton, being the measured width of the segment of road

[6] The image of the poles is taken from http://commons.wikimedia.org/wiki/Category:Flagpoles#mediaviewer/File:Cloetta_Center,_Flagpoles.jpg

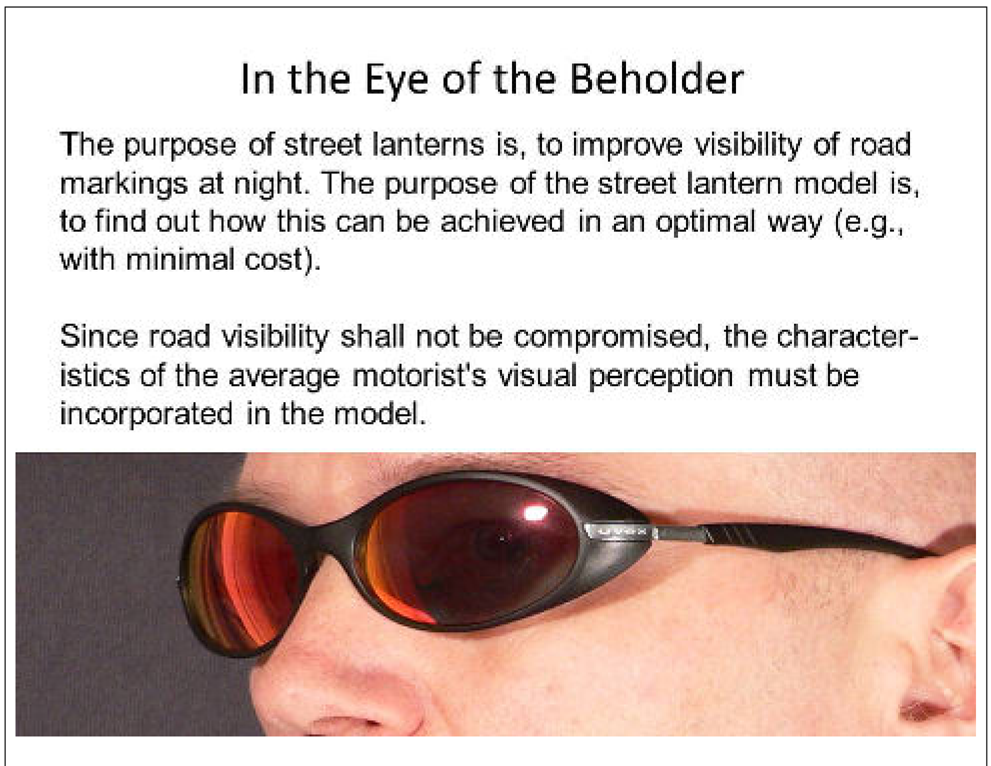
we need to illuminate. It is a constant in the model. When there is no risk for confusion, we may leave out the accolades in a singleton: instead of setting `road.width` $\in \{14.40\}$, we may set `road.width=14.40`. But if there is an uncertainty interval associated with a value, it is appropriate to write that `road.width` is an element of the set $\{14.30 \ldots 14.50\}$.

road.reflectivity: `reflectivity` - this means: we don't know yet the value or the value range for the reflectivity. It will follow from a separate model or from experiments. Therefore we denote it as a named type: `reflectivity`, which will be a singleton, or an uncertainty range if it results from a measurement.

car.height: $\{1.0 \ldots 3.0\}$m - this range of values may be used to check if the final solution is not sensitive to the actual height of the driver, as we supposed earlier.

traffic.speed: $\{20 \ldots 180\}$km/h - this indicates a range of speeds for which we should test the validity or applicability of the model. Does our model still make sense if cars go very fast?

driver.visualCapabilities: `driverView`.



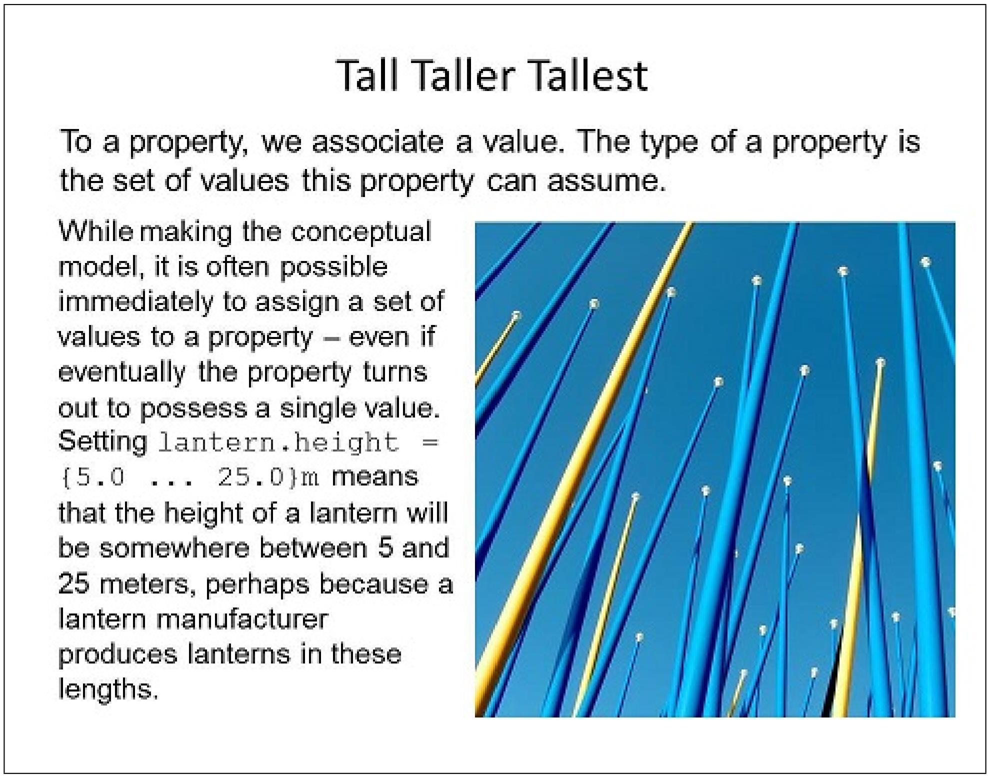
Everything we need to know about drivers' visual capabilities[7] cannot be captured in a single number. These capabilities include the minimal luminance so that road marking can be distinguished, and the maximal luminance so that blinding does not occur. We need a *new concept* that contains the perceptual characteristics of the average driver. This concept, that still is to be detailed, is called `driverView`. The property `visualCapabilities` has a COMPOUND type: its value is a concept with properties, such as `minimalLuminance`, `maximalLuminance`, and perhaps others. The opposite of a compound type is an ATOMIC type. Numeric values, strings and booleans are examples of atomic types.

traffic.density: $\{30\}$cars/minute. The value of this property may result from aggregation. Perhaps measurements of the actual traffic over a period of time are available.

authority.expenses: `real`. In this step we realize that we had forgotten a concept in our model, namely `authority`, with property `expenses`. Without this property in our model, we could not express the purpose of finding an *optimal* solution. We don't know the range of

[7] The image of sunglasses is taken from http://commons.wikimedia.org/wiki/Glasses#mediaviewer/File:Sonnenbrille_fcm.jpg

`authority.expenses` yet, hence the type `real` ▷36 . We want it to be as little as possible, though.



3: establish relations. [8]

Next we seek relations between concepts ▷37 . In principle, we could exhaustively check all pairs of concepts and ask 'is there a relation between these two concepts?' ▷38 . Below we give a list of relations that may emerge; other sets of relations may also be adequate:

- `illuminate(lantern(n), road(1))` - to express that the road is illuminated by multiple lanterns;
- `operatedBy(car(1), driver(1))` - to express that a car is operated by a driver. So the location of the driver will be fixed with respect to the location of the car;
- `consistsOf(traffic(1), car(n))` - to express that traffic is an aggregation of multiple cars;
- `ridesOn(car(n), road(1))` - to express that the location of any car is constrained to the road;
- `sees(driver(1), road (1), lantern(n))` - to express that the illumination, perceived by the driver, comes from light, emitted by lanterns, reflected on the road. Notice that this is a relation between 3 instead of 2 concepts;
- `pays(authority(1), lantern(n))` - to express that the costs of installing and operating the lanterns are to be paid by the authority responsible for lighting the motorways;
- `adjacent(lantern(1), lantern(2))` - to express that lanterns are adjacent to each other, in other words that each lantern has two adjacent lanterns ▷39 .
- `locatedOn(lantern(n), road(1))` - to express, for instance, if lanterns are located on the axis of the road, or at both sides, et cetera.

In this list, the numbers in brackets indicate the number of concepts involved in the relation. 'n' means '1 or more'. These so-called *arities* are further explained in Appendix **??**.

It usually requires several iterations before the lists of concepts, properties, values and relations are appropriate. At any time, we should check against the purpose of the eventual model.

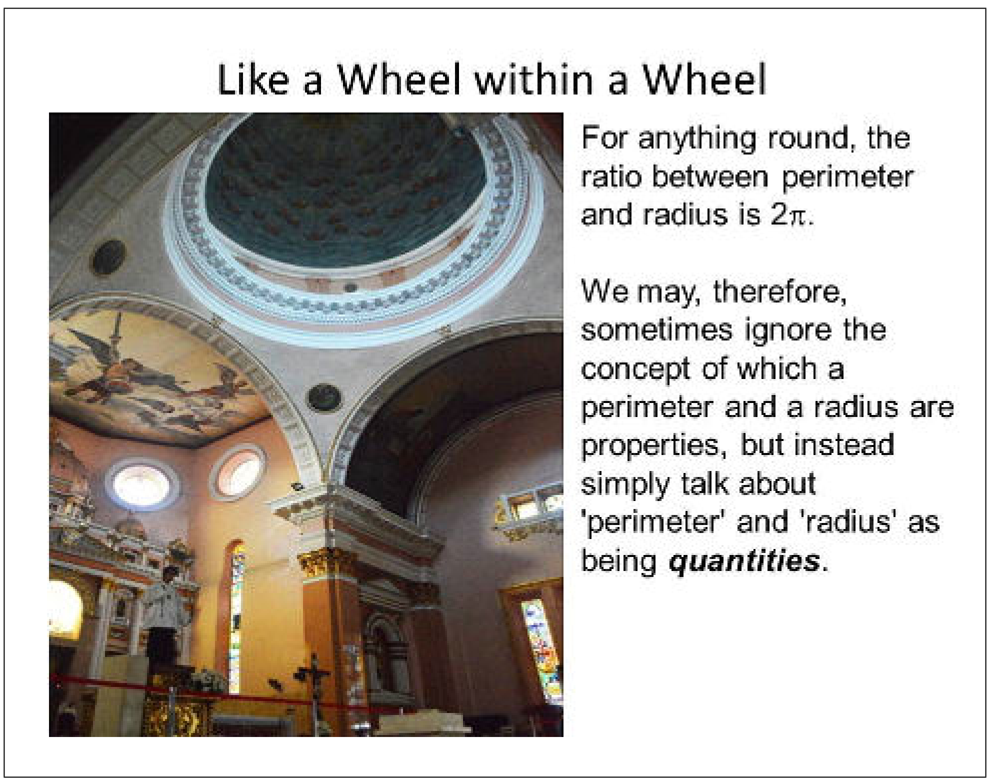
[8] The reproduction of 'The Jewish Bride' was taken from http://commons.wikimedia.org/wiki/File:Rembrandt_-_The_Jewish_Bride_-_WGA19158.jpg?uselang=nl

Using relations, a conceptual model can be graphically depicted as in Figure 2.1. Such a drawing is called an ENTITY-RELATION GRAPH. The nodes (usually drawn as boxes) in an entity-relation graph are the concepts from our conceptual model. Other terms for such graphical representations are SEMANTIC NETWORK or CONCEPT GRAPH.

Relations in an entity-relation graph need to be indicated by arrows, since all but symmetric relations have a direction: `R(A,B)` generally means something different from `R(B,A)`. Every node represents a concept and is depicted as a box; this box contains the name of a concepts and perhaps more information, such as properties and perhaps their types.

Concepts, properties, and value sets, although they form a natural perspective on the world, are quite subtle. In Appendix **??** we go somewhat deeper into some issues for dealing unambiguously with conceptual models.

2.6 Quantities



A property is always a property *of* some concept. We never encounter isolated properties. For some purposes, however, we don't need to know which concept some property is a property of[9]. For instance, to calculate the perimeter of a circle, it does not matter if this circle is the shape of a blood vessel, a piece of land or the lid of a bucket with paint. In all cases the same formula applies. Mathematicians commonly talk about QUANTITIES, disregarding the concept that the quantity is a property of. Knowing how to compute the perimeter of a circle can be applied to the concepts `bloodVessel`, `pieceOfLand` and `bucketWithPaint.lid` without further consideration of other properties [▷40] . Therefore it is adequate to talk about 'quantities'. We already know that every property has a type, therefore quantities have a type, too. The TYPE of a quantity is defined as the set of values the quantity can assume. Since we may not know what concept a quantity, such as `radius`, is a property of, the type of a quantity does not relate to something known of any concept. The

[9] The image of Binondo church is taken of http://upload.wikimedia.org/wikipedia/commons/1/15/Binondo_Church_Circular_Configuration.jpg?uselang=nl

type of a quantity, corresponding to property $C.P$ is therefore the union of all possible types of the property $C_i.P$ for all possible concepts C_i. In the case of `radius`, the type of the radius of a bicycle wheel is $\{0.3...1.8\}$m, but a radius in general can be any non-negative real number. Therefore the type of a quantity *radius* will be the non-negative reals.

To denote the difference between quantities that occur as properties of a concept versus quantities that appear without a conceptual context, we write the former in `this font`, whereas the second will be written in *this font*. So: "`myBicycle.frontWheel.perimeter= 6.28 * myBicycle.frontWheel.radius`" as opposed to "*perimeter*$=2\times\pi\times$*radius*".

We use the term 'quantity', where other texts would use words such as PARAMETER, VARIABLE, FACTOR, TERM, or COEFFICIENT.

The words 'parameter', 'variable', 'factor', 'term', or 'coefficient' all have slightly different meanings. These meanings vary over the disciplines; within one discipline the meaning can be different in different contexts. What is called 'coefficient' in one discipline might be 'variable' in another discipline, or in another context. To avoid confusion we stick with a single word, 'quantity', that will be used in all disciplines and all contexts. To stipulate differences in meaning we will introduce roles, or categories of quantities in Chapter 5.

Strings Attached

An often occurring elementary type is ***string***.

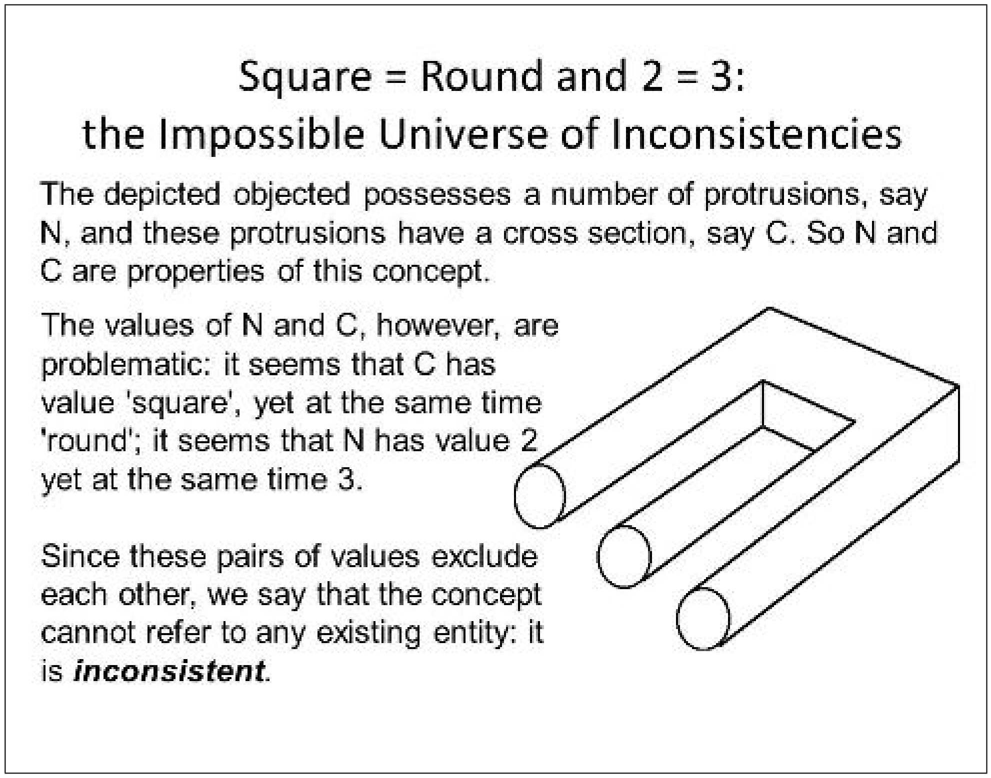
A string is a sequence of characters (letters, digits, interpunction), appearing within quotes. Just as with numbers, we can operate on strings. If p= 'pea', s='soup', then p+s='peasoup', whereas s+p='souppea'. Notice that here, p and s are not strings but names of quantities.

2.6.1 Types of Quantities

Since properties have types, quantities have types, too. We distinguish two sorts of types[10]: ELEMENTARY TYPES and COMPOUND TYPES. To know an element of an elementary type, we don't need any further properties. For instance, integer number 5 is fully known. All properties (such that it is an odd prime, smaller than 19, etc.) of the value 5 can be deduced using nothing else than 'the value is 5'. Other elementary types are BOOLEANS and strings such as 'aAaa', 'aaaa', 'pineapple' or '12345' (the latter not to be confused with the number 12345). We write 'pineapple' for the string consisting of the letters p,i,n,e, ...; `pineapple` without quotes is the name of a concept or a property. Also real numbers, e.g. including π, $\sqrt{2}$, and 12.7,

[10]The image of the neon sign was taken from http://commons.wikimedia.org/wiki/Neon_sign#mediaviewer/File:Neon_(12594495033).jpg

form an elementary type. Finally, an elementary type can be an enumerated list of constants. The type `material`, for instance, could be {`wood,metal,plastic,cement`} ▷41 ; the type `shirtNumbersOfFootballPlayers`={1,2,3,4,5,6,7,8,9,10,11}, which is a subset of the set integer numbers, is also a type.



An example of a compound type is the type `rectangle`: the set of all rectangles. To fully know a rectangle, we need further information. Its properties, such as length, width, area, perimeter, orientation, ... should be CONSISTENT[11] , that means: a concept with properties that have the indicated values should be LOGICALLY POSSIBLE. For instance, a rectangle with area 12 and perimeter 6 cannot exists: these two properties have inconsistent values. Further, for the sake of efficiency, properties should be INDEPENDENT, that is: the value of a property should not be derivable from the value of another property. So properties for defining a rectangle could be `width` and `height` both of type `real`. Every non-negative width and non-negative height determine a rectangle. Alternatively, we could give the properties `area` and `perimeter`: indeed, the width and height of a rectangle uniquely follow from the area and perimeter ▷42 .

We saw a notation for concepts in terms of their properties. This works for properties, both of elementary type and compound type.

We give the example of a rectangular box, that is a configuration of six rectangles. For brevity, we omit the units; all lengths are in cm. First, we could write this as Expression 2.1. This expression, although technically correct, gives no insight at all. It is difficult to see if the right dimensions are provided for the right edges.

A better way to write this, is Expression 2.2, together with the definition from Expression 2.3. We re-use the values `r1`, `r2`, `r3` so that we can easily see that top and bottom are CONGRUENT - that is, one is the result of translating, scaling or rotating the other. Similar for left and right, and for front and back.

This form still requires verification if the dimensions for matching edges are correct.

[11]The image of the impossible object is taken from http://commons.wikimedia.org/wiki/File:Impossible.png?uselang=nl

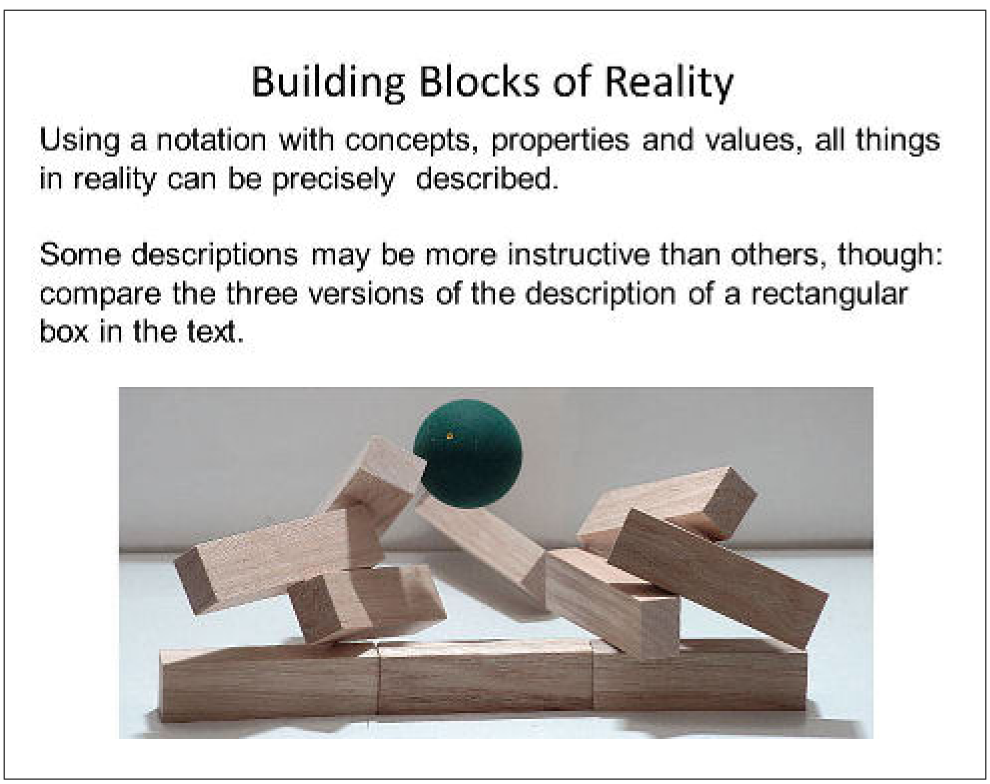
A third attempt is Expression 2.4, which is truly simpler than the earlier two versions.

```
rectBox = [bottom:[width:3, height:4], top:[width:3, height:4],
          left:[width:4, height:2], right:[width:4, height:2],
          front:[width:3, height:2], back:[width:3, height:2]].
```
(2.1)

```
rectBox = [bottom:r1, top:r1, left:r2, right:r2, front:r3, back:r3].
```
(2.2)

```
r1=[width:3, height:4],
r2=[width:4, height:2],
r3=[width:3, height:2].
```
(2.3)

```
rectBox=[width:3, height:2, depth:4],
```
(2.4)



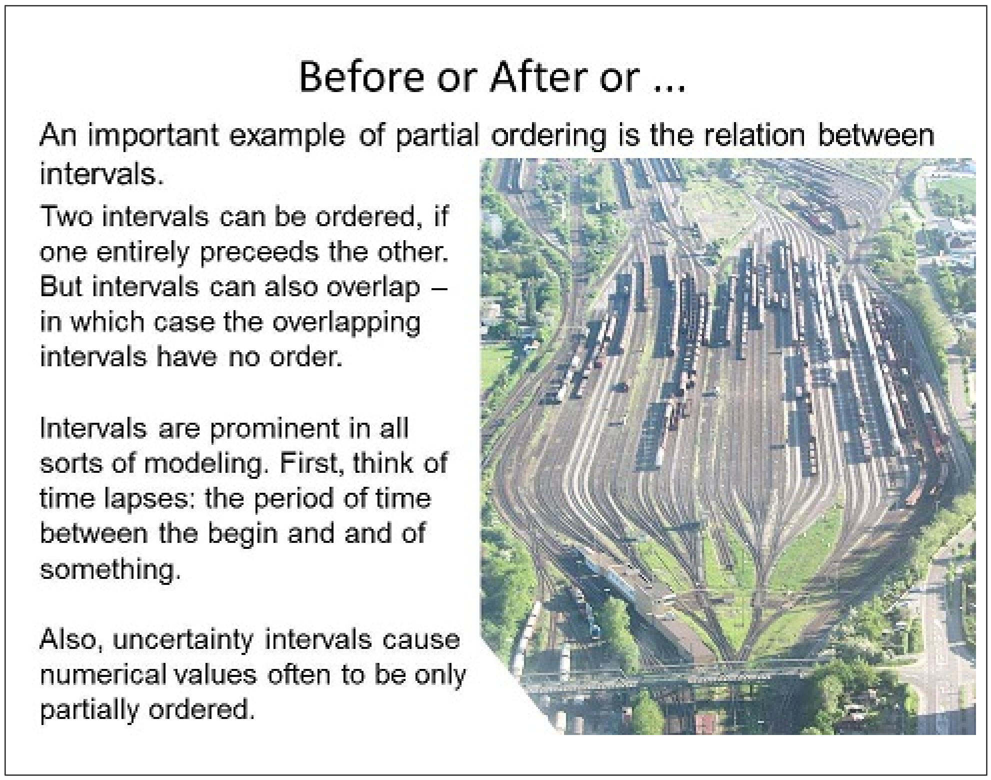
This example[12] shows that we should consciously choose properties and concepts such that the conceptual model represents the entity we want to model. In this case: if we define the box to consist of six rectangles, many CONSTRAINTS, i.e., limitations to the values of various properties, have to be fulfilled. Adjacent sides should have the same length for a shared edge, and a box has 12 such edges. The second version uses the symmetry of the rectangular box. The top and bottom rectangle are the same concept (rectangle r1), so we have to verify fewer constraints. Similar for the two other pairs of faces. The third version is the simplest of all. There are no constraints left, since the length, width and height of the box can be set independently.

[12] The image of the blocks was taken from http://commons.wikimedia.org/wiki/Category:Wooden_blocks#mediaviewer/File:Wood_Block_Break_Out.JPG

2.6.2 Operations on Quantities: Ordering

Conceptual modeling preceeds quantitative modeling. In quantitative modeling, we do mathematical operations on quantities. Which operations are allowed for which quantities, however, depends on the types of these quantities.

An important distinction between various types is, whether quantities can be ORDERED. Quantities that can be ordered, are called ORDINAL. Quantities that cannot be ordered are called NOMINAL.



A familiar ordering is the ordering of numbers: $3 < 7$. This relation is TRANSITIVE, that is, from $a < b$ and $b < c$ we have that $a < c$.

Further, it has an opposite relation: the opposite of $<$ is $>$.

Finally, it is defined for *any* two different numbers: always one or the other is the bigger one. This last condition is obvious for numbers, but in most other cases it does not hold[13]. For instance, we can look at family relations between people. My father is an ancestor of me, and his father was an ancestor of him and thereby also an ancestor of me: `ancestorOf` is transitive. The opposite is `descendantOf`. But most people are neither an ancestor nor a descendant of me. Therefore `ancestorOf` only applies to *some* pairs of individuals. This is called PARTIAL ORDERING.

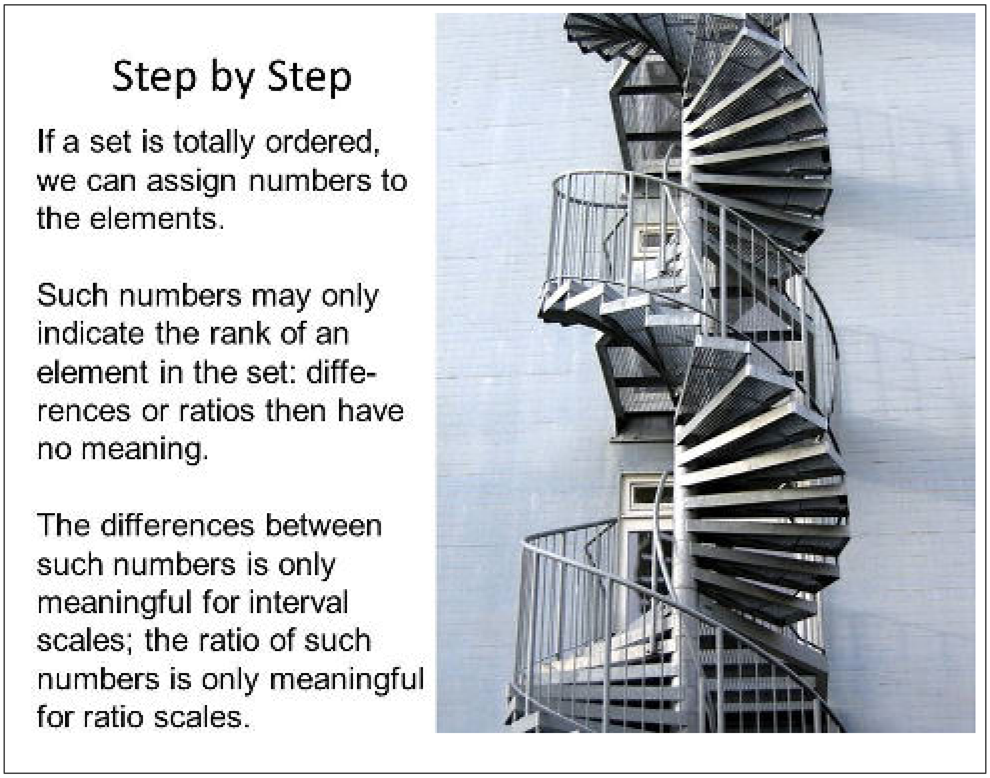
Another, perhaps more important example of concepts that are partially ordered, is *intervals*. Indeed, an interval with an upper bound that is entirely below the lower bound of a second interval is `below` this second interval, but if the two intervals overlap this is not the case. Intervals will play an important role in Chapter 6 when we investigate reliability, accuracy and sensitivity of models.

Sets that are totally ordered are called *totally ordered sets*; sets that are partially ordered are called partially ordered sets or POSETS. Sets that are not ordered at all are called NOMINAL. A set of countries is an example of a nominal set. If we consider a country as a concept with properties

[13] The image of the classification yard (Dutch: 'rangeerterrein') showing many overlapping and non-overlapping intervals (sequences of freight cars) was taken from http://commons.wikimedia.org/wiki/File:Rbfkornwestheim.jpg

such as area, population or gross national product, we can define an ordering, but what we order then is really a set of areas, a set of populations sizes or a set of amounts of money, and not a set of countries.

The various operations that are allowed for each type of sets are explained below[14]:



- *nominal sets:* In a nominal set, elements have no ordering. We can only assess if two elements are equal or not, and we can count how many elements (concepts) occur for which some property has some given value. We can say 'in this collection of cars, Volkswagens occur twice as much as Opels', but that does not mean that there is any ordering between a particular Volkswagen and a particular Opel.
- *partially ordered sets:* Transitive relations often give rise to partially ordered sets. Examples are `descendsFrom`, or `comesBefore`. The latter relation occurs for instance if we deal with processes where things happen at different times. See Section 3.3.1 for an example. Partial ordering may allow the verification of a design decision: 'alternative A is better than alternative B'. It occurs when dealing with preferences: you may like chess more than rugby, and chess also more than waterpolo, but the preference between rugby and waterpolo may be unknown.
- *totally ordered sets:* In a totally ordered set an ordering relation exists between *any* two elements in the set. An example is MOHS SCALE for mineral hardness ▷43 . As follows: take two samples of two different minerals; push one firmly onto the other and move. Only one of the two will receive a scratch. This introduces an ordering between any two minerals: `isSofter`, which is the opposite of `isHarder`.

  There are various sorts of scales associated to totally ordered sets ▷44 .

  1. *ordinal scale:* Mohs scale is an example of a totally ordered scale . It is possible, for a set of different minerals, to assign an integer number to each of them. This is also called a RANKING. But it is meaningless to ask if diamond is the same amount harder than copper oxide, as copper oxide is harder than chalk. So, taking averages of Mohs numbers to talk about 'the average hardness' is not allowed ▷45 . For a collection of minerals, however, it is allowed to search for the MEDIAN: the mineral for which the number of minerals that are less hard equals the number of

[14] The image of the spiral staircase is taken from http://www.rgbstock.nl/photo/mfjMtZA/wenteltrap+1

OK to compute ...	Nominal	Ordinal	Interval	Ratio
frequency distribution	yes	yes	yes	yes
median	no	total order:yes; partial order: no	yes	yes
add, subtract, mean	no	no	yes	yes
ratio	no	no	no	yes

Table 2.1: Operations allowed on various types of scales

minerals that are harder.

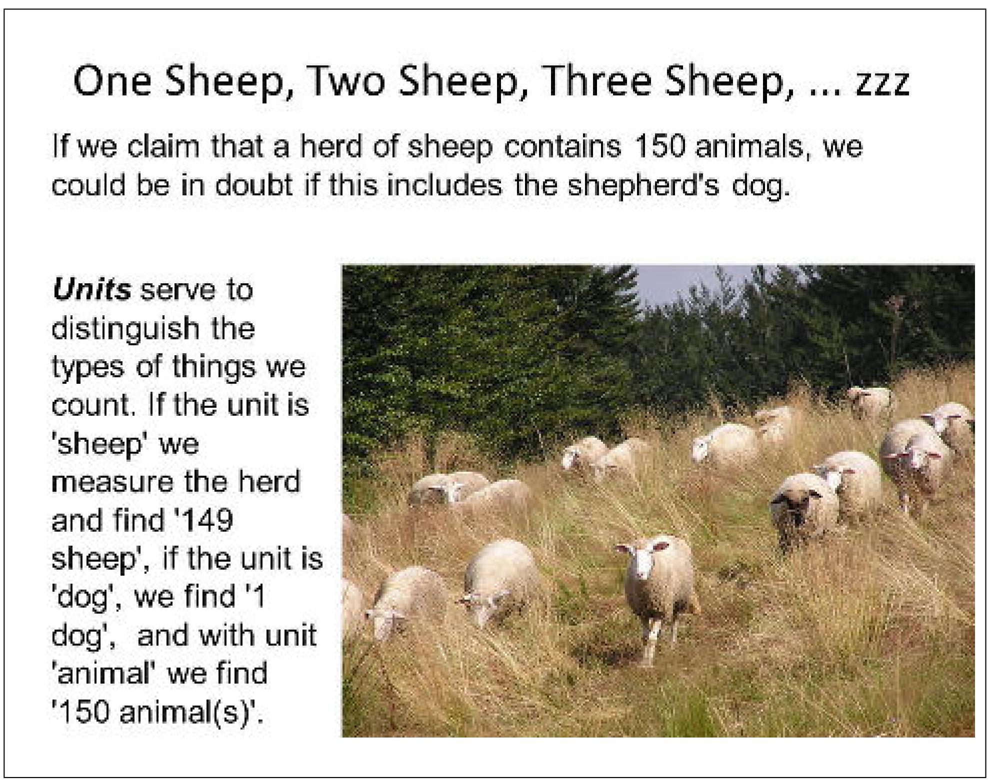
2. *interval scale:* The *difference* between two Mohs numbers has no meaning. For temperatures in a Celsius scale, however, the difference between 20 and 10 Centigrade has a meaning: it corresponds to an amount of energy, and the same amount of energy is needed to heat up something from 80 to 90 Centigrades. Scales that allow addition or subtraction, are called INTERVAL SCALES.

3. *ratio scale:* For a Celsius scale the ratio between, say 80 and 20 Centigrade does not correspond to something physical. For the Kelvin scale, though, a ratio between two temperatures corresponds to a ratio between energy contents: an amount of gas at 80 Kelvin contains 4 times as much energy than the same amount of gas at 20 Kelvin. From this, it follows that the energy contents of *any* amount of gas at 0 Kelvin is 0 Joule. The Kelvin scale is an example of a RATIO SCALE. A ratio scale has a meaningful zero, whereas the zero for a difference scale is arbitrary.

We summarize this in Table 2.1 ▷46 .

2.7 Units, Scales and Dimensions

2.7.1 Counting is Easier than Measuring

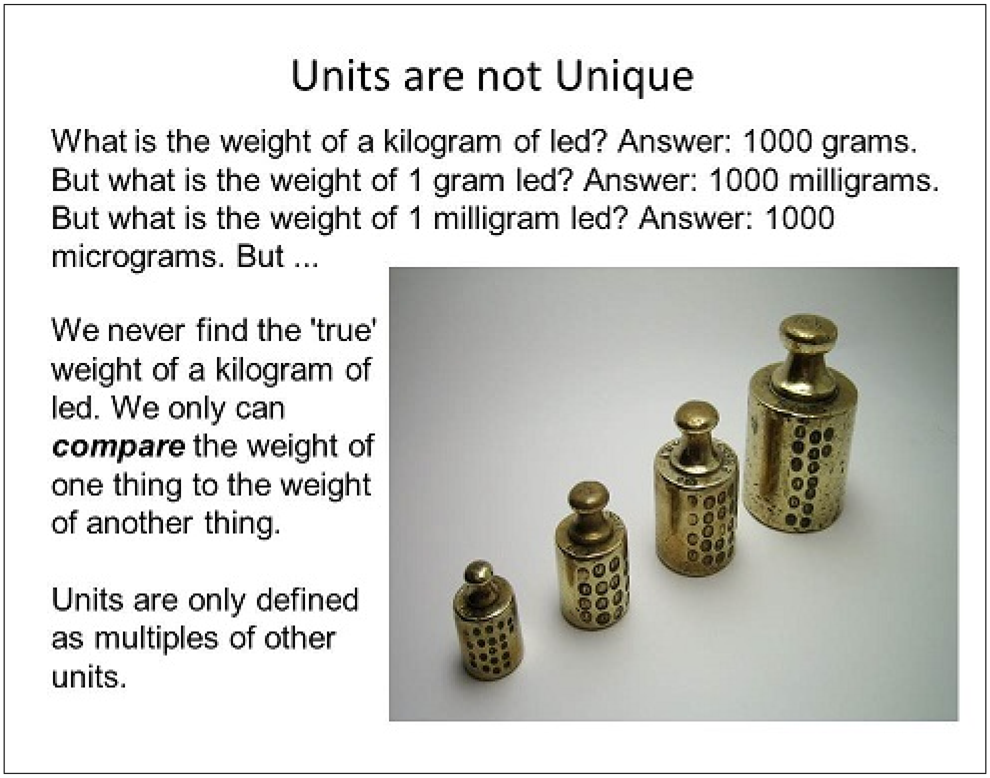


Quantities in a model often correspond to observations or measurements. The simplest form of quantitative observation is counting: answering the question 'how many units of sort X do I have?[15]'.

Suppose that we want to have the dimensions of a piece of land. Assume a known ASPECT RATIO for the piece of land, e.g. square. The piece of land is surrounded by barbed wire, spanned by poles 10 meter

[15]The image of the sheep was taken from http://www.rgbstock.nl/photo/mhilDzK/Schapen+in+de+bergen

apart. We can count the poles. If we find 400 poles, we conclude that the perimeter of the piece of land is 4000 meter, and the area amounts to 1 hectare.



Here, measuring is reduced to counting[16]. But we could doubt the accuracy. Are the poles really exactly 10 meter apart? To get more precise results, we use a measuring rod, say of 1 meter length. Again, measuring amounts to counting. The number of times the measuring rod fits in the perimeter is, say, 3998 times plus a bit.

To make our result even more accurate, we use a shorter unit, of one decimeter long. This time we find 39986 units plus a bit. Repeating the experiment with an even shorter unit (a centimeter) produces 399863 units plus a bit, and so on, until the unit is too small to assess if there still is a remaining bit, or until our curiosity is satisfied, or we have no smaller units at our disposal.

From this experiment, we learn the following. Suppose that we have two units, u_1 and u_2. They have a ratio $p_{1;2} = \frac{u_1}{u_2}$. That means: unit u_2 fits $p_{1;2}$ times in unit u_1. $p_{1;2}$ is an integer, it counts the number of times u_2 fits in u_1. Next, there is a quantity l, measured with u_1. This gives the number x_1; if we measure the same quantity with u_2 it gives x_2.

So we have $x_1 u_1 = x_2 u_2$, or $\frac{x_2}{x_1} = \frac{u_1}{u_2} = p_{1;2}$. We can use $p_{1;2}$ to predict x_2 if we have measured x_1 (namely: $x_2 = p_{1;2} x_1$), or the other way round.

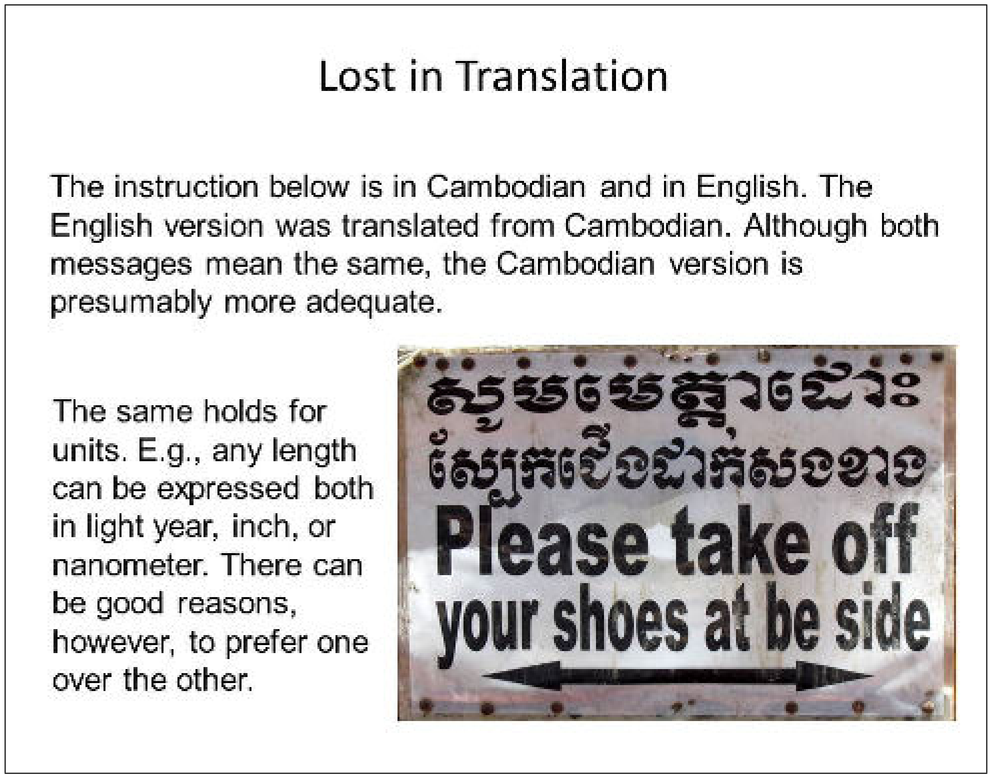
Now let there be a third unit u_3. Then we write

$$\begin{aligned} p_{1;2} &= \frac{u_1}{u_2} \\ &= \frac{u_1}{u_2}\frac{u_3}{u_3} \\ &= \frac{u_1}{u_3}\frac{u_3}{u_2} \\ &= p_{1;3}\ p_{3;2}. \end{aligned} \tag{2.5}$$

[16]The image of the weights is taken from http://www.rgbstock.nl/photo/mWjRlAm/Oma%27s+Old+Weights

So: in order to go from one unit to another unit[17], we use Expression 2.5.

For units u_i and u_j, the number $x_i\ p_{i;j} = x_i/p_{j;i}$ is an INVARIANT of the measured thing. That is: it does not change, if we measure the *same* thing with a *different* measuring unit.



We write that a length is 399 m instead of 399. The expression '399 m' is not just shorthand for '399 measured in meters'. Rather, it is a mathematical expression that involves a multiplication, in the same way that we write ab if we mean $a \times b$.

The factor 'm' in '399 m' is a multiplication with a number $p_{m;U}$ to obtain the invariant $x\ p_{m;U}$, where the number $p_{m;U}$ is different for any unit U. The symbol 'm' means $\times p_{m;U}$ or $p_{m;U}$ for short, where the numerical value of $p_{m;U}$ is not stated.

The *ratio* of the $p_{m;U}$ between two units, 'meter' being one of them, is defined when we know the other unit, U, we want to use. So we don't know what '100 m' means in absolute sense, but we know what it means in comparison with another length expressed in meters, or in comparison with another length expressed in other units. Notice that this ratio does not depend on anything we are measuring: it is purely a property of a pair of units.

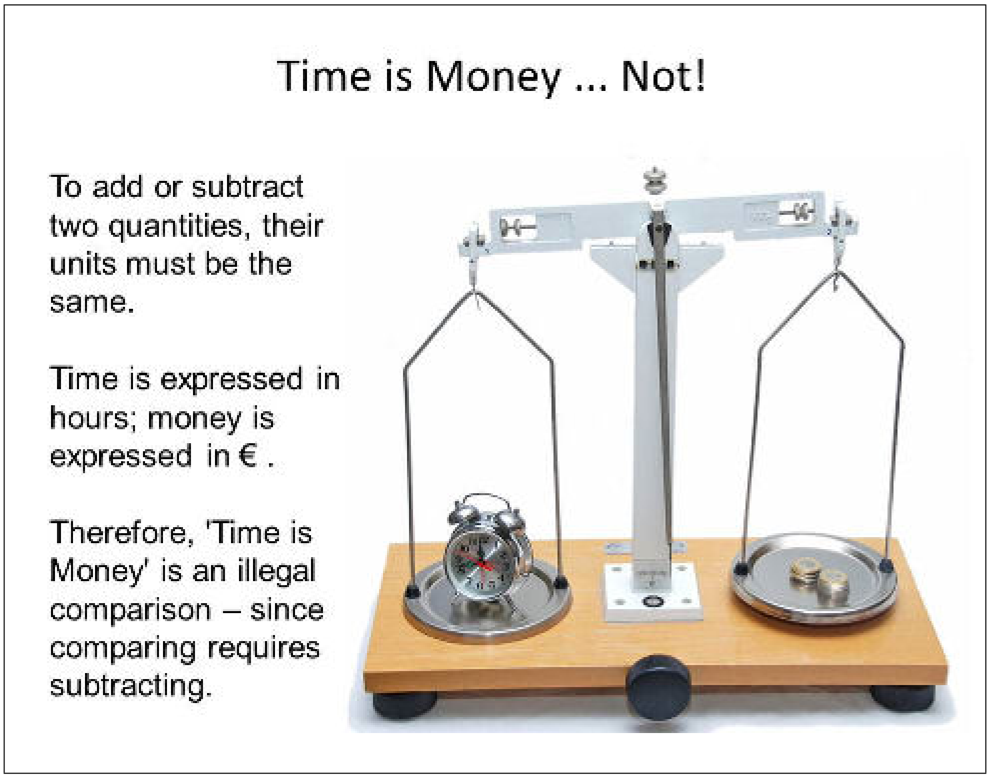
To work with quantities, we don't need to know the numerical value of $p_{m;U}$, or, in general, $p_{i;j}$. If we want to go to another unit, we use Expression 2.5:

$$\begin{aligned} x_m\ m &= x_m\ p_{m;U} \\ &= x_m\ p\ _{m;cm}\ p_{cm;U} \\ &= x_m\ \frac{u_m}{u_{cm}}\ p_{cm;U} \\ &= x_m\ 100\ p_{cm;U} \\ &= 100\ x_m\ cm. \end{aligned} \tag{2.6}$$

So, again, '1 m = 100 cm' states the equality of two algebraic products; the first one is the

[17]The image of the Cambodian-English instruction was taken from http://www.rgbstock.nl/photo/nKv2kXq/lost+in+translation

product of '1' and 'm', and the second one '100' and 'cm'.



Suppose[18] that we want to calculate the area of the piece of land. We measured one side in meters, giving 100 m (being shorthand for $100 \times m$) and for some reason, the other side in decimeters, giving 1000 dm. We know that the area of a rectangle is found by width $\times$ height. In this case we have $(100\times p_{m;U})\times(1000\times p_{dm;U}) = 100.000\times p_{m;U}p_{dm;U}$ or 100 m $\times$ 1000 dm=100.000 m dm. Although this is correct, we don't commonly write factors like 'm dm'. We use Expression 2.5 to write $p_{dm;U} = p_{dm;m}\ p_{m;U} = \frac{u_{dm}}{u_m}\ p_{m;U} = 0.1\ p_{m;U}$ to express this instead as 10.000 m m. This is again an algebraic product, and we write m m=m$\times$m =m^2. So the term 'm^2' is the consequence of consistent algebraic manipulation where symbols such as 'm' are treated as factors, representing unknown factors p with constant ratios.

We summarize:

Unit symbols are algebraic factors that are part of the expression, and should be manipulated as such when doing algebra with the quantities.

Algebraic operations include multiplications, but also additions or subtractions. To express that the perimeter of some piece of land equals twice the length (say, l_l=100 m) and twice the width (say, l_w=70 m), we write:

$$\begin{aligned} \text{perimeter} &= l_l \times p_{m;U} + l_w \times p_{m;U} + l_l \times p_{m;U} + l_w \times p_{m;U} \\ &= 2 \times (l_l + l_w) \times p_{m;U} \\ &= 2 \times 170 \times p_{m;U} \\ &= 340 \text{ m}, \end{aligned} \tag{2.7}$$

where we have explicitly taken the quantity $p_{m;U}$ 'out of the brackets'.

Taking a quantity out of the brackets is only allowed if the same quantity occurs in all the terms between the brackets. Suppose we want to express that we traveled 10 km in a taxi, after waiting half an hour. We might want to write the total waiting time as an addition: 10 km plus 30

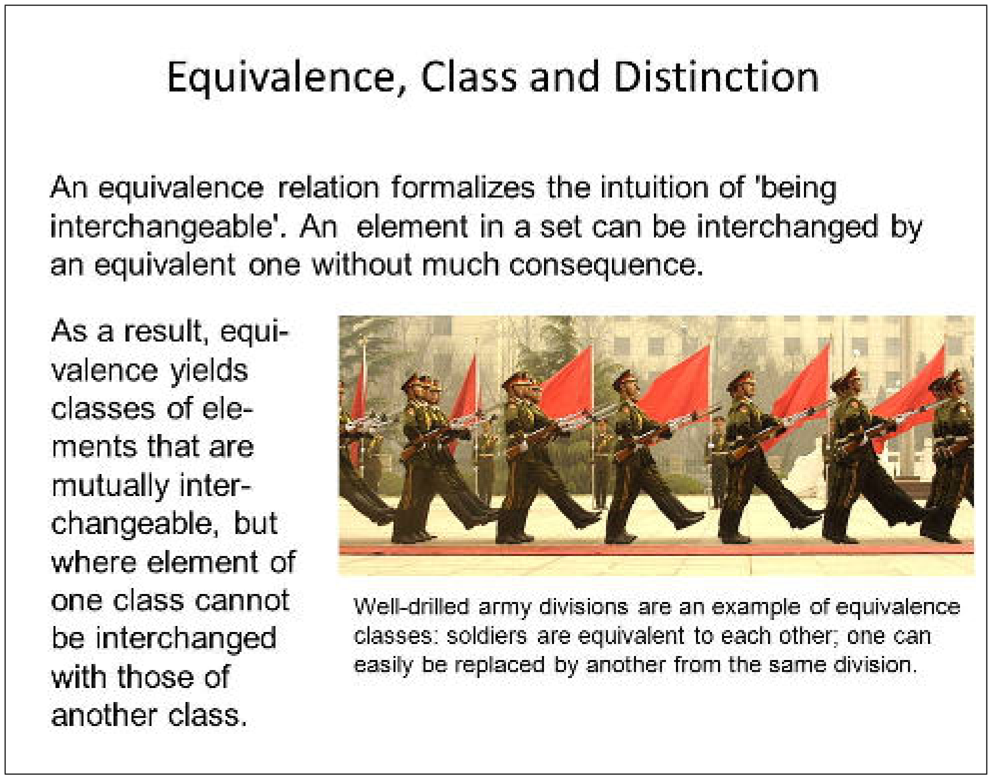
[18] The "time is money"-image was taken from http://www.rgbstock.nl/photo/mhYAppa/Clock+and+money+on+the+weighin

minutes. This is not wrong, but when we try to formalize it in the same way as calculating the perimeter of a piece of land we find:

$$\begin{aligned} \text{time elapsed} &= 10 \times p_{km;U} + 30 \times p_{minute;U} \\ &= \cdots , \end{aligned} \tag{2.8}$$

which cannot be simplified further. There is no common factor that we can take outside the brackets. There is no ratio between $p_{km;U}$ and $p_{minute;U}$ that is independent from any measurement, so we cannot express one as multiple of the other.

Therefore, adding quantities with different units is not forbidden pe se (when we calculated the perimeter of the piece of land we added meters and decimeters), but it is allowed only if we can take something out of the brackets, that is: if there are common factors.



2.7.2 Units and Dimensions

We have seen[19] an example of several units (m, dm, cm) that correspond to factors $p_{m;U}$, $p_{dm;U}$, $p_{cm;U}$, respectively; these units have constant ▷[47] ratio's that don't depend on anything measured: $p_{m;U} = 0.1 p_{dm;U}$, $p_{dm;U} = 0.1 p_{cm;U}$, and similar for units km, mm, μm, nm, etc.

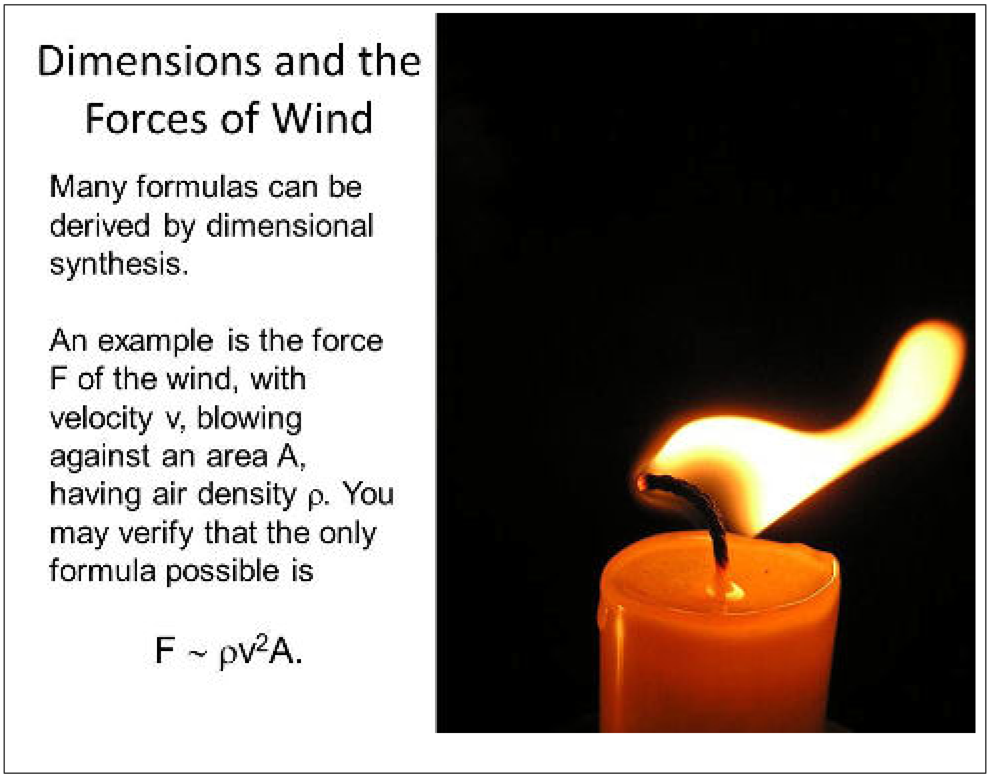
Two units, with p's that have constant ratio's, are called EQUIVALENT ▷[48] .

Things that are equivalent can be grouped in so called equivalence classes. Indeed: two things that are not equivalent cannot be in the same class. The equivalence classes, belonging to the relation 'has a constant ratio with' between two units, are called DIMENSIONS. To signify that 'length' is a dimension we usually use an abbreviation and a conspicuous font, like $\mathcal{L}$ for length. Examples of dimensions are length ($\mathcal{L}$), time ($\mathcal{T}$), mass, ($\mathcal{M}$) and many others.

A value is denoted by a number and a unit; the name of a quantity is sometimes annotated by the dimension in square brackets. So, if P.l is the property l of concept P, and it has dimension length, we may write P.l $[\mathcal{L}]$.

[19] The image of the Chinese army parade is taken from http://commons.wikimedia.org/wiki/File:Chinese_honor_guard_in_column_070322-F-0193C-014.JPEG?uselang=nl

Some dimensions correspond only to a single unit, for instance $\mathcal{SHEEP}$ (used to measure the size of a flock of sheep by counting; there are no fractional sheep, so units such as $\mu\mathcal{SHEEP}$ don't occur), and another one is $\mathcal{PIANO}$ (used to measure the size of a collection of piano's by counting them).



Some dimensions can be constructed from other dimensions. We have seen the example where the unit of area was expressed as the product m m or m^2. Other units for area are cm^2 or (light year)^2. These are also equivalent: indeed, m^2=10.000 cm^2, etc., so there is a constant ratio. Therefore we associate a dimension to these units, by the name of 'area' or $\mathcal{L}^2$.

The dimension area results by multiplying two equal dimensions ($\mathcal{L} \times \mathcal{L}$). Dimensions can also be constructed by multiplying or dividing *unequal* dimensions. For instance, the dimension 'speed' is $\mathcal{L}/\mathcal{T}$; its units could be km/h, light year/s or μm/month.

We have seen that units can be multiplied and divided, and therefore dimensions can also be multiplied and divided. Units can also sometimes be added, for instance 3m+5dm+7cm. The units m, dm and cm have constant ratio's, and ther4efore they are equivalent. So they correspond to the same dimension, $\mathcal{L}$. This dimension can again be taken outside the brackets: if we have a quantity q=3m+5dm+7cm, then the dimension of q is $\mathcal{L}$ [▷49] .

2.7.3 Dimensions and Formulas

If two quantities are equal, their dimensions are also equal [▷50] . This means that we can, to a large extent, guess the form of mathematical expressions, merely by observing dimensions.

We illustrate this for a pendulum[20].

A pendulum is a weight on a chord, subject to gravity. The oscillation time T of a pendulum is an amount of time T [$\mathcal{T}$]. It could depend on the mass of the weight, m [$\mathcal{M}$], the length of the chord, l [$\mathcal{L}$], and the gravity acceleration, g [$\mathcal{LT}^{-2}$]. Suppose that the expression we are looking

[20]Another example is illustrated in the preceding image box. The candle flame image is taken from http://www.rgbstock.nl/photo/mC2yjPI/kaarsen+3

for reads

$$T = m^{\alpha} l^{\beta} g^{\gamma}, \tag{2.9}$$

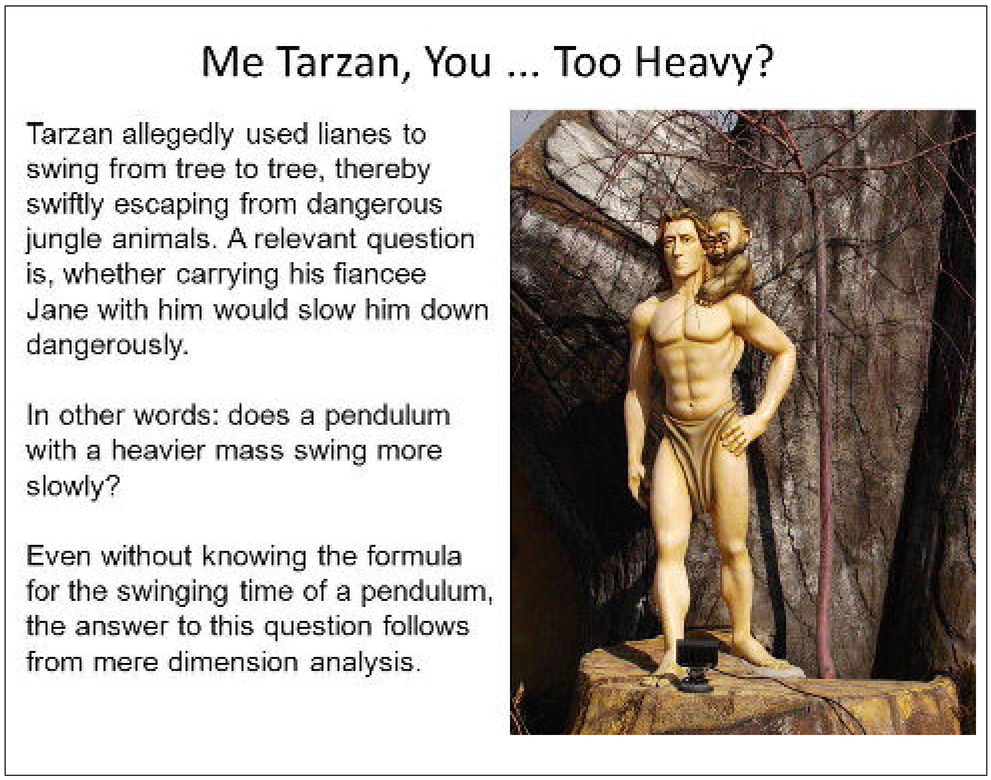
then we must find α, β, and γ. We equate the dimensions left and right; moreover, we use the fact that different dimensions are no multiples of each other ▷51 . Substituting the dimensions for mass, length and acceleration, we get[21]

$$\begin{aligned} \mathcal{T} &= \mathcal{M}^{\alpha}\mathcal{L}^{\beta}(\mathcal{L}\mathcal{T}^{-2})^{\gamma} \\ &= \mathcal{M}^{\alpha}\mathcal{L}^{(\beta+\gamma)}\mathcal{T}^{-2\gamma}. \end{aligned}$$

Equating the powers for $\mathcal{T}$, $\mathcal{M}$ and $\mathcal{L}$ we get:

$$\begin{aligned} \mathcal{T}: \quad 1 &= -2\gamma; \\ \mathcal{M}: \quad 0 &= \alpha; \\ \mathcal{L}: \quad 0 &= \beta + \gamma. \end{aligned}$$

So $\alpha = 0$, $\gamma = -\frac{1}{2}$, and $\beta = \frac{1}{2}$. The expression for the oscillation time therefore must have the form $T \propto \sqrt{\frac{l}{g}}$, consistent with the high-school formula $T = 2\pi\sqrt{\frac{l}{g}}$.



2.8 Mathematics: Functions of two Variables continued

In this chapter we discuss the conceptual model, consisting of concepts, properties and the relations between them. Relations often state that one quantity depends on several others. In Section 1.7 we introduced functions of multiple variables to help formalize such dependencies. In the same chapter, we saw that a common modeling purpose is *optimization*. We should be able, therefore, to optimize functions of several variables. In this section we will see that optimizing often entails to finding extreme values of functions, and we give some mathematical background for finding such extreme values.

[21] Tarzan's portrait was taken from http://commons.wikimedia.org/wiki/Category:Tarzan#mediaviewer/File:Harikalar_Diyari_Tarzan_06007_nevit.jpg

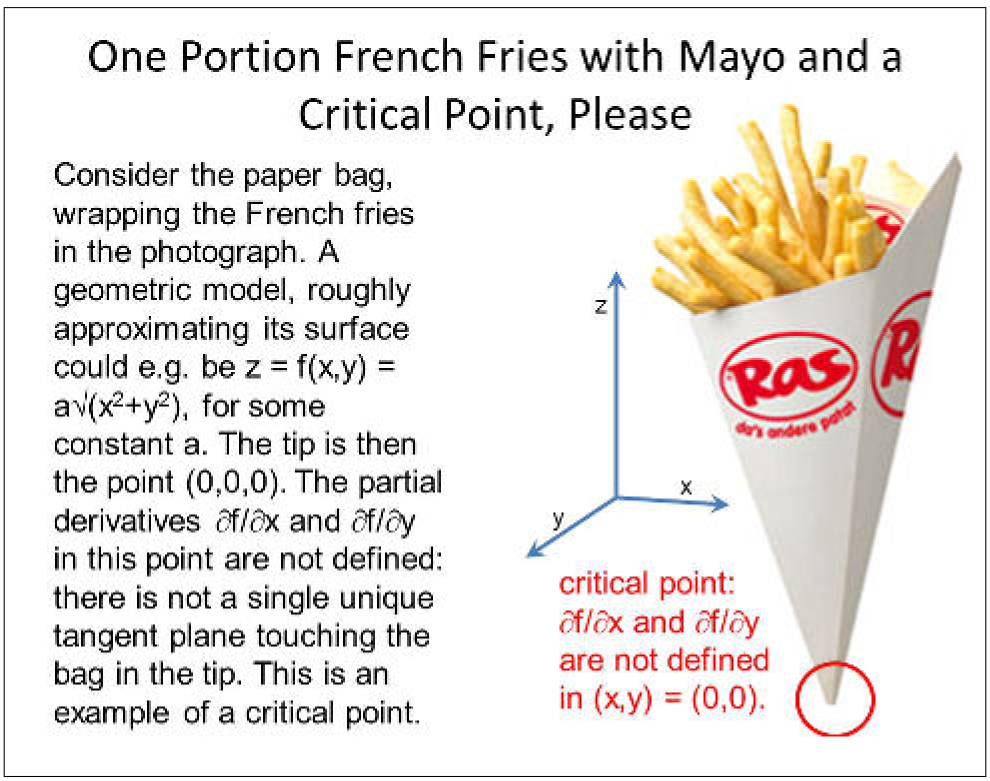
2.8.1 Extrema of Functions of several Variables

Local Extrema and Critical Points

First we recall some definitions and theorems for the case of a function of one variable.

A CONTINUOUS function f defined on a CLOSED, *bounded* interval $[a, b]$ attains both an absolute maximum and an absolute minimum on that interval ▷52 .

$f(c)$ is called a LOCAL (OR RELATIVE) MAXIMUM of f if $f(c) \geq f(x)$ for all x in some *open* interval containing c.
$f(c)$ is called a LOCAL (OR RELATIVE) MINIMUM of f if $f(c) \leq f(x)$ for all x in some *open* interval containing c.



A point ▷53 c in the domain of f is called a CRITICAL *point* of f if $f'(c) = 0$ or $f'(c)$ is undefined[22].

Suppose that $f(c)$ is a local extremum (maximum or minimum). Then c must be a critical point of f.

Suppose that f is continuous on the *closed* interval $[a, b]$. Then, each absolute extremum of f must occur at an endpoint (a or b) or at a critical point.

Bearing this in mind we can define also extrema for the case of a function of two variables.
The value $f(a, b)$ is called a LOCAL (OR RELATIVE) MAXIMUM of the function f if there is an OPEN DISK C centered at (a, b), for which $f(a, b) \geq f(x, y)$ for all $(x, y) \in C$.
The value $f(a, b)$ is called a LOCAL (OR RELATIVE) MINIMUM of the function f if there is an open disk C centered at (a, b), for which $f(a, b) \leq f(x, y)$ for all $(x, y) \in C$.

It can be seen in the Figure 'Contemplating a Local Extreme' that the tangent plane to a graph $z = f(x, y)$ at the local maximum or a local minimum is a horizontal plane. This means that both partial derivatives thee must be equal to zero. However, partial derivatives do not always exist. This reminds us of the case of single-variable functions: to refer to a point with zero derivative or undefined derivative, we used the term 'critical point'. We now generalize critical points to the

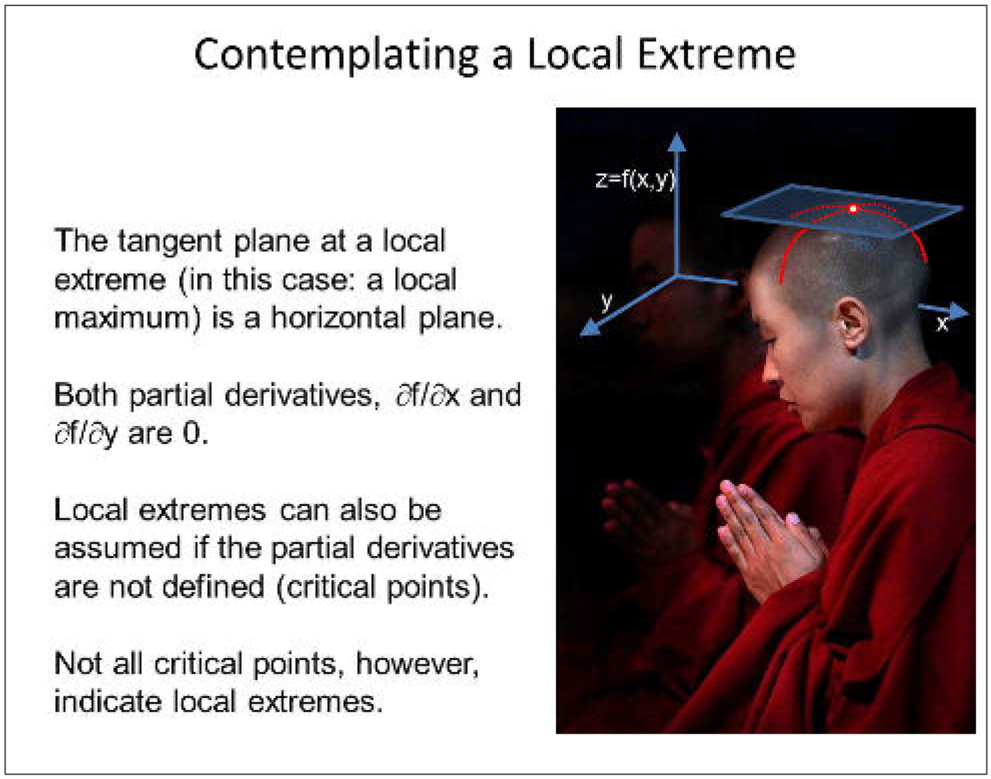
[22] The photograph of the French fries was taken from http://commons.wikimedia.org/wiki/File:RAS_frietzak.jpg?uselang=nl

case of functions with multiple variables:

The point (a, b) in the domain of f is called a *critical point* of the function $f(x, y)$ if either both partial derivatives are zero in the point (a, b) or at least one of the partial derivatives does not exist.

Go to this link to interactive explore a function surface $z = f(x, y)$ and investigate the relations between partial derivatives, tangents, the tangent plane and the normal vector.

Now we can formulate the following *theorem*[23].



If $f(x, y)$ has a local EXTREMUM at (a, b), then (a, b) must be a critical point. However, one should realize that critical points are only candidates to give an extremum. The theorem above is not 'if and only if', it is 'if..., then'. Some critical points are not EXTREMAL. The Figure 'There is Nothing Extreme in a Pringle' gives some examples. Here we can see a so-called SADDLE POINT. For instance, in the pringle-shaped surface $f(x, y) = x^2 - y^2$, both partial derivatives are equal to zero at $(0, 0)$. However, if we take the intersection of the function and the plane $x = 0$, then the function attains a maximum for $y = 0$. If we take the intersection of the function and the plane $y = 0$, then the function attains a minimum for $x = 0$. The same point acts both as a maximum (in dependence of x) and as a minimum (in dependence of y), and therefore is neither of the two according to our definition.

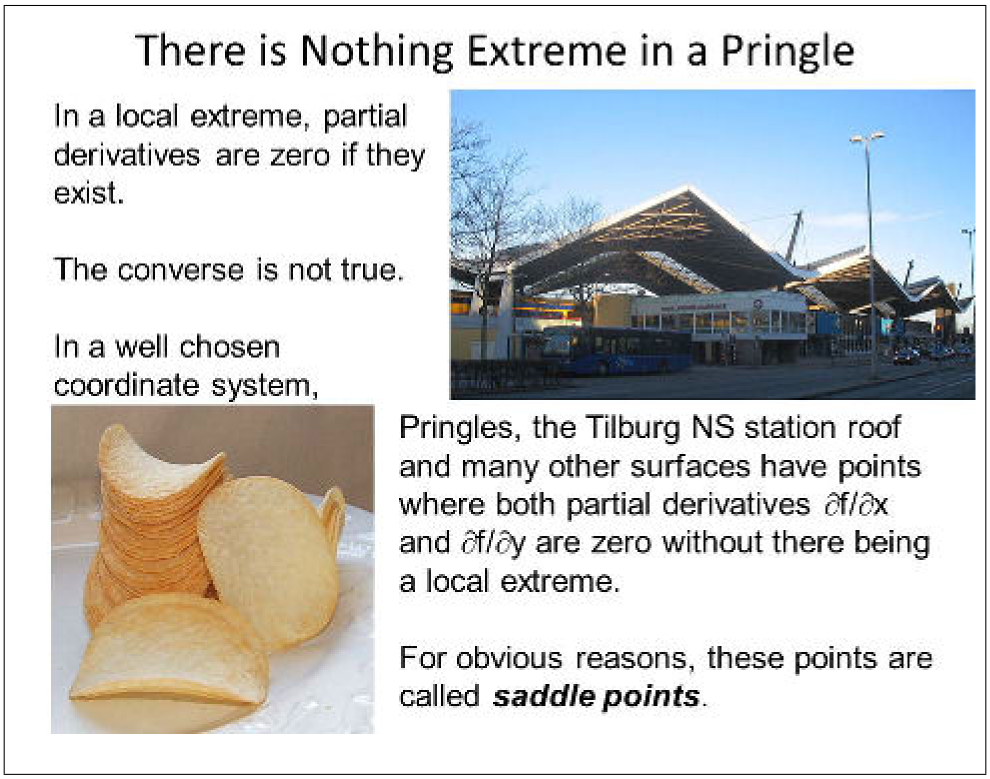
One can use visual inspection to infer whether a critical point is an extremum or not. There is also a mathematical test to determine whether a critical point with both partial derivatives equal to zero gives an extremum or not (the second derivatives test), but we will not discuss it in this section ▷54 . Sometimes also other arguments can be used (see Example 1 below).

To analyse the geometry of a saddle-type surface, click here. The shape of saddle surfaces can also be understood by studying its contour plot. To do so, click here. Presumably the simplest saddle-type surface is $f(x, y) = x^2 - y^2$; the part of this surface where $x^2 + y^2 < c$, for some constant c, could be an adequate model of a pringle. There are many more surfaces that feature saddle points, though. Practical examples occur in furniture such as chairs that have to give a

[23]The photograph of the meditating monk is taken from http://commons.wikimedia.org/wiki/File:Contemplative_Buddhist_monks_from_Bhutan_-_Flickr_-_babasteve.jpg?uselang=nl

comfortable fit to organic shapes found in the (human) body.

A simple example is the surface $f(x,y) = x^3 - pxy$ for various values of p. Click here to investigate the shape of the surface. To get a better understanding of the 3D shape of the surface[24], the orientation can be adjusted by dragging the mouse inside the image, rotating it over x, y or z-axes. The contour plot can be studied by clicking here.



Example 1 The point $(0,0)$ is the one and only critical point of the function $f(x,y) = \sqrt{x^2+y^2}$. Since $f(0,0) = 0$ and $f(x,y) \geq 0$, the value $f(0,0)$ is a minimum.

Example 2 The critical points of the function F in Expression 1.3 are $(-\frac{5}{3}, \frac{5}{3})$, $(-3,-3)$ (for these points both partial derivatives are zero) and all points (R_1, R_2) with $R_1 + R_2 = -3$ (for these points the partial derivatives do not exist). The value $f(-\frac{5}{3}, \frac{5}{3}) = \frac{112}{77}$ is a local minimum. The value $f(-3,-3) = 0$ is a local maximum. The points (R_1, R_2) with $R_1 + R_2 = -3$ do not yield finite extrema. This can be seen here using contourplots ▷55 . Notice that it may require quite some careful tweaking to find appropriate sets of contour values to actually see the behavior near the critical points: realize that the range of the function from Expression 1.3 is huge: for $R_1 + R_2$ close to -3 it varies from plus to minus infinity, whereas the values in $(-\frac{5}{3}, \frac{5}{3})$ and $(-3,-3)$ are close to 0. So the chance that an arbitrary level curve passes through an 'interesting' region of the domain is quite small. This illustrates a practical limitation to the use of numerically calculated level curves for investigating the behavior of functions of two variables.

Global Extrema

Next we will discuss global (or absolute) extrema.

We call $f(a,b)$ a GLOBAL (OR ABSOLUTE) MAXIMUM of f on the region R if $f(a,b) \geq f(x,y)$ for all $(x,y) \in R$.

We call $f(a,b)$ a GLOBAL (OR ABSOLUTE) MINIMUM of f on the region R if $f(a,b) \leq f(x,y)$

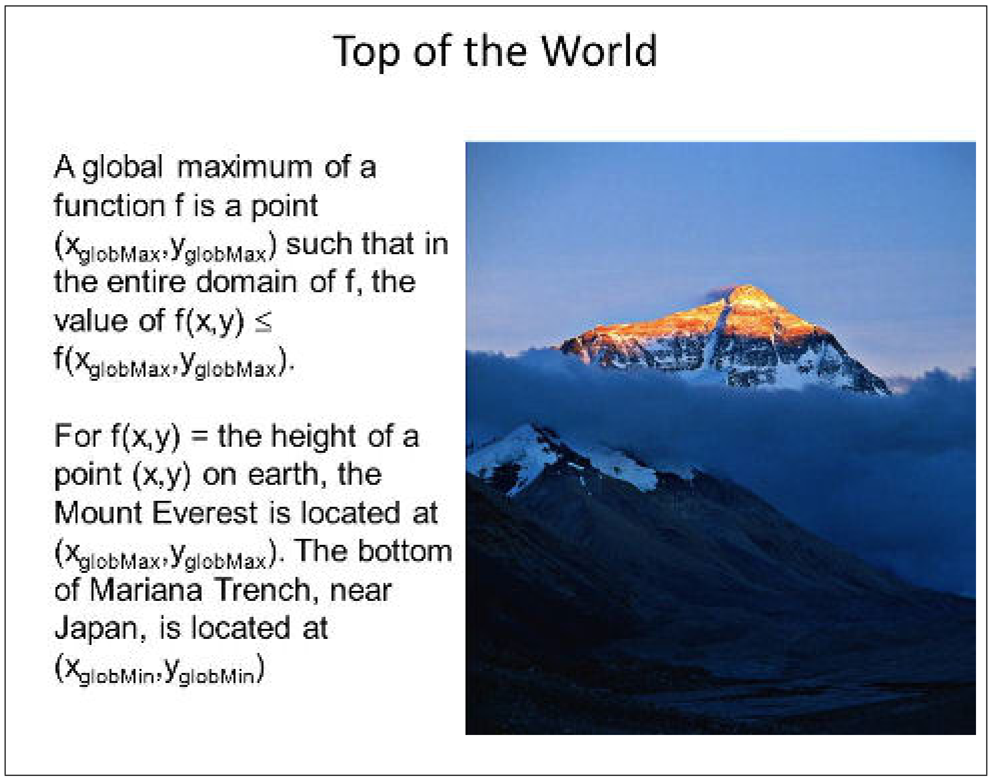
[24]The photograph of pringles is taken from http://commons.wikimedia.org/wiki/File:Pringles_chips.JPG?uselang=nl; the Tilburg station photograph is taken from http://commons.wikimedia.org/wiki/File:Centraal-Station_Spoorlaan_Tilburg_Nederland.JPG?uselang=nl

for all $(x, y) \in R$.

As for functions of one variable an *extreme value theorem* can be proven: Let $f(x, y)$ be acontinuous function, defined on a closed and bounded region R ▷56 . Then f possesses a global maximum and a global minimum, both either in critical points inside R or at its boundary[25].

2.8.2 Constrained Optimization

In Section 1.6 an example is given how the modeling process can be executed in practice. A crucial question that was mentioned in that section was 'with how little money can we safely illuminate the motorway?' The costs of illumination can be given as function of the quantities chosen. But the word 'safely' implies some constraints. The resulting illumination must satisfy the specification for a safe illumination.



There can be two kinds of constraints. The constraint might be an EQUALITY or an INEQUALITY.

An equality constraint is an equation; the found optimum should be such that it solves the equation.

An inequality constraint is an inequality; the found optimum should be such that the inequality is satisfied ▷57 .

For the case of *equality constraints* there exits a mathematical method called the method of Lagrange multipliers. This method is widely applicable ▷58 . It is a bit technical, though; here we will consider problems where a slightly simpler - though less general - technique is used. This technique amounts to substituting the constraint equality in the function that is to be optimized.

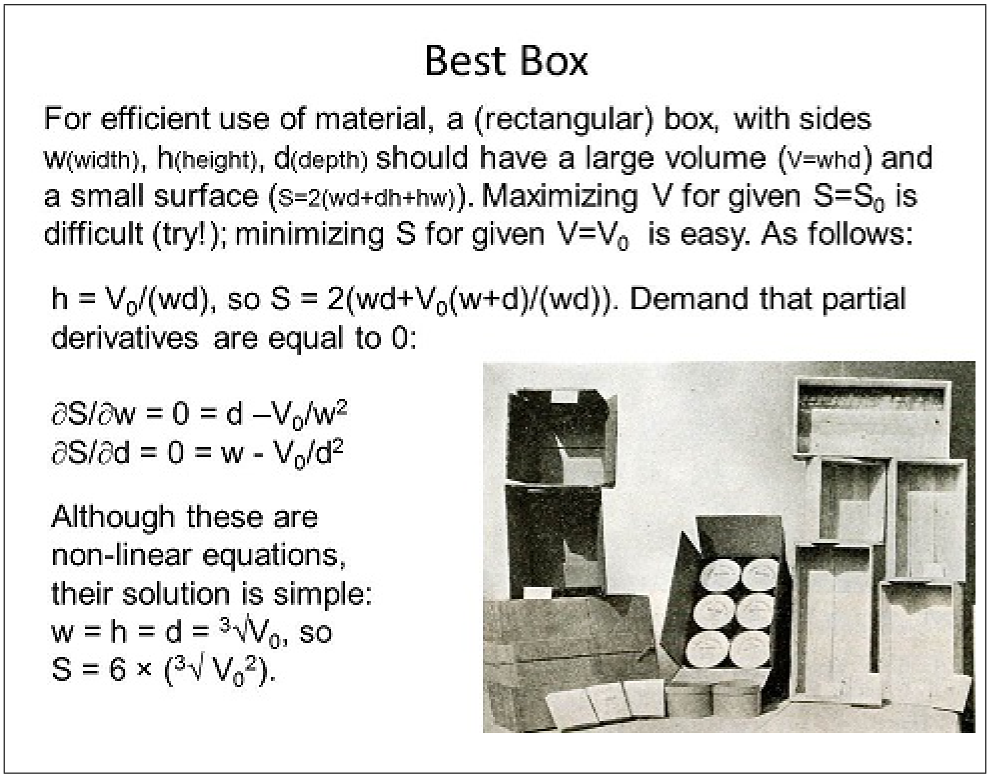
Optimization problems with *inequality constraints* are generally approached by attempting to transform the inequality constraints to equality constraints.

We proceed with some examples.

Example 1 (continued from Section Local Extrema and Critical Points): equality constraints. Consider again the function $f(x, y) = \sqrt{x^2 + y^2}$. We want to minimize this function under the constraint $2x + y = 3$. This can be solved us by substituting the equality constraint into the function

[25] The photograph of Mount Everest was taken from http://commons.wikimedia.org/wiki/Mount_Everest#mediaviewer/File:Mount_Everest_(topgold).jpg

to be minimized. Then the latter function reduces to a function of merely one variable. We find $g(x) = \sqrt{x^2 + (3-2x)^2} = \sqrt{5x^2 - 12x + 9}$. Minimizing $g(x)$ gives $x = \frac{6}{5}$ and so $y = \frac{3}{5}$. The minimum is equal to $\frac{3}{10}\sqrt{15}$; we easily verify that the solution indeed satisfies $2x+y = 3$. Another example is illustrated in [26] Figure "Best Box".



Example 2 (continued from Section Local Extrema and Critical Points): inequality constraints. Earlier, we introduced the function F in Expression 1.3, representing the average waiting time in the traffic lights model. Constraints in the traffic light model turn out to be inequalities. We will now introduce a rather general method to find an extremum in the case of inequality constraints. It amounts to reducing the optimization problem to a problem where inequality constraints are re-written in the form of equality constraints.

As follows.

For the traffic lights model we have the constraint that any traffic light should be green long enough to ensure that the entire queue of waiting cars is resolved at the moment that the traffic light turns red again. In Appendix **??** where Expression 1.2 is derived it is shown that this so-called 'no queue condition' leads to the following inequality constraints (f_0, f_1, f_2 in cars/minute; R_0, R_1, R_2 in minutes):

$$f_0 \geq f_1 + f_2, \quad f_0 R_2 \geq (2R_0 + R_1 + R_2) f_1 \quad \text{and} \quad f_0 R_1 \geq (2R_0 + R_1 + R_2) f_2. \tag{2.10}$$

For the values of $f_0 = 40$, $f_1 = \frac{40}{3}$, $f_2 = 5$, and $R_0 = 1.5$, used to derive Expression 1.3, this gives

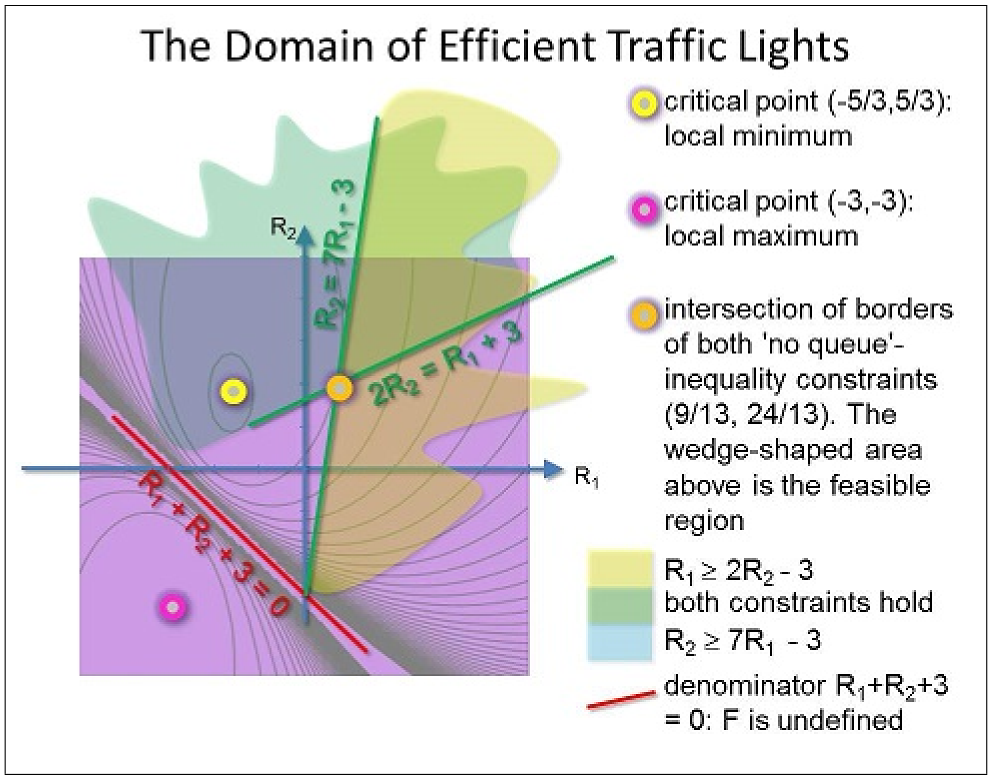
$$2R_2 \geq 3 + R_1 \quad \text{and} \quad 7R_1 \geq 3 + R_2. \tag{2.11}$$

In Figure 'The Domain of Efficient Traffic Lights' both lines, $2R_2 = 3 + R_1$ and $7R_1 = 3 + R_2$ are indicated. These lines are borders of the regions where each of the two inequalities hold (the yellow and blue zones); both inequalities hold in the overlapping wedge (green zone). This is the so-called FEASIBLE REGION. The corner point of the feasible region is the intersection of $2R_2 = 3 + R_1$ and $7R_1 = 3 + R_2$, that is the point $(\frac{9}{13}, \frac{24}{13})$.

[26] The photograph of various cardboard boxes is taken from http://commons.wikimedia.org/wiki/File:US_Dep_Agriculture_Bulletin_N_456_Marketing_Creamery_Butter_Fig_10.png

We see that none of the two critical points, calculated in Section 2.8.1, i.e. the local maximum of F in $(\frac{-5}{3}, \frac{5}{3})$ and the local minimum of F in $(-3, -3)$, are in the feasible region. There are no critical points at all within the feasible region.

The only place, therefore, where a local minimum of F possible could occur is *on the boundary of the feasible region*.



The feasible region is the region where all *inequality* constraints together hold; its boundaries are lines where *equality* constraints hold. So now we can apply the method that we used in example 1, that is: we substitute each of the equality constraints ($2R_2 = 3 + R_1$ and $7R_1 = 3 + R_2$), into F from expression 1.3, yielding the two 1-variable functions $F_1(R_1) = \frac{6}{77}\frac{105R_1^2+42R_1+63}{8R_1}$ and $F_2(R_2) = \frac{6}{77}\frac{30R_2^2+12R_2+18}{3R_2}$. For each of these two, critical points can simply be found, yielding the solutions $R_1 = \sqrt{\frac{3}{5}}$; $R_2 = 7\sqrt{\frac{3}{5}} - 3$ and $R_2 = \sqrt{\frac{3}{5}}$; $R_1 = 2\sqrt{\frac{3}{5}} - 3$.

Of these two candidate solutions, we see that the second one is entirely outside the feasible region, so the final solution for the traffic light problem is that $R_1 = \sqrt{\frac{3}{5}} \approx 0.77$ minutes and $R_2 = 7\sqrt{\frac{3}{5}} - 3 \approx 2.42$ minutes.

At this point, we may ask if we are certain that there is no way to achieve an even shorter average waiting time. As follows: EMIEL - TO BE COMPLETED (TO BE DECIDED: WILL THE EXPLANATION GO HERE OR IN AN ENDNOTE?).

The situation, as in the above traffic lights example, where constraints occur in the form of inequalities, is quite common. For instance, in the street illumination model, the height of the lamps should have a lower bound and an upper bound. The same holds for the illumination on the surface of the road.

It is instructive to recapitulate the approach to follow in such cases:

- *find feasible region* : what is the part of the domain where all inequality constraints are satisfied? Check if the feasible region is non-empty - otherwise, there is no solution to the problem. Each part (=segment) of the border of the feasible domain is the solution of one of the constraints, written as *equality* constraints.

- *for each of the segments, substitute the associated equality constraint into the function to be optimized* . The result is an unconstrained optimization problem, and we can search for critical points in the standard way by calculating (partial) derivatives and setting these to 0.
- *for each of the critical points found, check if they are in the feasible region* . If so, these are candidate solutions.
- *check if there are any critical points inside (=not on the borders of) the feasible region* , or prove that there aren't. If there any, add them to the set of candidate solutions.
- *check which of the candidates is the most extreme one* . The latter is the final solution of the constrained optimization problem ▷59 .

Finally, notice that, in general, many constraints can be given: unlike in the case of solving equations, where as a rule the number of equations should equal the number of unknowns to achieve a unique solution, there is no immediate relation between the number of inequality constraints and the number of occurring quantities.

This completes our treatment of analytic, i.e. symbolic, methods for finding local and global extrema, both in the case of equality and inequality constraints, for a function of multiple variables. These methods rely on the function being differentiable. In actual modeling situations, this assumption is often not valid. In later chapters, we will learn methods that also work in the case where the function to be optimized is not differentiable, or in the case where we have several functions that need to be optimized simultaneously.

2.9 Summary

- The *conceptual model* is constructed in stage 2 (conceptualization) of the modeling process;
- The conceptual model consists of *concepts*; *entities* in the modeled system are represented by concepts;
- A concept is a *bundle of properties*, every property consisting of a *name* and a *set of values*: this set is the *type* of the property;
- Concepts can have *relations*; the concepts and relations together form the *conceptual model*, usually drawn as an entity-relation graph. Relations can also exist between the properties of concepts. The conceptual model is constructed in 4 steps:
  - establish concepts;
  - establish properties;
  - establish types of properties;
  - establish relations.
- Sets of values can be *bound* in different ways to properties, e.g. as choices, as results from measurements, or as desired outcomes;
- Values, occurring in the type of a property, can be concepts of their own;
- *Quantities* are properties, where the concept they are properties of is disregarded;

- Allowed mathematical operations on quantities depend on their *ordering*; we distinguish *nominal* (no order), *partial ordering* or *total ordering*. For totally ordered scales, we further distinguish *interval scale* and *ratio scale*;

- *Measuring* amounts to *counting* the number of units of some sort fit in the measured item. Units can have constant ratio's (e.g., 1m=100cm);

- Sets of units that have a constant ratio are called equivalent. A *dimension* is an equivalence class on units;

- Operations on units follow the operations on quantities (dimensional analysis);

- Using the dimension of quantities, the form of a mathematical relation between them can often be derived.

- Functions of two variables:

  – The value $f(a, b)$ is called a *local (or relative) maximum* of the function f if there is an *open disk* C centered at (a, b), for which $f(a, b) \geq f(x, y)$ for all $(x, y) \in C$. The definition of a minimum is similar.

  – The point (a, b) in the domain of f is called a *critical point* of the function $f(x, y)$ if either both partial derivatives are zero in the point (a, b) or at least one of the partial derivatives does not exist.

  – If $f(x, y)$ has a local extremum at (a, b), then (a, b) must be a critical point.

  – We call $f(a, b)$ a *global (or absolute) maximum* of f on the region R if $f(a, b) \geq f(x, y)$ for all $(x, y) \in R$. Similar for a global minimum.

  – If $f(x, y)$ is continuous on the *closed and bounded region* $R \subset \mathbb{R}^2$, then the function f has both a global maximum and a global minimum on R. Moreover, a global extremum occurs only at a critical point or at a *boundary point* of R.

  – An optimization problem with an equality constraint can very often be solved by substitution of the equality constraint into the function.

  – An optimization problem with inequality constraints can often be reformulated such that it reduces to a optimization problem on a closed and bounded region, the boundaries of which are defined by the inequality constraints.

2.10 Learning goals

2.10.1 Knowledge

You should know the meaning of the terms concept, property, value, type, relation, quantity; the various kinds of ordering and the mathematical operations that are allowed for each form of ordering. You should know and understand the mathematical operations on units, and the various rules that apply. You should understand the concept dimension and dimension analysis / dimension synthesis. You should possess a working knowledge of the optimization of real functions of two variables as explained in Section 2.8 and of the relevant sections in the calculus book of either *Adams* or *Smith & Minton* (see below).

Adams: EMIEL: TO BE COMPLETED

Smith & Minton: §12.7 without the second derivative test and without the method op steepest ascent. §12.8 deals with constrained optimization. However the method of Lagrange multipliers is applied and this not part of the material for the exam. You must understand the basic ideas given in Section 2.8.2.

2.10.2 Skills

In this section, with 'problem' we mean: a problem that does not require domain-specific knowledge exceeding your present knowledge.

For a conceptual model, needed to solve a problem in some domain, you should be able to identify the most important concepts and their properties, and you should be able to denote the values for these properties using set notation. You should be able to construct an entity-relation graph depicting the concepts and the most important relations between them. You should be able to assess the types of the occurring quantities, determine if they can be ordered and in what sense, and you should be able to make a justified choice for the units to use. You should be able to convert arbitrary formulas from one unit system to another.

You should be able to check derivations and formulas using dimensional analysis, and in simple cases you should be able to derive formulas using dimension synthesis.

For a given function of two variables, you should be able to find the critical point of the function for a given region, classify the critical points and find the extrema. You should also be able to find the extrema when (in)equality constraints are present, applying the basic ideas given in Section 2.8.2.

2.10.3 Attitude

When confronted with a problem that might benefit from a formal approach, you should consider to use a model. When approaching a problem by using a model, you should have the attitude to build a conceptual model. You should have the inclination to formulate the properties of the occurring concepts in terms of quantities with well-defined types, and you should typically denote these in terms of set-notation. Whenever you encounter a formula, you should check its dimensional consistency. If you need to optimize a function, you should consider the possibility of using the analytical tools given in this chapter.

2.11 Questions

1. We use the terms 'entity' and 'concept'. They are almost the same. Explain in your own terms what the difference is.

2. (*) Is it possible to talk about an entity that is not a concept?

3. What can you say about the arity of `isA`, `hasA`, `specializesTo` and `partOf`?

4. Is it possible to access a concept that has no name?

5. Consider the words 'color', 'tomato', 'red'. Which is a concept, which is a property name, which is a value, which is a type, which is a function?

6. If 'red' is a concept, give some examples of properties and values. Can you think of a property of 'red' such that 'tomato' is a value?
7. We state: 'A property is a function of the concept it belongs to'. Explain.
8. Explain in your own words the meaning of 'intersubjective'.
9. Explain in your own words the meaning of 'segmentation'.
10. What is the difference between $\{4\}$ and 4 and '4'?
11. What do we mean by 'opposite relations'?
12. What is transitivity?
13. What is the meaning of a 'range', when we use this term in defining a type?
14. What are the 4 steps of constructing a conceptual model?
15. We say *most of* the construction of the conceptual model is part of the conceptualization stage (Section 2.5). This suggests that some part of the construction conceptual model is part of an other stage of the process model of Section 1.4. Which of the 4 steps is that, and which stage does it belong to?
16. We explain the 4 steps of constructing a conceptual model in Section 2.5. In step 3, we establish the types of properties. We see a number of different ways value sets are bound to properties. List them, and explain their differences.
17. Consider the relation `rel(a(3),b(n))`. What do the symbols in brackets mean?
18. What is an entity-relation graph? Give a reason why this name is wrong.
19. What is the difference between a quantity and a property?
20. What do we mean by the type of a quantity?
21. What is the difference between things, occurring in formal expressions, written in `this font` and in *this font*?
22. What is a compound type?
23. Why is the square of the perimeter in a rectangle at least 8 times the area?
24. What is the difference between an ordinal scale, an interval scale and a ratio scale?
25. What is an equivalence class?
26. What is the meaning of $p_{1;3}$ and $p_{3;2}$ in Expression 2.5?
27. We regularly encounter factors such as $p_{m;U}$. What is the meaning of U in these expressions?
28. Explain in your own words what a dimension is, try to avoid the word 'equivalence class'.

2.12 Exercises

1. 'Eiffeltower' is a concept. Give a bundle of properties, defining 'Eiffeltower' such that it is a singleton, and give a bundle properties such that it is a set with multiple elements.

2. In the street lamp example, given the problem and the purpose, add at least two concepts.

3. In the street lamp example, for each of the concepts, either

   - add one or more meaningful properties, given the purpose, or
   - give an argument why no more properties are necessary.

4. In Section 2.5, we introduce properties for the street lamp example. After having read Chapter 5, find the category numbers (I, II, III, or IV) for each of the properties.

5. In the street lamp example, give at least 5 relations between properties.

6. In the street lamp example, an important quantity is the distance between adjacent street lamps. This quantity could be a property of two concepts in our model. Which two? Give advantages and disadvantages of both choices.

7. In the street lamp example, there is a 3-ary relation, sees. Answer the following questions. Hint: think of a street in the rain, and and compare this with a dirt road ▷60 .

   (a) Explain exactly what the meaning of sees is. What should this relation calculate?

   (b) In the current version of the conceptual model, there are not enough properties to fully define sees. Name some of the properties we additionally need.

   (c) When is it possible to replace sees by two other relations that are both 2-ary?

   (d) When is it not possible to replace sees by other relations that are both 2-ary?

8. Give at least 4 examples of relations (not necessarily restricted to the street lamp casus) that are 3-ary, where two can be replaced by two 2-ary relations, and the other two cannot.

9. Similar to the street lamp casus, build a conceptual model for supporting the decision whether or not the owner of a private house should have solar panels installed. Go over all four steps in the construction of the conceptual model.

10. Similar to the street lamp casus, build a conceptual model for verifying if a governmental health agency has sufficient medicine in stock to remedy the outbreak of a contagious viral infection. Go over all four steps in the construction of the conceptual model.

11. Similar to the street lamp casus, build a conceptual model to help planning the supply of fresh vegetables to a super market. Go over all four steps in the construction of the conceptual model.

12. We discuss three types of totally ordered scales: a ratio scale, an interval scale, and a scale (such as Mohs' scale) which is not even interval scale. Give an example of a problem that can be approached by Mohs-type scale(s).

13. Give three examples of ratio scales, at most one of them from physics, and give an argument why they are ratio scales.

14. (*) Proposition: 'a ratio scale must consist of rational numbers or real numbers'. Give arguments in favor and against this proposition.

15. We derive the formula for the oscillation time of a pendulum from analyzing dimensions. Do the same for a mass-spring system.

16. The lens-makers formula (if necessary, consult Wikipedia) cannot be derived using dimensional analysis. Why not? What does this example teach you about the usefulness of dimensional analysis for deriving formulas?

17. You give a pizza-party: you invite N people. Everybody consumes S slices of pizza. A complete pizza contains P slices. The price for a pizza is E. Using dimensional analysis, give an expression of the amount of money every guest has to pay.

18. A farmer possesses chicken; these lay eggs and consume chicken food. Which quantities are needed to decide whether this farmer should purchase an additional chicken? Using dimensional analysis, derive a formula to answer the question whether or not the farmer should purchase another chicken.

19. Suggest a problem that can be solved in the same way as Problem 18, and solve it.

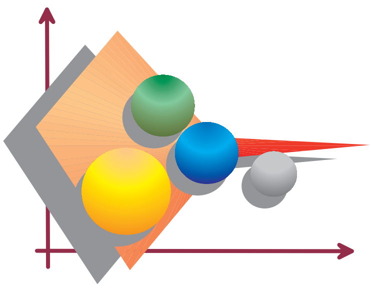
Exercises concerning Section 2.8

Adams:

Smith and Minton: §12.7: 3, 6, 12, 14 (?ACCEL?), 31, 35, 49, 50; §12.8: 9, 12, 15, 30, 47, 59, 60.

Chapter 3

Time for Change



'Now is past for the future and future for the past'

You see an empty green rectangle. Suddenly, from the left enters a white, shiny spherical object, followed by a red, similarly shaped one from the right. They get closer and closer, and then they collide. Their routes have drastically changed: the white ball leaves the scene at the top, whereas the red ball vanishes in downward direction. Next we see the same movie again, now played in reverse. The white ball comes in from above, the red one from below and after interaction they leave in horizontally opposite directions. We witness an equally plausible rendition of two colliding billiard balls. Apparently, for physical processes such as simple collisions between point masses or rigid spheres, the direction of time is irrelevant. Only when we look at a larger scale, say, of a complete carambole, there is a difference between past and future. The hit with the cue comes first, initiating the first ball's movement, and friction and collision losses gently slow everything down until the balls come to a standstill after a while. Time reversal at this scale would cause motionless balls gently to acquire speed, until they miraculously bump against a cue, held at exactly the right place by the billiard player, who then plans the shot ... which is clearly in conflict with our daily experience.

3.1 Change needs Time

Physical time at the micro scale is REVERSIBLE; this is not true if friction or other complex processes come into play. We say that time is MICRO-REVERSIBLE, and MACRO-IRREVERSIBLE. Physical time comes in a symmetric and a non-symmetric version. The same is true for human perception. To some extent, we can anticipate the future. Sometimes, we know what will happen, which enables us, for instance, to catch a ball. In most cases, however, the difference between past and future is obvious: we sense the difference between remembering and anticipating[1].

Both in physics and in the subjective experience, time often has a direction, sometimes called the *arrow of time*. The arrow of time relates to *cause* and *effect*. Causes, effects and the advance of time together form the ingredients of PROCESSES. A *process* is something that involves change over time. Events in processes may be linked by cause and effect-relations. An effect can never precede its cause, but not in every pair of events where one precedes the other, the first causes the second.

3.2 Introduction to Processes

In a process, things sometimes happen one after the other (I switch on the radio; next I switch on the light); sometimes one is a result of another (I switch on the light, so the room gets illuminated), sometimes unrelated (I sneeze, and somebody else switches on the light). A process may have conditional steps (if the sun shines brightly, I may decide not to switch on the light), and some conditions are purely time dependent (it is November, and it is four o'clock, so I switch on the light). Sometimes things need to be repeated (the radio is old, so I need to repeatedly slap until it starts working). Finally, sometimes we need to wait for things, either for a known amount of time (I wait until 20:00 to hear the news on the radio), or for an unknown amount of time (I wait to go out until the postman arrives).

All these sorts of things need to be considered if we want to describe a process.

3.2.1 States and State Charts

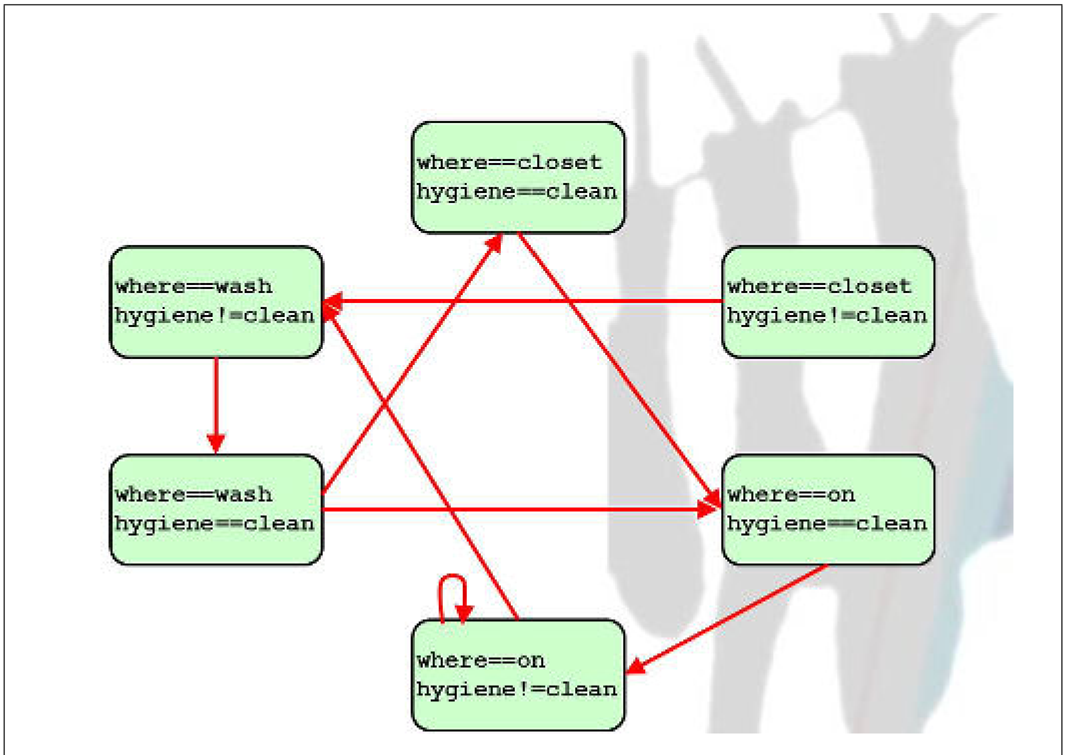
By means of a conceptual model, as described in Chapter 2, we describe the state of affairs of a modeled system in terms of concepts, properties, values and relations. This is fine as long as nothing changes. If something does change, we could start again and set up a conceptual model

[1]The image of the snail has been taken from http://www.rgbstock.nl/photo/mhXQ074/Slak+2

for the situation after the change, but this is obviously impractical in the case of many changes. We need another device to denote processes, involving changes in a modeled system.

We will use the idea of a so-called state chart to do so. A STATE CHART is a graphical means to denote things that take place over time. A state chart is a collection of STATES. A state is a snapshot of a system. That is: a representation of that system, containing all its concepts, their properties and the current values of these properties. In other words: conceptual models, as encountered in Chapter 2, are descriptions of states.

.



A state chart depicts the relevant transitions for a dynamical process.

Figure 3.1: A sock, described by the properties 'where' (with values 'on', 'closet', and 'wash'), and 'hygiene' (with values in between 'clean' and 'dirty'). (Illustration source: Kees van Overveld)

To illustrate 'state', consider the lifecycle of a sock. We characterize a sock by two properties. First, its location (called `where`) with values `closet` (stored in the closet), `on` (on a foot), and `wash` (being washed). Next its hygienic condition (called `hygiene` with a range of values, {`'clean'` ... `'dirty'`}.

States differ with respect to which values are currently assumed by the properties. This is called BINDING. 'Value v is currently bound to the property p' means that, at this time, v is the value of p. For a clean sock, laying in the closet, `where` is bound to `closet`, and `hygiene` is bound to `clean`.

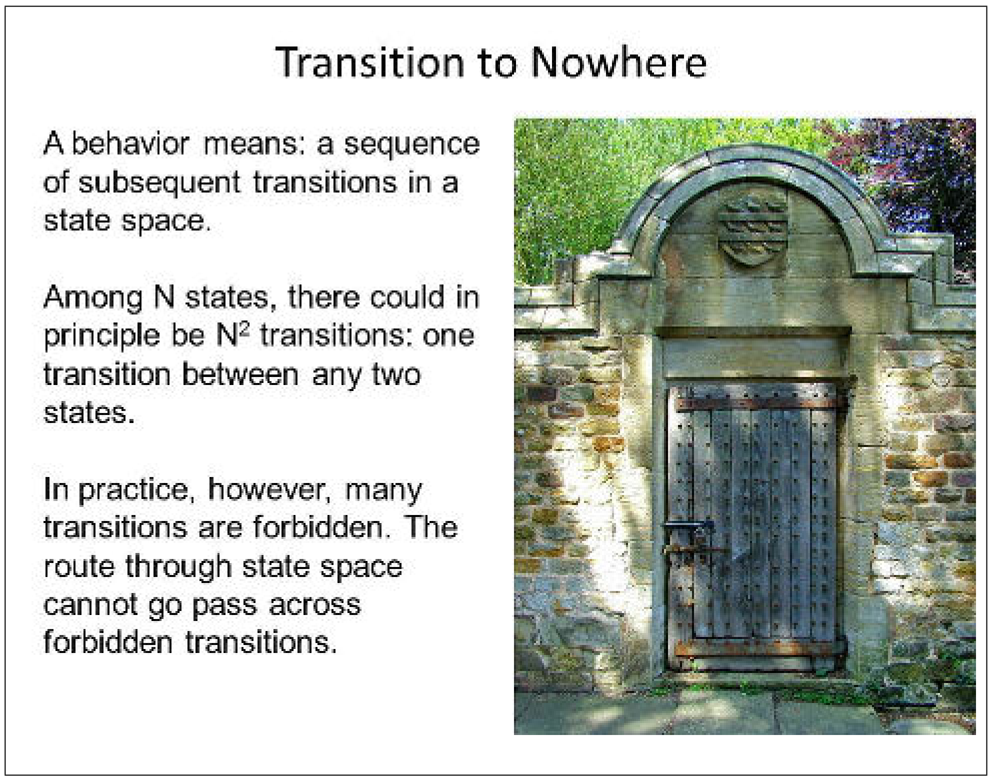
For any conceptual model there is a number of possible states. If, for the sock in our conceptual model, we only distinguish the hygiene values 'clean' or 'something else than clean', there is a total of 6 states. These correspond to all possible combinations of bindings, namely: `where==closet, hygiene==clean`, `where==closet, hygiene!=clean`, `where==on, hygiene==clean`, `where==on, hygiene!=clean`, `where==wash, hygiene==clean`, and `where==wash, hygiene!=clean`. We use the notation `a!=b` to say that property `a` is bound to some other value than `b`.

When a system goes from one state to another state, this corresponds to a change in binding. We call this a STATE TRANSITION ▷61 , or 'transition' for short.

Everything that happens to a sock, affecting its location or its hygienic condition is a transition.

We assume transitions to go instantaneously. In the sock example: there is one, indivisibly short instance where value `closet` for property `where` is replaced by `on`, et cetera. The fact that actually putting on a sock may take several seconds, is not accounted for in the state chart. The time *in between* subsequent transitions, however, can be indefinitely long. A state, characterized by `where==wash` may take an hour or so; a state characterized by `where==closet` may take arbitrarily long.

Some transitions may occur freely; most transitions, however, are subject to rules or conditions[2]. For instance, a transition from `hygiene==clean` to `hygiene!=clean` can only take place if `where==on`: socks in the closet don't get dirty, and socks being washed get from dirty to clean. Rules or conditions can forbid a transition to take place at all: the transition from (`where==on, hygiene!=clean`) to (`where==closet, hygiene!=clean`) is forbidden since we don't put dirty socks back into the closet.



Sometimes we may want to express that a transition takes place going from a state to *the same* state. For instance, we have a state that is defined by `where==on, hygiene!=clean`. If the sock gets dirtier, the value of `hygiene` changes, but since it was not clean before, it will stay not clean. So there is a transition where `hygiene` assumes a dirtier value, but since our choice of states only distinguishes the cases `hygiene==clean` and `hygiene!=clean`, this does not involve moving to another state. This trick helps to prevent the number of states becoming too large; in the next Section we will see that a large number of states is an often occurring problem in dealing with change in models.

State Space

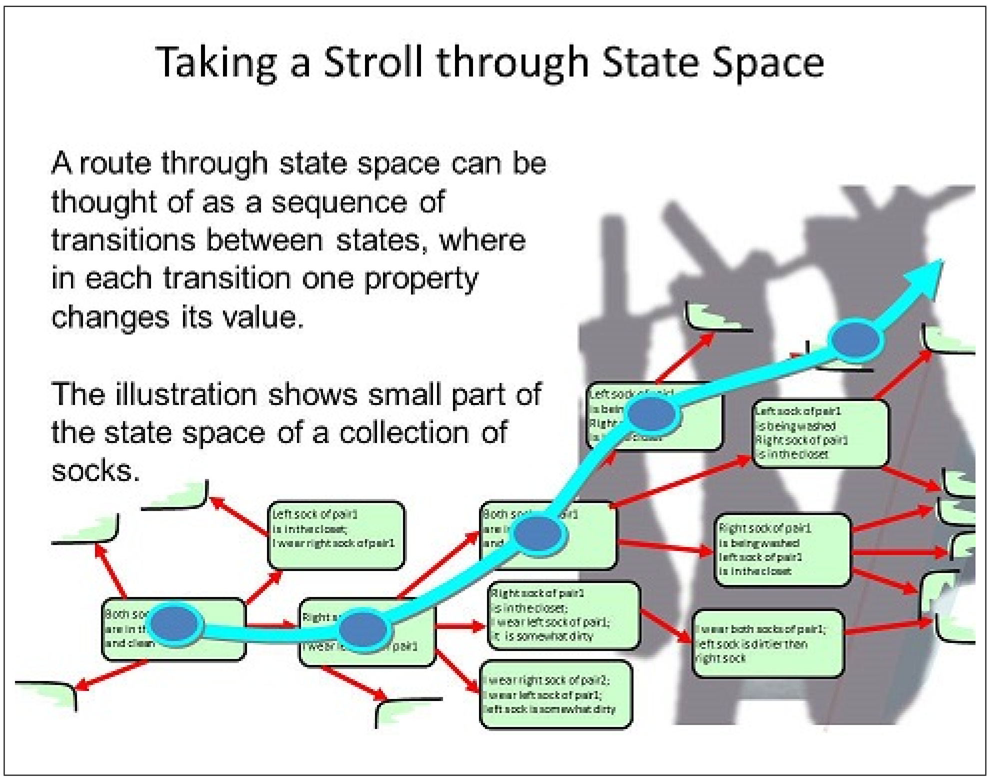
States, defined as the binding of values to properties, are numerous. Two properties, one with 3 values and the other one with 2 values gives $3 \times 2 = 6$ states. Adding one further property to a concept in the modeled system multiplies the number of possible states of that system with a factor equal to the number of different values this new property can assume. E.g., if we allow socks to have holes, i.e. introducing the property `nrHoles` with type $\{0, 1, 2, \texttt{many}\}$, the number

[2] The image of the closed gate is taken from http://www.rgbstock.nl/photo/mjQB46C/gesloten+hek

of states becomes $3 \times 2 \times 4 = 24$. If the conceptual model of a system contains N_p properties, and property i has m_i values, the number of states N_s is no less than

$$N_s = \Pi_{i=1}^{N_p} m_i. \tag{3.1}$$

All states of a system together are called the STATE SPACE. When a system develops over time, the changes it undergoes form a route through its state space, assuming one state after another. A route through state space typically visits only a limited number of states, going from some initial state to a final state. It is sometimes called TRACE or a BEHAVIOR ▷62 .



The maximal number of possible transitions for N_s states is N_s^2, including transitions leading from a state to the same state. Transitions may be forbidden, however, yielding a number of *allowed* transitions that is often much less than N_s^2. We saw the example in the life of a sock where the `location` of a dirty sock cannot change to value `closet`. Due to this restriction, and many others, the number of transitions actually allowed is merely 8 instead of the maximum of 6×6=36. In a physical system, such as colliding billiard balls, energy conservation is an example of a restriction. It forbids that the sum of kinetic energies of the balls after the collision exceeds their sum of kinetic energies before the collision. This is also an example of a forbidden transition.

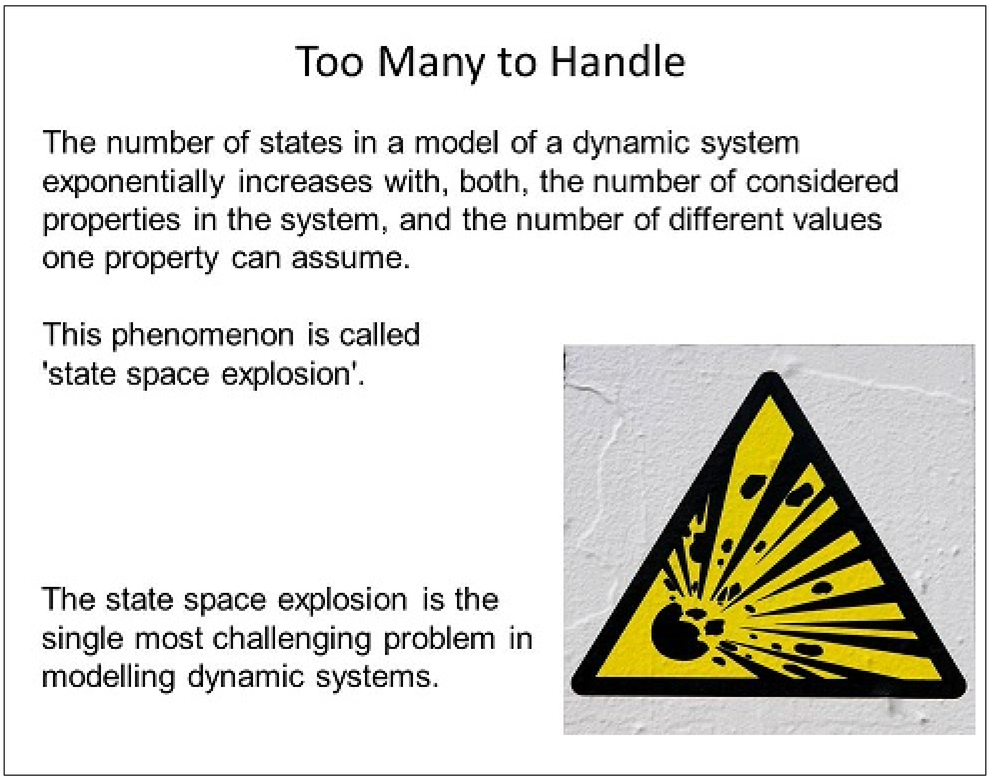
Transitions can be permitted or forbidden for various reasons: in physics constraints often are conservation laws; financial transitions may be constrained by credit limits; chemical reactions may not occur because of the absence of some reagens or catalyst, and rules of good housekeeping dictate what transitions for socks are forbidden.

Clicking this link starts the ACCEL modeling environment with a script running that allows you to interactively experiment with the state chart for the lifecycle of a sock.

Assuming that a state transition takes no time, we can define that at every transition, only one property changes its value, which makes the state chart easier to understand. In the case of a billiard ball collision: there is no difference between saying that the red ball and the white ball change their velocity at the same instant, or that there is an infinitesimal delay between the two changes ▷63 .

The size of the state space is immense; the collection of possible routes in it is even larger. Indeed, if we consider only routes with a length of N_t transitions, the number of routes, N_R is

$$
\begin{aligned}
N_R &= \Pi_{i=1}^{N_p} \Pi_{j=1}^{N_t} m_i \\
&= (\Pi_{i=1}^{N_p} m_i)^{N_t}
\end{aligned}
$$



This number grows explosively[3] both with N_p and with N_t; hence the name STATE SPACE EXPLOSION. The state space explosion is the most challenging problem in modeling dynamical systems. For any non-trivial system, it is intractable to account for all possible routes explicitly.

State Space Reduction: Symmetry

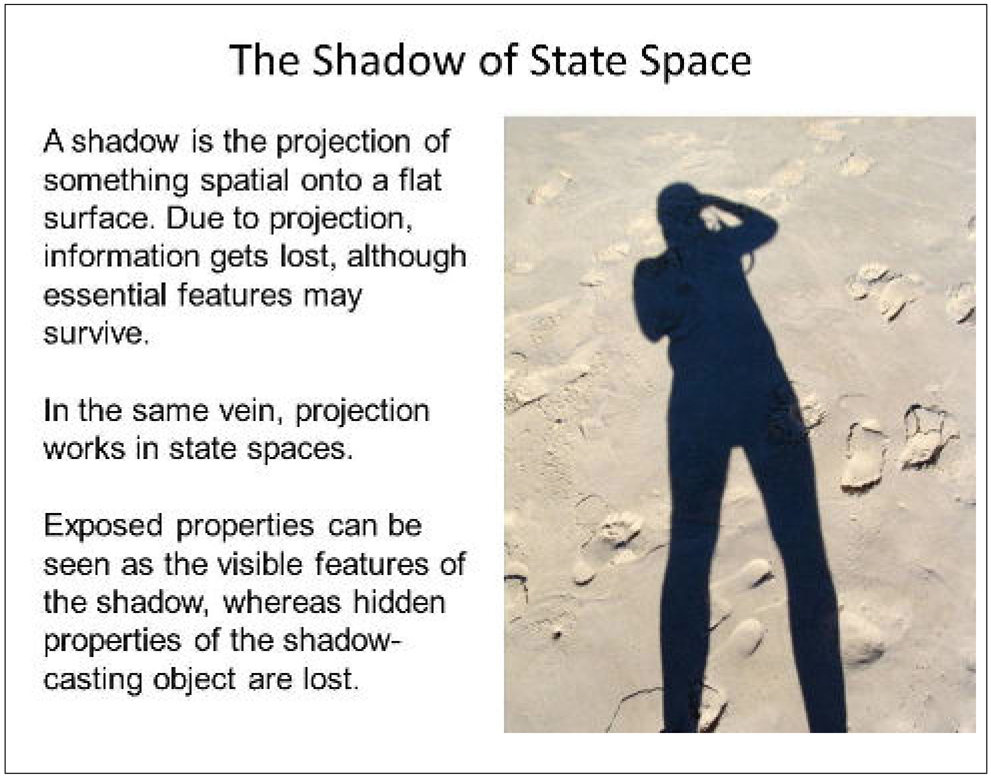
To reduce the size of the state space, SYMMETRY can sometimes be used. Symmetry is the condition that an entire system can be known even if only part of it is given. For instance: if only the left hand part of a mirror-symmetric piece of clothing is drawn, a capable tailor can make the entire piece. This is an example of SPATIAL symmetry. Temporal symmetry, for instance, applies in the example in the introduction of this chapter where the behavior of a billiard ball collision is the same when time is reversed (time reversal symmetry), or to express that the behavior of billiard balls in a carambole won't be different if that carambole would take place at some later moment in time (time shift symmetry). Temporal symmetry also applies to periodic phenomena. Knowing the motion of one swing of a friction-less pendulum is enough to know the entire behavior. Symmetries, other than spatial and temporal exist: permutation symmetry, for instance, occurs when we swap the two white bishops in a game of chess.

Symmetry can help reducing the size of the state space of a system. Suppose we want to verify the correctness of the Dutch railroad signaling system. A conceptual model contains representations for all signals, all trains and all railroad switches. 'Correctness' can be described in terms of states: it consists of requirements such as

[3] The image of the explosion hazard warning sign is taken from http://www.rgbstock.nl/photo/of8Lrbu/Gevaar

- two trains shall never occupy two adjacent railway segments;
- for a switch, the signal in at most one of the branches is green;
- signals in all railroad segments leading to one with a red signal carry orange signals;
- ...

If every REACHABLE state in the state space satisfies all conditions above, the signaling system is formally correct. A *reachable* state means: a state to which a permitted transition, departing from an other reachable state, leads. There is always at least one reachable state in a dynamic process; this is called the INITIAL STATE.



Symmetry helps reducing the state space of the railroad signaling system because the identity of trains does not matter for the verification of the signaling system (permutation symmetry). If correctness is verified for one possible set of trains, we can permute these trains arbitrarily, and correctness follows in the state with permuted trains as well.

State Space Reduction: Projection

PROJECTION[4] is a further means to help reduce state space. Projection means: limiting the number of properties, or the number of values for properties, considered in the model, to achieve the model's purpose with a reduced state space.

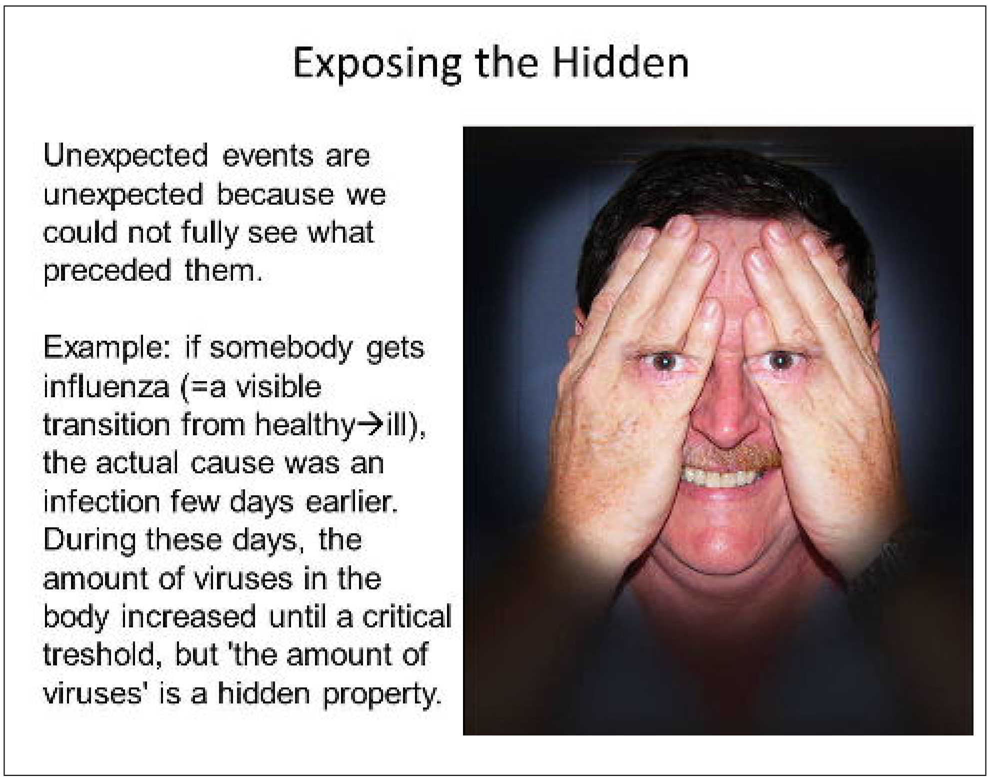
We illustrate this with an example.

In the lifecycle of a sock, there are many properties that could be taken into account. Apart from its location and its hygienic condition, we could keep track of its color (`color`), and the number of holes (`nrHoles`). The value of `color`, however takes a value that does not change over time. There are no transitions having effect on color, and therefore ignoring the property `color` from the state has no effect for the state space. For `nrHoles` this is different. Due to wear, `nrHoles` may increase, and due to repair it may decrease. Hence there are two transitions that have effect on `nrHoles`, and the total number of states of a sock is multiplied by the maximal number of holes we want to consider, according to Expression 3.1. Given the purpose of the model, it may be safe

[4] The image of the shadow is taken from http://www.rgbstock.nl/photo/mmeHudS/%3E+self-portrait

to ignore the occurrence of holes, yielding a considerably smaller state space. Also, we see that the size of the state space is determined by the number of values in {`'clean'... 'dirty'`}. If we, for instance, distinguish only 5 levels of dirtiness instead of 10, the number of states halves, and the number of transitions reduces roughly by a factor of 4 .

Exposed and Hidden Properties



This suggests the idea of EXPOSED PROPERTIES and HIDDEN PROPERTIES[5]. The exposed properties together determine the state transitions, observable from the 'outside' of the system. Changes in the values of hidden quantities go unnoticed. Hiding properties lowers the number of perceivable transitions.

Projecting can be done by leaving out quantities ▷64 , such as `nrHoles` in the sock example.

As a second example: remove the seconds-hand from a analogous clock, and the passing of seconds no longer leads to visible (exposed) transitions. The state of the clock is projected down from three quantities (hours, minutes, seconds) to the two quantities (hours,minutes). The inner (hidden) states of the clock, however, still change at least every second. This projection reduces the state space of the clock from $12 \times 60 \times 60 = 43200$ states to mere $12 \times 60 = 720$ states.

Projecting may mean, however, that the modeled system can no longer be fully understood. In the clock example: if we can't inspect the state of the second hand, any transition of the minute hand comes as a surprise.

So there is a trade-off: having many exposed properties gives a large state space; having few exposed properties gives a smaller state space, but the model may become incapable to explain all transitions.

Projecting may mean: hiding *properties*. More often, it means: hiding *values*. In the sock example, we ignore all grades of dirtiness other than `clean`. We only distinguish `clean` or *not*`clean`. Obviously, this is done with the goal of achieving a possibly small state space. So:

[5]The image of a hidden-exposed face was taken from http://www.rgbstock.nl/photo/nbtJ04e/Ik+zie+je+nog+steeds+...+2

the number of states of the modeled system is often huge, and the number transitions is therefore even much more so, but by clever choices for the hidden and exposed properties and values, the actual *purpose* of a model may be achieved with considerably fewer states.

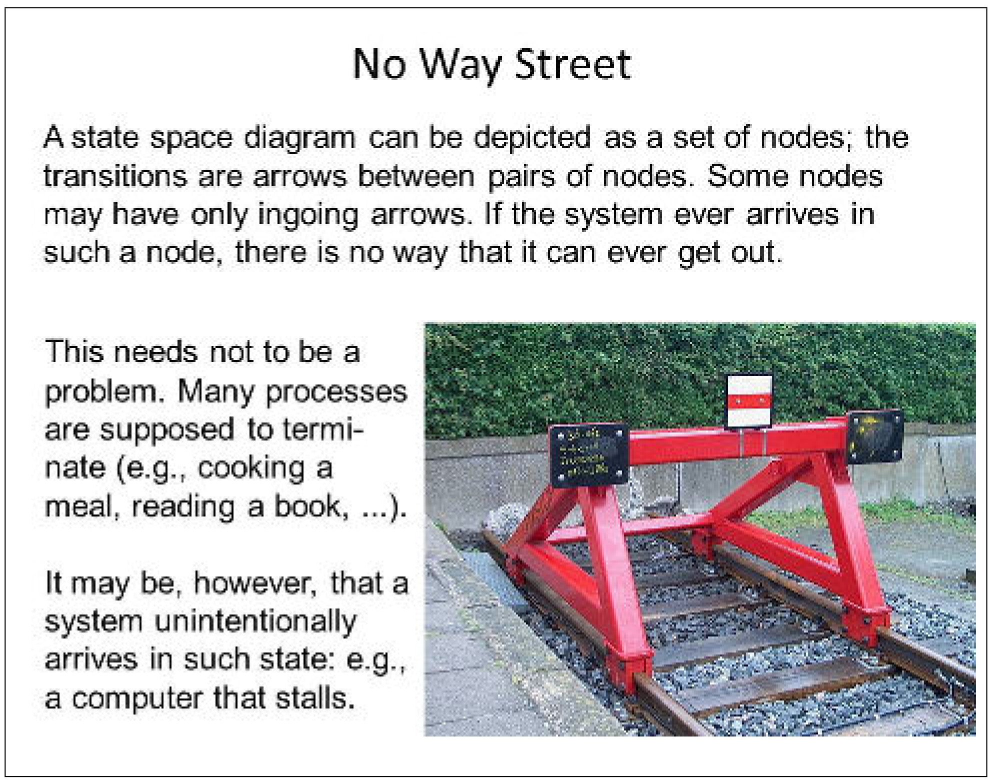
Projecting, with the purpose of the model in mind, is a powerful device to mitigate the state space explosion that would result from inadvertently adding more properties to the conceptual model.

There is no immediate right or wrong with respect to projection. The purpose of the model dictates which exposed behavior, and therefore which exposed properties we need.

Hierarchy and Orthogonality

State charts have been developed over the past decades into a powerful device for modeling dynamic systems, in particular by adding hierarchy. Hierarchy is a way to hide and expose properties, so that the internals of a dynamical system can be modeled without state space explosion. Also so called ORTHOGONAL or independent subspaces as part of present day state chart concept helps to mitigate state space explosion.

A formal treatment of these more advanced state chart concepts falls beyond the scope of these notes.



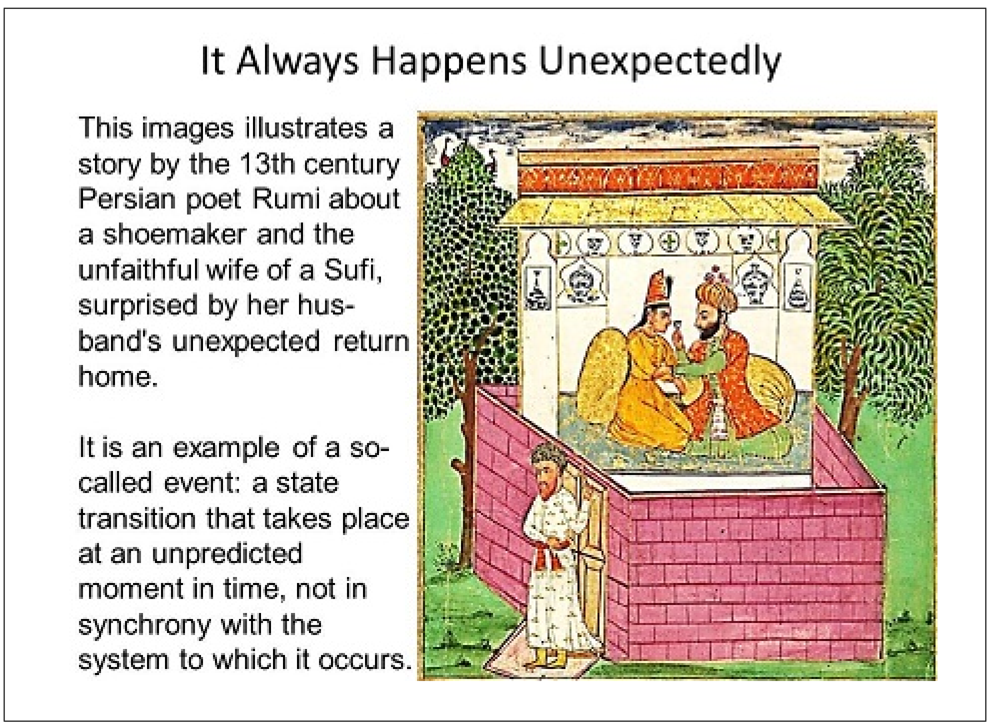
3.2.2 Applying State Charts

We mention a number of applications of state charts:

- verify if there are no DEADLOCKS[6] (=dead end states). A dead end state is a state that has no outgoing transitions. In such a state, the system has no way to make progress. An example from the railroad signalling case: two halted trains, waiting for each other's departure. A related anomaly is LIFELOCK. Then there is a small collection of permitted states, but none of these states has a transition to any state outside the collection. An example is after you-after you-blocking, occurring if two polite people both want to give way in a narrow passage;

[6]The image of the buffer stop is taken from http://commons.wikimedia.org/wiki/File:Stootblok_(staal).jpg?uselang=nl

- verify if states that should be reached can be reached. For instance, in a maintenance schedule (e.g., periodic control), it may be acceptable that, due to some high-priority exception, an occasional round of maintenance is skipped. If such interrupts occur too often, however, it could cause the complete maintenance scheme to break down. It may be necessary to verify that despite interrupts, maintenance at least takes place every once so often;
- verify if states that should be reached in some order are reached in that order. For instance, communication protocols as in computer networks (Internet, money transfer) need to be robust against network failures, out-of-order messages and perhaps against malicious attacks of the communication partner. A network failure or malicious attack is generally unpredictable, so the unwanted state transitions happen at unexpected instances, amidst the planned protocol;
- verify if occurring transitions are expected or admitted while monitoring a system;
- verify that *eventually* the behavior of a system will have certain properties. As an example, for a model for the Dutch railroad switching and signalling system: no train shall be held up *for ever* waiting for a red signal;
- argue about the synchronisation[7] of events. As follows. Things sometimes happen independent of anything else. A poor dancer may move his feet in a way that is not at all connected to the rhythm of the music. Something happening independently from the flow of events in some process P is called ASYNCHRONOUS with P. The opposite of asynchronous is SYNCHRONOUS. Synchronization means that the time order of events in one process is connected to the time order of events in another process. Example: when preparing a sandwich, the butter should be applied in between slicing the bread and putting on the topping. Applying butter is to be synchronized with the other two stages of the process.



As an example of the success of advanced use of projection: an automated parking garage in 's Hertogenbosch had to be verified for correct behavior. A straightforward state chart model of the system amounted to some 10^{80} states, clearly beyond the capability of any computer. Clever projection helped reduce the number of states, necessary for full verification, to a mere 10^6 - which

[7]The illustration from the Rumi story was taken from http://commons.wikimedia.org/wiki/File:Jalal_al-Din_Rumi,_Maulana_-_A_Shoemaker_and_the_Unfaithful_Wife_of_a_Sufi_Surprised_by_her_Husband%27s_Unexpected_Return_Home_-_Image_Detail.jpg?uselang=nl

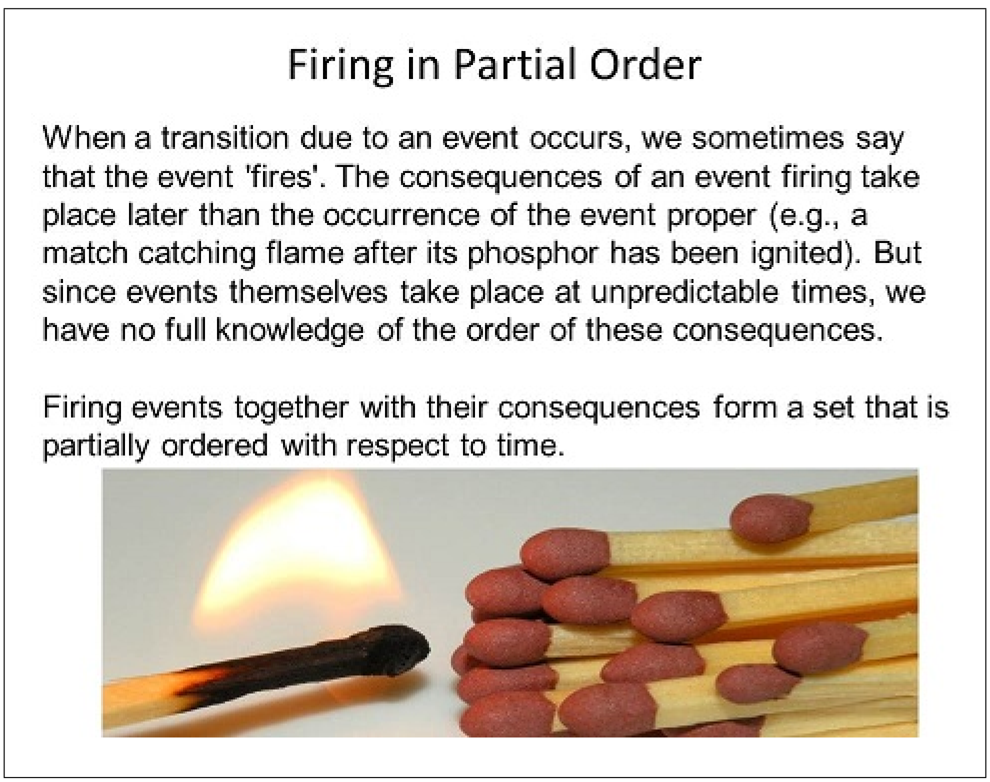
can be handled by a standard PC in a reasonable amount of time.

3.3 Time and State Transitions

We have not yet introduced the notion of time proper. In the sequel we consider three different ways of representing time in models.

3.3.1 Partially Ordered Time

Time *appears* to be totally ordered. For any two events, it appears possible to say which came first. There are exceptions, though: in the reconstruction of a crime it may be difficult to assess whether the prime suspect appeared at the crime scene before or after the fatal blow on the victim's head took place, and the distinction between these two may make the difference between imprisonment or acquittal on the ground of lacking evidence. A police inspector's report of the crime can be seen as a model with the purpose of documentation of the crime. In this model, the time order of the suspect's appearance at the crime scene and the assault could be undetermined: this is a model with *partially* ordered time.



In state charts, the time order for transitions that lead to and from some state are known[8]. Indeed, a transition leaving a state can only take place *after* the transition leading to that state has occurred. For other pairs of transitions, we don't know their order. In the sock example: we know that, with respect to the state corresponding to the sock being washed, the transition `getDirty` occurs before `getClean`, but we can't tell if the sock is put on before or after it had been stored away in the closet. In fact: the state chart describes a whole sequence of transitions, including many washes, many instances of storing away a sock, and many instances of putting on that same sock.

An arrow in a state chart is a transition that can be identified with a certain point in time. The

[8]The photograph of the flame igniting matches is taken from http://www.rgbstock.nl/download/Lajla/mxJIE2k.jpg

reason for a transition may be an external EVENT, that is: something that happens *outside* the modeled system which affects its state. The sock model does not represent the mechanism describing how and why socks get dirty, so the transition `getDirty` is due to an external event. Events can happen at any time, not necessarily synchronized with the process taking place in the dynamic system.

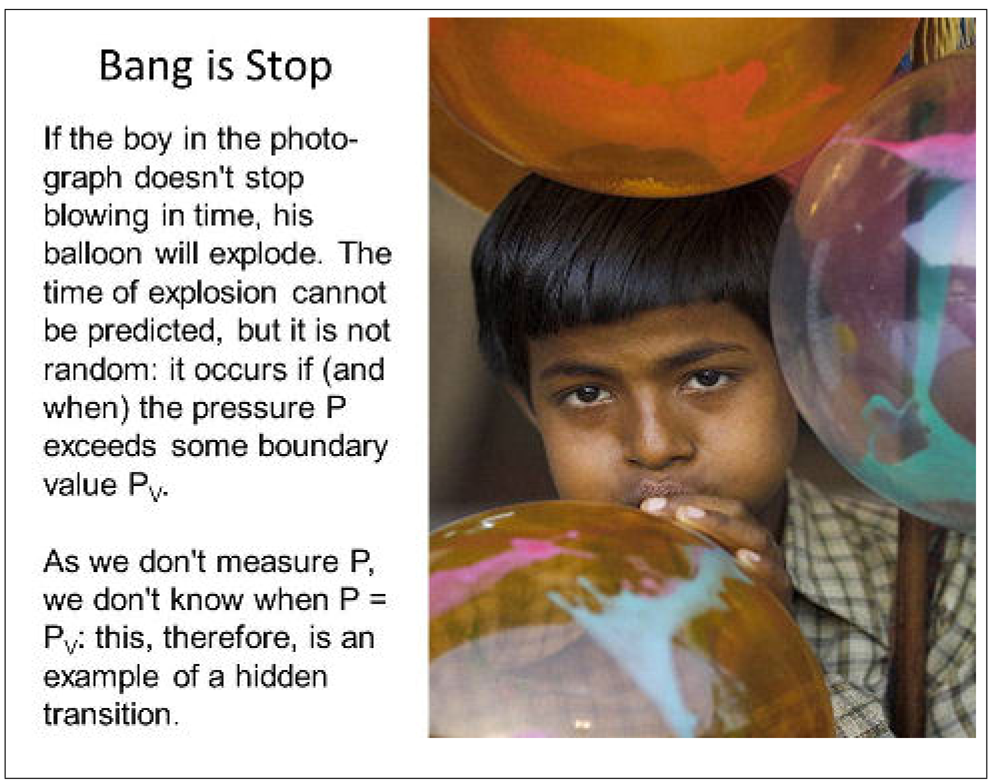
We give some examples of events:

1. *a telephone ringing:* the system consist of a process (say, cooking), and the ring is an asynchronous interrupt that may require SERVICING, such as putting the stove off before the phone is picked up;
2. *insertion of a coin into a coffee machine:* the system consists of a coffee machine, for instance busy controlling its internal water temperature. Irrespective of the current state of the coffee machine, the inserted coins have to be registered, since the display has to show the correct balance at any time;
3. *colliding billiard balls:* for the system consisting of a rolling billiard ball, a collision with something else is an external event.

Since events last infinitely short, we postulate that no two events take place at the same time. Occurring events are totally ordered in time ▷65 .

Next to events, there are INTERNAL TRANSITIONS ▷66 .

Internal transitions can occur as a result of a transition of a hidden property. For example: consider inflating a balloon[9]. Exposed properties are the amount of air we put into the balloon and its volume. Hidden properties are the stress in its skin, and the maximal stress it can endure. The moment where these two are equal, the balloon explodes. This is a transition without an external event, unlike when the explosion is caused by pinching the balloon with a needle.



Again in the three examples above, internal transitions are:

1. the cook stirs a pan of sauce, and as a result of

[9] The photo of the balloon-blowing boy is taken from http://upload.wikimedia.org/wikipedia/commons/a/a1/India_-_Varanasi_boy_balloon_-_2735.jpg?uselang=nl

that, the last lump in the sauce dissolves;

2. the water inside the coffee machine is warming up to the point where the heater is switched off;

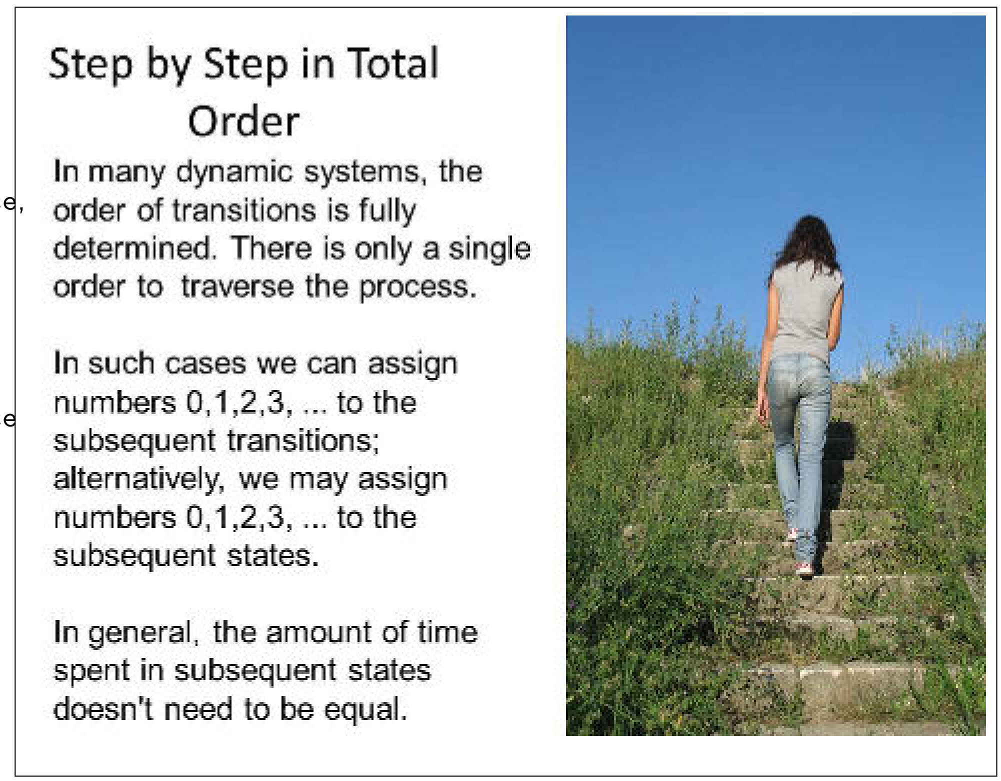
3. if locations of a rolling ball are measured with a ruler, subsequent marks on the ruler are passed in subsequent time points.

Although an actual stream of events and transitions in a process is totally ordered in time, we may only be interested in certain parts of their ordering. It may be that the model only involves events or transitions in a *partial* time order. It is also possible that the model represents *possible* transitions for the modeled system that may occur in more than one order. Then the assumption about total ordering no longer holds.

Again for the same three examples as above:

1. In a telephone communication, the events `dialA` and `answersTheCallB` have a fixed order. But after the conversation, both `A` or `B` can terminate the connection. So the order of the events `terminateA` and `terminateB` is irrelevant.

2. In a coffee machine, we identify four events: `insertCoin`, `makeChoice`, `startCoffeeMaking`, and `returnChange`. There is no fixed order between `insertCoin` and `makeChoice`, and neither between `start CoffeeMaking` and `return Change`. But both `start CoffeeMaking` and `return Change` occur after both `insertCoin` and `makeChoice` have happened. A coffee machine designer may use partial ordering to specify the coffee-buying process. For instance, if `comesBefore(start CoffeeMaking, give Coffee)` and `comesBefore (giveCoffee, returnChange)` in a situation where coffee beans have run out, the machine may not be able to give change.



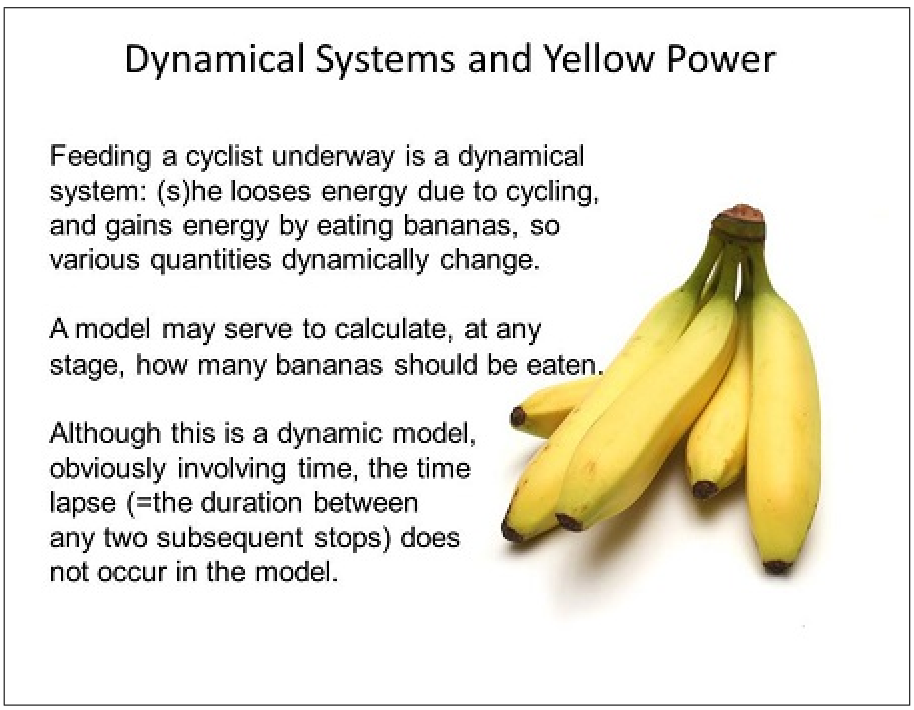
3. Depending on the variety of billiard game, the collisions between balls and cushions may or may not be relevant: in libre style billiard, there is no order constraint; in the variety known in the Netherlands as "tien over rood", a stroke is only valid if the events `hitFirstWhiteBall`, `hitRedBall`, `hitSecondWhiteBall` take place in this order.

3.3.2 Totally Ordered Time

To properly get up a flight of stairs[10], one should step on the stairs one by one in the order of increasing height. The time one spends on one step, however, could very well be different from the time one spends on another. The same is true for reading a book: the pages should be turned in the order of increasing page numbers, but reading one page could take more time than reading another one.

In this section we assume total time ordering. All occurring transitions can be uniquely labeled with increasing integers, 0,1,2,3.... For any two transitions t_i, t_j the relation `later`(t_i,t_j) holds when $i > j$.

Time Lapses



Since we define transitions to be instantaneous, two subsequent transitions can be identified with a TIME LAPSE. A time lapse, Δ, is the amount of time elapsing between two subsequent transitions. Δ is a function from two transitions to $\mathbf{R}^+$, such that $\Delta(t_i, t_j) + \Delta(t_j, t_k) = \Delta(t_i, t_k)$, where $i \le j$, $j \le k$, and $i \le k$. It follows that $\Delta(t_i, t_i) = 0$ for all transitions t_i. Indeed: transitions take no time.

Between t_i and t_{i+1}, nothing happens. None of the exposed quantities in the model undergoes a transition.

Transitions are labeled with increasing integers; states or state properties are labeled in the same way. For a state property Q, Q_i denotes its value after transition t_i. So we can also refer to state nr. i or 'state i' for short. This is the state during which Q assumes the value Q_i.

The time lapse of state i is $\Delta(t_i, t_{i+1})$, or Δ_i for short.

Total Ordering, Causality

In a model with totally ordered time, we don't require that the absolute length of a time lapse has a meaning. We give an example. We want a model to help calculate the number n_i of bananas[11], each containing B kCal, we should eat at each stop i in bicycle trip, where we have been given the

[10] The image of a lady climbing stairs is taken from http://www.rgbstock.nl/photo/mhAVAC6/de+trap+3

[11] The banana image is taken from http://www.rgbstock.nl/photo/mWjX0mK/Bananaaaaas%21

amount of effort E_i, between any two subsequent stops, as a list $\{E_i\}$ with values. We define Q_i being the amount of kCal we have in our body at stop i. When we start the tour we have Q_0 in our body. If we don't eat any bananas, the amount of kCal at stop $i+1$ is given by $Q_{i+1} = Q_i - E_i$. If we do eat n_i bananas, however, we get $Q_{i+1} = Q_i - E_i + n_i B$. We want to choose n_i the minimal value such that $Q_{i+1} \geq 0$. Indeed, we don't want to eat too many bananas, to avoid constipation; further, we eat an integer amount of bananas. So $Q_i - E_i + n_i B \geq 0$, and $n_i B \geq E_i - Q_i$. It follows that $n_i = \max(0, \lceil (E_i - Q_i)/B \rceil)$. So $Q_{i+1} = Q_i - E_i + B \times \max(\lceil (E_i - Q_i)/B \rceil)$, where $\lceil ... \rceil$ means: taking the rounded-up value. So, e.g., $\lceil 4.2 \rceil = 5$. Notice that the lengths of time lapses, Δ_i, doesn't enter into the calculation of Q_i.

Causality and Functions

In the banana example, we find $Q_{i+1} = Q_i - E_i + B \times \max(\lceil (E_i - Q_i)/B \rceil)$, or $Q_{i+1} = F_Q(Q_i, E_i, B)$. So: there exists a FUNCTION F_Q, so that we can write a quantity Q in subsequent states as a function of earlier values of Q, and, optionally, earlier values of other quantities (such as E_i), and/or constants (such as B).

The function F_Q represents our intuition of 'causal dependency'. It expresses the causal dependency of Q_{current} on Q_{previous} and optionally P_{previous}:

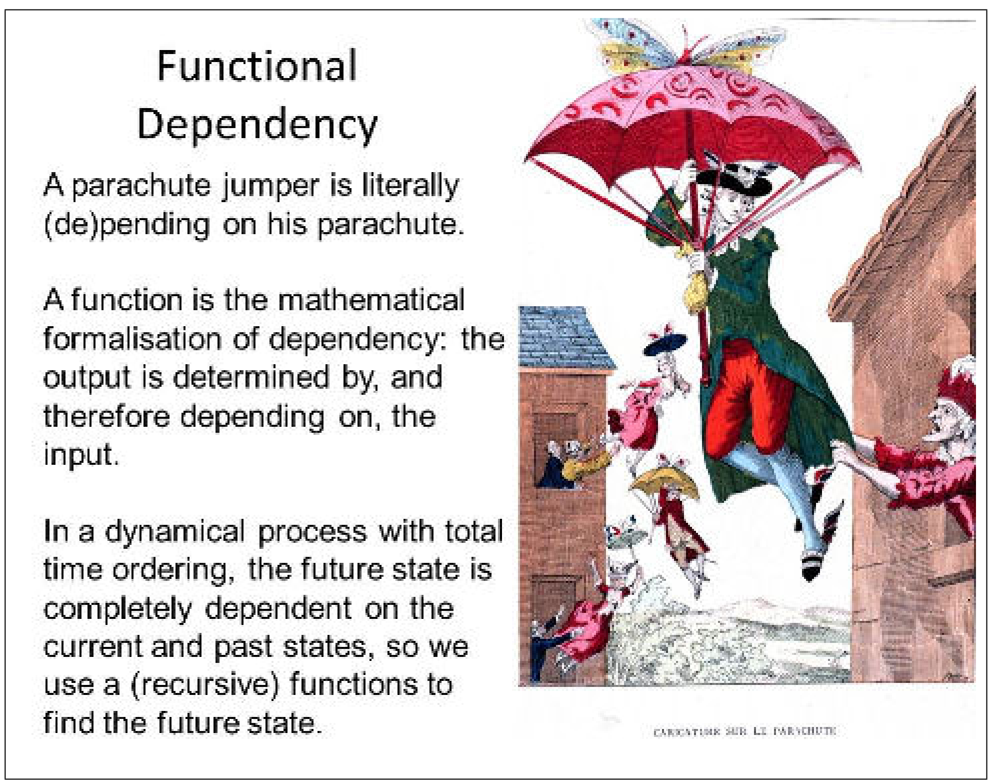
$$Q_{\text{current}} = F_Q(Q_{\text{previous}}, P_{\text{previous}}), \tag{3.2}$$

where P_{previous} holds the values of optional other quantities at a previous time step.

In the banana example these are the E_i and B; in a more elaborate model, Q could also depend[12] on perspiration, water intake, temperature and further quantities.

F_Q is a function, that is: given Q_{previous} and all P_{previous}, the value for Q_{current} follows from evaluating F_Q.

For F_Q to be a function, Q_{current} may only depend on state quantities of *earlier* states. F_Q being a function means that there must be causal relations leading



[12]The caricature on the invention of the parachute is taken from http://commons.wikimedia.org/wiki/Parachute#mediaviewer/File:Parachute_caricature.jpg

from earlier states j, with $j < i$, to state i ▷67 .

We may consider to allow:

$$Q_{\text{current}} = F_Q(Q_{\text{previous}}, P_{\text{current}}), \tag{3.3}$$

since this seems not to violate causality per se. But suppose that we also have

$$P_{\text{current}} = F_P(P_{\text{previous}}, Q_{\text{current}}), \tag{3.4}$$

then F_Q and F_P *together* violate causality. Hence the more strict requirement that *all* state quantities occurring as arguments in one of the functions F need to be taken at an earlier point in time.

Suppose that we have no hidden quantities, and that for every quantity Q,

$$Q_{\text{current}} = F_Q(P_{\text{previous}}), \tag{3.5}$$

is the only function in a process model. That is: Q does not depend on earlier values of itself. Then everything is static: the P don't depend on Q, and also not on previous values of themselves, so they must stay constant. Q depends functionally on the P, but with unvarying P, the Q also stays the same. We see that change is only possible if there is at least one quantity that depends on an earlier value of *itself*.

Evolution of Dynamical Systems

We summarize: if the purpose allows time to be totally ordered, and if need to evaluate quantities at subsequent time points, then we could use functions F_Q such that

$$Q_i = F_Q(Q_{i-1}, P_{i-1}). \tag{3.6}$$

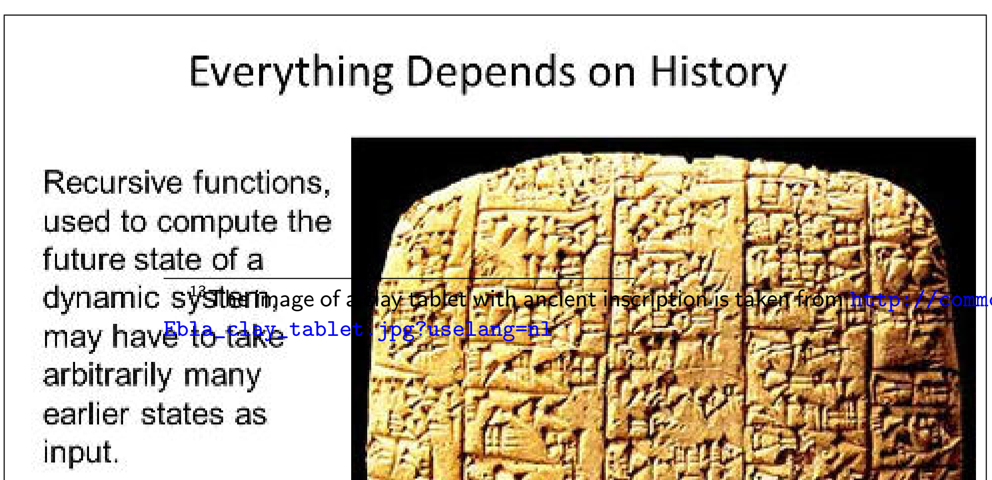
If Q_i depends on earlier states,

$$Q_i = F_Q(Q_{i-1}, P_{i-1}, Q_{i-2}, P_{i-2}, \cdots), \tag{3.7}$$

with as many i's as necessary.

Optional quantities P represent other quantities in the model.

If the oldest state[13] that occurs among the arguments of F_Q is $i - n$, we cannot use Equation 3.7 at the first n states. We need additional information to find the values of the quantities in these first n



[13] The image of a clay tablet with ancient inscription is taken from http://commons.wikimedia.org/wiki/File:Ebla_clay_tablet.jpg?uselang=nl

states. These are called INITIAL VALUES. The maximally occurring value of n is called the ORDER of the model. For example, for a financial transaction system where a new amount of money on an account is calculated from the current amount and a transferred sum, the order is 1. For the banana example, the order was also 1; in a little while we encounter systems that are order 2.

There are systems that require large values of n: for instance, a protocol where a product undergoes many different transitions, and a next transition may depend on each of these. A model for a system that has memory (such as a computer, an hourglass, or a living organism) may require large n. Most mechanical and otherwise physical systems, can be modeled with n no larger than 2. We say that systems with low order are (nearly) memory-free.

Instead of Equation 3.7 where i runs from 2 upwards, we could also write

$$Q_{i+1} = F_Q(Q_i, P_i, Q_{i-1}, P_{i-1}, \cdots), \tag{3.8}$$

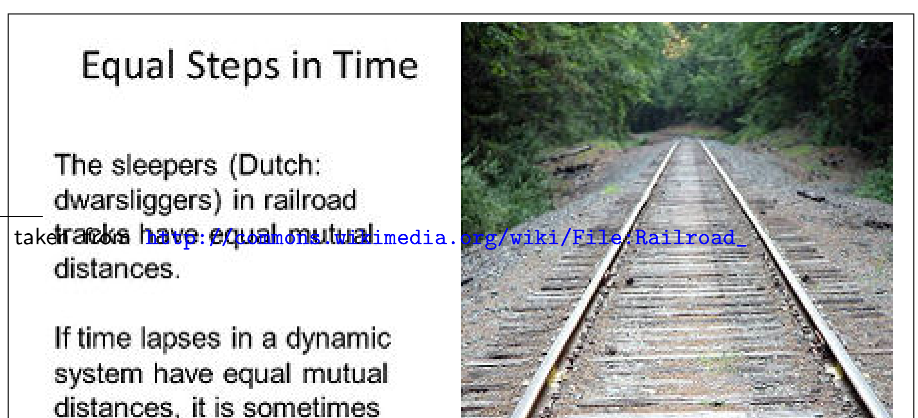
with i starting at 1, or even

$$Q_{i+2} = F_Q(Q_{i+1}, P_{i+1}, Q_i, P_i, \cdots), \tag{3.9}$$

with i starting at 0. Indeed, i is merely a dummy quantity that can be offset by any convenient integer number. Equation 3.8 and Equation 3.9 have the same order if the difference between the index on the left hand side (in this case $i+1$ or $i+2$) and the lowest occurring value in the right hand side is the same. The *order* of a dynamical model is the maximal number of earlier transitions we need to take into account in order to compute a transition.

3.3.3 Totally Ordered Time; Equal Intervals

Results thus far only required totally ordered time. Time lapses between subsequent transitions could have different lengths. Next we look to the special case where time lapses are equally long[14].



[14] The image of a railroad track was taken from http://commons.wikimedia.org/wiki/File:Railroad_Tracks_In_Woods.jpg?uselang=nl

Equal Intervals: Periodicity

Processes often assume PERIODIC time. Consider a pendulum to measure time; it works because of its periodic behavior with clearly marked transitions, *viz.*, the two extreme positions where the pendulum reverses its direction. Our calender is based on the annual seasons, driven by the earth's orbit round the sun. Also the financial world is influenced by this rhythm, as interest on a deposit is made payable every 1st of January.

Let us study the last example in some detail.

First, we ignore the periodicity. We take an example from *budgeting*: that is, planning expenses such that in the long run a sustainable financial situation occurs. There is an initial amount of money, A_0, and after an amount of time Δ_0 a fraction $s_0\Delta_0$ is spent. We expect to spend more in a longer period of time, hence the proportionality to Δ_0. We assume $0 < s_0\Delta_0 < 1$. Also we gain an amount $g_0\Delta_0$, again proportional with Δ_0. So after Δ_0, we have $A_1 = A_0(1 - s_0\Delta_0) + g_0\Delta_0$. This holds for every transition i:

$$A_{i+1} = A_i(1 - s_i\Delta_i) + g_i\Delta_i, \quad (3.10)$$

$$t_{i+1} = t_i + \Delta_i. \quad (3.11)$$

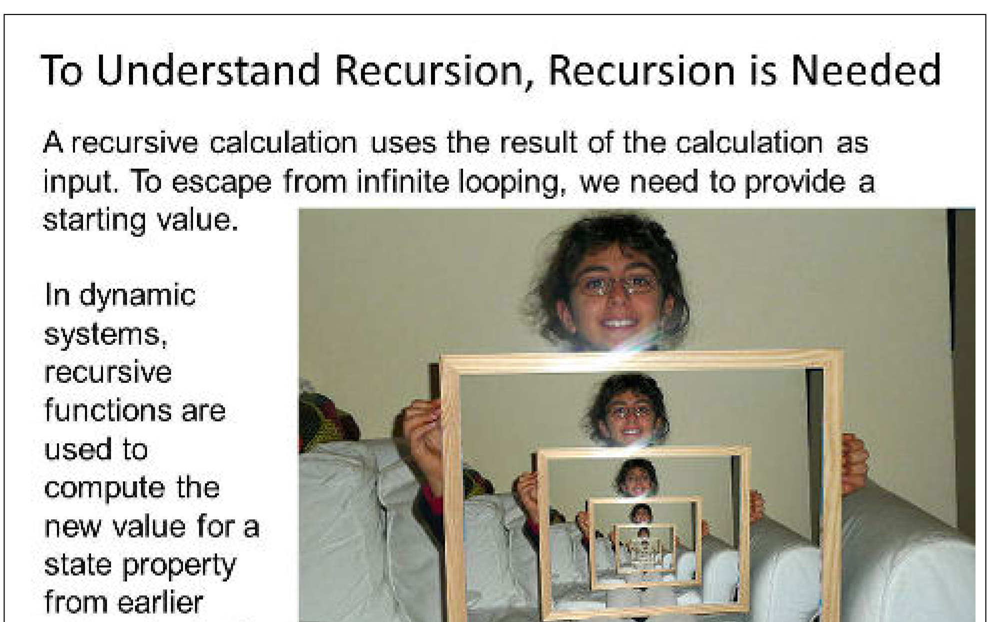
These are both in the form of Expression 3.2:

$$A_{i+1} = F_A(A_i, s_i, g_i, \Delta_i), \quad (3.12)$$

$$t_{i+1} = F_t(t_i, \Delta_i). \quad (3.13)$$

To know the amount left at time t, we seek the value for i_t with the largest t_{i_t} with $t_{i_t} \leq t$, since this was the transition where A reached its current value A_{i_t}. Next, just as we did with the banana example, evaluate Expression 3.12 for subsequently $i = 0$, $i = 1$, $i = 2$, $\cdots$, $i = i_t$. Eventually we find A_{i_t}.

This works for predicting the amount left, assuming



a series of expenses s_i, incomes g_i, and time intervals Δ_i. It doesn't work well for *deciding*, however. Suppose we want to decide, for a constant $g_i = g$, what expense rate $s_i = s$ we can afford for a positive result at time t. Then we can only re-calculate the entire sequence of A_i for every i until i_t. If we don't have any further information of s_i and g_i, there is no closed form expression for A_{i_t} when A_{i+1} is defined RECURSIVELY[15] as $A_i(1 - s_i\Delta_i) + g_i\Delta_i$.

Recursion

Recursion means: defining something in terms of earlier versions of itself. Our function F_Q in Expression 3.2 is a *recursive* manner to express the values of quantities in a dynamic system.

There is a brute force way to deal with recursion. That is to 'unroll' the sequence of applications of $Q_{i+1} = F_Q(Q_i, \cdots)$ for all subsequent i beginning with $i = 0$.

This resembles the way we calculate N factorial (written as $N!$), defined as

$$\begin{aligned} N! &= (N-1)! \times N \quad \text{if} \quad N > 0; \\ 0! &= 1. \end{aligned} \tag{3.14}$$

To find 5! we first need 4! and next multiply it with 5. To know 4! we must first know 3! etc., so we first calculate 0! which is 1; next 1!=1×0!, then 2!=2×1!, next 3!=3×2!, next 4!=4×3! and finally 5!=5×4!.

Evaluating subsequent states of a dynamic model by unrolling their recursive definition of Expression 3.2 is called SIMULATION. To simulate a dynamical system, we only need the F_Q's from Expression 3.2. Time intervals don't need to be equally long. With all Δ_i in Expression 3.12 different, however, we cannot do anything else but unroll the recursion. If Δ_i, s_i and g_i are constant, however, say Δ, s, g, we can compute the closed form result ▷68

[15] The photograph of the lady, holding that photograph is taken from http://commons.wikimedia.org/wiki/Category:Droste_effect#mediaviewer/File:Droste_1260359-nevit,_corrected.jpg

which turns out to be $A_i = (1-s\Delta)^i + \frac{(1-s\Delta)^{i-1}-1}{-s\Delta} g\Delta$, which is a useful result as it tells us exactly how much we possess at any point in the future. Moreover, we can study what happens if i goes to infinity[16]. Then $(1-s\Delta)^i$ vanishes, and so does $(1-s\Delta)^{i-1}$. What remains is

$$A_{\text{infinity}} = g/s. \tag{3.15}$$

This is dimensionally correct: g is an amount of money per unit of time, and s has the dimension of 1 over time. So g/s is indeed an amount of money. The value g/s means: if one spends, say, no more than 20% of one's assets per time period, in the long run ▷69 one will own at least 5 times g over that same period, *irrespective of* Δ.

This is an example of analyzing the ASYMPTOTIC BEHAVIOR, that is: the behavior of the system in the long run, thanks to the closed form expression $A_{\text{infinity}} = g/s$. It helps to fulfill modeling purposes such as deciding on acceptable spending rates.

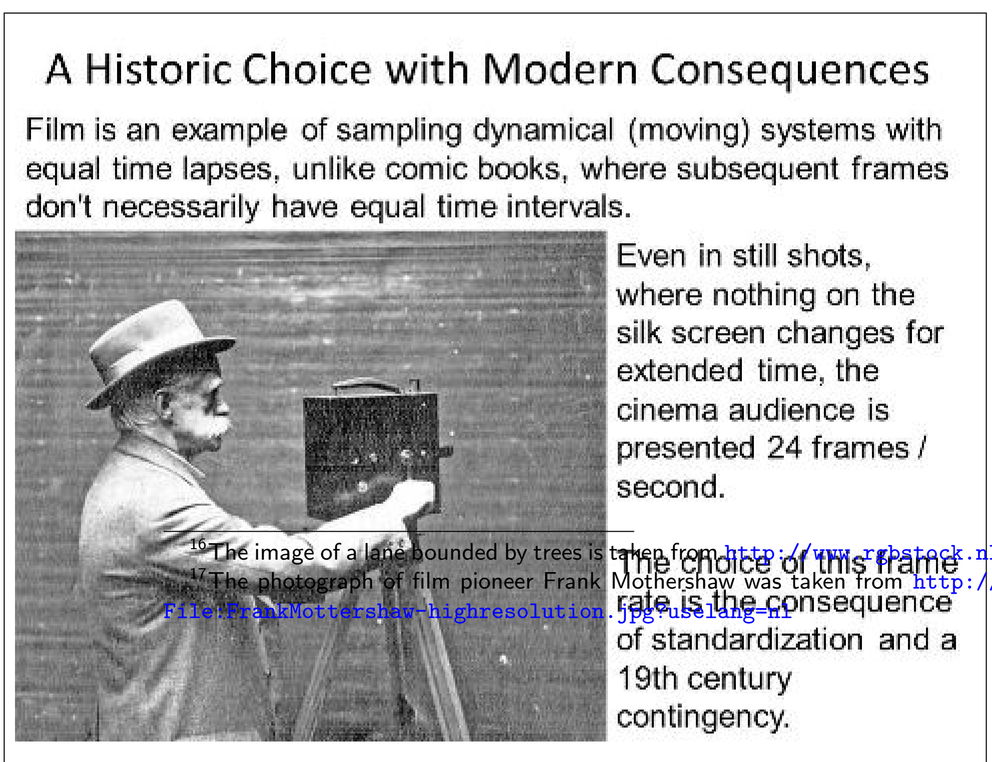
The result of Expression 3.15 shows that the actual length of the time lapse Δ is irrelevant. Whether we pay on daily, weekly or monthly base has no effect on the final situation.

This shows an interesting interpretation of our model. We can regard A as some time-dependent value, whose behavior is represented by a series A_i. We can, at any time t_t, ask what A is, just by looking up the most recent transition i_t and realizing that A at t_t equals A_{i_t}.

Time Lapses with Equal Length; Sampling

The values A_i, resulting from a recursive definition such as above, SAMPLE A as a function of time.

Sampling means: capturing the characteristics of a large set by looking at a small number of elements of this set. For instance, estimating the quality of a batch of oranges by testing the quality of a randomly selected handful (stochastic sampling), or recording an audio signal by means of 44100 distinct values taken per second (periodic sampling, used for recording audio on CD[17]). Unlike in



[16] The image of a lane bounded by trees is taken from http://www.rgbstock.nl/photo/meSmz1c/Oneindigheid

[17] The photograph of film pioneer Frank Mothershaw was taken from http://commons.wikimedia.org/wiki/File:FrankMottershaw-highresolution.jpg?uselang=nl

stochastic sampling, or audio sampling: when we sample in the form of Expression 3.12, the samples A_i are sufficient to know everything about the value of A at any desired point in time. By definition, A does not change in between subsequent transitions.

Developing a model with totally ordered time with constant Δ means *sampling* the modeled system. Unlike the banana example, the length of the sampling intervals Δ is usually relevant for the model outcome. Let us study the example of throwing a ball.

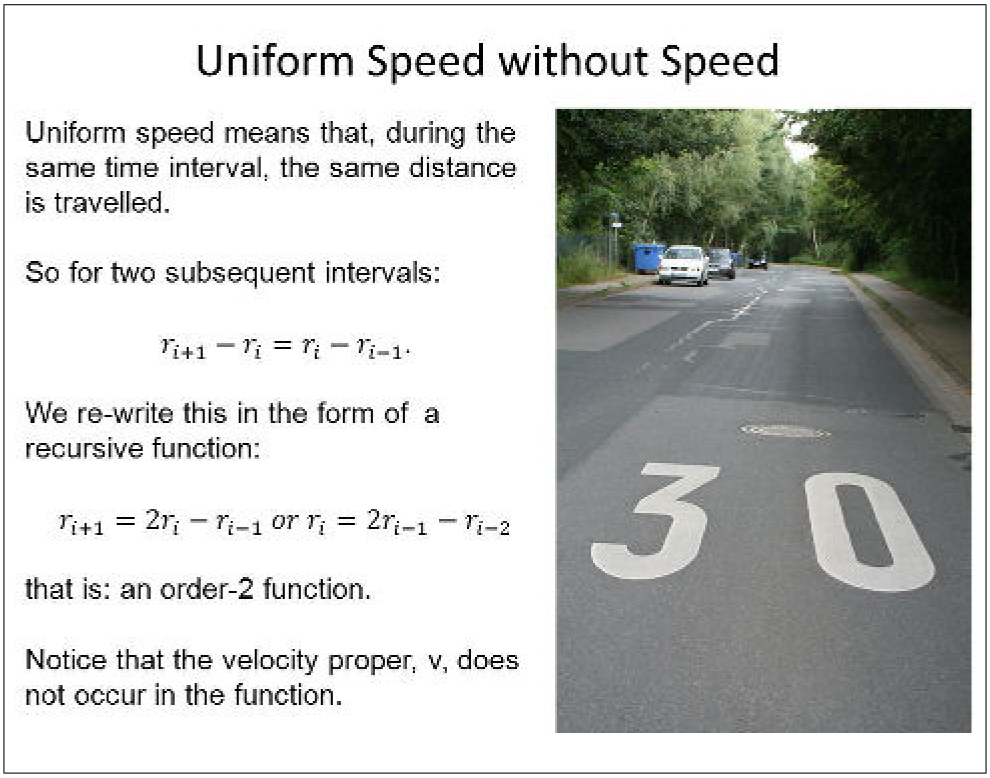
A ball is considered as a point mass: rotation is ignored. At time $t_i = \Delta i$ it has location r_i and velocity v_i. First assume no forces. Then velocity is constant: $v_i = v$. We propose a recursive process, with uniform velocity implying that equal lengths of distance are covered in equal time lapses:

$$r_{i+1} = F_r(r_i) = r_i + v\Delta. \tag{3.16}$$

Comparing the sampled locations r_i with the locations following from the secondary school physics formula[18], $r(t) = r_0 + vt$, we see full agreement. With $r_i = r(i\Delta)$ we get

$$\begin{aligned} r_{i+1} &= r_i + v\Delta \\ &= r(i\Delta) + v\Delta \\ &= r_0 + vi\Delta + v\Delta \\ &= r_0 + v(i+1)\Delta \\ &= r((i+1)\Delta). \end{aligned} \tag{3.17}$$

[18] The image of the road with speed limit is taken from http://www.rgbstock.nl/photo/mC2ICWS/30



So irrespective of Δ, sampled locations r_i agrees with true locations $r(i\Delta)$.

There is another recursive definition with the same result:

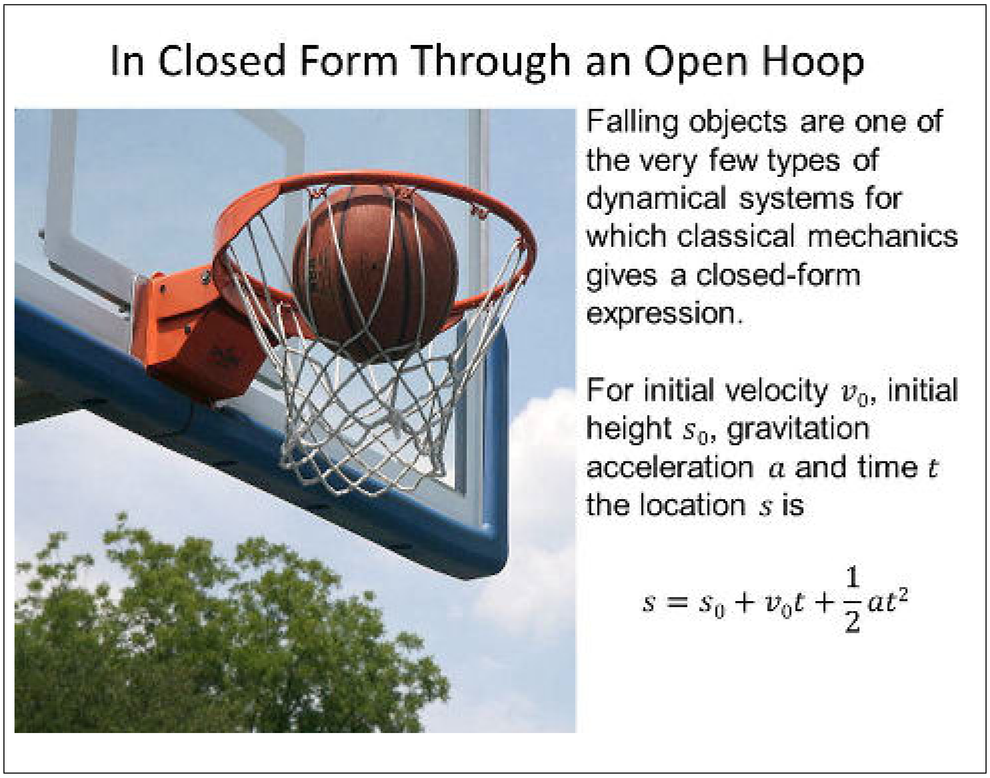
$$
\begin{aligned}
r_{i+1} &= F_r(r_i, r_i - 1) \\
&= 2r_i - r_{i-1}.
\end{aligned} \tag{3.18}
$$

This may be understood better if we re-write it: $r_{i+1} = r_i + r_i - r_{i-1}$, so

$$
r_{i+1} - r_i = r_i - r_{i-1}. \tag{3.19}
$$

This means: the displacement between i and $i + 1$ is the same as between $i - 1$ and i, which is an alternative way of defining uniform speed.

So we have both an order-1 version in Expression (3.16) and an order-2 version in Expression (3.18) of a recurrent model for uniform velocity. Expression 3.18 is interesting in that it does not contain the velocity: it holds for *any* (constant) velocity.



Next we apply a constant acceleration a, corresponding to a force ma where m is the mass[19]. In the absence of force, a material object moves with uniform velocity, as in Expression 3.18. In the presence of force, speed changes over time. The change of speed over a time interval is larger if the force is larger. The change of speed over time is *proportional* to the force and proportional to the time during which this force works.

From secondary school physics, the result should be:

$$r(t) = r_0 + v_0 t + \frac{1}{2}a^2 \tag{3.20}$$

By analogy with Expression 3.17, the sampled version could be tried as

$$\begin{aligned} r_{i+1} &= F_r(r_i, v_i) \\ &= r_i + v_i\Delta; \qquad (3.21) \\ v_{i+1} &= F_v(v_i) \\ &= v_i + a\Delta. \qquad (3.22) \end{aligned}$$

The solution for v_i from Expression 3.22 can be written down immediately, as this has the same form as Expression 3.17:

$$v_i = v_0 + ia\Delta, \tag{3.23}$$

because $v(t) = v_0 + at$. So we can take Expression 3.21 and Expression 3.22 together:

$$r_{i+1} = r_i + (v_0 + ia\Delta)\Delta. \tag{3.24}$$

[19]The image of the falling ball is taken from http://commons.wikimedia.org/wiki/Category:Falling#mediaviewer/File:2011-06-07_Basketball_in_hoop_still_shot.jpg

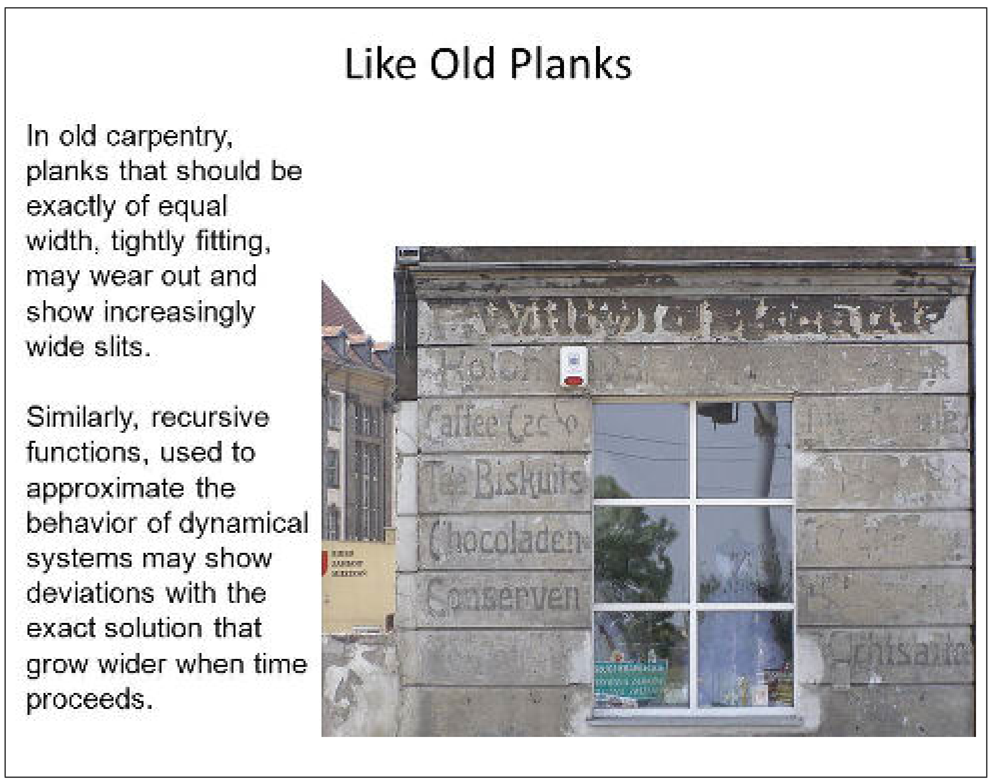
This is a so-called arithmetic series, of the form $r_{i+1} = r_i + x + yi$. It has a closed form solution:

$$r_i = r_0 + v_0 \Delta i + \frac{1}{2} a (\Delta i)^2 - \frac{1}{2} a \Delta^2 i, \tag{3.25}$$

which can be verified by subtracting the expressions for r_{i+1} and r_i. With $t_i = \Delta i$, we find

$$r_i = r_0 + v_0 t + \frac{1}{2} a t^2 - \frac{1}{2} a \Delta t. \tag{3.26}$$

The first three terms in the right hand part agree with the secondary school result (Expression 3.20). The last term, however, is wrong. Our simulation yields an error term $\frac{1}{2} a \Delta t$. This error increases proportionally with time. Moreover, the error is larger when Δ is larger.



This is a very common behavior for unrolling recursive functions that express model quantities in terms of their previous values[20]. For an accurate simulation, Δ should be sufficiently small, and the approximation deviates more when time proceeds. In Expression 3.26, halving Δ halves the error for any t. We should realize, however, that halving Δ also obliges us to do twice as many caluclations in unrolling Expression 3.21 and Expression 3.22 to arrive at r for a given time point t. There is a trade off between accuracy and time needed for the calculations. This is very common in simulations of all kinds.

Halving Δ in order to halve the error makes unrolling these recursive functions a so called 1ST ORDER APPROXIMATION. There are better schemes where halving the time step gives a reduction of 4 (so called 2nd order approximation) or even 8 (third order approximation).

For a pointmass, subject to a constant force, we can easily fix the flawed recursive definition from Expression 3.21. There is a recursion relation that gives the exact motion of a uniformly

[20]The image of the old Gdansk candy shop is taken from http://commons.wikimedia.org/wiki/File:Gdansk_Jana_z_Kolna_napis.jpg?uselang=nl

accelerated point mass. We set

$$\begin{aligned} r_{i+1} &= F_r(r_i, v_i) \\ &= r_i + v_i\Delta + \frac{1}{2}a\Delta^2; \\ v_{i+1} &= F_v(v_i) \\ &= v_i + a\Delta. \end{aligned} \tag{3.27}$$

We verify that the correct expression is reproduced. In hindsight, this is obvious. The contribution $\frac{1}{2}a\Delta^2$ is the displacement due to acceleration a during time lapse Δ.

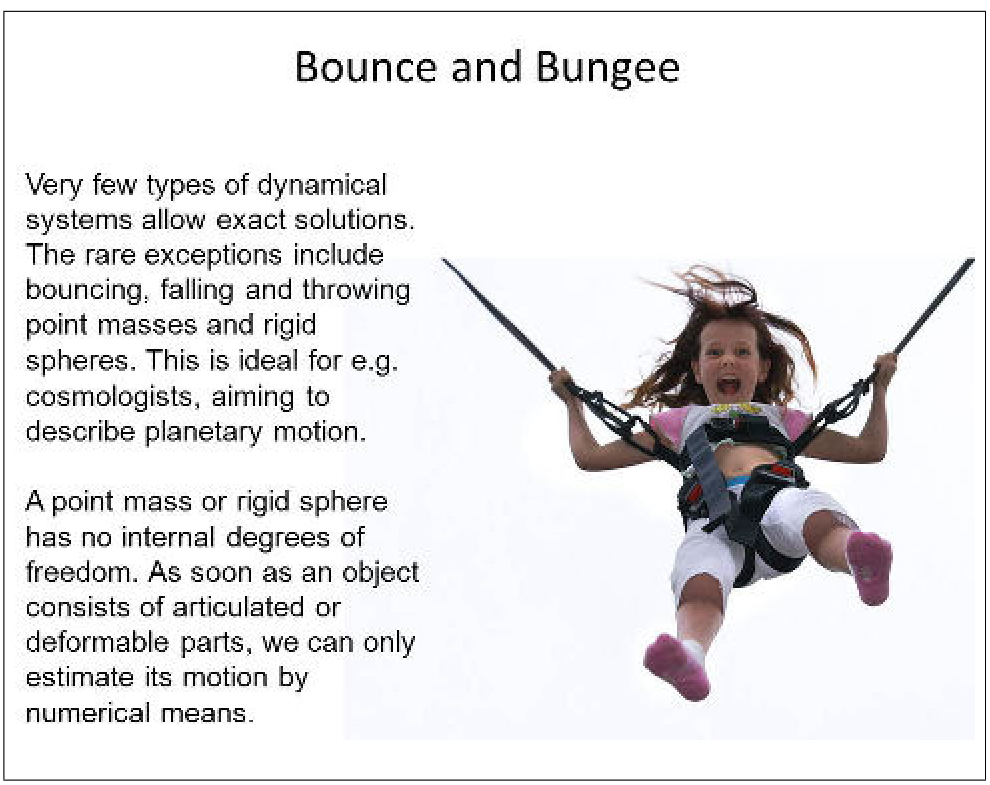
We can also achieve a correct result using our order-2 formula, Expression 3.18, as follows:

$$r_{i+1} = 2r_i - r_{i-1} + a\Delta^2. \tag{3.28}$$

Apparently, the order-2 recursive model of Expression 3.28 describes a point masses subject to acceleration a. For a=0 we get the correct result for a uniform motion. For a=constant, we get the correct motion of a uniformly accelerated point mass.

We postulate ▷70 that, in a recursive model, $r_{i+1} = 2r_i - r_{i-1} + a_i\Delta^2$, a_i is the acceleration at time step i. So whenever we should simulate the motion of a point mass, subject to a force, constant or varying, we try the recursive function Expression 3.28.

3.3.4 Totally Ordered Time; Equal, Infinitesimal Intervals: Continuous Time



The behavior for physical, financial, and many other types of systems can be modeled with recursively defined models[21]. Running such models amounts to repeated evaluation of the recursive functions to unroll the process. Sometimes this gives an exact result. This is possible if the modeled system only has transitions at discrete moments, such as financial systems. Also for physical systems recursive models can give an exact result, that is: a result that is independent of the choice of Δ and matches the results from physics.

We have also seen cases, however, where the result

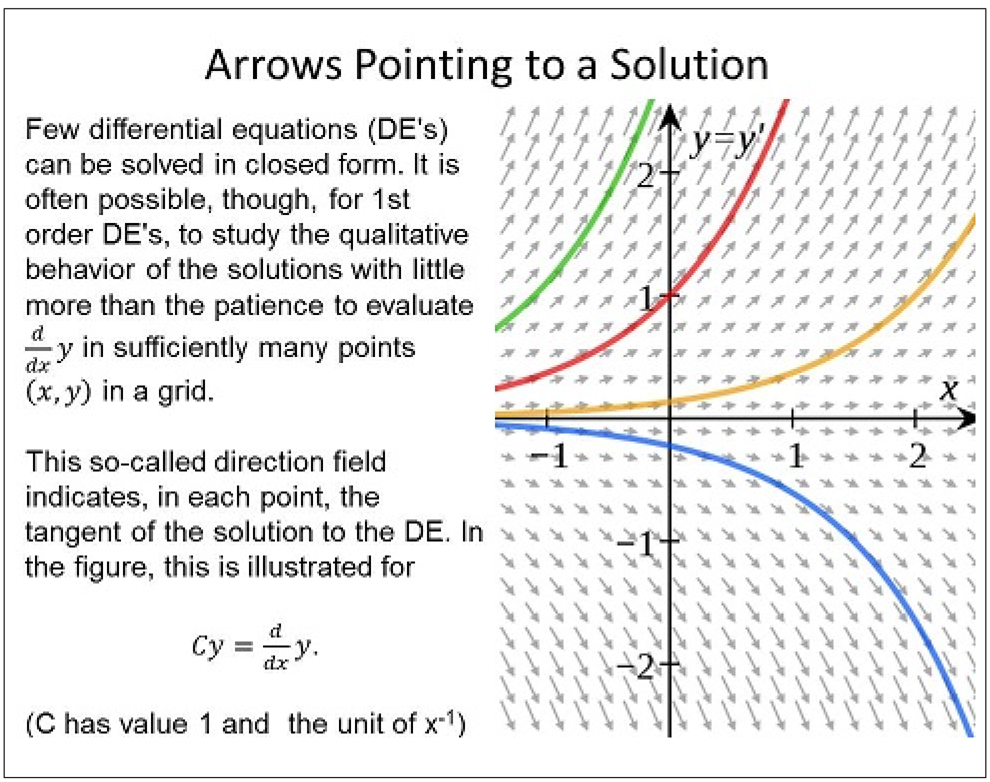
[21] The photograph of the bungee girl is taken from http://www.rgbstock.nl/download/BlueGum/mflfx8M.jpg

does depend on Δ, whereas the modeled system does not contain discrete transitions. In these cases, such as the rotating dumbbell and the mass-spring system, detailed in the end notes, we saw that the result may converge if we take $\Delta \to 0$.

This suggests that unrolling the recursive function is the numerical solution of a DIFFERENTIAL EQUATION. It turns out that numerical schemes using recursive functions can be devised ▷71 to help solving differential equations, also for more advanced dynamical systems.

Evolution of Dynamical Systems and Differential Equations

Some differential equations can be solved in closed form. That means that the solution has the form of a function rather than a list of numbers. We have seen an example of a closed-form solution: in the budget example, we got an expression for the asymptotic value for A in case of a constant income and a constant spending fraction (Expression 3.15). For modeling purposes such as *analysis*, a closed form is preferable.



Dynamical models in the form of differential equations may help us to learn something about the solution without having to unroll a simulation[22]. Similar as with the budget example, however, a warning is in place. Adding a non-linear term or otherwise changing a solvable differential equation, perhaps only marginally, may make it unsolvable with analytic techniques ▷72 . This is in contrast with unrolling the recursive version. As long as the recursive function can be evaluated, it doesn't matter what form it has. The risk with numerical evaluation, on the other hand, is the threat of large unaccuracy or even instability if Δ is too large.

We look at the connection between recursive functions and differential equations. Let us first consider a recursive function, where r_i is the unknown quantity, and p_i are given quantities.

$$r_i = F_r(r_{i-1}, p_{i-1}). \tag{3.29}$$

[22]The image of a direction field was taken from http://commons.wikimedia.org/wiki/Category:Differential_equations#mediaviewer/File:DGL_y-eq-dy.svg

With $t = \Delta i$, this is equivalent to $r(t+\Delta) = F_r(r(t), p(t))$. In the limit for $\Delta \to 0$, we can expand ▷73

$$r(t+\Delta) = r(t) + \Delta \frac{d}{dt} r(t) + O(\Delta^2) = F_r(r(t), p(t)). \tag{3.30}$$

So

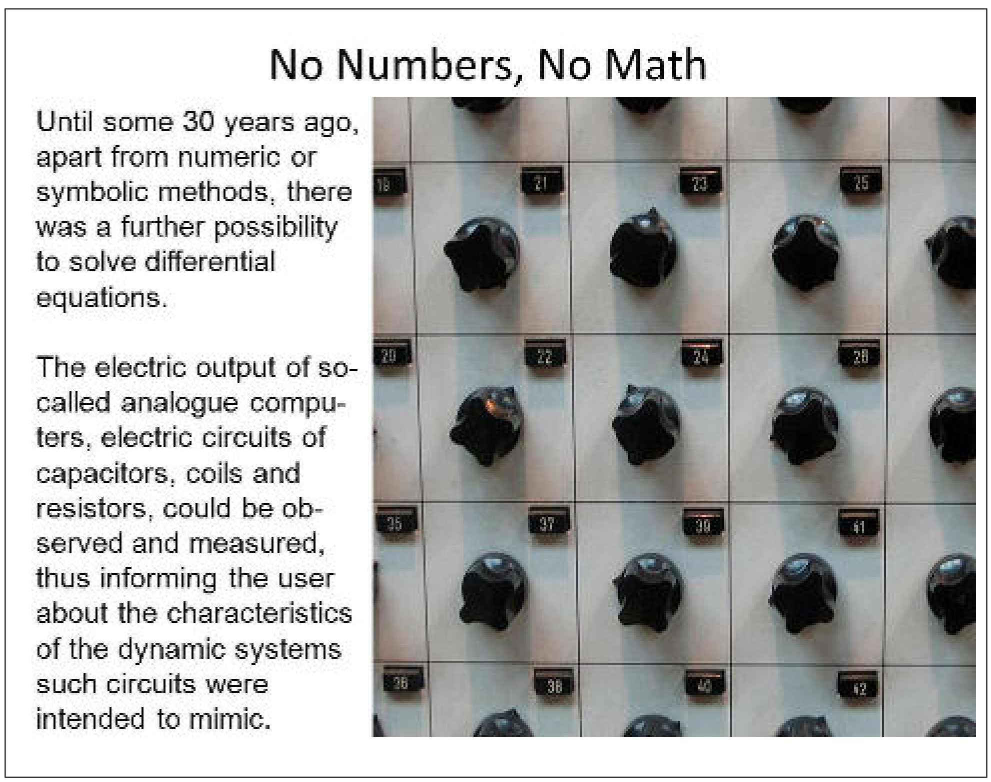
$$\frac{d}{dt} r(t) = \lim_{\Delta \to 0} \frac{F_r(r(t), p(t)) - r(t)}{\Delta}, \tag{3.31}$$

showing the correspondence between the derivative $\frac{d}{dt} r(t)$ and the recursive function F_r. The differential equations corresponding to Expression 3.10 and Expression 3.16 can be written as

$$\begin{aligned}
A_{i+1} &= A_i(1 - s_i\Delta_i) + g_i\Delta_i \\
&= A_i - s_i\Delta_i A_i + g_i\Delta_i \qquad \text{so} \\
A_{i+1} - A_i &= g_i\Delta_i - s_i\Delta_i A_i \qquad \text{or} \\
\frac{A_{i+1} - A_i}{\Delta_i} &= g_i - s_i A_i; \qquad \text{with } \Delta_i \to 0: \\
\frac{d}{dt} A(t) &= g(t) - s(t)A(t),
\end{aligned} \tag{3.32}$$

and $r_{i+1} = r_i + v\Delta$ or, equivalently $\frac{d}{dt} r(t) = v(t)$.

To model dynamic systems, we have various routes at our disposal ▷74 . We can try to describe their behavior by a differential equation and attempt to solve this by analytic means, or we can try to form a recursive function and unroll the recursion to simulate the system. A third, perhaps even more common approach is, to derive the differential equation and find an approximate solution by numeric means.

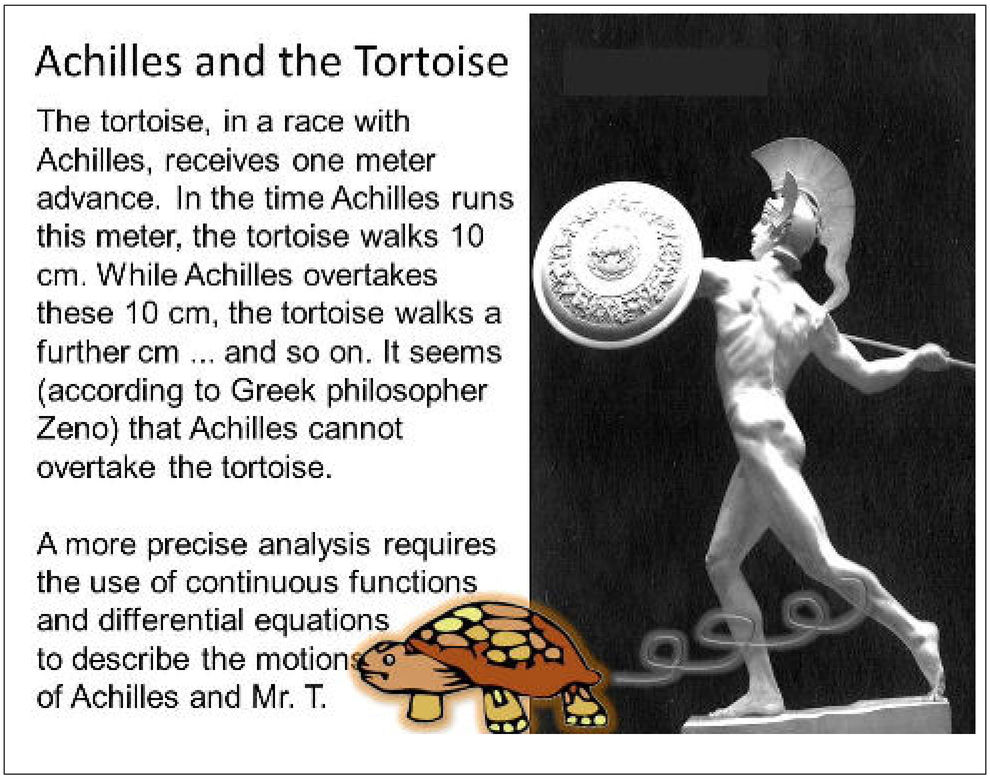


Both symbolic and numeric methods have their advantages and disadvantages[23]; see also Table 1.2. In case we are only interested in numerical estimates, and perhaps an indication of their

[23]The image of a control panel of an early analogue computer was taken from http://commons.wikimedia.org/wiki/File:MA-48L_analogue_computer,_pt._4_(2232406516)_(2).jpg?uselang=nl

accuracy, numerical methods are the best choice - although maybe at the expense of substantial computational effort. Numerical schemes are often simple to derive, and they are robust against small modifications of the precise form of the model. To the contrary, symbolic methods give deeper insight in the structure of the studied system, they typically don't require extensive computer processing, but they are much more selective as to their application: the smallest modification of a problem may render an analytic technique useless.

3.4 Mathematical Preliminaries - More about Differential Equations



In this chapter time and state transitions ar discussed[24]. We saw that recursive functions can be used to evaluate or unroll a behavior. If infinitesimal intervals (i.e. taking the limit) are chosen, recursive functions become differential equations (see Section 3.3.4). In the basic course Calculus differential equations were discussed. Some methods to solve differential equations were given. However, many differential equations cannot be solved exactly. In this section Euler's method to approximate the solution of a differential equation will be examined. To introduce this we will discuss direction fields.

3.4.1 Direction Fields

In Section 3.3.4 a direction field in the figure 'Arrows Pointing to a Solution' is given. This direction field can be found in the following way. Consider the differential equation given in Section 3.3.4

$$y'(t) = Cy(t), \tag{3.33}$$

[24]The image of Achilles was taken fom http://commons.wikimedia.org/wiki/Category:Achilles#mediaviewer/File:Wilhelm_Wandschneider_-_Achilles_(Modell).jpg; the tortoise was found at http://www.rgbstock.nl/photo/oGGAAfK/schildpad+clip+art

which describes the growth of bacteria, among other things. C is a constant with unit=1/unit of t. We know that the exact solution of this differential equation is given by $y(t) = Ae^{Ct}$, where the constant A can be found if an initial value is given (for example $y(0) = 2$). In a direction field an arrow starting at a point $(t_0, y(t_0))$ indicates the instantaneous rate of change of a specific solution for a given value t_0. This instantaneous rate of change is equal to the slope of the tangent of the tangent line to the solution curve at the point $(t_0, y(t_0))$. Therefore it equals $y'(t_0) = ACe^{Ct_0}$. So, for a given value of t_0 several arrows can be given by choosing different values of A. The arrows each are tangent to one of the exact solutions.

However, one does not need to know the exact solution of the differential equation to draw a direction field as can be seen by the following. Consider a point (t_0, y_0) and suppose that this point is part of a solution curve. So, there exists for some initial value a solution $y(t)$ of the differential equation with $y_0 = y(t_0)$. We know, using the differential equation, that $y'(t_0) = Cy(t_0)$ (see Expression 3.33) at $t = t_0$. Therefore, the slope of the tangent line to the solution curve at (t_0, y_0) is equal to $Cy(t_0)$. So, e.g. for $C = 1$, if we choose for example the point $(2, 4)$, we know that the instantaneous rate of change is equal to 4 and we can draw an arrow in the point $(2, 4)$.

The direction field gives an impression of the behavior of the solution curves. The solutions, however, cannot be derived from the direction field, neither numerically nor in closed form. The reason is that the arrow gives the direction of change for a given point (t_0, y_0), but we do not know which other points in the plane belong to the same solution curve. But we do know that the line through the arrow also represents a linear approximation to the function. So if we choose points on the line close to the point (t_0, y_0), these points might be close to points that belong to the actual solution. This idea is used in Euler's method.

Euler for Ever

Leonard Euler (1707-1783), namer of the famous numerical method for solving 1st order differential equations, is mainly known for his discovery of the fundamental relationship between the five most important numbers in mathematics: 0,1,e,π, and i (the imaginary number with property $i^2 = -1$)

One way, although perhaps not the least painful, to always remember it, is to have it scarified on one's shoulders.

3.4.2 Euler's method

Consider the following differential equation with given initial value

$$y'(t) = f(t, y), \qquad \text{with} \qquad y(t_0) = y_0. \tag{3.34}$$

The equation of the tangent line to the solution curve at the point (t_0, y_0) is given by

$$y = y_0 + y'(t_0)(t - t_0). \tag{3.35}$$

We can use this to find an approximate value y_1 of the solution at $t = t_1$. This approximate value is the y-coordinate of the point on the tangent line for $t = t_1$. So,

$$y(t_1) \approx y_1 = y_0 + y'(t_0)(t_1 - t_0). \tag{3.36}$$

This approximation is in general better if the value of t_1 is closer to the value of t_0. The value y_1 can be used to find an approximation of the value $y(t_2)$ at $t = t_2$

$$y(t_2) \approx y_2 = y_1 + y'(t_1)(t_2 - t_1). \tag{3.37}$$

This process can be repeated. The values $t_1 - t_0$ and $t_2 - t_1$ in Expressions 3.36 and 3.37 are called *step sizes*. Typically, the same step size is chosen throughout; it is denoted by h. The value of the derivative can be found with the differential equation given in Expression 3.34. Therefore

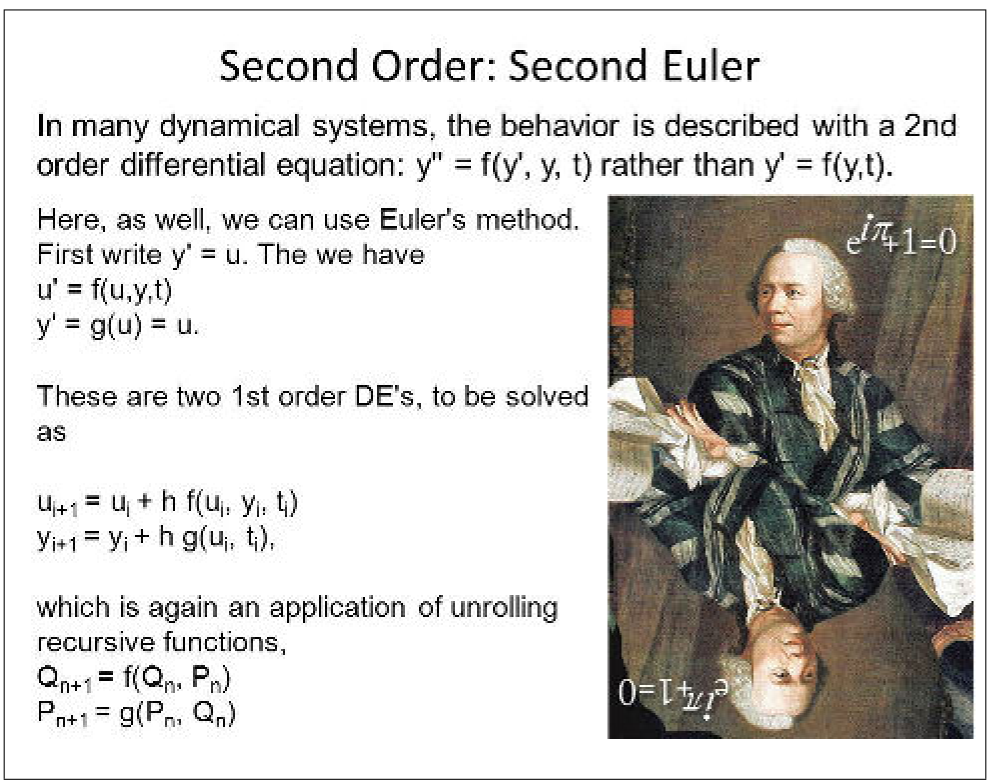
$$y(t_1) \approx y_1 = y_0 + h \cdot f(t_0, y_0), \qquad y(t_2) \approx y_2 = y_1 + h \cdot f(t_1, y_1). \tag{3.38}$$

Continuing in this way, Euler's method[25] finds the sequence of approximate values

$$y_{i+1} = y_i + h \cdot f(t_i, y_i), \tag{3.39}$$

approximating $y(t_{i+1})$, where $i = 0, 1, 2 \ldots$.

In Section 'Time Lapses with Equal Length; Sampling' (3.3.3) we have seen how the behavior of a dynamical system can be obtained by unrolling a recursive function $Q_n = F(Q_{n-1}, Q_{n-1})$. Notice that Euler's method has the exact form of unrolling a recursive function. The difference between the approach in earlier sections and Euler's method is, that in earlier sections we did



[25] The photograph of the scarified version of Euler's relation is taken from http://upload.wikimedia.org/wikipedia/commons/3/36/Euler%27s_identity_scarification%2C_3PiCon%2C_Springfield%2C_MA.jpg?uselang=nl. The original portrait of Euler was taken from http://upload.wikimedia.org/wikipedia/commons/6/60/Leonhard_Euler_2.jpg?uselang=nl

not attempt to write the dynamical model in the form of a differential equation: rather, we wrote the (glass box) mechanisms directly in the form of recursive functions.

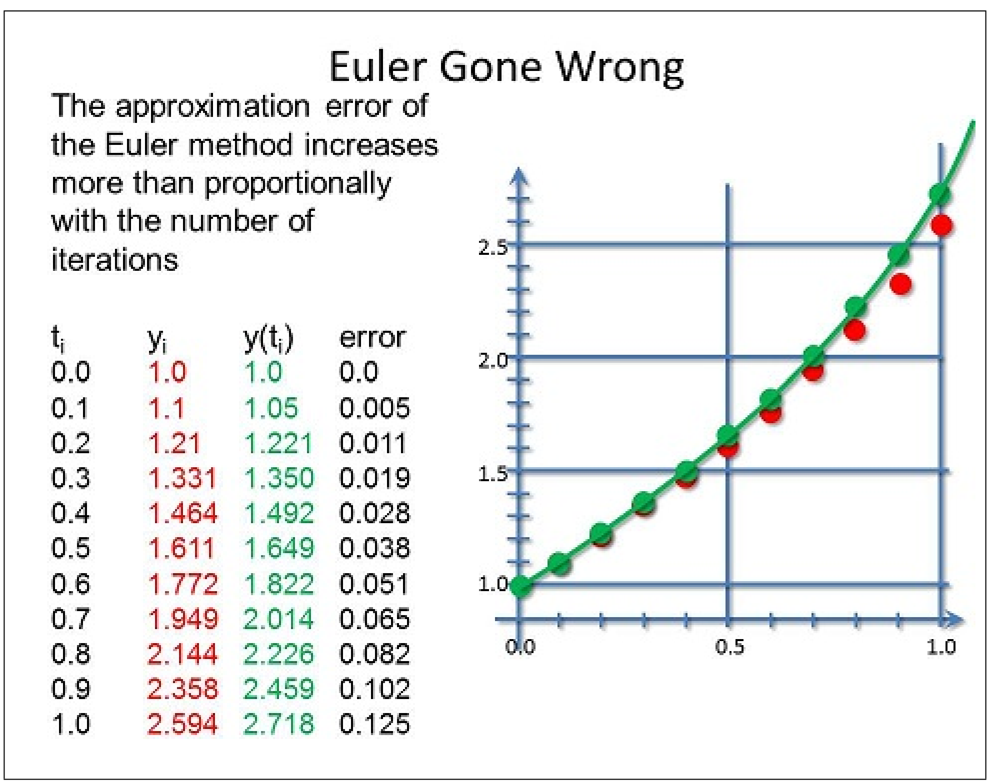
Both approaches have their strengths and weaknesses:

- *skipping the formulation of a differential equation* often gives a straightforward derivation of the recursive functions. The solution (=finding the dynamical behavior of the modelled system) does not require any analytical skills or knowledge of numerical methods. The disadvantage is, that the step size needed for sufficiently accurate approximations may need to be extremely small, leading to very large numbers of iterations.
- *formulating the dynamical system in the form of a differential equation* may require more advanced physical, mechanical or chemical reasoning. It may also, however, in some cases, lead to differential equations that can be solved in closed form by analytical means. And if this is not possible, there usually exist numerical methods that are more efficient than Euler's method.

In the following we use Euler's method to find an approximate solution of the differential equation given in Expression 3.33, where we take C=1, with initial value $y(0) = 1$. Since in this case the exact solution is known, one can compare the approximate solution with the exact solution.

With step size $h = 0.1$ one finds

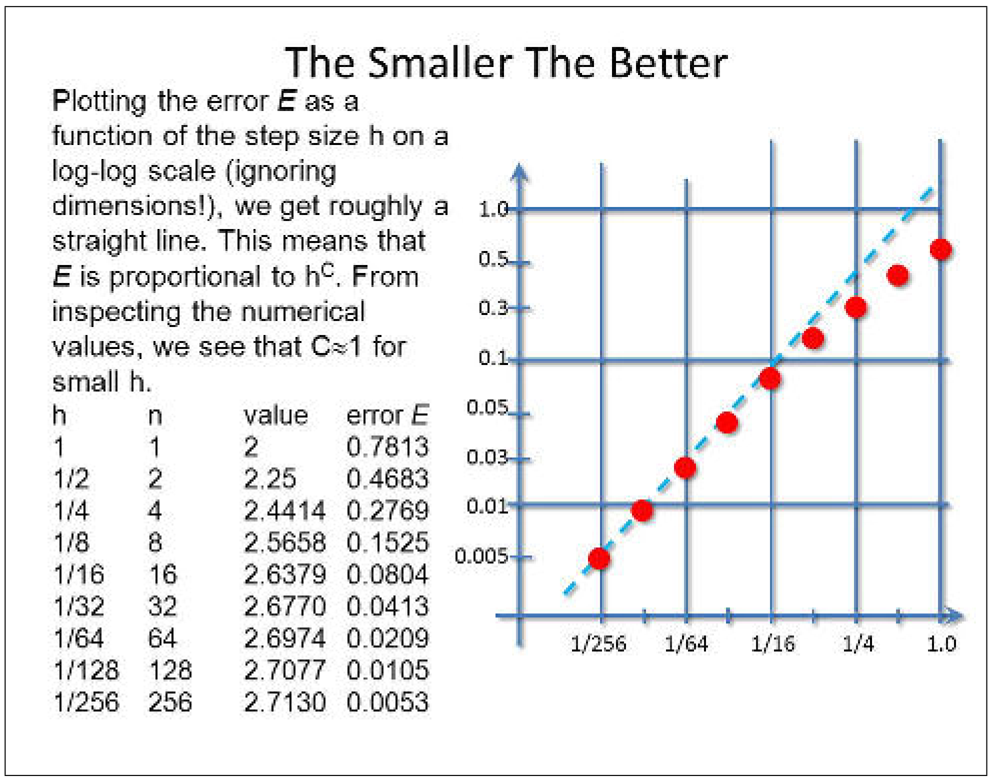
$$\begin{aligned} y(t_1) &\approx y_1 = y_0 + h \cdot f(t_0, y_0) = 1 + 0.1 \cdot 1 = 1.1, \\ y(t_2) &\approx y_2 = y_1 + h \cdot f(t_1, y_1) = 1.1 + 0.1 \cdot 1.1 = 1.21, \\ y(t_3) &\approx y_3 = y_2 + h \cdot f(t_2, y_2) = 1.21 + 0.1 \cdot 1.21 = 1.331. \end{aligned}$$



In Figure 'Euler Gone Wrong', the first ten values are given. We see that the error increases with the value of the argument t. In Figure 'The Smaller The Better', the result is shown for other step sizes. The error increases slower if the step size is smaller. In fact, the error in estimating $y(1)$ is seen to be more or less proportional to the step size.

One could (and should!) of course ask 'which step size is small enough'. This first depends on the desired accuracy of the approximated function value; this will be discussed further in Chapter 6. It also depends, however, on the rate of change of the solution we try to approximate. If the solution varies slowly, step sizes may be larger compared to the case where a solution shows rapid changes. In general, the step size should be small compared to the CHARACTERISTIC TIME. Indeed, many dynamical processes can be associated to a characteristic amount of time. For periodic processes, this is the period (e.g., for a beating heart, the characteristic time is the average duration of one heart beat). For (exponentially) increasing or decreasing behavior, it is e.g. the time of doubling or halving. For the behavior to be reproduced at least qualitatively, there should be several time steps within the characteristic time of the modeled process. So, for a differential equation with periodic solutions, the time step should be (much) smaller than the period.

Notice that this relates to dimensional analysis: the unit-less ratio $\frac{h}{T_C}$ of the step size h and the characteristic time, say T_C should be (much) smaller than 1. The number n of iterations needed to calculate an approximation to the solution for a given time t should therefore be inversely proportional to T_C. Moreover, for the Euler method we see that the error, for a given h, increases more than proportional to the duration of time for which we want to find the solution. For the purpose *prediction* we could call this the PREDICTION TIME, T_P. The number of iterations is therefore at least proportional to T_P. Combining the latter two results, the number of required iterations is at least proportional to $\frac{T_P}{T_C}$.



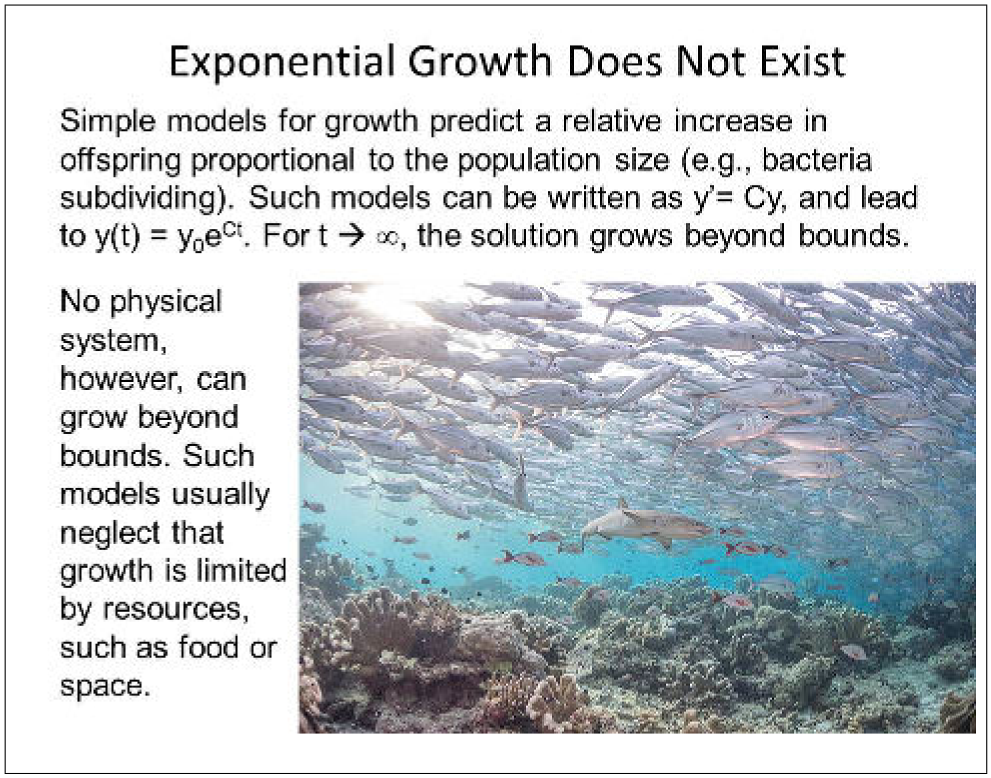
3.4.3 Equilibrium Solutions and Stability

A solution $y(t)$ of the differential equation given in Expression 3.34 is called an equilibrium solution if it is a constant function. So, $y'(t) = 0$ and from Expression 3.34 it follows that $f(t, y(t)) = 0$ for all t.

Consider the differential equation

$$y'(t) = k \cdot y(M - y), \tag{3.40}$$

which is a model for so called LOGISTIC GROWTH[26] where M is the maximum population determined by resources (with unit the unit of y) and k is the growth rate, with the unit of y per unit of time. For equilibrium solutions we must have $y'(t) = 0$ and so $y(t)(M - y(t)) = 0$. This gives $y(t) = 0$ or $y(t) = M$ for all t. Go to this link to experiment with line element fields for $y'(t) = k \cdot y(M - y)$. Adjust k and M at will; click somewhere in the image to follow a solution, starting in the clicked location.



One can see that for the line $y = M$ all arrows near the line point towards it, whereas the arrows at the line $y = 0$ diverge. Therefore the solution $y(t) = M$ is a so-called STABLE SOLUTION. An equilibrium solution is **stable** if solutions close to the equilibrium approach the equilibrium solution for $t \to \infty$. An equilibrium is **unstable** is solutions close to the equilibrium solution tend to get further away from the equilibrium if $t \to \infty$.
Mathematical definitions for these concepts are not given here. The (un)stability is checked by use of the direction field.

Material for the exam

The relevant material that one should study for this chapter is
Adams:
or
Smith and Minton:
§7.3

Exercises

Adams:
or

[26] The photograph of the school of fish was taken from http://commons.wikimedia.org/wiki/File:Reef_shark_beneath_a_school_of_jack_fish.jpg?uselang=nl

Smith and Minton:
§7.3:

3.5 Summary

- A *state* is a snapshot of a conceptual model at some time point;
- The *state space* is the collection of all states of a modeled system;
- Change comes in the form of *transitions* between states; a *state chart* is a graph where nodes are states and arrows are transitions;
- A *behavior* is a path through state space; a process is the set of all behaviors. The size of state spaces is huge; this is called *state space explosion*. Two methods to mitigate state space explosion:
  - *symmetry*: some parts of state space are identical and therefore redundant;
  - *projection*: distinguish *exposed* and *hidden* properties or value sets;
- Multiple flavors of time:
  - *partially ordered time*, for instance for specification and verification;
  - *totally ordered time*, for instance for prediction, steering and control;
  - A *recursive* function of the form $Q_{i+1} = F(Q_i, Q_{i-1}, Q_{i-2}, ..., P_i, P_{i-1}, P_{i-2}, ...)$ is used to evaluate or *unroll* a behavior;
    * *equal intervals*: the possibility for closed form evaluation (example: periodic financial transactions); *sampling*;
    * *equal, small intervals*: approximation, sampling error (examples: moving point mass, rotating dumbbell, mass-spring system, dissipation);
    * *infinitesimal intervals*: *continuous* time, differential equations (examples: motion of a point mass with or without force); contrast between numerical and symbolic approach

3.6 Learning goals

3.6.1 Knowledge

You should know the meaning of the terms state, state diagram, transition, process, behavior, event, state space explosion, deadlock, projection, hidden and exposed properties; the various types of time models (partially ordered time, total ordered time, with or without equal time lapses); you should know what recursive functions are, and how to use them, and what sampling is. You should know the main ideas behind numerical and symbolic solution of differential equations, their use, application, advantages and disadvantages.

Regarding the mathematical notions in this chapter, Emiel: aanvullen svp.

3.6.2 Skills

This chapter contains numerous examples. For each of the examples, you should be able to invent at least two very different other ones. You should be able to set up a state chart for an arbitrary discrete dynamic system (with a limited amount of states). You should be able to distinguish permitted and forbidden transitions; you should be able to choose among various alternative choices for state properties. You should be able to calculate the number of states in a state chart. You should be able to choose hidden and exposed properties, and reason about the consequences of your choice. You should be able to recognize and, for simple dynamic systems, devise a recursive function. Also for simple dynamical systems, you should be able to unroll a recursive function (using a compute environment of your own choice); you should be able to choose a stepsize and understand its consequence in terms of accuracy and computation time. Even if you may not be able to actually solve a differential equation by analytic means, you should be able to make a substantiated choice between numerical and symbolic methods for modeling a dynamic system, given the purpose you try to achieve. Notice: the mathematical treatment of the rotating dumbbell and the mass-spring system, elaborated in the end notes, are not part of the mandatory material.

For the mathematical notions in this chapter, Emiel: aanvullen svp.

3.6.3 Attitude

When confronted with a problem that involves change over time, you should be inclined to argue about it in terms of states and state transitions. For discrete systems, you should be inclided to think in terms of state charts; you should have the tendency to try and find representations that lead to transparent and possibly small state spaces. For systems with totally ordered time, you should be inclined to search for recursive functions; when asked to implement a simulation of a dynamical system, you should either be tempted to seek help to find a closed-form solution, or try to implement an algorithm to unroll the appropriate recursive function, both depending on the purpose of your model. When using a numerical method, you should be inclined to think about accuracy in terms of step size, and you should do experiments to try and estimate the accuracy of the numerical outcomes.

3.7 Questions

1. In the second paragraph of Section 3.2, a large number of occurrences in relation to lights and radios are mentioned. Which of these are states, which are events?

2. For each process mentioned in the second paragraph of Section 3.2, draw a little state chart.

3. We say 'Applying butter is to be synchronized with the other two stages of the process'. Say this in your own words.

4. What is the use of a state chart?

5. Draw a state chart for making tea that takes into account that at any time the telephone might ring.

6. What do we mean by 'binding'?

7. How long does a state transition take?

8. A system has N_p properties, each of type `Boolean`. How many behaviors of b steps are possible?

9. What is the difference between a behavior and a process?

10. Describe one shot in a game of billiard as a state chart.

11. Explain the notion of 'symmetry' in your own terms.

12. Give an example of symmetry for the 4 color ballpoint.

13. We explain symmetry by means of an example, taken from railroad safety context. We give four requirements ('two trains shall never occupy [...] waiting for a red signal'). The last one of the list has a different character from the earlier three. Explain.

14. What is the meaning of 'reachable' in state spaces?

15. Hiding properties can cause a system to behave seemingly random. Explain.

16. If the ink levels in the 4 color ballpoint are only distinguished between 'empty', 'full', and 'half', how many states does the system have?

17. The 4-color ballpoint could also be modeled with 5 properties as follows: `[levelRed:{0...100}%, levelGreen:{0...100}%, levelBlue:{0...100}%, levelBlack:{0...100}%, whichIsOut: {red,grn,blu,blk,none}]`. How many states does this system have? Try describing the behavior of the pen by means of this model. Compare with the initial model. What behavior can be expressed, what what cannot be expressed in state charts? Why?

18. Explain in your own words that the 9 nodes in Figure 3.1 indeed represent the full state space of the four color ballpoint.

19. What is lifelock?

20. What is the difference between an event and a transition?

21. Why is it allowed to assume that no two events can occur at the same time?

22. What is an internal transition?

23. In each of the three examples of internal transitions in Section 3.3.1, think of one or more hidden properties that should be made exposed so that the transitions are no longer internal.

24. For the three examples of partial ordering in Section 3.3.1, draw part of a state chart that illustrates the partial ordering.

25. In the case of totally ordered time, we enumerate both states and transitions. Why is this allowed?

26. In a dynamic model with totally-ordered time, t_i, t_j and t_k are times where transitions occur. From $\Delta(t_i, t_j) + \Delta(t_j, t_k) = \Delta(t_i, t_k)$ it follows that transitions take no time. Explain.

27. In the paragraph 'Total Ordering and Causality' we discuss the profitability of a bus route. We use the quantities E_i and L_i, for the number of entering and leaving passengers for every stop. Re-write the example with the number of passengers in the bus.

28. Explain in your own words: what is recursion?

29. (*) Causality forbids that the cause of some occurrence takes place after the occurrence. Is causality a sufficient condition for the existence of a recursive function to simulate a dynamic process?

30. Explain why a recurrent function of the form $Q_{i+1} = F(Q_i, P_{i+1})$ is forbidden.

31. Explain the meaning of 'order' in 'the order of a dynamical system'.

32. (*) You may know what the 'order of a differential equation' means. If so, explain the relation between this meaning of 'order' and the 'order of a dynamical system described by a recurrent function'.

33. We claim that for a system, described by a recurrent function of the form $Q_i = F(P_{i-1})$, everything is static. Explain.

34. Give an example of a system with order > 2. (Hint: don't think of a physical system).

35. In Section 3.3.3 we say that a fraction $s_0\Delta_0$ is spent in time interval Δ_0. Explain the product $s_0\Delta_0$.

36. In the financial example in Section 3.3.3, we calculate how much we are allowed to spend for a given income. We could also ask the question: how much do we need to earn for a given spending rate. Re-do the example to answer this latter question.

37. We given an example of recursion to calculate $N!$. How many multiplications does this recipe take? How can you do it quicker?

38. In the closed-form calculation of the financial example, we spend a fraction $s\Delta$. Redo the example where we *save* a fraction $s\Delta$.

39. Schemes for unrolling recursive functions may introduce an error. The example for a moving point mass with constant velocity we derived in the text had order 1. What does this mean?

40. Schemes for unrolling recursive functions may introduce an error. The example for a moving point mass with constant velocity we derived in the text had order 1. For what reason is a scheme for unrolling recursive functions that has order 2 better than a scheme with order 1?

41. Verify that the scheme from Expression 3.28 for unrolling recursive functions for the accelerated point mass introduces no error.

42. Verify Expression 3.25.

43. Explain the step leading to Expression 7.12 and Expression 7.13.

44. Verify Expression 7.22 and Expression 7.22.

45. In the example of the rotating dumbbell, the solution Expression 7.21 contains a factor $\frac{1}{2}$. You may remember the formula for centrifugal force, $F = m\omega^2 r$, which contains no such factor. Explain why both formulas are consistent.

46. In problems such as the rotating dumbbell, there is a maximum time step Δ. Relate this maximum time step to the rotation frequency.

47. We propose, in Expression 7.28, to represent damping by $\epsilon(r_i - r_{i-1})$. Show that this indeed gives a reduction of $(1-\epsilon)^2$ per time step of the kinetic energy.

3.8 Exercises

1. In at least two different ways, describe one shot in a game of billiard as a state chart (for both, formulate a purpose - the purposes don't need to be very plausible).

2. In the text, we give three examples where symmetry helps reducing the size of state space. Which three? Give a fourth example.

3. Consider the four color ballpoint.

   (a) The number of permitted states of the four color ballpoint is 5/32 of the number of possible states. Explain.

   (b) Change the conceptual model of the four color ballpoint so that all possible states are permitted.

4. The externally visible behavior of a baby can be characterized by two properties: `sound:[cry,silent]` and `consciousness:[sleep,awake]`

   (a) Make a state chart with the states and transitions among these properties.

   (b) For each transition, indicate whether it can be causally understood or not.

   (c) Babies often cry because their stomach is empty. Add a (hidden) property `stomach:[full, empty]`, and extend the state chart accordingly.

   (d) Give an interpretation of each transition: is it caused by internal events or external events?

   (e) Think of a purpose for which the latter state chart would be a useful model.

   (f) Think of an extension of the model such that the purpose of Question 4e is fulfilled better (you don't need to give the full state chart for your extended model).

5. Consider a coin-operated coffee machine

   (a) Propose a purpose for a dynamic model.

   (b) What properties of the coffee machine need to be taken into account, and what are their value sets?

   (c) Set up a state chart for the coffee machine.

   (d) Propose a property to be hidden. Which part of the behavior then becomes 'random' (=unpredictable)?

   (e) Propose an extension of the model such that your purpose is fulfilled better (you don't need to give the full state chart for your extended model).

6. (a) Propose a system (other than a baby, a four color ballpoint and a coffee machine), and a purpose that would merit a dynamical model.
   (b) Propose a small, but meaningful set of properties for your system such that a statechart model would help to realize the purpose.
   (c) Propose the smallest, but meaningful sets of values for these properties.
   (d) Give a state chart for the dynamical behavior of the model.
   (e) Think of an extension that would better serve the purpose (you don't need to give the full state chart for your extended model).

7. Give a state chart of the game of tick tack toe; make use of symmetry to reduce the number of states.

8. In Section 3.3.2 we introduce a simple pricing model for a bus route. This model is naive. Why? Suggest at least two improvements. Elaborate these improvements in the form of recursive functions.

9. Give an example of a dynamical system with an order larger than 2.

10. The Watt regulator is a system to control the speed of a steam engine. It consists of two weights, connected on a vertical shaft that rotates with a speed equal to the speed of the flywheel. Call the rotation speed ω. The weights are mounted on hinges that allow them to move outward due to centrifugal force, and at the same time go up. The height of the weights is h, and h depends on ω. The height of the weights is used to control a valve, which regulates the steam pressure P: when the weights go up, the steam pressure is lowered and vice versa. So P depends on h. Finally, the rotation speed ω depends on P and on the external load L: ω increases with P and decreases with L.

    (a) Give the simplest possible functions that express how h depends on ω, how ω depends on P and L, and how P depends on h.
    (b) Show that these three functions are not in the form of a set of recursive functions of Expression 3.2.
    (c) Give a function that expresses a *change* in ω in terms of a *change* in P and L, and similar for h and P.
    (d) Show that the latter functions are in the form of Expression 3.2.
    (e) (*) Show that the Watt regulator stabilizes the rotation speed of a steam engine - that is: when the rotation speed changes due to a change in L, the rotation speed will soon get back to the original speed.
    (f) Discuss limitations in the functioning of the Watt regulator according to the model.

11. In Section **??** we give a model for a mass spring system using discrete time steps. Derive a similar model for a pendulum.

    - First, assume that the deviation angle ϕ is small enough so that we may approximate $\sin(\phi)$ by ϕ.

- Show that the approximated solution is independent of the maximum deviation.
- Show that, for sufficiently small time step Δ, the approximated solution converges to the closed form solution.
- Next, drop the assumption of small maximal deviation. Again, give the discretely approximated solution.
- (*) In the example of an oscillating mass-spring system, we saw that there is a condition relating the time lapse Δ to the frequency of resonance, $\sqrt{\frac{C}{m}}$ of the mass spring system: the approximation only works id Δ is sufficiently small compared to $\sqrt{\frac{m}{C}}$. Can you find a similar condition for the case of the pendulum?

12. There are two armies; their numbers of soldiers initially are S_1 and S_2. When they engage in a battle, it takes e_1 soldiers of army 1 to kill one soldier of army 2, whereas it takes e_2 soldiers of army 2 to kill one soldier of army 1. Devise a dynamical model, in the form of recursive formulas for S_1 and S_2 that depend on time, such that we can predict the outcome of the battle.

13. In a biotope, there lives a population of fox and a population of rabbit. The sizes of the two populations are F and R; both depend on time, and their initial values are F_0 and R_0. We want to predict if the two populations can continue to live together in the same biotope.

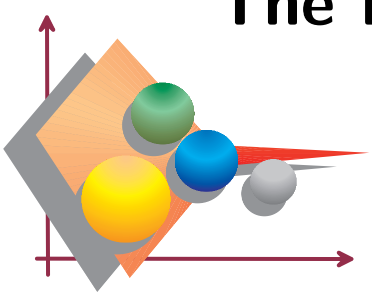
    (a) Which factors influence next years population sizes?

    (b) How does the simplest mathematical relations look like to account for these factors?

    (c) Devise a discrete dynamical model, using recursive functions, to describe the evolution of F and R over time.

Chapter 4

The Function of Functions



'Who put the 'fun' in 'function'?

The last rays of sun color the ancient city walls with bronzen fire. A full moon is rising in the east, and its pale light adds a tint of blue to the deep shadows of the ziggurat. Soon the million diamonds of the desert night will shine their silent light into the cold, but now the sand is still warm. A small group of merchants seek shelter for the heat beneath the city gates. "How much"? A young trader with a foreign accent weights his purse. The other man does a quick calculation on his fingers. "Ten and two for the camel, and three more for each of the donkeys". He squints his eyes. The foreigner takes a deep breath. "Ten and eight. You ask a lot. For the sake of Hourmazd, they better be healthy!" Little do they know that at the same time, high above their heads, two astronomers dispute the planets. "I tell you, twelve degrees over the last month: in two more weeks, three degrees per week: Ishtar, be her beauty forever blessed, will have gained another six, eighteen in all!" No fingers are counted, as astronomers read the heavens as they read the circles of the zodiac, three hundred sixty degrees in a full circle it was and always will be.

4.1 What is a Formal Model?

Babylonian merchants nor astronomers-priests would recognize the homework of a twenty-first century schoolboy, doing arithmetic by writing numbers on paper, aligning the digits, and adding them one by one, starting from the right. But neither would ancient astronomers understand the finger-counting of the merchants of their days, nor would merchants comprehend the ritual circle-shifting of astronomers.

Ancient astronomers, antique merchants and present day schoolboys all perform exactly the same operation: addition. Whether the payment for livestock, the inclination of planets or just numbers: addition is addition. Or ... is it?

More than Pilsener

A marketeer, needing a catchy brand name to help a product stand out from the average competitor, may consider the affix 'plus'.

'Plus' has the connotation of completion, growth, increase or abundance.

But only in rare cases, completion, growth, increase or abundance are actually formalised adequately by the mathematical operation that takes 1 and 1 and yields 2.

Assyrian and Babylonian astronomers used a hexadecimal number system to represent angles. In the same era, and the same area, merchants counted on their fingers, using a decimal system without realizing. Since adding angles, or adding amounts of coins at first sight seem different, since the two social groups were very much separated, and since numbers systems used for the two applications differed, unification of the two types of operations had to wait some thousand years[1].

This oversight seems hard to imagine. How could it be that intelligent people did not arrive at the idea of addition, as children in our days learn the technique at age 7?

Addition is an abstract manipulation with numbers, to represent the intuition of 'bringing together'. To see the subtleties of this representation, we look at some further examples.

The chance of throwing 6 with a dice is 0.1667. The chance of throwing a 12 with two dice is not 0.1667+0.1667=0.3333. Stronger: the chance of throwing 6 with two attempts is also not 0.1667+0.1667.

In electric circuits: the combined effect of two resistors, R_1 and R_2 can be $R_1 + R_2$, but it can also be $\frac{R_1 R_2}{R_1+R_2}$, depending on whether they are switched in series or in parallel.

In mechanical systems: the combined effect of two springs, with spring constants C_1 and C_2 can be $C_1 + C_2$, but it can also be $\frac{C_1 C_2}{C_1+C_2}$, depending on whether they are mounted in parallel or in series (exactly opposite to resistors!).

In geometry: when p and q are two locations in space, the addition $\frac{1}{2}p + \frac{1}{2}q$ is the point right in the middle between p and q. Similarly, $\frac{1}{3}p + \frac{2}{3}q$ is the point twice as close to q as to p. But the addition $\frac{1}{3}p + \frac{1}{2}q$ has no meaning ▷75 .

So: in some cases 'bringing together one thing to another thing' can be represented by the mathematical operation '+', and in other cases this is not allowed. The question is: in which situations do we use mathematical addition? Or, conversely: which mathematical operation do we need in a given situation?

This question lies at the heart of the modeling process.

[1]The photograph of the beer bottle was taken from http://commons.wikimedia.org/wiki/File:Pils_Plus.JPG?uselang=nl

In Chapter 2.1 we learned to develop a conceptual model, that is: make an inventory the terms and notions we need to consider to serve our purpose, but don't do any mathematics yet.

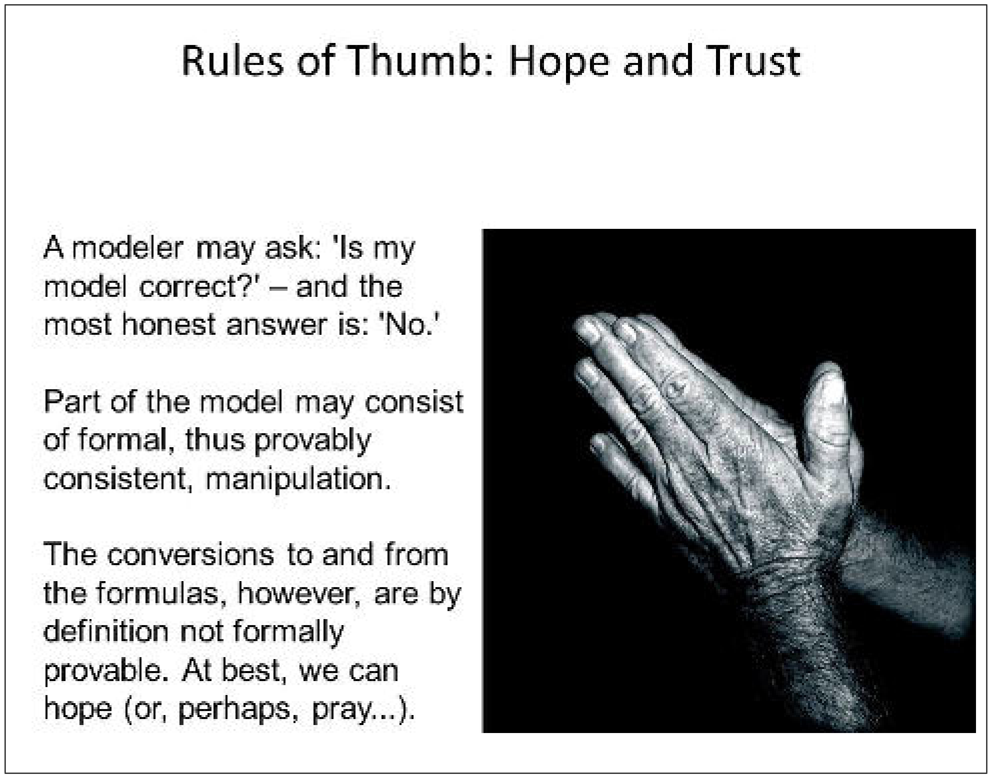
In the formalization stage we develop the *conceptual* model into a *formal* model. Intuitions expressed in natural language are to be translated into mathematical relations.

With respect to this formalization, the following holds: *There is no formal, provable correct way to translate intuitions about things in the material world into mathematical constructions* ▷76 [2].

There are, however, rules of thumb ▷77 .

In the sequel, we explore a practical route to constructing models. This route is a HEURISTIC. That is: it is not a *proven* method. There is no guarantee that it leads to a purposeful model in any given situation. If it does, there is no guarantee that the model is in some sense optimal. If it fails there may not be other routes that do succeed.

It is, however, a practical route.



4.2 A Practical Route to Formal Models

The heuristic route to formalization comprises the following elements: *naming*, *structure*, and *formation of mathematical expressions*.

4.2.1 Naming

In Chapter 2.1, we encountered naming schemes for the properties of concepts. We saw various ways to access the values of properties, for instance using the dot-notation. But the dot notation, and similar forms, can be cumbersome when we do mathematics. Therefore we introduced quantities, so we not always need to know the concepts the quantities belong to. A simple name can be just as good. In naming, the following rules of thumb should be taken into account:

- names should be unambiguous;

[2]The photograph of the praying hands is taken from http://www.rgbstock.nl/photo/2dyX5Gs/Praying+Hands+-+Duotone

- names should not conflict with standard usage (e.g., don't call a quantity `SIN` because SIN will be associated with trigonometry);
- names should not rely on the distinction between lowercase and uppercase, or the distinction between o (letter) and 0 (digit), etc.;
- names should not start with a digit;
- if Greek or other non-Western alphabets are used, a transliteration is needed to convert formulas into computer script. For instance, ρ might become `rho`;
- self-explanatory names help to understand formulas[3]. This leads to long strings, however, perhaps including multiple words in one name. In computer scripts this is fine, as long as no spaces occur; to improve readability, a device called CAMEL CASING is attractive, so `alongandunreadablename` becomes `aLongAndReadableName`.

What's in a Name?

Except to people reading Welsh, the sign below, found at a small Welsh railway station, means little or nothing to most. It counts as the longest name in the world, and it is actually an accurate description (in Welsh) of the geographical location it refers to.

Although camel casing could have helped somewhat, it is recommended to use strings that are more meaningful to a preferably broader audience when naming quantities. And that are perhaps shorter, too.

It often happens that quantities are introduced halfway during constructing a model. In Section 1.4 we recommend that the conceptualization stage yields an inventory of all concepts and their properties. These properties were to become quantities in the formalization stage. Introducing a new quantity halfway through the formalization stage does not mean, however, that we always need to go back to the conceptualization stage. If the number of quantities in the model is less than, say, 100, we may just work with some 'loose' quantities. For these 'back of the envelope'-models, the conceptualization stage is merely a systematic way to collect enough relevant factors to get going; later we may not need the conceptual model anymore.

For larger models the conceptual model will figure in other contexts as well. It will, for instance, document the model, or it enables the model be analyzed by automated means. In that case the full consistency between the conceptual model and the formal model is to be maintained, and the introduction of a new quantity during the formalization stage needs to be backed up by adjustments in the conceptual model.

---

[3] The photo of the Welsh railway sign was taken from http://commons.wikimedia.org/wiki/File:Llanfair_railway_station_sign_banner.jpg?uselang=nl

4.2.2 Structure

There are two, related, devices to add structure to the formalization: the *chain of dependencies* and the *to-do-list*.

Chain of Dependencies

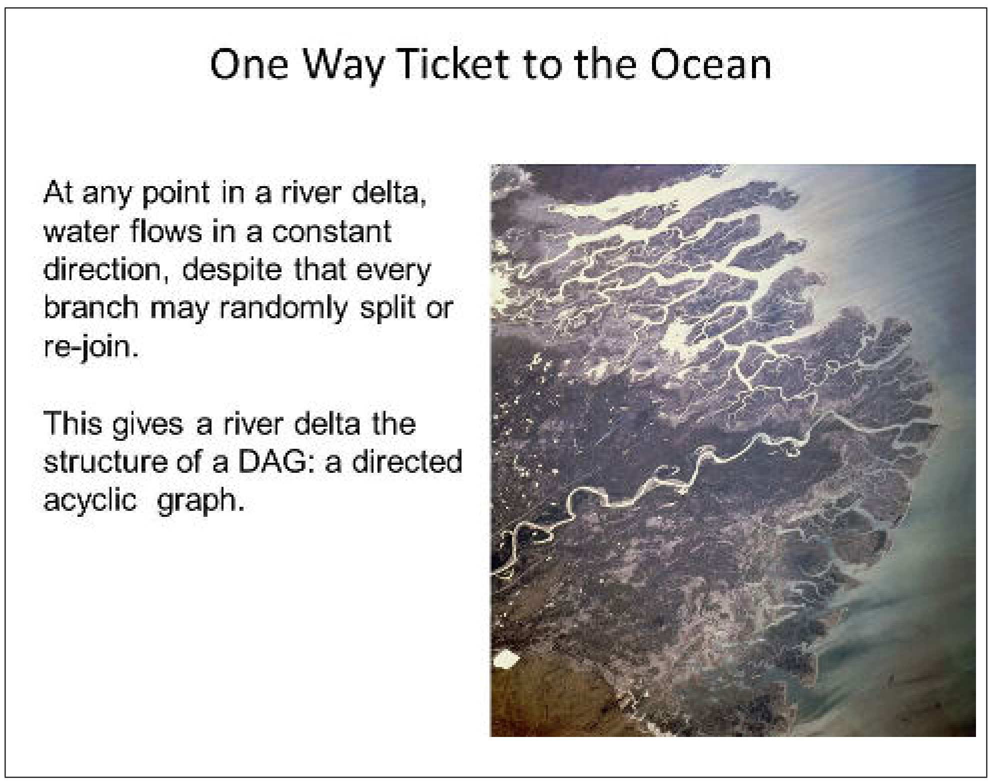
Quantities depend on other quantities as given by a *directed, a-cyclic graph*. We abbreviate this by DAG.

A directed graph is collection of nodes, connected by arrows rather than arcs.

A graph is A-CYCLIC if it is not possible to make a closed path along the nodes, following arrows only.

Every arrow in the DAG designates the relation `dependsOn`. With a relation `isArgumentTo`, we would get a graph where every arrow has the opposite direction.

Nodes with only outgoing `dependsOn`-arrows are called ROOTS.



To every node in the DAG we associate a quantity.

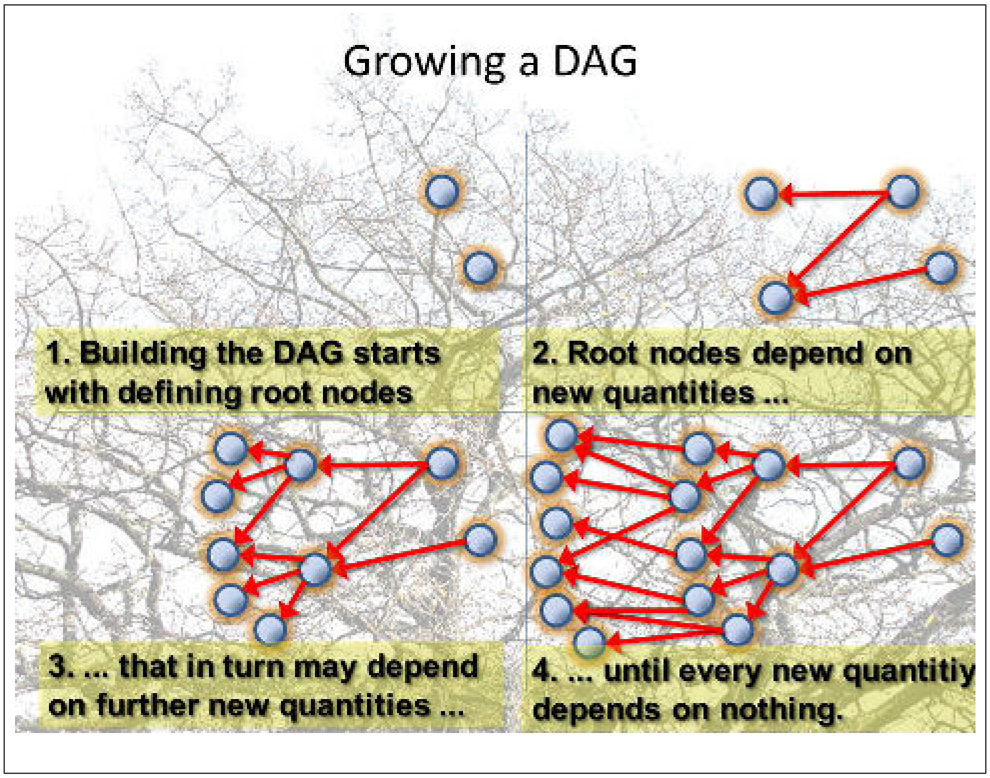
Each root node is a quantity we are *interested in*, depending on the purpose of our model. The purpose of the model defines the root nodes of the DAG[4]. A root node is, for instance, a quantity for which we need to predict its value, or something we need to optimize.

The DAG expresses this interesting quantity in terms of things it depends on, which are again expressed in terms of things *they* depend on, and so on. This goes on until we reach nodes in the DAG that need no further expansion. Nodes that need no expansion are nodes without outgoing `dependsOn`-arrows. They are called LEAVES, and they depend on nothing. A leaf, for instance, can be a constant number or a quantity that can be freely chosen.

In Chapter 5, the nodes in the DAG will be labeled with categories. The root nodes will be labeled as category II; a leaf that is a constant is category III, and a leaf that can be freely chosen is

[4]The photograph of the Indus delta is taken from http://upload.wikimedia.org/wikipedia/commons/e/ef/Indus_River_Delta.jpg?uselang=nl

category I. All other nodes are category IV.



Expanding the DAG, that is: adding nodes to a DAG, starts in an existing node which represents some quantity, say x, with a value that is as yet unknown. Expansion consists of four steps:

- *1* think about the simplest mechanism that causes x to take its value;
- *2* express this mechanism in terms of other quantities (perhaps new ones);
- *3* state qualitatively how this dependency looks like (e.g., is it increasing? is it oscillatory?);
- *4* translate this qualitative statement into a quantitative statement (i.e., a mathematical expression).

An example may illustrate this.

Consider the relation `sees` in the road illumination problem, introduced in Section 2.5. The quantity we are interested in is l_p, the perceived light intensity.

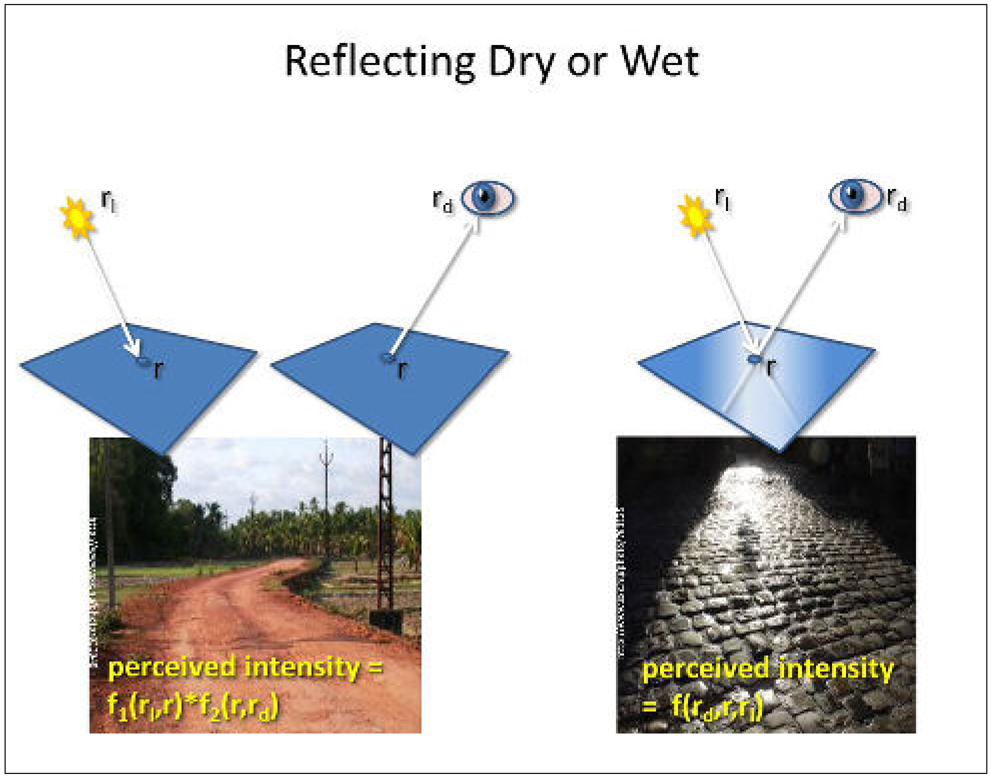
The above steps are:

- *1. What is the mechanism?* l_p gets its value as a consequence of the light of multiple lanterns, reflecting on the road, shining into the eyes of the driver. There are three involved concepts (`lantern`,`driver`,`road`), hence a 3-ary relation.

How to formalize `sees` depends on whether we deal with a dry or with a wet road.

Let us first consider the simple case of a dry road. Then the road is a *diffuse* reflector. Every point of the road surface acts as a miniature light source (point source), shining equally bright in all directions. The intensity of this point source is the result of the incoming light of all contributing lanterns. The value of l_p for a diffuse reflector is independent of the angles of ingoing and outgoing light rays. So l_p only depends on the light intensity on the road in the point where the driver is looking at. Therefore the relation `sees` can be written as (1) a relation between the driver and the road and (2) a relation between the road and the lanterns. The 3-ary relation `sees` separates into two 2-ary relations.

This differs from the situation of a wet, and therefore *specularly* reflecting road. For a mirror, the reflection is high if the angles of incoming and reflecting rays are equal and opposite. A wet road resembles a mirror. Individual lamps can be distinguished, each reflected in the road surface, whereas on a dry road the light of all lanterns blends together into one smooth distribution of light. So for specular reflection, l_p again depends on the point of the road where the driver is looking at, and it also depends on the locations and intensities of all contributing light lanterns.



- *2. What are the quantities this mechanism depends on?* In the diffuse case, l_p only depends on the light intensity, R_I, in the point r of the road where the driver is looking at, and the location of the driver, r_d.

  For the specularly reflective case, every tiny piece TP of shiny road surface acts as a small mirror. It causes a relation between the direction of the light ray, coming in onto TP, and the direction of the ray leaving TP due to reflection. If the angles of these directions are equal and opposite, the amount of reflection in TP is maximal. It decreases when the angles are more different. The quantity l_p therefore depends on the driver's location, r_d, *but also* on the locations $r_{l;i}$ and intensities P_i of each of the individual, contributing lanterns. The lanterns are distinguished by i, $i = 1, 2, 3, \cdots$. Further, l_p depends on the reflectivity of the wet road surface, ρ, and a quantity β that determines how glossy the road is. Large β: like a mirror; medium β: like plastic; low β: like plaster or clay.

- *3. What is the qualitative form of the mechanism?* This is the same in the diffuse case and the specular case. It is the amount of light that can be seen when looking, from the point r_d, in the direction of the point r of the road. In the diffusely reflecting case it is independent of the looking direction, $r - r_d$; in the specularly reflecting case, it does depend on the looking direction.

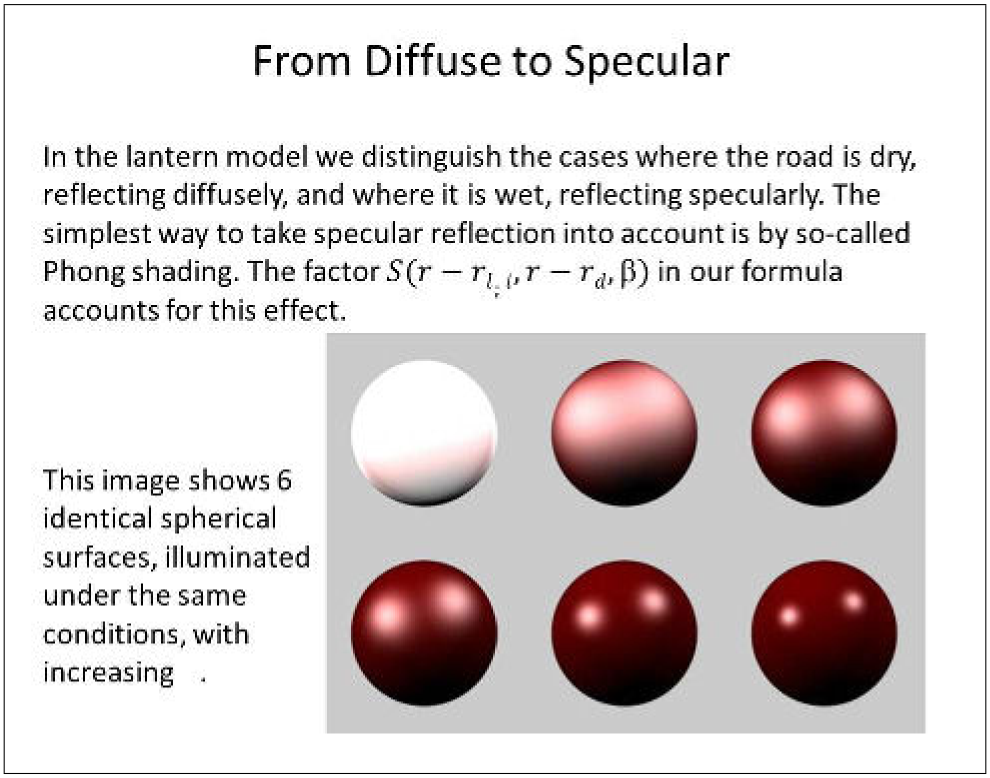
- *4. What is the quantitative form of the mechanism?* In the diffuse case:

  We defined R_I as the emitted brightness in the point r. The perceived brightness in the location of the driver, r_d depends on R_I and on the distance D_{driver} between r and r_d; it decreases with the square of this distance.

$$l_p = \frac{R_I}{D_{\text{driver}}(r, r_d)^2}. \tag{4.1}$$

In two next expansion steps, we expand D_{driver} and R_I. Quantity D_{driver} follows from Pythagoras' theorem: $D_{\text{driver}} = \sqrt{||r_d - r||^2}$. R_I depends on the location r on the road, and on the positions and intensities of the contributing lanterns. It is given by a summation over all lanterns:

$$R_I(r) = \Sigma_{i:\text{contributing lanterns}} \frac{P_i \rho}{D_{\text{lantern}}(r, r_{l;i})^2}. \tag{4.2}$$



As we said, lanterns are distinguished by i, P_i is the intensity of lantern i, $r_{l;i}$ is its location, and ρ is the reflection efficiency of the road surface.

We notice that indeed sees is replaced by two 2-ary relations: it is written as the product of a first factor: a sum of terms that each take the relation between $r_{l;i}$ and r into account, and a second factor that takes the relation between r and r_d into account.

For the specularly reflective case[5], we cannot split the relation into two simple relations. We get one single, more complicated expression. It involves the location of the driver r_d, the locations of the contributing lanterns $r_{l;i}$ *and* the locations where specular reflection on the road occurs, r. We must summate over i and account for the location r_d in one expression; the two cannot be separated. The result is

$$l_p = \Sigma_{i:\text{contributing lanterns}} \frac{P_i \rho \; S(r - r_{l;i}, r - r_d, \beta)}{D_{\text{driver-lantern}}(r_d, r, r_{l;i})^2}. \tag{4.3}$$

Here, S and $D_{\text{driver-lantern}(r_d, r, r_{l;i})}$ are two quantities that have to be defined in two subsequent expansion steps.

Their meaning is: S is a factor that determines the amount of reflection given the directions of incoming $(r - r_{l;i})$ and outgoing $(r - r_d)$ light rays and the glossiness β ▷78 ; $D_{\text{driver-lantern}}(r_d, r, r_{l;i})$ is the length of the light path from lantern at $r_{l;i}$, via the reflecting location r, to the driver's location r_d, that is: $||r_{l;i} - r|| + ||r_d - r||$.

Clicking this link starts the ACCEL modeling environment with a script running that allows you to interactively experiment with various illumination conditions.

[5] The image of the specularly illuminated balls is taken from http://commons.wikimedia.org/wiki/Category:Phong_shading#mediaviewer/File:Phong_shading-balls.jpg

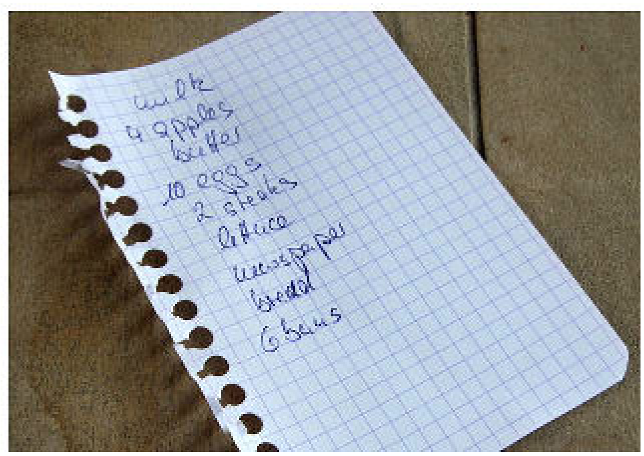
Apart from the mathematical details of the physics of illumination, this example illustrates some general aspects of translating qualitative relations into quantitative expressions:

- Assumptions are crucial in determining the arity of a relation. Assuming that the road acts as a diffuse reflector allows to calculate R_I independent of the location of the driver. This is simpler than having to take the driver's location into account: we can deal with a relation only involving driver and road, ignoring the relation between the road and the lanterns. In a later step, the relation between the lanterns and the road are studied.
- The assumption of a dry road gives conceptual simplicity, it also means that the *execution* of the model is more efficient. Indeed, values of R_I can be sampled in a sufficient number of values r and stored; they can be reused for every driver position. If the assumption of diffuse reflection does not hold, we have to recalculate, for every driver position, the contribution of all involved lanterns again.
- The difference between the two cases (diffuse, specular) is very large, both conceptually and in terms of performance. The modeler should therefore investigate whether simplifying assumptions are acceptable.
- Conversely, if the modeler can only work with a simple model, because of conceptual difficulties or because of performance issues, (s)he should ensure that the simplifying assumptions are communicated and validated. It is possible that the assumptions are so restrictive that the model's outcomes have nothing to say for the problem situation at hand. In that case the entire modeling effort is in vain.

Not Ready Yet

When running errands in a supermarket, we may be perplexed by the variety of offered goods, forgetting what we wanted to buy in the first place. An errand list aids our memory: every time we put an item in our cart, we scratch off the item from the list.

Building a formal model as a DAG is very similar, except that defining one quantity may need new quantities-to-be-defined, thereby temporarily extending our errand list.



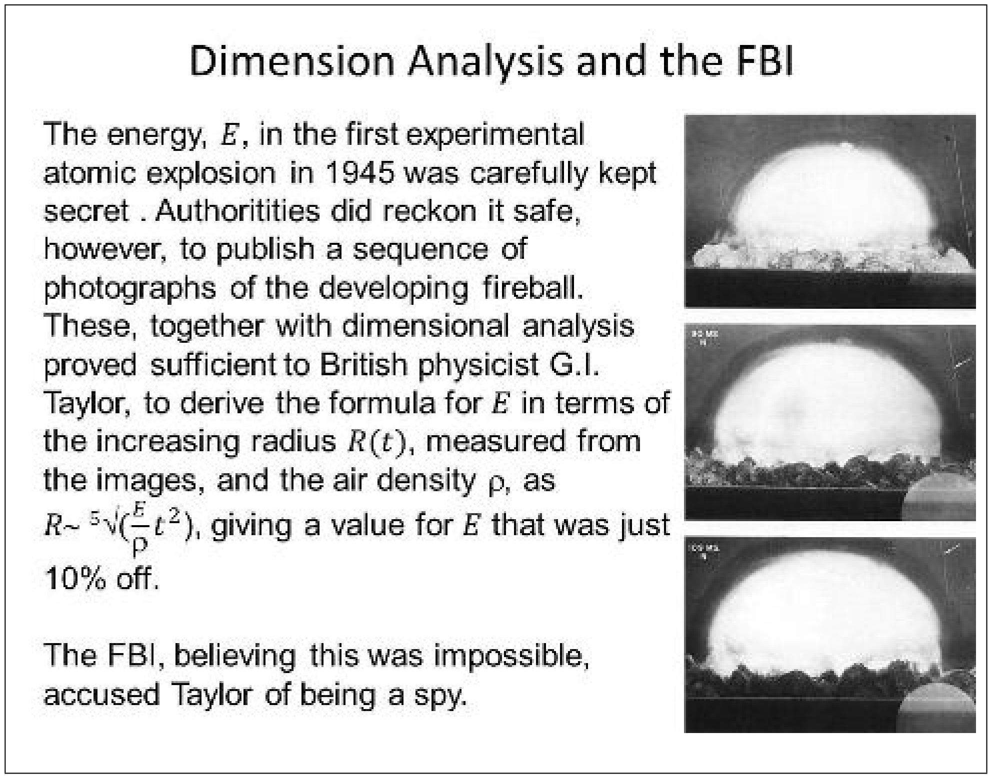
To-do-list

Expressing quantities in terms of other quantities can lead to many intermediate quantities[6]. The TO-DO-LIST is a simple device to keep track of intermediate quantities. It helps building the DAG in a systematic way. Further, it indicates the moment when a model is EXECUTABLE, that is: when all occurring values can be computed from known data.

[6]The photograph of an errand list was taken from http://www.rgbstock.nl/download/gesinek/nVgdnpm.jpg

4.2.3 Forming Mathematical Expressions

We discuss several ways to get mathematical expressions: dimensional analysis, the relation wizard and the function selector, and the heuristics of using multiple models. Furthermore, we give a method to find estimates for constants that are not readily available in published sources.



Dimensional Analysis

For expressions involving only additions, multiplications, subtractions, divisions, and tractional powers[7], dimensional analysis can be a powerful means to obtain formulas (see Section 2.7). This is even true for non-algebraic dependencies (such as sine or log-functions). These can only be applied to dimensionless arguments. From $y = \sin(x)$, dimensional analysis implies that both x and y have dimension 1.

Relation Wizard and Function Selector

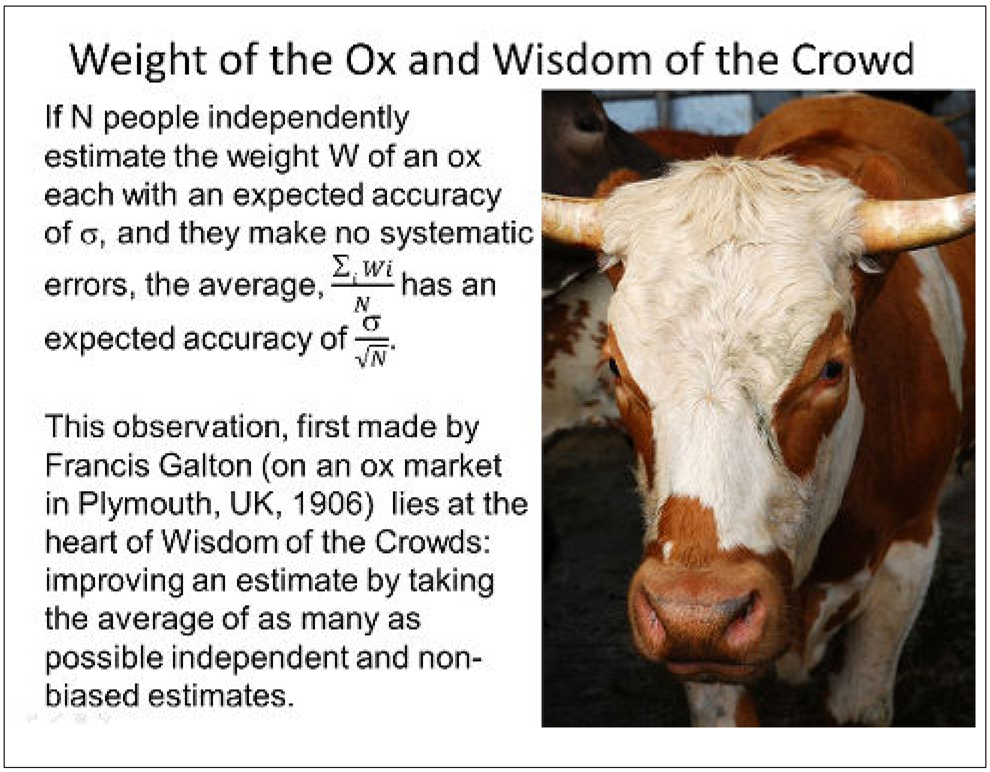
These are two interactive tools, to help identify useful material from a small repository of calculus results. Both are complementary to standard calculus texts. In standard calculus texts, techniques are explained, proven and trained in isolation. The complementary route starts from features from a problem or situation, phrased in non-mathematical terms, and next to find a possible formalization. The RELATION WIZARD(click here) starts from a qualitative relation and helps to find appropriate mathematical terminology using a number of question and answer steps; the FUNCTION SELECTOR(click here) uses the visual features of a graph or a data set as input and suggests a possible parameterization to reproduce these.

[7] The photographs of the Trinity experiment were taken from http://commons.wikimedia.org/wiki/Trinity_test#mediaviewer/File:Trinity_explosion_film_strip.jpg

Two models is better than one model

A model is never perfect, but it may be difficult to assess its reliability. An indication can be obtained by comparing the outcome of a model with the result of a second, independent model.

Wisdom of the Crowds



Quantitative models require values for quantities. Some result from empirical measurements[8], including an estimate of their accuracy. Internet is also a rich source when the value of a constant is sought.

Sometimes, however, we need to *guess* values. For instance for a quantity for which no reliable source can be found, and where empirical experiment is impractical. For instance, if quantities relate to demographic or psychological phenomena that many are familiar with, but that not (yet) have been formally measured. In these cases we may have to resort to mere guessing. Examples are: what is the fraction of families in Holland, possessing a trampoline (a hamster, an expresso machine, ...)

A simple yet powerful way both to improve the accuracy and to estimate the uncertainty of guesses is to use the WISDOM OF THE CROWDS. This method was first described by statistician FRANCIS GALTON. When members of a group of people make independent, unbiased estimates of some quantity the *average* of their guesses will be much closer to the actual value than the independent guesses. Further, the spread in their ratings gives an idea of the reliability of the average.

In the above examples, the method would amount to people in a group each guessing the desired number by writing down the first 20 families they think of, and next identifying which of these posses a trampoline (a hamster, an expresso machine, ...).

4.3 Examples

We illustrate the above elements with three examples.

[8]The photograph of an ox was taken from http://www.rgbstock.nl/photo/mjN1vGe/cow

4.3.1 Example 1: the Detergent Problem

Detergent Dump and Family Size

Is the amount of detergent, annually dumped in the environment indeed a linear function of the average family size?

Both are mere numbers. In order to assess their proportionality, both should be given e.g. in dependence of time. And even then it is impossible to say which depends on which. The function is a mere assumption of the modeller, made for some purpose.

To demonstrate the use of dimensional analysis, the to-do-list, and the wisdom of the crowds, we look at the question [▷79] :

'what is the total amount of detergent, annually dumped by Dutch households into the environment[9]'.

The model to answer this question is developed in Table 4.3.1. In this table, every row represents one quantity. It corresponds to one expansion step in the 'chain of dependencies', explained above. In such a step, a quantity is defined in terms of other quantities or values. The column 'Relations' gives this definition. The qualitative meaning of the quantity is given in parenthesis. The column 'Dimensions' gives the dimension of the quantities. The column 'Assumptions' lists the assumptions that underlay each expression.

Rows in tables such as Table 4.1 are added one by one. The first row corresponds to the quantity that represents the answer to the question our model attempts to solve. This is a root node in the DAG. Every time an expression introduces new quantities, these are added to the to-do-list. Every time we take a quantity from the to-do-list, this adds a row to the table.

The to-do-list starts with the quantity `amAnDetDmp`, the amount of annually dumped detergent in Holland. There is no other choice: it is the only thing our model should calculate. The first expression is a product of two quantities: `nrAnWshs` and `detPWsh`, for the number of annual washes in Holland, and the amount of detergent per wash, respectively.

This expression is not the only possibility. We could also express `amAnDetDmp` as the product of the number of months in a year and the average monthly dumped detergent, or the product of the average amount of dumped detergent per province per year multiplied by the number of provinces.

We are guided, however, by the unit of `amAnDetDmp`: [kg/year]. If we try a product expression, say, using quantities `x` and `y`, with units [dimX] and [dimY], we demand that [dimX]*[dimY]=[kg/year]. So a plausible choice is [dimX]=[kg/Δ] and [dimY]=[Δ/year] with some unit Δ.

We choose for Δ the unit 'wash'. This corresponds to our assumption, listed in the column 'assumptions', that detergent only ends up in the environment via washes. An expression involving

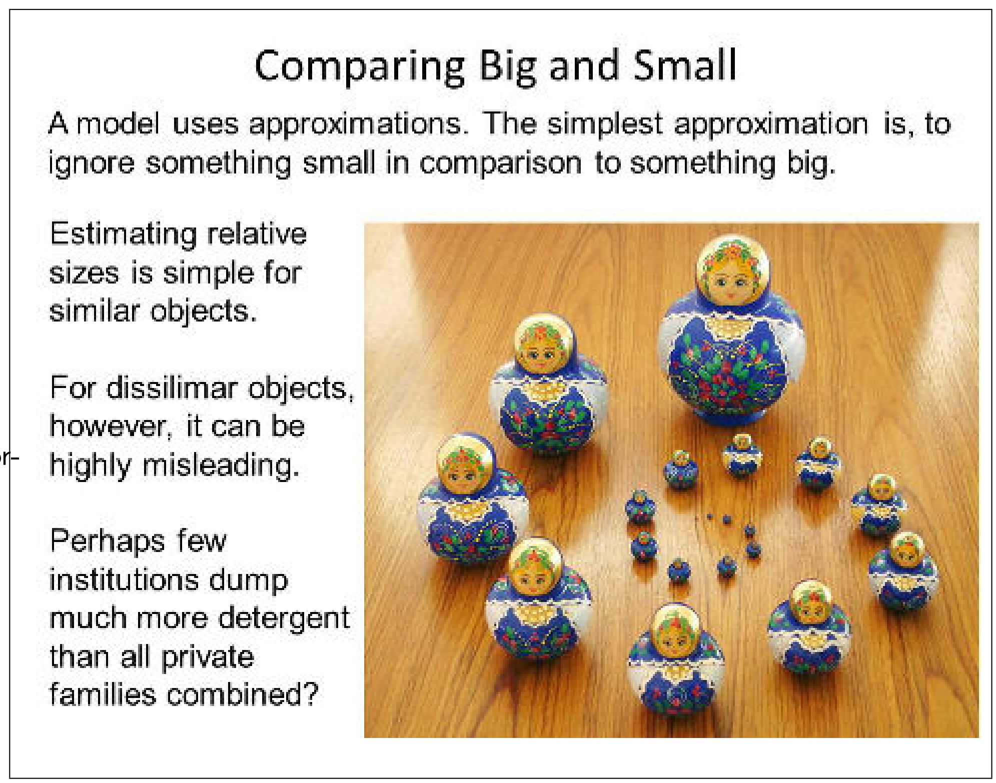
[9]The photograph of a family of swans is taken from http://www.rgbstock.nl/photo/oHwcMqA/de+zwanen

the number of washes is valid if we assume no spillage by detergent factories or otherwise. Another consequence of this choice is that we assume *all* detergent used in washes ends up in the environment. If we would doubt any of these assumptions, this would give additional quantities, for instance, `amAnDetDmp = (nrAnWshs - nrCarefulWshs) * detPWsh + amSpilledDet` where `nrCarefulWshs` and `amSpilledDet` stand for the number of washes per year without spilled detergent, and the amount of detergent spilled via other routes than washes. For now we choose not to have these mechanisms in ▷80 .

As stated in Section 2.7.2, anything can be a dimension, that is: an equivalence class on units. Here we use 'wash' as a unit. It is only equivalent to itself, so the unit 'wash' yields a dimension $\mathcal{WASH}$. Deriving formulas from *units* rather than from *dimensions* has two advantages: we ensure dimensional consistency, and different units with the same dimension are converted correctly.

The unit 'wash' is not a standardized physical unit. By treating it as a unit, though, instead of a dimensionless number, it will help deriving a consistent set of expressions. Suppose we would treat 'wash' as a dimensionless quantity, we could add it e.g. to a number of people per family. To protect us against such absurdities, we use a very stringent system of units, and stick to the algebraic rules to deal with them.

After proposing the formula `amAnDetDmp = nrAnWshs * detPWsh`, we eliminate `amAnDetDmp` from the to-do-list, and we add `nrAnWshs` and `detPWsh`. This completes the definition of `amAnDetDmp`. From now on, we can compute its value without bothering about its meaning.

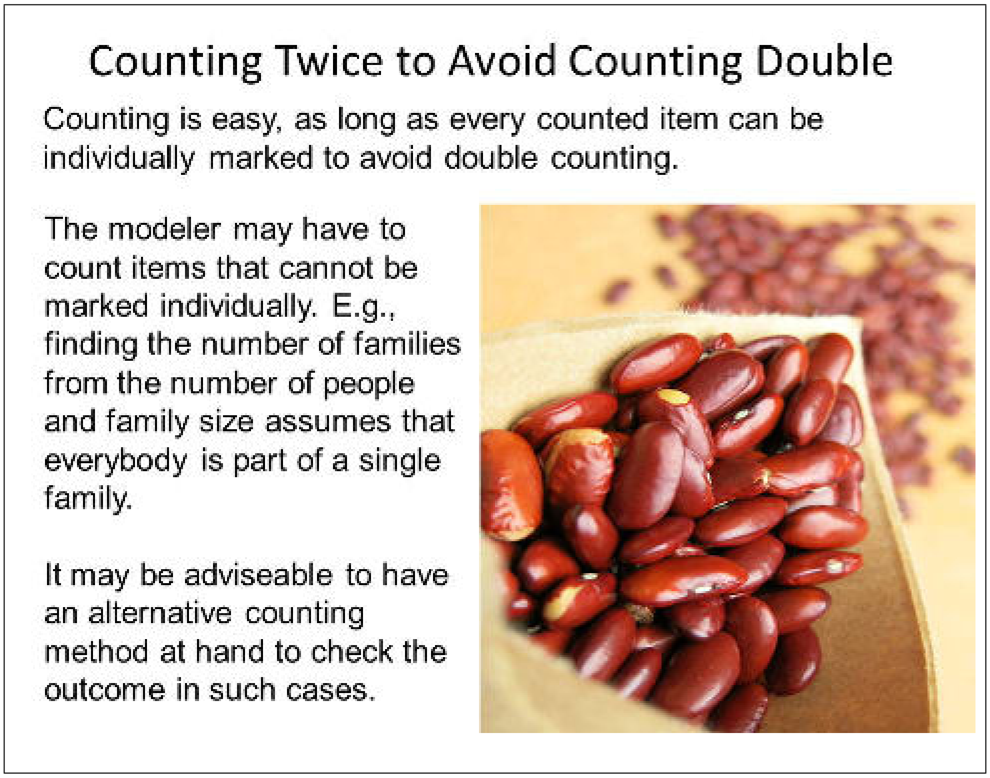


The next step (row 2 in the Table 4.3.1) arbitrarily chooses `nrAnWshs` from the to-do-list. Searching for quantities that determine its value, we need a mechanism where families, by doing washes, cause detergent to end up in the environment. So we introduce the concept *family*, again by expanding a unit [wash/year] as [wash/(fam*year)]*[fam]. Correct algebraic manipulation of the units 'wash', 'year' and 'fam' automatically yields the correct formula. Moreover, the formula introduces the assumption that we only consider family washes: we exclude institutional wash- and cleaning businesses[10],

[10]The photograph of big and small dolls was taken from http://www.rgbstock.nl/photo/2dR8Cnn/Russian+Dolls

etc., from our model.

Row 3 proceeds along the same scheme; again we picked an arbitrary quantity from the to-do-list. This time the assumption involved is that everybody belongs to exactly one family.



Next we pick `nrPIH` (=number of people in Holland)[11] from the to-do-list, and create a new row for it in Table 4.3.1. This is row nr. 4. The number of people in Holland has resulted from numerous historic and demographic mechanisms. For instance, it is the sum, over all provinces, of the number of people per province. We don't see that such an elaboration would help us, though ▷81 . Therefore, we consider it as a constant. It does not require an expression, so there are no new entries in the to-do-list. Its value can be assumed to belong to the 'common knowledge': most people, interested in this model, can be assumed to know the value 17± 0.5 million.

Row nr. 5 is devoted to `nrPPFam` (=the number of people per family). This is also a constant that does not depend on anything else in the model. The value can be found by consulting statistical yearbooks or other publicized sources. We work with a value of 2.2 ± 0.2.

This leaves two more entries in the to-do-list: `nrAnWshsPFam` (=the number of annual washes per family) and `detPWsh` (=the amount of detergent per wash). They correspond to rows nr. 6 and 7 in the table. Again they are considered constants: we don't want to elaborate if these quantities depend on anything else in our model. The values of `nrAnWshsPFam` and `detPWsh` don't belong to the common knowledge. It also might be difficult to find reliable values in publicized sources. Therefore we use the *wisdom of the crowds*, explained before. We find values `nrAnWshPFam`=100± 20 wash/year, and `detPWsh`=0.17 ±0.03 kg/wash.

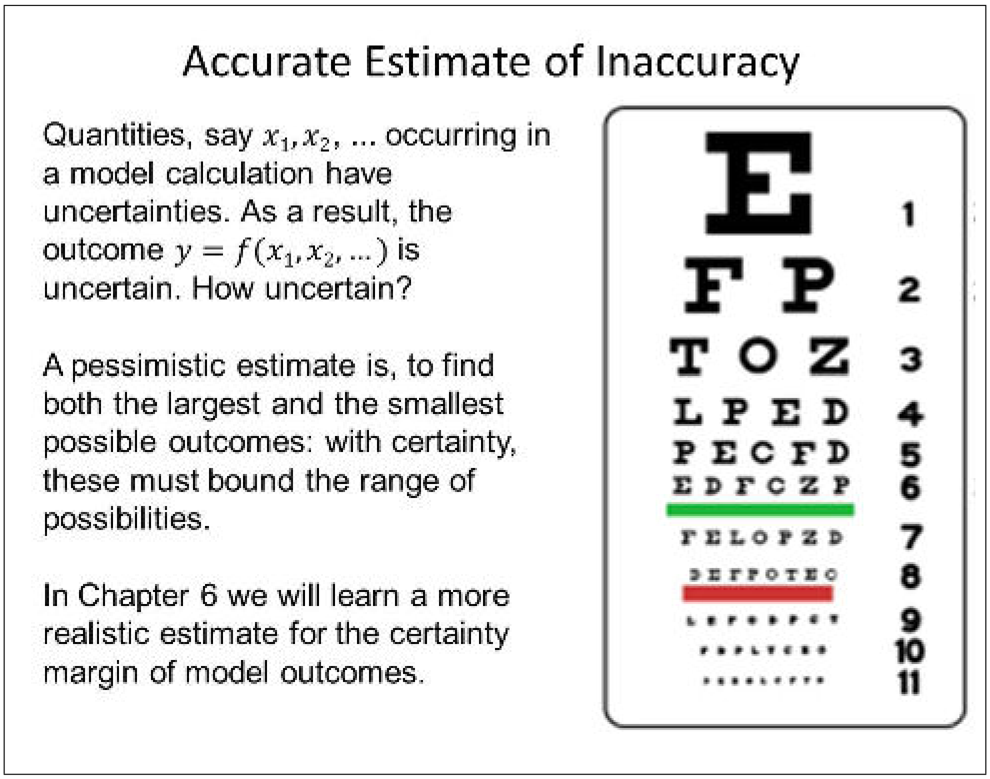
This completes the development of the model; we can now substitute the values of all the constants in the expressions, and find the answer of 130 million kilogram.

[11] The image 'counting beans' was taken from http://www.rgbstock.nl/photo/o2hoxsi/beans

Estimating Accuracy

Substituting all numerical estimates into the model gives a number of 130 million kilogram. To estimate the accuracy, we could search the most extreme values admitted by the intervals of the input quantities. That is, taking both `detPWsh`, `nrAnWshsPFam`, and `nrPIH` to their maximum and `nrPPFam` to its minimum (yielding 210 million kg); the opposite extreme is 77 million kg.

If we assume the uncertainties of the input quantities to be independent, though, the width of this range of $\{77 \cdots 210\}$million kg is too pessimistic.



Indeed, a more realistic estimation method[12] is explained in Section 6.4.2. In brief, it amounts to, first, obtaining the margins in each of the input quantities, and next calculate the *relative uncertainties*. We list these in Table 4.2.

In this example we use that the entire model consists of multiplications and divisions only. This means that for all the partial derivates, we need to compute according to Expression 6.18, we get $|\frac{\partial}{\partial x_i} y| = |\frac{y}{x_i}|$; $|\frac{x_i}{y} \frac{\partial}{\partial x_i} y|$ therefore equals 1 for all x_i and for y being the output quantity of our model. So we can approximate the relative accuracy $\frac{\Delta y}{y}$, again according to Expression 6.18, by $\sqrt{\Sigma_i \frac{\Delta x_i}{x_i}^2}$.

So, to estimate the combined contribution of the margins in Table 4.2, we take the square root of the sum of squares of the separate relative accuracies, yielding 0.3. We therefore claim that our result of 130 million kg/year is about 30% reliable. According to the website http://wiki.watmooi.nl/pages/Wassen_en_onderhoud and various other websites, the actual amount should be 150 million kg.

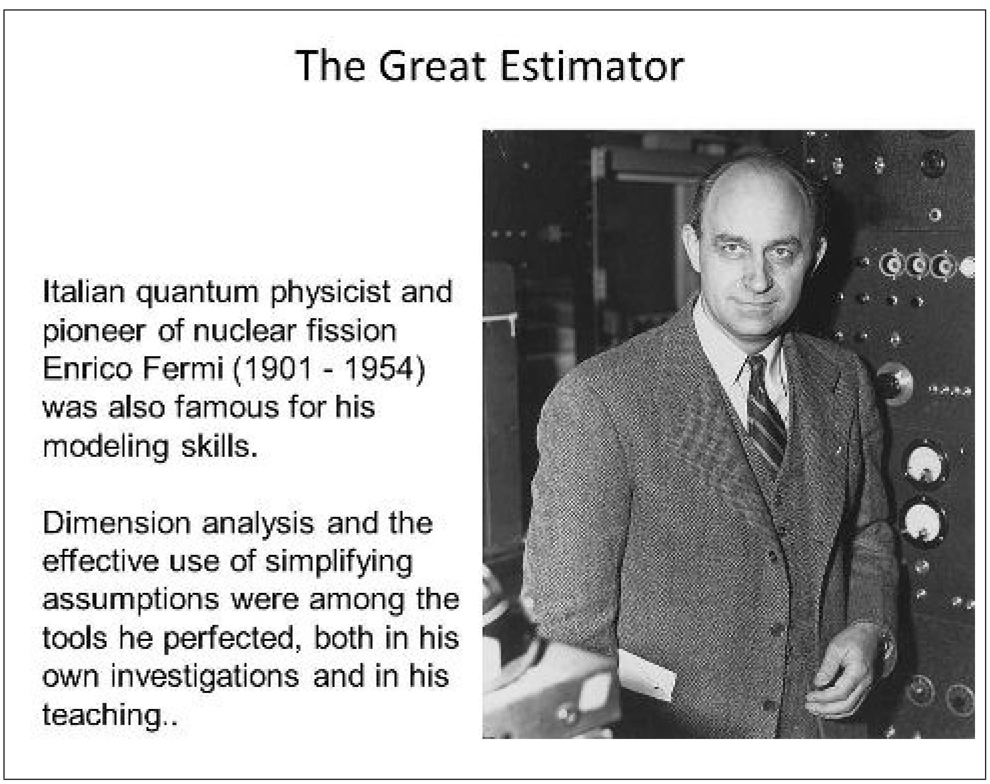
[12] The image of the Snellen chart is taken from http://upload.wikimedia.org/wikipedia/commons/e/e7/Snellen06.png?uselang=nl

	Relations	Dimensions	Assumptions
1	`amAnDetDmp = nrAnWshs * detPWsh` (amount of annually dumped detergent)	[kg / year] = [wash / year] * [kg / wash]	Washing laundry is the only way detergent gets into the environment
2	`nrAnWshs = nrAnWshsPFam * nrFamIH` (annual number of washes)	[wash / year] = [wash / (fam * year)] * [fam]	No institutional laundry washing, only families
3	`nrFamIH = nrPIH / nrPPFam` (number of families in Holland)	[fam] = [people] / [people / fam]	families are disjoint: everybody belongs to exactly one family
4	`nrPIH` = 17 $\pm$ 0.5 million (number of people in Holland)	[people]	common knowledge
5	`nrPPFam` = 2.2 $\pm$ 0.2 (number of people per family)	[people / fam]	public domain
6	`nrAnWshsPFam` = 100 $\pm$ 20 (number of annual washes per family)	[wash / year]	wisdom of the crowds
7	`detPWsh` = 0.17 $\pm$ 0.03 (amount of detergent per wash)	[kg / wash]	wisdom of the crowds

Table 4.1: How much detergent do Dutch households annually dump into the environment?

quantity	value	margin	relative uncertainty
`detPWsh`	0.17	0.03	0.2
`nrAnWshPFam`	100	20	0.2
`nrPIH`	17	0.5	0.03
`nrPPFam`	2.2	0.2	0.1

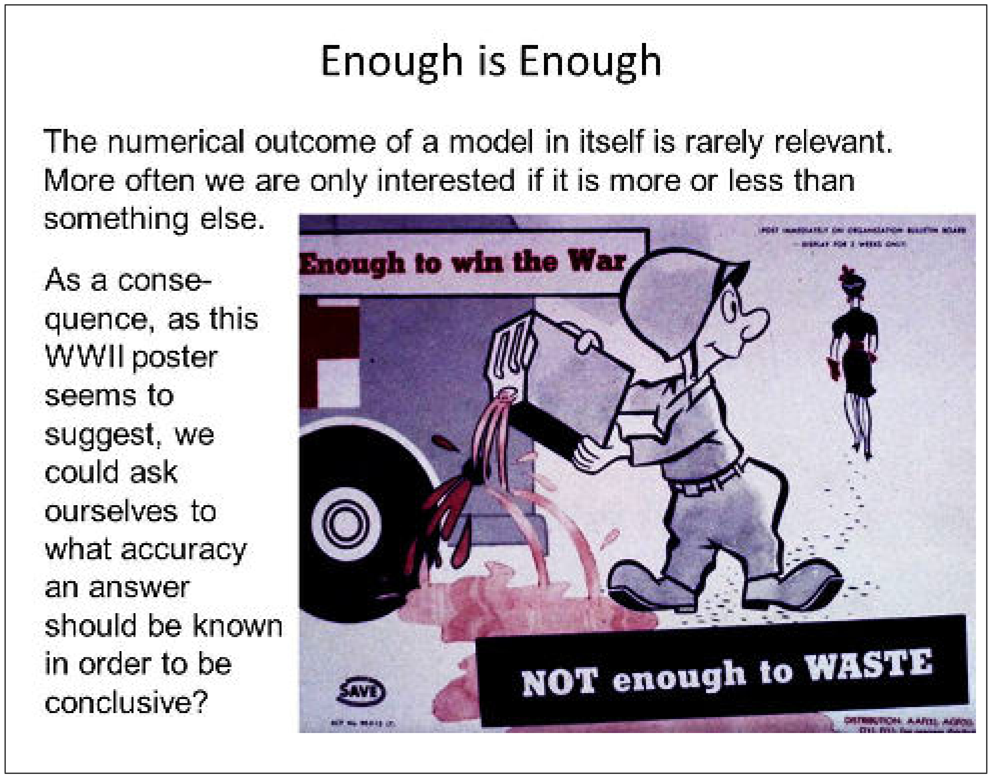
Table 4.2: Uncertainties in the final result due to uncertainties in each of the input quantities



4.3.2 Example 2: The Chimney Sweepers Problem

Enrico Fermi[13], renowned physics Nobel laureate and pioneer of nuclear fission, was also a proficient modeler.

He is known to have estimated, within 50% accuracy, the explosion force of the first atomic bomb by observing the effect its shockwave had on a handful of falling confetti. For didactic purposes, he invented the problem to estimate the number of piano tuners in Chicago using nothing but common knowledge and simple arithmetic reasoning.



He used the same method to estimate the number of planets in the universe carrying intelligent life, an estimate still called the Fermi paradox: why, since there should be so many aliens, have we never seen one? As an hommage to Enrico Fermi, we now focus on the question 'how many chimney sweepers work in Eindhoven'.

The Consequence of Purpose for the Applicability of a Model

Before developing the model, we must first formulate its purpose[14]. We will pick three:

- *1.* To assess if there are not more than 50, so that the next annual chimney sweepers convention can be held in Restaurant 'the Swinging Sweeper', having a maximum capacity of 50 guests. Our model should produce an upper bound; underestimating is not problematic.
- *2.* To see if there are at least 200, so that it is financially feasible to begin a special interest group magazine 'Sweeper's Bazaar'. The model should produce a lower bound; overestimation is not problematic.
- *3.* To see if there are about the same number of chimney sweepers as sewage cleaners (23) to decide whether we could form couples each consisting of one sweepers and one cleaner, offering

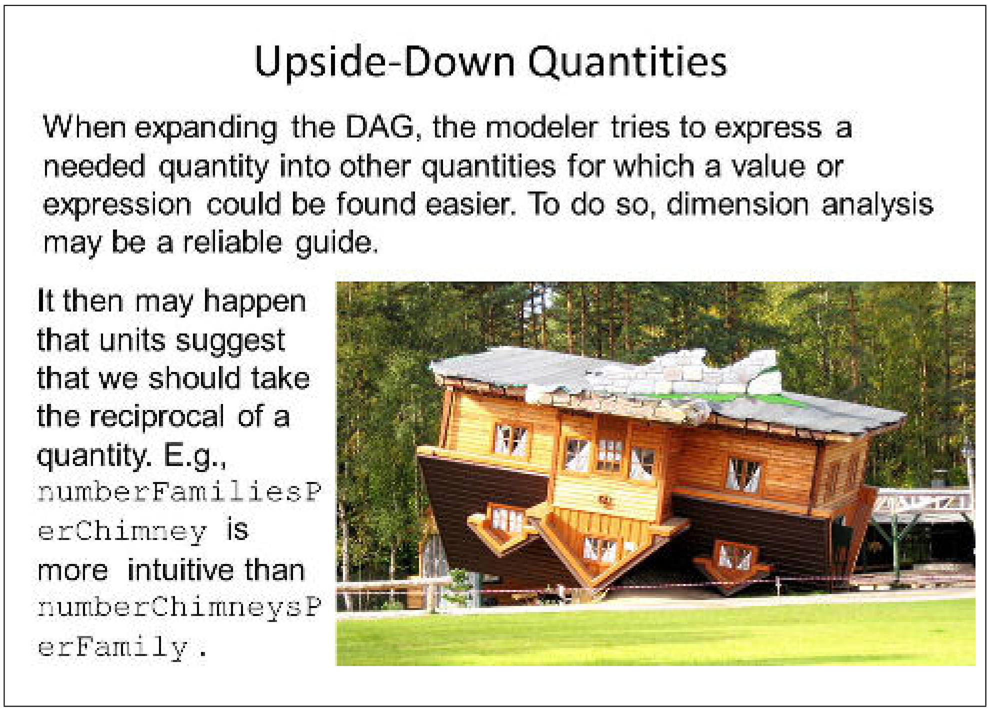
[13] The Enrico Fermi photograph was taken from http://commons.wikimedia.org/wiki/Enrico_Fermi#mediaviewer/File:Enrico_Fermi_1943-49.jpg

[14] The wartime poster image, inviting to reflect on when enough is enough, is taken from http://commons.wikimedia.org/wiki/File:%22Enough_to_Win_the_War,_Not_Enough_to_Waste%22_-_NARA_-_514215.tif?uselang=nl

a new form of professional cleaning services. Now the outcome needs both an upper- and a lower bound.

The model is elaborated in Table 4.3.

We explain non-trivial steps only.



Line 1. We use the label 'E' in the expressions for units (referring to 'Eindhoven') as if it is a unit. We call this a PSEUDO UNIT. 'Sw' is a unit to be used to indicate a number of sweepers; 'E' indicates that we restrict the number of sweepers to Eindhoven. 'E' should occur in the denominator of only one of the terms `nrChIE` (number of chimneys in Eindhoven) and `nrSwPCh` (number of sweepers per chimney). This may seem to be wrong, since both quantities depend on the circumstance that we deal with Eindhoven. Only one of them, `nrChIE`, however, is proportional to some property of Eindhoven. For instance, the area of Eindhoven, or the number of houses in Eindhoven. For this reason we have unit [Ch/A] for `nrChIE`, and we don't have 'E' in the denominator of the unit of `nrSwPCh`.

Line 2. The same argument explains why 'E' occurs in the unit for `nrFamIE` (number of families in Eindhoven) and not in the unit for `nrChPFam` (number of chimneys per family). Even though the number of chimneys per family in Eindhoven may be different from the number of chimneys per family elsewhere: this dependence is not a proportionality.

Line 6. 'The number of chimneys per family' (`nrChPFam`) is somewhat counterintuitive[15]. We may express it as 1/(the number of families per chimney), with a consistently adjusted unit. Next we can relate this number, for instance, to the ratio of the number of family houses and the number of apartment houses.

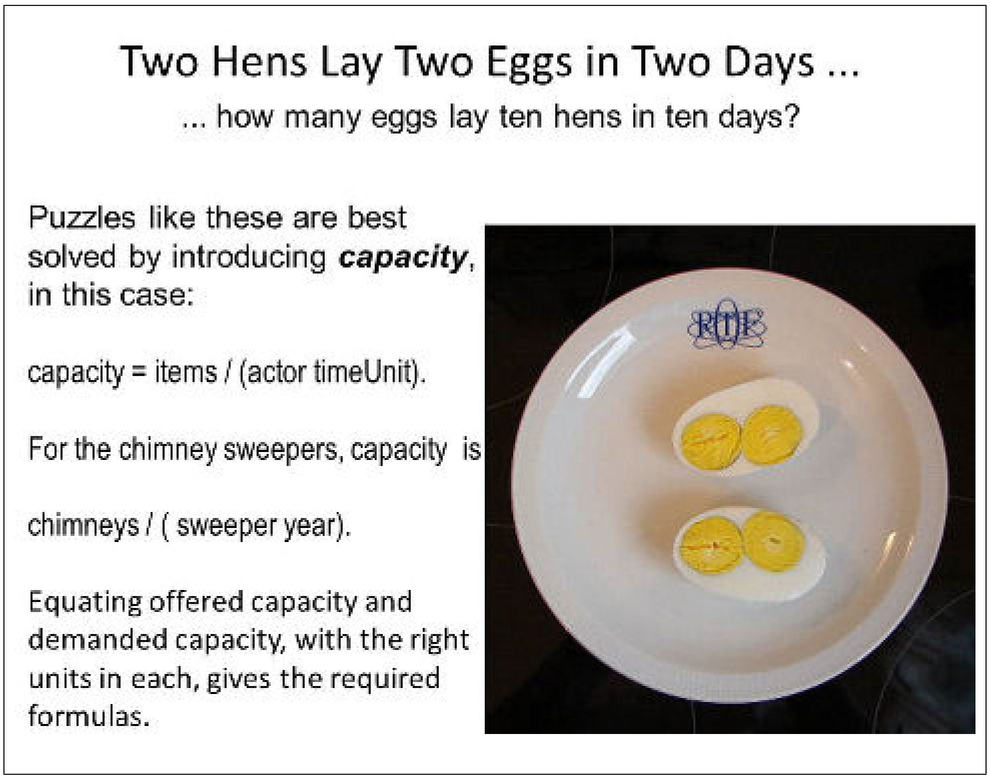
Line 7. To find an expression for `nrSwPCh`, we ask: 'what links the nr of sweepers to the number of chimneys?'. The answer is: sweepers *service* chimneys. 'How many services' relates to the number of available sweepers and to what a single sweeper can do per unit of time. At the same time, 'how many services' relates to the amount of service-time a chimney needs.

We assume that sweepers have no specialized roles in the process. That is: 'sweeping' is a process that can be distributed over sweepers: 'if 5 chimney sweeper sweep 20 chimneys per day, 4 chimney

[15] The photograph of the upside-down house is taken from http://commons.wikimedia.org/wiki/File:An_%27upside-down_house%27_in_open-air_museum,_Szybmark,_Poland..jpg?uselang=nl

sweepers sweep 16 chimneys per day'. Chimney sweepers behave different from musicians in a string quartet: if it takes 4 violinists to play a sonata in 10 minutes, 8 violinists can't play the same sonate in 5 minutes. The efforts of musicians can not simply be added and exchanged.

The homogeneity of a task such as chimney sweeping permits to talk about CAPACITY as a quantity proportional to the amount of resources (=chimney sweepers) and to the amount of time. An example is 'man hours' as unit for the capacity of human labor. The capacity of a sweeper is proportional to the time he works ([Sw * year]). The 'need' of a chimney is expressed in ([Ch * year]). The unit 'Se' is an abbreviation of 'service', so `nrSwPSe` is the number of sweepers per service.



Line 8. Notice that we have two units (hour and year)[16] . These both have the dimension of time. Distinguishing the units makes the analysis more precise. It ensures that conversion factors, such as the number of hours per year, end up in the right place in formulas.

Filling in all values gives a number of 19 chimney sweepers. We use the same method to estimate the accuracy of the outcome as in the previous example. Relative accuracies are listed in Table 4.4.

The square root of the sum of squares of these relative uncertainties is 0.27, so our final answer is 19$\pm$ 5. Compare this to the number of chimney-sweeping firms, found in the 2012 yellow page

[16]The photograph of the one (or two?) egg(s) is taken from http://commons.wikimedia.org/wiki/Category:Chicken_eggs#mediaviewer/File:2010_double_eggs_twins_002.JPG

	Relations	Dimensions	Assumptions
1	`nrChSwIE =nrChIE * nrSwPCh` (number of chimney sweepers in Eindhoven)	[Sw / A] = [Ch / A] * [SW / Ch]	Eindhoven ch.-sweepers sweep Eindhoven chimneys only
2	`nrChIE = nrChPFam * nrFamIE` (number of chimneys in Eindhoven)	[Ch / A] = [Ch / Fam] * [Fam / A]	ch.-sweepers sweep only chimneys on family houses
3	`nrFamIE = nrPIE / nrPPFam` (number of families in Eindhoven)	[Fam / A] = [P / A] / [P / Fam]	families are disjoint: everybody belongs to one family
4	`nrPI` = 250000 ± 5000 (number of people in Eindhoven)	[P]	common knowledge
5	`nrPPFam` = 2.2 ± 0.2 (number of people per family)	[P / Fam]	public domain
6	`nrChPFam` (= 1/`nrFamPCh`) =0.1 ± 0.02 (number of chimneys per family)	[Ch / Fam]	wisdom of the crowds
7	`nrSwPCh = nrSwPSe * nrSePCh` (number of sweepers per chimney)	[Sw / Ch] = [Sw * year / Se] * [Se / (Ch * year)]	Introduce time to associate sweepers capacity to chimneys need
8	`nrSwPSe = timeP1Se / timeP1Sw` (number of sweepers per service)	[Sw * year / Se] = [hour / Se] / [hour / (Sw * year)]	assume same time for every service
10	`timeP1Se` = 2 ± 0.25 (needed time per service)	[hour /Se]	wisdom of the crowds
11	`timeP1Sw` = 1200±100 (available time per sweeper per year)	[hour /(Sw * year)]	sweepers also do administration etc.
12	`nrSePCh` = 1 (nr services per year)	[Se / (Ch * year)]	insurance require chimneys to be swept once per year

Table 4.3: How many chimney sweepers work in Eindhoven?

quantity	value	margin	relative uncertainty
`nrPIE`	250000	5000	0.006
`nrPPFam`	2.2	0.2	0.1
`nrChPFam`	0.1	0.02	0.2
`timeP1Se`	2	0.25	0.12
`timeP1Sw`	1200	100	0.08
`nrSePCh`	1	0	0

Table 4.4: Uncertainties in the final result due to uncertainties in each of the input quantities

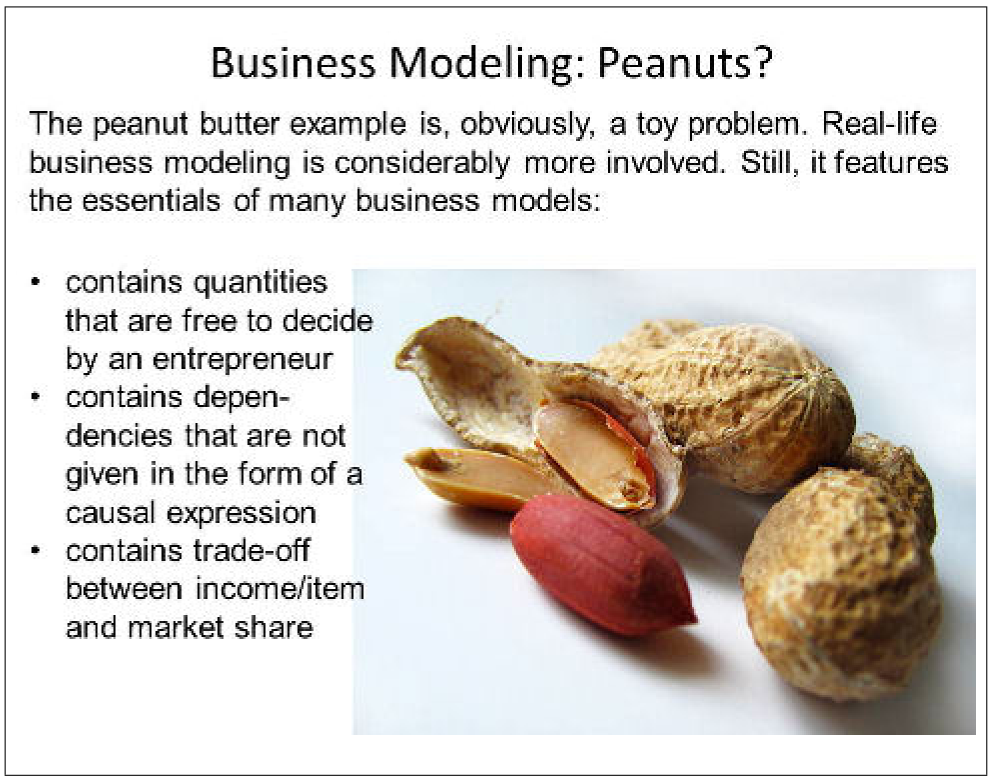
dictionary of Eindhoven: there are 7 companies listed. If, on average, these employ 3 chimney sweepers each our results are quite close.

Now we go back to our purposes. The result 19± 5 is sufficient to fulfill purposes 2 and 3; it is not accurate enough to fulfill purpose 1.

The Benefits of a Second Model

To further confirm our model, we may develop a second model, possibly independent from the first one. We could, for instance:

- Collect data from the Eindhoven garbage collection services, and find the amount of disposed ashes and soot. Assuming that all ashes and soot come from chimney sweeping, and if estimating the amount of ashes and soot a chimney sweeper collects per service, we can deduce the number of chimney sweepers.
- Using the wisdom of the crowds, we may estimate how often we see a chimney sweeper in Eindhoven per time unit. If we also estimate how many other professionals we see in the same period in the same place, we can find out the ratio of chimney sweepers to the total number of Eindhoven professionals that occur in the streets.
- Using again the wisdom of the crowds, we may find out what sweeping a chimney on average costs. If we relate this to a plausible annual salary for a chimney sweeper, and to the amount of money paid for sweeping a chimney, we again can do the calculation ▷82 .



4.3.3 Example 3: the Peanut Butter Problem

All mathematical operations in the previous two models[17] were multiplications and divisions. They all resulted from reasoning with units or dimensions. Dimension analysis would have also helped in case of additions and subtractions. Indeed, in $a = b + c$, all three quantities must have the same unit.

[17] The photograph of peanuts is taken from http://www.rgbstock.nl/download/michaelaw/mGIoDb4.jpg

Not all models, however, consist of additions, multiplication, divisions or subtractions only. Other functions are needed to express the dependency of one quantity on another.

To illustrate the development of a more sophisticated formal model, we look at the problem of making profit selling a new brand of peanut butter. We try to find a selling price so as to earn as much as possible. We see an immediate dilemma: if the selling price per item is too low, we may sell a large volume but the profit may be low; if the selling price per item is too high we attract too few customers, and the profit will be low as well.

Analyzing 'what depends on what' reveals some ambiguity. One view is to regard the ideal price per item[18] as the quantity we are most interested in. So we should write it as a function, depending on other quantities until we have quantities with known values so that everything can be computed. But it is not clear what the price per item depends on, other than the desire to become as rich as possible.

It seems simpler to regard the profit as a quantity that depends on something else. Profit depends on the number of customers, but also on the selling price per item. We don't know the price per item, though, and we see no route to find it.

At this point we are stuck again. We may get help from the Relation Wizard.

The Relation Wizard

The Relation Wizard is an organized list of questions and answers. It resembles a flora used by botanists. A flora helps determining a plant by asking questions on the visible features of the plant to be determined. All questions are multiple choice questions and each possible answer leads to a following question. At the end the identity of the plant is found.

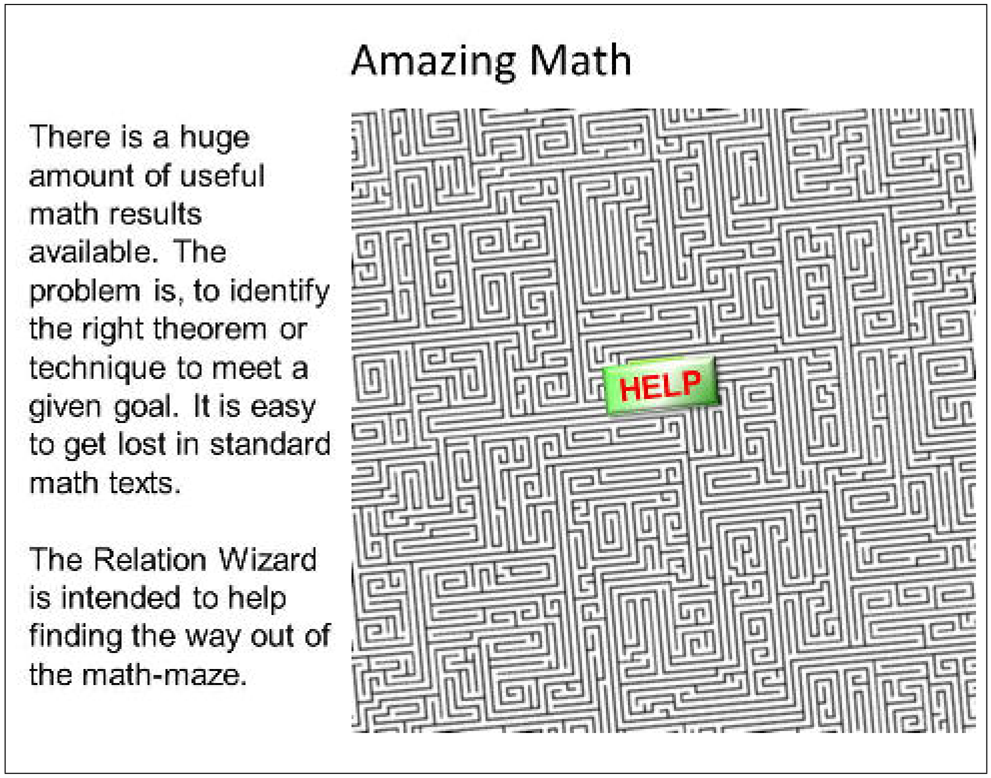
The Relation Wizard helps to identify mathematical notions that may be needed when modeling. It assumes that one has a *relation* in mind. This relation was formulated in the conceptual model. Here we are looking for the relation between profit and price per item. The Relation Wizard does

[18]The photograph of the florist's shop istaken from http://www.rgbstock.nl/download/sundstrom/miftZPW.jpg

suggestions as to what kind of mathematical notion might translate this relation into things that can be computed.

This contrasts with most math textbooks, where the order of presentation is inspired by mathematical structure. In the Relation Wizard, the order comes from non-mathematical features of the modeled system at hand, and it will eventually suggest fragments of mathematical equipment.

The first question of the Relation Wizard[19] is: 'does probability play an important role in the relation? → stochastic' and the alternative: 'if not: → deterministic'. In the peanut butter case, perhaps stochastics may play a role, but we think that by setting a price, we determine the resulting profit. So we choose 'deterministic'.



Next question: 'is one quantity given in dependency of the other(s)? → 'functional', or 'if not: → non-functional'. If we were clear about what would depend on what, we would not have consulted the Relation Wizard in the first place, so we choose 'non-functional'.

In the section labeled 'non-functional', one of the options is 'should we find a value of x such that some y is minimal or maximal? → optimality'. This is exactly the situation at hand. Our problem is a problem of optimality. The section on 'optimality' states our case: 'There is a recipe to find y from x, y=f(x); we need to know x such that y is optimal (minimal or maximal), often subject to additional conditions.'.

According to the Relation Wizard, we should express the profit (=the quantity we want to be optimal) in terms of quantities that it depends on. This is very similar to the earlier two models, so we now know what to do.

We set up the model analogously to Sections 4.3.1 and 4.3.2, building the Table 4.5 as we go. Everything runs smoothly until we arrive at line 6.

At this point, we get stuck again: we need to express *how* `marketShare` depends on `pricePerItem`, so we need a function. But we cannot deduce the form of the function from dimension analysis.

[19]The image of the maze is taken from http://www.rgbstock.nl/download/micromoth/ouH4Wey.jpg

	Relations	Dimensions	Assumptions
1	`profit=income-expenses` (profit from selling peanut butter)	[Euro / year]	no taxes, no inflation
2	`income = pricePerItem * nrSoldItems` (income by selling peanut butter)	[Euro / year] = [Euro/ourPB] * [ourPB/year]	no discount with larger quantities per purchase; the price per items stays constant over an entire year
3	`nrSoldItems = nrSoldTotal * marketShare` (number of sold items in a year)	[ourPB / year] = [allPB / year] * [ourPB/allPB]	our peanut butter will not affect the total market, that is: our peanut butter will only attract people that already decided to purchase peanut butter - who now will choose for our peanut butter rather than the competitor's product
4	`nrSoldTotal` $=\cdots$ (the total number of sold peanut butter items per year)	[allPB / year]	this amount can be found by consulting a trustworthy marketing bureau
5	`pricePerItem` $=\cdots$ (price to be charged for our peanut butter per unit)	[Euro/ourPB]	free to choose
6	`marketShare` $=\cdots$ (fraction of all sold units of peanut butter that will be sold by me)	[ourPB/allPB]	this depends on `pricePerItem`

Table 4.5: How to get rich with peanut butter?

We again consult the Relation Wizard. We realize that the functional dependency we are looking for is a *numerical dependency* (both `marketShare` and `pricePerItem` are numeric quantities), and in a number of steps (via MONOTONOUS , SMOOTH, NON-RATIONAL , and 'other') we find, as a suggestion, the *logistic function*. The logistic function is explained below.

There is another approach[20], though. Functions have two aspects. First, there is a mathematical expression defining y in terms of x, involving operations on x. Second, there is a

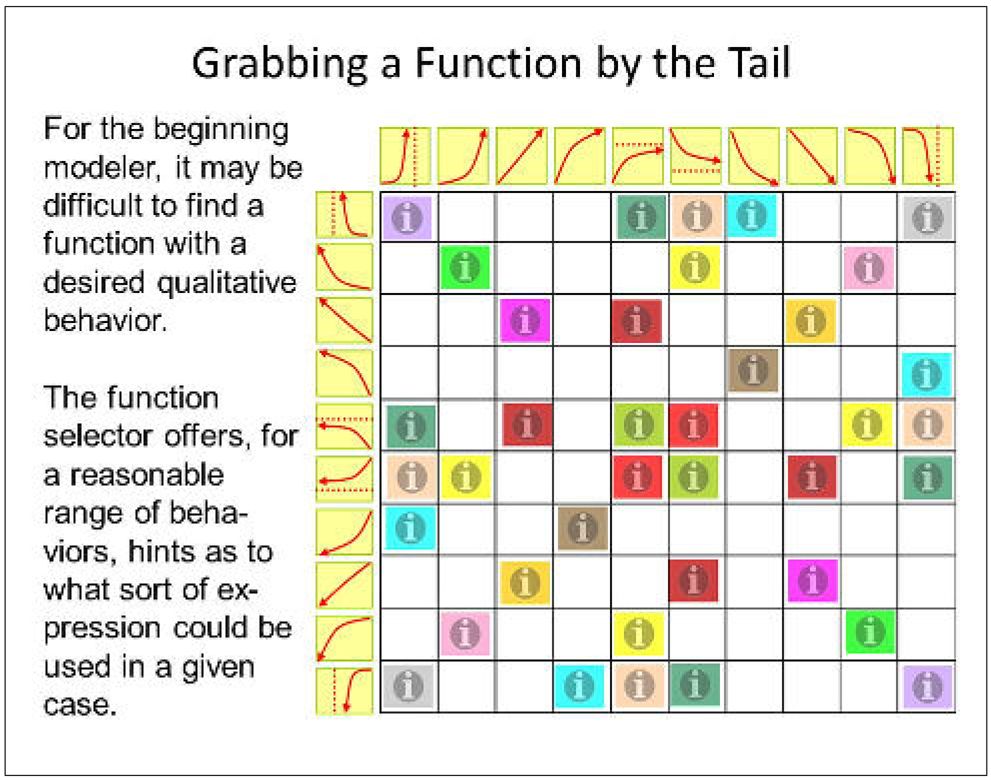
[20]The photograph of the multi-slides is taken from http://upload.wikimedia.org/wikipedia/commons/9/96/Multi_slides.jpg?uselang=nl

graph with features such as asymptotes, inflection points and AXIS INTERCEPTS.

Finding the graph from the function expression is simple: merely calculate a sufficient number of pairs $[x_i, f(x_i)]$ and plot these in the $x-y$-plane. Typically we do some analysis first. Does the graph have *asymptotes*, does it have $x-$axis intercepts or does it have local minima or maxima? All these tests are, in most cases, straightforward. The opposite route, where a graph is *given* and we need a function expression, is much more challenging. This is where the Function Selector comes in.

Function Selector

When we need a mathematical expression that is a function, we may have intuition of how the function should look like. This may happen if the function should represent empirically obtained data. The data points seem to indicate the shape of the graph. It is as if the graph should pass through the points.



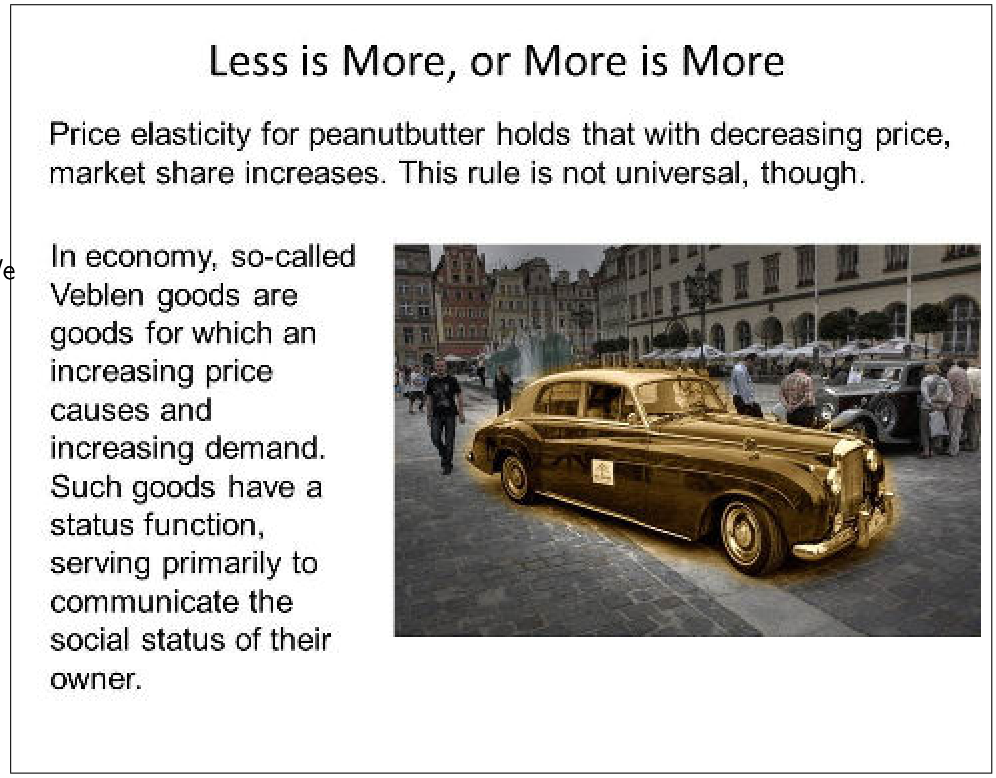
In the present situation, where we try to find a function that tells how `marketShare` depends on `pricePerIem`, we don't have any data, but we nevertheless have quite some knowledge about the shape of the function at hand:

- The function will be decreasing: if the price per item is higher, fewer people will choose our peanut butter;
- The function will be monotonically decreasing: we assume there won't be price ranges where the attractiveness *increases* with the price per item;
- The function does not have to be smooth. Nevertheless, for simplicity, we may choose a smooth function. The absence of jumps in the graph has many advantages, for instance for numerical methods to solve the optimization problem;
- If the price per item increases, eventually `marketShare` will become 0. We say that the behavior SATURATES, that is: it will never exceed a certain value, no matter how large or how small the argument becomes. In this case, the rightmost saturation value is 0;
- If the price per item decreases, perhaps even assuming negative values, eventually we will attract

100% of the market. So `marketShare` as function of `pricePerItem` has a leftmost saturation value of 1 ▷83 ;

- From the above requirements it follows that the range of the function is between 0 and 1.

We know quite a bit about this so called PRICE ELASTICITY(=the relation between market share and the selling price per item), but we don't know what function expression should be used to express `fractionFormMe(pricePerItem)`. We don't even know if there is *a* single correct mathematical expression[21].



To pick a *plausible* function that meets with intuitive behavior, the Function Selector takes 'both ends' of a behavior as input, and suggests one or more expressions that exhibit this behavior. 'Both ends' here means: both for increasing and decreasing argument. The Function Selector helps to find ▷84 a function $y = f(x)$.

For `marketShare`$=f_{\mathsf{MarketShare}}$(`pricePerItem`) the Function Selector suggests three possibilities: a function consisting of three linear segments (the so-called *ramp* function), an arctan function, and a so-called *logistic* function.

In modeling, RATIONAL and exponential or logarithmic functions seem to be more preferred than (arc)tangents. This may relate to the ease of calculation of rational functions (involving only additions, subtractions, multiplications and divisions), and the close relation of exponentials and logarithms to the 'natural' operations, addition and multiplication: $e^{x_1+x_2} = e^{x_1}e^{x_2}$ and $\ln(x_1x_2) = \ln(x_1) + \ln(x_2)$. There are no similarly simple expressions involving (arc)tangent functions.

For the sequel we will only discuss the ramp function and the logistic function.

[21] The image of a Rolls Royce was taken from http://commons.wikimedia.org/wiki/File:Rolls-Royce_(3661656309).jpg?uselang=nl

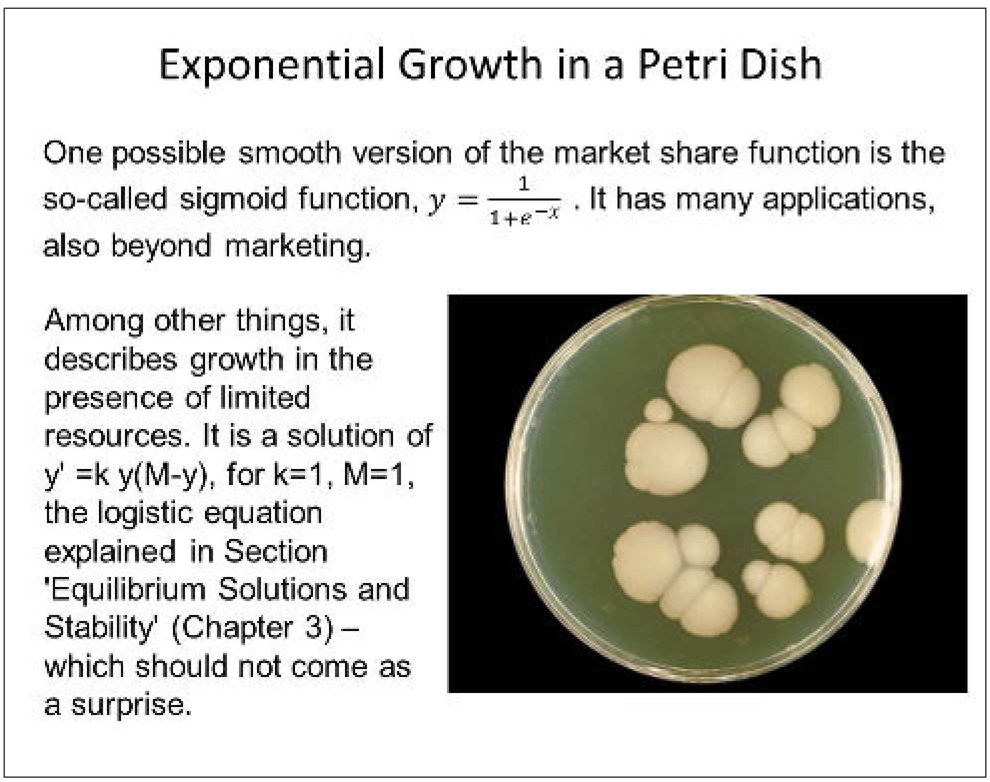
The Ramp Function A so-called RAMP function $y = f_{ramp}(x)$ consists of two piecewise constant segments and a linear segment between them. It is parameterized by 4 quantities

$$
\begin{aligned}
y &= f_{ramp}(x, x_0, x_1, y_0, y_1) && (4.4)\\
&= y_0 \quad \text{if} \quad x < x_0 \\
&= y_1 \quad \text{if} \quad x > x_1 \\
&= y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0) \ \text{ else,} && (4.5)
\end{aligned}
$$

assuming that $x_0 \neq x_1$.

Choosing this function asks for an interpretation of the quantities ▷85 $x_0 \cdots y_1$. For y_0 and y_1, obvious values are 1 and 0, respectively.

From dimension analysis, the quantities x_0 and x_1 must be prices per item, as they are compared against x which is a price per item. A plausible interpretation is, that they are the prices of the cheapest and most expensive competing peanut butter. Then the ramp function formalizes the intuition that nobody will choose our peanut butter if we are more expensive than the most expensive competitor, and everybody will choose our peanut butter if we are cheaper than the cheapest competitor.



This is a simplistic assumption indeed, ignoring all other factors that affect a customer's choice. But it shows the connection between the mathematical expression of a function and its interpretation.

The Logistic Function A so-called LOGISTIC function, also called SIGMOID function[22], unlike the ramp function, is smooth. The logistic function approximates a linear behavior for a limited part of the domain, but it saturates if arguments get very large or very small.

The simplest expression is $y = f_{logistic}(x) = \frac{1}{1+e^{-x}}$. We see that this has the right range of values: y ranges from 0 to 1. It takes its middle value (=0.5) for $x = 0$, though, and we have no control over the slope.

[22]The photograph from the petri dish was taken from http://commons.wikimedia.org/wiki/Petri_dish#mediaviewer/File:Candida_albicans_PHIL_3192_lores.jpg

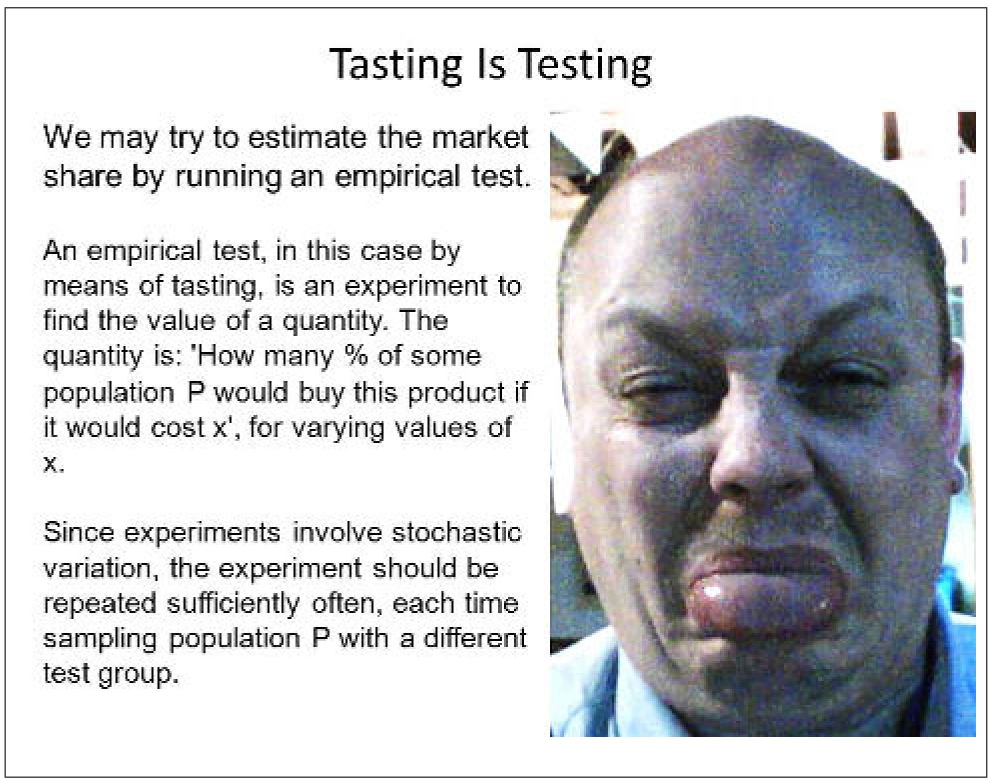
This occurs often: we have found a functional expression that more or less has the right behavior, but we need to 'tweak' it in order to make it do what we want. A modeler should develop some skills in tweaking the expressions for functions. 'Tweaking' can amount to shifting or scaling the argument, shifting or scaling the range, or more advanced transformations. In some cases, (s)he might benefit from existing work. For instance, for the logistic function, a 'tweaked' version exists, called the RICHARDS CURVE:

$$y = f_{Richards}(x, K, A, Q, B, M, \nu) = A + \frac{K - A}{(1 + Qe^{-B(x-M)})^{1/\nu}}, \tag{4.6}$$

where

- A: the lower asymptote;
- K: the upper asymptote;
- B: the rate of increase;
- $\nu > 0$: affects near which asymptote maximum growth occurs;
- Q: depends on the value y(0);
- M: the x-value of maximum growth if $Q{=}\nu$.

This example illustrates a trade-off between simplicity and accuracy. Indeed, although Richards curve can be tuned to take many logistic shapes, it is not trivial to find meaningful interpretations of the quantities A … M in the peanut butter case.



Glass Box Models, Black Box Models

The ramp function has a parameterization that can be related to causal mechanism. A prospect buyer[23], having to decide if (s)he would buy our peanut butter for `pricePerItem` might make a decision by hypothetically selecting a random number between the minimum and maximum price of competing peanut butters. If our `pricePerItem` is lower than the random number, (s)he buys our peanut butter.

[23] The photograph of a prospect non-buyer is taken from http://commons.wikimedia.org/wiki/Category:Taste#mediaviewer/File:Bad_taste.jpg

Although this is a very simplistic mechanism, for a sufficiently large population of prospect buyers, the outcome would be exactly as given by the ramp function.

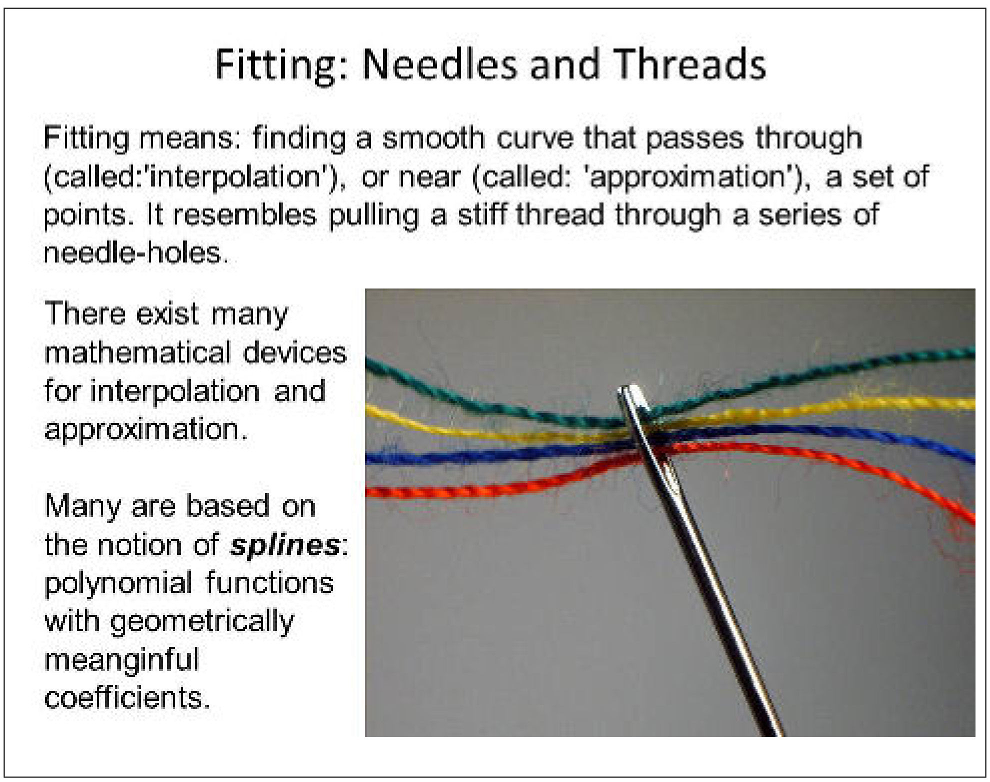
The ramp function approach therefore can be called a glass box model. It is a formalization of an postulated mechanism.

For the logistic function, this is different.

The version with no further quantities, $\frac{1}{1+e^{-x}}$, certainly does not suit our needs. When we modify the expression to make it more applicable, for instance to Expression 4.6, we introduce quantities $(A,\ K,\ B,\ \nu,\ Q,\ M)$ of which only A and K have a relation to a hypothetical mechanism of choosing peanut butter in dependency of its price [▷86] .

The logistic function, in this example, would classify as a *black box function*. In Section 1.3.8 we discussed the difference between glass box and black box models. We learned that a black box function commonly serves to give a mathematical expression for relating empirical data.

In the peanut butter example, however, there is no data.



So: if we should use a black box approach, we must ensure that we have data[24].

The standard way in examples like this is, to involve a panel of m people. These should be taken so as to represent the population of all potential peanut butter buyers.

Next, assemble a range of n prices, ranging from values that are certainly too low to certainly too high.

Then, for every price from the list, ask every member of the panel whether (s)he would buy our peanut butter for that price. This means that at the end we have a list $L = [[p_0, \phi_0], [p_1, \phi_1], \cdots [p_{n-1}, \phi_{n-1}]]$. Here ϕ_i is the fraction of people that would buy our peanut butter for price p_i.

In an elaborate version of the experiment, we repeat the interrogation a number of times, each time with a different panel, to get an idea of the precision of the ϕ_i. Then the list is $L' = [[p_0, \phi_{0,min}, \phi_{0,max}], [p_1, \phi_{1,min}, \phi_{1,max}], \cdots [p_{n-1}, \phi_{n-1,min}, \phi_{n-1,max}]]$, where $\phi_{i,min}$ and $\phi_{i,max}$ give a range for the fraction of people that would buy our peanut butter

[24] The photograph from a needle and some threads is taken from http://upload.wikimedia.org/wikipedia/commons/d/da/Aguja_Hilo_1.jpg?uselang=nl

for price p_i.

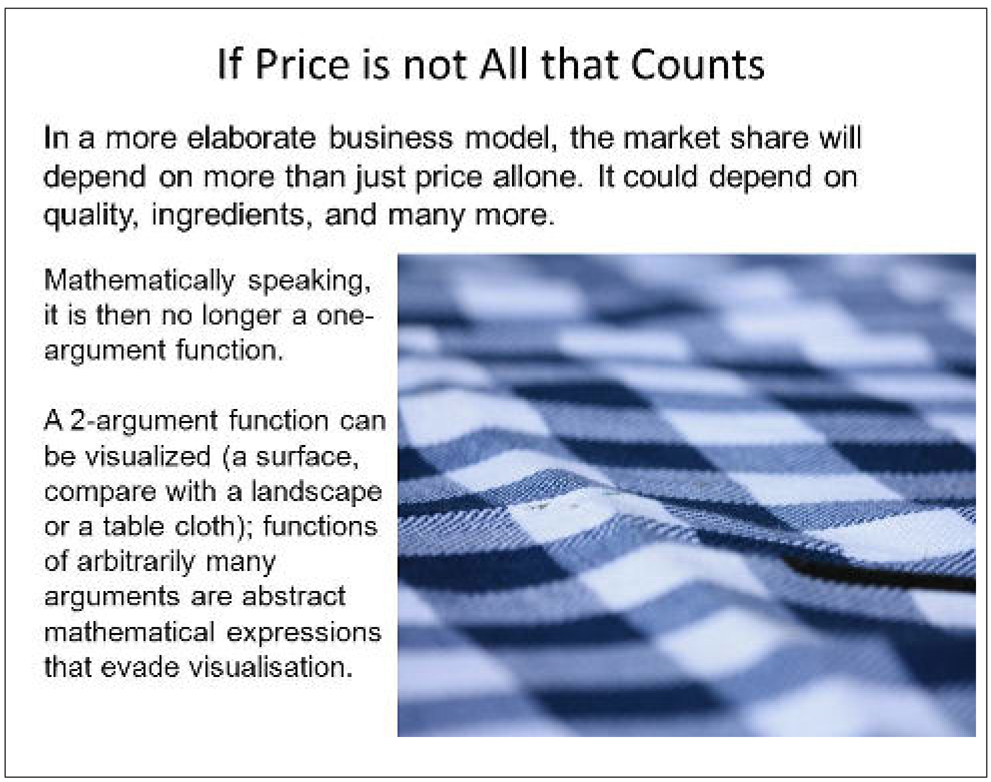
Such a list, either L or L', is a function. Given a value for p_i it gives us the fraction ϕ_i, or it gives us lower and upper bounds for such a fraction.

It cannot, however, give a value for a value $p_{notInList}$ of `pricePerItem` that is not one of the p_i. Still, there must be a fraction for *any* value of `pricePerItem`. To get a function f_ϕ such that $\phi = f_\phi(p)$ gives the market share for *any* price p, we can use a standard technique called FITTING.

Linear least squares, to be explained in Section 6.3.2 is one of many different forms of fitting. Tools like MS-Excel offer many different procedures for finding best-fit functions to sets of data. In Appendix **??** we give a brief elaboration on some issues related to fitting.

Functions of One or More Arguments

We return to the peanut butter case.



Suppose we have done the panel interrogation in the luxurious version[25]. We obtained data $[[x_i, y_i]]$ where y_i is the market share, given a price per item of x_i. The y_i have uncertainty ranges. We may have noticed in passing that some uncertain ranges were quite large, but since we used an advanced fitting scheme, we obtained a function f_{fit} to calculate the expected market share for arbitrary selling price p. Next we multiply p with $f_{\text{fit}}(p)$, and search for the value for $p = p_{opt}$ that optimizes $pf_{\text{fit}}(p)$. We bring the peanut butter to the market for p_{opt}.

Despite our efforts, however, the model may fail dramatically: the realized profit may be much lower than predicted.

We analyse how that can be possible.

We started with the expressions `income = pricePerItem * nrSoldItems` and `nrSoldItems = nrSoldTotal * marketShare`. We abbreviate these quantities as follows:

- $inc = pPI \times nSI$ (inc=`income`; pPI=`pricePerItem`; nSI=`nrSoldItems`);

[25] The image of a tablecloth was taken from http://www.rgbstock.nl/photo/ojkz2F6/picknickkleed+patroon

- $nSI = nST \times fFM$ (nST=`nrSoldTotal`; mSh=`marketShare`);
- $mSh = f_{\text{fit}}(pPI)$,

so the expected income is

$$inc = pPI \times nST \times f_{\text{fit}}(pPI). \tag{4.7}$$

This model disregards that customers buy different amounts of peanut butter. Let c distinguish customers, and let customer c buy $nST(c)$ units of peanut butter. Then Expression 4.7 becomes:

$$inc = pPI \times \Sigma_c nST(c) \times f_{\text{fit}}(pPI, c). \tag{4.8}$$

The difference[26] between Expressions 4.7 and 4.8 is, that first we had $f_{\text{fit}}(pPI)$ and next we set $f_{\text{fit}}(pPI, c)$. When f_{fit} is a function of both pPI and c, we can express that not every customer finds the same price appropriate for our peanut butter. If f_{fit} only depends on pPI, we cannot make this distinction ▷87 .

Working with Expression 4.7 or with Expression 4.8 can make a major difference.

Suppose that parents with large families, buying large amounts of peanut butter find our peanut butter too expensive. This is quite plausible: when needing much peanut butter, you can save a lot by buying the cheaper varieties. These peanut butter buyers, because for them nST is large, have a relatively large contribution to $\Sigma_c nST(c) \times f_{\text{fit}}(pPI, c)$. In this scenario, the value for inc according to Expression 4.8 can be much lower than with Expression 4.7.

The crucial mistake we made was, to fit the entire data set with a function with *a single argument* only. Using a single argument-function, it is impossible to make a distinction between individual customers, or even between two or more distinct customer groups.

[26]The photographs of the lady and the shopping cart were taken from `http://commons.wikimedia.org/wiki/Category:Shopping_bags#mediaviewer/File:Tetra_Pak_housewife_with_shopping_net,_1950s.jpg` and `http://commons.wikimedia.org/wiki/Category:Shopping_carts#mediaviewer/File:Shopping_cart.png`, respectively

We explore this idea a bit. Suppose that there are two types of customers: one, called type BIG, representing big families, buying peanut butter only if it is cheaper than 1 Euro/jar. The other type is called SMALL. They buy peanut butter if it is cheaper than 2 Euro/jar.

If the panel would only consist of type BIG, the fitted data gives a step function that is 1 for $pPI < 1$ Euro/jar and 0 for $pPI > 1$ Euro/jar. In case the panel would consist only of type SMALL, also a step function results, at 2 Euro/jar.

In reality, the panel will consist of a mixture of BIG and SMALL type customers. Since we didn't think of the difference between BIG and SMALL beforehand, we didn't consider to select panel members beforehand. The data therefore will be a mixture of points belonging to the one step function and those belonging to the other step function. If we ignore this mixing, we use a 1-argument function to fit the entire collection. The result is neither of the two step functions. We get a vague graph that doesn't do well in approximating either of the data separately.

The lesson to learn is, that, before postulating a function to fit data, we first should ask whether a 1-argument function would be sufficiently expressive. The branch of statistics called MULTIPLE REGRESSION ANALYSIS studies this topic in depth.

4.4 Mathematical preliminaries

materiaal aan te vullen door Emiel ...

...

...

...

4.5 Summary

- Translating a conceptual model to a *formal model* cannot be done in a formally provable correct way;
- A heuristic approach has the following elements:
  - appropriate *naming*
  - structure

    chain of dependencies: the formal model as a *directed acyclic graph*;
    - What *mechanism*?
    - What *quantities* drive this mechanism?
    - What is the *qualitative* behavior of the mechanism?
    - What is the mathematical expression to describe this mechanism?

    to-do-list to ensure that all intermediate quantities are found and elaborated in turn;
  - *Formation of mathematical expressions*:

    dimensional analysis can give mathematical expressions, e.g for proportionality and inverse proportionality;

    the *Relation Wizard* can help finding appropriate fragments of mathematics;

the *Function Selector* can help finding an appropriate expression for a desired (visual) behavior;

* The notion of *ramp functions* to formalize *monotonic*, *saturating* behavior;
* The *logistic function* as a smooth approximation to the ramp function;

wisdom of the crowds can help improve the accuracy of guessed values;

4.6 Learning goals

4.6.1 Knowledge

You should know about the necessity of translating a conceptual model to a formal model; you should know that this translation is not a formally provable step which always requires interpretation and imagination. You should know about four heuristics to be used in this translation: the good practice of naming convention, the use of dimensional analysis to derive plausible formulas, the use of the to-do list and the role of Wisdom of the Crowds. You need to know the definition of a DAG, and how it is used in relation to the to-do list; you should know the working of the Relation Wizard and the Function Selector.

Emiel: aanvullen svp.

4.6.2 Skills

To build a formal model, you should be able to choose names for quantities according to common naming conventions; you should be able to pick the quantities to start a DAG with, based on the purpose of the model, and you should be able to construct this DAG using the to-do list. You should be able to understand the derivation of the expression for l_p in the street lamp example, both in the diffuse case (two binary relations) and in the specular case (one ternary relation; you don't need to reproduce the derivation in the case of the specular case, though.) In developing the DAG, you should be able to construct formulas using dimension synthesis; when capacity plays a role in your model, you should be able to formulate capacity-related relations. You should be able to make use of pseudo units; you should be able to make use of Wisdom of the Crowds. You should be able to make a first estimate of the error bounds of your model (the correct use of Expression 6.18 is not required as part of the learning goals for this chapter). If you encounter dependencies that cannot be obtained from dimension analysis, you should be able to formulate requirements for the function you need; in simple cases, you should be able to consult the Relation Wizard and the Function Selector for help. You should be able to 'tweak' expressions for functions in order to make their graph comply with requirements such as asymptotes, zeroes, etc. In case you cannot find an expicit function, expressing the quantity you are interested in in terms of further quantities, you should be able to try an alternative route, formulating one or more equations and, in simple cases (e.g., linear equations), solve these. You should know how to choose between a glass box and a black box approach.

Emiel: aanvullen svp.

4.6.3 Attitude

When confronted with a problem that might benefit from a formal model, you should consider to use the techniques based on the construction of the DAG as explained in this chapter. When you introduce a functional dependency, you should be inclined to assess if you have found all arguments for the function at hand. You should be inclined to perform an initial error analysis and estimate the upper and lower bounds of the numerical outcome. When using Wisdom of the Crowds, you should be inclined to consult a possible large number of peers for a possibly accurate estimate.

4.7 Questions

1. In Section 4.1, we give some examples involving dice, resistors and springs. What is the point we try to make?
2. $\frac{1}{3}p + \frac{1}{2}q$, where p and q are points in the plane, doesn't mean anything, whereas $\frac{1}{2}p + \frac{1}{2}q$ has a well-defined meaning. Explain.
3. $\frac{1}{4}p + \frac{1}{4}q + xr$ is a point in space, and p, q, and r are points in space. What is x?
4. We have a lens with focal length f_1 and a second lens with focal length f_2. We build an optical system containing both lenses. What is the focal length of the optical system?
5. Irrespective of the dimension of X and Y (assume they are equal), show that $X + Y$ and $\frac{XY}{X+Y}$ have the same dimension. What does this mean for the usefulness of dimensional analysis to find the right formula for 'addition' ?
6. What do we mean by 'there is no formal provable correct way to translate intuitions [..] into mathematical constructions'? Does this mean that any mathematical construction is equally 'good' to represent some intuition? Discuss.
7. The heuristic approach to formalization comprises of three elements. Which?
8. What conventions should apply to naming quantities?
9. What is meant by 'camel casing'?
10. The element 'structure' of the formalization consists of 'chain of dependencies' and 'to-do-list'. Explain both in your own words.
11. What is a DAG?
12. We build a DAG as a formalization of a conceptual model. What are root nodes, and what do they stand for?
13. What do we mean by 'expansion' in relation to the construction of the 'chain of dependencies'?
14. What are the 'leaves' in DAG?
15. Explain that the graph, constructed using the to-do-list, will always be a-cyclic.

16. Expansion consists of four steps. Which?
17. In Section 4.2.2, we explain that the relation `sees`, which first occurs as one 3-ary relation, under the assumption of a dry (=diffusely reflecting) road, can be replaced by two 2-ary relations. Give the essence of the argument, without elaborating the mathematical details.
18. Describe, in your own words, the working of the to-do list.
19. When is a model executable?
20. Explain in what sense, the recommendation 'two models is better than one model', resembles the 'wisdom of the crowds'.
21. When should you apply the heuristic 'wisdom of the crowds'?
22. Develop the detergent-model, making use of the number of provinces in Holland (=12).
23. Develop the detergent-model including a term `amSpilledDet`. Is the new model outcome better? Discuss.
24. Why do we develop a model (such as the detergent model) based on units rather then on dimensions?
25. Think of a purpose for the detergent model. Given this purpose, analyze whether our solution (130 million kg/year, with an error bound of 30%), is sufficient or not.
26. After completing the detergent model, we give a discussion of how to deal with pseudo units.
    (a) In your own words, explain what *pseudo units* are.
    (b) Give an example where pseudo units help yo find the right formula.
    (c) Give an example where pseudo units don't help to find the right formula.
27. What do we mean by 'proportional to Eindhoven', in the chimney sweepers' problem?
28. We expand the quantity `nrSwPCh` by introducing *capacity*. Explain what capacity is.
29. The derivation of the chimney sweepers' model uses two time units. Why? Derive the model, using only one time unit.
30. Develop the chimney sweepers' model using estimates for the annually deposited amounts of ashes and soot. How accurate is the outcome of this version of the model?
31. Develop the chimney sweepers' model using estimates for the costs (salary of a chimney sweeper). How accurate is this version of the model?
32. In the peanut butter problem, we need a function to represent price elasticity.
    (a) What is price elasticity?
    (b) What do you know of a function that represents the behavior of price elasticity?
    (c) Give two functions that reproduce this behavior; for both, give an advantage and a disadvantage.

33. What is fitting?

34. Explain in your own words which problem with the peanut-butter model causes the 'dramatic failure' in the text.

35. Under which assumption are Expressions 4.8 and 4.7 identical?

36. We made a mistake fitting the data set $[(pPI_i, mSh_i)]$ with a function with a single argument only. Which argument? If we would take a function with two arguments, what should the second argument be?

37. In which circumstance would a modeler need to resort to multiple regression analysis?

4.8 Exercises

1. In Section 4.2.2, we show that the arity of a relation may depend on assumptions: on the assumption of a diffusely reflecting road, sees can be written as two 2-ary relations that can each be formalized independently; for a specularly reflecting road, this is not possible.
   (a) The number of calculations to deal with the 3-ary version of sees is more than the number of calculations to deal with the version consisting of 2-ary relations. Why? What determines this difference?
   (b) Suppose that the street lamps would serve to better see other cars. What would then be the meaning of the relation sees? What would be its arity?
   (c) Can the arity of the relation sees be reduced for street lamps that are supposed to illuminate other cars? If yes, how? If no: why not?
   (d) Think of another example of a problem where a relation, similar to the sees relation in the street lamp problem, can take different arities, depending on assumptions.

2. When developing a formalization with a to-do-list, the order in which we take quantities from the to-do-list is arbitrary.
   (a) Give an argument that the resulting formal model, in case of a DAG with one root, will not depend on the order.
   (b) What is the situation in case of a DAG with two or more roots?
   (c) Give an argument that, even if the resulting formal model will be the same, one order for processing the to-do-list may be more convenient than another.
   (d) Compare the DAG that follows from using the to-do-list with the entity-relation graph that follows from the conceptualization stage. These are to some extent independent of each other, but to another extent they are related.
      i. In what respect are they independent?
      ii. In what respects (more than one!) are they related?

3. We made the model for the detergent problem without first making a conceptual model. Make a conceptual model for the detergent problem.

4. We made the model for the chimney sweepers problem without first making a conceptual model. Make a conceptual model for the chimney sweepers problem.

5. In the solution of the detergent-problem, critically analyse each step (=each line in Table 4.1).

   (a) See if there are assumptions, other than the ones we give in the derivation.

   (b) For such additional assumptions, find out how the model should be adjusted in case the assumption does not hold (hint: in the text, we give one example where we introduce `nrCarefulWshs`). Notice: you are not supposed to rework the entire model, only do one or some adjustments per step of the derivation; for the next step you may assume again the original model.

   (c) Do you find steps in the derivation of the model where no further assumptions apply? If so, how do you know?

6. In the solution of the chimney sweepers-problem, critically analyse each step (=each line in Table 4.3).

   (a) See if there are assumptions, other than the ones we give in the derivation.

   (b) For such additional assumptions, find out how the model should be adjusted in case the assumption does not hold. Notice: you are not supposed to rework the entire model, only do one or some adjustments per step of the derivation; for the next step you may assume again the original model.

   (c) Do you find steps in the derivation of the model where no further assumptions apply? If so, how do you know?

7. Develop a model, similar to the chimney sweepers model, to find the number of clock-makers working in Eindhoven. First define a purpose; compare the accuracy of the outcome of your model to the requirements imposed by your purpose. If possible, compare the outcome of your model to the actual number.

8. Develop a model, similar to the chimney sweepers model, to the find which of the three: planes, cars, trains, consume more fuel. Restrict yourself to the situation in one country (choose the country you feel most confident with). First define a purpose; compare the accuracy of the outcome of your model with the requirements imposed by your purpose. If possible, compare the outcome of your model to the actual number.

9. Develop a model, similar to the chimney sweepers model, to see how much area should be reserved in Holland to grow trees for sustainable paper production on behalf of temporarily used paper (newspapers, toilet paper, ...). First define a purpose; compare the accuracy of the outcome of your model with the requirements imposed by your purpose. If possible, compare the outcome of your model to the actual number.

10. Develop a model, similar to the chimney sweepers model, to see how much area you would need to grow your own food. First define a purpose; compare the accuracy of the outcome of your model with the requirements imposed by your purpose. If possible, compare the outcome of your model to the actual number.

11. A popular belief has it, that your lungs contain some air that was also in Napoleon's lungs.

    (a) Formulated in this form, the belief cannot be verified. Make a reformulation, to be called R, that can be made subject to verification by a model.

(b) Develop a model, similar to the chimney sweepers model, to verify R.

(c) With respect to the estimates, needed to calculate (execute) your model, there is a difference in whether you want to want to verify if R could be true or whether you want to verify if R could be false. Set up two models, one for each purpose. (Hint: we formulated three purposes for the chimney sweepers problem. In one purpose we needed to find an upper bound; in another purpose we needed to find a lower bound for the number of chimney sweepers working in Eindhoven. Something similar applies in this case.)

12. A popular belief has it, that if all people of the world would simultaneously start walking in Eastward direction, that this, due to conservation of angular momentum, would slow down the rotation of the earth. Call this belief R_1. A popular counter argument is that this cannot be true, for the fluctuations of the motions of the masses of air in the atmosphere correspond to a much larger angular momentum, and these also (seem to) have no net effect to the rotation of the earth, as the timing of sunset and sun down can be predicted within a minute accuracy for many years ahead. This counter argument is called R_2.

    (a) Verify R_1 using a model, similar to the chimney sweepers model for its first or second purpose. That is: first think if we need an upper bound or a lower bound (of what?) to achieve our purpose, and next see if a simple model can be made that is sufficiently accurate to give the required upper or lower bound.

    (b) Verify if R_2 is a sufficient counter argument, also by using a model similar to the chimney sweepers model for its first or second purpose.

Chapter 5

Roles of Quantities in a Functional Model

'Four is the only number that can be written as $x + x$, $x * x$ or x^x'

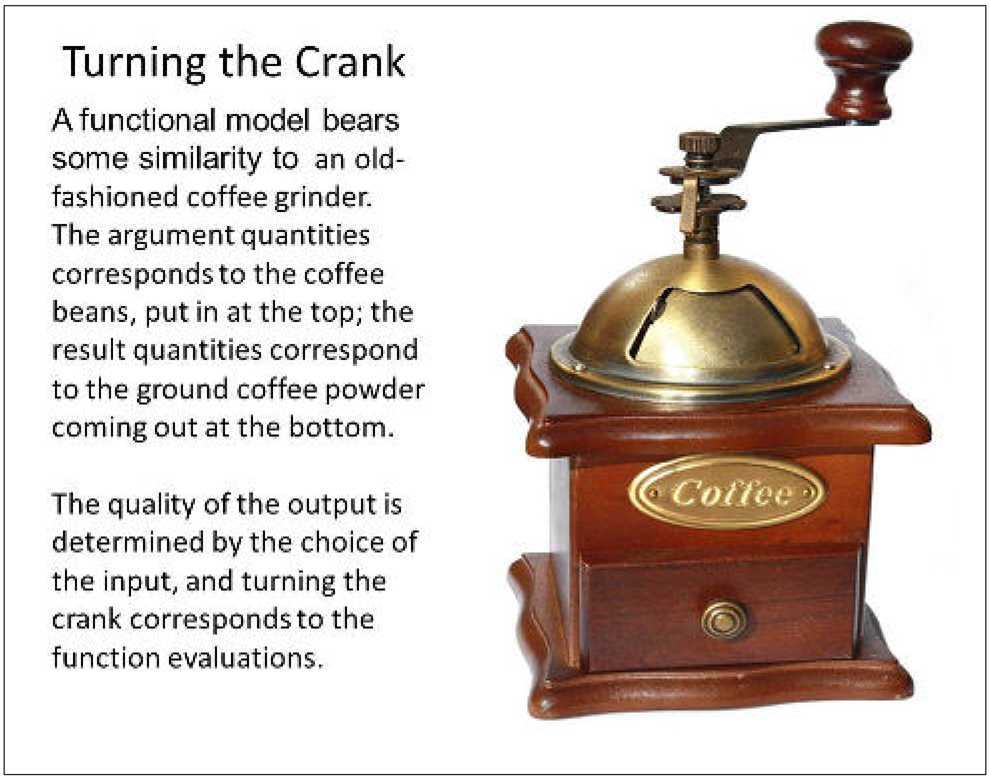
The weather has been fine for some days now, and the forecast for this afternoon is no less than splendid. Tom, Dick and Harriet decide to spend the afternoon outdoors, and the idea of a bicycle trip soon pops up. Tom unfolds the map. His finger follows a wavy brown line in the middle of large blotches of green. "Let's take the old forrest trail and then follow the river to ...". "Hang on", Dick interrupts, "we can't leave for too long as I have to do my laundry, and ...". Harriet agrees with him. "Yes, and I don't want to get dead tired like last time. So at least we should drive slowly ...". Tom frowns. "But driving slowly we don't get far and we won't see much. Unless ... ". He glances at Dick. "Perhaps we could get back a little later?". Dick shrugs. "Or drive a bit faster. Moreover, I don't insist on seeing much, I have been over there last week, and ...". Harriet seizes the map. "I don't mind seeing much, but then we'll need more time. Perhaps we should ...". And then, as if by agreement, all three look up from the map. "Perhaps we should make a model to help us find out!"

5.1 Functional Models

A functional model is a model where input is mapped to output. The input is something that is under control of the modeler. The output represents something that is interesting, relevant or valuable in relation to the purpose of the model. In Chapter 4 we introduced a simple strategy, involving chains of dependencies and a to-do-list, to construct functional models. According to this strategy, we need to express the quantities we are most interested in in terms of other quantities by thinking of mechanisms, determining how the interesting quantities receive their values.

Functional models occur for many different purposes[1]. We give some examples:

predict (1) 'when ...': The input is empty. The output is a time point in the future. The prediction, for instance, of the next solar eclipse has no free quantities: the constellation of the solar system cannot be influenced.



predict (2) 'what if ...': The input corresponds to the condition after the 'if'. For instance, in a street lamp simulator one can investigate what the reflected light (=the output) would be if the lanterns would be closer together. The distance between the lanterns then is an input quantity.

decide: A *decision* means: binding values to quantities that we have the authority to define. In a game of billiard, a decision would be the position and orientation of the cue for a shot. The output is the consequence of the decision: whether there will be a carambole or not. Decisions are an important application of modeling. Designing can be seen as a process of decision making, and the role of modeling in design amounts to supporting these decisions. That is: suggesting which decisions should be made (=which values should be bound to decision quantities) in order to achieve certain, desired consequences.

optimize: Optimisation is a special case of decision making. Input quantities are the free quantities that can be varied with the intention that some ordinal output quantities are optimal.

verify: Verification is similar to prediction (1), where we want to know whether something with certainty will or will not be the case. So the input is empty, and the output is limited to the values true and false.

steer/control: A model for steering or controlling a system should keep the difference between a desired value and a realized value small. A simple example is a thermostat, that should maintain the temperature close to a set value. The output is the achieved difference between desired and realized values; the input may be empty. Alternatively, the input consists of perturbations (say, whether a door is opened or closed).

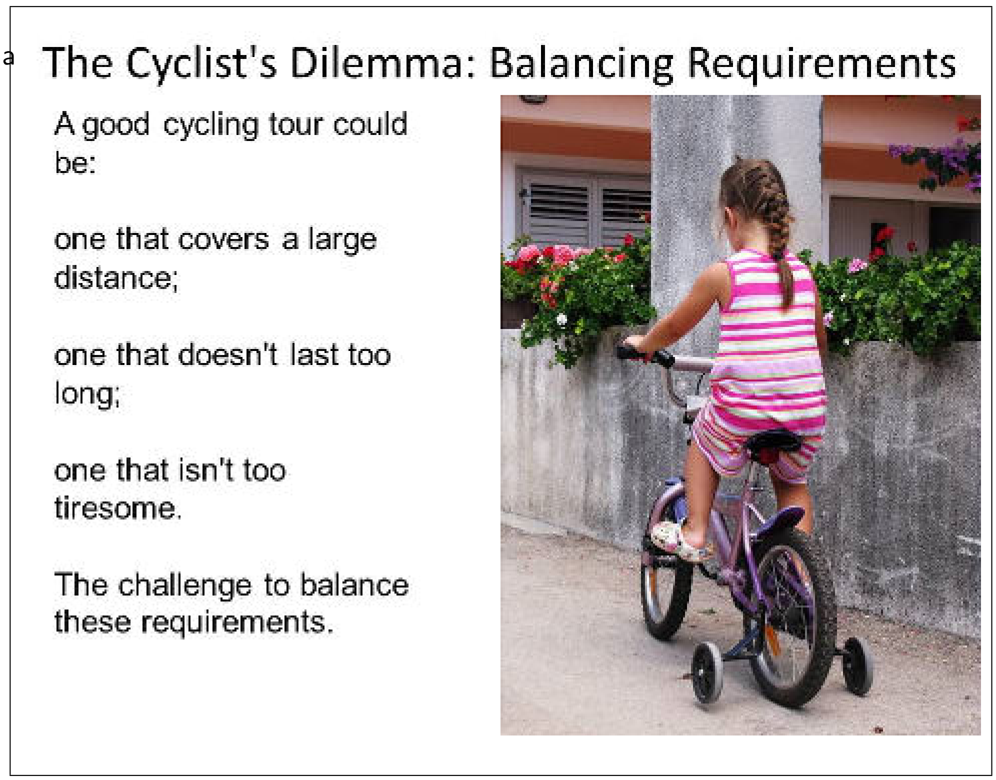
[1] The photograph of the coffee grinder is taken from http://www.rgbstock.nl/photo/2dRf6qp/koffiemolen

The output of a functional model often represents some intention. In particular this applies to deciding, optimizing, and steering / controlling. In cases where we need to distinguish more and less successful outputs, output quantities need to be ordinal: it must be possible to *compare* one outcome to another outcome.

5.2 The 4-Categories Approach to Building Functional Models

5.2.1 The Structure of a Functional Model

We develop the structure for a functional model by analyzing the cyclist's dilemma ▷88 [2] from the introduction of this chapter. We have three quantities that express how good a cycle trip is. Tom is interested in seeing as much as possible; for him, the distance cycled, s (in kilometers), should be as large as possible. Dick doesn't have much time for cycling. To him, the time spent cycling, t (in minutes or hours), should be as short as possible. To Harriet, the amount of effort, W (in Joules) should be minimal.



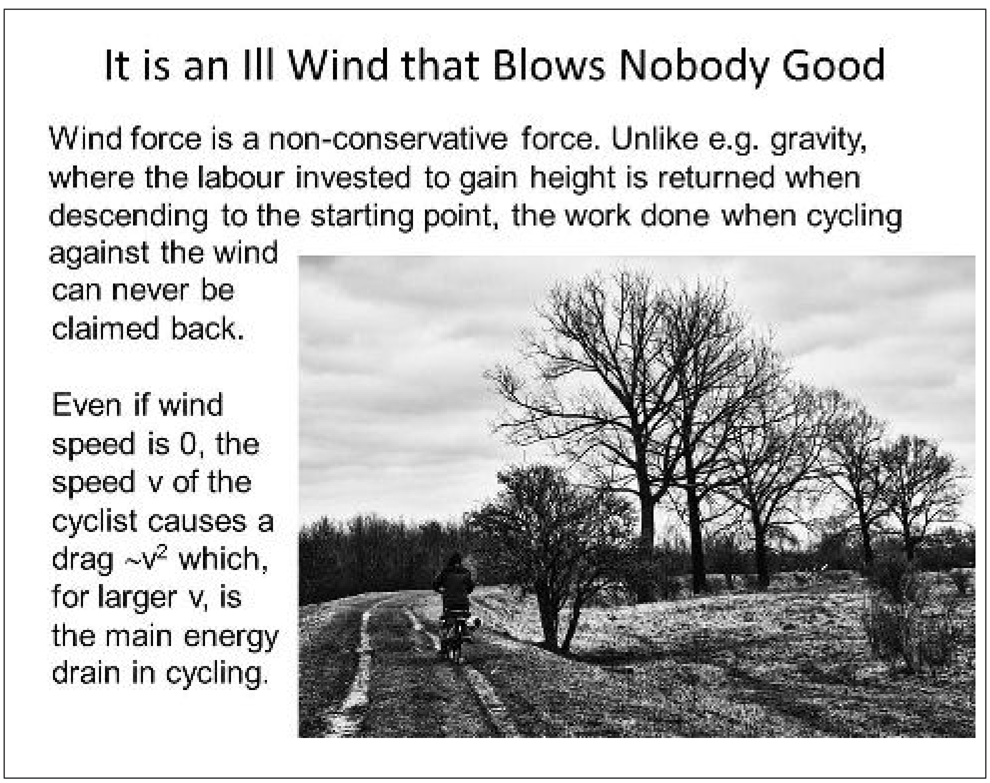
It is not a priori clear which functional model we need: it could be $s = f_s(t, W, ...)$, or $t = f_t(W, s, ...)$, or $W = f_W(s, t, ...)$. We use the strategy from Section 4.2, making use of the to-do-list. We recapitulate the strategy:

- identify a quantity you need for the purpose
- put this on the to-do-list
- while the todo list is not empty:
  - take a quantity from the todo list
  - think: what does it depend on?

[2] The photograph of the girl on a tricycle is taken from http://www.rgbstock.nl/photo/n5mqrrq/op+een+fiets

  - if depends on nothing ?
    * are we free to choose? $\rightarrow$ choose
    * are we not free to choose? $\rightarrow$ substitute value from context
  - else
    * give a formula for the quantity, expressing it into the quantities it depends on
      · if possible, use dimensional analysis to find an expression;
      · otherwise, propose a suitable mathematical expression (perhaps use the Relation Wizard or the Function Selector)
      · think about assumptions
      · in any case, verify dimensions
    * add newly introduced quantities to the todo list
  - end if

- todo list is empty: evaluate your model
- check if purpose is satisfied



Since the cyclist's dilemma is relatively simple, we only elaborate the most crucial steps. To fully develop it, we need some quantities from the context of the model. First, we should realize that the effort W can be expressed as Fs, F being the force[3] excerted by a cyclist ▷89 . For now we take F constant; otherwise the effort should be expressed as $W = \int F ds$.

We assume that the main contribution to the force is the wind drag, and in chapter 2 we derived the formula $F = c\rho A v^2$. Here c is a dimensionless constant (taken to be about 1, according to experiments); A is the projected area of the cyclist, ρ is the air density and v is the cyclist's velocity relative to the air.

Before starting the formal model derivation, we make some observations.

[3]The photograph of the strong lonely cyclist is taken from http://commons.wikimedia.org/wiki/File:Hoe_sterk_is_de_eenzame_fietser_How_strong_is_the_lone_rider_(5475908107).jpg?uselang=nl

Before Developing a Model: look out for Trivial Cases

To achieve each of the goals independently is trivial. Minimize t or minimize W is realized by setting s to zero. Maximizing s is done by setting t to infinity.

Let us now look at attempts to achieve two goals simultaneously.

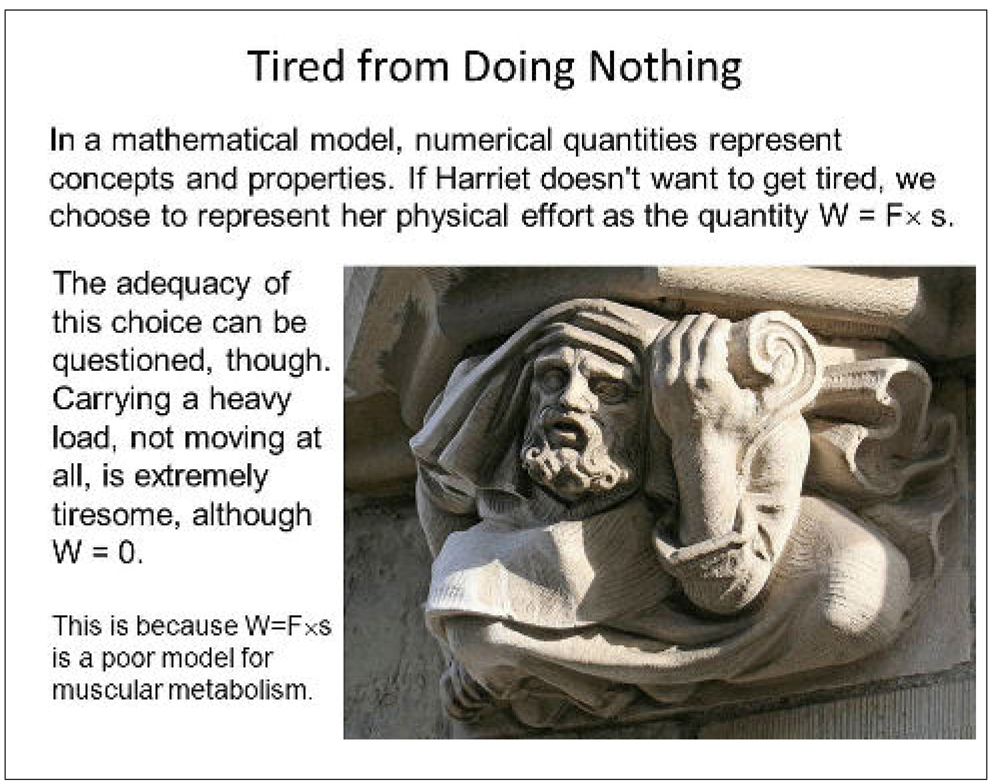
minimize t and minimize W : again achieved by setting s to zero.

minimize t and maximize s : this has no other solution than letting v go to infinity - which is not very realistic.

minimize W and maximize s : this is not trivial. It is an example of a trade-off. We will see more on trade-offs in Section 5.3.2. For now, we define a trade-off as a situation where we have one or more quantities (t and v) that we are free to choose, but that each have opposite effects with respect to what we want to achieve. Indeed, increasing t with fixed v, causes an increase of s (good!), but also an increase in W (bad!)[4]. Increasing v with fixed t will increase s (good!), but it also increases W (bad!). We could ask if there is some combination of t and v that is in some sense optimal.

Finally we might look at all three the desired quantities: what happens if we take s, t, and W into account. It can be seen that this, for the same reason, leads to a trade-off; the complicating factor here is that we are not sure if we should choose the combination [s,v], [v,t], or [t,s] as quantities we are free to decide upon.

To discover a common pattern in formal models, we develop the strategy from Section 4.2 for all three of the above cases where we seek to comply with two simultaneous goals.



Case 1: maximize s, minimize W

quantity needed for purpose: s and W; put these on to-do-list

pick s from to do list: s depends on t and v

Formula: $s = vt$

pick W from to do list: W depends on F and s

Formula: $W = Fs$

pick F from to do list: F depends on c, ρ, A and v

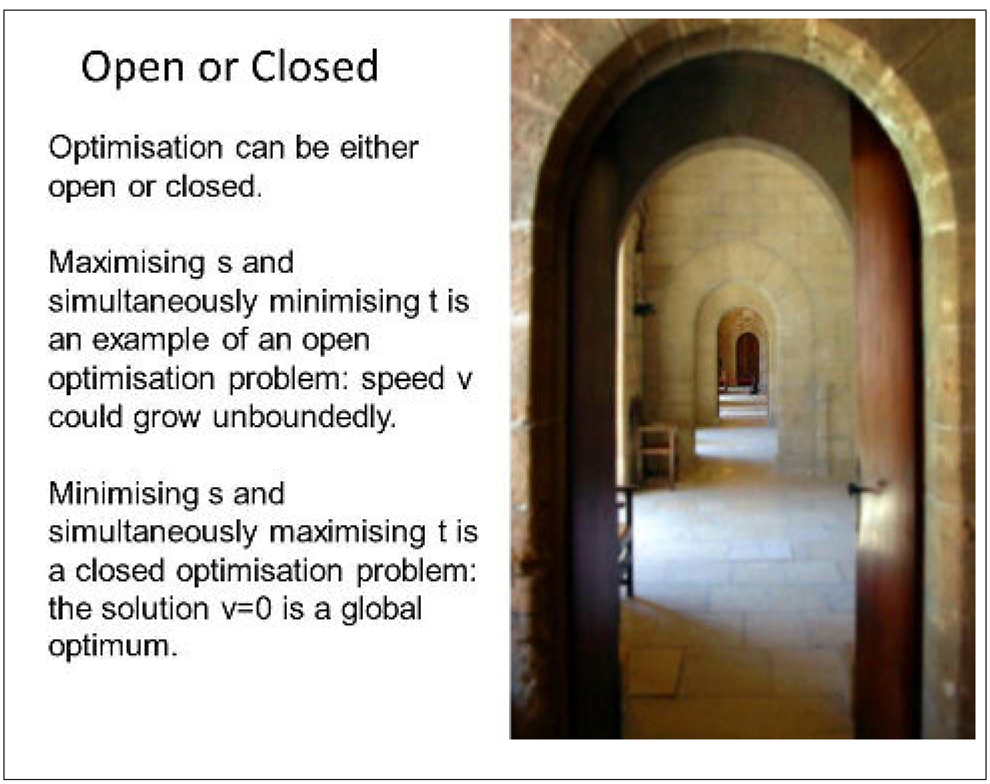
Formula: $F = c\rho Av^2$

[4]The photograph of the statue of a working labourer is taken from http://www.rgbstock.nl/photo/mmaI4He/Labourer

pick c from list $\rightarrow$ constant
pick ρ from list $\rightarrow$ constant
pick A from list $\rightarrow$ constant
pick t from list $\rightarrow$ choose
pick v from list $\rightarrow$ choose

Case 2: minimize W, minimize t

quantity needed for purpose: W and t; put these on to-do-list
pick W from to do list: W depends on F and s
Formula: $W = Fs$
pick F from to do list: F depends on c, ρ, A and v
Formula: $F = c\rho A v^2$
pick c from list $\rightarrow$ constant
pick ρ from list $\rightarrow$ constant
pick A from list $\rightarrow$ constant
pick t from list: depends on s and v
Formula: $t = s/v$
pick s from list $\rightarrow$ choose
pick v from list $\rightarrow$ choose

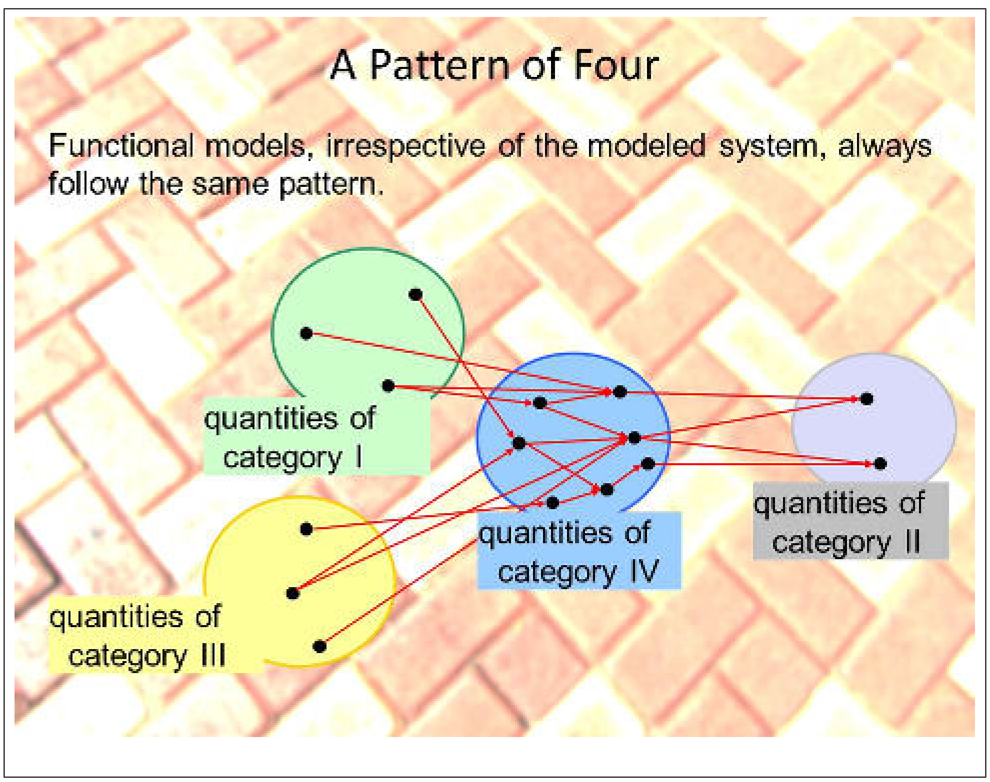


Case 3: minimize t, maximize s

quantity needed for purpose[5]: t and s; put these on to-do-list
pick t from to do list: t depends on v and s
Formula: $t = s/v$
pick v from list $\rightarrow$ choose
pick s from list $\rightarrow$ choose
(notice: W is irrelevant, and so are ρ, A, c and F)

Analysis

We developed 3 different functional models for one and the same system (=designing a bicycle trip, characterized by s, t, W), that should meet with various combinations of requirements on s, t and W. The occurring functional expressions, and some of the quantities are different. F, c, ρ, A and v occur in some models, and they are absent in others. Still, there is a number of similarities between the three models:



- All three models consist of functional definitions. Quantities are expressed as a function of other quantities (e.g., F), or as constants (ρ, A, c), or as quantities that can be freely chosen (s, t, v);
- in all three, there are quantities that we are interested in. In the three models these are, respectively, $[s, W]$, $[W, t]$ and $[t, s]$;
- in all three, the 'interesting' quantities are expressed in one or more quantities. Some of these (e.g., F) are in turn expressed in yet other quantities;
- eventually, the chains of dependencies terminate in constants or quantities that can be freely chosen. Namely, in case 1: t, v; in case 2: s, v; in case 3: s, v.

The above patterns[6]occur in all functional models. Quantities in functional models can be at-

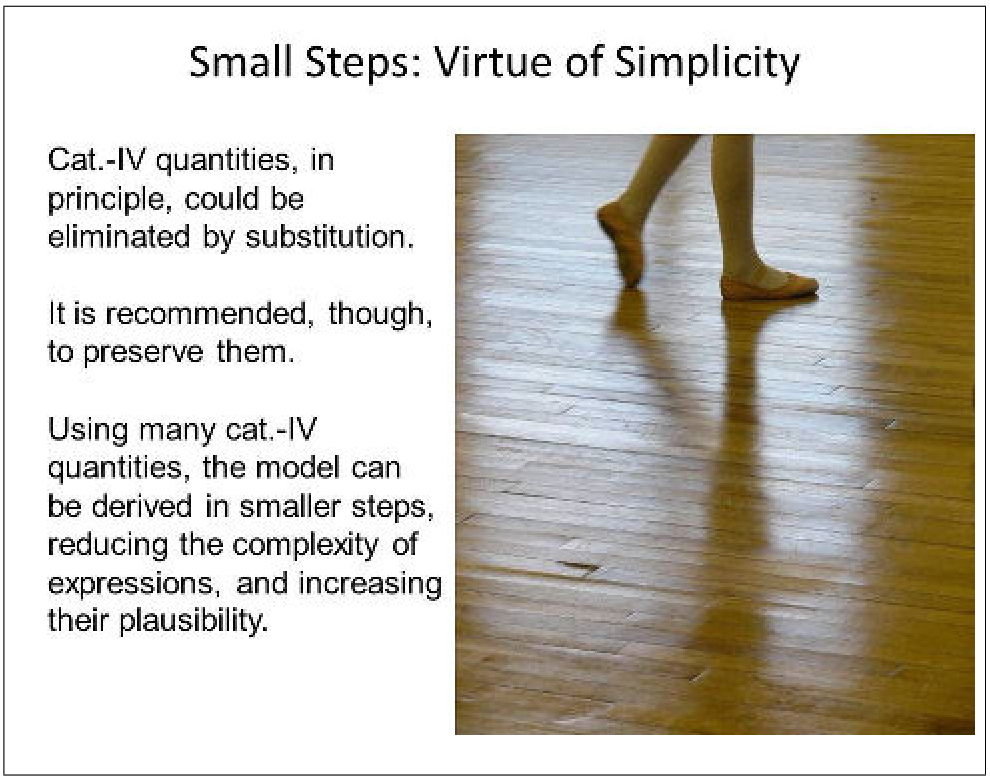
[5]The photograph of an open door in a closed monastery is taken from http://www.rgbstock.nl/photo/ook23xc/deuren

[6]The photograph of the tile pattern with four types of bricks was taken from http://www.rgbstock.nl/photo/

tributed to one of four categories:

- whether they are free to be chosen (so called CATEGORY I, abbreviated by cat.-I),
- whether they are output quantities (so called CATEGORY II, abbreviated by cat.-II),
- whether they come from the context of the modeled system (CATEGORY III, abbreviated by cat.-III), or
- whether they are intermediate quantities (CATEGORY IV, abbreviated cat.-IV).

Cat.-IV quantities result from stepwise developing the model. They could be eliminated from the model without affecting its purpose. For instance, in case 1: $W = Fs$, $s = vt$ and $F = c\rho Av^2$, so we could immediately write $W = c\rho Av^3t$. It is advantageous, however, to keep s and perhaps F in the model.



Firstly, because they represent meaningful properties of the system we are modeling, even though the eventual purpose may not need their value.

Secondly, they keep all occurring expressions as simple as possible[7]. Indeed, each of the functions $W = Fs$, $s = vt$ and $F = c\rho Av^2$ is easier to understand than $W = c\rho Av^3t$.

Thirdly, they are useful if the model needs to be checked for correctness. It can be helpful to inspect the value of category-IV quantities.

Finally, they can help improving the performance of the model. As follows. It is possible that a single cat.-IV quantity q occurs in the chain of dependencies for more than one cat.-II quantity. Once it is evaluated for one chain, its value can be reused in other chains. Eliminating q would require all chains to be fully evaluated for each cat.-II quantity, which may involve redundant computations.

So, although eliminating cat.-IV quantities by substitutions such as above reduces the number of functional definitions in the model, the benefit of a slightly more compact model usually does not outweigh the benefits of having simple functional definitions.

We conclude that the procedure for stepwise construction of functional models in a natural way

n7mAPHo/bestrating+patronen

[7]The photograph of the girls feet is taken from http://www.rgbstock.nl/photo/mg1RmNo/Nastreven+van+een+droom

leads to models containing quantities that occur in one of four categories.

In the sequel, we look a bit closer to each of these categories.

5.2.2 Input of the Functional Model: Category I

All decision quantities together form inputs of the functional model. The input of the functional model is the cartesian product of the sets of values for each of the decision quantities. A CARTESIAN PRODUCT of two sets is the set of all tuples where the first element is taken from the first set and the second element is taken from the second set. Cartesian product straightforwardly generalizes to more than two sets. For instance, the cartesian product of {1,2,3} and {a,b} is {[1,a], [1,b], [2,a], [2,b], [3,a], [3,b]}.

The cartesian product[8] of all value sets of all cat.-I quantities is called cat.-I space. If the functional model serves to support design decisions, it is called the DESIGN SPACE. Every point in cat.-I space is one possible set of decisions. It determines one possible outcome of the functional model.

Suppose we are designing a box. The decision quantities are: width (w), height (h), depth (d) and material (m). One possible box is the 4- tuple `[w:6cm, h:3cm, d:1cm, m:cardboard]`, denoting a cardboard box with dimensions 6cm, 3cm, 1 cm. This cat.-I space is 4-dimensional. Three of its dimensions are numbers, that is: ordinal quantities, and one dimension is nominal. Such mixes between ordinal and nominal often occur in design. Perhaps we are interested in the stability of the box and therefore we may need to distinguish both various materials and various sizes in a single design.

The input of the functional model represents all possible decisions, given a set of decision quantities. Some decisions yield plausible artifacts (e.g., `[w:6cm,h:3cm,d:1cm,m:cardboard]`), some are implausible (e.g., `[w:600cm,h:300cm,d:100cm,m:paper]`). 'Plausible' and 'implausible', however, are qualifications of the resulting artefact. Stakeholders will be less happy with implausible artifacts. In our approach, where we identify the design process with a function, the

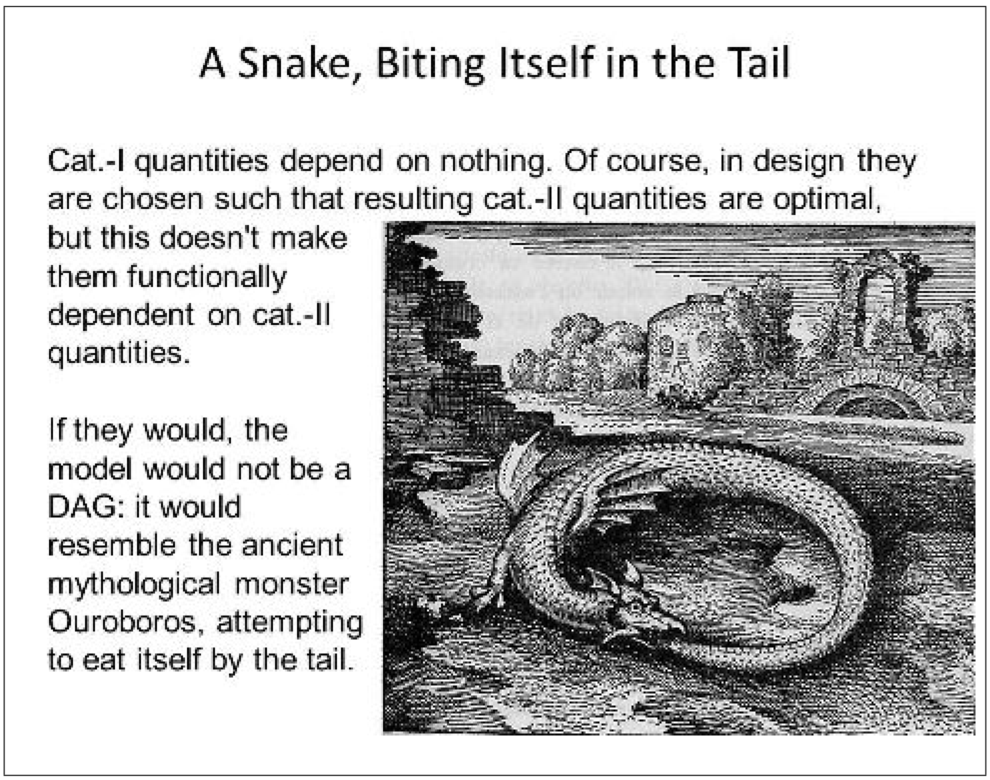
[8]The photograph of the barred window is taken from http://www.rgbstock.nl/photo/mm3Mw3w/window

appreciation of an artefact entirely belongs to the *output* of the functional model.

No Dependency-by-Anticipation

Many designers 'think ahead' when they make design decisions. A designer could argue: 'if I choose paper as a value for m, the values of w, d and h should be limited to at most a decimeter. Since my box must be larger, I should choose aluminum'. This argumentation seems to introduce a dependency between quantities in the input of the functional model. Material seems to depend on size on size. The choice for aluminium is not a *free* choice anymore.

We analyze this a bit deeper[9].



Let the cat.-I quantity be x (say, the material), and the cat.-II quantity be y (say, stability). Causally, y depends on x, but by anticipation, x depends on y. There is a causal loop between x and y: they mutually depend on each other. In a functional model of the design process, however, we cannot include such DEPENDENCIES-BY-ANTICIPATION. The model only represents dependencies for *causal* relations. We consider all quantities in the input of the functional model to be independent: they only depend on the free choice of the modeler / designer.

The modeler can propose *any* combination of values for cat.-I quantities. This causes a large part of cat.-I space to be inhabited by invalid artifacts, that is: artifacts that fail certain requirements. But whether an artefact *meets* or *fails* a requirement is an issue with respect to output quantities. Success or failure is to be assessed, based on quantities representing stakeholders' value. These are the quantities in cat.-II space.

In the example of a box, 'stability' is a cat.-II quantity. We may appreciate a stable box over a floppy one. We may even impose a constraint or a requirement, stating that stability should be at least some minimal value. The stability of a box, however, is a *consequence* of the decisions that lead to that box. We cannot choose stability as a value for some cat.-I quantity. It is a property on the output-side of the functional model. Therefore any constraint is to be formulated on the

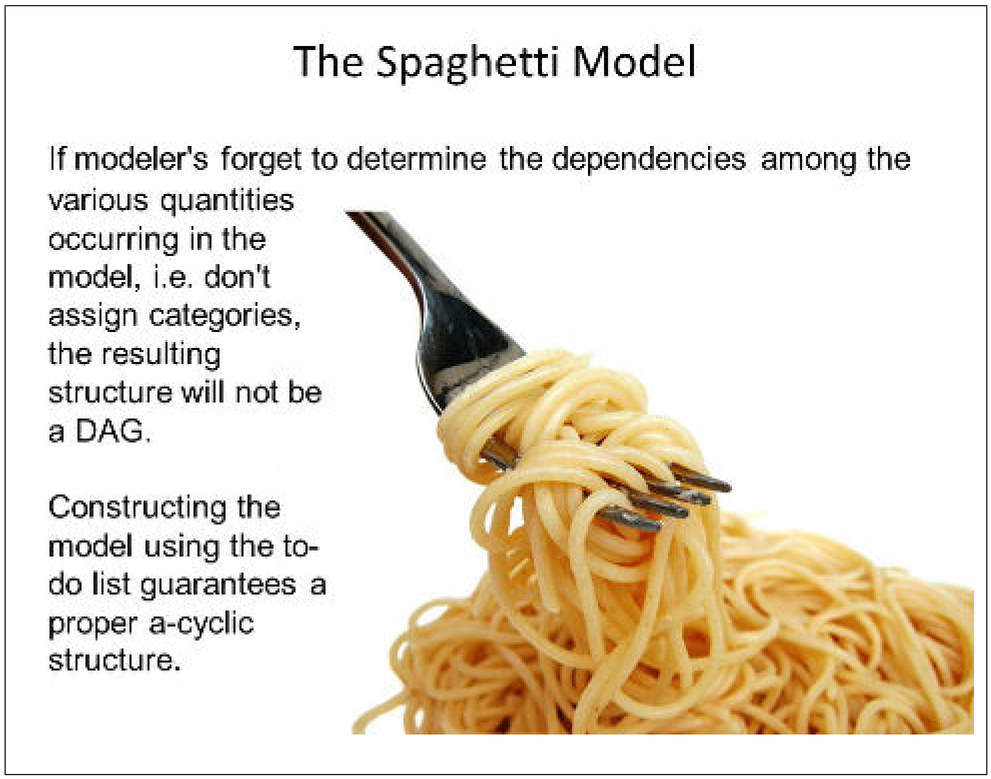
[9] The image of Ouroboros was taken from http://commons.wikimedia.org/wiki/Ouroboros#mediaviewer/File:Michael_Maier_Atalanta_Fugiens_Emblem_14.jpeg

quantities constituting the *output* of the functional model rather than the input.

Considering all decision quantities as strictly independent, we hope to achieve two things:

- There are no CAUSAL LOOPS in a functional model. A causal loop is a circular dependency. Functions ▷90 cannot be used to deal with circular dependencies[10]. The absence of causal loops has large computational advantages: if all functional dependencies are computable, any cat.-II value can be evaluated for any possible point in cat.-I space.

This means that a design process becomes an optimisation process. We can, in principle, search the entire cat.-I space to find the 'best' solution. 'Best' is defined in terms of conditions on cat.-II quantities. In practice this may not be feasible ▷91 , but we can still devise meaningful heuristics to search for 'good' solutions; furthermore, we can do 'what if' evaluations of the design.



- Perhaps even more importantly: the strict separation between input and output of the functional model helps in avoiding confusion when designers in a team have different intuitions about dependencies. In the example of the cyclist's dilemma, Tom, Dick and Harriet may agree that $s = vt$ holds. But some will think of $t = f_f(s, v) = s/v$ ('how long does it take me to cycle a certain distance'), and others think of $s = f_s(t, v) = tv$ ('how far can I cycle during a certain amount of time'). Using a functional model enforces to choose which of the three, s, t or v will be in cat.-I, and which will be in cat.-II of the functional model. This invites to make *explicit* what the purpose of the artefact will be in the context where it is to function.

5.2.3 Output of the Functional Model: Category II

In design, the output of a functional model represents the effect the ATBD will have on the stakeholders. Any wish, desire or requirement should be represented as conditions on quantities in the output of the functional model. Each wish, desire or requirement typically will regard at least one quantity. For instance, the requirement that the weight should be less than 1 kg regards

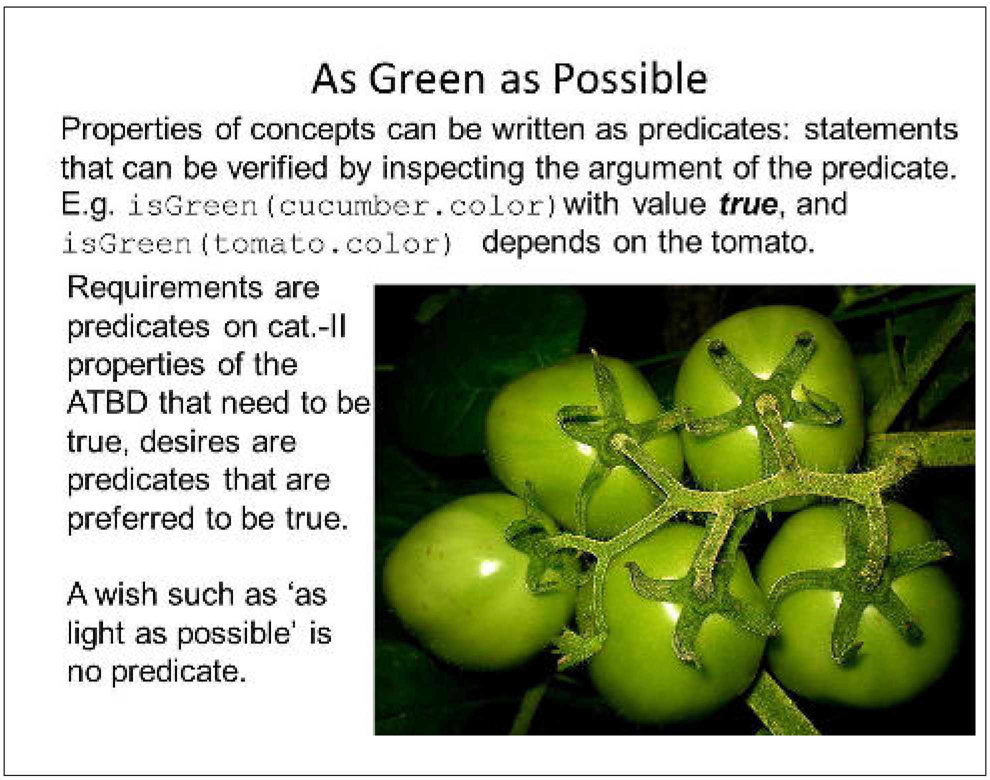
[10] The photograph of the plate with spaghetti is taken from http://www.rgbstock.nl/photo/2dk0tEY/Spaghetti

the quantity `weight`. We form *one tuple* of all result quantities, in the same way as we group the value-sets of all cat.-I quantities into the cat.-I space. The tuples form result quantities form what is called the cat.-II space. An element from cat.-II space is therefore a tuple. Every cat.-II quantity, that is: each property of such a tuple, represents an aspect to compare one candidate ATBD with other candidate ATBDs. If, for instance, a stakeholder prefers a lighter artefact above a heavier one, the artefact's weight will be one of the cat.-II quantities.

In Appendix **??** we elaborate on the choice for appropriate category-II quantities, in particular for models for design.

Using Cat.-II Quantities: Requirements, Desires or Wishes

Quantities in cat.-II express the effect of design decisions on stakeholders' values. Everything we *want* from an ATBD must therefore be expressed in cat.-II quantities.



There are three types of 'wants': *requirements*, *desires* and *wishes*. To understand these we need some terminology[11].

About Propositions and Predicates A PROPOSITION is a sentence that is either true or false. 'It is raining' is a proposition that is true if droplets of water fall from the sky. 'close the window' or 'is the window closed?' are no propositions.

A PREDICATE is a proposition with an argument, or: a function with {`TRUE`, `FALSE`} as range. Example: `isGreen(x)` is a predicate, which is `TRUE` if we substitute for `x` the term `cucumber`, `FALSE` if we substitute `canary`, and depending on the particular tomato if we substitute `tomato`.

Concepts are bundles of properties, and with every property of a concept, we can associate predicates. For instance, a cucumber being green can be expressed either as `cucumber.color=green`, `color(cucumber)=green`, or `isGreen(cucumber)=TRUE`. The fragment '`=TRUE`' in the latter expression can be omitted. Indeed the addition of '`=TRUE`' neither to a true nor to a false proposition

[11] The photograph of green tomatoes was taken from http://commons.wikimedia.org/wiki/File:Green_Tomatoes_5.jpg

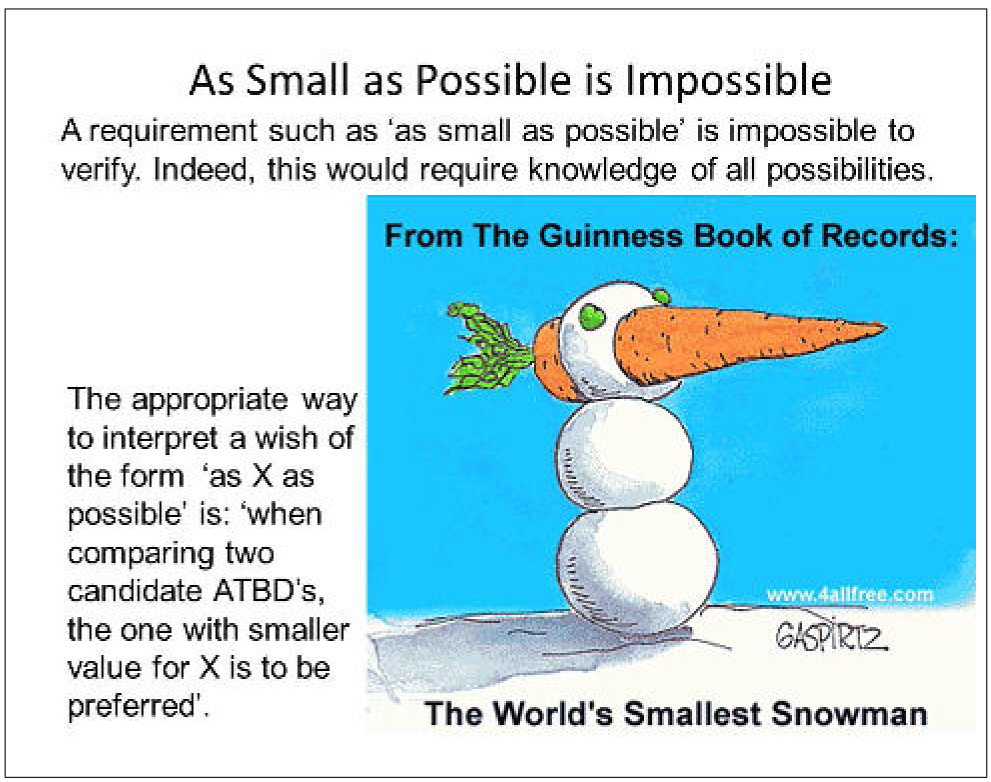
does give any further information.

Requirements: Predicates on Cat.-II Quantities The cat.-II quantities of an ATBD correspond to conditions for the appreciation of the ATBD. Some of these conditions can be expressed by predicates. Some of these predicates are the REQUIREMENTS. Examples of requirements for an umbrella are, that it is to be operated by hand, that it is waterproof or that is weights less than 3 kg. They can be expressed as predicates over the ATBD: `operatedByHand(ATBD)`, `waterproof(ATBD)`, and `weigthsLessThan3kg(ATBD)`, respectively.

Alternatively, requirements can be expressed as predicates over cat.-II quantities such as `operationMode`, `quality` and `weight`: `isManual(ATBD.operationMode)`, `isWaterproof(ATBD.quality)`, and `isLessThan3kg(ATBD.weight)` ▷[92] .

Cat.-II quantities for design need to be ordinal. This is also true for cat.-II quantities that are given by predicates. Their values are `TRUE` or `FALSE`, and `TRUE` is better than `FALSE`. Any requirement that yields `FALSE` renders the ATBD unacceptable. So requirements reduce the design space to a limited, and perhaps empty sub-space.

Desires: False is Acceptable A second type of conditions, expressed as predicates, is the DESIRE. There is no logical difference between requirement and desire. Both are predicates, and for both `TRUE` is better than `FALSE`. For a desire, however, it is acceptable that the predicate is `FALSE`.



Wishes: Conditions that are No Predicates The third type of conditions can not be expressed as a predicate. Examples are: 'as cheap as possible', 'uses as little energy as possible', 'produces as little toxic waste as possible', 'as much as possible profit'[12].

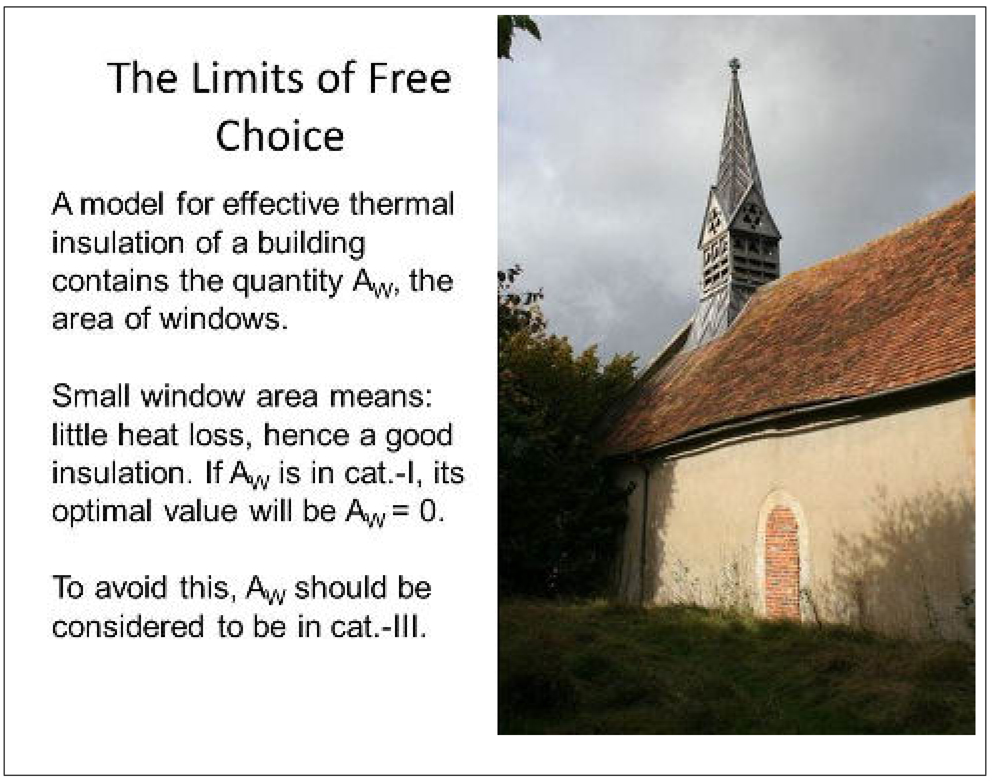
These conditions, when taken literally, are impossible to demand for an ATBD. In general it will be impossible to show that some design cannot be made, say, cheaper in any way ▷[93] . To do so would require (i) a *complete* design space, that is: a

[12] The smallest snowman-cartoon was taken from http://commons.wikimedia.org/wiki/File:World%27s_Smallest_Snowman.gif

design space that contains *all possible* cat.-I quantities, including those for ATBD's not invented yet. Further, it would require (ii) an exhaustive search over this complete design space. Both (i) and (ii) are clearly impossible.

Still, it is reasonable to demand that something is very cheap, uses very little energy, etc., without casting it in the form of requirements or desires ▷94 . The condition that something should be minimal or maximal over the current design space will be called a WISH. In Section 5.3.1 we see how we can deal with wishes.

5.2.4 Limitations from Context: Category III



Cat.-III quantities represent quantities in the context[13] of the model, which don't depend on any decisions. Quantities of category III are determined by external circumstances, e.g. physical constants, demography or vendor's catalogues. The modeler has no opportunity to adjust the values of such quantities.

Quantities in category III do not depend on any cat.-I decision, but values of cat.-II quantities may depend on them.

In many modeling processes, the demarcation between cat.-I and cat.-III quantities is interesting. Consider designing the thermal insulation of a house. Part of the heat leaks through the windows. A house without windows is cheaper to get warm than one with windows. The question is, whether it is acceptable to reduce the area of windows for insulation purposes. If it is, the area of windows is a cat.-I quantity; if not, it is in category III.

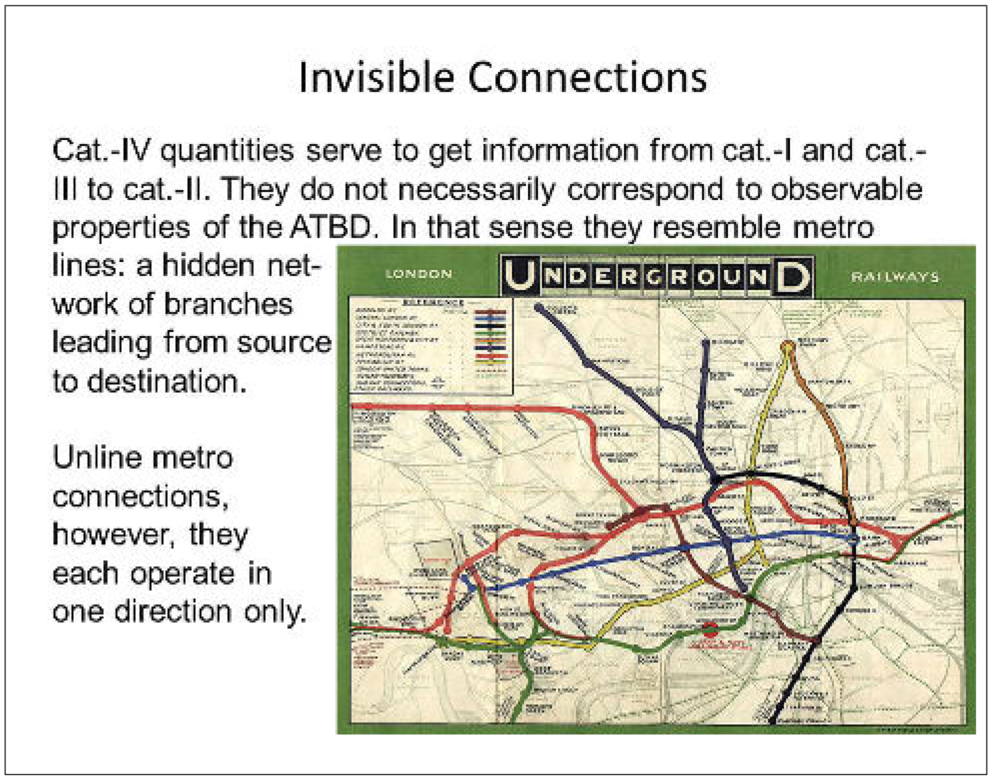
If it would be acceptable to adjust the area of windows, we immediately see that setting this area to 0 is optimal with respect to heating costs. If the model contains nothing to prevent this, the eventually selected solution is one where all windows are removed. If we don't allow this, there must be a cat.-II quantity, and an associated condition, causing window area to be large. This condition is either a requirement ('window area is at least W m^2') or a wish ('window area is as large as possible').

[13] The photograph of the chapel with no windows was taken from http://commons.wikimedia.org/wiki/File:South_side_of_the_church_-_geograph.org.uk_-_1561421.jpg

A functional model whith window area in category III is simpler; for this reason, design problems are limited to restricted design space. Nevertheless, challenging the boundary between category I and category III is a promising route to arrive at innovative design.

5.2.5 Constructing the Functional Model: Category IV

When we build a model as a collection of functional dependencies[14], cf. Section 5.2.1, a collection of intermediate quantities emerges. For each of them, we have to say if it depends on something. It may be practical to provisionally say that it isn't. We illustrate this: in a business model for a taxi company, elaborated in Appendix **??**, we encounter `flPpl`, the fuel price per liter. This fuel price depends very little on cat.-I quantities. Indeed, if the taxi company operates cars that consume much fuel, the world fuel consumption will increase, and accordingly the price will increase. But this influence is extremely little. Therefore, rather than developing this dependency and making `flPpl` a cat.-IV quantity, we ignore the dependency and treat it as a constant in cat.-III.



This pattern often occurs. A quantity, say `q`, enters a model in cat.-III, where it is given a provisional value. Once the model is developed far enough to do sensitivity analysis (see Section 6.4.2), we can estimate whether the model substantially improves if we include the mechanism that causes the value of `q` to depend on something. If so, `q` migrates from cat.-III to cat.-IV; otherwise, it stays in cat.-III.

The above explanation of categories I ... IV, together with the scheme presented in Chapter 4 using the to-do-list, is the theoretical framework for constructing functional models. In order to become able to construct an actual function model for supporting design decisions, it is recommended that the reader consults the example, elaborated in Appendix **??**.

[14] The London underground map was taken from http://commons.wikimedia.org/wiki/File:Tube_map_1908-2.jpg

	Category	May depend on	Meaning w.r.t. the design decisions	Type	Example
I	independent decision quantity	nothing; autonomy of the designer	Represent the design decisions	any	physical dimensions, materials, configurations of the ATBD
II	objective quantity	categories I, III, IV	Represent goals; the quantities in category I should be such that quantities in this category are optimized	Ordinal (numeric)	optimized performance quality, profit related to the ATBD
III	contextual quantity	nothing; beyond the autonomy of the designer		any	social and geographical circumstances; legislature; physical constants
IV	auxiliary quantity	categories I, III, IV.		any	internal properties of the ATBD; quantities that are affected by the ATBD but that are not mentioned the objectives.

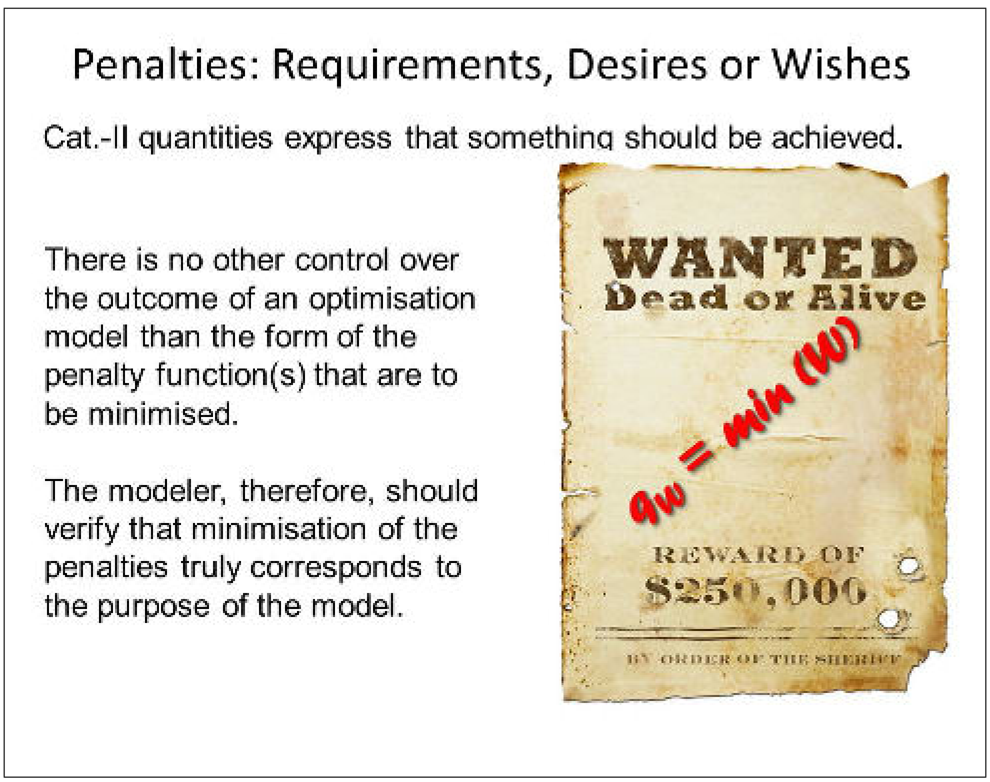
Table 5.1: A taxonomy of the four categories of quantities in a model

5.3 Executing a Functional Model

With the construction of a functional model using the four categories, we complete the formalisation phase. The model is now ready to start being executed. For some purposes, such as prediction or numerical verification, execution amounts to simply evaluating the DAG, obtaining values for the cat.-II quantities. For optimisation and supporting decisions, however, we should find a set of values for cat.-I quantities that yield cat.-II quantities that are in some sense optimal.

5.3.1 Optimisation in a Functional Model

For functional models of particular form and structure, we can use mathematical techniques. For instance, if the model comprises only a single cat.-II quantity, say q, and if all the occurring functions are differentiable on the entire cat.-I space, we can compute the partial derivatives $\frac{\partial}{\partial x_i}q$ with respect to all cat.-I quantities x_i. Next we demand all these partial derivatives to be zero, and solve for the unknown x_i.



This approach only works in rare circumstances, though: in modeling, the occurrence of non-differentiable functions is very common, and most sets of equations cannot readily be solved.

A somewhat more generic approach[15], again for a single cat.-II quantity, is to resort to numerical techniques such as STEEPEST DESCENT. This usually avoids having to do symbolic manipulation involving partial derivatives and the solution of sets of equations.

Since we don't want to rely on strong assumptions regarding cat.-II space, however, we seek for optimisation methods that don't assume anything about the number of cat.-II quantities.

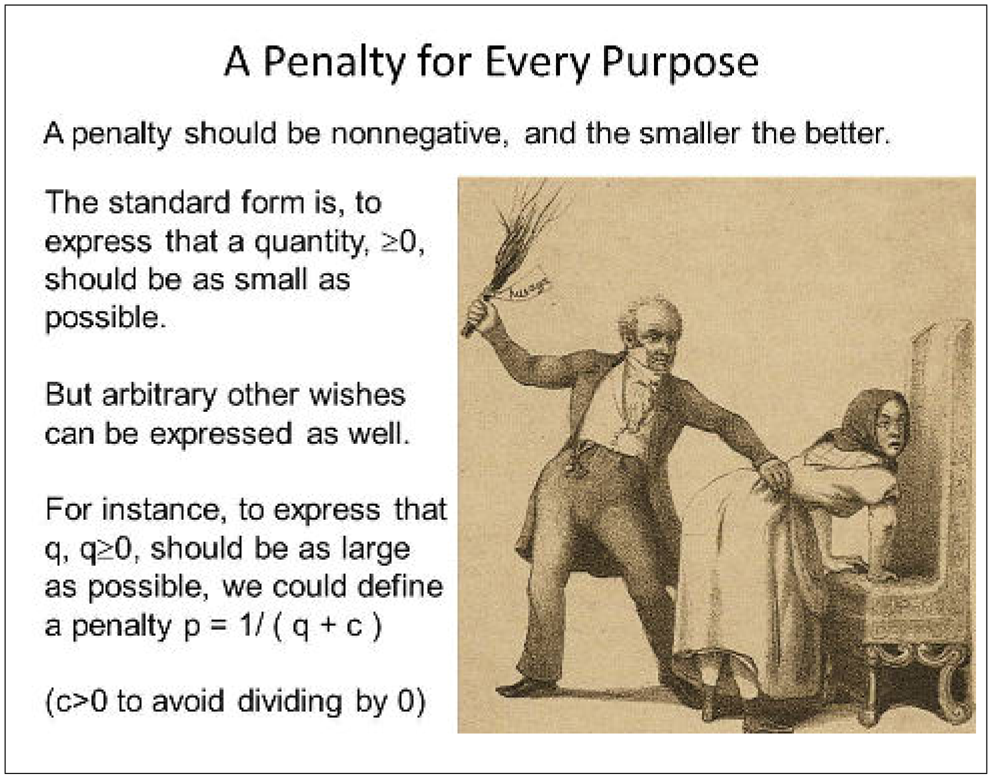
We will first look into lumping; next we will investigate methods that are suited to deal with trade-offs.

[15]The image of the dead-or-alive poster was taken from http://www.rgbstock.nl/photo/nVrincc/wanted+poster

Lumping and Penalties

In modeling for design, optimisation means that we seek a set of values for cat.-I quantities that make the ATBD 'as good as possible'. To express that an ATBD should be 'as good as possible', we introduce the notion of PENALTY FUNCTIONS .

A penalty (function) q_i is a quantity[16]depending on cat.-I quantities, that expresses how successful our ATBD is, in some respect. Penalty functions need to be nonnegative, and when their value is larger, this expresses that the ATBD performs worse with respect to the proprty q_i stands for.



Suppose, in a design case, that we want to express that the ATBD should be as light as possible, or that it should consume possibly little energy, that it imits possibly little CO_2, that it makes possibly little noise, ... All these requirements correspond to cat.-II quantities that are numerical, non-negative, and should be as small as possible. We call them q_1, q_2, q_3, etc. They all express some stakeholder's value for the ATBD. We see that all q_i are indeed penalties. We then can express the desire for optimality of the entire ATBD by replacing all separate q_i by one single Q, defined as $Q = \Sigma_i q_i$.

By requiring Q to be as small as possible, all q_i, to some extend, are required to be small. In the lucky case that $Q = 0$, the ATBD is optimal. This approach allows for arbitrary many preferred conditions. With each condition we associate a penalty function and hence a cat.-II quantity q_i.

In this way we can accommodate dilemma's like the cyclist's dilemma. There are various possible scenario's. In each of the following scenario two of the quantities (s,t,v) are in cat.-I. Also, there are two cat.-II quantities:

cat.-I: v, t; cat.-II: $q_s = \min(1/s)$; $q_W = \min(W)$; ; favors Tom and Harriet
cat.-I: s, v; cat.-II: $q_W = \min(W)$; $q_t = \min(t)$; ; favors Harriet and Dick
cat.-I: s, v; cat.-II: $q_t = \min(t)$; $q_s = \min(1/s)$; ; favors Dick and Tom,

[16]The drawing of a landlord, spanking his servant (as was legally allowed in Sweden until in the 19th centrury) was taken from http://commons.wikimedia.org/wiki/File:Husaga_(teckning_av_Fritz_von_Dardel).jpg?uselang=nl

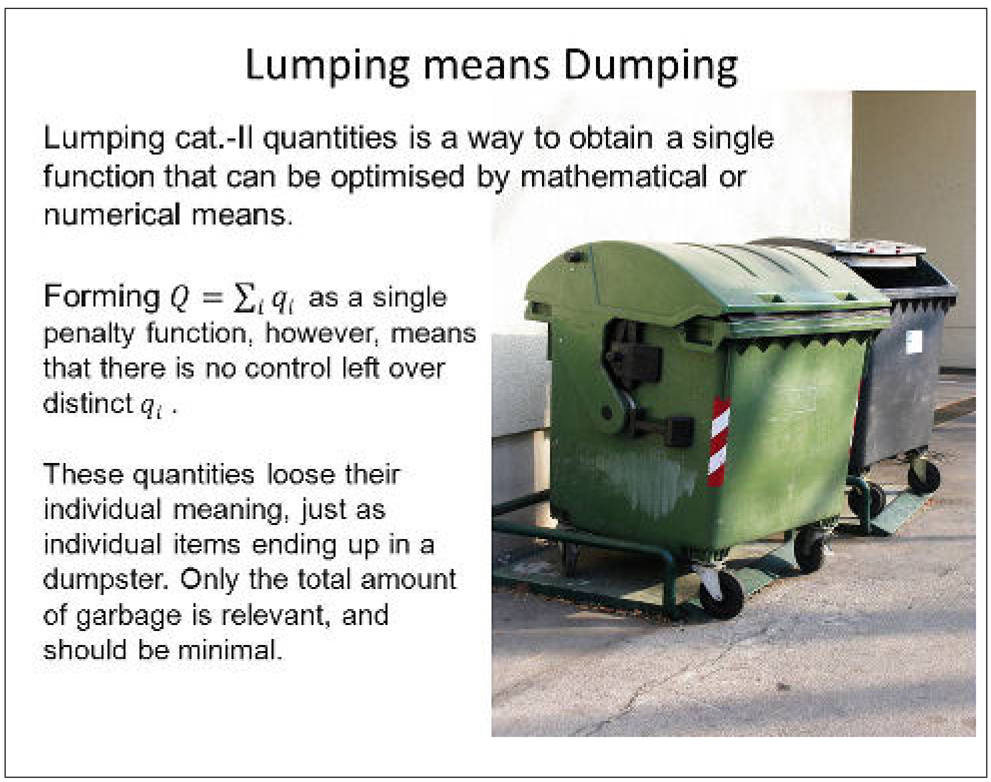
where each scenario favors one group of stakeholders. There is even a fourth scenarion that we might consider, if we want to please all stakeholders:

cat.-I: t, s, v; cat.-II: $q_s = \min(1/s);\ q_t = \min(t);\ q_W = \min(W)$; favors Tom, Dick and Harriet

We observe two peculiarities:

- A penalty should be such that 'the smaller, the better'. If for a quantity we want to express the intuition 'the bigger the better' (such as for the distance s), we can form a penalty by taking $\frac{1}{s}$ instead of s. With similar, usually simple transforms, we can express arbitrary wishes regarding the values of quantities in the form of penalties.
- It may seem strange that a quantity appears both as cat.-I and as cat.-II, as in the case where all three stake holders' preferences are taken into account, and s is both cat.-I and cat.-II. We should realize, however, that an identity function is also a function. In other words, by setting $s_1 = s_2$, we very well can have that s_1 is a cat.-I quantity and s_2 is a cat.-II quantity, despite that their values are identical.

Now in principle we could deal with all sorts of dilemma's and trade-offs by forming a penalty $Q = \Sigma_i q_i$. Let us examine if this is a good idea[17].



First, the substitution $Q = \Sigma_i q_i$ is not unique. We could also have set $Q = \Sigma_i q_i^2$, or some other non-negative combination. Further, combining penalties may be not mathematically meaningful. Indeed, some q_i may have dimensions. If dimensions of the q_i differ, adding is undefined. We might fix this with a form such as $Q = \Sigma_i w_i q_i$, with non-negative w_i. If the w_i are well-chosen, the terms $w_i q_i$ can be made to have equal dimensions. But this asks for the values of the w_i. Indeed, the w_i express the relative contribution of each of the q_i to Q. When we translate the design requirements to cat.-II quantities q_i, we only state that we want them all to be as small as possible. We don't say which are the most important ones. Now suppose we design an aircraft, and q_1 is the chance of a crash, expressed as expected times per year and q_2 is the number of times per year we need to do maintenance. Both are costly and

[17] The photograph of a dumpster is taken from http://www.rgbstock.nl/download/lusi/nnGlrma.jpg

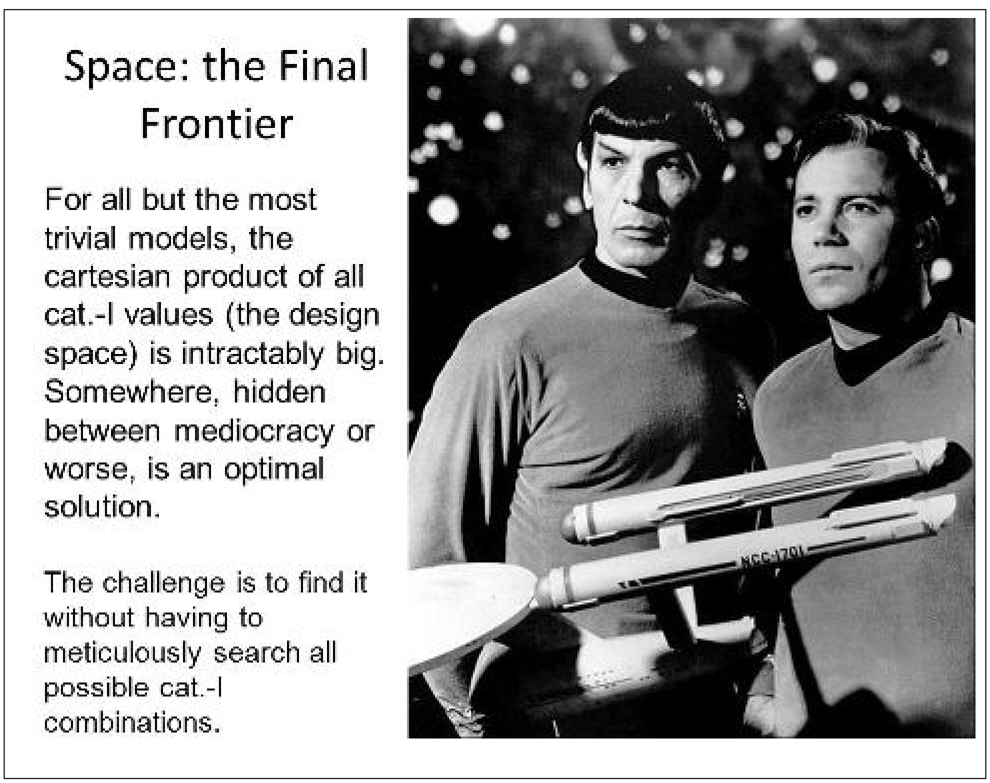
hence should be minimal. But an airline with $Q = 1000 \times q_1 + q_2$ is very different from one with $Q = q_1 + 1000 \times q_2$.

Nevertheless, LUMPING is common practice. In particular in cat.-II quantities relating to money. In the example of thermal insulation, we lump together the costs for heating and the costs for investment. We are only interested in the total cost. It seems obvious that both involved weights a_i are equal: both amounts are expressed in Euro's. But even in this simple case lumpting is problematic. As follows: investment is a number of Euro's; heating cost is a number of Euro's per year. Even if we know the duration of our stay in the house, we don't know how inflation will develop. A Euro of 2012 may be much more or less than a Euro of 2022.

Lumping always introduces additional assumptions. Each of the weight factors w_i corresponds to at least an assumption. These assumptions are difficult to formulate and even more difficult to verify.

The only alternative to lumping is having more cat.-II quantities. Unfortunately, having many cat.-II quantities is also disadvantageous, as we will elaborate in Section 5.3.1.

Searching Cat.-I Space



Whether a cat.-II condition is a *requirement*, a *desire* or a *wish* relates to the way we search design space[18]. One strategy (see Section 5.3.1) is to generate a population of candidate ATBD's and next incrementally improve subsequent generations of such populations. If conditions are requirements, this may decimate a population if individuals that fail for any requirement are taken out. It could even cause an entire population to vanish if it contains no individuals meeting all requirements, thereby prematurely stopping the process. Instead, we may consider to first replace a requirement by a wish. For example, a requirement $a_i = b_i$, for some cat.-II quantity a_i and constant b_i, can be replaced by a penalty term $c_i \times (a_i - b_i)^2$. This penalty term will try to get a_i closer to b_i. It is the wish that a_i and b_i be as close together as

[18]The photograph if renowned space explorers, captain James T. Kirk and dr. Spock, was taken from http://commons.wikimedia.org/wiki/Star_Trek:_The_Original_Series#mediaviewer/File:Leonard_Nimoy_William_Shatner_Star_Trek_1968.JPG

possible. The relative importance of this wish is given by c_i. Perhaps if we are lucky, there may be a candidate with $a_i = b_i$.

Wishes, expressed as penalties, rather than requirements, steer the search process into directions where $|a_i - b_i|$ is small. The strength of the tendency to approach solutions with small $|a_i - b_i|$ is controlled by the relative size of the c_i. Too many of such terms, however, will make the entire process less predictable. The advantage of an approach with penalties is, that we don't discard solutions 'underway'. There is, however, no guarantee that at the end the condition $a_i = b_i$ will hold exactly ▷95 .

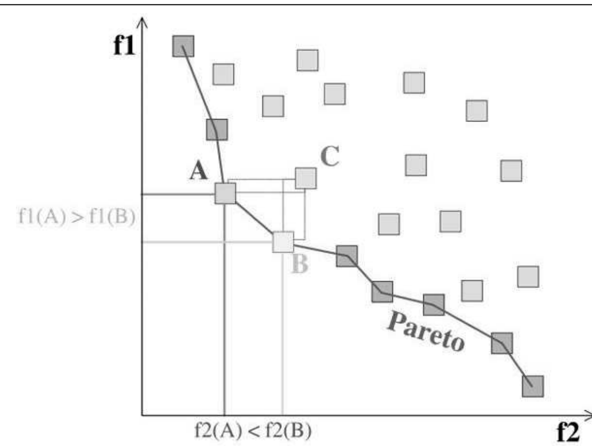
5.3.2 An Alternative to Lumping: Dominance

Every point in design space has n_1 coordinates where n_1 is the number of cat.-I quantities. A point in design space is a tuple with a property for each decision quantity. It corresponds to one possible ATBD. There are no other ATBD's than points in the design space.

We can also make a space spanned by the cat.-II quantities. Every point in this space is a tuple with n_2 coordinates, where n_2 is the number of cat.-II quantities. A point in the cat.-II space is a tuple with one property for each stakeholders' value. It corresponds to one ATBD. It may happen, though, that two different points in Cat.-I space correspond to the same point in cat.-II space. In general, not every possible element of the cartesian product of all cat.-II values is reached. We say that the functional model is not INJECTIVE, and it is not SURJECTIVE. An injective function preserves distinctiveness (any two different argument values produce different result values) and a surjective function reaches all values in the range set.

.

In cat.-II space, we compare ATBD's. To compare them, we use that all coordinates of cat.-II space are ordinal. So, for any cat.-II quantity, and for any two ATBD's, we assess if one of the two is better. Since the functional model is not injective, two different ATBD's, for some cat.-II quantity, can be equally good. For instance, for all cat.-II quantities that correspond to requirements, all acceptable solutions have the value TRUE.



The Pareto front.

Figure 5.1: A cat.-II space with two cat.-II quantities, f1 and f2. Both need to be minimal. All squares correspond to ATBD's. Items A and B are on the Pareto front; they are not dominated. Item C is dominated. (Illustration source: Wikipedia)

Comparing ATBD's relates to DOMINANCE. One ATBD, say x, *dominates* another ATBD, say y, if, for all cat.-II quantities p_i, $x.p_i$ is better than $y.p_i$. For instance, for cat.-II quantities speed and comfort, a Rolls Royce dominates a wheelbarrow.

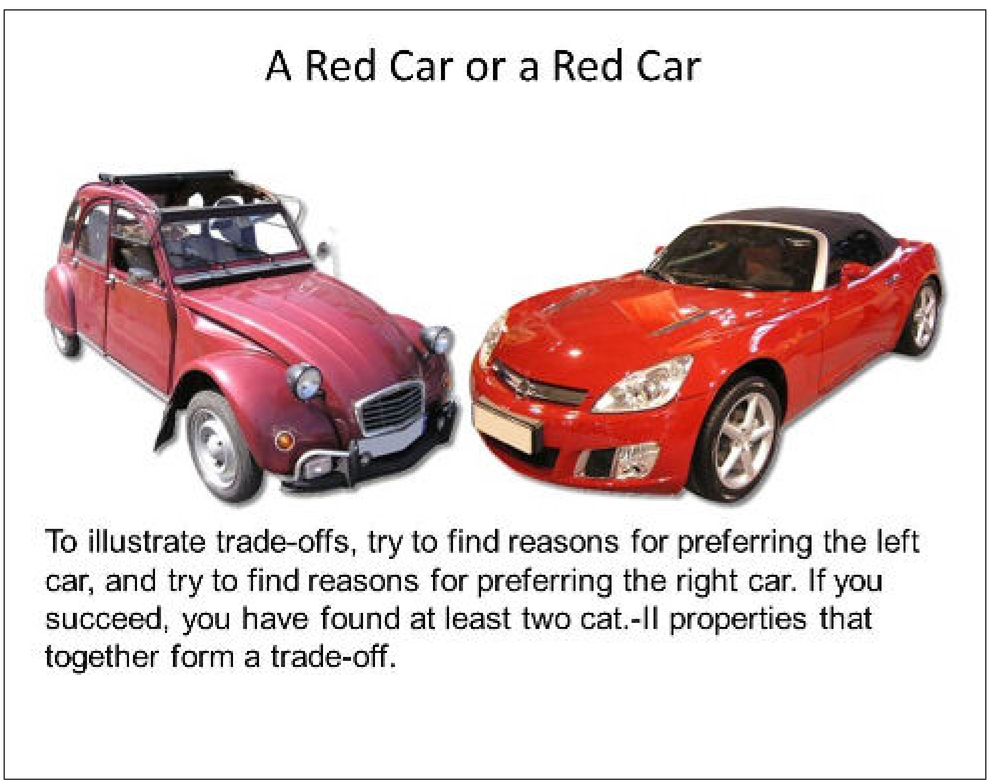
Dominance helps to reduce the collection of ATBD's to be considered. Obviously, an ATBD that is dominated should no tbe considered for realization. The design space contains another ATBD that is better *in all respects*. If cat.-II quantities capture the important stakeholders' values, a dominated solution cannot be the preferred solution.

So we should focus on the non-dominated solutions. The larger the number of cat.-II quantities, however, the larger the chance that any ATBD is non-dominated. Indeed, domination involves *all* cat.-II quantities. If ATBD's would be uniform randomly scattered in cat.-II space, the chance for any two ATBD's x and y that for some cat.-II quantity p, $x.p$ is better than $y.p$ is 0.5. So if we have n independent cat.-II quantities, the chance for an ATBD to be dominated is 2^{-n}. So, if n increases, the percentage of non-dominated solutions rapidly goes to 0. For sufficiently many cat.-II quantities, *all solutions are non-dominated*. This is one poignant reason to limit the number of cat.-II quantities.

We have to give good consideration about which, and how many, cat.-II quantities occur in the functional model. With too many, most solutions are non-dominated. Then there is no simple way to limit searching for the most appropriate solution to a restricted part of design space. On the other hand, reducing cat.-II space to few cat.-II quantities involves lumping, and the introduction of arbitrary weight factors and hard-to-verify assumptions. It is challenging to find the *few most relevant* cat.-II quantities in any design model.

For a first design model, 2 well chosen cat.-II quantities is often a recommendable start. With the additional advantage that a 2-D cat.-II space can be conveniently visualized.

Trade-offs and the Pareto Front



In practice, ATBD's are not scattered uniformly over cat.-II space. Rather, they form complex clusters. In most situations, cat.-II quantities will depend non-continuously on one or more of the cat.-I quantities. For instance, for the design of the box: the material choice is a *nominal* cat.-I quantity, so there is no notion of 'nearby' values. Therefore the mapping from cat.-I space to cat.-II space is by no means smooth. Two nearby points in cat.-I space may be mapped to remote points in cat.-II space, and vice versa.

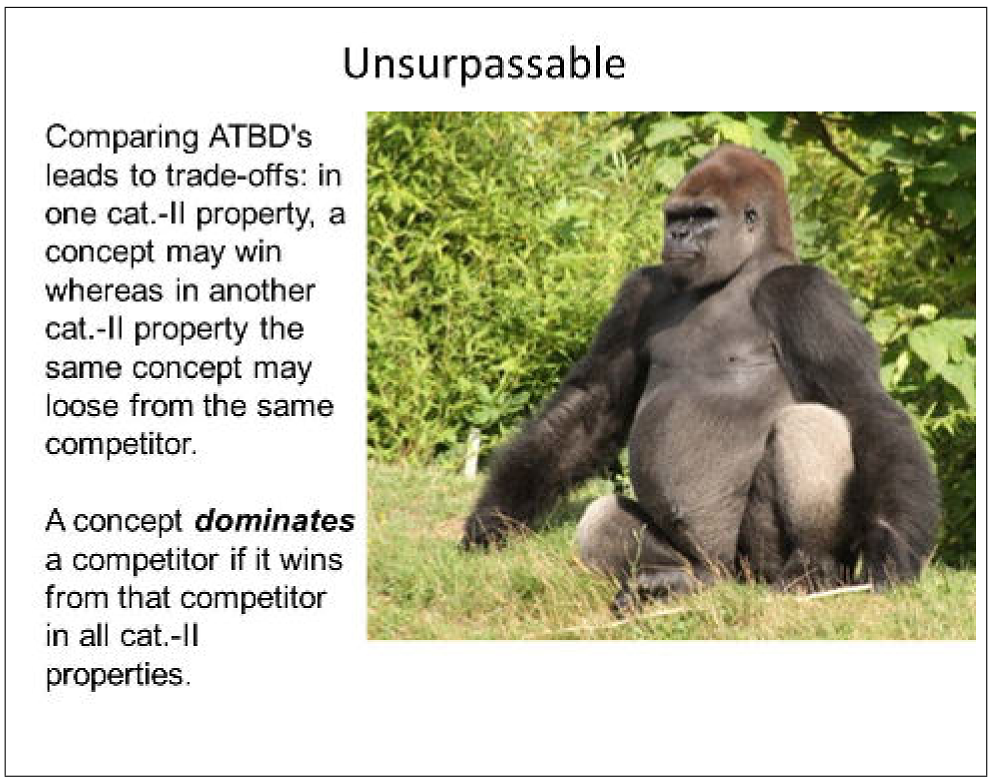
Still, there are some interesting geometric features in cat.-II space[19]. First assume that we have 2 real-valued cat.-II quantities, $f1$ and $f2$ that both should be minimal. Assume a large fraction of

[19] The photographs of the two red cars are taken from http://www.rgbstock.nl/photo/mhizZKo/Mooie+auto+2CV and http://www.rgbstock.nl/photo/mpSHQSC/Rode+auto

dominated solutions. There are few non-dominated solutions. These form the so-called PARETO FRONT, named after VILFREDO PARETO, an Italian economist and philosopher who studied the distribution of income and economic efficiency.

In Figure 5.1 we see some ATBD's, indicated by small squares. Most of them, such as C, are dominated. These need not to be considered for selection. ATBD's A and B are not dominated, and none of them is better than the other. They demonstrate the notion of TRADE-OFF: the situation where in A, $f2$ is better than $f1$ and in B, $f1$ is better than $f2$.

The 'curve', plotted in Figure 5.1, connecting the solutions A and B and others has no mathematical meaning. Since the functional model is generally not smooth, an environment in cat.-I space does not map to an environment in cat.-II space. A Pareto front such as in Figure 5.1 is merely a subset of the collection of ATBD's.



For any functional model there is a theoretical Pareto front. If we would inspect *all* points in cat.-I space, and for each ATBD compute the corresponding point in cat.-II space, we find points on a curve that forms a boundary of the attainable area in cat.-II space ▷96 . In case of n cat.-II quantities, these points are on a $n-1$-dimensional hyper surface in cat.-II space. Therefore it is better to think of the Pareto front as a hyper surface, even though for $n=2$ it is a 1-dimensional 'curve'.

The Pareto front contains solutions that are not dominated[20]. Moving *within* the Pareto front means exploring trade-offs. Moving in the direction *perpendicular to* the Pareto front means either considering solutions that are inferior or that are superior. The challenge is to construct successive approximations to the Pareto front that incrementally move in the direction of superiority.

Approximating the Pareto Front: Genetic Programming

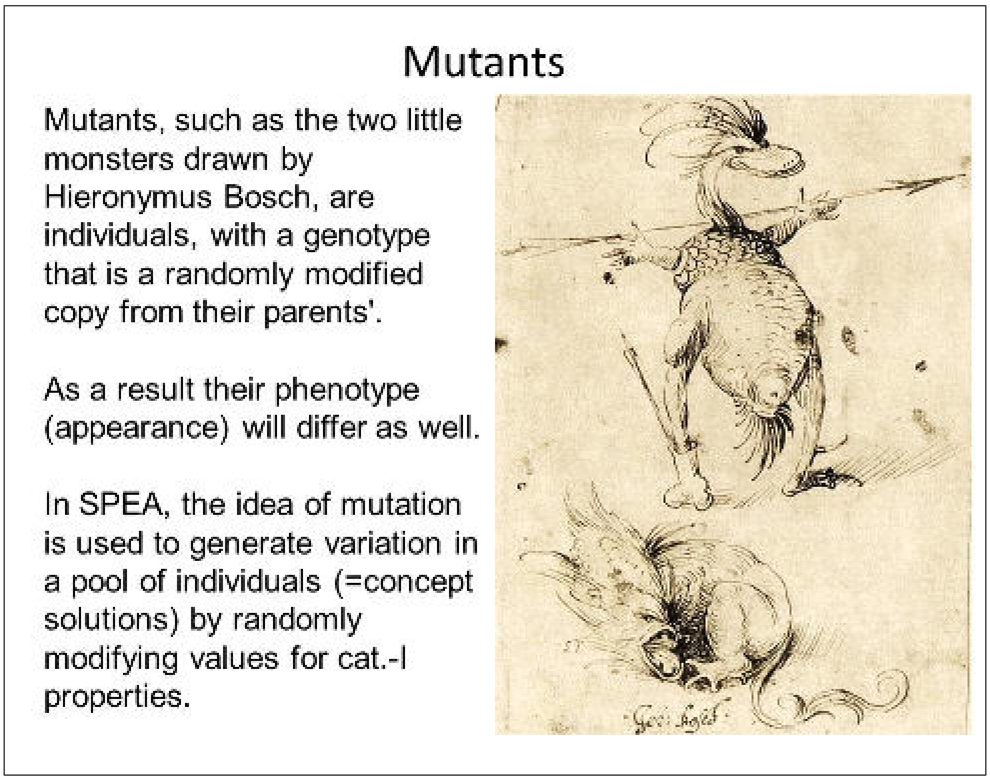
The number of possible model outcomes is typically very large and it is impossible to analyze them all. Furthermore, with multiple cat.-II quantities, we cannot totally order them. Therefore we resort to a technique called PARETO OPTIMISATION.

[20]The gorilla photograph was taken from http://www.rgbstock.nl/photo/o8WWgf2/Gorilla+Wildlife+5

Some features of Pareto optimisation are:

- It helps to reduce the number of possible solutions to those that are non-dominated;
- No weights have to be thought of, that is: much subjectivity is avoided and therefore discussions related to weight-factors can be postponed or avoided.

Although the Pareto technique has significant advantages for optimizing a model for obtaining 'best' cat.-I quantities, there is a drawback that should be well understood:



The purpose of the Pareto technique[21] is to reduce the collection of candidate solutions as much as possible. The smaller the fraction of non-dominated candidate ATBD's, the more benefit we have from the Pareto technique. As seen before, however, the chance that a solution is non-dominated increases with the number of cat.-II quantities. So the effectiveness of the Pareto technique reduces with increasing number of category-II quantities. In practice, with collections of solutions of no more than few hundreds, the number of category-II quantities should not exceed, say, 4.

Once an approximation to the theoretical Pareto front is constructed, negotiations with the problem owner are in place to agree on trade-offs. During negotiations, subjective preferences may occur, but at least the discussions can focus on the few solutions that are non-dominated.

The idea of Pareto optimality has been introduced as early as the 19th century. It has recently gained much interest in the design community ▷97 in the form of the SPEA-algorithm, short for Strength Pareto Evolutionary Algorithm. SPEA exploits the connection with EVOLUTIONARY DESIGN.

Evolutionary design is based on the metaphor of biological evolution.

All observable properties of an biological individual (or, in our case: the ATBD) are part of the PHENOTYPE of that individual. This is its manifestation in the biotope where it will have to compete for fitness. Therefore, the phenotype determines the individual's chance of survival. The attributes of the phenotype can be considered as the cat.-II quantities in a design case.

[21] The Bosch monsters were taken from http://commons.wikimedia.org/wiki/Hieronymus_Bosch#mediaviewer/File:TwoMonstersBosch.jpg

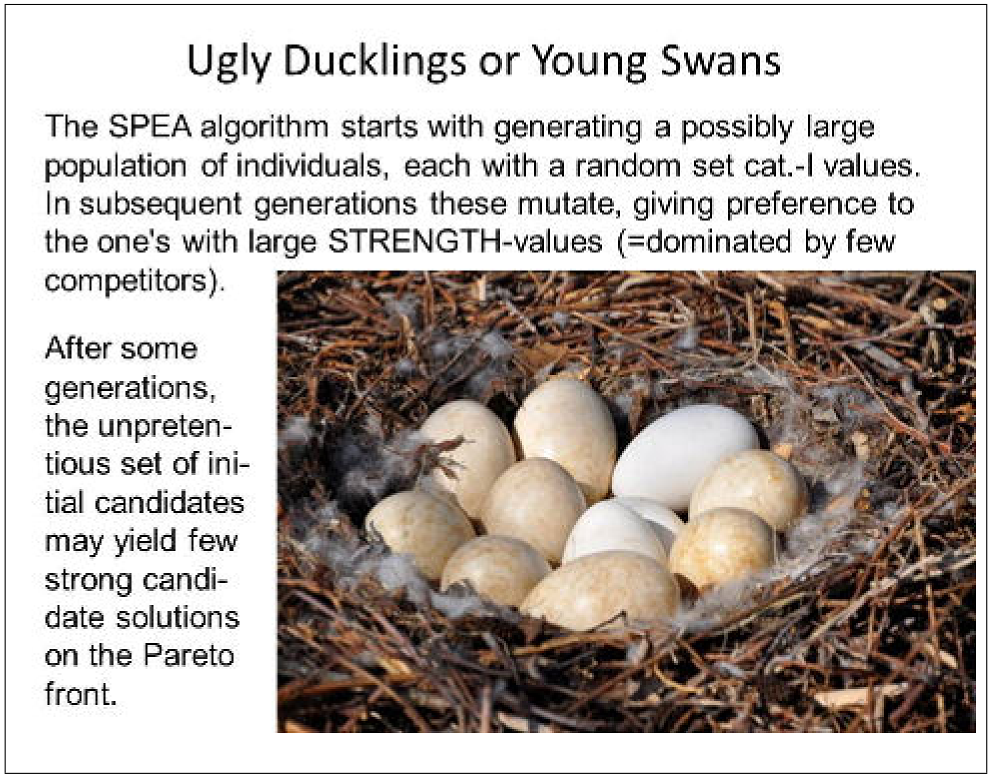
The phenotype is determined by the GENOTYPE and the surroundings of the individual, in much the same way as cat.-II quantities are determined by cat.-I quantities (representing the genotype) and cat.-III quantities (the biotope, or the context).

Evolution is driven by survival of the fittest. Better and stronger design solution(s) will survive after some selection procedure. In order to do so, Pareto-genetic optimisation follows the next steps, inspired by evolution theory:

1 Create a set of individuals, called a POPULATION, in a random way. That is: choose, for each individual, arbitrary values for all cat.-I quantities. Every individual is a candidate ATBD[22].

2 The fitness of individual X is related to the number of individuals by whom it is dominated. The lower this number, the 'fitter' we regard solution X. The number of dominating competitors relates to the quantity STRENGTH of which SPEA derives its name.

The fittest individuals are those that are not dominated at all, that is: those on the Pareto front. They have 0 dominating competitors. The number of dominating competitors for x is found by calculating the values of the cat.-II quantities for all individuals in the population.



3 Eliminate, randomly, a certain percentage of the non-fit ones. Do not eliminate them all, since there might be some latent strong properties, that may lead to fit individuals in future generation when combined with other values for other cat.-I quantities.

4 Generate additional new individuals using a number of different schemes ▷98 :

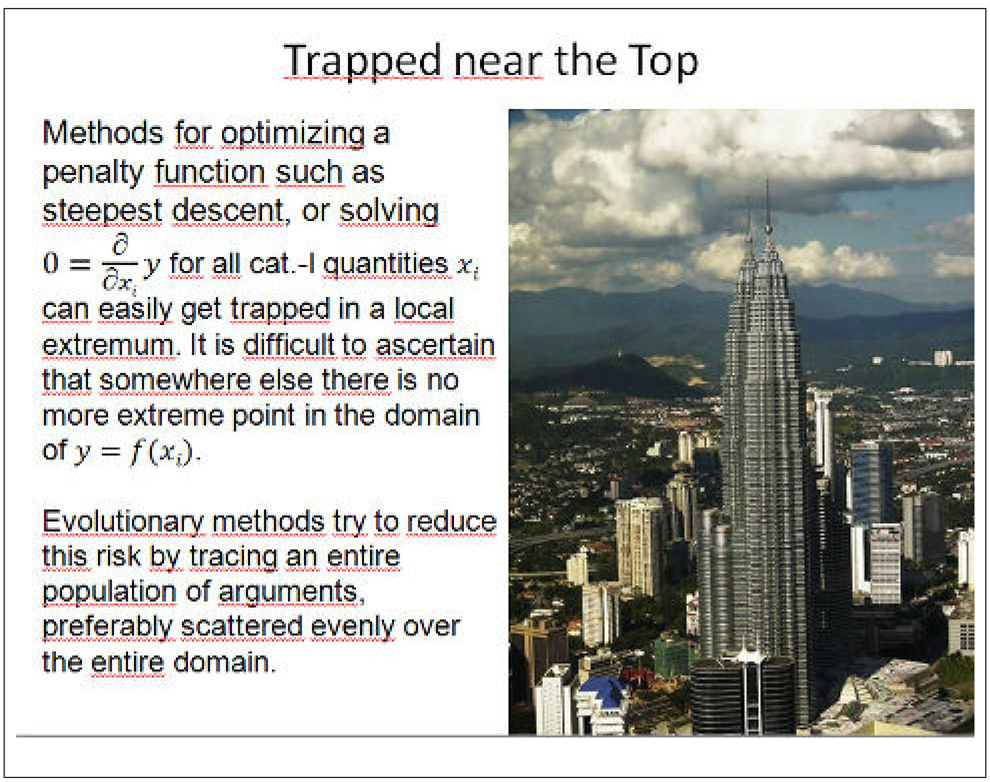
- Random mutations;
- Crossing over: combine successful genes (=values for cat.-I quantities) from two individuals (i.e., take the values of a number of cat.-I quantities from one successful individual and the values of the complementary ones from another successful individual);
- Re-introduce few of the eliminated candidate solutions: some of their attributes are given different values in the next population and therefore it is possible that an eliminated solution can come out better with these new attribute values.

[22] The photograph of a nest with swan's eggs was taken from http://www.rgbstock.nl/photo/o8LWirq/Swan+Nest+5

After step 4 return to step 1.

During the first few generations, the Pareto front can move as a whole in the direction of superior solutions. A new individual is labeled as being on the Pareto front only if it is not dominated. Individuals only leave the Pareto front when replaced by something better. Therefore, there is monotonous progress. As long as a sufficiently many individuals are not on the Pareto front, there is room for mutations by means of various mechanisms. The front then may continue to improve. There are three reasons for the front to cease improving:

- The approximated front is sufficiently close to the theoretical front. Further improvement cannot be obtained.
- If the majority of the solutions is on the Pareto front, there is not enough 'livestock' left for generating new mutants. Then we might consider increasing the population, or eliminating some of the individuals on the Pareto front. The latter jeopardizes the property of monotonous improvement. It is possible that we eliminate an individual that causes a 'hole' in the Pareto front. As a result, the Pareto front might locally deteriorate. In a more careful approach we should ensure that an individual is taken from the Pareto front *only* if there are other individuals that are sufficiently nearby to avoid significant deterioration.



- A third reason for the Pareto front to cease improving is, that the entire population occupies a 'niche' of the cat.-I space. This can be compared with a numerical algorithm for function optimisation that gets trapped in a local extreme[23], or a mountaineer that gets stuck on a local top.

Although the usage of an entire population rather than a single individual reduces the chance to get trapped in a local optimum, the size of the entire solution space could be so large that the used population gives insufficient coverage. Then we should either increase the size of the population ▷99, or increase the relative probability of random mutations.

[23] The photograph of the twin towers was taken from http://www.rgbstock.nl/photo/mh98QvE/Twin+Towers

Pareto: Post Processing

At the end of SPEA, there is a Pareto front that no longer improves. The Pareto front *as a whole* does not move in the direction of superior solutions. Improvements of *individual solutions* on the Pareto front, however, could still be possible.

Indeed, call an individual, resulting from a completed SPEA process, s. All the cat.-II quantities for s have values such that s is not dominated. These values are functions of all cat.-I quantities. For now, assume that all cat.-I quantities are real numbers, and also assume that the dependency of s on the cat.-I quantities is continuous. Then we can apply, for a cat.-I quantity v with value x, a small change to $x \to x + \delta x$. Next we check if solution $s(v = x + \delta x)$ dominates solution $s(v = x)$. If so, the solution with $v = x + \delta x$ is a *better* solution than the solution with $v = x$, and we replace the value for v by the new one. If not, we try $v = x - \delta x$.

If none of the two[24] gives a new s that dominates the old s, s is optimal with respect to v. It cannot be improved ▷100 . We repeat this process with all cat.-I quantities, until no further improvements can be obtained. An 'improvements' here means that a new solution dominates a previous one. We then have reached the situation where s is locally optimal. We remember that prior to the local optimisation, we could not ascertain that.

After applying the above process to all cat.-I quantities for all Pareto front-solutions may produce new cat.-II values to some of them. Further local optimisation is no longer possible ▷101 . We have *lost* the guarantee, however, that the collection of improved solutions are all non-dominated.

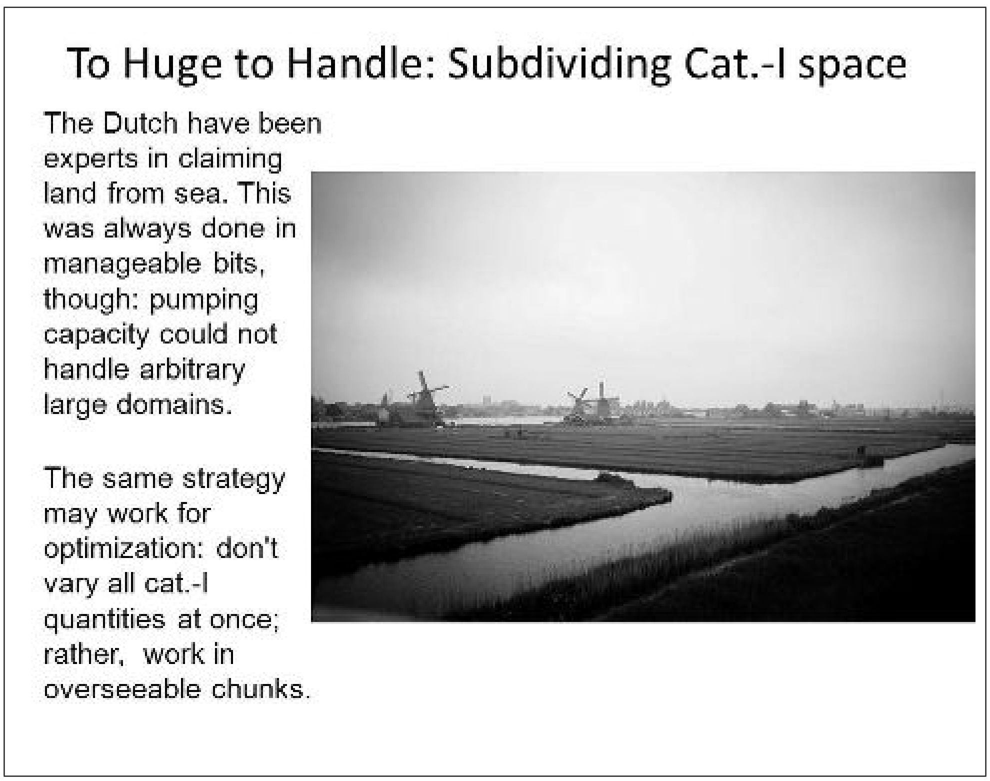
Indeed, it is possible that a single solution, say, s_1 has improved to the extent that it now dominates one of its former Pareto-colleagues s_2 that did not benefit from attempted improvement. Obviously, s_2 should no longer be labeled as being on the Pareto front. So after the local optimisation post process, we should do at least one further round of SPEA to ensure that the claimed Pareto-solutions are indeed non-dominated.

[24]The photograph of a hairdresser trying to outperform himself is taken from http://commons.wikimedia.org/wiki/File:Wartime_Hair_Dresser-_the_work_of_Steiner%27s_Salon,_Grosvenor_Street,_London,_England,_UK,_1944_D18216.jpg?uselang=nl

Practical experience shows that the application of this post process can give a spectacular improvement of the outcomes.

Last Resorts

It is possible that Pareto-evolutionary optimisation gives insufficient results. One route is to implement more advanced versions of the SPEA algorithm, or to apply more sophisticated heuristics for generating new individuals. This requires advanced, specialized theoretical skills that may fall beyond the average modeler. Before doing so, therefore, two simpler strategies should be tried:



Splitting cat.-I space Sometimes the scope of the functional model is too large[25]. That is: the functional model is supposed to govern a very large design space. This is in particular the case if we try to model INTEGRAL DESIGN. Integral design means that decisions in many different facets of the ATBD are taken simultaneously. In the example of thermal insulation of a house, we may want one model to estimate the savings due to insulating the roof, the windows and the walls. To deal correctly with trade-offs involving choices between these remedies, a single, integrated model is necessary.

It may be, however, that such an integral model is too complex to give a conclusive result with SPEA. An alternative is, to have three separate models. One deals with the insulation of windows; a second deals with the roof, and a third one with the walls. By comparing the results of these three separate models, we may find that, for instance, one of the three gives little contribution. Models with limited scope are simpler to build and debug, and it is simpler to ascertain their reliability. Furthermore, they may save the work of building a sophisticated model if a simple one already gives sufficient advantage.

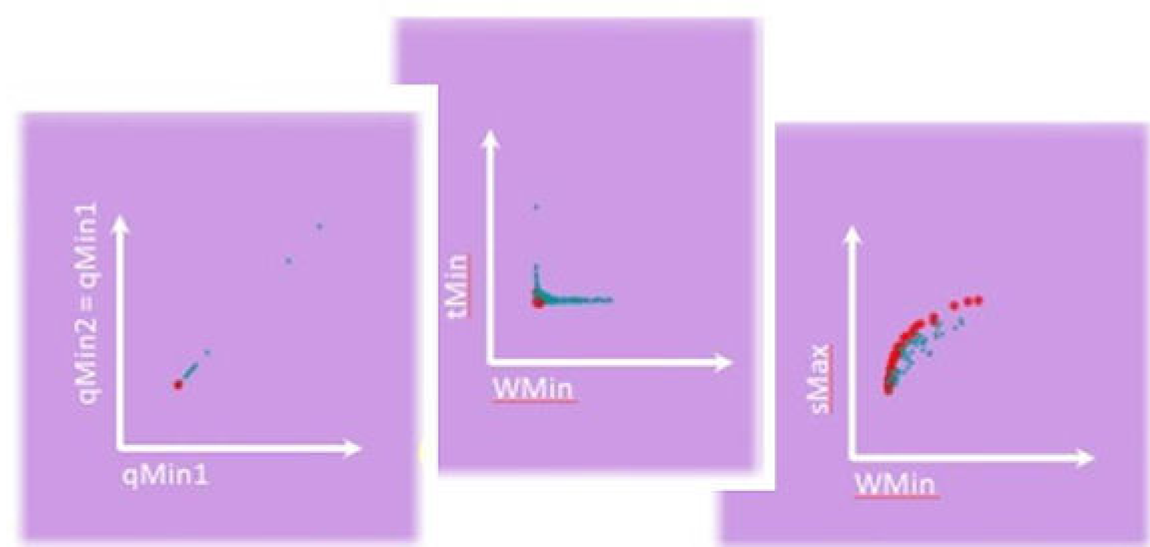
Fixing some cat.-IV quantities Cat.-IV quantities propagate changes in cat.-I quantities to cat.-II. The combined effect of various routes of influence can be difficult to comprehend, and it may lead to unexpected or unwanted behavior in evolutionary optimisation.

[25]The photograph of Dutch polder area 'de Zaanse schans' was taken from http://commons.wikimedia.org/wiki/Category:Polders#mediaviewer/File:Zaanse_Schans_(14320970627).jpg

It is sometimes a feasible to 'borrow' the value for a cat.-IV quantity from an existing situation, assuming that the behavior of the ATBD will be similar. In this way, we pretend that the quantities that depend on it are not affected by designer's choices. Formally, we treat such quantities as if they are in category III. If this strategy solves the optimisation problem, we can go back to the original dependency afterwards, and re-calculate the value of the original cat.-IV quantities, to get a more accurate estimate and redo the entire calculation.

What about the Optimal Cycling Trip?

We started this chapter with analysing the problem of finding an optimal cycling trip. We have learned that 'optimal' is defined in terms of penalties; the choice of penalties determines the resulting solution, or, in case of trade-offs, the resulting Pareto front. Let us see what the results are in various scenarios.



SPEA results for the Optimal Cycle Trip

.

Left: when we apply lumping, there is only a single optimum, but the outcome depends on the (arbitrary) weights. Middle: in some cases, e.g. by demanding t and W both to be minimal, there is no trade-off, and the Pareto front reduces to a single point (in this case, $t = W = 0$). Right: in case of a true trade-off (e.g., demanding s maximal and W minimal), a collection of solutions on the Pareto front results. These are the bigger dots; the smaller dots are dominated candidate solutions. The points on the Pareto front vary with respect to cat.-I quantities, v and t. Obviously, the choice t=v=0 corresponds to a point on the Pareto front: this is the lower left point with W=s=0.

In general it will be impossible to give an analytic expression for the shape of the Pareto front. In this (simple) case, however, we can find e.g. the minimal W for a given s. Indeed, $W = Fs$, and F does not depend on s, so the minimum is attained for minimal $F = c\rho Av^2$. That is, the slowest speed we can cycle gives us the least effort. This is, in hindsight, not too unexpected. Notice that we cannot drive with $v = 0$; the model does not contain any mechanisms, though, that dicate what the lowest value for v depends on. v_{min} is therefore in cat.-III.

5.4 Mathematical preliminaries

materiaal aan te vullen door Emiel ...

...

...

...

5.5 Summary

- A *functional model* helps to distinguish input and output;
  - Desired *output* follows from purpose;
  - *Input*, if present, represents modeler's or designer's choices;
- Building a functional model in the form of a DAG according to Section 5.2.1 reveals the *roles* of quantities. These are:
  - *Cat.-I*: free to choose;
    - Models for (design) decision support: the notion of *design space*;
    - Choice of cat.-I quantities: no *dependency-by-anticipation*;
    - The necessity of avoiding causal loops;
  - *Cat.-II*: represents the intended output;
    - The advantages and disadvantages of *lumping* and *penalty functions*;
    - The distinction between *requirements*, *desires*, and *wishes*;
    - The notion of *dominance* to express multi-criteria comparison;
    - The *Pareto-front* as the hyper surface of non-dominated points in cat.-II space;
  - *Cat.-III*: represents constraints from context;
  - *Cat.-IV*: intermediate quantities;
- For purpose *optimisation*, an *evolutionary approach* can be used to deal with arbitrary functional models and multiple cat.-II quantities;
  - The Pareto front can be approximated by evolutionary approach using the *SPEA*-algorithm;
  - Local search can be used for post-processing.

5.6 Learning goals

5.6.1 Knowledge

You should know the meaning of the terms cat.-I, cat.-II, cat.-III, cat.-IV; you should know the various ways a model can be used to express an intention (requirements, desires, wishes). You should know what penalties and penalty functions are, and the advantages and disadvantages of lumping. You should know what dominance means, and you should the role of the Pareto front in relation to optimisation. You should know what to do to approximate the Pareto front by means of genetic programming, you should know the advantages and limitations of genetic programming, and you should know some measures to deal with common issues in SPEA.

Regarding the mathematical notions in this chapter, Emiel: aanvullen svp.

5.6.2 Skills

Once you have obtained a formal model that requires optimisation, you should be able to identify the roles (cat.-I ... cat.-IV) for each of the occurring quantities. You should be able to make a choice for a method to execute the optimization (analytic means, numerical means using lumping, genetic programming), and you should be able to implement this method. You should be able to identify trade-offs, and you should be able to interpret (numerical) outcomes in terms of trade-offs, and in terms of the initial purpose of your model. You should be able to execute a genetic algorithm (you don't need to be able to implement one yourself), and to apply post processing once it converges. You should be able to modify the model (e.g., split cat.-I space or temporarily fice the values of some cat.-IV quantities) in case the results are not sufficient.

For the mathematical notions in this chapter, Emiel: aanvullen svp.

5.6.3 Attitude

When confronted with an optimisation problem, you should be inclined to seek for the network of dependencies among quantities, and you should be inclined to choose the appropriate ones for optimization. You should think abount the problem in terms of trade-offs, and various definitions of optimality. In a design problem, you should challenge the values cat.-III quantities, and attempt to move some from cat.-III to cat.-I. You should be inclined, after having executed an optimization algorithm, to try and improve the results by some of the techniques introduced above.

5.7 Questions

1. Explain in your own words what a functional model is.
2. In decision models, what is the input and what is the output?
3. Explain in which circumstance the purposes 'verification' and 'prediction' are similar with respect to input and output.
4. We say that 'often the output of a functional model represents an intention'. Explain.
5. When is it necessary that outputs, corresponding to intentions, are ordinal?
6. Explain why the book printers' dilemma is a dilemma (or better: a trilemma).
7. The book printers' example starts from the relation $TS^2 = PA$. This relation is only valid under a number of assumptions and approximations.
   (a) List at least three assumptions.
   (b) List at least three approximations.

(c) What is the difference between an assumption and an approximation?

(d) Verify if (1) assumptions can always be replaced by approximations, or (2) if approximations always can be replaced by assumptions.

8. Redo the analysis of the 3 cases in the book printers' dilemma where S has the dimension of area in stead of length.

9. In your own words, define category I.

10. In your own words, define category II.

11. In your own words, define category III.

12. In your own words, define category IV.

13. Explain that cat.-IV quantities always can be eliminated from a functional model. Why is this often not a good idea?

14. Cat.-II quantities are always immaterial. Explain.

15. What is a cartesian product? Explain that cat.-I space is a cartesian product.

16. For the sets with elements 1,2,3 and with elements a,b, the cartesian product is {[1,a], [1,b], [2,a], [2,b], [3,a], [3,b]}. Notice we use both square brackets and curly brackets. Explain.

17. What is 'dependency by anticipation'?

18. In Section 5.3.1 we say that an airline with $Q = 100q_1 + q_2$ is very different from one with $Q = q_1 + 1000q_2$. Explain the difference.

19. In Section 5.3.1 we say that an airline with $Q = 100q_1 + q_2$ is very different from one with $Q = q_1 + 1000q_2$. Is this always true?

20. What is a causal loop?

21. What is a proposition?

22. What is a predicate?

23. Explain the difference between requirement, desire, and wish.

24. Why is there a dilemma between lumping or not lumping cat.-II quantities?

25. Give a formal reason why a penalty function of the form $\Sigma_i q_i$ is problematic.

26. In a logical expression, $P = TRUE$ we can leave out the part '$= TRUE$'. Why?

27. A banana is yellow. Write this as a proposition, as a predicate and as a relation between an attribute and a value.

28. Why do cat.-II quantities for design need to be ordinal?

29. We want to impose a condition $q = 4$ on a cat.-II quantity q. Express this as a penalty function.

30. We want to impose a condition $q > 4$ on a cat.-II quantity q. Express this as a penalty function.

31. We want to impose the condition that q should be as large as possible. Express this as a penalty function.

32. We want to impose a condition $q_1 = 4$ and a simultaneous condition $q_2 = 5$ on cat.-II quantities q_1 and q_2. Express this as a single penalty function in at least two different ways. Discuss which penalty function should be used to achieve what.

33. What is cat.-II space?

34. Which of the two, cat.-I space or cat.-II space is meant by 'design space'? Why?

35. Explain in your own words what 'dominance' means.

36. For an increasing number of cat.-II quantities, the chance to get dominated decreases. Explain.

37. What is the Pareto front?

38. We state that 'an environment in cat.-I space does, in general, not map to an environment in cat.-II space'.

    (a) Why not?

    (b) Under what condition *does* an environment in cat.-I space map to an environment in cat.-II space?

39. We state: 'moving in the surface [of the Pareto front] means exploring trade-offs'. Restate this in your own words.

40. We introduce an example of thermally insulating a house. What is the point we try to make with this example?

41. Give two reasons why we want few non-dominated points in cat.-II space.

42. What is the fitness function in SPEA?

43. Explain the necessity for post processing in approximating the Pareto front.

44. Explain how post processing works.

45. Explain why, after post processing, there should be at least one more iteration of SPEA.

46. Post processing only works for certain cat.-I quantities. Which?

47. In case SPEA gives no satisfying results, we suggest two 'last resorts'. Rephrase them in your own words.

5.8 Exercises

1. Think of a situation where the detergent problem from Section 4.3.1 is a model with purpose 'optimize' in the context of designing.
   (a) Formulate the required cat.-II quantities;
   (b) what could be reasonable cat.-I quantities? Motivate your choice;
   (c) formulate the model (this may involve minimal extension of the original detergent-model);
   (d) implement the model, e.g. in ACCEL and run SPEA.
   (e)
2. Think of a situation where the chimney sweepers problem from Section 4.3.2 is a model with purpose 'optimize' in the context of designing.
   (a) Formulate the required cat.-II quantities;
   (b) what could be reasonable cat.-I quantities? Motivate your choice;
   (c) formulate the model (this may involve minimal extension of the original detergent-model);
   (d) implement the model, e.g. in ACCEL and run SPEA.
   (e)
3. The peanut butter problem from Section 4.3.3 can be seen as a model with purpose 'optimize' in the context of designing.
   (a) What quantity is in cat.-II?;
   (b) what quantity is in cat.-I?;
   (c) in the original form, there is only a single cat.-II quantity. Use standard mathematical optimisation to solve the optimisation problem:
      i. choose a ramp function for the price elasticity and use symbolic mathematics;
      ii. choose a suitably adjusted logistic function and use numerical mathematics;
      iii. compare the outcomes of the two methods. Which do you trust more, and why?
   (d) make a simple but meaningful extension to the model such that there will be a second cat.-II quantity;
   (e) implement the extended model, e.g. in ACCEL and run SPEA.
   (f)
4. Implement a model for the printer's dilemma:
   (a) use a penalty function that combines the interests of the three stakeholders (backpackers, senior citizens and publisher);
   (b) use a set-up with 3 cat.-II quantities and SPEA;
   (c) draw conclusions from comparing the outcomes of the two approaches.

5. We introduce the problem of thermal insulation, but we don't develop a full model in the text.

   (a) develop a model where the window size is in cat.-III;
   (b) develop a model where the window size is in cat.-I;
   (c) implement and run both models, assuming a single cat.-II quantity (=the total cost for heating, including costs for gas and costs for investment)
   (d) Compare the outcomes of the two models; interpret the results.

6. Cat.-II quantities for design decisions need to be ordinal. What about other purposes? Give arguments for your answers, possibly with examples.

7. We claim that a Rolls Royce doesn't dominate a wheelbarrow. Suggest a minimal set of cat.-II quantities such that this claim holds.

8. The number of cat.-II quantities is a delicate balance between two arguments. Give these these two arguments.

9. Simultaneous optimisation of multiple cat.-II quantities is not supported by standard numerical optimisation, because such methods assume a single function. What is the quintessence of SPEA such that we can do such optimisation? What are the disadvantages of SPEA compared to numerical optimisation of a single function?

10. What is the quintessence of the SPEA algorithm?

11. In case the fraction of dominated solutions in cat.-II space is too small, we have to intervene.

    (a) Why?
    (b) What does this intervention amount to?
    (c) There is a potential risk with this intervention. Which?

12. When we describe the method to approximate the Pareto front by genetic means, we propose three different schemes.

    (a) Which schemes?
    (b) For each scheme, consider that it would be the only one. What would be the disadvantage of only using this scheme?
    (c) For each scheme, think of a reason why you would want this scheme to have a high(er) probability.

13. We give three reasons why the SPEA algorithm could cease improving the Pareto front.

    (a) Which are they?
    (b) How can you assess which of these reasons applies in a concrete situation?

14. The taxi model is fully spelled out in Appendix **??**; it is available as one of the demo's in ACCEL.

    (a) Run SPEA on the original model; summarize the outcome in few characteristic solutions;

(b) find relations between cat.-I quantities such that the corresponding cat.-II quantities are on the (approximated) Pareto front. To do so, you may have to fix values for the nominal cat.-I quantities. Express these relations by of fitting the simples possible dependencies (preferably linear dependencies).

(c) extend the model by including some hybrid car types and/or a simple maintenance scheme. Make sure that the terms you add have an influence to `profit` via at least two routes with opposite effects. For example: maintenance gives longer life expectancy for the cars, but it also costs money.

(d) see if, with the extended model, you can find a taxi company that dominates the original one.

Chapter 6

Models and Confidence

'Trust in Providence, but tie up your camel' (inspired by an ancient Arabic proverb)

Since the beginning of human culture, man has looked up to the heavens and wondered about what he saw. Soon he began telling stories, relating the sun, the moon and the stars to more earthly things, in an ongoing attempt to get more satisfying answers to the questions that came up. He started making models of the kosmos, and these models got increasingly advanced. They can be seen as answers to questions, the questions being the purposes of the models. Throughout the ages, both questions and answers progressed. Unfortunately, the questions developed faster than the answers. We have now reached the stage where only for some 20% (some authors claim: 4%) of what we believe constitutes the kosmos, we have a rational explanation; the rest is mysterious dark matter or dark energy. One could therefore say that, as far as our confidence in kosmos models is concerned, despite the growing sophistication of thinkers through the ages, the ratio between satisfying answers and open questions has constantly gone down to reach a historical low at the beginning of the 21st century.

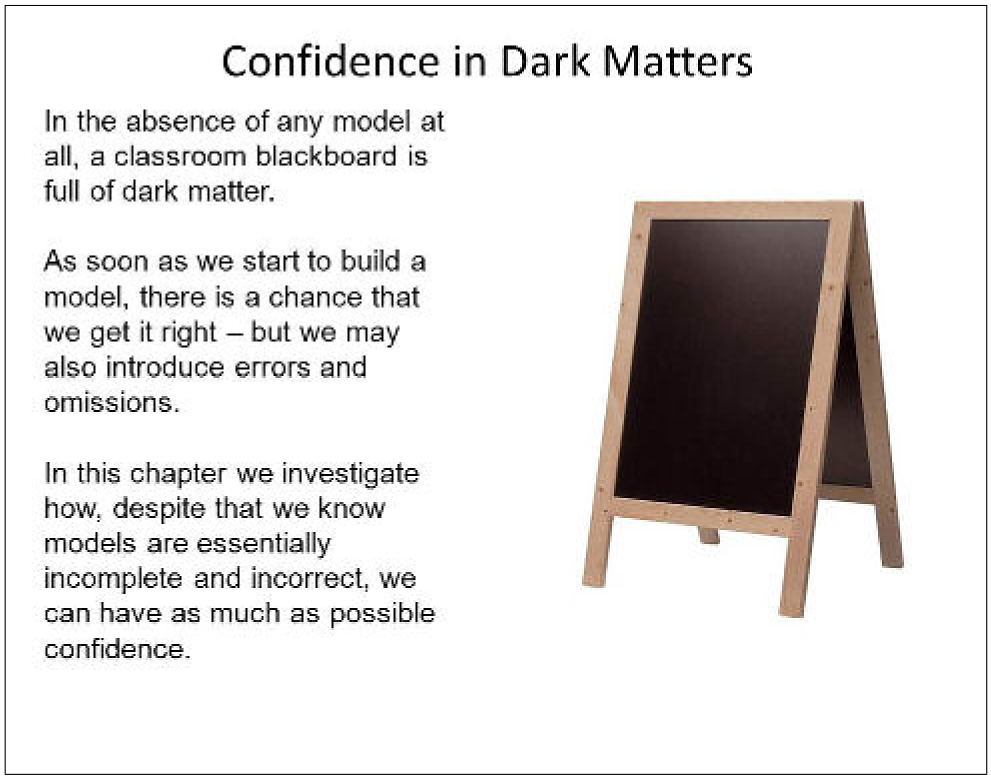
6.1 What do we mean by 'Confidence' in a Model?

The above argument is a biased view to the advances of cosmology. For instance, present theories about the motion of celestial bodies lead to accurate predictions of astronomical constellations, and we believe to understand much about the formation of planets and stars. It seems unfair to judge the ingenuity of current cosmological models and theories merely by their inability to account for the amount of assumed matter in the universe [▷102]. We must be more careful with defining the qualities of a model. A model, such as an astronomical model can have *multiple* qualities. Qualities of models, as discussed in depth in Chapter 7, include distinctiveness, genericity and scalability. Astronomical models score very good on these criteria despite that they fail to explain most of the matter in the Universe.

We see that we should not judge a model in isolation. Whether we have confidence in the outcome of a model depends on:

- *the model* ;
- *the modeled system* ;
- *the purpose of the model* with respect to the modeled system.

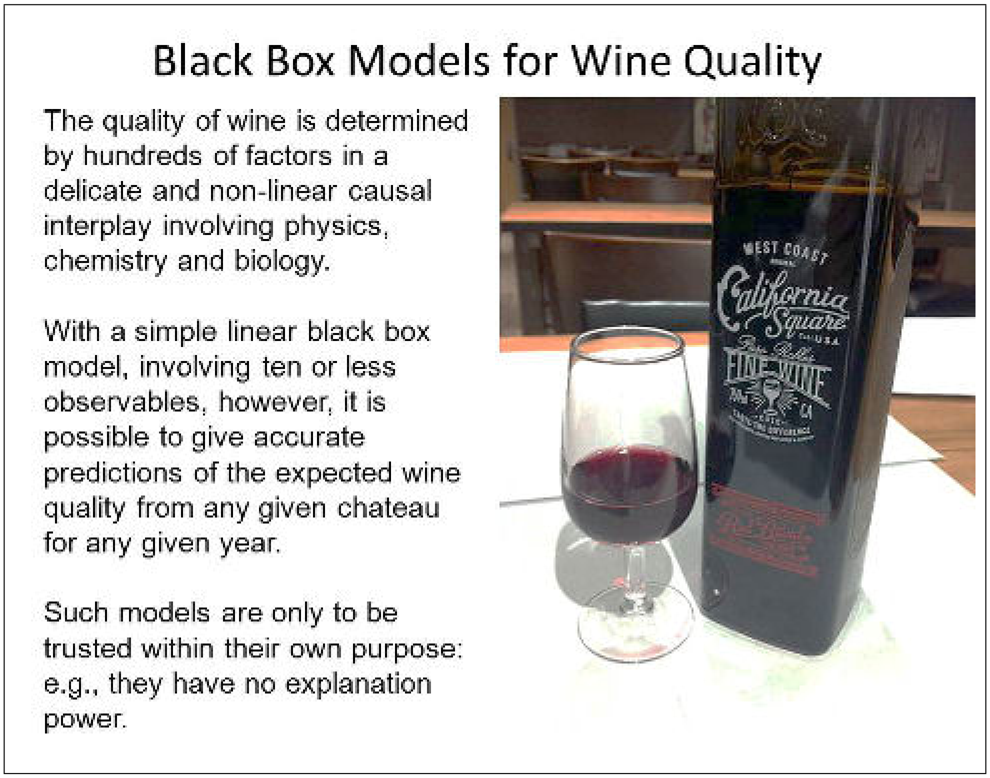
To discuss the merits of a model[1] all three perspectives need to be taken into account. We give examples:



- A model can be easy to develop, e.g. because no advanced mathematics are needed to get a model outcome. This is of little worth if the correspondence of the model to the modeled system is low, or if the outcomes have little practical use. This applies to most traditional secondary school physic problems, involving idealized systems like rigid slopes and massless pulleys, or the static liquids in communicating vessels. With simple mathematics, we achieve exact values for the time a point mass moves down a rigid slope, but there are few systems for which such a model is relevant. If the purpose of such a model is to explain basic principles of physics the purpose is fulfilled. If the purpose is to ascertain e.g. the safe launch of a multi million kg ship from the slide in a shipyard, the purpose is not fulfilled.
- A database containing the recorded events of a group of Internet users, e.g. while communicating with a server over some period of time may be fully accurate. That is: it corresponds with large precision to real, historical occurrences. If we merely want to archive these data to answer as much as possible future questions ▷103 , the purpose may be fulfilled. If, on the other hand, we need to analyze the cause for an acute performance bottleneck in Internet traffic the approach may be inadequate, and we should look into condensed, aggregated representations of the data instead.
- A simple linear black box model, aggregating a small number of empirical data points, can be a

[1]The photograph of a blackboard was taken from http://www.rgbstock.nl/photo/nzcgvbQ/Schoolbord

useful aid to taking decisions, even though it contains nothing that represents the internals of the modeled system. For example ▷104 , wine making is governed by subtle biochemistry, involving hundreds of enzymes, bacteria and process quantities such as soil, weather and harvest conditions. A glass box model would involve hundreds of coupled partial differential equations. A simple black box model, restricted to ten or less carefully chosen quantities statistically outperforms experienced connoisseurs in predicting the quality of a wine from a given chateau in a given year. But such a black box model fails if we try to explain the occurrence of a vanilla flavor in that same wine[2].



To further discuss 'confidence', we have to have some terminology.

6.1.1 Verification and Validation, Accuracy and Precision

In the literature, we encounter two terms to talk about the confidence in models: VERIFICATION and VALIDATION.

From etymology, the Latin root 'veritas' (for 'truth') occurs in 'verification', and the root 'valiant' (for 'strength') occurs in 'validation'. The core notions of the words 'validation' and 'verification', are 'strength' and 'truth'.

Verification: is model X correct?

Verification amounts to answering: 'Is model X correct?', in other words: is it 'true' enough to be convincing?

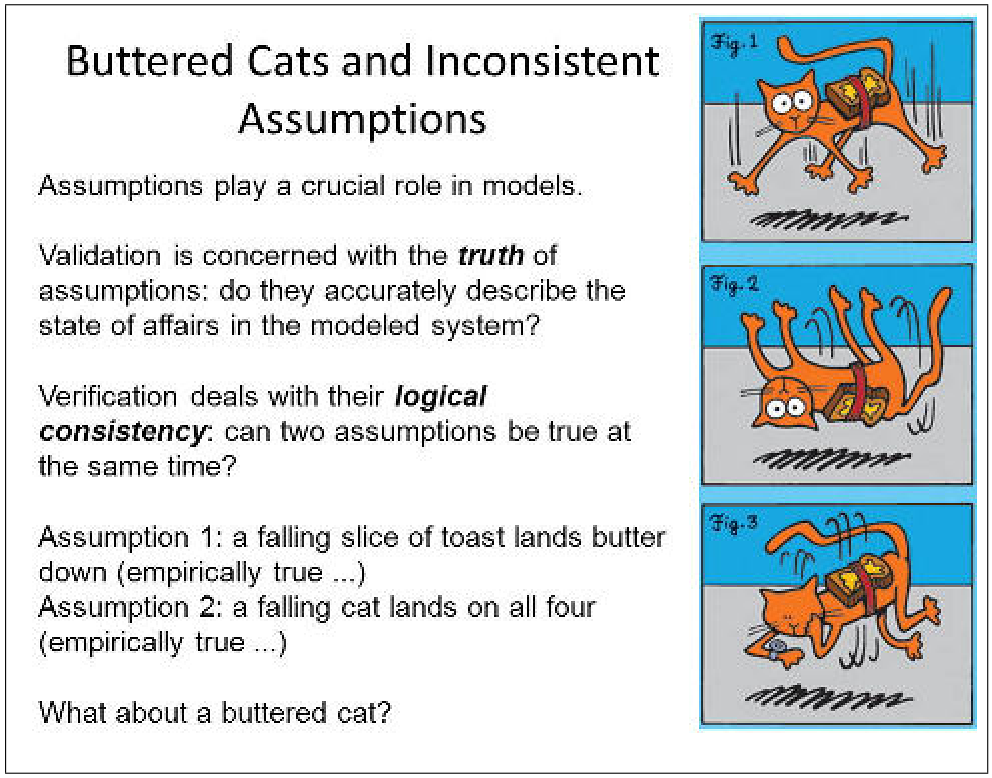
Verification takes the internal consistency of the model into account, in particular the consistency between the conceptual model and the formal model.

Examples of consistencies, checked in verification, are the dimensional correctness of expressions, and the check that values of quantities, obtained by execution of the model, fall within the admitted sets of values (types) given during the conceptualization. For instance, if a property

[2]The photograph of a box of quality wine was taken from http://commons.wikimedia.org/wiki/File:Paso_Robles_red_blend_unique_wine_bottle.jpg?uselang=nl

`x.p` of concept `x` represents a probability in the conceptual model, its type is the range $\{0 \cdots 1\}$. If the numerical value of `x.p` turns out to be 1.5, the *verification* fails. If it is 0.95, but we have empirical information from the modeled system that the chance should be less than half, the *verification* may succeed, but *validation* fails.

A further example of verification is, that a functional model should be a DAG.



Finally, verification relates to checking the logical structure of assumptions[3]. That is: checking if conclusions are logically justified from the assumptions. Here, two necessary conditions apply: (1) all assumptions must be true - which is part of the validation ▷105 , and (2), there should be sufficiently many assumptions such that formal expressions uniquely follow from the assumptions - which is part of the verification.

Both conditions are easily fulfilled separately: condition (1) can be fulfilled by having only unproblematic assumptions, and (2) can be fulfilled by having assumptions of a form equal to the formal expressions. For instance: 'if we assume that x is the sum of y and z, then we may set $x = y + z$'.

The problem, of course, is that (1) and (2) *together* may be difficult to achieve. As an example, consider the peanut butter problem from Section 4.3.3. One of the functions in the model is the expression for the market share,

$$\texttt{marketShare} = \texttt{fRamp(price,cheapestCompetitorsPrice,mostExpensiveCompetitorsPrice,1,0)}.$$

If we include this expression in the model under the assumption 'the market share is a decreasing function of `price`', condition (1) is rather unproblematic, but condition (2) clearly does not hold: from the fact tht a function should be decreasing between 1 and 0, it does nof follow that it is the proposed ramp function, as there are many other functions than the ramp function that are decreasing between 1 and 0. The assumption is clearly not sufficient.

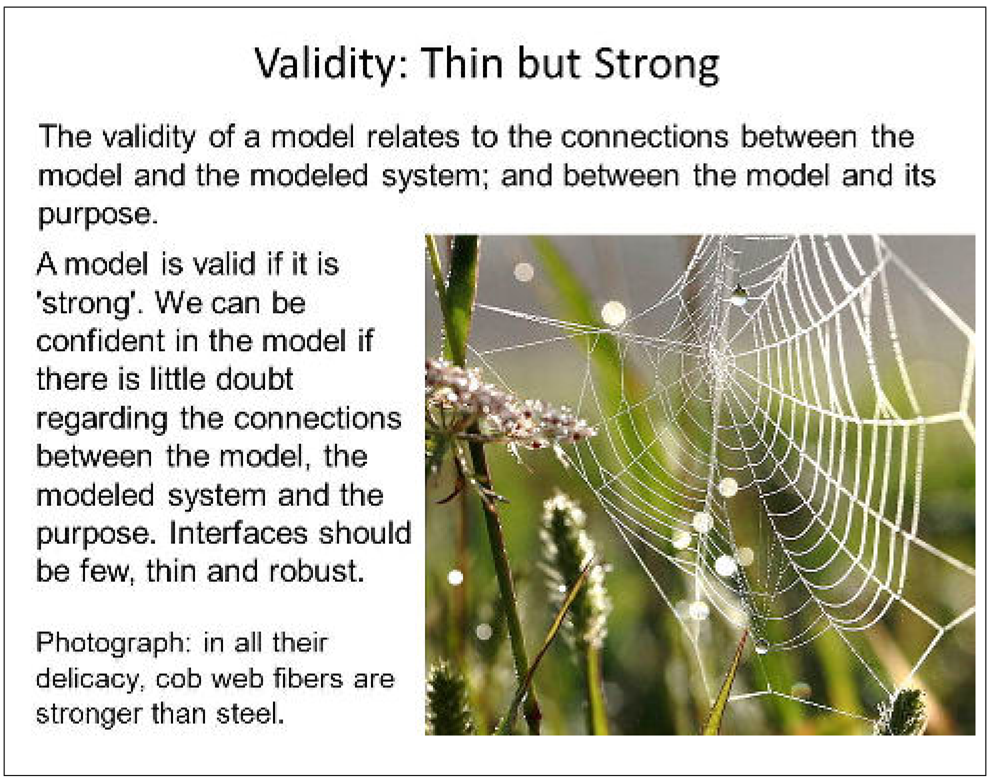
[3]The image of the paradoxical falling buttered cat is taken from http://commons.wikimedia.org/wiki/Category:Buttered_cat_paradox#mediaviewer/File:Buttered_cat.png

On the other hand, if we include the above expression in the model under the assumption 'the market share is a ramp function that equals 1 if `price` is less than the price of the cheapest competitor and that equals 0 if `price` is larger than the price of the most expensive competitor', condition (2) is trivially fulfilled, but condition (1) is problematic.

It will be next to impossible to find an assumption such that both (1) and (2) hold.

The problem of verifying the assumptions in a model is, in general, unsolved. In the philosophy of science this problem is known as the problem of VERISIMILITUDE, and it is debated in the literature until today.

Validation: is X the correct model?



Validation amounts to answering: 'Is X the correct model?', in other words: is it 'strong' enough to satisfy the purpose[4]?

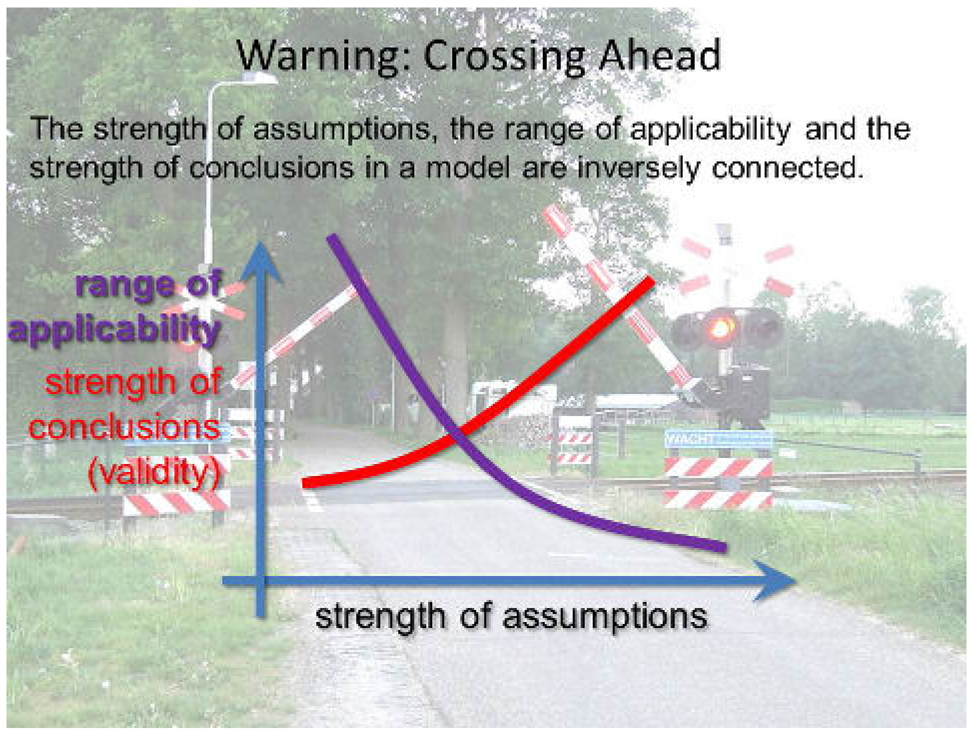
Validation takes the connection between a model and the modeled system, and the connection between a model and its purpose into account. It concerns the external interfaces of the model. It deals with issues such as: 'are cat.-III quantities reliable, given the modeled system?', 'are the obtained values for cat.-II quantities conclusive, given the purpose?', 'do predictions of the model match with empirical observations within tolerance?', 'are assumptions about the modeled system true?'.

We give an example of a model outcome that is invalid because it is not conclusive. Recall the chimney sweepers problem from Section 4.3.2, for the purpose to verify if there are not more than 50 Eindhoven chimney sweepers. Suppose that the outcome of the model would be 45 chimney sweepers with a margin of $\pm 20\%$. The number of sweepers then can be anywhere between 36 and 54, so the model outcome is not suitable to serve the purpose. We then say that the model outcome is *invalid*. An outcome of 45 chimney sweepers with a margin of $\pm 3\%$, according to a similar reasoning, would be a valid outcome.

[4]The cob web photograph was taken from http://commons.wikimedia.org/wiki/File:Strength_3472.jpg?uselang=nl

Verification and Validation: Weak and Strong, Assumptions and Results

We saw two conditions with respect to assumptions: (1) all assumptions must hold and (2) there must be enough assumptions so that the formal expressions logically follow from these assumptions. Consider a trivial example[5]: the wallet example from Chapter 1. Say we have 30 Euro in our wallet, and the book we consider to buy costs 25 Euro, will we have enough money left for a meal?



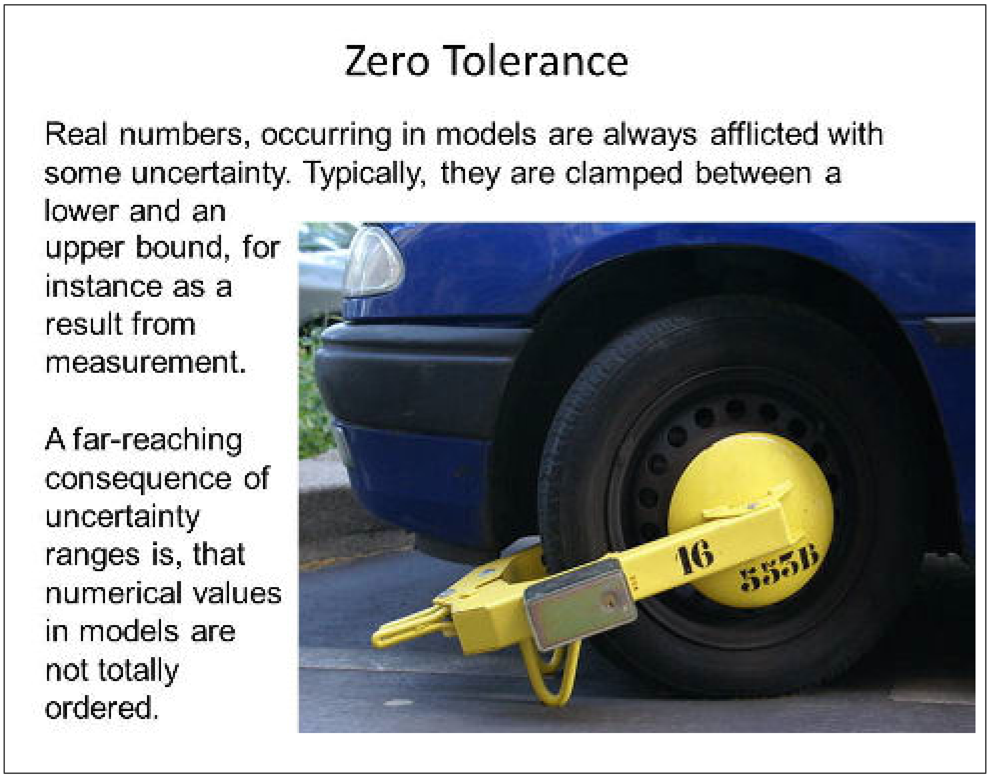
To make the model work, we must make an assumption about what a meal will cost. We now see a trade off between conditions of type (1) and of type (2). If we allow a larger TOLERANCE on the assumed price for a meal, that is: if we assume the price of a meal to lay in a broader range of values, this assumption will have a bigger chance to be true. There will be a bigger class of cases where the price of an actual meal indeed lies in this interval. Condition (1) is easier fulfilled.

We use the words WEAK and STRONG to reason about assumptions and their chances to be true.

A stronger assumption ('a meal costs between 4 and 5 Euro's') implies a weaker assumption ('a meal costs between 3 and 8 Euro's). A weaker assumption applies to a larger collection of possible situations than a stronger assumption on the same topic. So with weaker assumptions, the model has true outcomes in a larger class of situations.

The flip side, however, is that with weaker assumptions, the outcome of the model may not be precise enough to fulfill the purpose. Then the outcome of the model 'yes you can afford to buy the book', or 'no you cannot afford to buy the book' cannot be rigorously obtained from the model in the case of the weaker assumption. The outcomes are not conclusive. Condition (2) is failing for the weaker of the two assumptions. With the strong assumption ('a meal costs between 4 and 5 Euro's'), the model is capable to to a valid suggestion ('after buying the book, you will have 5 Euro left, enough to buy a meal'). With the weak suggestion, it isn't ('after buying the book, you will have 5 Euro left, which may or may not be enough to buy a meal').

[5]The photograph of the railroad crossing was taken from http://commons.wikimedia.org/wiki/File:Spoorwegovergang.jpg?uselang=nl



This is a common pattern in verifying and validating a model.

Verification can often be made easier at the expense of the validity of the model, that is: at the expense of the strength of the model outcome. And vice versa: if we need a very strong model outcome, the assumptions needed may be so strong that their verification is problematic ▷106 .

Trade-offs between verification and validation are often related to tolerance intervals[6]. Given the purpose, we should find out what the required tolerance in the outcome of the model needs to be, so that we can produce a valid output. This required tolerance should be not narrower than strictly necessary. Indeed, the narrower the output tolerance needs to be, the narrower the tolerances occurring in the assumptions need to be, and the more difficult (or: the less plausible) it is to ascertain that the assumptions hold. Verification of the assumptions of the model will be harder if a valid output demands narrower tolerances.

In Table 6.1 we list some model purposes from Section 1.2.2. For each we give examples of validation and verification. We also give examples of failing assumptions, and outcomes that may be verifiable, but have little value with respect to the purpose.

Verification and validation[7] are often based on ACCURACY and PRECISION ▷108

To be valid, the accuracy of a model outcome must be sufficiently high, and it must be sufficiently precise. Furthermore, errors should be sufficiently small, all with respect to a given purpose. As an example: the moon's gravity can be safely left out of a model to calculate if somebody following a diet has lost weight; the moon's gravity is essential in a model to predict the next high tide.

Accuracy relates to BIAS, SYSTEMATIC ERROR and OFFSET. For example: if one attempts to measure one's weight by stepping on a scales with shoes on, the weight of the shoes causes the read value to be too large. This error won't get smaller by repeating the measurement: it is a *systematic* error.

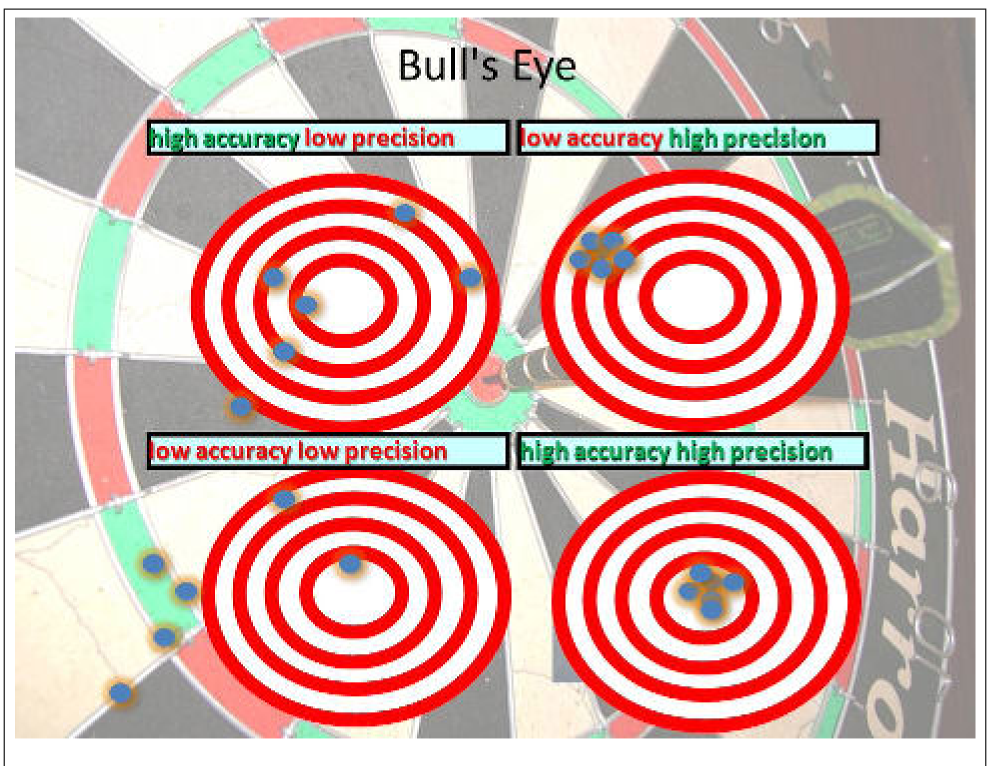
Precision relates to NOISE, RANDOM ERROR, STATISTICAL SPREADING.

[6]The photograph of the wheel clamp was taken from http://www.rgbstock.nl/photo/meTaQ0q/No+parking

[7]The photograph of the bull's eye was taken from http://commons.wikimedia.org/wiki/File:Harrows_Bristle_Board_Bullseye.JPG?uselang=nl

Purpose	Validation	Verification	Failing assumption	Verifiable but little value for the purpose
explanation	explanation relates to the phenomenon to be explained	explanation contains no inconsistencies	forgetting to check if the intended audience is familiar with the used terminology	a fragment of formal text without clarifying examples
prediction (weather forecast)	a prediction with sufficiently certainty to justify, e.g., cancellation of an event due to upcoming thunderstorm)	check that a predicted wind speed does not exceed 100 km/u, if the model uses approximations only valid < 100 km/u,	forgetting to check if solar radiation is indeed constant	correct prediction of the weather of the next five minutes
compression	lossy compressed scanned text that, although noisy, is still readable	assessing that compressed data has identical statistical behavior as original data	for a linear least squares fit, forgetting to check if errors in the input data have a normal distribution ▷107	linear least squares fit of 2 data points
optimization	an optimization producing a solution that is significantly better than previously known solutions	an optimization producing an optimum with 100% certainty bacause it does exhaustive search ofthe search space	forgetting to check if the entire search space is indeed accessible in the modeled system	an optimization that reports that no solutions could be found satisfying all constraints

Table 6.1: Validation and Verification of models, given their purpose



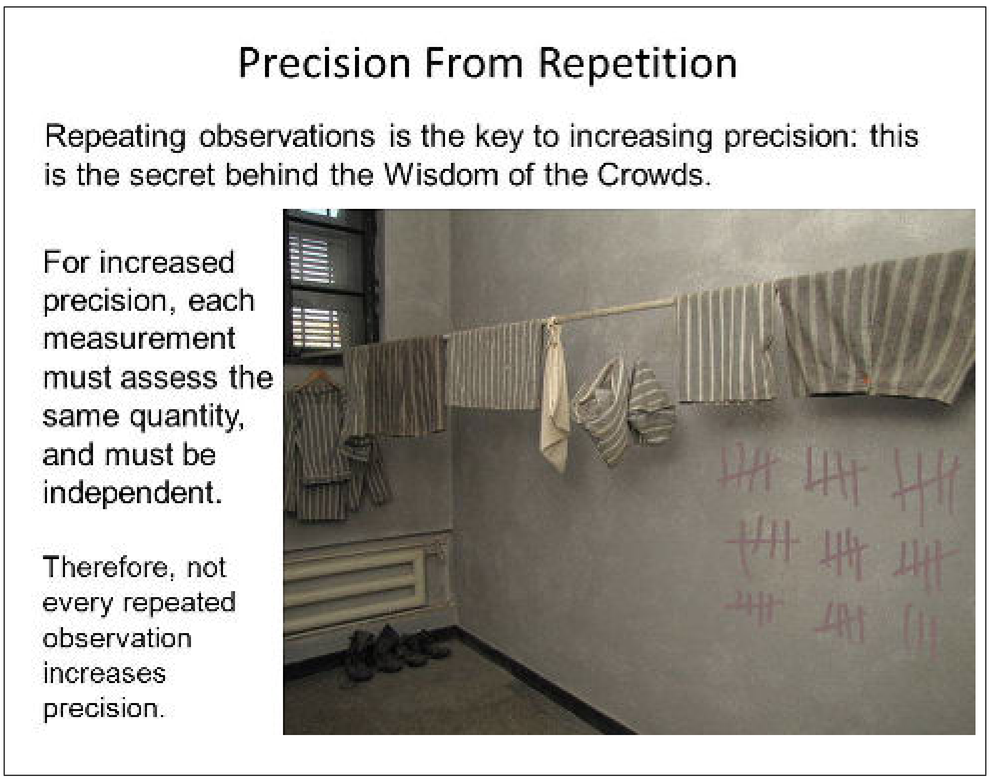
Precision can typically be improved by repeating measurements; accuracy only gets better if sources of systematic deviation are removed. Therefore, precision is closely related to REPRODUCIBILITY, that is: the extent to which repeated executions of the same measurement yield results that are similar.

Both terms relate to the degree of certainty ▷109 of the (numerical) outcomes of the model.

Accuracy The term 'accuracy' is used to discuss the deviation between values v_E and v_T. The estimated value, v_E, is the model outcome. The true value v_T is measured from the modeled system.

Sometimes, obtaining v_T is unproblematic. For instance, counting the number of letters in a text is easy ▷110 . Counting the number of obese children in a population, however, is problematic since the definition of 'obese' involves a threshold to distinguish children that are obese from those that are not. If we alter the threshold, this will effect the outcome of out measurement. But do we know what the 'correct' value for the threshold is?

When obtaining v_T amounts to measuring with non-integer outcome ▷111 , we must deal with UNCERTAINTY in measurements ▷112 . Usually measurements are repeated, and we assume that measuring outcomes are independent and that we are measuring the same thing every time. Then the sample average A_S, defined as $A_S = \frac{1}{n}\Sigma_{i=1\cdots n}a_i$ where a_i are n independent measurements of the same real quantity, in the limit for n to ∞ converges to a constant value. This value is v_T.



Precision The term 'precision' does not assume a 'true value'[8]. A measurement can be made arbitrarily precise by sufficiently many repetitions, irrespective of the quality of the instrument for measuring. The only assumption is that the quantity we are measuring really exists, independend of our measurements ▷113 , and is constant during the period of repeated measurements. The precision of a value is indicated by the number of significant decimals. Repeating can make the number of decimals arbitrary large. Given the quality of the instrument, however, all but the first few decimals are meaningless.

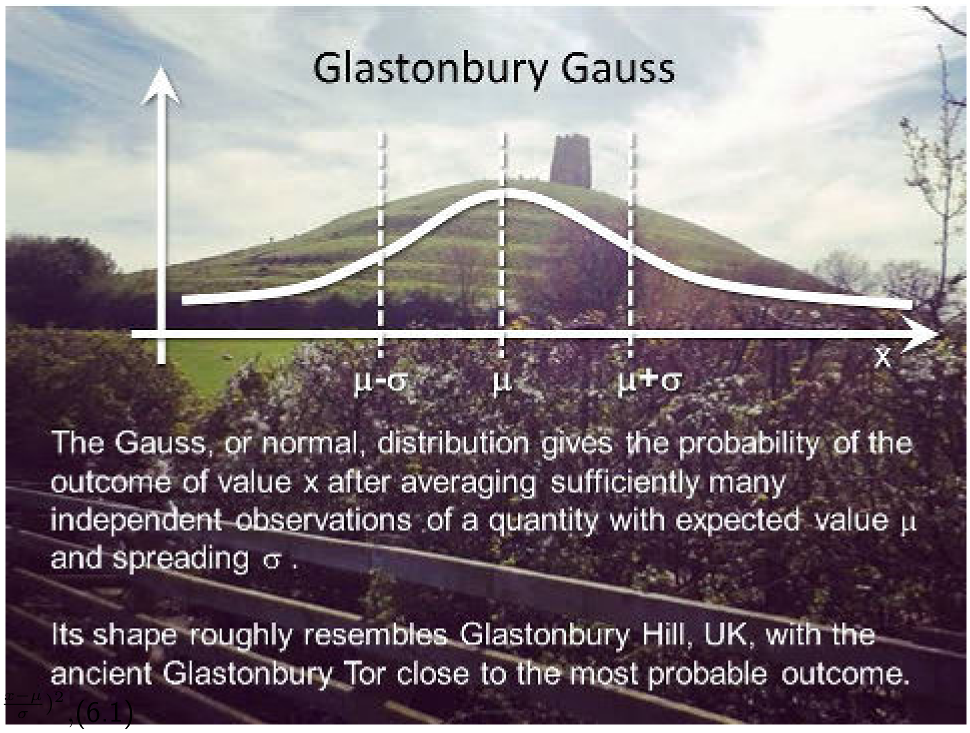
In numerical models, something similar applies. For instance, many numerical algorithms make use of iteration. When we approximate an integral or the solution of an equation by an iterative algorithm, we use a step size, say h. Typically, the smaller the step size, the more precise the answer ▷114 . For numerical models, just as for measurements, however, we must realize that increased precision does not imply increased accuracy. With *failing* precision, however, we are guaranteed to find a decreased accuracy.

[8]The photograph of the jail wall with tallies was taken from http://www.rgbstock.nl/photo/mB5G4Po/state-o

6.1.2 Distributions to Indicate Uncertainty

Lacking accuracy and/or lacking precision cause uncertainty to increase, and hence decrease confidence in model outcome. Instead of a single value for any model quantity, we have a DISTRIBUTION of *possible* values.

The precise form[9] of the distribution is usually unknown: does it have a peak, is it symmetric? In general, however, a broader distribution means less certainty and lower confidence. A value with much certainty gives a narrow distribution. The limiting case of absolute certainty is a distribution with width 0 (=a peak distribution).



An often occurring uncertainty distribution is the NORMAL distribution. The normal distribution is given by the Gaussian or NORMAL DISTRIBUTION

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (6.1)$$

that is a function with a single maximum for $x = \mu$ and decreasing, but positive, values for x going to plus and minus infinity. It is symmetric round $x = \mu$. Integrating the function over x between minus and plus infinity yields 1. This is equivalent to saying the chance for any outcome between minus and plus infinity is 1.

The Gaussian has many properties relevant for statistics. In particular, a sum of sufficiently many uncorrelated numbers with centre value μ and *standard deviation* σ has a normal distribution. The standard deviation is a measure for the spreading of the values. This result is the so-called CENTRAL LIMIT THEOREM. Its consequence is, that if measurement of a constant value is repeated sufficiently often, where all measurements are independent, the measurement outcomes are distributed normally, and the best estimate of the real value is the center value μ of the associated normal distribution.

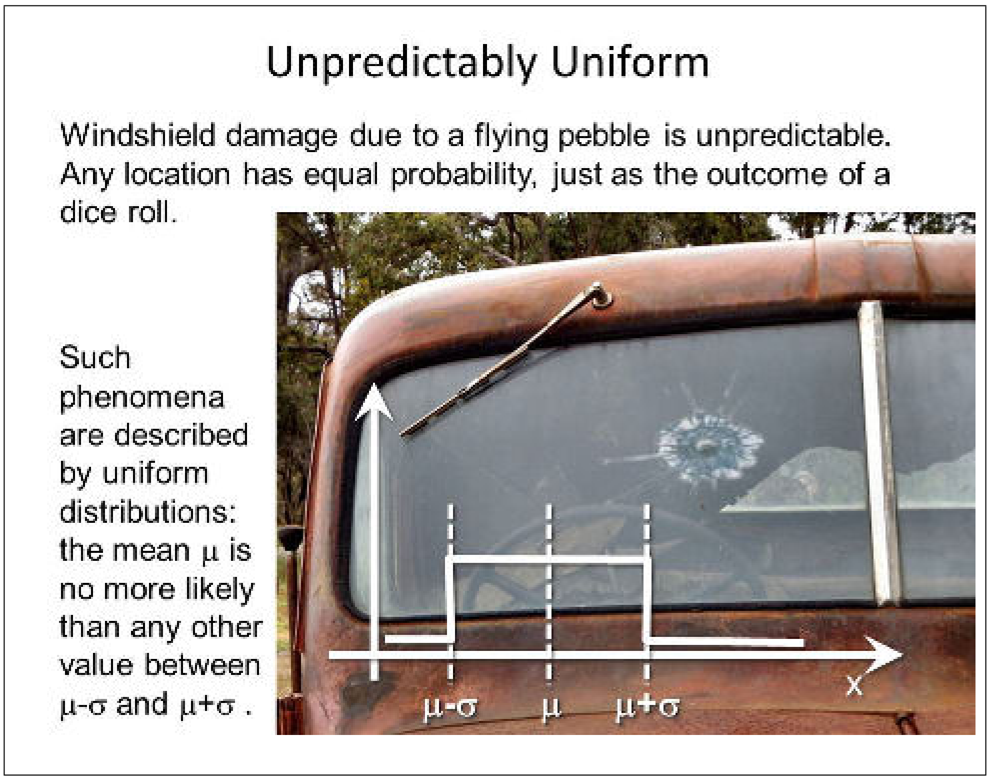
We may know beforehand that quantity x can only occur in a limited interval, say between σ_{low} and σ_{high}. The Gaussian, in that case, is clearly not adequate. We then often assume the

[9] The photograph of Glastonbury Tor is taken from http://commons.wikimedia.org/wiki/File:Glastonbury_tor.jpg?uselang=nl

distribution to be UNIFORM, given by

$$
\begin{aligned}
f(x;\sigma_{\mathsf{low}},\sigma_{\mathsf{high}}) &= \frac{1}{\sigma_{\mathsf{high}}-\sigma_{\mathsf{low}}} \quad \text{for} \quad \sigma_{\mathsf{low}} \leq x \leq \sigma_{\mathsf{high}}, \\
&= 0 \quad \text{elsewhere.}
\end{aligned} \tag{6.2}
$$

The uniform distribution applies[10], for instance, to the throws of a single dice ▷115 . The outcome for each 'measurement' is equally probable. Unlike with the Gaussian, there is no central value corresponding with 'the' true outcome; hence the flat distribution in stead of the peak in the Gaussian.



The purpose of a model may be formulated in terms of hard upper- and lower boundaries without specification of the distribution in between. Then the uniform distribution is the safest choice to work with.

Both the normal distribution and the uniform distribution have a characteristic width and a center value. In the normal distribution, μ is the center value, and about 68% of the accumulated probability is between $\mu - \sigma$ and $\mu + \sigma$. For the uniform distribution, the lower and upper borders of the interval bound 100% of the probability density; the probability of having a value outside the interval is 0. It has its central value precisely in the middle between σ_{low} and σ_{high} and it is normalized just as the Gaussian.

Distributions and the Usefulness of a Model Outcome

Let us now look at how distributions can help assessing the usefulness of a model outcome.

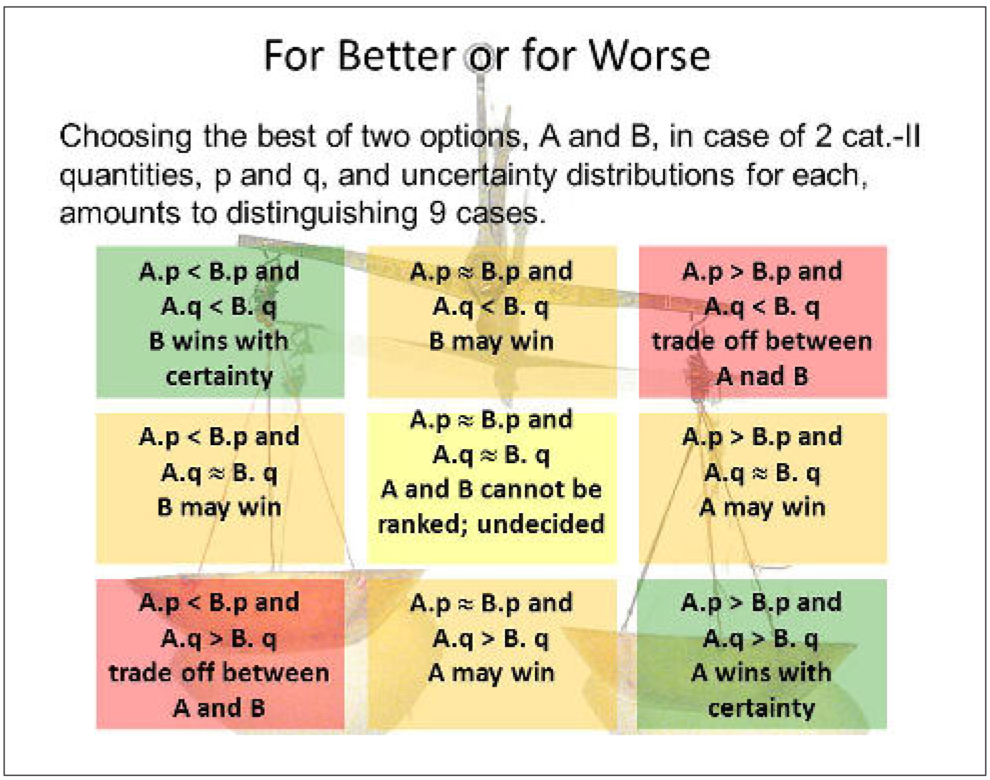
We take the example where a model should support a decision. The simplest case is a BINARY DECISION. This occurs for example in diagnosis ('is there yes or no a malign tumor in this CT image?'), or to substantiate a design decision ('which of two materials, A or B will perform best with respect to price and stability?').

[10]The photograph of the windshield was taken from http://www.rgbstock.nl/photo/otnK1Eq/verlaten+oude+truck3; the star was taken from http://www.autobedrijfvanelst.nl/Diensten-en-Services/reparatie-and-onderhoud/Ruitenservice/

In the first case, we must associate a threshold to distinguish between malign or not malign. This threshold could be, for instance, on the estimated area of a tumor. Larger than T is interpreted as malign. In the CT image are, of course, no tumors. There only is an area of pixels with certain properties. These properties are calculated from detected CT intensities at some location, and this calculation yields an uncertainty interval. This means that pixels have some probability to be erroneously classified as 'belonging to the tumor', and similar, pixels can be erroneously classified as 'not belonging to the tumor'.

An algorithm to estimate the area of a tumor works with uncertain pixel values. Hence, the calculated area of the 'tumor' also will be a distribution of values rather than a single value. The diagnosis not so much amounts to the question whether a single value (=the calculated are of a tumor) is above or below threshold. Rather, it assesses if a *distribution* is above or below threshold. For normal distributions, the answer is a confidence value. That is a percentage of the area underneath the Gaussian distribution that is above or below the threshold. If this percentage is close to 0% or close to 100%, the model gives us a high confidence. Assuming that *all* sources of uncertainty are accounted for in the Gaussian distribution, then close-to-0% or close-to-100% means that the model gives much confidence. Values closer to 50% give little information.

The narrower the distribution[11], the fewer cases there are where the outcome is close to 50%. To get outcomes that are conclusive in as many as possible cases, we should have a maximally high precision; the spreading in the outcomes should be possible small.



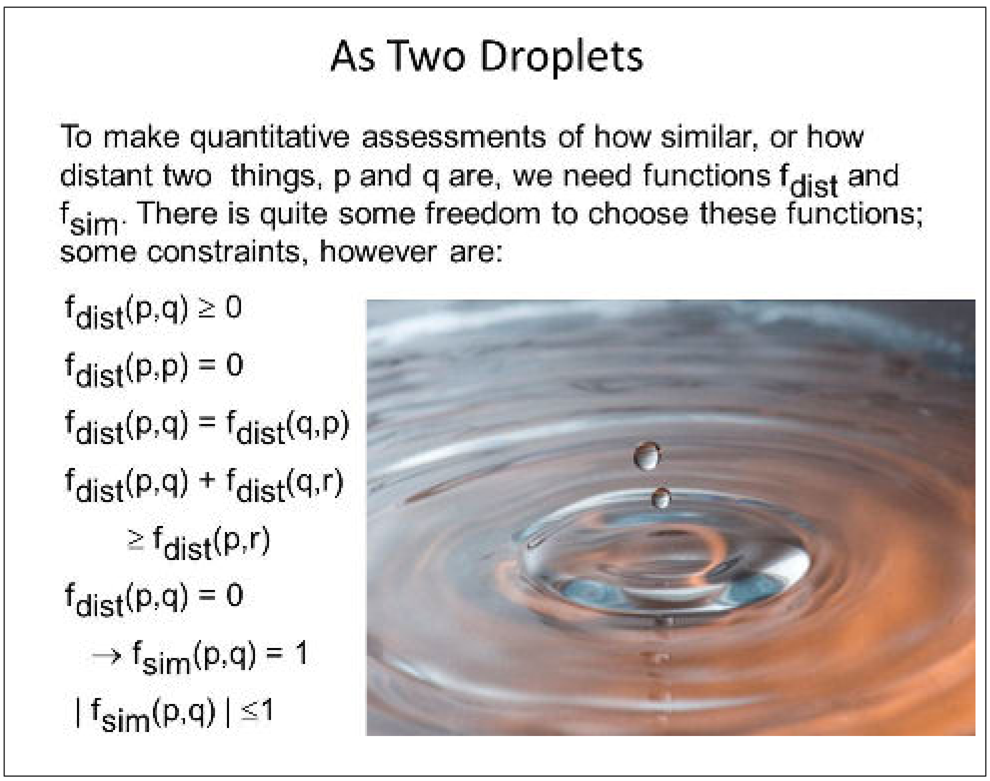
The second case, selecting the best from two materials, is more subtle. First, since the choice is based on two instead of one cat.-II quantity, we may see trade-offs: material A could be better on stability, and material B could win on price, or vice versa. But even in a situation where, say, A could dominate B, we need to be careful. As with the diagnostics example, cat.-II quantities will be distributions rather than single values. So rather than 4 cases, we actually have 9 cases. For each of the two cat.-II quantities, we can have that A is better than B, that B is better than A, or that the two distributions overlap too much and none can be said to be better. These undetermined cases reduce the confidence, and therefore the usefulness, of the model as a

[11]The photograph of balance scales was taken from http://commons.wikimedia.org/wiki/Category:Balance_scales#mediaviewer/File:Samenwaage.jpg

decision support tool.

6.2 Distance and Similarity

SIMILARITY is related to DISTANCE. The distance between two numbers is the absolute value of their difference. In the plane, the (shortest) distance between two points is usually defined as the length of a straight line between the two. Notice that a distance is always 0 or larger. The distance between something and itself is 0, and the distance between P and Q plus the distance between Q and R cannot be less than the distance between P and R. The latter is called the TRIANGLE INEQUALITY.



One way to say that two things are similar[12] is to say that their distance is small. In penalty functions, cf. Section 5.3.1, we used this idea. The point with penalty value 0 is optimal; the smaller the (non-negative) penalty value, the closer we are to this optimum.

Both similarity and distance are functions of two arguments. The minimal distance between two things is 0; their similarity is then maximal. By convention, we say that the similarity function than gives the value 1. When the two arguments are further apart, the distance function increases, perhaps to infinity, whereas the similarity decreases.

For similarity, we can accept negative outcomes. Negative outcomes correspond to the intuition of 'opposite'. We encountered an example when we explained *directions* in Section 1.3.5. The similarity in directions of two people both walking northward is maximal (1); if one of the two walks eastward their similarity is 0, and if one walks northward whereas the other one walks southward, similarity is -1. We have the idea that a difference in directions cannot be larger than between northward and southward, hence the assignment of the maximally negative value to 'similarity' in this case.

Further, both symmetry and distance should be symmetric functions: when we swap the two

[12]The photograph of the water droplets was taken from http://www.rgbstock.nl/photo/ooyFwWE/twee+druppels

'things', both their distance and their similarity should not change.

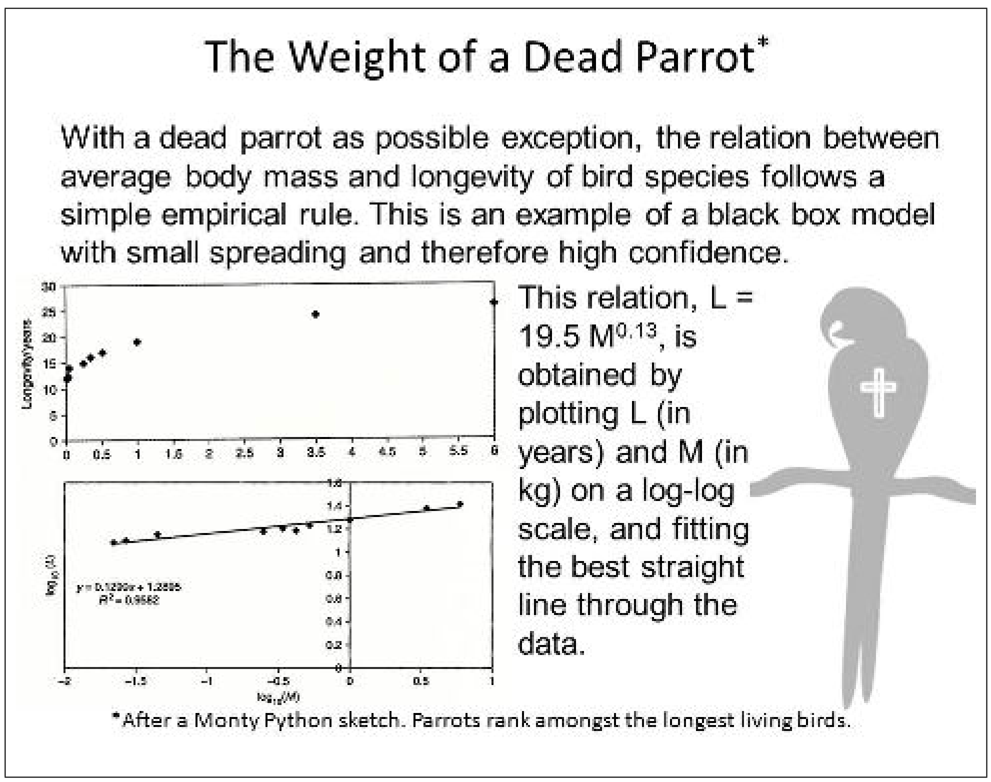
Unlike distance, the addition of similarities does not give something intuitive. Therefore, there is no equivalent for the triangle inequality in similarity measures.

In Appendix **??** we give a more detailed account of the calculation of distance and similarity functions in various cases.

6.3 Confidence in Black Box Models

a set of empirical data. The model derives from this data a representation, typically consisting of some quantities. These quantities can be interpreted as features in the initial set. For instance, the study of the relation between mass and longevity in birds[13], cf. Section 1.3.8, hinges on values for 'the average mass' and 'the average longevity' of birds of species X. 'Average mass' and 'average longevity' are the quantities that together form the black box model of the set of data.

The AVERAGE is an example of a quantity, obtained from empirical data, used in a black box model. Other examples are the STANDARD DEVIATION and the CORRELATION.



6.3.1 Features from Data Sets and from Distributions

We must distinguish quantities that can be obtained directly from a set of data $S = \{s_1, s_2, s_3, ...s_n\}$, and quantities that are properties of the distribution, of which S is drawn.

For instance, the average, formed by adding the s_i and dividing by n, is a property from S. S, however, could have been taken from a (hypothetical) collection, characterized by some distribution, say D. This distribution tells what the chance for each possible outcome is when drawing samples. The distribution D could for instance be a uniform distribution, that is that every outcome between a lower bound and an upper bound has equal chance. Such a distribution is

[13]The silhouette of a (living) parrot was taken from `http://www.rgbstock.nl/photo/mNueNQG/silhouette+parrot`

characterized by some quantities, including its mean. For D being a uniform distribution, the mean is halfway between the lower and the upper bound. If the collection S is large enough, the average $\frac{1}{n}\Sigma_{i=1,\ldots n}s_i$ will be a good approximation for the mean of D.

Similar to the mean of D, related to the average from S, there is the standard deviation of D that is related to the spreading among values in S.

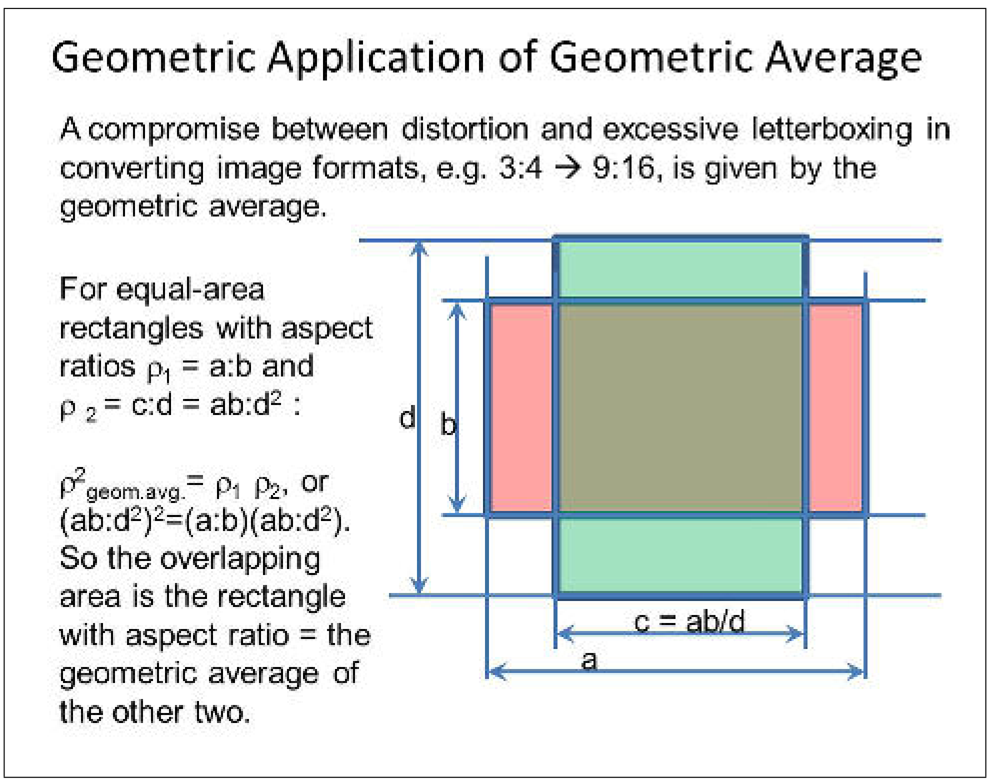
Mean

The mean, sometimes called μ, of a distribution, sampled by a set of numerical values can be estimated as the sum over these samples, divided by the number of samples,

$$\mu = \frac{1}{n}\Sigma_{i=1\cdots n}s_i, \tag{6.3}$$

(the so-called ARITHMETIC MEAN), or as the n-th root of their product [▷116] ,

$$\mu = (\Pi_{i=1\cdots n}s_i)^{1/n}, \tag{6.4}$$



(the so-called GEOMETRIC MEAN). In both cases the dimension of μ equals the dimension of the values s_i. The definitions only make sense if all s_i have the same dimension.

An application of the geometric mean is finding 'average' aspect ratio's. If a film with width:height ratio s_1 needs to be displayed on a screen with width:height ratio s_2, there is a dilemma between stretching, which distorts the image, or letter boxing, i.e. cutting of parts of the image, which is a waste of resolution. A compromise between these is, to map the image to an aspect ratio that is the geometric means of s_1 and s_2. This gives not too much distortion and doesn't waste too much screen resolution. It can be proven that two rectangles with same center and same area but different width:height ratio's overlap in a rectangle with a width:height ratio that is the geometric means of the two.

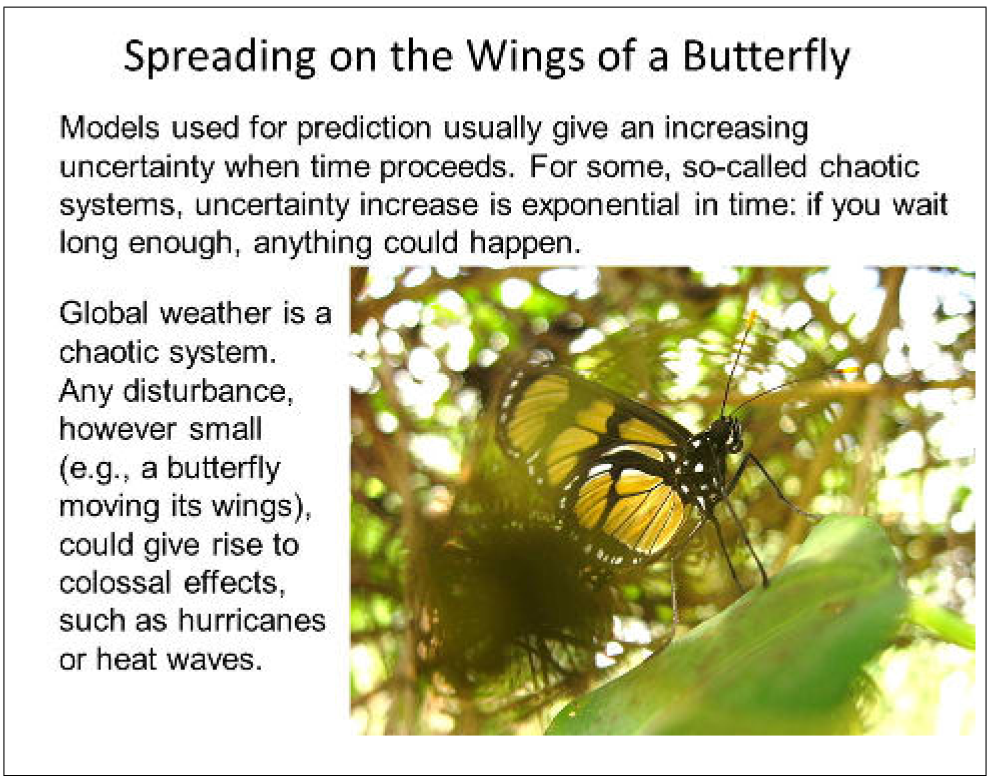
These definitions behave identical if all data values s_i are the same, say s: then $\mu = s$ for both. If one of the s_i approaches 0, however, there is a difference. The arithmetic mean then produces a value that is just a little bit smaller; the geometric mean approaches zero. This gives a first cue as to which definition to use: how should the calculated mean behave in case of a zero among the input data?

The intuition of the mean is the 'central value' of a distribution, that is the value around which all values in any randomly sampled set should vary. The amount of variation in the distribution is expressed by the so-called *standard deviation*.

Standard Deviation

Suppose that we have a distribution with arithmetic mean $\mu = 0$. Then a reasonable measure, say σ_0, of the amount of spreading[14] is the square root of the (arithmetic) mean of the squares of the values ▷117 ,

$$\sigma_0 = \sqrt{\frac{1}{n}\Sigma_{i=1\cdots n} s_i^2}. \tag{6.5}$$



This quantity has the dimension of the s_i, and it is only zero if all s_i are zero. We can now derive the standard deviation σ of the original set by subtracting the arithmetic mean from each of the s_i:

$$\begin{aligned}\sigma &= \sqrt{\frac{1}{n}\Sigma_{i=1\cdots n}(s_i - \mu)^2} \\ &= \sqrt{\frac{1}{n}\Sigma_{i=1\cdots n}(s_i - \frac{1}{n}\Sigma_{j=1\cdots n} s_j)^2}. \end{aligned} \tag{6.6}$$

[14]The photograph of butterfly, on its way causing a thunderstorm, is taken from http://commons.wikimedia.org/wiki/File:Efeito_borboleta.JPG?uselang=nl

It can be shown ▷118 that the standard deviation can also be calculated as

$$\sigma = \sqrt{\frac{\Sigma_{i=1\cdots n} s_i^2}{n} - \mu^2}, \tag{6.7}$$

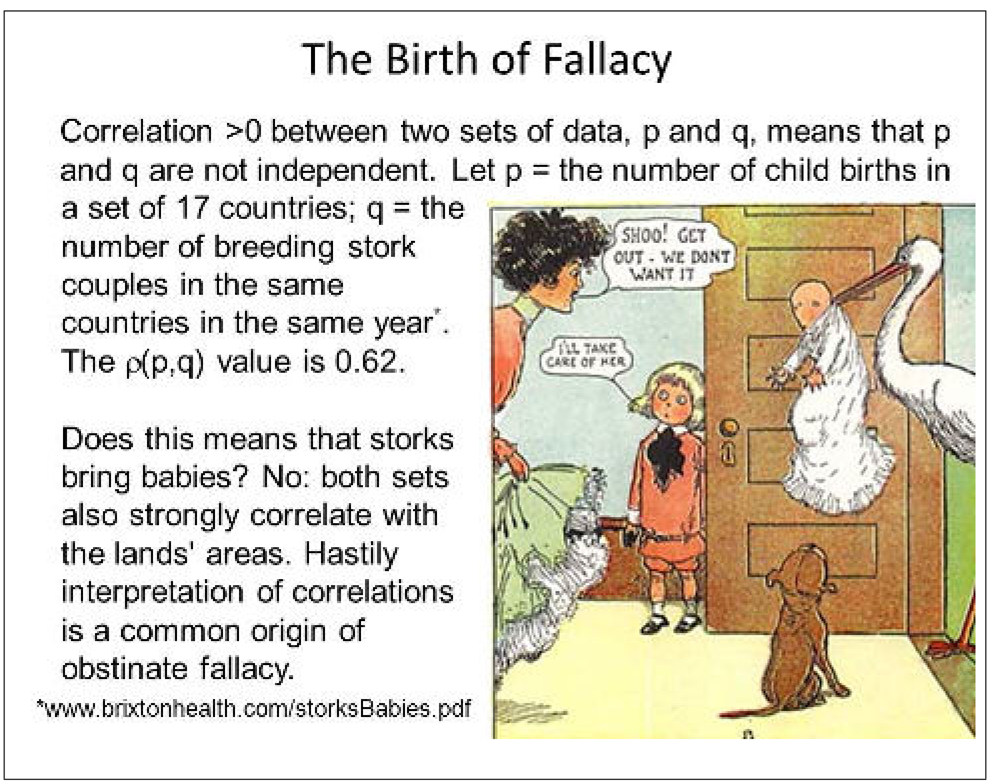
that is: the square root of (the means of the squares, minus the square of the means). It is sometimes useful to work with σ^2 instead of σ. The quantity σ^2 is called the VARIANCE.

The standard deviation is 0 if and only if all the s_i are equal. Furthermore, for a set with two numbers s_1 and s_2 it gives $\sigma = |\frac{s_1 - s_2}{2}|$. Indeed, the standard deviation is the spread in the data. The larger the standard deviation, the less cohesion in the data. The smallest standard deviation ($=0$) occurs if all values are equal. Therefore standard deviations cannot be negative.

Correlation

The correlation ρ, for data points (x_i, y_i), is defined as the normalized inner product of the two vectors $x - \bar{x}$ and $y - \bar{y}$, where $\bar{p}$ is a vector where all elements equal the arithmetic mean of the elements of p. So

$$\rho = \frac{(x - \bar{x}, y - \bar{y})}{\sqrt{(x - \bar{x}, x - \bar{x})}\sqrt{(y - \bar{y}, y - \bar{y})}}. \tag{6.8}$$



The correlation[15] informs about the similarity between vectors x and y. This definition of similarity is immune ▷119 against offsetting each vector by a constant amount, i.e., setting $x_i \to x_i + c$. This happens thanks to the subtraction of the means. If both the x and the y would have mean 0, the correlation would be simply

$$\rho = \frac{(x, y)}{\sqrt{(x, x)}\sqrt{(y, y)}}, \tag{6.9}$$

i.e., the cosine of the angle between x and y, regarded as vectors in an n-dimensional space.

The intuition of correlation is the similarity between two lists of values $[x_1, x_2, x_3, \cdots]$ and $[y_1, y_2, y_3, \cdots]$.

[15] The stork cartoon is taken from http://commons.wikimedia.org/wiki/File:Buster_Brown_baby.jpg?uselang=nl

When these have less in common, the correlation is less. 'Having something in common' means: given the i-th element of the x-es, $x[i]$, with how much certainty can we predict $y[i]$? If the correlation is 0, the $y[i]$ corresponding to given $x[i]$ can be anything. For correlation=1, it means that $y[i]$ is fully known.

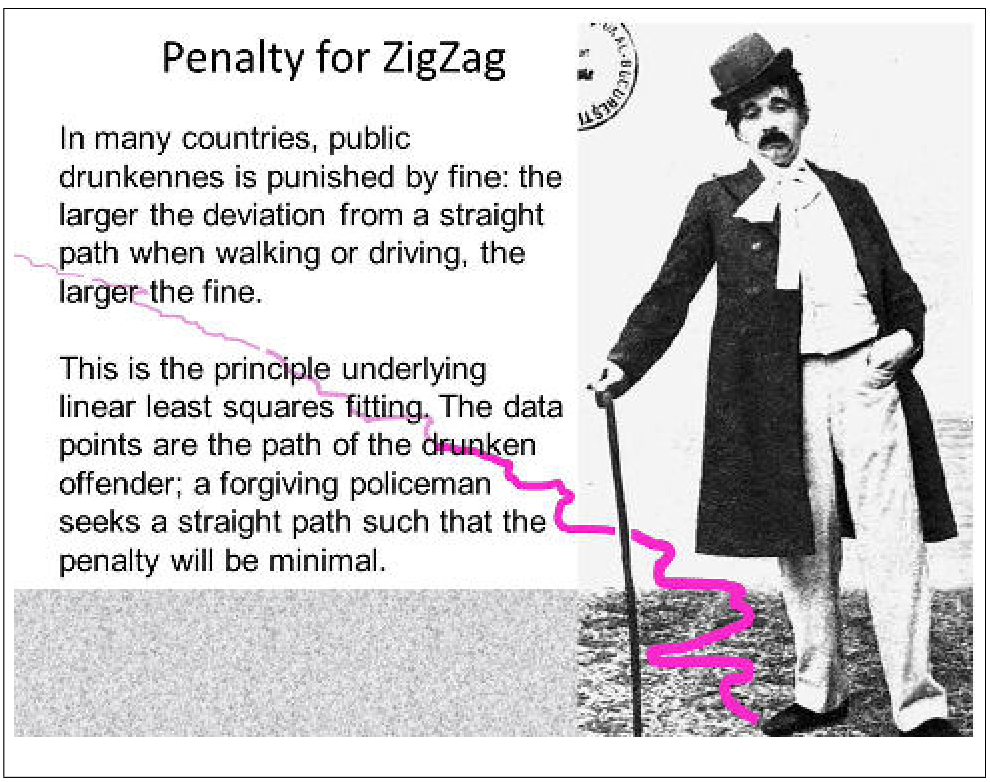
From Expression 6.9, but also from Expression 6.8, we can prove that correlations are between -1 and +1. Indeed, correlation can be negative. This is obvious from the interpretation as the cosine of the angle between vectors x and y. It means that, with negative correlation, if $x[i]$ is larger than $x[j]$, it is likely that $y[i]$ is *less* than $y[j]$ and vice versa.

Apart from average, standard deviation and correlation, their are many other quantities that can be used to represent (*compress*) empirical data. Examples are clustering and principal component analysis. All these approaches yield quantities that inform us about global characteristics the initial set.

6.3.2 Example of the Value of a Black Box Model: Linear Least Squares

The value of a black box model is larger if the quantities, obtained from the model, are more easily interpretable in terms of features of the original set. DATA MINING is the name for a number of modern techniques where patterns in large volumes of data are sought using black box models.

We give an example: the LINEAR LEAST SQUARES fit of a set of data $[[x_1, y_1], \cdots [x_n, y_n]]$. We search a linear function[16], $y = f_y(x) = ax+b$, and we seek the quantities a and b such that this straight line best represents the original data. To derive a and b we first shift the x_i and y_i so that both averages are 0. Let $x'_i = x_i - \bar{x}$ and $y'_i = y_i - \bar{y}$. Then the line we try to find passes through the origin, and $b' = 0$. Shifting does not affect the slope of the line, so $a' = a$.



To find a, we use again the idea of a penalty function. In this case a plausible penalty function is

[16]The photograph of the tramp was taken from http://commons.wikimedia.org/wiki/File:Brezeanu_este_Cet%C4%83%C5%A3eanul_turmentat.JPG?uselang=nl

$$\begin{aligned} p(a) &= \Sigma_{i=1\cdots n}(y'_i - ax'_i)^2 \\ &= (y' - ax', y' - ax') \\ &= (y', y') - 2a(y', x') + a^2(x', x') \end{aligned}$$

To minimize the penalty function we differentiate with respect to a and require the derivative to be 0. We get ▷120

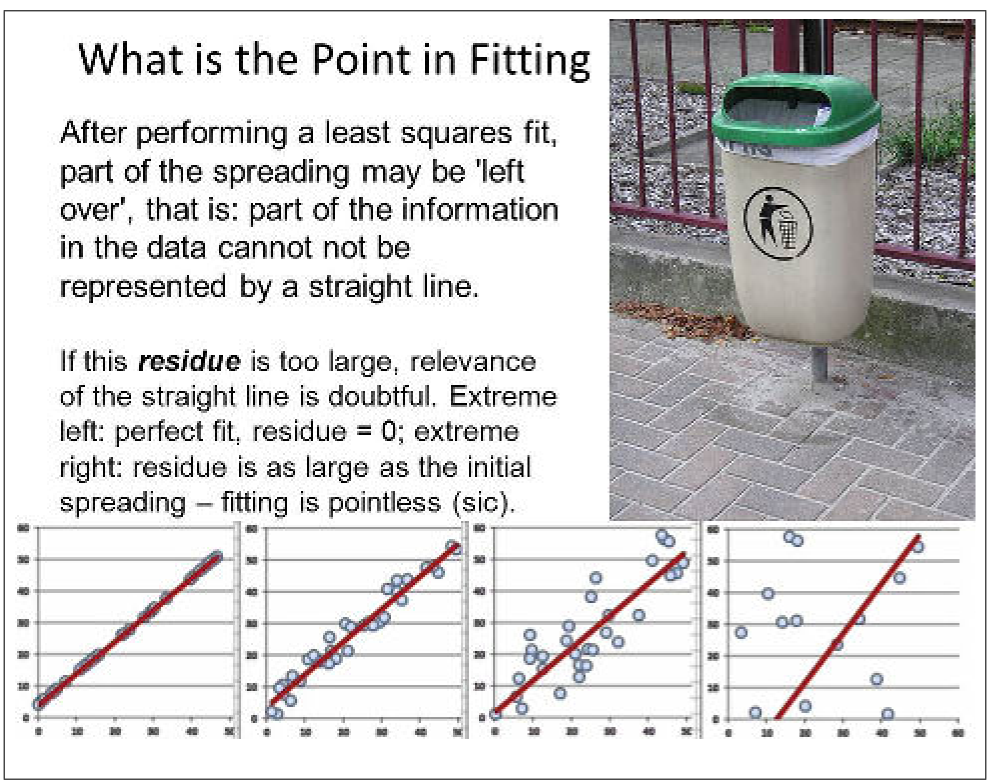
$$0 = -2(y', x') + 2a(x', x'), \tag{6.10}$$

so

$$\begin{aligned} a &= \frac{(x', y')}{(x', x')} \\ &= \frac{(x - \bar{x}, y - \bar{y})}{(x - \bar{x}, x - \bar{x})} \end{aligned} \tag{6.11}$$

For the initial line function, $y = ax + b$, we still have to find b. This follows as $b = \bar{y} - a\bar{x}$.

6.3.3 Validating a Black Box Model



A black box model yields values for quantities that may represent information of the initial data set. The interpretation of these quantities may be subtle, though[17]. We illustrate this by further elaboration of the example of linear least squares fit.

Residual Error

The derivation of a linear least squares fit works for any collection data, provided that $(x', x') \neq 0$. This doesn't mean, however, that any data set could be approximated by a straight line.

The obtained values for a and b don't reveal if the fit

[17]The photograph of the litter bin was taken from http://commons.wikimedia.org/wiki/File:Vuilnisbak-Lebbeke.JPG?uselang=nl

is appropriate. An *appropriate* fit gives a good approximation to the data. To assess the appropriateness of a fit, we look at the RESIDUAL ERROR . The residual error is the sum of squares of the differences between the initial data value and its estimate. For linear least squares fit:

$$RSS = \Sigma_{i=1\cdots n}(y_i - f(x_i))^2, \tag{6.12}$$

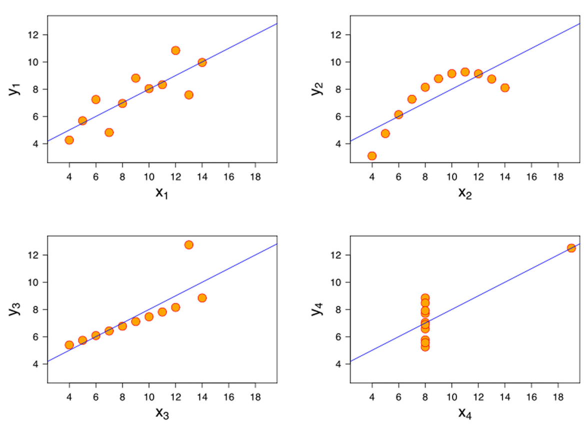
where $f(x_i)$ is the y-value of the fitted line for abscissa x_i. If the fit were perfect, the line would pass through all of the data points, and RSS would be zero. If RSS is not zero, it can be compared against the variance, that is: the square of the standard deviation of the initial y. So we should inspect the number

$$\frac{\Sigma_{i=1\cdots n}(y_i - f(x_i))^2}{\Sigma_{i=1\cdots n}(y_i - \bar{y})^2}. \tag{6.13}$$

This number should be considerably smaller than 1. If it would be close to 1, the spread of the y_i around the line is almost as wild as the initial spread. The line does hardly explain the behavior of y_i as a function of x_i. We just as well could have refrained from making the fit.

Distinctiveness

.



Examples of data sets that are all optimally fit by the same linear function.

Figure 6.1: Four data sets with very different qualitative behavior. If each of these data sets is approached with a linear least square fit, though, the same linear function results.

A black box model is a compact representation of a set of data. When it is chosen well, it captures the essence of the data. There is, however, no guarantee that a black box model does this.

Francis Anscombe in 1973 constructed 4 sets of data points $[[x_1, y_1], \cdots, [x_n, y_n]]$ that correspond to very different distributions in the plane. The distributions nevertheless yield exactly the same statistical quantities for the averages of the x_i and the y_i, for the standard deviations, and for the correlation. If we compute linear least squares fits for all four data sets (see Figure 6.1), we get the same numerical results. It is tempting to conclude that these data sets are in some sense similar, and that they should resemble somewhat a straight line.

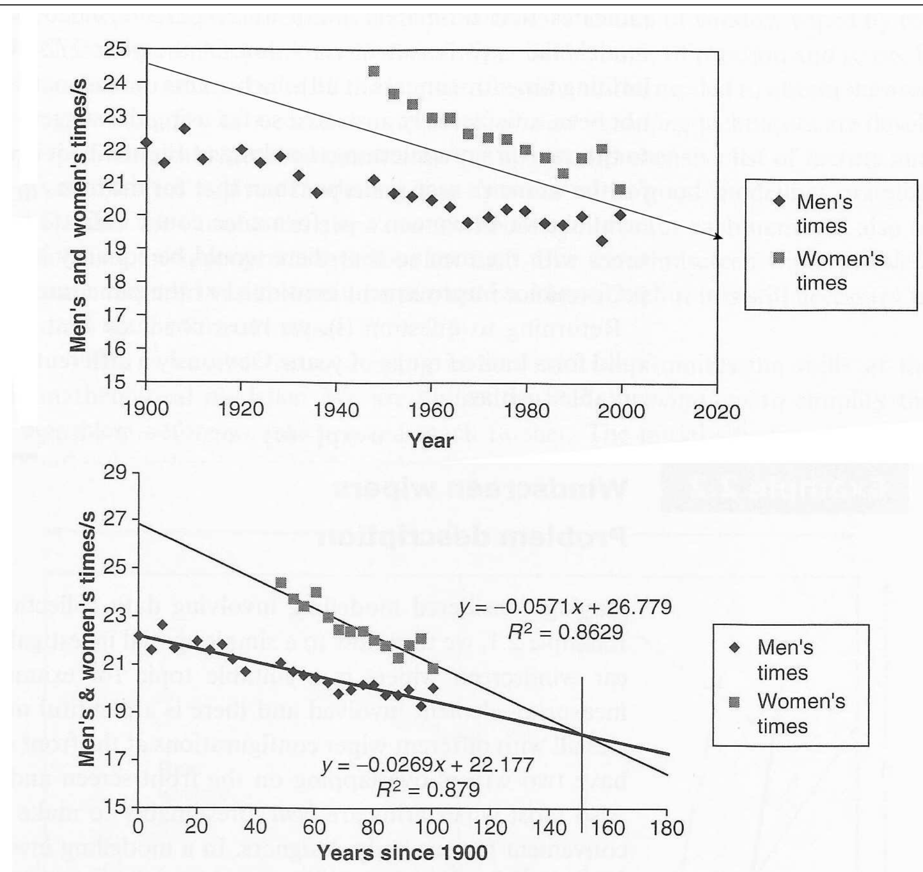
But from visual inspection it is clear that they don't.

The lesson to learn from Anscombe's example is, that (1) mathematical procedures to extract black box models from data should not be followed blindly, and (2) it is adviseable to visually inspect data.

Common Sense

.

Suppose that we have a data set, such as in Figure 6.2 (taken from Edwards and Hamson), top. We fit this data with a single linear function. The result has a bad residual error. That is because the data is a merger of two data sets. Indeed, there is a hidden quantity: data points are either a man's Olympic gold medal or a woman's Olympic gold medal for running 200m. The x-coordinates are the year; the y-coordinate are the time of the winner. In this example, the 'hidden' quantity was not really hidden, but in practice we do have to deal with hidden quantities of this kind. Statistics can help to identify CLUSTERS. That is: to attribute each data point to one of several clusters, where each cluster corresponds to a single black box model ▷121 .



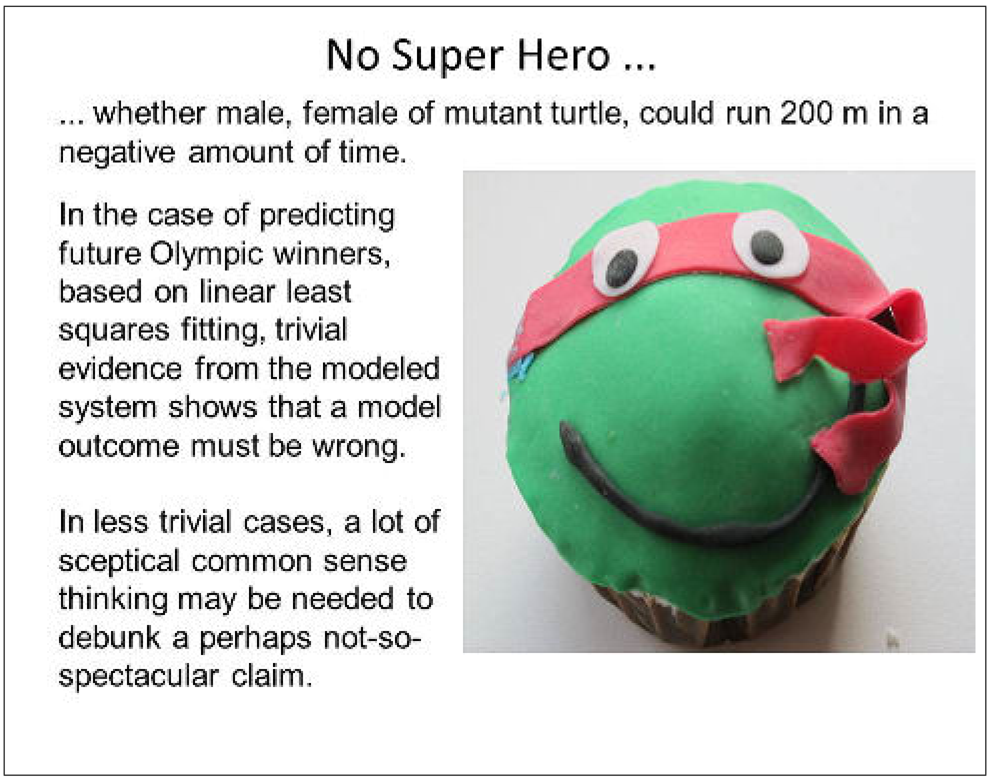
Fitting and running

Figure 6.2: Raw data that could be subject to fitting by straight lines (top). The result would give a poor so-called residual error. This is caused by the fact that the data in fact consists of two populations; when each of them if fit separately, two much better fits result (bottom). The two separate fits are depicted in the lower most image. They correspond to linear functions, describing the Olympic gold medal times for running 200 meter for men and women.

In this case, the two clusters, of course, are given: the men data and the women data. Both clusters are analyzed with a linear least squares fit, and the residual errors for both data sub-sets are sufficiently small. Also, visual inspection shows that both linear models look reasonable.

Still, both models give rise to two implausible conclusions. First, that in due course the two lines would intersect, roughly in the year 2050. So women from then on will run faster than men. Second, that a few decades later the Olympic gold medal time will reach the value 0, to become negative in the years to follow.

These two conclusions would be justified on the basis of the two straight lines, but not on the basis of the data that underlie them. A good analysis should take the uncertainty distributions of the four involved quantities into account. These quantities are the two slopes and the two intercepts of the two TREND LINES. Such analysis would show that the uncertainty of the coordinates of the intersection point are very large. Perhaps women overtake men in 2040 but it is much more likely that this will occur 1000 or more years later, or never. Indeed, the slopes of the lines differ little, so within the margins there is a chance that they will be parallel or even that the slope of the 'women-line' is shallower than the slope of the 'men-line'. The latter would mean that in due course women will take relatively longer times for the 200m than men.



The more careful analysis[18] shows that we should not try to conclude anything beyond, say, 2010. Moreover, for every conclusion we should give a probability distribution.

At a more fundamental level, all conclusions are based on the assumption that the Olympic winner's times of both men and women depend *linearly* on time. There is no real reason that this should be the case.

Finally there is a strong argument from physics that running times can never become negative.

So despite mathematically plausible results for the quality of the linear least square fits: if we fail to state, and justify, the main underlying assumptions, we cannot draw meaningful conclusions with a black box model.

6.4 Confidence in Glass Box Models

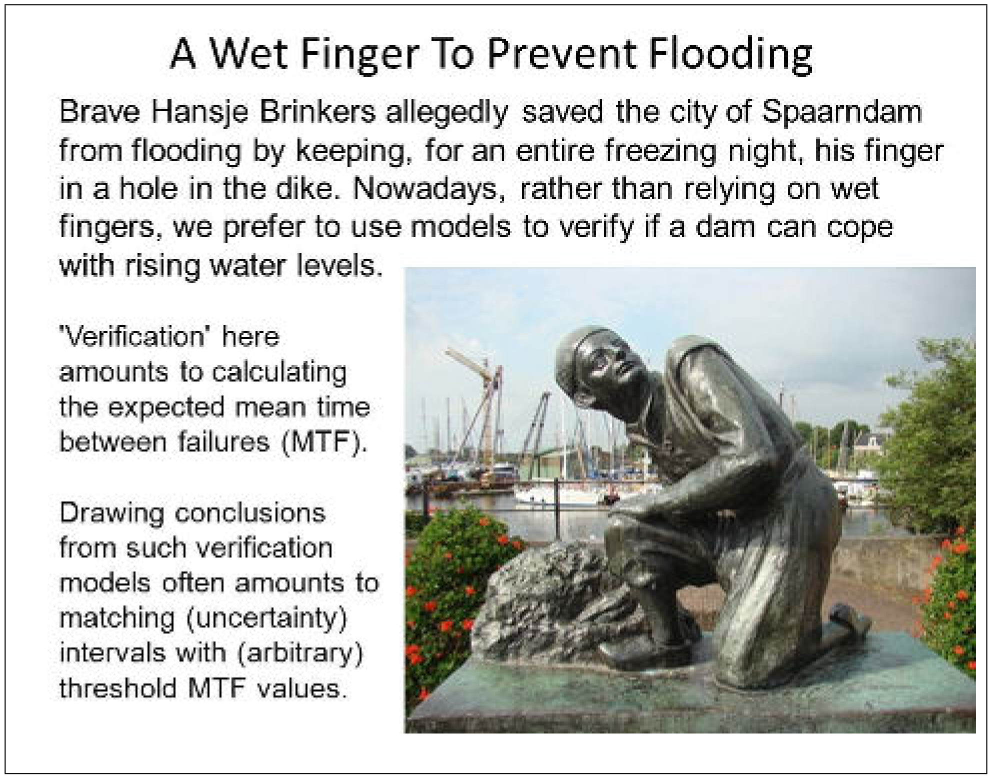
We consider glass box models that compute the value of output quantities as function of input quantities. The purpose is defined in terms of the output quantities. In order to fulfill the purpose, the uncertainty of the output quantities should be not too large.

An example is the functional model of Section 5.2. The output quantities are the cat.-II quantities; the input quantities are cat.-I and cat.-III. The purpose of the functional model is to assess of

[18] The photograph of the Ninja Turtle Cupcake was taken from http://commons.wikimedia.org/wiki/File:Ninja_turtle_cake_(8683170838).jpg?uselang=nl

two ATBD's (=two sets of cat.-I quantities) which dominates the other. Domination can only be assessed if, for all cat.-II quantities, uncertainty distributions don't significantly overlap. For uniform distributions this could mean that intervals are disjoint. A weaker form is to demand that they do not overlap more than some percentage. For normal distributions, we work with confidence intervals expressed as a number of times σ. A modern trend in quality management in industrial manufacturing is to strive for 6σ or 99,99966% confidence intervals. By far most other modeling applications are less ambitious.

An other example is the purpose of *prediction*. A prediction is useless if it is too uncertain. But how certain does it need to be? A prediction is rarely the ultimate purpose of a model. The outcome of the prediction often also has a purpose. For instance to falsify a hypothesis, or the select one hypothesis over another hypothesis. Again we relate the width of the uncertainty distribution of the output quantities to the extent in which the model satisfies the purpose. The uncertainty distribution of the output quantities should be narrow enough to distinguish the two competing hypotheses.



A final example is where the model should verify something[19]. Then the acceptable uncertainty distribution is often dictated by convention. For instance, a system can be called 'safe' if the chance for something bad happening is less then 1 incident per 10000 year (the so-called MTF, or *Mean Time between Failures*). The number '10000' follows from the common feeling that this is low enough risk to consider everything else more dangerous. So here the output quantity is a distribution of expected mean times between failures. The lower bound of this distribution here should not be less than 10000 year.

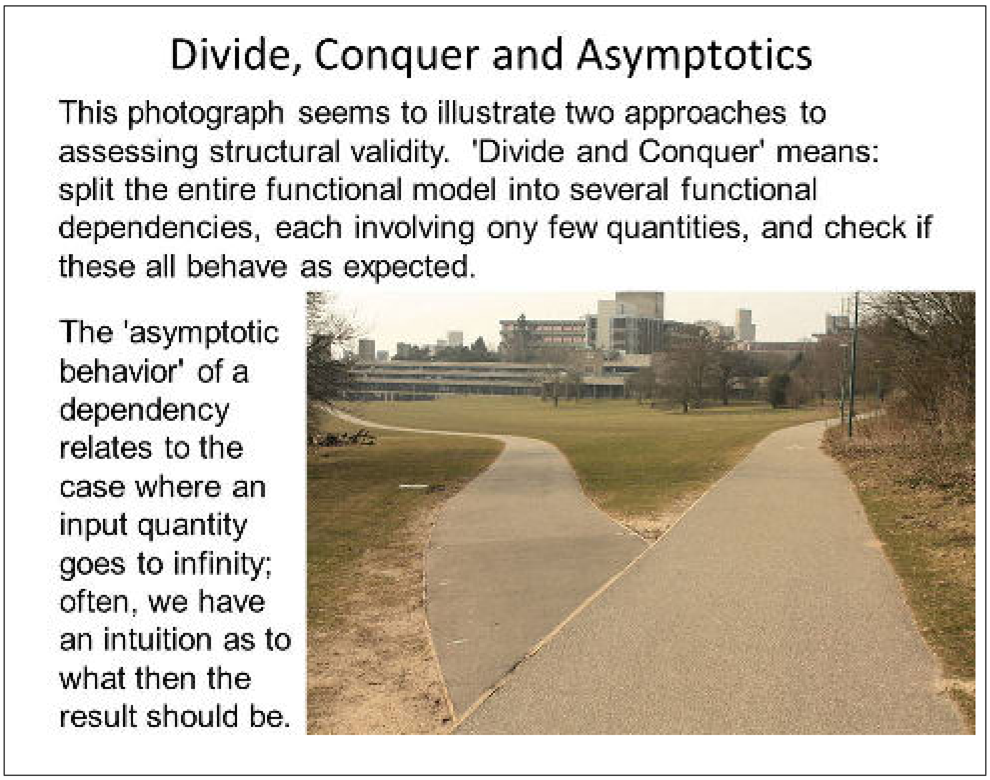
From now on we assume that we have an upper boundary for the uncertainty interval of the output quantities so that the outcome satisfies the purpose. We define a glass box model as 'valid' if (1) the possible outcomes of the modeled system don't fall outside the uncertainty interval of the output quantities of the model, and (2) that the uncertainty interval of the output quantities is narrow enough to fulfill the purpose.

[19]The photograph of the Hansje Brinkers statue was taken from http://upload.wikimedia.org/wikipedia/commons/5/52/Spaarndam_hans_brinker.jpg?uselang=nl

6.4.1 Structural Validity Assessment

There are two conditions for a glass box functional model to be valid. The calculation of the uncertainty intervals must be correct, and these uncertainty intervals must be sufficiently narrow to satisfy the purpose. The second check is a quantitative check; we explain it in Section 6.4.2. The first check is more subtle: how to ensure that the model contains the right functional dependencies?

There is no rigorous, procedural answer. There is, however, a number of efficient heuristics.



divide and conquer: The model as a whole is a function[20]. It maps values for input quantities to values for output quantities. This function will rarely be a single expression. Most often, the function will be composed of several functions, the argument of one function being the output value of a previous function. Although the entire function will typically be too complex for intuition, this is may not the case for the separate, individual functions or small groups of functions.

We illustrate this with the taxi example from Appendix **??**.

Consider the profit as a function of the world fuel price (Euro / liter). There is a number of cat.-IV quantities that need to be computed in order to express the profit in terms of the world fuel price. Yet we are certain that the net behavior will be decreasing. If fuel gets more expensive, the profit will not go up. If the model does not reproduce this behavior, there must be an error somewhere in the chain of dependencies between the cat.-III quantity `flPpl` and cat.-II quantity `profit`. Conversely, `flPpl` will certainly *not* depend on `profit`.

By examining a sufficient number of pairs of quantities from the model, with an intuitive notion of what the dependency *should* be, we can increase our confidence in the validity of the model. 'Examining' here means: drawing a plot of one quantity as function of the other on a meaningful interval of argument values.

The term 'divide and conquer' means, that by partitioning the multitude of separate functions into groups, the entire model can be be diagnosed without examining all $n(n-1)$ possible functional

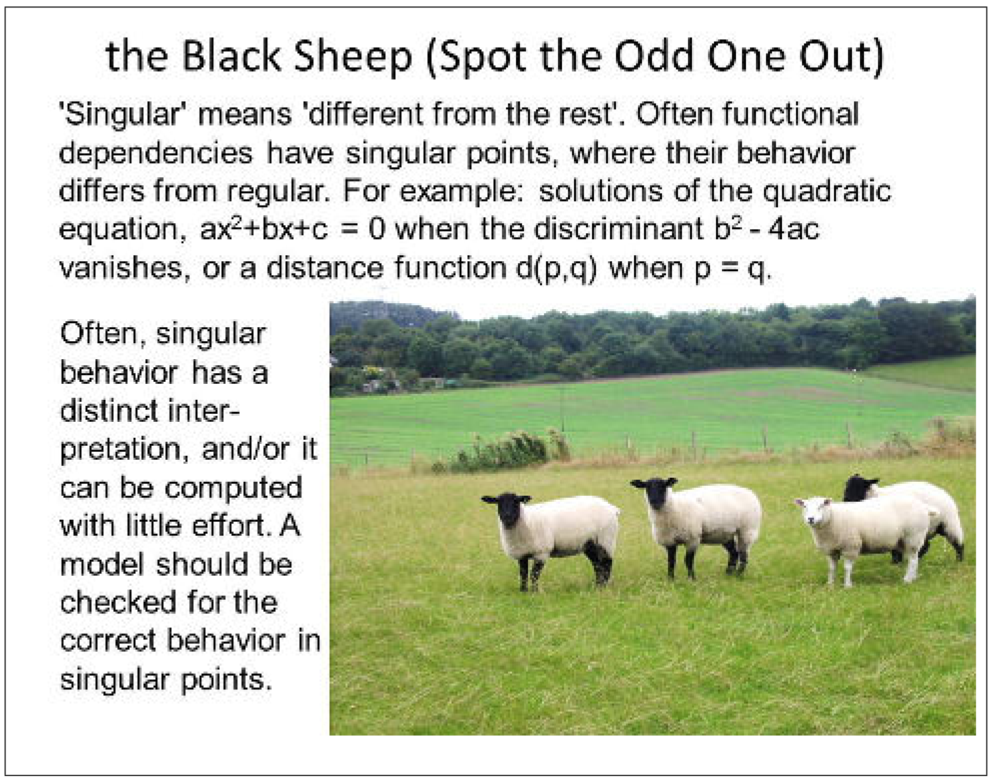
[20] The photograph of the road junction is taken from http://www.rgbstock.nl/download/ColinBrough/o3LeJ9g.jpg

dependencies for n quantities in the model. A visual scheme of the graph of all dependencies is a great help in choosing these groups.

asymptotic analysis: Every functional model has a REGIME, that is: a range of values for the input quantities where a model should work properly (for the taxi example: kilometer price somewhere between 10 cents and 10 Euro per kilometer).

If we move far outside this regime, the model typically will be simpler. Indeed, for a kilometer price of 100 Euro per kilometer we sell no rides; with 1 cent per kilometer, we get all potential customers. These simpler configurations can sometimes be verified with little effort. The model should be such that in these cases the expected outcomes are reproduced.

singular cases: When looking at a graph for a dependency, there is a big chance that one or more zero-crossings occur[21]. A zero-crossing means that there is a value x_0 and a function f such that $f(x_0) = 0$. Also the point $(0, f(0))$ is a zero crossing. A zero-crossing often allows easy interpretation: what does it mean that $f(x_0) = 0$, can we understand the value of x_0, etc.. In the taxi model: a value of 0 for the profit means that our income exactly cancels the expenses. What happens, in this case, if there would be one ride more? This may be simple to calculate by hand; the outcome should be reproduced by the model.



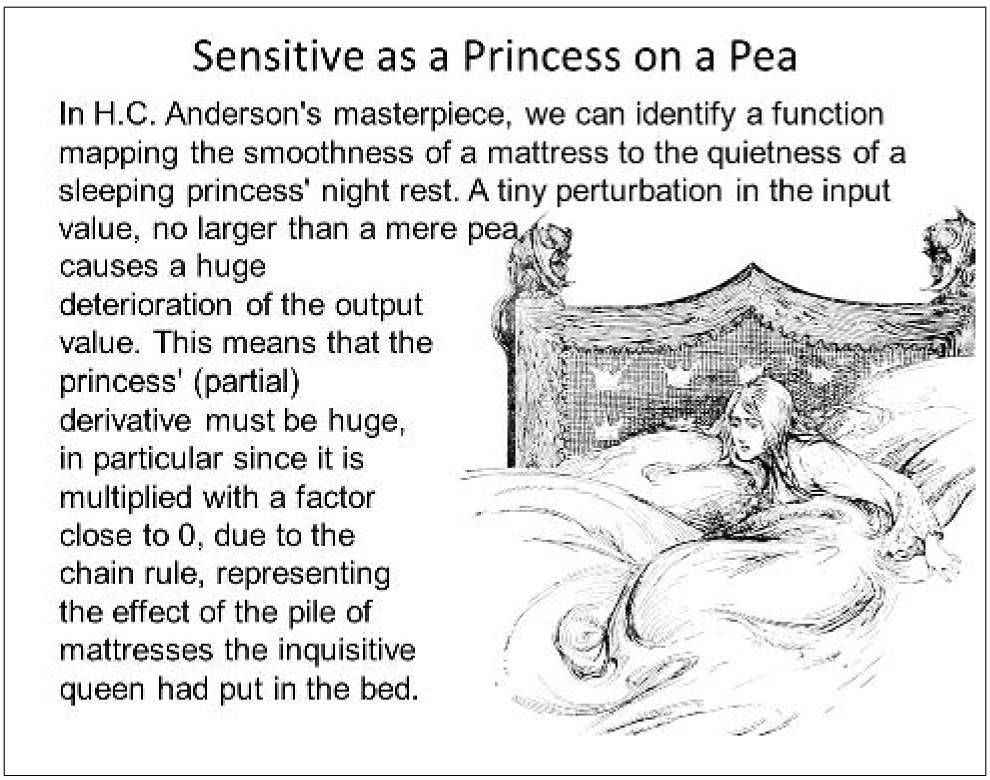
check convergence: Parts of a function evaluation may involve iterative calculation. Iterative calculation requires setting a number of steps and/or a step size. When the number of steps is insufficient, or the step size is too large, the result will be imprecise. To find out if this is the case, the calculation should be redone with a larger number of steps and/or with a smaller step size. In case the obtained answer significantly differs from the previous calculation, we must doubt convergence.

[21]The photograph of the singular sheep istaken from http://commons.wikimedia.org/wiki/File:Spot_the_odd_one_out%5E_-_geograph.org.uk_-_966073.jpg?uselang=nl

6.4.2 Quantitative Validity Assessment

If input values are fully precise, and for a model M_1 we have no better model M_2 such that M_1 is an approximation of M_2, there is no way to estimate the uncertainty of the outcome. The only source for uncertainty is then lacking accuracy, that is: the difference between the model and the modeled system. If the modeled system can be interrogated we can estimate the accuracy of the model. For example, a weather prediction model 'predicts' yesterday's weather, and we compare the 'predictions' with the actual data. This informs us about the uncertainty of our model. If such predictions show a *systematic* error (say, the predictions are always 1 centigrade too cold), we can even improve the model by adding 1 centigrade to the predicted value. Even though we don't know what causes this deviation.

Most often, however, we cannot asses the uncertainty in a model. For instance for a functional model for taking design decisions: the ATBD does not exist yet, so there is no way to assess the modeled system.



In cases where the uncertainty cannot be assessed by comparison with the modeled system, we still may be able to *estimate* this uncertainty.

Sensitivity Analysis

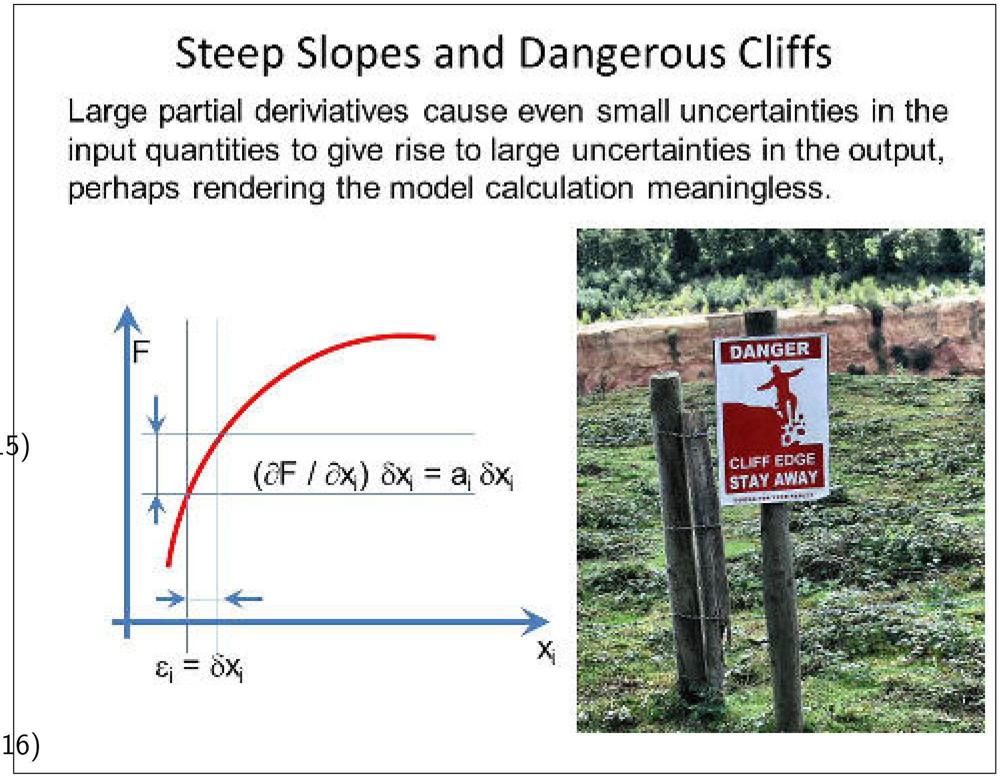
Suppose we have a number of input quantities x_i, $i = 1 \cdots n$ with uncertainty distributions with standard deviation σ_i. By x we denote the vector of all x_i: $x = [x_1, x_2, x_3, \cdots]$. Assume that the uncertainty distributions for these quantities are normal distributions with means $\mu_i = x_i$. Further assume that the function F, calculating one of the output quantities y, is differentiable in all arguments, and further assume that all the standard deviations σ_i are so small that we can ignore $\frac{\sigma_i^2}{\mu_i^2}$ in comparison with $\frac{\sigma_i}{\mu_i}$. This means[22] that we can develop a first order approximation to ERROR PROPAGATION. Error propagation regards the way uncertainties in inputs determine the uncertainties in outputs of a functional model.

[22]The image from 'The Princess on the Pea' was taken from http://commons.wikimedia.org/wiki/File:Page_190_illustration_b_in_fairy_tales_of_Andersen_(Stratton).png?uselang=nl

For one particular set of samples $\mu_i + \epsilon_i$ taken from the uncertainty distributions of the x_i, we get

$$\begin{aligned} F(\mu+\epsilon) &= F(\mu) + \frac{\partial}{\partial\mu_1}F(x)\epsilon_1 + \frac{\partial}{\partial\mu_2}F(x)\epsilon_2 + ... + O(\epsilon^2) \\ &= F(\mu) + \Sigma_{i=1\cdots n}\frac{\partial}{\partial\mu_i}F(x)\epsilon_i + O(\epsilon^2). \end{aligned} \quad (6.14)$$

We want to know[23] $\Delta F = |F(\mu+\epsilon) - F(\mu)|$. Indeed, the EXPECTATION VALUE of ΔF, when all of the $\mu_i + \epsilon_i$ are uncorrelated samples from the uncertainty ranges of the μ_i, is the standard deviation from the uncertainty distribution of y.



In first order in ϵ, ΔF is a linear function of all of the ϵ_i,

$$\Delta F = \Sigma_{i=1\cdots n} a_i \epsilon_i, \quad (6.15)$$

where $a_i = \frac{\partial}{\partial x_i}F(\mu)$. For ΔF linearly depending on normally distributed, uncorrelated quantities ϵ_i, such as Expression 6.15, the variance is given by

$$\sigma_F^2 = \Sigma_{i=1\cdots n} a_i^2 \sigma_{x_i}^2, \quad (6.16)$$

so

$$\sigma_F = \sqrt{\Sigma_{i=1\cdots n} a_i^2 \sigma_{x_i}^2}, \quad (6.17)$$

where σ_{x_i} is the standard deviation of quantity x_i, and σ_F is the standard deviation of y.

Often, we want σ_F as a percentage of y, and the σ_{x_i} are also commonly given as percentages of x_i. It does not mean much that an absolute uncertainty is 0.005 or 10^4. But we intuitively know that an uncertainty range of 0.005% is very accurate, whereas 100% uncertainty means that we have very little information. So instead of Expression 6.17, we calculate

$$\begin{aligned} \frac{\sigma_F}{y} &= \sqrt{\Sigma_{i=1\cdots n} a_i^2 \frac{1}{y^2}(\frac{\sigma_{x_i}}{\mu_i})^2 \mu_i^2} \\ &= \sqrt{\Sigma_{i=1\cdots n} c_i^2 (\frac{\sigma_{x_i}}{\mu_i})^2}, \end{aligned} \quad (6.18)$$

[23]The photograph from the warning sign is taken from http://www.rgbstock.nl/photo/mBZeK06/Gevaar+-+klif

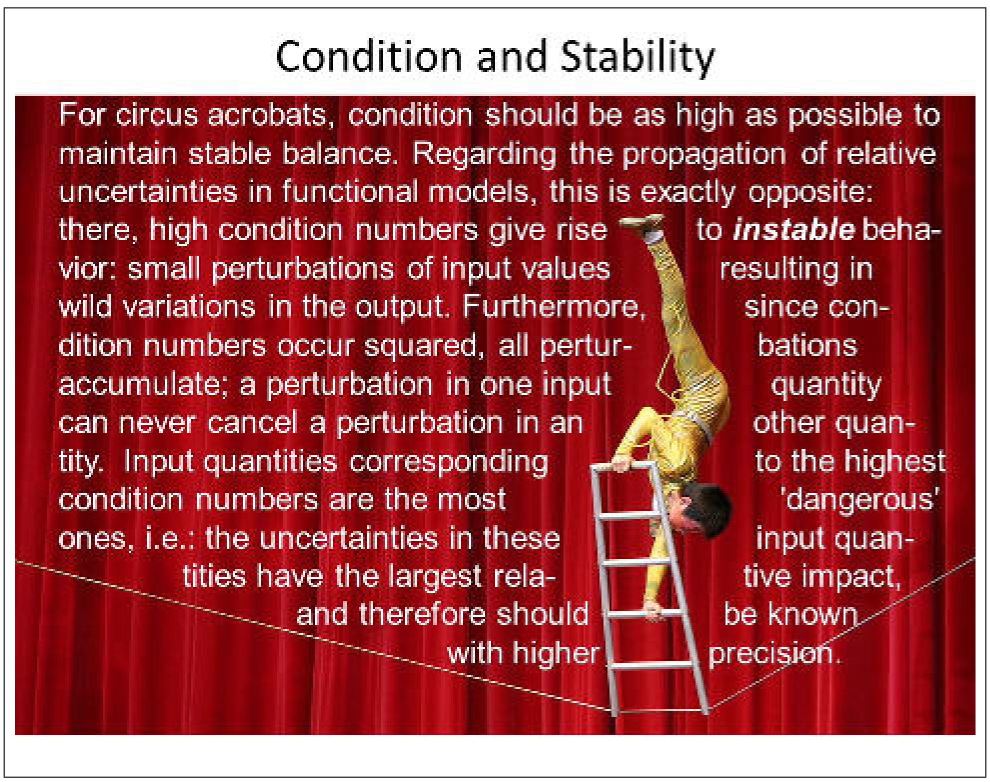
where $c_i = a_i \frac{x_i}{y} = \frac{x_i}{y} \frac{\partial}{\partial x_i} F$.

These $|c_i|$ are called the CONDITION NUMBERS. They are a measure for the sensitivity of F for a relative change in x_i. The calculation of the $|c_i|$ is called SENSITIVITY ANALYSIS.

A large value for a condition number indicates that the behavior of the model could have a problem with STABILITY, that is: a small change in the input could cause a problematic large change in the output. Since input quantities, e.g. as results of measurements, often have uncertainties, the numerical output of instable models should be treated with suspicion.

The result in Expression 6.18 is in some respects interesting:

Less conservative than adding absolute values: All relative error terms, σ_{x_i}/μ_i, occur squared, whereas the ϵ_i occur as linear terms in Expression 6.15. That means that their total contribution is smaller than when added linearly. Indeed, add n equal terms and compare this with taking the square root of their squares. The former is a factor $\sqrt{n}$ bigger. So the 'square root of squares' mitigates the effect of having to add n terms to each other: the total relative error does not grow proportional to n but to $\sqrt{n}$.



No cancellation of errors: All condition numbers, c_i, occur squared[24]. Sign cancelation cannot occur. Suppose we have two quantities, x_1 and x_2, one occurring with positive partial derivative and the other one with negative partial derivative. If the condition numbers instead of their squares would occur, the uncertainty contributions of both terms could partially compensate. If x_1 would be very uncertain, we could improve the certainty of the result by *increasing* the uncertainty of x_2. This is of course impossible, and it cannot occur because of the squared condition numbers, and the σ^2's being always positive.

One rotten apple ...: As a consequence, even a single very large relative error term could deteriorate the certainty of the entire model.

Search for large condition numbers: In order for a quantity to contribute much to the total uncertainty, its own relative uncertainty should be large and its condition number should be large. When seeking improvements for the uncertainty of a model, we should look for the quantities with

[24]The photograph of the equilibrist was taken from http://commons.wikimedia.org/wiki/Category:Acrobatics#mediaviewer/File:Acrobat.jpg

large uncertainties and large condition numbers.

In order to calculate Expression 6.18, we need estimated values for the $|c_i|$. In general, $|c_i|$ depends on x_i. The condition numbers can vary significantly over the input space. It is possible that in one regime, e.g., one part of the design space, the output is not at all sensitive to a relative change in some input quantity, whereas in another part it is highly sensitive.

Condition numbers require partial derivatives; these can be estimated as

$$\frac{\partial}{\partial x_1}F(x) \approx \frac{F(x1+h, x2, x3, \cdots) - F(x1, x2, x3, \cdots)}{h}, \tag{6.19}$$

and similar for the other x_i, where h is small.

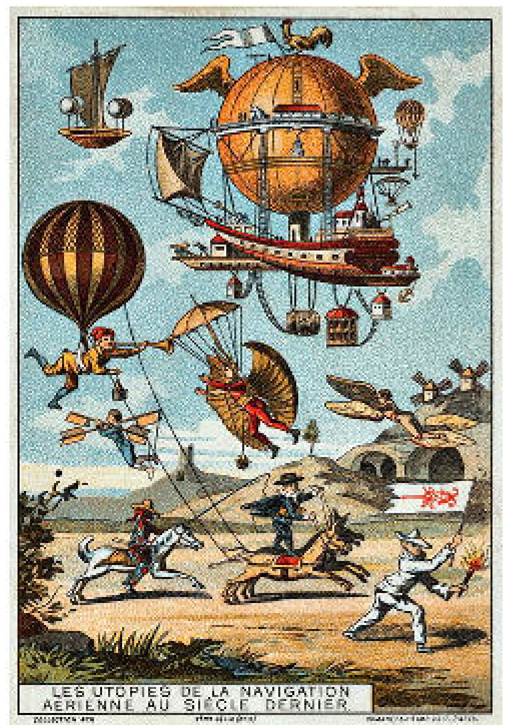
'Small' h means: the outcome of Expression 6.19 with h replaced by $h/2$ should not be different ▹122 .

In the numerical estimation of $\frac{\partial}{\partial x_i}F(x)$ we make no distinction whether F is differentiable or not. Always a numerical value of $\frac{\partial}{\partial x_i}F(x)$ results. If F is not differentiable in the interval $x_i \cdots x_1 + h$, the estimated value of $\frac{\partial}{\partial x_i}F(x)$ typically will be very large. Then the relative contribution to the uncertainty in x_i gets very large. This is a warning not to trust the model around this value of x_i. Most non-differentiable functions are only non-differentiable in few singular points: for instance, $|x|$ is everywhere differentiable except in $x = 0$. This means that in practice we can work with functions without having to bother about them being differentiable. Only when we by accident evaluate a function in a singularity, we get a very large uncertainty ▹123 .

Improving Models: With Enough Trust, Anything Flies

Developing a model is an iterative process. To reduce uncertainty in an early version of a model, it may be worthwhile to do a sensitivity analysis, and focus on cat.-III quantities corresponding to large condition numbers.

Providing improvements for those quantities has relatively large impact on the overall accuracy of the model.



The Contribution of Model Errors to Uncertainty; Model Improvement

With sensitivity analysis we find out how accurate a quantity needs to be to ensure a model outcome with better uncertainty than some given threshold[25].

Suppose that we want to know if some approximation is allowed when we need to meet a purpose. As before, we assume that the 'meeting the purpose' can be means that there is an upper bound

[25] The image with various improved balloon models is taken from http://upload.wikimedia.org/wikipedia/commons/e/e1/Early_flight_02561u_%282%29.jpg?uselang=nl

on the acceptable uncertainty of the output quantity(s). Further, suppose that the approximation we have in mind consists of replacing an expression by a constant, as is often the case. Perhaps the 'constant' is zero, for an additive expression, or 1, for a multiplicative expression ▹124 .

Call the new constant C, and pretend that it is an input quantity for the functional model. Assign an uncertainty range to it, corresponding to the range of values that the expression, for which C is a stand-in, could have. For example, consider a model for the efficiency of solar panels. An important input quantity is the solar radiation, but this varies during the day. We have an expression where the solar radiation depends on time. If we replace this expression by a constant C, the uncertainty range of C must correspond to the variation of solar radiation during the day. Than, using Expression 6.18, we calculate the uncertainty range of the entire model with the inclusion of the uncertainty range in C. It might be that the resulting uncertainty is still sufficient for the model's purpose, for instance, to verify if solar panels would be an effective substitute for wind energy. If so, there is no need to improve the model with the actual inclusion of the solar radiation *as a function of time*: the simplified model with C plus its uncertainty range is sufficient for our purpose.

In this way, sensitivity analysis can be used to check if model improvements are worth while.

6.5 Mathematical preliminaries

materiaal aan te vullen door Emiel ...

...

...

...

6.6 Summary

- Modeling involves *uncertainty* because of different causes:
  - *Accuracy* relates to the agreement between values, resulting from a model, and from the modeled system;
  - *Precision* relates to something that can be improved with effort, e.g., repeating a measurement more often, or taking more digits or more terms into account;
- All forms of uncertainty cause *distributions* of values rather than a single value;
  - The *normal distribution*: smooth, peaked, and with infinite wide support;
  - The *uniform distribution*: non-smooth, non-peaked, and with finite support;
- The notions of *distance* and *similarity*;
- The meaning of confidence for black box models:
  - Commonly occurring features of aggregation: *average*, *standard deviation* and *correlation*;
  - *Linear least squares* as an example of *aggregation*;

  - *Validation* of a black box model:
    * *Residual error*: how much of the behavior of the data is captured in the model?
    * *Distinctiveness*: to what extent can the model distinguish between different modeled systems?
    * *Common sense*: how plausible are conclusions, drawn from a black box model?

- The meaning of confidence for glass box models:
  - *Structural validity*: do we believe the behavior of the mechanism inside the glass box?
  - *Quantitative validity*: what is the numerical uncertainty of the model outcome?
    * *Sensitivity analysis* and the propagation of uncertainty in input data;
    * Sensitivity analysis to decide if a model should be improved.

6.7 Learning goals

6.7.1 Knowledge

You should know the meaning of the terms validation, verification, (uncertainty) distribution, normal and uniform distributions, accuracy, precision, distance, similarity, average, standard deviation, correlation, linear least squares fit, residue, sensitivity and condition number. You should know the relevance of these terms for assessing the confidence of black box models. You should know four tests for the confidence in glass box models.

Regarding the mathematical notions in this chapter, Emiel: aanvullen svp.

6.7.2 Skills

You should be able to distinguish accuracy and precision and to improve precision (e.g., of measurements). You should be able to calculate averages and standard deviations of data sets, and to perform linear least squares fits. You should be able to interpret these in the context of a model purpose. You should be able to assess the confidence in a glass box model of your own making using divide-and-conquer, asymptotic analysis, analysis of singular cases, and (when applicable) analysis of convergence. You should be able to perform a sensitivity analysis of a glass box model using (1st order) error propagation, both for absolute and for relative (percentual) errors.

For the mathematical notions in this chapter, Emiel: aanvullen svp.

6.7.3 Attitude

You should be inclined, for the numerical results of your model, to assess their accuracy and precision. You should be inclided to assess their plausibility by means of sanity checks (the Olympic gold medals-model is an example). You should be inclined, for the numerical outcomes of a glass box model, to perform both structural and quantitative confidence assessments (using the methods from Section 6.4.1). You should be inclined to check if your model requires improvement in order to fulfill its purpose, and if so, you should be inclined to use sensitivity analysis to find out which aspect(s) of your model should be improved.

6.8 Questions

1. In the opening paragraph, we say that 'we believe to understand much[...]'. Why are we so cautious, and don't we say 'we understand much [...]'?
2. What do we mean by 'condensed, aggregated' representations in the context of understanding bottlenecks in Internet traffic?
3. In your own words, why is verification more difficult than validation?
4. To explain verification, we introduce two conditions, called (1) and (2). Explain these conditions in your own words.
5. What does it mean that a mathematical expression follows uniquely from assumptions?
6. What do we need to do to verify if a prediction model holds in a particular situation?
7. In your own words, explain the difference between accuracy, precision and error.
8. Consider repeating an experiment.
   (a) What can we improve by repeating, the accuracy, the precision, or the error?
   (b) Under what assumption does the precision improve due to repeating?
   (c) Give a necessary (but not sufficient) condition on the outcomes of repeated measurements so that we can check if this assumption holds.
9. We state that a risk analysis always underestimates a risk. This claim is valid under the assumption that ...
10. Why can we say something about the direction of the error in risk estimation, involving multiple risk-scenarios, whereas we cannot in chemical or physical processes in case of multiple reaction pathways?
11. We talk about uncertainties in terms of distributions. What is a distribution? How does this use of the word 'distribution' relate to everyday use (as in distributing a bag of candy over a group of children)?
12. What is a normal distribution?
13. When does a normal distribution for uncertainties occur?
14. In a normal distribution, what can you say about the dimension of σ?
15. In a uniform distribution, what can you say about the dimension of σ_{low} and σ_{high}?
16. Compare the normal distribution and the uniform distribution: similarities, differences.
17. Why do we make the distinction between distributions and sets of samples in Section 6.3.1?
18. Discuss the behavior of Expression 6.5 for $n = 1$.
19. In your own words, explain the difference between average and mean.

20. In your own words, explain the difference between standard deviation and spreading.
21. Explain the difference(s) between using uncertainty distribution(s) for assessing if a threshold is exceeded, and for assessing which of two uncertain quantities is larger.
22. We compare a material A and a material B on two properties. Properties have uncertainties, and according to the text, we arrive at 9 cases.
    (a) Which 9 cases?
    (b) How many out of these 9 cases are conclusive?
    (c) What are the reasons for inconclusiveness in the other cases?
23. Give the numerical properties of distance and similarity as given in Section 6.2.
24. What is the triangle inequality for distances?
25. Which three versions of mean do you know?
26. In your own words, explain the meaning of the standard deviation.
27. Correlation is a measure for similarity. What does that mean?
28. Prove that correlations are between -1 and $+1$.
29. Explain the relation between linear least squares fitting and the method of residuals from Section 6.3.3.
30. Before we use the quantities a and b, obtained from a least squares calculation, we recommend in Section 6.3.3 to apply one of three methods for ascertaining the validity of these quantities. Which are these methods?
31. Explain in your own words the quintessence of using residual error as a means for validation of black box quantities.
32. What is the point of the Francis Anscombe data set?
33. In discussing Figure 6.2, we bring in an argument regarding clustering. Rephrase this argument in your own words.
34. What do we mean by 'striving for confidence intervals of 6 sigma'?
35. What do we mean by 'structural validity assessment' for glass box models?
36. Regarding structural validity assessment, we use the term 'divide and conquer'. What does that mean?
37. What do we mean by 'regime' in the context of structural validity assessment? Give examples of regimes in the taxi company example.
38. Rephrase in your own terms what we mean by asymptotic analysis.
39. What is the purpose of a condition number?

40. Explain, without using mathematics, why the relative error when adding n uncertain terms is not proportional to n.

41. What do we mean by 'cancellation' in Section 6.4.2?

42. To calculate condition numbers, we need partial derivatives.

    - What is a partial derivative?
    - In order to numerically estimate a partial derivative, the independent quantity (=the argument of the considered function) needs to satisfy a condition. Which?
    - For a numerical estimate, we need a stepsize h. How can we assess if a value for h is adequate?

43. A functional depedency $y = f(x)$ has a constant condition number C, for all x. What do you know about f?

44. Verify the expressions for the condition number for the case of exponential and logarithmic functions.

6.9 Exercises

1. Give an example of a model and a modeled system and two purposes, P_I and P_{II}, such that we can be confident for purpose P_I whereas we cannot be confident for purpose P_{II}.

2. Consider a secondary school physics assignment about a statical configuration with ropes and pulleys. In such assignments, the pulleys are often said to be massless. In case of static balance, all accelerations are 0, and all masses sum up to zero for each pulley. Why is the addition 'massless' for the pulleys still necessary? What if pulleys are ideally smooth (no friction)?

3. We claim that a simple rigid slope model is insufficient to conclude anything about the safety for a ship launching from a slide in a shipyard. Give at least two important factors or effects that are disregarded in this model when applied to launching a ship.

4. We present a meterological model as an example of a model that can be fit for purpose, but nevertheless gives no confidence when applied in some context (e.g., to predict next month's weather). Give another example.

5. We present the chimney sweepers problem in a particular context as an example of a model that gives too weak results, and therefore is invalid. Give a similar example, taken from the peanut butter problem.

6. Extend Table 6.1 for other purposes.

7. Give an example of a modeled system, a model and a purpose where accuracy, precision and eror are three diffrent aspects of uncertainty.

8. How do we extrapolate to x_∞ given a sequence of approximations $x_p = f_x(h/2^p)$. How do we check if this extrapolation is allowed?

9. Adding a term or a factor to a model to account for yet another mechanism does not necessary reduce the uncertainty of the model outcome. Give an example.

10. Show that the normal distribution from Section 6.1.2 is normalized to 1.

11. Look up (in the Internet) the definition of the binomial distribution. Compare the normal distribution and the binomial distribution: similarities, differences, use, assumptions ...

12. Given two quantities, each with a given (not necessarily equal) normal distribution. Give an expression for the chance that the expectation value of one is larger than the other.

13. Consider two partially overlapping rectangles with equal area but different width:height ratio's. Show that their intersection has a width:height ratio that equals the geometric average of the two width:height ratio's.

14. We discuss three definitions of average. Discuss their differences and similarities. Which should be used when?

15. In the derivation of the standard deviation, we use the arithmetic average. There are two other types of averages, though. Give definitions of the standard deviation using the other two definitions.

16. We derive the formula for linear least square fitting for a function of the form $y = ax + b$. Make a derivation for a fit of the form $y = a_1x_1 + a_2x_2 + b$.

17. In the derivation of the least square fit, we assume that the y-errors in each of the data points (x_i, y_i) have equal impact. If the data points result from measurements, hoewever, it could be that some points are measured more accurately than others. Give an improved least squares fit that takes this into account.

18. There is a number of conclusions that seem to follow from Figure 6.2 but that don't make sense from circumstantial arguments. Give at least two conclusions that *do* hold.

19. When discussing the role of clustering in the context of Figure 6.2, we raise the issue about fairness. Give arguments in favour of fine-grained clustering and arguments against it.

20. Safety is sometimes defined in terms of the chance for an incident happening in X years. Give arguments to assign a substantiated value to X.

21. The section on structural validity assessments for glass box models contains an example regarding the quantities `flPpl` and `profit` from the taxi company model.

    (a) Perform the experiment as explained in the text for a number of different taxi company configurations (a taxi model configuration is a series of assignments of values to ca.t-I quantities in the taxi company model), and verify if the model behaves as expected.

    (b) Find an other pair of quantities to perform such an experiment and do the experiment with that pair.

22. Find and demonstrate at least three examples of asymptotic analysis in the taxi company model.

23. Explain that the detergent problem and the chimney sweepers problem have no non-trivial singular cases singular cases, where a trivial singular case means that some cat.-II quantity is zero or infinity. Think of a minimal extension to the chimney sweepers model so that it has a singular case, and analyse this case.

24. In Chapters 1, 2 and 4 we introduce the street illumination example. In a full elaboration of this model, the issue of convergence would occur.

    (a) Why?

    (b) How would you check if convergence is achieved?

25. We consider improving the accuracy of a model by replacing a cat.-III quantity by a cat.-IV quantity. Give a heuristic for selecting a candidate cat.-III quantity to replace.

26. Condition numbers in an all-positive additive expression are all less than one. This suggests that we can make any additive model arbitrary precise by replacing all terms by summations of other terms. Discuss this suggestion.

Chapter 7

A Working Model - and then?

'Starting is easier than stopping'

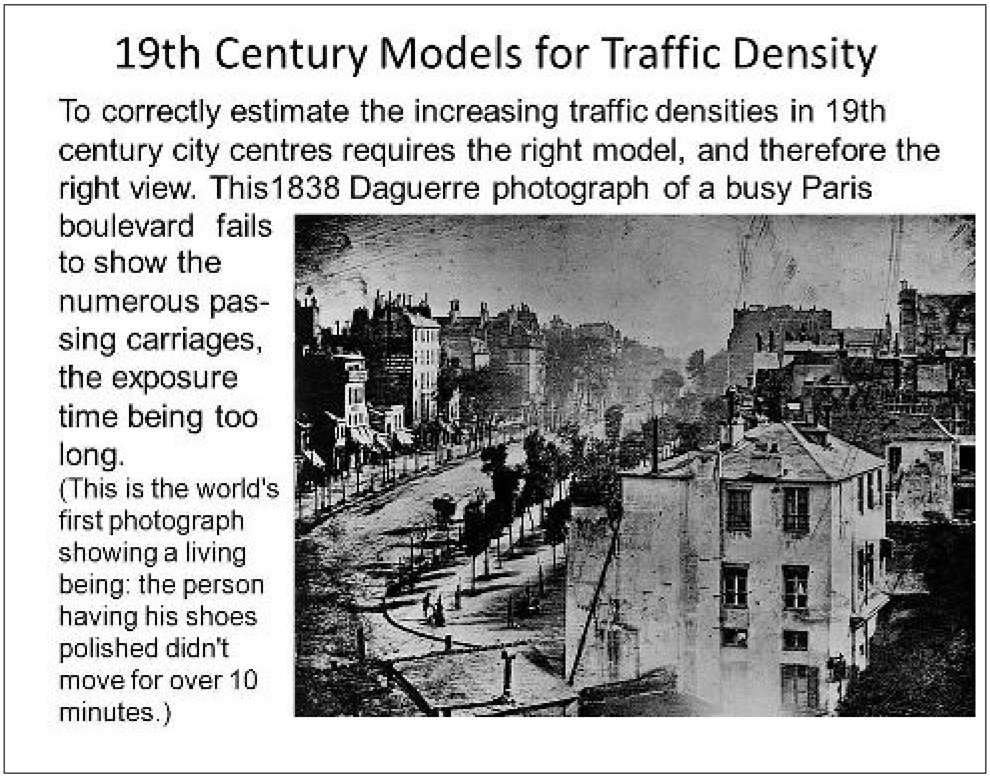
At the end of the 19th century, the world's largest cities saw a devastating disaster coming up in the foreseeable future. Models showed, for instance, that in as little as 50 years London's streets would be covered by no less than 9 feet (about 3 meters) of insects-attracting and diseases-spreading horse manure. The area to stable all horses would use up ever increasing amounts of farmland, and even more would be needed to grow hay for feeding them. Indeed, London at that time hosted some 11000 cabs, and their amount was carefully recorded over the preceding few decades; extrapolation was easy, and the results spelled inevitable doom over urban civilization. In New York, where in 1898 the first international urban planning conference was held, delegates concluded, after only three days of conferring, that such problems were insolvable, gloomily skipping the remaining 7 days of the planned program [▷125] , perhaps going home to prepare for disaster.

7.1 The Need for Interpretation

In the above example, there was presumably little wrong with the mathematics. The outcomes of the equations might be accurate and precise. The model developed was a black box model, not too dissimilar from the model for the development of men's and women's world records on 200 meter sprint, discussed in Chapter 6. The mathematical techniques to deal with least squares approximations were fully understood by the 19th century.

What was wrong was the *interpretation* of the outcomes of the model. The number of horses would *not* continue to increase exponentially the same way it did the previous decades. It would rather see a sharp decline due to the unforeseen advance of automobiles, only recently invented and at that time an insignificant factor in the arena of public transport.

So the outcome of the calculations of the London city clerks was accurate and precise. But, as we stated in Section 1.4.5, the outcome of a model calculation is not the solution of the problem for which the model was developed. There is always the need for *interpretation*: what is the implication of the numbers for the *initial problem*?



In this chapter we address the conclusion stage of the modeling process, connecting the model and the model context which contain the modeled system, the stakeholders and the model's purpose.

A problem is not solved by a model alone[1]. We only have solved a problem atfer completion of the interpretation stage. Moreover, a solution is never absolutely right, because models are not absolutely right. Instead, we may say that one solution is *better* or *worse* than another. Or we can say that the model needs an improvement, in case the model is not good enough to fulfill its purpose. An improvement may make a present solution *better* without guarantee that the new solution will be good.

For all of this, we must be able to *compare* solutions.

Comparison needs *criteria*. It will be our aim to propose a set of criteria to compare models and model outcomes by.

7.2 Criteria for Modeling

Proposing a meaningful set of criteria for modeling is difficult because the variety of modeling approaches is immense. This variety is spanned both by the variety in purposes in Section 1.2 and the variety of modeling approaches from Section 1.3.

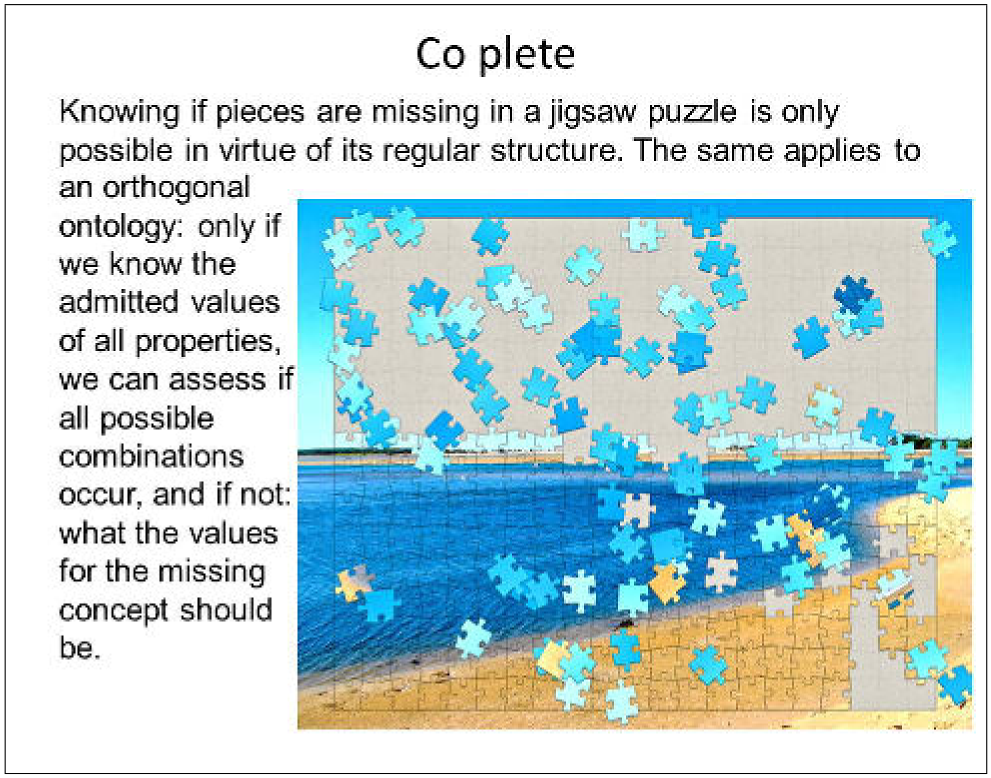
For clarity, we build the collection of criteria-*for-the-modeling-process-as-a-whole* ▷126 using few orthogonal properties. This is an exercise in applying Section **??** from Appendix **??**.

We seek few meaningful properties, that apply to *any* modeling situation, and that are ENUMERABLE, that is: all possible values of the properties can be enumerated. For example, boolean

[1]The Daguerre photograph eas taken from http://commons.wikimedia.org/wiki/File:Boulevard_du_Temple.jpg?uselang=nl

properties and integer properties are enumerable; the `aggregation-phase` of matter is an enumerable property because it can only take the values `gas`, `liquid`, `solid`. A property such as `shape` is not enumerable.

A taxonomy[2]build with only enumerable properties has the advantage of being complete COMPLETE. In a complete taxonomy it is possible to give any concept a unique location.



The first distinction is, whether the criterion relates to the *definition stage* of the modeling process (stage 1) or to the *conclusion stage* of the modeling process (stage 5, see Table 1.3).

The definition stage contains the problem definition and the intended stakeholders: who are they, what do they want and expect. The conclusion stage contains the presentation of the model outcome and its interpretation, use and usefulness in the modeling context: to what extent is the purpose of the model fulfilled.

A second distinction is, whether a criterion relates to the model and the modeled system (to be called 'inside') or to the stakeholders and the context (to be called 'outside').

A third distinction is, whether this criterion relates to a qualitative or a quantitative aspect of the model.

Any criterion serves to distinguish 'better' and 'worse' models. A criterion therefore needs to be ordinal: the type of any criterion cannot be a nominal set. By 'quantitative' *vs* 'qualitative', therefore, we don't refer to the type of values for this criterion, but rather to the aspect of the modeling process this quantity represents.

This gives a total of 8 combinations of (definition *vs* conclusion), (inside *vs* outside), and (qualitative *vs* quantitative). Table 7.1 lists the 8 resulting criteria. They are further explained in Sections 7.2.1 to 7.2.8 ▷127 .

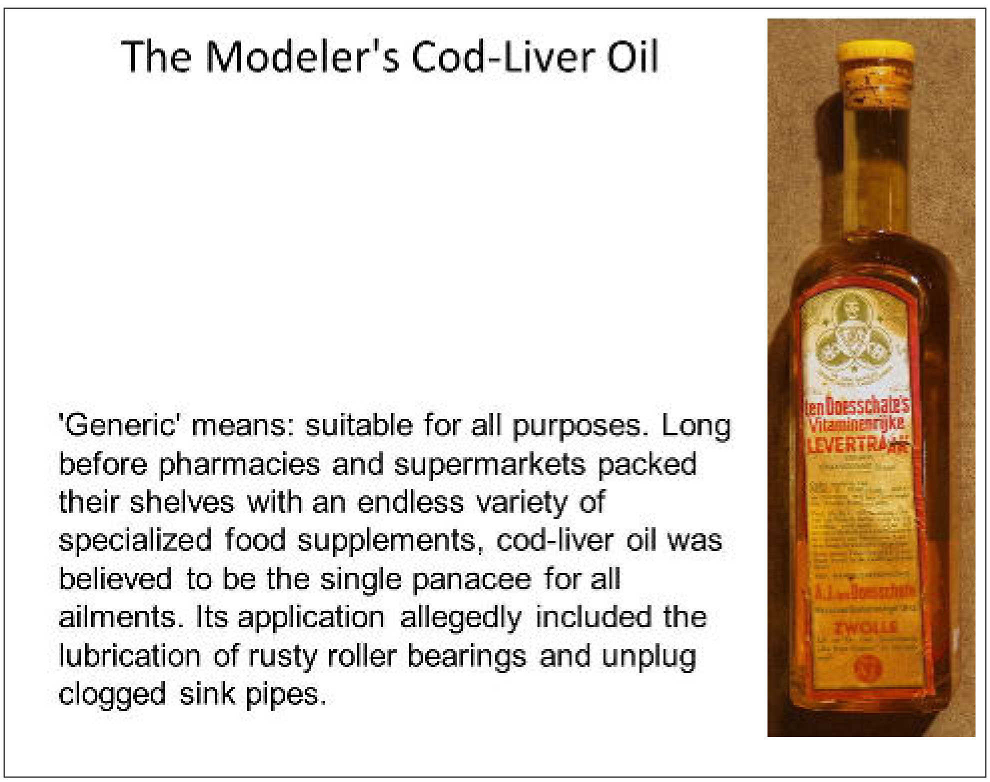
[2]The jigsaw puzzle image was taken from http://www.rgbstock.nl/photo/n68c9Zy/Holiday+Jigsaw+Puzzle

Definition or Conclusion stage	Inside or Outside	Qualitative or Quantitative	Criterion
Definition stage	Inside	Qualitative	Genericity
		Quantitative	Scalability
	Outside	Qualitative	Specialization
		Quantitative	Size
Conclusion stage	Inside	Qualitative	Convincingness
		Quantitative	Distinctiveness
	Outside	Qualitative	Surprise
		Quantitative	Impact

Table 7.1: Criteria for the Modeling Process

7.2.1 Definition Stage, Inside, Qualitative: Genericity

A GENERIC model can be used for a large variety of modeled systems[3], and perhaps even for a large variety of purposes. With respect to the criterion 'genericity', model A is better than model B if model A works for more different kinds of situations.



We give some examples:

- in data fitting, a model A is more generic when it has additional adjustable quantities compared to a model B. A quadratic model, $y = ax^2 + bx + c$ has three adjustable quantities (a, b and c), and it can fit both parabolic dependencies and linear dependencies. The model therefore is more generic than a linear model, $y = ax + b$. Whether a particular data set *should* be fitted by a linear or a quadratic model is a different matter. It may be that a linear model has a larger residual error, but still the improvement gained by a quadratic model is not significant because, e.g., there is no intuitive explanation for the quadratic term, or because the residual error with another 3-quantity model such as $y = ax^3 + bx + c$ would be much less.

Whether modifications of a model do increase its genericity depends on the purpose of the model.

- in data models, a so-called TRIPLE STORE is more generic than a table-based database, because any given table can be represented by triples, but an arbitrary collection of triples can

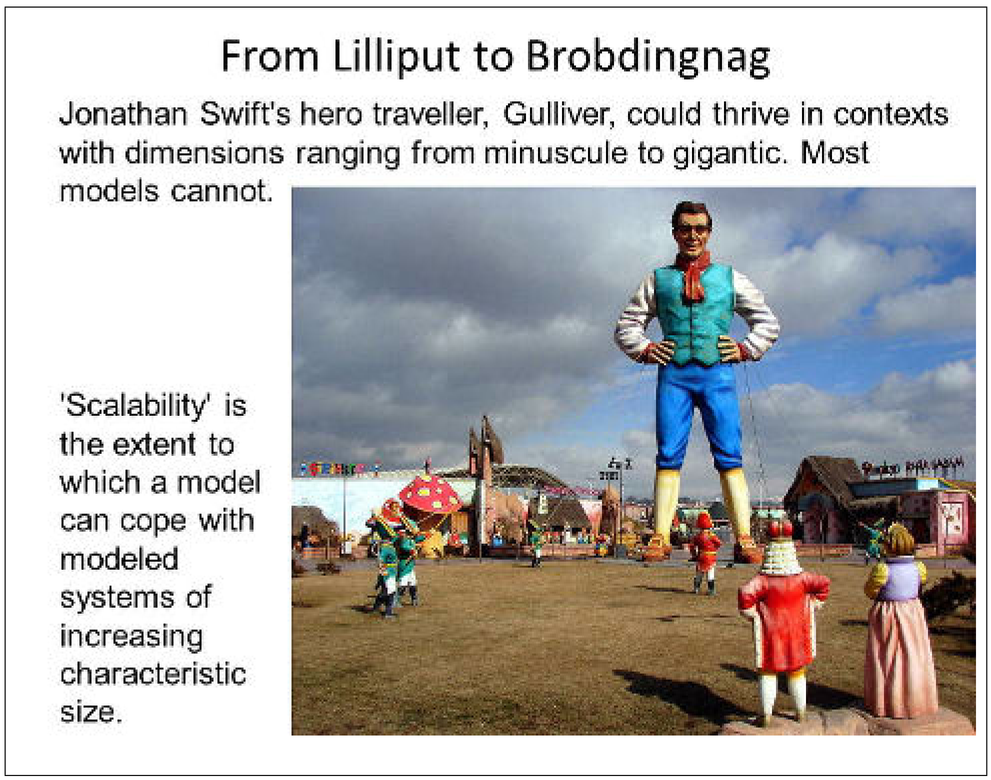
[3]The photograph of the cod-liver oil bottle was taken from http://upload.wikimedia.org/wikipedia/commons/a/ac/Ten_Doesschate%27s_Vitaminerijke_Levertraan.JPG?uselang=nl

not be represented by a given table structure.

- in physical models describing the flow of continuous media (liquids, gases), a model which allows the medium to be compressible is more generic than one where the density is constant. The latter can be adequate for the flow of water through pipes but it cannot be used to calculate energy loss in a hydraulic system where part of the energy is converted into heat due to compressing or decompressing of the oil.
- weather models are generic because the physical phenomena, represented in the equations, range from flow, thermal effects, to radiation. For efficiency, however, such models have much optimizations built in. For instance, the geometry of the earth may be hard-coded, and therefore a weather model cannot solve a problem involving airflow through a set of pipes, even though the physics is simpler. The two models (weather models and a model for air flow through pipes) therefore cannot be compared with respect to genericity.

7.2.2 Definition Stage, Inside, Quantitative: Scalability

A SCALABLE model can be used for a range of sizes or SCALES for a modeled system. The distinction between Scalability and Genericity is, that Scalability refers to homogenous quantity (many items, all of the same sort), where Genericity relates to the heterogenous quantity (many items, all different sorts).



The *scale* of a problem[4] is a number, say n, characteristic for the size of the problem. For instance, for database-related problems, the scale could be the number of tuples in the database; for a simulation of an avalanche, the scale could be the number of modeled particles.

Rather than in the precise value of n, modelers are interested in the behavior of the solution time of the problem when n increases. Instead of the solution time, the determining number can also be the amount of required memory space in a computer. Limitations on the scale of a modeled system relate to the PERFORMANCE: sometimes, the modeled system is so big that execution of

[4] The image of Gulliver in Lilliput was taken from http://commons.wikimedia.org/wiki/Category:Gulliver#mediaviewer/File:Harikalar_Diyari_Gulliver_Lilliputians_06034_nevit.jpg

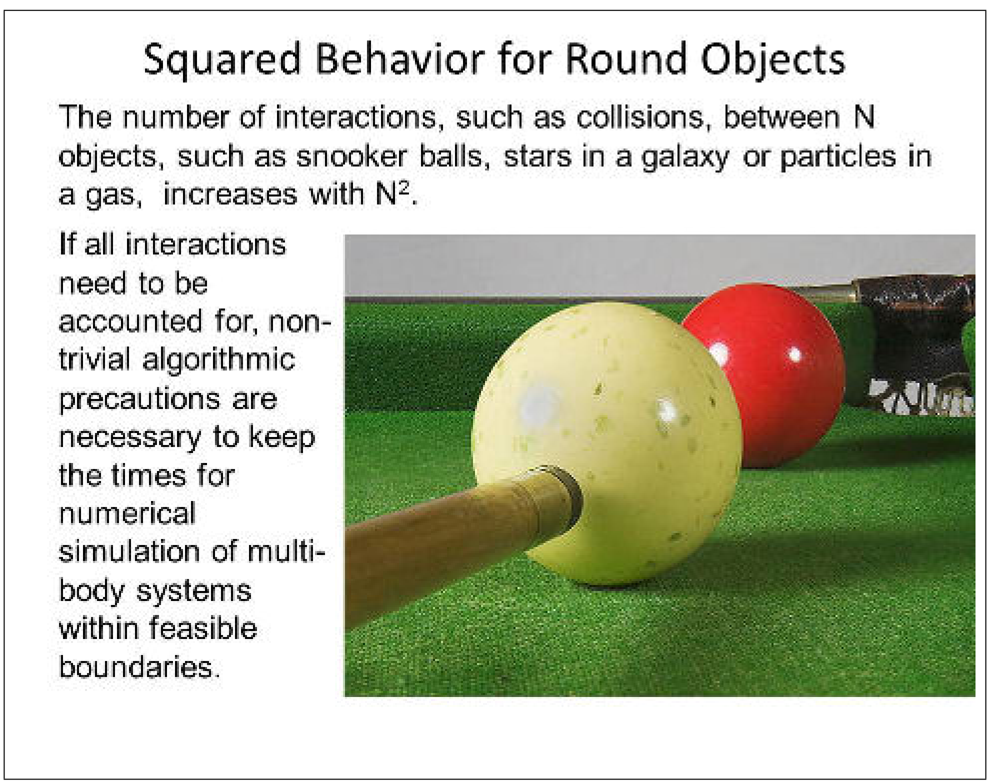
the model takes unacceptably long; in other cases, it may require too much computer memory.

Most models require at least $O(n)$, for instance to input n items of data. The notation $O(n)$ means that, if a problem with n_1 items takes time or memory size t_1, a problem with n_2 items requires $t_2 = \frac{n_2}{n_1} t_1$. For $O(n^2)$ this would be $t_2 = (\frac{n_2}{n_1})^2 t_1$, and similar for other expressions in n.

Sorting n items can be done in $O(n \log n)$ steps. Inverting an $n \times n$ matrix takes $O(n^3)$ steps when using a straightforward method, but faster approaches exist. Many queries on data require $O(n^2)$ or worse unless the data is sorted or INDEXED, when performance drops down to $O(\log n)$ or even $O(1)$. The latter means that the query time is independent of the number of stored items. Google queries don't get slower over time, despite that the amount of data, stored in the Google repositories rapidly grows. *Indexing* means that additional data is kept to speed up often occurring queries. For example, to speed up free text search in a large body of documents, an index is often generated that consists of all occurring words, with, for each word, references to the documents where this word occurs.

With respect to the criterion 'scalability', model A is better than model B if model A works well for modeled systems with a larger number of components.

We give some examples:



- an approach to the simulation of continuous systems is to use particles[5]. Every particle represents a small amount of the continuous medium. Particles interact with each other, for instance they attract nearby particles to imitate viscosity, and collide with them to represent pressure. The number of interactions between n particles is $\frac{n(n-1)}{2}$, so a straightforward implementation of collision detection has a running time proportional to $O(n^2)$. There is a serious performance limit: the number of particles cannot be too large.

  A more advanced approach exists where the space in which particles reside is partitioned in cubic cells. Every cell contains only few particles. Interactions between particles in remote cells are neglected. Then the complexity reduces to $O(n)$ and the advanced model can accommodate a much larger number of particles: it is a much more *scalable* model.

[5]The photograph of snooker bal is taken from http://commons.wikimedia.org/wiki/Snooker#mediaviewer/File:To_pot_the_red.jpg

- there is a large class of problems for which solution take an amount of time, which increases faster than polynomial in the size of the problems. Examples are $O(2^n)$ or $O(n!)$. A well-known instance is the so called TRAVELING SALESMAN PROBLEM: given n points on a 2D map; what is the order in which they should be visited so that the total route has minimum length. It is believed ▷128 that the traveling salesman problem cannot be solved in an amount of time that grows only polynomially in the size of the problem. This approach scales very badly. Increasing n with 1 means more than *doubling* of the amount of time. If we drop the requirement that the route should be minimal, however, there are many heuristics that perform well. Much research in information science and computer science is devoted to approximate, but scalable, solutions of problems of this nature.
- a modern class of problems where scalability is an issue relates to the Internet and social media. The linkedIn network, for instance, is a data structure that maintains relations between people. These relations can be seen as a graph. Nodes are people and edges are relations between two people. If we assume that the average number of connections for a single individual is constant, this is a so-called SCALE-FREE network. If the network grows, the average number of connections between any two nodes then stays constant. This is an extreme form of scalability.
- models for discrete dynamic systems often are based on state charts (see Section 3.2.1). For the verification of such systems, a naive approach is to traverse the entire state space. Such algorithms scale very poorly. The number of states in a state space with n quantities is proportional to $O(2^n)$.

Only for the Initiated

According to Christmas folklore, three kings followed the star to Bethlehem. In more scholarly substantiated versions, it is believed that they were in fact oriental magi or astrologers.

In this view, the wandering star was a (material) model with the purpose to communicate. Since it conveyed only a single message, it had low genericity; since it could only be understood by the magi, its level of specialization was extremely high.

7.2.3 Definition Stage, Outside, Qualitative: Specialization

The term 'SPECIALIZATION' here is an abbreviation of 'Level of specialization of the intended problem owner[6]'.

The amount of specialization of the intended stake holder to benefit from the model outcomes can work in two, opposite directions.

[6]The image of the magi was taken from http://upload.wikimedia.org/wikipedia/commons/a/a8/DPAG-1997-Weihnachten-AnbetungDerKoenige.jpg?uselang=nl

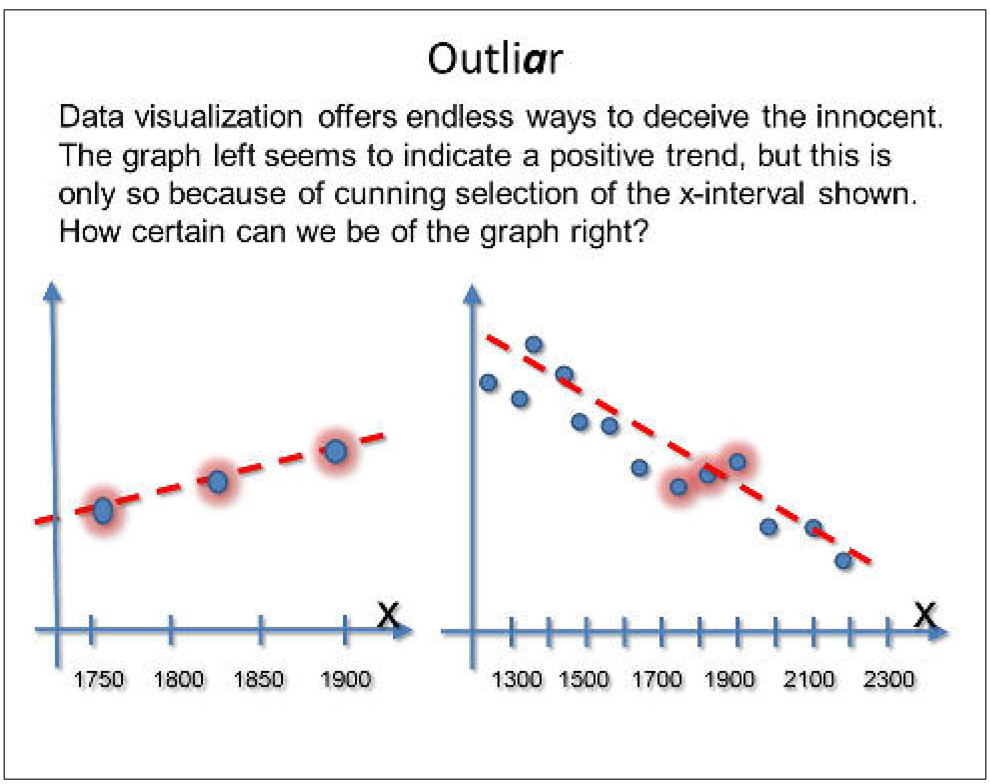
A first interpretation is, that if a model is capable of satisfying a more specialized audience, it performs better. Then the distinction between the modeler and the problem owner blurs. If the problem owner is very much a specialist in modeling, the problem owner him- or herself may do the modeling. Then the requirement that the model should comply to the knowledge level of the problem owner becomes equivalent to the requirement that the model should satisfy its own maker. This is a trivial requirement.

Therefore we will use the second interpretation. If a model fulfills the purpose and does so to the appreciation of laymen, it is better than a model which needs a certain amount of expertise from its users. So a model, with respect to the criterion 'level of specialization of the problem owner', performs better if users of the model need to know less.

The latter meaning of the criterion 'specialization' also implies that the model plus the interpretation of the outcomes must safeguard against common misconceptions. For instance: a weather map shows the regions of low pressure. Specialists can deduce from such maps the direction of the wind ▷129 . Leaving out wind arrows makes the map cleaner and provides room for other information, useful to meteorologists. The weather maps used in the public television weather reports, however, should show wind arrows.

For the modeler, to make his model outcomes excel with respect to the criterion 'specialization', there are a number of guidelines.

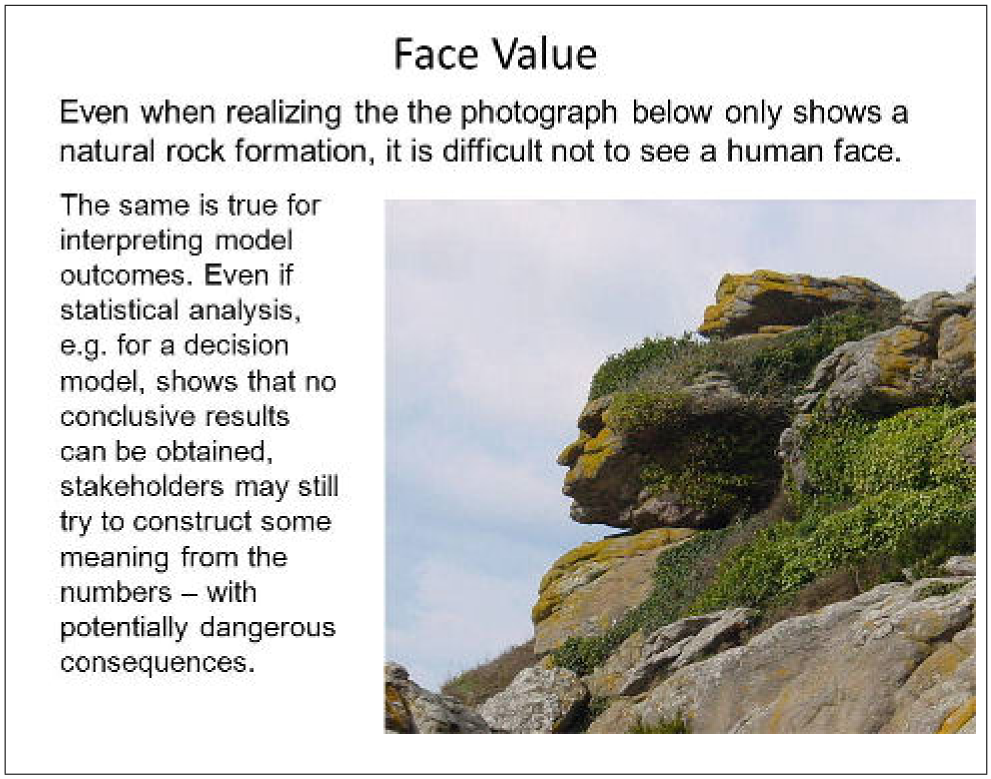
invest in presentation: DATA VISUALIZATION is a sub discipline from communication science, focussing on the effective design of visuals to convey the meaning of data. Pie charts, bar charts, 3D plots and more advanced forms of mapping data to images are part of data visualization.



The modeler does not need to be an expert in data visualization, but (s)he should be familiar with the basic principles. These include selecting the appropriate type of graphic style, the proper choice of scales, labels and caption placement, the use of error bars, the pro's and con's of combining multiple data in a single graph, etc.. For those who are not familiar with these notions, the 1954 Daniel Huff classic 'How to lie with statistics', and the more recent series of texts by Edward Tüfte on data visualization are recommended.

The quality of data visualization is determined by tens of detailed choices. A creative idea to present data may help understanding patterns in a body of data. The essence of data visualization, however, amounts to answering a single question: 'which *message* should this graph communicate?'. Often this will *not* be a numerical message. It is not interesting that a particular value is 17.3; it can be interesting that this value is not larger than 18. Or it can be interesting that it has been 17.3 already for the last 5 days. Or it can be interesting that it is exactly half of the value of some other quantity. Or it can be interesting that everybody assumed that it should be 17.4, and numerous other messages. Once the message is clear, the choice for the right visualization style, and all other choices usually follows with little effort.

when in doubt, limit the claims: Many people have some difficulty in interpreting uncertainty[7]. What does it mean, for instance, if a weather prediction announces 15% chance of overcast weather: does that mean that the sky everywhere will be overcast for 15% of the area; will it be overcast for 15% of the day, or will 15% of the country have a grey day?



Since models use approximations, they rarely give exact answers. Many models show the reliability of their answers using some indication of uncertainty. Taking uncertainty distributions into account, it is sometimes possible to stretch the distinctive power (see Section 7.2.6), and reach a stronger model outcome. It is wise, however, not to emphasize such effects if the chance of misinterpretation is large.

For instance, consider a model to compare two products, A and B with respect to two features, p and q. For both features, we have 'the larger the better'. Suppose the expected value of $A.p$ exceeds the expected value of $B.p$ by 20%, and the expected value $B.q$ exceeds the expected value of $A.q$ by 1%. Many would prefer A over B.

If the standard deviation of p, however, is 50 whereas the standard deviation of q is 0.0005, the conclusion might just as well be reversed. In such situations it is safer to declare the model's outcome to be inconclusive. Even though a statistician might give a meaningful interpretation to the outcome ▷130 .

The general advice is: be more conservative with the interpretation of the model when the problem

[7]The photograph of the rock formation was taken from http://upload.wikimedia.org/wikipedia/commons/1/1e/Visage_dans_un_rocher.jpg?uselang=nl

owner is less experienced. Being CONSERVATIVE here means: tune down the distinctive power of the model outcome to safeguard less-informed problem owners against overly extreme conclusions.

be warned for biased stakeholders: The guideline 'when in doubt, limit the claims' applies to less-informed, but neutral stakeholders. Stakeholders, however, are not always neutral with respect to the problem outcome[8]. They may have interest in one or the other outcome.

For instance, an entrepreneur who wants his business model to be approved by a bank, and asks a business consultant to analyse his proposal using a model, will hope for optimistic profit estimates. An airline company ordering a risk analysis for its fleet will be glad if the risk for incidents turns out low.

Most model outcomes contain assumptions that affect both the model outcome, and the interpretation of the outcome. In an ideal world, the modeler would build a model and give an interpretation that are VALUE-FREE. 'Value free' means that no assumptions are biased in favor of, or opposed to stakeholders' interests. So all estimates and approximations should be such that a neutral outcome results. In practice this is impossible. Not all assumptions can be known, consequences with respect to stakeholders' interests cannot be fully overseen, and not all stakeholders agree with respect to their interests.

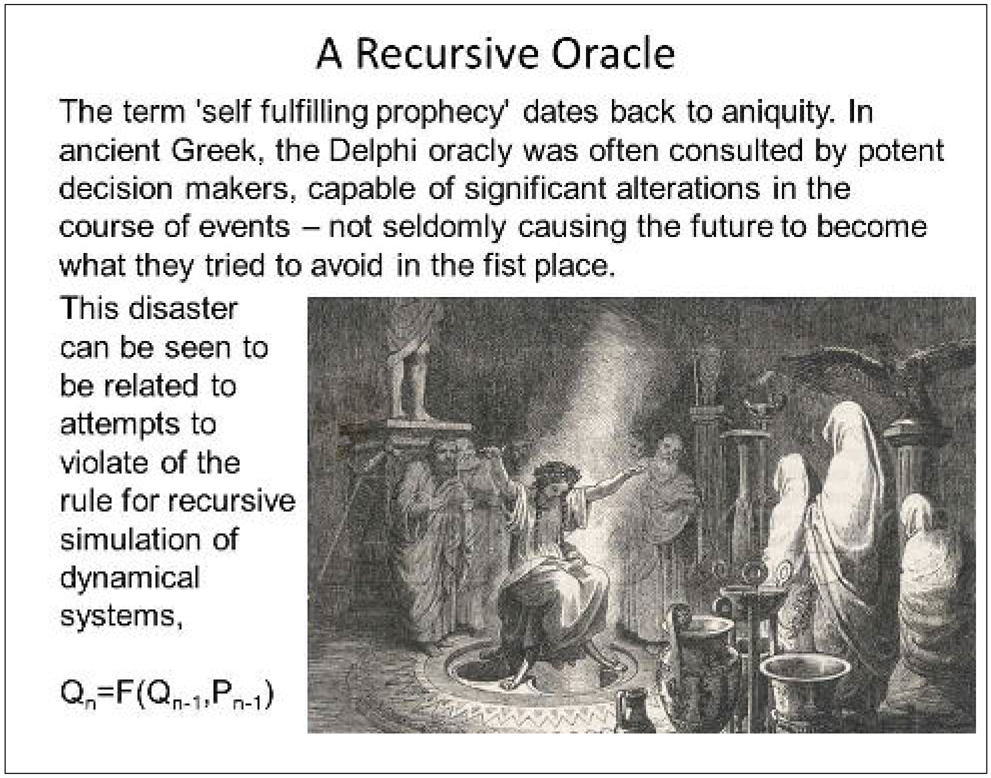
To deal with stakeholders' bias, we don't advocate a strategy of deliberately counteracting any known stakeholders' biases, for instance by tweaking uncertainties of certain assumptions and estimates. Indeed, the modeler then would also be guilty of 'lying with statistics'. Rather, a modeler should be aware of the risk that a model output is given a non-neutral interpretation, and (s)he should practice damage control by

- giving a detailed account of all assumptions and approximations in the final report;
- insisting that the entire report is kept as an indivisible document;
- refrain from providing graphs or other 'easy-to-misapprehend' results that might support known stakeholders' biases;
- seek assistance from a neutral colleague or ask an independent second opinion.

[8] The photograph of terrace tables in the snow was taken from http://commons.wikimedia.org/wiki/File:Eternal_optimism.jpg?uselang=nl

self-fulfilling and self-denying predictions: So-called WICKED PROBLEMS are problems that are changed by their own solution.

For example[9]: beginning in 1939, the introduction of DDT in agriculture aimed at solving the problem of crop size reduction due to insects feeding. The reasoning was, that by decrease of insects populations, the damage done to harvests would be mitigated. According to the knowledge, present at that time, this reasoning was sound, and models based on it predicted succes. It was not known, though, that insects mutation rate is very high. In few generations insect varieties emerged that were resistant to DDT. Bird, feeding on DDT-resistant insects, accumulated the poison in their body tissue, and were killed: since higher animals don't mutate at the same pace insects do, resistant bird species did not evolve. So, not only did insects no longer die from DDT, their natural predator gradually lost terrain, causing insect populations to grow rather than shrink. DDT turned into a pest rather than a pesticide in few decades. The net effect of deploying DDT was crop size reduction instead of crop size increase.



Such causality loops are very common: they give rise to predictions either causing their own fulfillment, or their own denial. Unfortunately, we read mostly of predictions of doom that are self-fulfilling, and predictions of glory that are self-denying. Modelers ought to be aware of a potential of self-influence in their models, in particular when model outcomes may have substantial impact (see Section 7.2.8).

7.2.4 Definition Stage, Outside, Quantitative: Audience

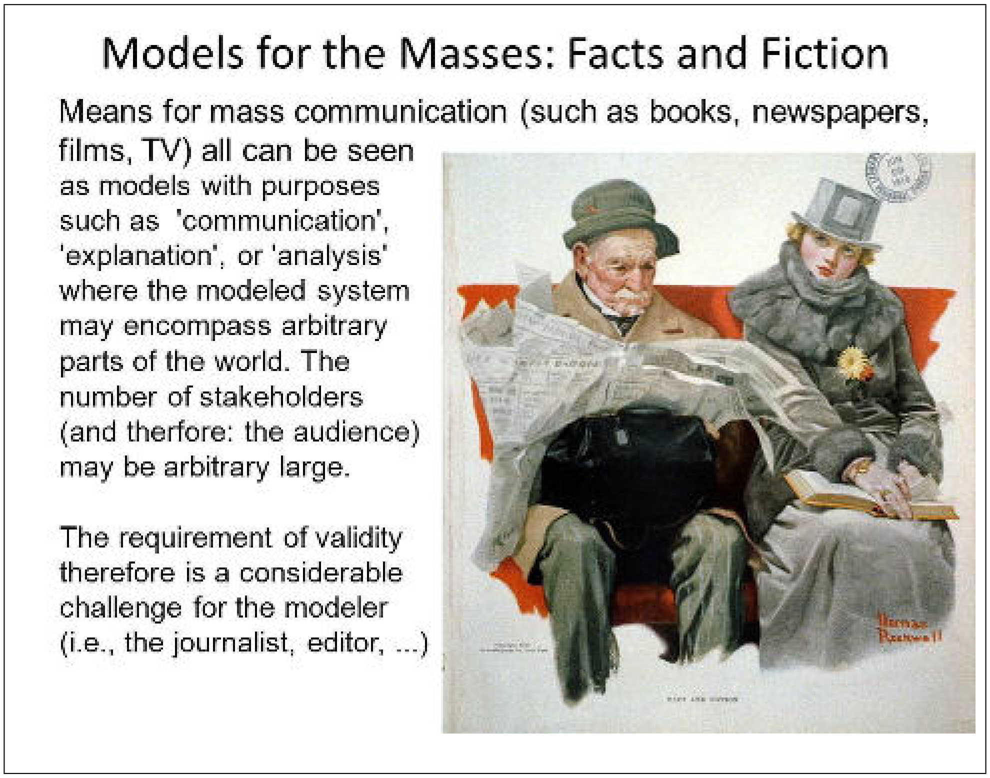
' AUDIENCE' here is an abbreviation of 'number of intended stakeholders'.

This criterion can be compared to 'Specialization' in the second interpretation. The size of a audience, possessing a certain expertise decreases when the required expertise level increases. So when a model should serve a large size audience, it can only do so if it assumes a low level of specialization.

[9]The image of the Delphi Oracle wastaken from http://upload.wikimedia.org/wikipedia/commons/3/3f/The_Oracle_of_Delphi_Entranced.jpg?uselang=nl

The same recommendations therefore apply to 'Specialization' and to 'Audience'.

There is, however, one additional aspect.



When a model serves the purpose of few stakeholders, the presentation and interpretation can be fine-tuned to meet the needs of precisely these stakeholders, to the extent where a *vis-a-vis* discussion can help clarify the model outcomes[10].

For a large audience, this is no longer possible, and alternative communication channels are needed.

An example is a recent innovation in Dutch democracy called 'de stemwijzer' (the vote advisor). In the weeks prior to an election of the Dutch national Parliament, a website is opened, where visitors can fill in their multiple choice reactions to a number of statements, representing issues of political controversy between the parties. The vote advisor then suggests which party to vote. The model, in essence, is a black box model that operates on the data, entered by the user.

The vote advisor, serving hundreds of thousands of users to a reasonable level of satisfaction, scores high on the criterion 'audience'. The same is true for an increasing number of Internet sites, including route planners, sites for comparing consumer products, and weather prediction applications. For a model to score high on 'audience', special attention has to be paid to robustness, ease-of-use, and performance.

7.2.5 Conclusion Stage, Inside, Qualitative: Convincingness

In the taxi model, detailed in Appendix **??** , there are two cat.-II quantities: `profit` and `fun`. The first one is defined as `profit=income-expenses`, the second one as `fun=ffMe*ffmC*ffmP*10`. The three quantities `ffMe`, `ffmC`, and `ffmP` mean, respectively 'fun for me', 'fun for my clients', and 'fun for my personnel'. All three are numbers between 0 and 1, and their product means that the total fun is 0 if any of the three contributing terms is 0.

There is a large difference in CONVINCINGNESS between the two quantities `profit` and `fun`.

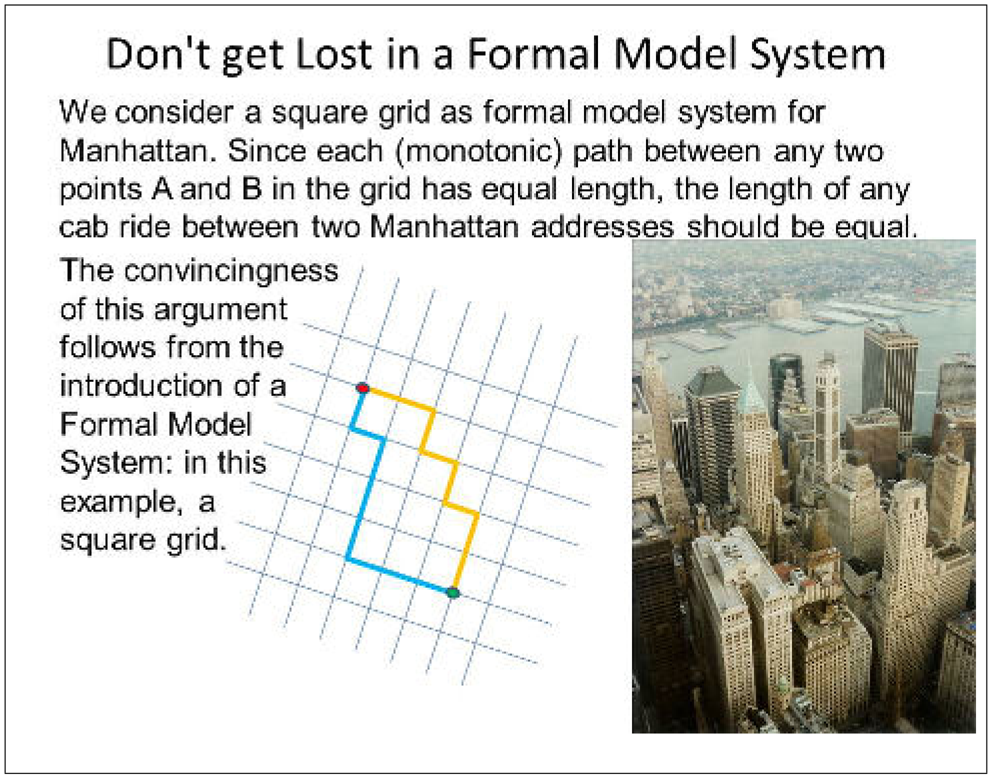
Imagine two different models, one for each cat.-II quantities, say the 'taxi-profit' model and the

[10]The 1917 Normal Rockwell painting 'Fact or Fiction' was taken from http://commons.wikimedia.org/wiki/Category:Newspapers_in_art#mediaviewer/File:Fact_%26_Fiction_by_Norman_Rockwell_1917.jpg

'taxi-fun' model. The taxi-profit model contains assumptions, but most of them are not far-fetched. They are, for instance, that tips can be ignored in the calculation of the company's income, and that the variation of a car's fuel consumption per kilometer is relatively small. Most people, including experienced workers in the field of taxi companies, will find such assumptions plausible.

The choice to model *fun* as a product of three contributing factors, however, is not obvious at all. Even if there is no immediate reason why this should be wrong: the step `fun=ffMe*ffmC*ffmP*10` is not evident. The same holds for other steps in the taxi-fun model. The taxi-fun model is less convincing than the taxi-profit model.

Convincingness hinges on plausibility of assumptions. This can be related to a ordinal scale.



Highest plausibility: are assumptions logically deducible from other assumptions that are not problematic? The argument takes the form: 'since earlier we assumed P, and since Q is implied by P, we can now assume Q'. If P is not problematic, and the implication is evident, Q is not problematic either.

Second best: is there a *law* or *theory* in a well-accepted discipline (physics, economy, ...), such that the current assumption is an instance of that law or theory? For instance: 'since this machine part can be considered to be rigid, it acts as a lever, and therefore the torque on the left part equals the torque on the right part'. The weakest part is: '... can be considered rigid', but once the problem owner accepts this step, the conclusion is accepted by anyone who accepts the physical laws governing torques in levers.

Third best: is there a plausible FORMAL MODEL SYSTEM (FMS, for short)[11] to which the current system may be compared?

As an example, consider the tidal waves in the world's oceans. These result from interplay between gravity and the inertia and viscosity of all the world's open water. A FIRST PRINCIPLES MODEL would take the geometry of all the world's coast lines and sea floors into account. It would bring in partial differential equations to describe how water moves within these boundaries. Tidal motion,

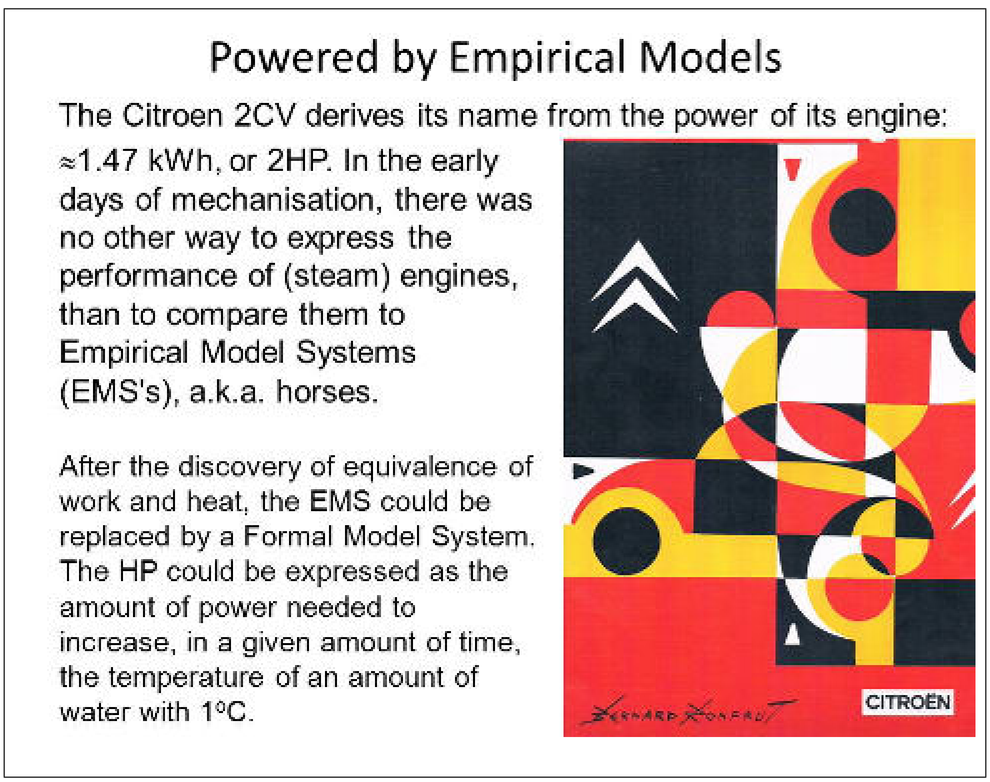
[11] The Manhattan photograph was taken from http://commons.wikimedia.org/wiki/Manhattan#mediaviewer/File:Financial_District_Manhattan.jpg

however, can also be described by an FMS: Laplace's 1776 tidal model, containing just a handful of quantities to represent all of the involved water, plus an external force to account for gravity.

It does the following. It takes the world as a perfect sphere; the water is considered incompressible; the geometry of the sea-floors is left out. Next, this FMS is treated using first principles. We use physical laws that hold for a uniform sheet of an ideal incompressible fluid on a perfect sphere. These laws may not at all be applicable to the original system. The extent to which the FMS and the original system correspond, however, is not discussed. There is no attempt to map all internal (cat.-IV) quantities of the FMS to quantities in the original system. Instead, the confidence in the FMS-approach comes from the facts that (1) the FMS is physically consistent, and (2) the *predictions* from the FMS can be empirically verified. So, rather than verifying the *construction* of the FMS, which may be impossible if the original modeled system is too complicated, we base its plausibility on the correspondence between the *outcomes* of the FMS and the *empirical data* obtained from the *original system*.

As long as

- the numerical correspondence is good enough for the purpose, and
- the application of the first-principles laws to the FMS is appropriate, and
- there is no a priori inconsistency between the original system and the FMS,



non-logically justifiable assumptions are often accepted. Many examples of lumping and 'emergent behavior', as discussed in Sections 1.3.6 and 1.4.3 use formalizations of physical behavior that cannot be logically justified by first principles; they also use this notion of FMS instead.

Fourth best: We may seek support for an assumption in our model from an EMPIRICAL MODEL SYSTEM (EMS for short), together with a similarity argument[12]. This approach is sometimes followed for soft quantities, such as 'fun', 'comfort' etc. ▷131 .

We illustrate the working of an EMS this for the quantity `fun` in the taxi model. Suppose we have empirical data from a number of existing taxi companies, including interview

[12]The Bernard Ronfaut 2CV poster was taken from http://commons.wikimedia.org/wiki/File:Affiche_2CV.jpg?uselang=nl. Notice that, in the example of calculating power, the replacement of EMS by FMS has the advantage of standardization: horses have different powers, whereas water has always the same specific heat.

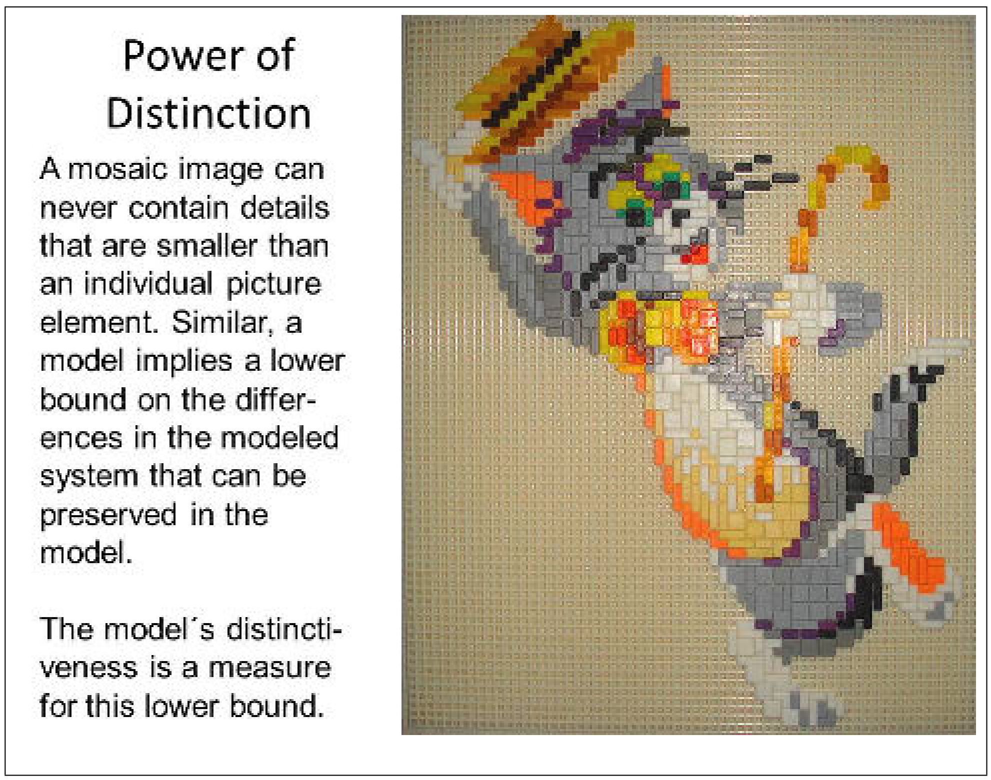
data of company owners, personnel and customers regarding their perceived 'fun'. The scores on interview forms are the definition of 'fun': the amount of perceived fun is the value that interviewees fill in when asked how much fun they have. Next we do a statistical analysis on these fun scores in relation to other factors, say $x_i, i = 1 \cdots n$, pertaining to all of the taxi companies involved. For the x_i we take cat.-I quantities, similar to those in Appendix **??**. If the correlations (see Section 6.3) are sufficiently strong, we use these to construct an EMS: a black-box model that predicts, for a given vector x_i, what the scores for fun *would* be if we would have a taxi company, corresponding to vector $x = (x_i, i = 1 \cdots n)$. We have thus made an *empirical* model to predict 'fun' that is more convincing than the dubious expression `fun=ffMe*ffmC*ffmP*10`.

Another example of an EMS is the SCALE MODEL. A scale model is, e.g., a small version of an aircraft put in a wind tunnel with the purpose to study the aerodynamical behavior of the original. The proverbial guinea pig is also an empirical model system.

Fifth best (=weakest): If we can not empirically justify our formulas, and we cannot derive it from a suitable FMS, there is little support for a formula such as `fun=ffMe*ffmC*ffmP*10`.

There is only one, weak, argument for its plausibility: the choice for the product of three fun factors is consistent with the intuition that 'fun' should increase when each of three factors increases. But there are numerous other formalizations that achieve the same.

The above 5 steps give a scale for the convincingness of a model outcome.



7.2.6 Conclusion Stage, Inside, Quantitative: Distinctiveness

Most purposes as listed in Table 1.1 have to do with 'distinction'. The criterion should be called 'the power of distinction'; for brevity we use the term 'distinctiveness'[13].

explanation: often an model for explanation says that phenomenon X is caused by mechanism Y. The DISTINCTIVENESS, or *distinctive power*, relates to the strength of the argument. An argument is more distinctive in explaining that X and not X' is caused by Y, when X and X'

[13]The photograph of the mosaic cartoon cat is taken from http://commons.wikimedia.org/wiki/Category:Ministeck#mediaviewer/File:Ministeck_Tom_%26_Jerry.jpg

are more similar. Also, the model is more distinctive in explaining that X is caused by Y and not by Y', when Y and Y' are more similar;

prediction 1(*): a prediction about when something will happen (=at time T, and not at time T') has larger distinctive power if T and T' are closer together;

prediction 2: a prediction about what is going to happen under certain circumstances (namely, X and not X') has larger distinctive power if X and X' are more similar;

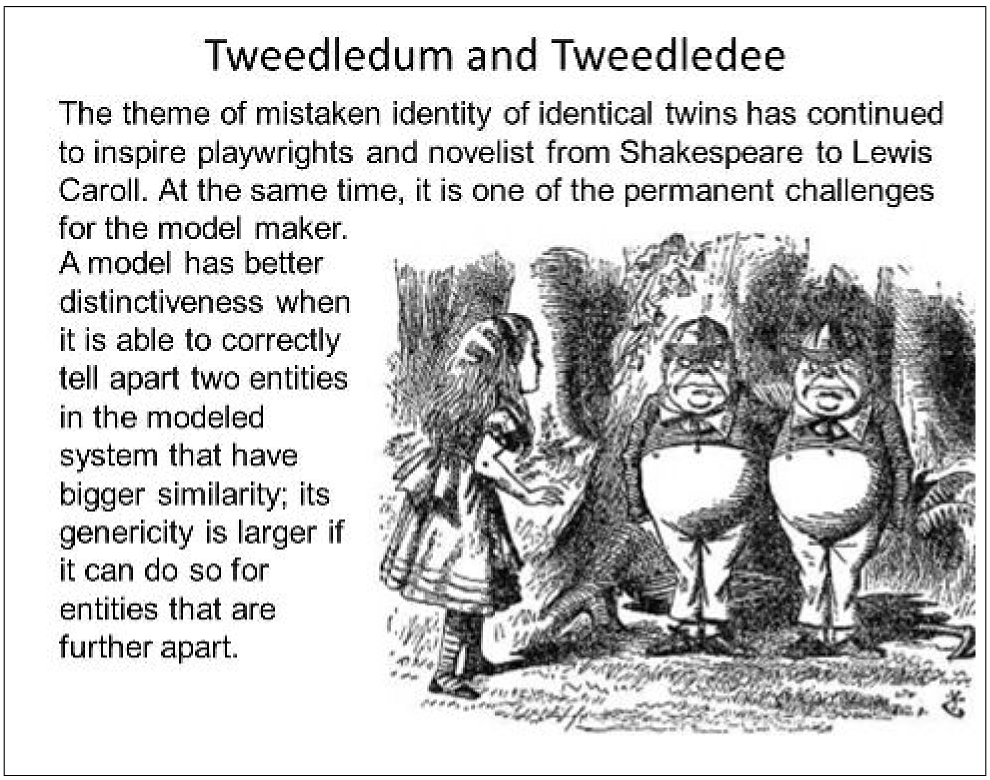
compression (*): let data set A be compressed to c_A. The compression model is more distinctive if the difference between A and A', such that c_A and c'_A are just distinguishable, can be smaller ▷132;

inspiration: a model, used for inspiration (for instance, a taxonomy or an ontology such as a MORPHOLOGICAL BOX) is more distinctive if it is capable to represent two items A and A' as different entries, where A and A' are more similar;

communication: a model, used for communication, has bigger distinctive power if, for two messages M and M', the receiver perceives the received versions r_M and r'_M as just being different, for M and M' being more similar;

unification: the quality of a model for 'unification' does not relate to distinctiveness;

abstraction: the quality of a model for 'abstraction' does not relate to distinctiveness;



analysis: A model[14] for analyzing phenomenon P produces a result a_P. The model is more distinctive if, for a_P and a'_P being just distinguishable, P and P' can be more similar;

verification: a model for verification has greater distinctive power if, for only one of X and X' being true, and the model successfully identifying the correct one, X can be more similar to X';

exploration: the quality of 'exploration' does not relate to distinctiveness;

decision: a model for decision making has larger distinctiveness if, for two alternatives A and A' where it suggests the right alternative, A and A' can be more similar;

optimization (*): a model, used for finding an optimal X is more distinctive when it rightly prefers X over X' whereas X and X' are more similar;

[14] The Jonh Tenniel drawing from the twin brothers Tweedledum and Tweedledee was taken from http://upload.wikimedia.org/wikipedia/commons/d/d5/Tennieldumdee.jpg?uselang=nl

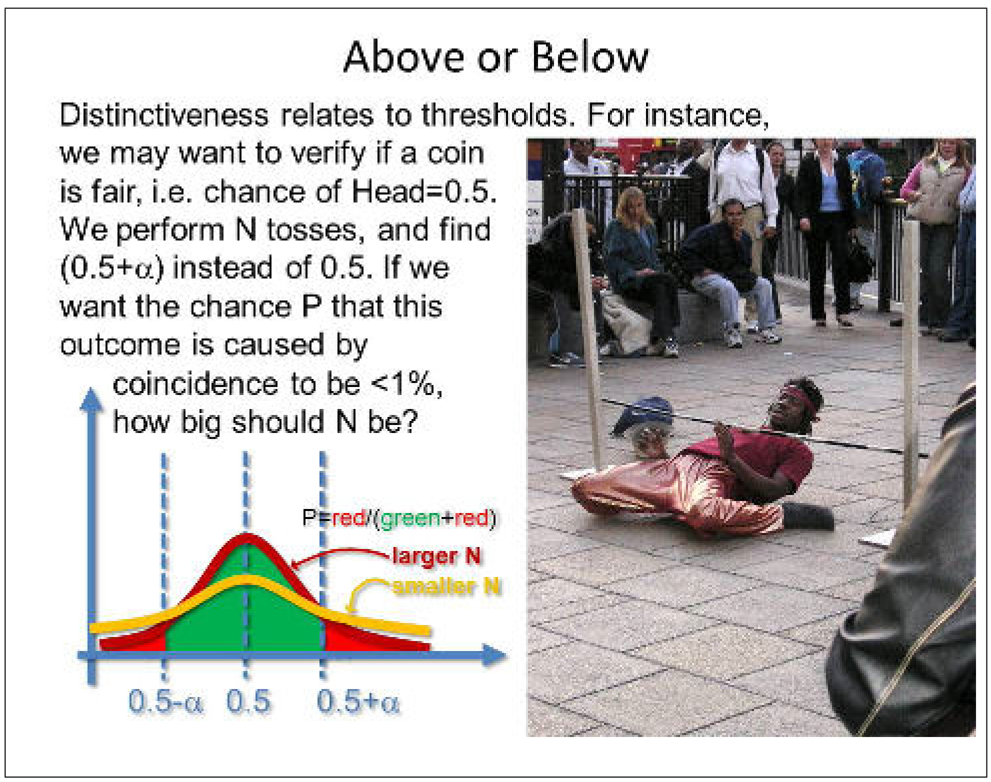
specification: a model for specifying S (e.g., the formal definition of the behavior of a machine) has greater distinctiveness when it can specify both S and S', preserving their differences, where S and S' are more similar;

realization: a model for realizing R (e.g., a blueprint for a built object) has greater distinctiveness when it can cause the realization of either R or R' at will, where R and R' are more similar;

steering and control: a model for steering and control has greater distinctive power if, for two targets T and T' it can ensure the right outcome in each case when T and T' are more similar.

For the purposes labeled with (*), the distinctiveness of the model relates to accuracy. A prediction is more accurate if the predicted number p has more significant digits, that is, if the interval of possible values for p that are not distinguishable by the model, is narrower.

All occurrences of 'distinctiveness' in the above list rely on *similarity* and *distance*, for numbers, vectors, non-numerical objects. For this reason we elaborate 'similarity'and 'distance' in Appendix **??**.



Distinctiveness[15] always relates to *thresholds*. For instance, for a decision, where the model has to favor one of two alternatives. There is a quantity, say p, computed by the model, and a value, say p_0 such that $p < p_0$ favors alternative 1, and $p > p_0$ favors alternative 2. Since the model contains approximations and assumptions, the calculated value of p is an uncertainty distribution, cf. Section 6.1.1. Assessing whether $p < p_0$ or $p > p_0$ amounts to assessing whether the bulk of the distribution of p lies on one or the other side of p_0.

To see what this means, we recall the example of the storm warning from Section 1.1. Suppose that p is the calculated chance for a storm of some force, and p_0 is the threshold such that for $p > p_0$ the pop festival must be canceled. If the entire distribution of p is above p_0 there will be cancelation; if the entire distribution is below p_0 the festival will go on.

In all other cases, the outcome is not unequivocal. The festival management, however, can only make a binary decision: either the festival is canceled, or it is not canceled.

The consequence of erroneously canceling the pop festival are bad (financial losses, loss of prestige,

[15] The photograph of a limbo dancer is taken from http://commons.wikimedia.org/wiki/Category:Limbo_(dance)#mediaviewer/File:Limbo.jpg

frustrated audience, law suits). The consequences of erroneously *not* canceling the pop festival are bad too, but different (casualties or injuries, material damage).

This scenario occurs often. Whenever there is the option of doing something on the basis of a model-based analysis, we can make two types of errors.

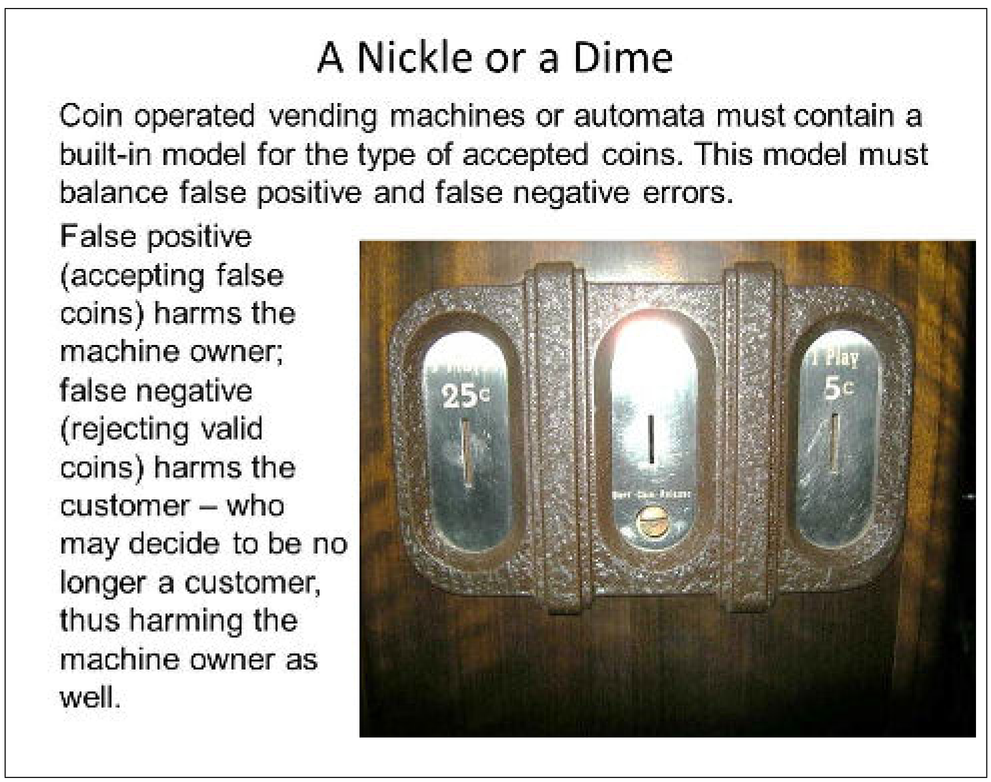
false positive : the situation where we take an action that should not have been taken;

false negative : the situation where we should take action, but omit to do so.

Errors in decision models with a single threshold fall in one of these two categories. The impact of false positive and false negative errors can differ significantly. For the pop festival it is the difference between financial loss and casualties; for the decision to buy the 25 Euro book (see Section 1.1), it is the difference between not owning the book or a sober meal.

False positive / false negative errors[16] often occur in medical context. A patient has some symptoms, and a diagnosis is a statement of the patient's condition based on these symptoms. This conclusion can be wrong in two ways. A pathological condition is not recognized, so no action will be taken where an action should be taken (false negative), or a pathological condition is assumed where no such condition exists (false positive).

Again the consequences are very different: false positive causes anxiety in the patient, costs for the unnecessary treatment and perhaps harmful effects of this treatment; false negative may cause the continuation or even deterioration of the pathological condition.



The difference in consequences of a false positive and a false negative may be accounted for by taking the break-even point not halfway the distribution of p. The type of error that has the largest adverted consequence should have the smaller chance. This can be realized by offsetting the value of p_0, or by asymmetric stretching the uncertainty distribution for p.

To *quantify* distinctiveness, we look into the distribution of uncertainty in the output quantities in models. The narrower these distributions, the larger the distinctiveness.

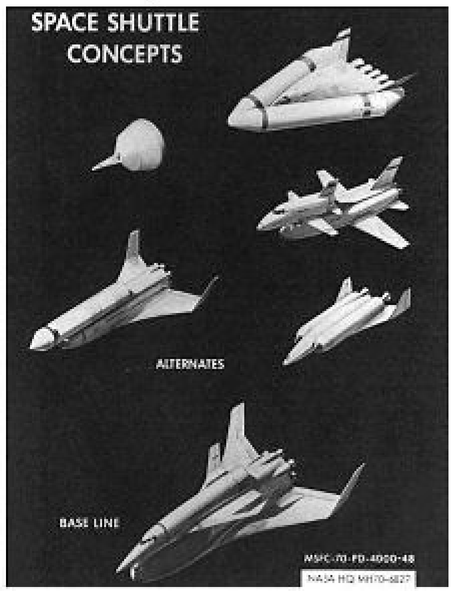
[16]The photograph of an old coin-operated juke box is taken from http://commons.wikimedia.org/wiki/File:Mills_Studio_jukebox_coin_slots.jpg?uselang=nl

7.2.7 Conclusion Stage, Outside, Qualitative: Surprise

Outcomes of many models are boring. Often it is a priori clear that a model will produce a value, say, between 0 and 1 (the chance of something happening), or between 0:00 and 23:59 (the time of the day of something happening), or any other ranges with a priori known boundaries.

Expecting the Unexpected

In radical design, the largest possible variety of ideas should be obtained. One way to achieve this is to form an ***ontology*** of the design space. That is: identify a collection of cat.-I quantities and their values such that any combination yields a potential design solution. In the 1960-s, prototype space shuttle concepts were generated in this way 'by hand'; present day ontologies, containing perhaps millions of concepts, are implemented in software to be executed on computers.



There are, however, model purposes that ask for models with SURPRISING outcomes. We give some examples[17].

- A model using principal components analysis (abbreviated by PCA) works as follows. There is a number of vectors, every vector containing a number of factors that represent 'inputs' and a number of factors that represent 'outputs'. For agriculture, 'inputs' could be amounts of water and sunlight, insecticides, nutrients, minerals etc.; 'outputs' are biomass and various qualities of a plant. There is no glass box model describing the outputs as functions of the inputs.

By means of PCA we may assess which combination(s) of input quantities account(s) for the largest part of the variance in the output. The number of possible combinations is often so big, that we cannot anticipate all possible outcomes. Moreover, the combination of the most crucial factors may shed light on hitherto unseen mechanisms in the studied system. The interpretation of the dominant principle components in terms of such a mechanism therefore is often the most interesting, and indeed *surprising* result of the model. Principal component analysis is increasingly used for instance in life sciences and social sciences.

- Ontologies describe items in a domain in terms of a number of attributes. Swapping values for these attributes can produce surprising new items. Ontological models therefore are used in exploration (e.g., a morphological box) and inspiration.

Principal component analysis is a black box modeling strategy; ontological models classify as glass box models. Both are potentially surprising because their output, with n different input terms (in PCA: the input factors; in ontology models: the attributes), the number of possible outputs is $O(2^n)$. This grows rapidly with n, and soon escapes human capabilities of anticipation.

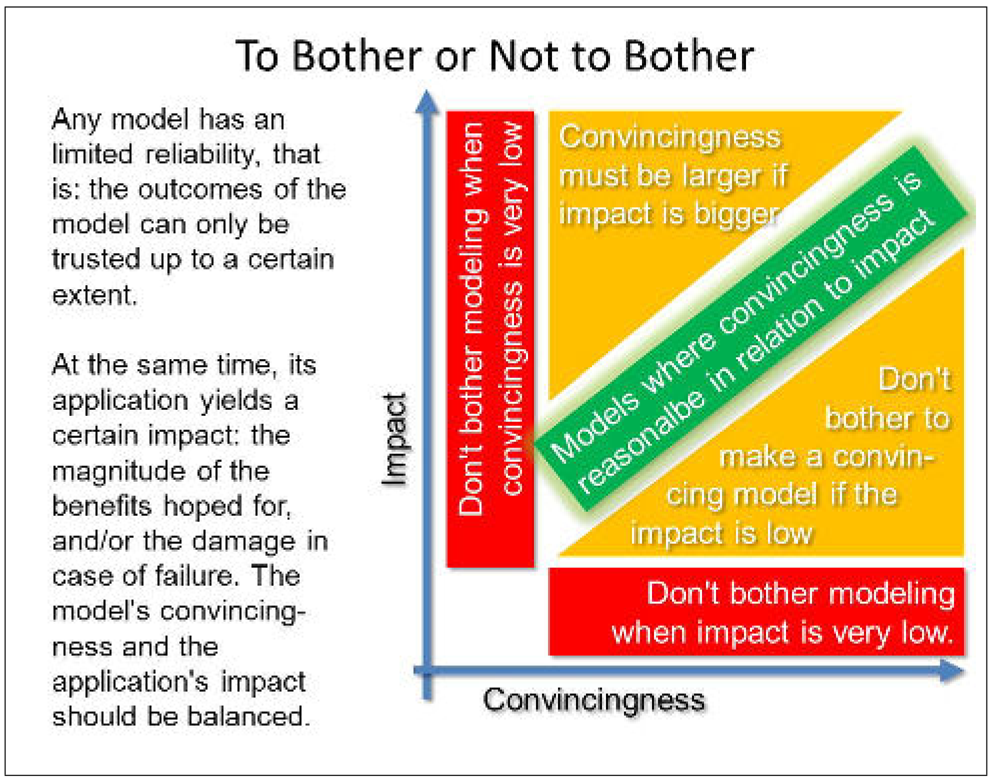
[17] The design sketches for the space shuttle weretaken from http://commons.wikimedia.org/wiki/File:Space_Shuttle_concepts.jpg?uselang=nl

Surprise can also come from a different type of purpose:

- In abstraction and unification, we build a formal model that captures the essence of some given system. In Section 1.2.1 we presented gravity as a model that unites the motion of heavily bodies and things on the earth. When gravity was formalized in the form of a simple formula involving only masses, distances and one scalar constant, it spawned deep questions: what would happen if the value for the exponent of the distance would not be 2? What can we say about the interaction between 3 bodies? What can we say about the dynamics of planet - or galaxy formation? The potential of surprise was -and still is- large.

7.2.8 Conclusion Stage, Outside, Quantitative: Impact

The weather model and the wallet model from Section 1.1 both contain a form of prediction The wallet model was the most certain, the weather model the most uncertain. We can look at the IMPACT of both models. The wallet model has foreseeable impact: the consequences of failure don't exceed one missed evening meal. The impact of the weather model are much larger: when it fails in one of two directions the impact can range from hundreds of thousands of euro's to injured or even killed festival goers.



Models can range from trivially reliable to far off. The consequences of *applying their results* can range from futile to immense. This suggests the diagram 'To Bother or Not to Bother'. We could make a similar diagram where impact is balanced against distinctiveness.

To express impact as a number, we use four quantities:

r_1 the profit or income in the present situation, without the model outcome; r_2 the profit or income with the model outcome in place; c_1 the cost of ownership in the current situation; c_2 the cost of ownership with the model outcome in place.

For all quantities, the same time scale is taken (e.g., lifetime or yearly amounts). The dimensionless quantity ρ,

$$\rho = \frac{((r_2 - r_1) - (c_2 - c_1))}{(|r_2 - r_1| + |c_2 - c_1|)} \tag{7.1}$$

is a number between -1 and 1. Positive values mean a beneficial contribution, negative values mean that the impact is adversely. The absolute value, $|\rho|$ indicates the size of the impact.

A model cannot always prevent that $\rho < 0$. For instance, a model to predict the financial consequences of sea level rising will find negative values for ρ.

'Impact' is also determined by the chance and consequence of an incident. Then the magnitude of the impact is proportional to the amount of damage (claims, cost for rebuilding, perhaps insurance premium) and to the chance for an incident.

The diagram 'To Bother or Not to Bother' can be summarized as: balance the impact of the model's outcome against its convincingness.

For two models, both meeting with their expectations, and everything else being equal, the one with larger impact is a better model, if is is the modeler's purpose to sell or publish the model, or to derive prestige from it. A successful model with a larger impact adds more to a modelers palmares.

If, however, the modeler's purpose is to play safe, the model with the lesser impact is the better one. The lesser the impact, the smaller the consequences of failure.

7.3 Criteria for Modeling and Purposes

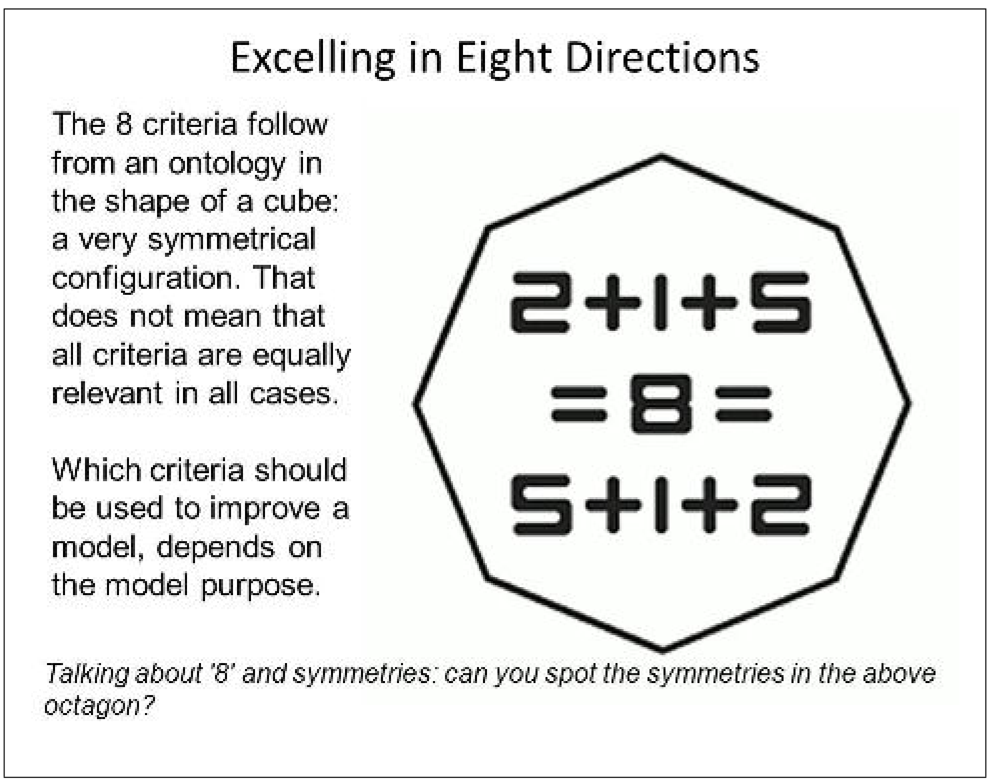


Table 7.2 gives, for each purposes from Table 1.1, the relevance of each of the criteria from Table 7.1. In the table, '-' means that a criterion usually has little relevance for a modeling purpose; 'XX' means much relevance, and 'X' is somewhere in between[18].

Purposes 'explanation', 'communication', and 'analysis' don't make sharp distinction between one or the other criterion. This is because all three apply to many different modeling techniques. It says very little that a model should be used for explaining, communicating or analyzing, as virtually any model can do that. Explanation, communication, or analysis are

[18] The image from the ambigram featuring 'symmetries of 8' was taken from http://commons.wikimedia.org/wiki/File:Ambigram-8-eight-math-2-1-5-rotation-mirror-basile-morin.gif?uselang=nl

rarely the *only* purpose of a model. The modeler, then, should seek a purpose 'behind the purpose', e.g. by asking 'to what aim should the analysis be done', or 'what should be the effect of the explanation'. Often an underlying purpose exists, such as deciding, predicting or optimizing.

Once the *real* purpose is clear, we have the following list of most relevant criteria per purpose:

- *Prediction (1 and 2):*

Convincingness (predictions should be plausible: they should not contain too wild assumptions),

Distinctiveness (predictions should be accurate),

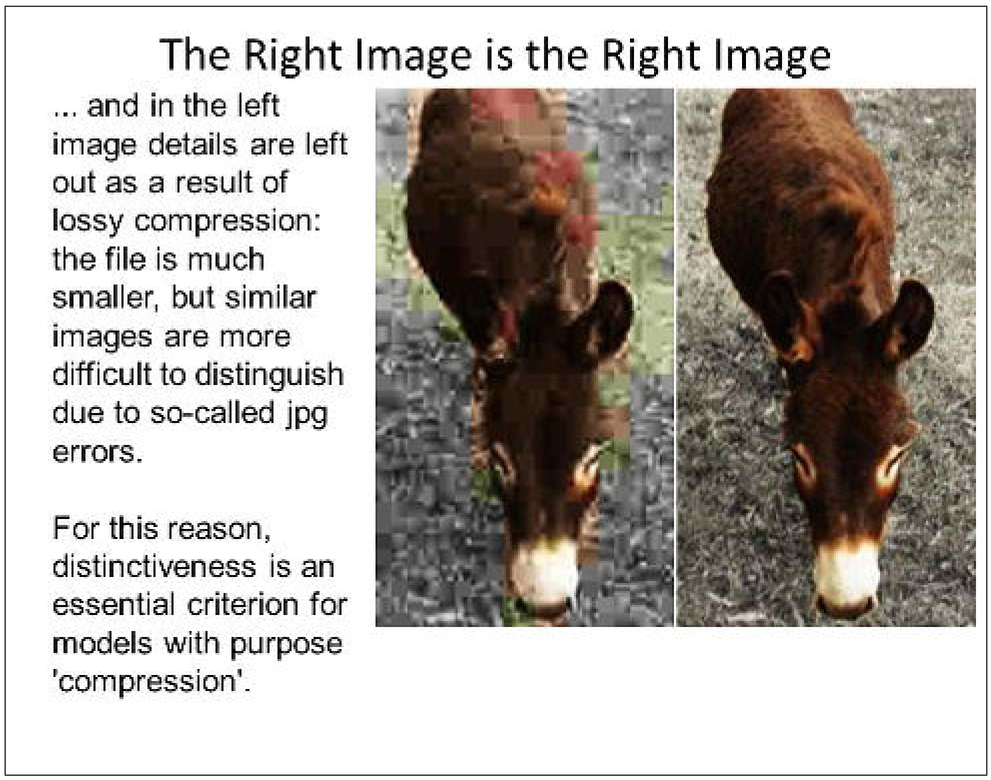
Impact (predictions with little impact don't merit the effort of running a model);

- *Compression:* [19]

Scalability (input data should be allowed to be voluminous - small input size doesn't merit compression),

Audience (public domain applications heavily rely on compression - e.g. images, video and sound; this asks for robust approaches),

Distinctiveness (the more distinctive a compression model, the more information of the input data is preserved);



- *Inspiration:*

Surprise (an inspirational model should be able to give new ideas - irrespective if these are guaranteed correct (convincingness) or sharply different (distinctiveness));

- *Unification:*

Genericity (unification is more challenging, and more powerful if it can relate to more varied areas),

Convincingness (to unify different areas needs assumptions from different domains. These are not necessarily logically obvious),

Surprise (if the unification brings no new and unexpected insights, it is of little use);

- *Abstraction:*

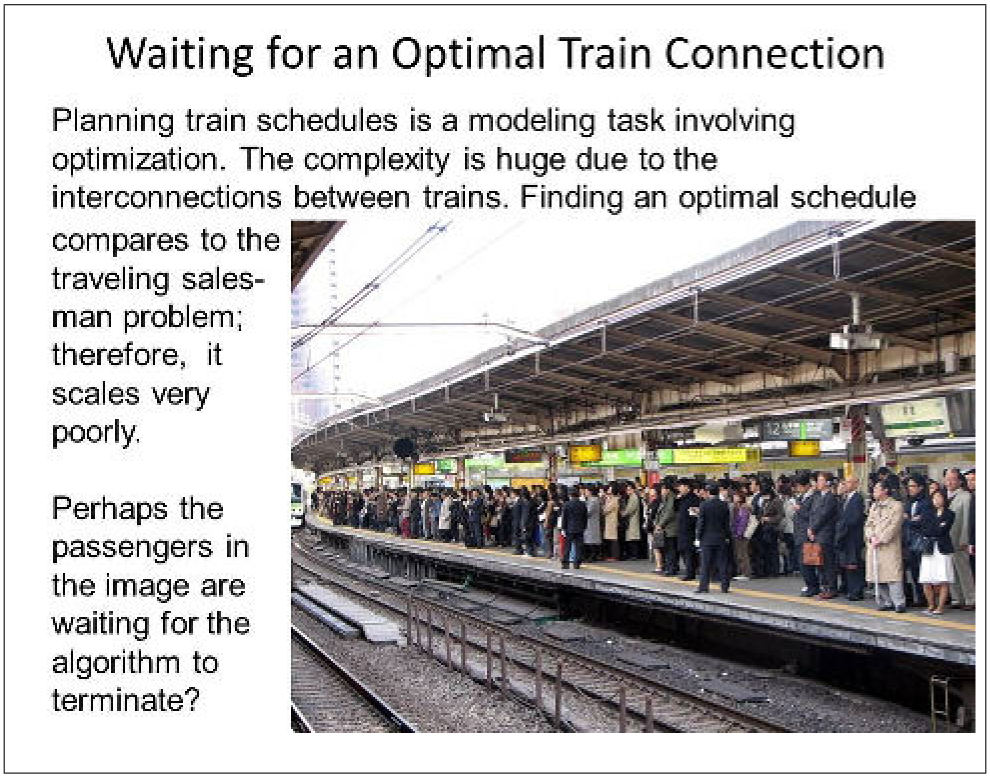
Convincingness and

Surprise (both similar to Unification; unlike Unification, however, genericity is no crucial criterion

[19] The two images demonstrating compression artefacts are taken from http://commons.wikimedia.org/wiki/File:JPEG_example_donkey_010.jpg?uselang=nl and http://commons.wikimedia.org/wiki/File:JPEG_example_donkey_100.jpg?uselang=nl

as abstraction applies to a single area). On the other hand,

Distinctiveness is relevant to safeguard against models that loose the ability to distinguish one thing from another;



- *Verification:*

Scalability (small scale problems don't require sophisticated models; they can be verified 'by hand'. The challenge for verification models is to cope with, e.g., large state spaces),

Convincingness (a model used for verification needs sufficiently convincing arguments. Indeed, the propositions that are to be verified are usually not one-to-one mapped onto formal quantities),

Impact (formal verification, being cumbersome and resource-intensive, is reserved for vital application areas);

- *Exploration:*

Genericity (the added value for models for exploration, comes from their ability to cross borders between domains),

Surprise (the main reason for exploring a domain is the hope to find something new and unexpected);

- *Decision:*

Convincingness (as with verification, the outcome of a decision is rarely a logical argument. The role of non-logically necessary assumptions is crucial),

Distinctiveness (see the argumentation in Section 7.2.6, elaborating the case for decisions),

Impact (same argument as for verification);

- *Optimization:* [20]

Genericity (optimization algorithms occur as 'stand alone' functional modules, like a separate branch of technology. Developing such technology is more cost effective if deployed for a larger variety of domains),

Scalability (same argument as Verification),

Convincingness (same argument as Decision),

[20] The photograph of waiting passengers on Shinjuku station is taken from `http://commons.wikimedia.org/wiki/File:Rush_hour_at_Shinjuku_04.JPG?uselang=nl`

Impact (same argument as Decision);

- *Specification:*

Genericity (formal specification languages appear, entirely devoted to specification of systems. To developing such languages, the same observation applies as in Optimization),

Distinctiveness (a specification model needs to be sufficiently precise to allow for the expression of small differences between the specified systems or machines);

- *Realization:* [21]

Genericity (same as with Specification),

Distinctiveness (same as with Specification),

Impact (if the realization of an artifact can be done purely by automated execution of a model, such artifacts can proliferate at high rates (e.g., computer viruses, modified DNA). For an other argument as to the relevance of the impact criterion for realization: see the poster 'we can't win with blueprints'.);

Impact of Blueprints

Blueprints are detailed specifications of material artefacts – for instance weapons. This WW-II poster tries to make the point that the impact of blueprints, provided their purpose (realization) can be achieved fast, can be the difference between wining or losing a battle.

- *Steering and Control:*

Distinctiveness (the reason to steer and control a system is, to achieve small distance between the value of some quantity as it should be, and the value of the same quantity as it actually is).

By using Table 7.2 and the clarification in the above list, we can interpret model outcomes as follows:

1 from the problem definition, analyse which purpose the model should fulfill (stage 1 from Section 1.4);

2 from this purpose, and from Table 7.2, identify which criteria are most central to the current problem setting;

3 develop the model (stages 2,3,4 from Section 1.4);

4 using the selected criteria, assess if the model outcome fulfills the purpose. If not, the criterion for which the model outcome is insufficient may hint at a direction for improvement;

[21] The WW-II poster was taken from http://commons.wikimedia.org/wiki/Category:Blueprints#mediaviewer/File:%22WE_CAN_WIN_WITH_BLUE_PRINTS%22_-_NARA_-_516071.jpg

purpose	Genericity	Scalability	Special-ization	Audience	Convincing-ness	Distinctive-ness	Surprise	Impact
Explanation	X		X	X	X	X	X	X
Prediction 1	-	X	-	X	XX	XX	-	XX
Prediction 2	-	X	-	X	XX	XX	X	XX
Compres- sion	X	XX	-	XX	-	XX	-	X
Inspiration	-	-	X	-	-	-	XX	-
Communi-cation	XX	-	XX	XX	XX	XX	X	XX
Unification	XX	-	-	-	XX	X	XX	-
Abstraction	-	-	-	-	XX	XX	XX	-
Analysis	XX	XX	XX	-	XX	XX	X	X
Verification	X	XX	-	-	XX	XX	-	XX
Exploration	XX	-	X	-	-	X	XX	-
Decision	X	X	X	-	XX	XX	-	XX
Optimi- zation	XX	XX	-	-	XX	X	-	XX
Specifi- cation	XX	X	-	-	-	XX	-	X
Realization	XX	X	-	-	-	XX	-	XX
Steering and Control	-	X	X	X	X	XX	-	X

Table 7.2: Category-II quantities for the modeling process

5 interpret the model outcome and formulate a solution of the initial problem (stage 5 from Section 1.4); pay special care to those aspects of the model that account for the selected criteria.

7.4 Mathematical preliminaries

materiaal aan te vullen door Emiel ...

...

...

...

7.5 Summary

- Leading question: to what extent has the initial problem been solved?
- Approach: define criteria to assess the quality of the modeling process *as a whole*
- A taxonomy to make criteria for modeling quality:
  - *Input* or *output* side?
  - *Inside* (=model, modeled system) or *outside* (=context, stakeholders)?
  - *Qualitative* or *quantitative*?
- Resulting criteria:
  - *Genericity*: how many different modeled systems can we handle?
  - *Scalability*: how large can the size of the problem be?
  - *Specialization*: how much should the intended audience know?

        * The importance of presentation;
        * Limiting the claims;
        * Biases;
        * Self-fulfilling and self-denying prophecies;
    - *Audience*: how large can the intended audience be?
    - *Convincingness*: how plausible are the assumptions?
    - *Distinctiveness*: e.g., how accurate, how certain, how decisive can the model outcome be?
        * The case of *false positive* and *false negative* outcomes;
    - *Surprise*: to what extent can the model outcome give new insight?
    - *Impact*: how big can the consequences of the model outcome be?

- Criteria for modeling quality are related to purposes; every purpose has its own dominant criteria.

7.6 Learning goals

7.6.1 Knowledge

You should know the meaning of the terms genericity, scalability, specialization, audience, convincingness, distinctiveness, surprise, impact, in the way they are used in this chapter. You should know the regimes of scalability, and examples of each regime; you should know 4 aspects of specialization (presentation, limiting claims, biased stakeholders, self-fulfilling and self-denying prophecies); you should know the 5 discussed levels of convingcingness with examples of each level; you should know the formula for impact. For each modeling purpose, you should know at least one criterion that is relevant to improve models for that purpose.

Emiel: aanvullen svp.

7.6.2 Skills

You should be able to compare two models on each of the 8 criteria. You should be able to assess, for a given model, its scalability regime. You should be able to choose an adequate data visualisation technique for presenting model outcomes. You should be able to indicate the level of convincingness for a given model. You should be able to estimate the impact fo a given model. In a concrete case, starting from a model of your own making, you should be able to analyse direction(s) for improvement, and make concrete recommendations.

Emiel: aanvullen svp.

7.6.3 Attitude

When given a model, a modeled system, and a model purpose, you should have the attitude to analyse if the purpose is adequately fulfilled; and if not, you should have the attitude to analyse (using the criteria discussed in this chapter) a direction for improvement; you should have the attitude to improve your model accordingly.

7.7 Questions

1. The 19th century prediction the big cities were to be covered by thick layers of horse manure was based on a large number of assumptions. Name at least three.

2. This chapter is mainly devoted to formulating criteria for modeling. What should these criteria do?

3. We say 'the problem of proposing [...] criteria [...] is that the variety of modeling approaches is immense'. Why is this a problem?

4. We say 'a property such as shape is not enumerable'. Why not?

5. Explain in your own words what 'building a taxonomy' means.

6. What does it mean that a taxonomy is complete?

7. We define 'input' and 'output' of modeling processes in terms of the stages of the modeling process of Section 1.4. Which stages?

8. A distinction between criteria is 'qualitative' and 'quantitative'. We stress that we don't refer to the type of the criterion.

   (a) Why not?

   (b) What *do* we refer to, then?

9. Explain, in your own words, the meaning of 'genericity'.

10. Why is a black box model of the form $y = ax^3 + bx + c$ more generic than $y = ax + b$?

11. Why is a triple store more generic than a table-based database? (If necessary, look up what a triple store is.)

12. A physical model for compressible fluid is more generic than a physical model for incompressible fluid. Why?

13. Give an example of an aspect of weather models that would count as being non-generic.

14. Explain, in your own words, the meaning of 'scalability'.

15. In your own words, explain what $O(n^2)$ means.

16. Why is $O(1)$ the same as $O(2)$?

17. A model with required processing time $O(n)$ is less scalable than a model with processing time $O(\log n)$. Why?

18. A jig-saw puzzle contains white pieces only; there are no two pieces with the same border: at each of the four sides of any piece, there is at most one fitting piece. The puzzle contains n pieces. What is the amount of time needed to assemble this puzzle?

19. Explain, in your own words, the meaning of 'scale free'.

20. In Chapter 3, we studied the state chart of a four color ballpoint.

(a) Find a property in the ballpoint's conceptual model, to be called n, such that the size of the state chart is polynomial in n, that is: the size of the model is $O(n^m)$ for some m.

(b) What is m in this case?

21. Explain, in your own words, the meaning of 'specialization'.
22. What do we mean by 'invest in presentation'?
23. What do we mean by 'when in doubt, limit the claims'?
24. What do we mean by 'be warned for biased users?'
25. What do we mean by a 'wicked problem'?
26. Explain, in your own words, the meaning of 'audience'.
27. Explain, in your own words, the meaning of 'convincingness'.
28. We give five levels of convincingness. Which?
29. Formulate the difference between 'third best' and 'fourth best' in the levels of convincingness in your own words.
30. In the description of the 'third best' level of convincingness, we state that there should be no 'a priori inconsistency'. Why the addition 'a priori'?
31. Explain, in your own words, the meaning of 'distinctiveness'.
32. We illustrate 'distinctiveness' for a number of different purposes. Some have an asterisk (*), some don't. Explain the asterisk.
33. Explain the meaning of the terms 'false positive' and 'false negative'.
34. Explain, in your own words, the meaning of 'surprise' as we use it.
35. Explain how surprise can result from working with ontologies.
36. Explain, in your own words, the meaning of 'impact'.
37. Explain Expression 7.1.
38. Why are there two ways to interpret 'impact'?

7.8 Exercises

1. This chapter is mainly devoted to formulating criteria for modeling. What constitutes 'good' criteria?
2. The combinations of purposes and modeling styles give a large variety of modeling contexts. There are factors that increase the diversity of modeling situations even further'. Name some.

3. We say 'a property such as shape is not enumerable'. The value of 'shape', regarded as an attribute, is a concept in its own right. Formulate at least 5 properties of concept 'shape' with attributes that *are* enumerable.

4. Considering complete taxonomies:

   (a) Choose a domain.

   (b) Give an example of a non-complete taxonomy for this domain.

   (c) Give an example of a complete taxonomy for this domain.

5. Regarding *genericity*:

   (a) Give a concrete example of a modeled system, a purpose and a model that you would call not good enough with respect to generic. (If the following two questions are too difficult, your example is not concrete enough. Choose a way to make it more concrete. The same advice holds for subsequent exercises.)

   (b) Suggest a modification that makes this model more generic.

   (c) Discuss if, with respect to any of the other criteria, the modified model has deteriorated.

6. Consider genericity for black box models.

   (a) Defend or dispute: 'a black box model is more generic when it has more freely adjustable quantities'.

   (b) Relate this issue to the notion of 'residual error', discussed in Section 6.3.3.

7. Regarding *scalability*:

   (a) Give a concrete example of a modeled system, a purpose and a model that you would call not good enough with respect to scalability.

   (b) Suggest a modification that makes this model more scalable.

   (c) Discuss if, with respect to any of the other criteria, the modified model has deteriorated.

8. Describe as precise as possible how a traveling salesman should find the shortest possible route visiting n addresses. Show that the complexity of your approach is $O(2^n)$.

9. Regarding *specialization*:

   (a) Give a concrete example of a modeled system, a purpose and a model that you would call not good enough with respect to specialization.

   (b) Suggest a modification that makes this model more suitable with respect to specialization.

   (c) Discuss if, with respect to any of the other criteria, the modified model has deteriorated.

10. Regarding the guideline: 'when in doubt, limit the claims':

    (a) Think of an example (either a real one - give a reference! - or an imaginary one) of a modeled system, a model and a purpose where the modeler should limit claims.

    (b) Think of a worst-case scenario of what could happen otherwise.

11. We explain the example of DDT as an example of self-fulfilling and self-denying predictions.
    (a) Where is the 'prediction' in the story?
    (b) Which of the two categories (self fulfilling - self denying) applies here?
    (c) Think of two examples of your own where a model is a self fulfilling and one were a model does a self-denying prediction.
12. Regarding *audience*:
    (a) Give a concrete example of a modeled system, a purpose and a model that you would call not good enough with respect to audience.
    (b) Suggest a modification that makes this model more suitable with respect to audience.
    (c) Discuss if, with respect to any of the other criteria, the modified model has deteriorated.
13. Regarding *convincingness*:
    (a) Give a concrete example of a modeled system, a purpose and a model that you would call not good enough with respect to convincingness.
    (b) Suggest a modification that makes this model more suitable with respect to convincingness.
    (c) Discuss if, with respect to any of the other criteria, the modified model has deteriorated.
14. We give the example of the Laplace tidal wave model as an FMS. Give another example.
15. We give five levels of convincingness.
    (a) For each of the five, give an example (different from the examples we provide in the text).
    (b) For each of the five, give an example where the assumption does not hold despite that it is 'convincing', and give an example of an modeled situation plus a purpose where this failing assumption is fatal, and one where it is not.
16. Regarding *distinctiveness*:
    (a) Give a concrete example of a modeled system, a purpose and a model that you would call not good enough with respect to distinctiveness.
    (b) Suggest a modification that makes this model more suitable with respect to distinctiveness.
    (c) Discuss if, with respect to any of the other criteria, the modified model has deteriorated.
17. Give an example of the application of the criterion 'distinctiveness' for the purpose 'explanation'.
18. Consider the diagram 'To Bother or Not to Bother'. There are quite a few lessons hidden in this figure. Formulate them in your own words.
19. In Chapter 1, we discuss a model for predicting if you can afford to buy a 25 Euro book. One of the input quantities was a 'measurement' of the present amount of money. This measurement could be inaccurate, e.g. if you ignore the copper coins. Analyse this inaccuracy in terms of false positives/ false negatives.

20. Think of another example, similar to the wallet-example, of a simple model used to make a prediction, and analyse the false positive / false negative scenario's.

21. Discuss the pro's and con's of lumping to define cat.-II quantities (cf. Section 5.3.1) in relation to false positives and false negatives.

22. Regarding *surprise*:

    (a) Give a concrete example of a modeled system, a purpose and a model that you would call not good enough with respect to surprise.

    (b) Suggest a modification that makes this model more suitable with respect to surprise.

    (c) Discuss if, with respect to any of the other criteria, the modified model has deteriorated.

23. Regarding *impact*:

    (a) Give a concrete example of a modeled system, a purpose and a model that you would call not good enough with respect to impact. Clarify which of the two views to 'impact', discussed in the text, you have chosen: favoring low impact or favoring high impact.

    (b) Suggest a modification that makes this model more suitable with respect to impact.

    (c) Discuss if, with respect to any of the other criteria, the modified model has deteriorated.

24. Give an example (imaginary or existing - in the latter case, give a reference) where Expression 7.1 can be used to estimate the impact of a model.

25. (*) Think of a scenario where the use of models in stock exchange markets could contribute to the occurrence of a crisis.

Notes, Index, and Glossary

Notes

[1]◁In a good approximation, planet's orbits are ellipses - at least, when viewed from a position that is fixed with respect to the sun. An ellips can be approximated by a circle, where the midpoint runs over another circle. In this respect, the Ptolemaic view was not unreasonable. When seen from the earth, however, planet's orbits are more complicated: they may contain loops and intersections. Ptolmy catered for these anomalies by having more and more advanced systems of circles running over circles, so called epicycloids

[2]◁Formulated as: 'A line joining a planet and the Sun sweeps out equal areas during equal intervals of time', This is the second of Keppler's three laws; the first stating the orbits are ellipses with the sun in a focal points; the third stating that 'The square of the orbital period of a planet is directly proportional to the cube of the semi-major axis of its orbit'.

[3]◁A consequence of the missing closed-form solution, however, is that it is not possible to prove that for instance the Solar System as we know it, is a stable configuration. So their is no mathematically rigorous proof that, say next year, the earth will still be in its familiar annual orbit. It is possible to run *numerical simulations*, though - and this of course has been done extensively. These show that there is no serious need for worry the first 100 million years or so - but the deeper question is: what is the amount of certainty that can be obtained from numerical simulations?

[4]◁It is interesting to think of what the criteria are for an answer to be acceptable in some given community.

[5]◁The distinction between a 'model' and a 'theory' is subtle. Most authors would call Newton's treatment of gravity a theory. It is usually understood that a theory has a wider scope of application than a model. So we speak of 'the theory of gravitation', from which 'a model for the Solar System' can be derived.

[6]◁Obviously, there is no end to possible 'why'-questions. One could (and physicist do) ask 'why does mass produce gravity?'. Interestingly, the communities who are satisfied by the answers to the increasingly fundamental questions rapidly get smaller.

[7]◁Communication is a process where a MESSAGE is conveyed from a sender to a receiver. When we view a user requirement document as a message, the sender is the prospect user, and the receiver is the designer.

[8]◁Exploration models may comprise of tables or catalogues. For instance, to aid the choice 'what sort of material should we choose for X', a catalogue of materials with their properties, offered by some vendor, or even the Periodic Table of Elements. Regarding this last example: it seems that the set of all chemical elements forms a closed set: we can enumerate all chemical elements. This may be true now, but it was certainly not the case when Mendelejev, the inventor of the Table, began his work. In fact, his exploration model helped to identify numerous as yet *undiscovered* chemical elements. So a model that is intended to aid exploration can serve to do predictions.

[9]◁This set of options can be COMPLETE or incomplete. For instance: if the decision is 'this rod should have some length X', or 'how long should this rod be', the set of possible outcomes is fully known: it will be some positive number of millimeters. In this case we are deciding with a complete set of options. An incomplete set could be associated with the question 'from what material do we make this rod': perhaps glass, carbon fibre or ceramic do not occur on our list simply because we have forgotten them, perhaps because we skipped the stage of doing a systematic exploration.

[10]◁Formal models contrast with INFORMAL MODELS. Until the 17th Century, mathematical notation was not standardized and little developed. Mathematical reasoning and non-mathematical reasoning, such as philosophical debates, roughly used the same type of language. Recent times see an increasing degree of rigor in mathematical vernacular, to the extent where computers verify mathematical argumentation. But even today there is room for THOUGHT EXPERIMENTS, most often formulated in natural language. Thought experiments, not expressed in any formal language such as mathematics, classify as informal models. In these lecture notes, our main focus will be on formal models.

[11]◁We don't take relativity theory into account here.

[12]◁We give an example where both continuous and discrete modeling takes place in one application. Consider a game of billiard. We want to develop a model to help a billiard player choosing the right cue position and angle to launch a bal such that the other two balls will be hit. We want to approach this problem by means of a dynamic SIMULATION - that is, we want to model the motions of the balls and the collisions as these take place in time. The set of collisions in one billiard shot form a discrete set of events. Each of them is determined by the initial state of the two colliding balls (that is, their position, velocity and rotational velocity). All these quantities are continuous, and the state after the collision follows from evaluating a mathematical function taking the initial states as input (see Appendix **??**). This yields a new set of values, and the balls will continue to roll on the billiard table after the collision. One problem, however, is to find out *where* a next collision will take place. There are two possible approaches here.

1. We ignore friction and rolling, and assume uniform motion:

   $$s(t) = s_0 + v \times t. \tag{7.2}$$

   Here, s and v are, respectively, a location and a vector in 2D space. Using formulas of this form for all three balls, we can compute if two balls, i and j will collide. This gives rise to the following condition:

   $$||s_{i,0} + v_i \times t - s_{j,0} + v_j \times t||^2 = (2\rho)^2, \tag{7.3}$$

   where ρ is the radius of a ball. The above formula simply asserts that the squared distance between the centers of two balls equals the squared sum of the radii - in other words: two balls touch. If this equation has a solution for some t, this is the time point of collision of the two balls i and j. Following this approach we work with continuous quantities only, and we don't need to sample anything.

2. The second approach is that we sample time t. That is, we replace the continuous quantity by a series of discrete values $t_1,\ t_2,\ t_3, \cdots$ where $t_{i+1} = t_i + \Delta t$ for some small Δt. In every new time step we recompute the location of all the three balls, and we check if there is a state where two balls are close enough to have touched. It turns out that this approach is in some respect simpler, in particular if we want to take friction and rolling into account. Sampling introduces a problem, however: since t only takes the discrete values, it is only possible to detect that two balls are close enough to collide if this occurs at one of the t_i. Suppose Δt was chosen too large: we then run the risk that we miss the event of a collision. On the other hand, choosing Δt very small is a sure waste of effort: we know beforehand that most of the tiny simulated displacements of the balls will show no collision.

The above dilemma is characteristic in virtual all modeling situations where sampling is used. When the sampling distance, sometimes called STEP SIZE, is too small we waste computational effort. If the step size is too large, however, it may introduce artifacts (such as missing crucial events like the collision in the above example).

For this reason, two modern trends in modeling are:

1. Use ADAPTIVE step sizes. Step sizes only need to be small in areas where the calculated quantity varies rapidly. In 'boring' regions stepsize can be big.

2. Using hybrid approaches, that is: mixing continuum calculations and sampling in one model.

[13]◁Wolfram's Mathematica, see http://www.wolfram.com, is a famous exception.

[14]◁The reader may wish to consider this table (again) after having read Section 3.3.4

[15]◁Intuitively, two items can be close together, or far apart. This can be expressed in a number: 0 means that two items coincide, and the larger the number, the further they are apart. If items have coordinates (x_1, y_1, z_1) and (x_2, y_2, z_2), an expression such as Pythagoras' theorem, $\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}$ follows this intuition. We could have chosen other expressions as well - each with their own properties. The current choice, for instance, has the property of being the same after arbitrary rotation of the configuration of the points in space.

[16]◁Numbers, operations, and sets are examples of mathematical objects. Mathematical objects are formal CONSTRUCTS. That is: they are ideas, invented by man, to help working with intuitive notions. Indeed, the mathematical objects we introduced here are all closely related to our intuitive feelings about space; they help to talk in a more precise way of what we experience when we perceive the space around us.

[17]◁Notice that that a deterministic model is not necessarily correct: if we ignore the copper coins when assessing the current contents of our wallet, we know that our finding is an underestimate - but we know this with absolute certainty. Further, inaccuracies may result from non-perfect measurements of quantities that occur in the modeled system.

[18]◁Increasing an ensemble size will not always reduce the magnitude of fluctuations. No matter how often we throw a fair die, the fluctuations will always stay between 1 and 6, with equal chance for each outcome. But repeating a measurement of each time the same quantity (say, repeatedly weighing a given object) will make the average deviation decrease proportional to $1/\sqrt{n}$ with n measurements. Fluctuations cannot get arbitrary small, though: for fundamental physical reasons (quantum mechanics) there are lower limits to the achievable accuracy in any measurement.

[19]◁The amount of molecules in one mole is roughly 6×10^{23}, Avogadro's number, and at 273K, 1 mole of gas under 1 atmosphere occupies about 22 liter.

[20]◁Emergent behavior not always results from lumping: sometimes the items in the system that together form the emergent behavior are not as similar as molecules in a gas. For instance: the fact that a riding bicycle, despite its two wheels, does not topple over is an *emerging* result of the interplay of many forces, torques, and momenta. But since the various forces and torques all apply to different kinds of parts of the bicycle, we would not call the bicycle's emergent stability a bulk-property of the bicycle.

[21]◁It turns out that they do

[22]◁The existence of a causal mechanism seems to be related to the amount of compressibility of data. Data that cannot be compressed appears random, and there can be no mechanism inside something that is random. On the other hand, data that compresses extremely well - such as a long series of all equal numbers - also seem to contain at best a trivial mechanism. Apparently, interesting mechanisms prevail in these regime where there is some, but not too much compressibility in data. The existence of a non-trivial mechanism is sometimes seen as an indication for the presence of *meaning*.

[23]◁This is a slightly extended version from the modeling process as described in Edwards and Hamson.

[24]◁It is tempting to believe that such data exists outside the context of any model. This is generally not the case, however. Raw data is always the result of a measurement procedure. To complicate things, this is even true for

sensorial data. What we think to perceive with 'unarmed eye' is a construct, produced by millions of visual receptors in our retina and several hundreds of millions neurons performing aggregations and other processing thereafter. The only thing that we know for certain is that all this processing leaves its traces in the produced measurement result; we can merely wonder if there is something there that could be called 'the original' signal that triggers the entire cascade of transformations.

A measurement can be a physical process such as a thermometer measuring the temperature of the liquid in a bottle; it can also be a social process such as a questionnaire asking interviewees their appreciation of a new brand X of lemonade. Such measurements, however, only can be said to produce, respectively, 'the temperature of the liquid' or 'the average preference for lemonade X' if we believe the model that underlies this procedure. Let us focus on measuring temperature first. What we call '$39.5°$C' is the level of Mercury (Hg) half way between two marks on a glass tube, one mark at 39 and the next mark at 40. For this Hg level to signify something else than *just* a level of Hg in a glass tube, we need an *explanatory model* that links the construct 'temperature' (=that which we want to know) to the motion of molecules, and next the motion of molecules to the volume of an amount of Hg. This model involves a large number of assumptions, many of which are known to be only rough approximations. Rather we should say that the number '39.5' is generated by a procedure (involving, in this case, a physical instrument), and that perhaps, by reasoning 'backwards' through a model for this procedure, we may find a construct 'temperature' that has some meaning. For the questionnaire this is even clearer: every social scientist knows that the answers given in a street interview are determined, among other things, by the weather, the looks and attitude of the interviewer, the phrasing of the questions and hundreds of other factors. So: it requires significant modeling - and hence: assuming and believing a lot of implausible things - to explain that there exists this construct of 'preference for lemonade X', *independent* from the measurement that is used to assess it. 'Obtaining data' therefore is not merely 'obtaining data': incorporating a procedure of measurement forms a significant part of the modeling process.

[25]◁This is similar to the impossibility to assess if the lock on your door is sufficient. You can only empirically conclude the opposite: if somebody breaks in it was obviously *not* good enough. This observation generalizes to the attempts to experimentally assess the truth of a hypothesis: this is logically impossible. The best thing to do is to seriously attempt to invalidate the hypothesis. The longer it withstands serious attempts of invalidation, the more confidence it gains in communities of practitioners.

[26]◁To adjust the view direction, the mouse cursor should be dragged (=moved, holding the mouse button down) in the image. When doing so, a circle appears. If the cursor movements stay outside the circle, rotation takes place around the axis, perpendicular onto the image. It is then as if the image plane is grabbed by the mouse, and re-oriented. If dragging takes place inside the circle, however, the view is re-oriented as if one is re-orienting a ball by dragging a point on the surface of the ball.

[27]◁This demo uses the following method to adjust the view: the three sliders labeled 'yaw', 'roll' and 'pitch' can be used to rotate the view in 3D.

[28]◁According to some authors, the Indians did know that horses were animals. Most likely, *some* Indians had this knowledge, whereas others were ignorant of the fact.

[29]◁For an entity that is represented in a conceptual model we have a corresponding concept. Often, we refer to the entity by using the name of the concept. An entity that has *no* representation in any conceptual model has no distinct name. As soon as we talk about some entity, we give it a name, and we implicitly build a conceptual model for that entity: namely, a language fragment featuring that entity. So we can talk about, say, a table, but it is questionable if we really refer to a thing that is beyond any conceptual model. Some philosophers believe that anything that we can talk about is a concept in our own, temporary, conceptual model of the world, made out of language constructs. Even merely thinking about a table is likely to make words appear in your mind. And for pictorial thinking, one might say that a mental picture also is a concept in some conceptual model, and not the real table. Therefore things that are not concepts are, at best, extremely elusive.

[30]◁According to Genesis, the first thing Adam was allowed to do after he settled in the freshly installed Paradise was to give *names* to all the animals and plants he saw. In the process, he gave names to *groups* or *classes* of animals on the basis of their intrinsic appearance. He did not, for instance, called animals walking to the left the

leftWalkers to distinguish them from rightWalkers. Indeed, such distinction would be meaningless, as individual animals would constantly shift from being one concept to another. Also, he did not give names to the individual lions (Simba, Clarence, ...), but rather named them as a species: lion. For animals that became domesticated, however, he may have preferred individual names: the cow named Clara III used to produce more milk, but she was more recalcitrant when being milked, than Clara IV, and for the milkmaid it was good to know this difference. Names - and words in general - seem to be related to *purpose*.

31◁For this reason, the act of naming, in Christian tradition, is considered to be an essential, or symbolic act. Hence the protocol of baptizing, where the name-giving is lifted to the spiritual level.

32◁To distinguish concepts from entities, we will use a different font. Entities are written in the standard font, whereas concept, as well as other formal items, are written in this font.

33◁Formally, P(C) and C['P'] mean the same thing, provided that C is an aggregation possessing a property P. In computer implementation, however, the first version will only work if there is a function defined with the name P, returning C['P'] when called with argument C. With this function, we can retrieve the P-value for any concept that possesses a property P

34◁There is some loss of information, though. By omitting the particles 'a' and 'the', we can no longer distinguish 'a dog is near a lantern' from 'the dog is near the lantern' from 'the dog is near a lantern' from 'a dog is near the lantern'. Notice how the meaning of these four statements shifts.

35◁It is a good habit to work with concepts that are singular (lantern in stead of lanterns). That there may be a multitude of lanterns in the conceptual model will come later.

36◁The set of real numbers in mathematics is denoted by the symbol $\mathbb{R}$. Computers prefer standard letters, hence real. The issue of to what extent a computer can actually represent a real number is subtle, and is not discussed here.

37◁In some cases, we are more detailed, and establish not only relations between concepts, but also between *properties of* concepts.

38◁In practice, it may suffice just to report relations following from our immediate understanding of the modeled situation. On the other hand, there also may be relations that involve 3 or more things.

39◁This, of course, is only true if there is no first and no last lantern

40◁Cf. Appendix **??**, this is indeed a form of abstraction. In this view, 'quantity' is an abstract class, and properties inherit from this class.

41◁Here, wood etc. are concepts. For many purposes, however, we don't need to know the properties of these concepts. This is similar to the use of the word 'wood' in natural language without explicitly mentioning the sort of wood, the color, the density, etc.

42◁Provided that the square of the perimeter is at least 8 times the area. Indeed, not every pair of values for area and perimeter define a valid rectangle.

43◁http:en.wikipedia.orgwikiMohs_scale_of_mineral_hardness.

44◁Scales themselves are also ordered: operations allowed to a nominal scale are allowed to a partially ordered scale; operations allowed to a partially ordered scale are allowed to a totally ordered scale; operations allowed to a totally ordered scale are allowed to an interval scale, and operations allowed to an interval scale are allowed to a ratio scale.

45◁Compare this to computing the average of two telephone numbers to find out the telephone number of somebody living in between.

46◁This table is a simplified version of the table in `http://www.graphpad.com/faq/viewfaq.cfm?faq=1089`.

47◁'Constant' here means: not depending on what is measured.

48◁Equivalence is a relation that is *reflexive*, *symmetric* and *transitive*. We encountered transitivity before when we discussed ordering in Section 2.6.2. The contstant-ratio-relation for units is transitive: if units u_1 and u_2 have a constant ratio, and so do u_2 and u_3, then u_1 and u_3 also have a constant ratio. 'REFLEXIVE', means that the relation applies between an item and itself. An example is `hasSameFatherAs`. The constant-ratio-relation for units is indeed reflexive: the ratio between a unit and itself is 1, which is constant. A relation is SYMMETRIC if the relation between `A` and `B` also applies between `B` and `A`. An example of a symmetric relation is `marriedTo`. The constant-ratio-relation for units is symmetric: if the ratio between u_1 and u_2 is constant (i.e., does not depend on what is being measured), then this holds as well for the reciprocal.

Concepts that are equivalent can be seen to belong to an EQUIVALENCE CLASS. Indeed, 'having a constant ratio to' is an equivalence relation. So there is an equivalence class consisting of all units such a m (meter), cm (centimeter), inch, km (kilometer), light year, etc. This equivalence class is called 'length'.

Similarly, units such as s (second), h (hour), d (day), μs (micro second), y (year), etc. are equivalent. They also form an equivalence class; this equivalence class is called 'time'.

49◁The fact that we write $\mathcal{L}$ and not $\mathcal{L} + \mathcal{L} + \mathcal{L}$ follows from reasoning by analogy with units. We write 3cm+5cm=8cm rather than 3cm+5cm=8(cm+cm) because this follows from applying the algebra both for the numbers and for the units, and the notation for dimensions follows from this rule.

50◁If two quantities are equal, their units need not to be equal: 3dm=30cm.

51◁Otherwise they would be the same dimension, by definition.

52◁Don't confuse the notation of a closed interval, $[a, b]$ with the aggregation of two non-named concepts or values, as in `[a,b]`. The first indicates a set of real numbers that typically has infinity many elements; it could also be denoted as $\{a...b\}$. The second is an ordered list with two elements, `a` and `b`.

53◁If c is an element of the domain of a single-variable function, it is a number. It may seem strange to call this 'a point'. Consider the following argument, however. Elements of the domain of a function of multiple variables correspond to multiple numbers, e.g. an element of the domain of the function $f(x, y) = x + y$ would be the pair $(3, 4)$. These form the coordinates of a two-dimensional point. Similar for functions of three variables (an element of the domain is a three-dimensional point), et cetera. Therefore it is common to call the elements of the domain of a function in general 'points' in stead of numbers - even if the function has only a single argument.

54◁There is a mathematical test that helps to determine if an extremum occurs at a critical point. This test uses second-order partial derivatives. The partial derivative of a function $f(x, y)$ again is a function of two variables. Therefore one can take partial derivatives with respect to x or y. Here we denote them by $f_{xx}(x, y)$, $f_{xy}(x, y)$, $f_{yx}(x, y)$ and $f_{yy}(x, y)$. If some conditions are satisfied one has $f_{xy}(x, y) = f_{yx}(x, y)$. Now we can formulate the *second derivative test*. Suppose that $f(x, y)$ has continuous second-order partial derivatives in some open disk containing the point (a, b) and that the first-order partial derivatives are zero in (a, b). Define the *discriminant* $D(a, b)$ for the point by $D(a, b) = f_{xx}(a, b)f_{yy}(a, b) - (f_{xy}(a, b))^2$. The test gives: (i) If $D(a, b) > 0$ and

$f_{xx}(a,b) > 0$, then f has a local minimum at (a,b); (ii) If $D(a,b) > 0$ and $f_{xx}(a,b) < 0$, then f has a local maximum at (a,b); (iii) If $D(a,b) < 0$, then f has a saddle point at (a,b); (iv) If $D(a,b) = 0$, then no conclusion can be drawn.

[55]◁To alayse the behavior of a function, we could of course just substitute different values for its arguments (=operate the sliders in ACCEL, specifying the values of all relevant quantities) and see what the calculated result is. A more efficient way is to go to the analysis tab in ACCEL and get a contour plot (click the checkbox 'graph'- this will force a contour plot instead of a graph). We want to visualize f (click 'f' in the right most table) in dependency of R1 and R2. Click the column 'H'(for horizontal) behind R1 and 'V'(for vertical) behind R2 in the leftmost table. The contourplot will be automatically drawn. All its properties, such as the range of the contour values, can be adjusted by filling in appropriate values in the text boxes. By moving the cursor in the figure, clicking, and dragging the cursor while holding the button down, the local values of R1, R2 and f are given.

[56]◁We clarify the notions of *closed* and *bounded*. A region $R \subset \mathbb{R}^2$ is BOUNDED if there is a disk that completely contains R. A region $R \subset \mathbb{R}^2$ is *closed* if it contains all its BOUNDARY POINTS. A point (a,b) is a boundary point of a region R if *every* open disk centered at (a,b) contains points in R and points outside R.

[57]◁More precisely stated: For an optimization problem, say to maximize $f(x,y)$, subject to inequality constraint $g(x,y) < 0$, a solution (x_S, y_S) is a point that satisfies $g(x_S, y_S) < 0$, whereas in other points (x_P, y_P) for which $g(x_P, y_P) < 0$, we have that $f(x_S, y_S) > f(x_P, y_P)$.

[58]◁The method of Lagrange multipliers can be used for optimization under constraints (TO DO: a concise explanation of the intuition behind the Lagrange Multipliers).

[59]◁This scheme is rather naive. It can be made more efficient in many ways. In particular, if the constrained inequalities are linear expressions in the unknowns, and if the function to be optimized is sufficiently well-behaved, there are powerful methods to solve the constrained optimization problem. These are the so-called SIMPLEX-METHODS .

[60]◁For further hints, consult Section 4.2.2.

[61]◁Examples of state transitions, not relating to socks, are scoring a goal in soccer, having your birthday, or switching on the light.

[62]◁A PROCESS is a behavior, such as in physics, or a collection of behaviors, such as in computer science. There the word 'process' refers to 'a running program', which can show different behaviors, for instance in dependence of varying inputs

[63]◁There is a subtlety: in a collision of two billiard balls, energy, momentum and angular momentum are only conserved after the colliding balls both have assumed their new motion states. For this reason, in physical calculations, we sometimes act as if various properties change simultaneously.

[64]◁'Leaving out quantities' is not always a deliberate action on the side of the modeler. Often, quantities are left out where the modeler is ignorant that these quantities possibly could belong in the system. The modeled system behaves in a way that seems random: things happen without an observable cause. Consider the case of somebody getting sick - that is: suffering from certain symptoms. It is possible that this person contracted a virus some days earlier. Since the 'quantities' associated with the process of this virus multiplying are hidden, large parts of the process go unnoticed. Only by the time that 'exposed' quantities change their value (say, body temperature rising), the chain of causes and effects becomes manifest. Major parts of fundamental research in disciplines as varied as quantum physics and human biology are in constant search for the existence of HIDDEN QUANTITIES that may help explain the unexplained.

[65]◁We remember that *intervals* are not totally ordered. Therefore, if events would take a finite amount of time, we could not assume them to be totally ordered. Processes, taking a finite amount of time, for the same reason are partially ordered.

[66]◁The distinction between (external) events and (internal) transitions is not always clear-cut. To a large extent, it is determined by system borders, which may be arbitrary. For instance, consider somebody typing at a computer keyboard. If the typist and the keyboard are considered to be one system, all key strokes are internal transitions, caused by hidden properties, such as the brain states of the typing person. If the typing person is thought to be *outside* the system, the key strokes are external events. But assume that the keyboard is old and worn out, and at some point a key stroke causes it to break, producing some faulty signal to the computer. It requires quite a bit of detailed cause-and-effect reasoning to find out what, in this case, causes the fault: is it the typist's action (an external event), or is it the slowly degradation of the structural stability of the keyboard that reaches a state where it no longer can withstand the action of typing (an internal transition, taking place between values of one or more hidden quantities)? An interesting case study in this respect, with a complex interplay between external events and internal transitions is the analysis of the precise cause of the Harold of Free Enterprise disaster in 1987, see `http://rzv113.rz.tu-bs.de/Bieleschweig/pdfB2/deStefano_Bieleschweig.pdf`.

[67]◁The fact that transitions, and hence: states, are totally ordered is no sufficient condition, but it is at least a good indication. Indeed, if there would be transitions that *could* occur in arbitrary order (as with partially ordered time), these certainly could *not* be causally related, and then a functional dependency would not be possible.

[68]◁ If we assume all Δ_i, s_i and g_i to be constant, say Δ, s and g, Expression 3.12 reduces to

$$A_{i+1} = A_i(1 - s\Delta) + g\Delta, \tag{7.4}$$
$$t_{i+1} = t_i + \Delta. \tag{7.5}$$

To calculate A for some t_t we need A_i for $i = t_t/\Delta$. Here we assume that t_t is an integer multiple of Δ. Next we use that Expression 7.4 is an arithmetic - geometric series: for recursive relations

$$A_{i+1} = aA_i + g\Delta, \tag{7.6}$$

we have

$$A_i = a^i + \frac{a^{i-1} - 1}{a - 1} g\Delta, \tag{7.7}$$

which is easily verified by evaluating $aA_i + g\Delta$ and checking that this yields the expression for A_i with i replaced by $i + 1$.

So, for $a = 1 - s\Delta$ we have

$$A_i = (1 - s\Delta)^i + \frac{(1 - s\Delta)^{i-1} - 1}{-s\Delta} g\Delta. \tag{7.8}$$

We can use this result to calculate $A_{t_t/\Delta}$ for arbitrary t_t: we found a closed form solution for A_i for any t without unrolling the recursion. Closed form results not always exist, however. With little additional terms added to Expression 7.4 we have to resort again to unrolling the recursion. For example: suppose that the bank gives an amount of interest ρ depending on the saved amount, $\rho_i = f_\rho(A_i)$. Then Expression 7.4 changes into

$$A_{i+1} = A_i(1 - s\Delta) + g\Delta + f_\rho(A_i), \tag{7.9}$$

and for most functions f_ρ there is no closed form solution for A_i, nor could we calculate the asymptotic value A_{infinity}. For *unrolling*, the addition of any extra terms is no complication: we just need to evaluate the recursive definition for the model quantities.

[69]◁This is the type of reasoning governmental agencies use when they decide on healthy national economies in the light of national debts etc.

[70]◁Such a step is sometimes called 'Ansatz', the German word for 'approach'. It refers to a heuristic procedure that often works, without guarantee for success.

[71]◁We have seen some cases where a recursive function gives an exact behavior of the motion of a point mass. There are many situations, however, where a recursive model does not hold exactly. Consider the case where two point masses are connected by a massless rod of length ρ. So the points keep at a fixed distance ρ from each other. They take the shape of a dumbbell. We write the recursive model for both:

$$r_{1;i+1} = 2r_{1;i} - r_{1;i-1} + a_{1\to 2;i}\Delta^2; \quad (7.10)$$

$$r_{2;i+1} = 2r_{2;i} - r_{2;i-1} + a_{2\to 1;i}\Delta^2. \quad (7.11)$$

The accelerations $a_{1\to 2;i}$ and $a_{2\to 1;i}$ have a physical meaning. They result from the forces working between the two point masses. These are equal and opposite (action = - reaction), and they can vary in time (hence the index i). We don't know how big they are. We do know, however, that they are exactly big enough to keep the distance between the point masses constant and equal to ρ. Further, since they are equal and opposite, they must work along the line $r_{1;i} - r_{2;i}$. So we can replace the vectors $a_{1\to 2;i}$ and $a_{2\to 1;i}$ by a single unknown SCALAR quantity β:

$$r_{1;i+1} = 2r_{1;i} - r_{1;i-1} + \beta(r_{1;i} - r_{2;i})\Delta^2; \quad (7.12)$$

$$r_{2;i+1} = 2r_{2;i} - r_{2;i-1} + \beta(r_{2;i} - r_{1;i})\Delta^2. \quad (7.13)$$

We are interested in the relative motion of the two point masses. Therefore we form the difference $R_i = r_{1;i} - r_{2;i}$. Subtracting Expression 7.12 and Expression 7.13 we get a recursive model for the relative movement:

$$\begin{aligned} R_{i+1} &= 2R_i - R_{i-1} + 2\beta R_i\Delta^2 && (7.14)\\ &= 2(1+\beta\Delta^2)R_i - R_{i-1} && (7.15) \end{aligned}$$

To find the value of β, we demand that $||R_{i+1}|| = ||R_{i-1}|| = ||R_i|| = \rho$, the distance between the point masses. To calculate $||x||$ we recall that $||x||^2 = (x, x)$, so:

$$\begin{aligned} \rho^2 &= ||R_{i+1}||^2 \\ &= (R_{i+1}, R_{i+1}) \\ &= (2(1+\beta\Delta^2)R_i - R_{i-1}, 2(1+\beta\Delta^2)R_i - R_{i-1}) \\ &= 4(1+\beta\Delta^2)^2(R_i, R_i) + (R_{i-1}, R_{i-1}) - 4(1+\beta\Delta^2)(R_i, R_{i-1}) \\ &= 4(1+\beta\Delta^2)^2\rho^2 + \rho^2 - 4(1+\beta\Delta^2)(R_i, R_{i-1}). && (7.16) \end{aligned}$$

This reduces to

$$(1+\beta\Delta^2)^2\rho^2 = (1+\beta\Delta^2)(R_i, R_{i-1}). \quad (7.17)$$

Write $(R_i, R_{i-1}) = \rho^2\cos(\phi_\Delta)$, then ϕ_Δ is the rotation of the dumbbell over time lapse Δ.

Equation 7.17 has two solutions. The trivial solution is

$$\beta = -\frac{1}{\Delta^2}. \quad (7.18)$$

If we substitute this back into Expression 7.14, we get $R_{i+1} = -R_{i-1}$. This means that for any three subsequent states, $i-1$ and i and $i+1$, the orientation of R flips 180° from $i-1$ to $i+1$.

The other solution is $1+\beta\Delta^2 = \cos(\phi_\Delta)$. If Δ is small, which also means that ϕ_Δ is small, the cosine can be approximated by a Taylor series:

$$\cos(\phi_\Delta) = 1 - \phi_\Delta^2/2 + \phi_\Delta^4/4! - \cdots. \quad (7.19)$$

Since for small rotation ϕ_Δ is proportional to Δ, we write $\phi_\Delta = \omega\Delta$, where ω equals the current rotational velocity. Then, up to first order in Δ^2:

$$1+\beta\Delta^2 = 1 - \omega^2\Delta^2/2 + \omega^4\Delta^4/4! - \cdots, \quad (7.20)$$

or

$$\beta = -\frac{1}{2}\omega^2 + O(\Delta^2), \tag{7.21}$$

which is again a secondary school result: the centripetal force is proportional to the square of the rotational velocity.

This result is not exact. We make an error proportional to Δ^2. With sampling step size twice as small, the error gets 4 times as small.

We look again at the example of the rotating dumbbell. We started with a recursive model with time lapse Δ. If Δ is sufficiently small we see that the simulation for the relative motion is a uniform rotation. The rotation speed is closer to the secondary school result when Δ is smaller.

Moreover, we looked at the *relative* locations only. The absolute locations don't occur in Expression 7.14. Any constant velocity can be added to $r_i - r_{i-1}$, for all i, and the simulation still holds. To check that the general solution for Expression 7.10 consists of a rotation plus a uniform velocity we substitute

$$\begin{aligned} r_{1;i} &= r_0 + v_0\Delta i + \frac{1}{2}\rho\begin{pmatrix}\cos(\omega\Delta i)\\ \sin(\omega\Delta i)\end{pmatrix}, \\ r_{2;i} &= r_0 + v_0\Delta i - \frac{1}{2}\rho\begin{pmatrix}\cos(\omega\Delta i)\\ \sin(\omega\Delta i)\end{pmatrix} \end{aligned}$$

into Expression 7.12 and Expression 7.13.

When sampling periodic phenomena, such as a rotating dumbbell, for given time lapse Δ, there is an upper limit to the frequency of the phenomenon that can be represented.

In this light we look again at Expression 7.18. Suppose the dumbbell rotates increasingly faster, where the motion is sampled with constant Δ. When ω gets so large that, in between two subsequent transitions, the rotation is a full turn, this cannot be distinguished from a situation where it rotates two full turns, or three full turns, etc.. This is also what happens e.g. in Western movies where rotating spoke wheels seem to rotate backwards. The sampling rate of 24 frames/second is not enough to capture the real rotation of the wheels.

In Expression 7.21, β is only approximately proportional to ω^2 if $\omega^2\Delta^2$ is small compared to 1. This is a fundamental limitation to sampling. It is called ALIASING. Aliasing means that there is a phenomenon, periodic with frequency ν_p, that is sampled with a frequency ν_s. (Re)constructing the phenomenon from the samples works well if ν_s is sufficiently high compared to ν_p. If ν_s is too low, the reconstruction differs from the original phenomenon. It assumes an 'alternative identity', which is the literal meaning of the word 'alias'.

We study one more example of unrolling recursive definitions. We make a recursive model for a mass spring system. Again we start from the order-2 version. The force is proportional to the deviation from a rest position, r_0:

$$r_{i+1} = 2r_i - r_{i-1} - \omega^2\Delta^2(r_i - r_0), \tag{7.22}$$

where $\omega^2 = \frac{C}{m}$, C the spring constant, and m the mass.

We try a solution of the form

$$r_i = A\cos(\Delta\phi i) + r_0, \tag{7.23}$$

for unknown ϕ and A. So

$$r_{i+1} = A(\cos(\Delta\phi i)\cos(\Delta\phi) - \sin(\Delta\phi i)\sin(\Delta\phi)) + r_0; \tag{7.24}$$
$$r_{i-1} = A(\cos(\Delta\phi i)\cos(\Delta\phi) + \sin(\Delta\phi i)\sin(\Delta\phi)) + r_0. \tag{7.25}$$

Form $r_{i+1} - 2r_i + r_{i-1}$ and equate this to $\omega^2\Delta^2(r_i - r_0)$. The result must hold for any i; this gives

$$2 - 2\cos\Delta\phi = \omega^2\Delta^2 \tag{7.26}$$

and arbitrary A. As with the rotating dumbbell, we find

$$\phi = \omega(1 + O(\Delta^2\phi^4\omega^{-2})). \tag{7.27}$$

We interpret this as follows. For Δ^2 small compared to $\omega^2\phi^{-4} = \frac{C}{m}^2\phi^{-4} \approx \frac{m}{C}^2$, that is, a spring with small spring constant C or large mass m, the solution oscillates with a frequency $\phi \approx \omega$. When Δ goes to 0, ϕ goes the correct value ω.

If, on the other hand, for a given Δ, the value $\frac{C}{m}$ is large, something else happens.

In the case of the dumbbell we had built in that the distance between the point masses stay constant. The simulation goes into a 'trivial' mode when Δ is too large. It then flips its orientation every two transitions.

In the mass-spring example there is no built-in mechanism that keeps the solution within bounds. The repeated evaluation of Expression 7.22, for $\omega\Delta$ too large, becomes UNSTABLE.

If a recursive model of the form $r_{i+1} = 2r_i - r_{i-1} + K\Delta^2$ gets unstable, the subsequent distances $||r_i - r_{i-1}||$ get increasingly larger. In other words, the kinetic energy of the moving point masses ($\frac{1}{2}mv^2$, or $\frac{1}{2}m||r_i - r_{i-1}||^2\Delta^{-2}$) gets increasingly larger. We can understand this as follows. For a system to be stable, its kinetic energy needs to be more or less constant. So any force should contribute, on average, no work to the system. The dot product $(K_i, r_i - r_{i-1})$, on average, should be zero, where K_i is the force in state i. For a spring, for instance, the point mass moves in the direction of the force for half of the time, and in the opposite direction for the other half. If Δ is too large, there will be evaluation errors in the recursive model. Then there is no reason that the calculated $(K_i, r_i - r_{i-1})$ also is zero on average. The simulated system acquires kinetic energy. Indeed: the kinetic energy can only increase, since it is quadratic in the velocity. So whether the error in the velocity is positive or negative, the error in the kinetic energy will always be positive. As a result, differences $r_i - r_{i-1}$ will be slightly too large, and kinetic energy increases even further. The errors reinforce each other, and soon the simulation goes 'out of hand'.

In physical reality, just the opposite occurs. If we have an oscillating mass spring system, after a while it slows down. This is because of friction or damping. There is always a force that works in *opposite direction* of the movement. In the absence of such force, we would have $||r_{i+1} - r_i|| = ||r_i - r_{i-1}||$ (according to Expression 3.19), and hence conservation of energy.

If, on the other hand, $||r_{i+1} - r_i|| = \sigma||r_i - r_{i-1}||$, with $0 < \sigma < 1$, velocities get increasingly smaller. A simple way to build in such DISSIPATION (=loss of energy in a dynamic system, typically due to friction or damping) is by introducing a force that always counteracts the movement. Since forces, and hence accelerations, can be added together, such a counteractive contribution can be simply added to other forces in the recursive model. So we write

$$r_{i+1} = 2r_i - r_i + \Delta^2 K - \epsilon(r_i - r_{i-1}) \tag{7.28}$$

instead of

$$r_{i+1} = 2r_i - r_i + \Delta^2 K, \tag{7.29}$$

for whatever acceleration K we want to consider. Indeed, for $K = 0$ we get

$$r_{i+1} = 2r_i - r_i - \epsilon(r_i - r_{i-1}), \tag{7.30}$$

so

$$\begin{aligned} r_{i+1} - r_i &= r_i - r_{i-1} - \epsilon(r_i - r_{i-1}) \\ &= (1-\epsilon)(r_i - r_{i-1}). \end{aligned} \tag{7.31}$$

For small positive ϵ, in the absence of other forces, the kinetic energy decreases in time with a factor $(1-\epsilon)^2$ per time step. So adding a term $-\epsilon(r_i - r_{i-1})$ imitates the effect of dissipation in a physical system, and stabilizes our calculations. *Stabilizing* means: protecting against 'going out of hand'. The addition of a term $-\epsilon(r_i - r_{i-1})$ is trivial for the evaluation of the recursive functions in unrolling the simulation. Approaching this solution using differential equations (see Expression 3.3.4) is much more elaborate.

[72]◁For linear differential equations with a sufficiently small non-linear additional term, there is a technique called LINEARIZATION, which boils down to doing a first-order Taylor expansion to the non-linear contribution.

[73]◁This assumes that r, as a function of time t, behaves well. 'Well behaving' means: r first needs to be defined in any t, in a sufficiently narrow surrounding of t it should should not vary more than proportional to the size of the surrounding, and, most importantly, its derivative should exist.

[74]◁To see the corresponding differential equation for a recurrent function of order 2, we realize that

$$r(t+\Delta) = r(t) + \Delta \frac{d}{dt} r(t) + \Delta^2 \frac{d^2}{dt^2} r(t)/2! + \Delta^3 \frac{d^3}{dt^3} r(t)/3! + O(\Delta^4) \quad (7.32)$$
$$r(t) = r(t) \quad (7.33)$$
$$r(t-\Delta) = r(t) - \Delta \frac{d}{dt} r(t) + \Delta^2 \frac{d^2}{dt^2} r(t)/2! - \Delta^3 \frac{d^3}{dt^3} r(t)/3! + O(\Delta^4). \quad (7.34)$$

By adding two times Expression 7.33 and subtracting once Expression 7.34 we get

$$\frac{d^2}{dt^2} r(t) = \frac{r(t+\Delta) - 2r(t) + r(t-\Delta)}{\Delta^2} + O(\Delta^4), \quad (7.35)$$

or

$$r(t+\Delta) = 2r(t) - r(t-\Delta) + \Delta^2 \frac{d^2}{dt^2} r(t) + O(\Delta^4). \quad (7.36)$$

This helps to understand why we encountered the expression '$2r_i - r_{i-1}$' in Expression 3.28, Expression 7.10 and Expression 7.22. In the form of differential equations, these models read, respectively:

$$\frac{d^2}{dt^2} r(t) = a \quad \text{(constant acceleration);} \quad (7.37)$$

$$\frac{d^2}{dt^2} r_1(t) = a_{1\to 2}(t), \quad \text{(massless rod)}$$
$$\frac{d^2}{dt^2} r_2(t) = a_{2\to 1}(t); \quad (7.38)$$

and

$$\frac{d^2}{dt^2} r(t) = -\omega^2 (r(t) - r_0) \quad \text{(mass spring system).} \quad (7.39)$$

As an illustration, we give the closed-form solution of the last one. We verify that $r(t) = r_0 + A\cos(\omega t)$ for arbitrary A solves Equation 7.39. Similar as with recurrent functions, these differential equations have order two: named after the highest occurring derivative in the right hand side.

When we introduced dissipation in the recurrent functions, we added a term $-\epsilon(r_i - r_{i-1})$. In the case where $\Delta \to 0$, the difference $r_i - r_{i-1}$ plays the role of $\Delta \frac{d}{dt} r(t)$. This is inconvenient, since Δ goes to 0. But we remember that the interpretation of ϵ was a reduction of speed *per time step*. So we still can work with a dissipation term $-\epsilon \frac{d}{dt} r(t)$, but the dimension of ϵ has to be 1/time. Indeed, the dissipation term needs to be proportional to 1/time^2 in order to add it to other acceleration terms.

Including dissipation the mass spring system becomes

$$\frac{d^2}{dt^2} r(t) = -\omega^2 (r(t) - r_0) - \epsilon \frac{d}{dt} r(t) \quad \text{(dampened mass spring system).} \quad (7.40)$$

This illustrates our remark about the advantages and disadvantages of using differential equations versus recurrent functions. Adding dissipation causes no additional complication for unrolling the recursion. The solution for the combined 1st- and 2nd order differential equation Expression 7.40, however, is harder to guess than the solution for Expression 7.39. A closed-form solution is possible; a full treatment of linear second order differential equations, however, falls outside the scope of these lecture notes.

[75]◁Adding 'points' is a somewhat sloppy abbreviation of adding *vectors*. Let us choose an origin O; then $p+q$, where p and q are both vectors starting in O, is a vector, which extends from O to the far point of a parallellogram spanned by O, p and q. The vector $\frac{1}{2}(p+q)$ is the vector, extending from O to the midpoint of the parallellogram. If we would have chosen another origin, say O', where $O' = O + \delta$, the summation $p' + q'$ is equivalent to $p - \delta + q - \delta$

since $p' = p - \delta$ and $q' = q - \delta$. So $M' = \frac{1}{2}(p' + q') = \frac{1}{2}(p + q - 2\delta) = \frac{1}{2}(p + q) - \delta = M - \delta$ where M is the midpoint when expressed with respect to O. So we have $M' + \delta = M$, in other words: M with respect to O refers to the same point in space as M' with respect to O'. So, the addition of $\alpha_p p + \alpha_q q$ is only independent of the choice of the origin if the sum of the coefficients α_p and α_q of p and q is one. In all other cases, the outcome will change if the origin changes. If we omit to explicitly define the location of the origin, the result would be undefined altogether.

To complicate things further, $p + q$ can also mean that p is a point, and q is a vector, in which case the result is a translated version of p, and hence a point. In this case, the rule that coefficients need to add to 1 does not apply. By denoting points and vectors both as a column of numbers, without further context, confusions such as these are difficult to avoid.

[76]◁This is sad, but obvious: suppose that there *would* be a theory of translating non-mathematical ideas into mathematics. For this theory to be a formal theory, we must be able to unambiguously *prove* its results. So these results must be written in terms of formulas with a unique, interpretation-less meaning. But the things that need to be written in these formulas are ... intuitions that have not yet been translated. So: there is a huge chicken-egg problem, forbidding that we can formally argue about the correctness of the translation of some non-mathematical intuition into something mathematic.

[77]◁Also, there are many application domains where the mapping between 'reality' and mathematics is not so problematic after all. Interestingly, this is often the case because, in fact, the mapping works the other way round. For instance, in the world of finance: financial transactions are often *defined* by mathematical operations, money being just another representation of numbers. So if I give you a 50-Euro note, and I expect five 10-Euro notes in return, the mathematical expression '50 = 5 × 10' is not so much a translation of the perceived exchange of banknotes; rather, the banknotes are exchanged in order to do justice to the equality of 50 and 5 × 10. Similar in the design of artifacts: mathematical representations are often used to *specify* the relations between quantities rather than to *analyze* them.

In physics (and natural sciences as a whole), however, the situation is much more subtle.

For instance, many phenomena in physics are described using something called the SUPERPOSITION PRINCIPLE. This principle holds that the combined effect of two or more actions or interactions is described by the mathematical sum of two quantities, each quantity representing one of the two actions or interactions.

The simples example is *force*: when two forces, applied to a point mass, work simultaneously, this is indistinguishable from a single force which is the vector sum of the two original forces. Principles such as the superposition principle first have been observed empirically. They were verified with experimental observation, and time after time superposition seems to hold within experimental accuracy. This, after a while, has inspired physicist to search for theories such that the superposition principle is *built in*. That means: any calculation with such a theory will guarantee that the outcome satisfies superposition. Such theories often center around so-called LINEAR DIFFERENTIAL EQUATIONS. A linear (differential) equation is an equation with the property that, if f_1 and f_2 are solutions, any combination $\alpha_1 f_1 + \alpha_2 f_2$ is also a solution. Important examples are the equations for mechanical and electrical systems, electromagnetism, and quantum mechanics. We must realize, however, that the superposition principle, and hence the predominant role of linearity in all of physical theory, cannot be mathematically proven. It is only a practical way for describing reality, that thus far seems to comply with empirical measurements, and in particular with the empirically observed superposition principle.

Linearity is a very desirable property for quantitative theories of physical phenomena - but also for economical, biological or other contexts. The mathematical treatment of non-linear problems is tremendously more complicated. Still, there is a fundamental problem with the representation of a system as a linear system. In terms of the input-output behavior of the studied system, it means that $f(\lambda x) = \lambda f(x)$ and $f(x_1 + x_2) = f(x_1) + f(x_2)$, where x is some input signal. f represents the system, and λ is an arbitrary real number. The first requirement, $f(\lambda x) = \lambda f(x)$, must hold for arbitrary large or small λ, both negative and positive. There exists no material system, however, that can accept an infinitely wide range of input values, including negative ones, and that at the same time is sufficiently sensitive to respond to any relative change in the input, irrespective how small it is. Indeed: a relative change of. say, 0.000000001% in the input should produce a relative change of 0.000000001% in the output. It is impossible to empirically verify that, for instance, the laws of electromagnetic interaction have these properties; it is next to certain that all known mechanisms in economy, biology etc. *don't* have these properties.

[78]◁A form for S that gives a reasonable match with empirical data is $S(r_1,r_2,\beta) = (\frac{(r'_1+r'_2,n)}{||r'_1+r'_2||})^\beta$, where the vector n is the vector (0,0,-1), that is: the unit vector perpendicular to the reflecting surface. The vectors r'_1 and r'_2 are the normalized vectors $r_1/||r_1||$ and $r_2/||r_2||$, respectively; $r'_1+r'_2$ is the so-called halfway vector. If r_1 and r_2 have equal but opposite angles with respect to n, the dot product $(r'_1+r'_2,n)$ equals the length $||r'_1+r'_2||$, so $(\frac{(r'_1+r'_2,n)}{||r'_1+r'_2||}) = 1$, and $S=1$. S drops off to 0 if the angles differ more; dropping off is sharped if β is larger. This model was based on a proposal in 1973 by Computer Graphics researcher Bui Tuong Phong, and modified by Jim Blinn in 1977, and is named Phong-Blinn reflection ever since.

[79]◁For educational reasons, we deliberately formulate no purpose in this first example.

[80]◁By having separate units such as 'wash' and 'fam', we avoid mistakes such as attempting to add washes and families.

But now consider the following. We ask for the amount of detergent dumped by *French* households. Suppose that average number of people per Dutch family is different from the number of people in a French family. We then could do the calculation with the French family size. At some point we need to correct for the ratio between the Dutch and the French family size. This ratio is formally dimensionless, but we could give it the unit: [people/(fam France)]/[people/(fam Holland)]. By pretending that 'France' and 'Holland' are also units, and respecting the algebraic rules, this ratio between French and Dutch family size would be labeled as [Holland/France]. We use this heuristic idea in the chimney sweepers example.

[81]◁Unless, for instance, we would happen to know the list of population numbers of each of the provinces.

[82]◁The third option is not completely independent from the model we developed earlier. It also uses estimates of the number of people or the number of families having their chimneys swept in Eindhoven. So any problems with the accuracy of `nrChPFam` will not show up as a discrepancy between the two model outcomes.

[83]◁We ignore the possibility that our peanut butter might increase the total market for peanut butter.

[84]◁When a behavior is needed that involves multiple arguments, the Function Selector can be used for one argument at a time. For instance, in Physics the Boyle-Gay Lussac-Avogadro law states that the pressure P of an amount n of confined gas (in moles) at temperature T has the following behavior:

- P increases proportionally with n;
- P increases proportionally with T;
- P decreases inversely proportional with V.

The Function Selector would suggest here three functions:

- $P = C_1 n$;
- $P = C_2 T$;
- $P = C_3/V$,

and by assuming that there are no other factors that should be taken into account, a modeler might write

$$P = \frac{nRT}{V}, \tag{7.41}$$

where R is a constant. Notice that this works in virtue of the fact that $P = f_P(n,T,V)$ is a SEPARABLE FUNCTION, that is: it can be written as a product of functions that each depend on only one quantity, and that all three functions are positive. This latter condition means that whether or not $P = f_P(x)$ is increasing or decreasing in x (x being either n, T or V) does not depend on any of the other quantities. Constructing functions of multiple quantities in a more general case is tricky, and the Function Selector should be used with great care in these circumstances. In particular the case of non-separable functions is not covered well by the Function Selector.

[85]◁We were looking for a function of the form `marketShare`$=f($`pricePerItem`$)$. This function has only one argument. Taking for f a concrete expression, such as the ramp function, may introduce further arguments. In this

case, four of them. Formally, there is no difference between quantities that act as arguments and those that have other meaning. All occurring quantities need to be known for the function to evaluate.

86◁There are many applications of the Richards function where the quantities do have a clear interpretation in terms of a causal process. For instance, in systems biology or physics.

87◁Suppose that our panel would have consisted of *all* potential peanut butter buyers, and suppose that we would have kept the data for each panel member in all detail. In that case, we could have performed a fit for every panel member, and found the function $f_{\text{fit}}(pPI, c)$.

88◁Actually, it is a trilemma, as there are *three* options to consider.

89◁The model we will be developing here is extremely simplistic. It ignores many crucial features, such as the cyclist's metabolism. Actual models to calculate efforts for cycling with a precision that it can help taking decisions, e.g. regarding the strategy for a Tour de France runner, involve several tens of quantities.

90◁Functions cannot represent causal loops, unlike equations. We can easily solve problems of the sort 'x=4y-1' and at the same time 'y=x+2'. However, the two functional definitions 'y=x+2' and 'x=4y-1' cannot be expressed simultaneously in the 4-categories DAG. When using the 4 categories DAG-approach, we therefore always seek to describe the modeled system as a functional system without causal loops. For two categories of systems this is always possible: for *design*, where a designer takes independent decisions, as we do not model 'dependency-by-anticipation', and for *dynamical systems* for which there is a recursive function (see Section 3.3.2).

91◁In any non-trivial design problem, cat.-I space has a very high dimensionality, that is: there are tens or even hundreds of quantities in cat.-I, many of them assuming many possible values.

92◁The former example illustrates the rare situation where a *condition on the ATBD* is expressed in terms of a cat.-I quantity. Indeed, the operation mode of an artifact is probably something that is free for the designer to decide, so `ATBD.operationMode` could be a cat.-I quantity with value `manual`; the requirement that the operation mode should be `manual` means that any realization of the artifact where the operation mode is not manual, is, with respect to the condition regarding `operationMode`, inferior to one where the operation mode is `manual`.

93◁In some cases there are theoretical boundaries on some cat.-II quantities. For instance, thermodynamics can dictate that some machine cannot work unless it consumes X energy; in information theory, there are results that forbid more than Y bit/second of information to be transported over a channel. These boundaries, in general, are not very tight in the sense that technologically feasible solutions perform considerably less than the theoretical limits.

94◁Combinations of requirements, desires and wishes such as the following occur: an ATBD must weight less than X1 (requirement), it would be nice if it would weight less than X2 (X2<X1; a desire), and it would also be nice to find or approximate the point in the current design space where the weight is least.

95◁There are more sophisticated approaches to explore the design space of a functional model in the presence of requirements. These amount to applying transformations to the cat.-I quantities, such that constraints are automatically fulfilled. One example: suppose we are looking for tuples (x, y) that should fulfill the constraint $x^2 + y^2 = 1$. We could imply a constraint $x^2 + y^2 - 1 = 0$, or we could impose a penalty function $|x^2 + y^2 - 1|$. The first approach has the disadvantage that solutions in a population of candidate solutions, for which $x^2 + y^2$ differs from 1 are discarded; the second approach is a heuristic that might fail, for instance if there are many simultaneous penalty functions. A much more robust method is, to transform the tuple (x, y) to $(\cos\phi, \sin\phi)$, that is: to replace the *two* cat.-I quantities x and y by *one*, ϕ, such that the requirement is automatically fulfilled. In general, it may be very difficult, however, to find such transformations.

[96]◁Assuming some conditions on the mathematical properties of the functional model

[97]◁In particular, the work of Eckart Zitzler should be mentioned: http://www.ekki.ch/WWW/Eckart_Zitzler.html.

[98]◁If all new individuals would result from random assignments of values to cat.-I quantities, we achieve broad coverage of the solution space. Slow convergence may result, though, since the population won't benefit from fortunate interbreeding of two half-successful individuals. On the other hand, if all new individuals result from crossing over, we maximally benefit from interbreeding, but the diversity of the population stays limited. The art of tuning a genetic design process is, to have a good balance between the various mechanisms.

[99]◁At the expense of computation time: the amount of effort to do one iteration is proportional to the size of the population

[100]◁We assume that δ is sufficiently small, and disregard other numerical subtleties.

[101]◁We have not yet explained what to do with non-numerical cat.-I quantities, or with non-continuous dependencies. We give no exhaustive treatment of these cases here.

[102]◁Notice that the amount of matter in the universe is estimated on the basis of outcomes of *other* cosmological models. The problem, of course, could also be in the other models.

[103]◁An application is to serve in forensic contexts, to gather evidence for criminal offences. There is a serious dilemma between privacy and safety issues: unlimited saving, storing and coupling of data may help to resolve crimes, but it also may form a major challenge on privacy.

[104]◁Taken from 'Super Crunchers' by Ian Ayers, Random House.

[105]◁An assumption that does not hold does not necessarily mean that a model outcome is wrong. It merely means that we have not (yet) presented an argument of logical necessity showing that it should be correct.

[106]◁We discuss the example of weak and strong assumptions and outcomes for the wallet example in the simple case where strength or weakness of an assumption is directly expressed as an interval of tolerance. As follows: the assumption that a quantity is between 4 and 5 is a stronger assumption than that it is between 3 and 8; the former assumption implies the latter. For expressions, a similar argument can be sometimes given. Indeed, the assumption that $x = y + 1$ is a stronger assumption than the assumption that $y + 0.5 < x < y + 1.5$. This is fine for expressions that evaluate to single values; it needs non-trivial generalization if we want to express, for instance, that the assumption that y is an increasing function of x is a stronger assumption than that y is a monotonous function of x. Fur such cases, it is not sufficient to argue about tolerance intervals.

[107]◁Linear least squares technique and normal distribution will be explained later in this chapter

[108]◁Next to accuracy and precision, sometimes a third category of deviations is distinguished: the MODEL ERROR. 'Error' relates the outcome of a measurement or a model to a hypothetical better measurement or a more complete model. 'Better' means that the measurement or the model contain less known failures. A simple example is that in an additive model, where a value is calculated as the sum of a number of contributions, one or more contributions are omitted. For instance, in the taxi model of Appendix **??**, the advertisement costs are not included in the term `expenses`). We know that this is wrong, because advertisement costs do exist in the modeled system. A model that would incorporate advertisement costs would be better.

[109]◁Modeling always deals with reducing uncertainty. This can be illustrated as follows. Let p be a cat.-II quantity. When introduced in the conceptual model, it has a type, which is a set V_P of values. Let S_0 be the size of this set. After completing the execution of the formal model, P can occupy a range of values that form a subset of V_P. The size of the subset, say S_1 will be not larger than S_0. This is true whether V_P is a discrete set or a continuous set. The information increase, as a result of the execution of the model, is determined by the ratio of $\frac{S_1}{S_0}$. Suppose $\frac{S_1}{S_0} = 0.5$. It is then meaningful to say that the increase of information (=the reduction of uncertainty) corresponds to 1 bit. For instance: V_P was the set $\{a_1, a_2\}$, and the model calculation has revealed that in fact $P = a_1$ and not $P = a_2$, this can be expressed by one bit. Similarly, if $\frac{S_1}{S_0} = 0.25$, as in the case where $V_P = \{a_1, a_2, a_3, a_4\}$, and after the model we know which of the four has resulted, we have gained 2 bit. In general therefore we can define $-\log(\frac{S_1}{S_0})$ the information gain, or the increase in certainty, as a result of executing a model. With multiple cat.-II quantities, the certainty increases with $-\log(\frac{S_{1;i}}{S_{0;i}})$ bits for each cat.-II quantity P_i. Halving an uncertainty interval corresponds, according to this definition, to one additional bit of certainty increase.

[110]◁Provided there is a unique definition of what a 'letter' is. Do interpunction characters count as letters? And spaces? Is there a space between two letters if they are more than ... mm apart? As is always the case, counting starts with a unique and unambiguous classification process that determines if something should be counted as one.

[111]◁With measurements there are two caveats.

First, the assumptions of independency and 'measuring-the-same-thing-every-time' either require additional models or theories, or these assumptions are wrong. We encountered an example in Section 1.4.3. The 'average mass' of a blackbird has no meaning if the blackbird samples are taken, unknowingly, from two different subspecies of blackbird that have different average longevities.

The second caveat regards producing a single measured value. This requires the evaluation of a model of the measuring instrument. Consider measuring the value of an electric resistor by reading off the deviation of a needle in a multimeter. The value observed is an angle (in degrees), not a resistor (in Ohm). The angle of the needle is the result of a balance between a spring force and a magnetic force. The magnetic force is the result of a current in a coil. The current in the coil is the result of a voltage applied to a resistor. So there are no less than four model steps involved, getting from the measured value (the resistance) to the observed value (the angle of deviation). All these model steps involve assumptions and approximations. In practice, the number of model steps is much larger, and it can be very difficult to see the consequence of all involved assumptions and approximations.

[112]◁If contributions in an additive model all have the same sign, omission of a non-negative term leads to overestimate or underestimate. In a model for risk analysis: the chance of failure due to a number of independent processes is the sum of the chances of failure due to each of the processes, and the chance for each of the processes to fail is larger than 0. The more fail-processes we can think of, the better the model and the larger the chance for failure. A risk analysis model therefore always will be too optimistic. There are other additive processes, however, for instance in chemistry and particle physics, where unknown terms can be both positive and negative. In such cases the sign of the error is not a priori known. Furthermore, the result of failing mechanisms can be more complicated than in additive models.

[113]◁The condition that a quantity that is measured actually exists is seldom challenged. Still, this is a non trivial issue. Physical quantities like temperature or pressure only 'exist' in statistical sense, so measuring the temperature of an ensemble of 5 molecules is dubious. In social science, there are many quantities that result from a procedure, but of which it is questionable if there is an independent cause for their existence. Intelligence has long been a notorious example: is there a definition of 'intelligence' that is independent of a procedure of measurement?

[114]◁It is even possible to extrapolate to x_∞, giving a sequence $x_0 = f_x(h)$, $x_1 = f_x(h/2)$, $x_2 = f_x(h/4)$, ..., $x_p = f_x(h/2^p)$, assuming that the truncation error is given by $x_p = x_\infty + a(\frac{h}{2^p})^q$

[115]◁The sum of throws for n dice, converges for n to infinity to the normal distribution according to the central limit theorem

[116]◁There are other definitions for *mean*, however. For instance, the mean speed μ over a certain distance is the inverse of the arithmetic mean of the inverses of the speeds s_i for a set of equal sub-distances. This is called the HARMONIC MEAN, as follows:

$$\mu = \frac{n}{\Sigma_{i=1\cdots n} s_i^{-1}}. \tag{7.42}$$

Suppose two stretches of road of equal length; the first we drive 50 km/h and the second we drive 100 km/h. The harmonic mean $= \frac{2}{\frac{1}{50}+\frac{1}{100}} = 66.7$ km/h. This is indeed correct: let the length of one stretch is l, then the first stretch took $l/50$ hours and the second stretch took $l/100$ hours. The total time is $\frac{l}{50} + \frac{l}{100} = \frac{3l}{100}$ hours; the total stretch is $2l$, so the mean speed is $2l/\frac{3l}{100} = 66.7$km/h. The result is irrespective of the actual length of the stretches.

[117]◁It can be shown that the estimator from Expression 6.5 is a little bit too low for small n. Moreover, if $n = 1$ the standard deviation should be undefined. For these reasons, Bessel has proposed a correction of the form $\sigma_0 = \sqrt{\frac{1}{n-1}\Sigma_{i=1\cdots n} s_i^2}$. Careful analysis shows that this estimator gives the correct outcome, for all n, in case the samples are taken from a normal distribution.

[118]◁ In all summations below, the range of the index is from 1 to n:

$$\begin{aligned}
\sigma &= \sqrt{\frac{1}{n}\Sigma_i(s_i - \frac{1}{n}\Sigma_j s_j)^2} \\
&= \sqrt{\frac{1}{n}\Sigma_i(s_i^2 + \frac{1}{n^2}\Sigma_j\Sigma_k s_j s_k - \frac{2}{n}s_i\Sigma_j s_j)} \\
&= \sqrt{\frac{1}{n}\Sigma_i s_i^2 + \frac{1}{n^3}\Sigma_i\Sigma_j\Sigma_k s_j s_k - \frac{2}{n^2}\Sigma_i\Sigma_j s_i s_j} \\
&= \sqrt{\frac{1}{n}\Sigma_i s_i^2 + \frac{1}{n^2}\Sigma_j\Sigma_k s_j s_k - \frac{2}{n^2}\Sigma_i\Sigma_j s_i s_j} \\
&= \sqrt{\frac{1}{n}\Sigma_i s_i^2 + \frac{1}{n^2}\Sigma_j\Sigma_k s_j s_k - \frac{2}{n^2}\Sigma_j\Sigma_k s_j s_k} \\
&= \sqrt{\frac{1}{n}\Sigma_i s_i^2 - \frac{1}{n^2}\Sigma_j\Sigma_k s_j s_k} \\
&= \sqrt{\frac{1}{n}\Sigma_i s_i^2 - (\frac{1}{n}\Sigma_i s_i)^2}
\end{aligned} \tag{7.43}$$

[119]◁The mathematical term is 'invariant'.

[120]◁We notice that the result for a is closely related to the correlation ρ according to Expression 6.9):

$$a = \rho\frac{||y - \bar{y}||}{||x - \bar{x}||}. \tag{7.44}$$

It is a good habit, for results such as Expression 7.44, to spend a minute and ask if such result can be intuitively understood. As follows: consider the case where there is no statistical noise, and (for simplicity) assume that $\bar{x} = \bar{y} = 0$. Then, for all i, $y_i = ax_i$, and $|\rho| = 1$. So we then have that $|a| = \frac{||y||}{||x||}$. Is this plausible? We can understand this expression as follows. Suppose that we apply a scale factor s to all y_i: $\tilde{y_i} = sy_i$. Then $||\tilde{y}|| = s||y||$, but also $\tilde{y_i} = sax_i$, so the line equation becomes $\tilde{y} = sax = \tilde{a}x$, or $\tilde{a} = sa$. So indeed a and ρ are proportional, and (without noise and $\bar{x} = \bar{y} = 0$), $|a|$ is proportional to $||y||$. Similar for scaling the x_i. So, at least for the special case of noiseless, 0-average data, we have an independent argument that Expression 7.44 could be correct. We should now proceed to show that the argument also holds if $\bar{x}$ or $\bar{y}$ differ from 0, and perhaps even to consider the case where noise is present. These investigations are left as an exercise to the reader.

[121]◁In general it is not a priori known how many clusters there are. Obviously, we can get a perfect match if the number of clusters equals the number of data points. Then the residual error is 0, irrespective of the quality of the black box models.

The challenge in finding clusters therefore is, to balance the number of free quantities and the residual error. This poses relevant questions. Recently, Jamaican athletes perform exceptionally well in running. Suppose that detailed cluster analysis shows that we get a significant lower residue by not having two, but four clusters: Jamaican men, non-Jamaican men, Jamaican women and non-Jamaican women. Is this sufficient ground to split the Olympic categories, in other words, is it fair to have Jamaican runners compete with non-Jamaican runners? In paralympics(=an international sports competition with athletes with physical handicaps), what categories of handicaps should be defined such that competition is fair? Obviously, we enter a subtle debate were conclusions drawn from modeling have consequences with ethical bearing. This happens in sports, but also in social sciences, involving models for demography, work and income, schooling and education and many others.

[122]◁On the other hand, h should not be unnecessarily small, because then the difference between $F(x1+h, x2, x3, \cdots)$ and $F(x1, x2, x3, \cdots)$ may drop below MACHINE PRECISION, that is: the computer cannot tell $F(x1+h, x2, x3, \cdots)$ and $F(x1, x2, x3, \cdots)$ apart. So it is good practice to evaluate Expression 6.19 a couple of times with h, $h/2$, $h/4$, to see if the estimate for $\frac{\partial}{\partial x_1}F(x)$ can be trusted.

[123]◁For some functional dependencies condition numbers don't depend on the argument values. First, consider a MULTIPLICATIVE expression:

$$F(x) = \Pi_{j=1\cdots n} x_j^{\beta_j}. \tag{7.45}$$

Then

$$\begin{aligned} |c_i| &= |\frac{\partial}{\partial x_i}F(x)\frac{x_i}{y}| \\ &= |\beta_i \Pi_{j=1\cdots n} x_j^{\beta_j} x_i^{-1}\frac{x_i}{y}| \\ &= |\beta_i \Pi_{j=1\cdots n} x_j^{\beta_j} x_i^{-1}\frac{x_i}{\Pi_{k=1\cdots n} x_k^{\beta_k}}| \\ &= |\beta_i|. \end{aligned} \tag{7.46}$$

So, if the exponent of some argument x_i is 1 or -1, the condition number for x_i equals 1, irrespective of the value of x_i or any of the other x_j. Also, if x_i occurs as a factor with some exponent β_i, as $x_i^{\beta_i}$, the condition number does not depend on any x_j. We already made use of this result in Section 4.3.1.

Models such as the ones derived in Sections 4.3.1 and 4.3.2, and most models that are derived using dimension analysis have a form such as Expression 7.45.

Next consider an ADDITIVE expression where all terms have the same sign:

$$F(x) = \Sigma_{j=1\cdots n} \beta_j x_j. \tag{7.47}$$

Then

$$\begin{aligned} |c_i| &= |\frac{\partial}{\partial x_i}F(x)\frac{x_i}{y}| \\ &= |\beta_i \frac{x_i}{y}|. \end{aligned} \tag{7.48}$$

Adding all condition numbers gives 1, and since all condition numbers are positive, each of them is less than 1. Condition number =1 means that a 1% variation in an input quantity causes a 1% variation in the output. So in an all-positive additive model, condition numbers are all less than 1; the more contributions there are, the smaller the condition numbers.

Moreover, condition numbers are proportional to $\beta_i x_i$, that is: the condition numbers for bigger terms are bigger. This should be kept in mind when adding terms to an additive expression: smaller terms may better be left out as their contribution soon will be negligible to the 'bulk' of the expression. This argument does not hold for additive expressions where both positive and negative terms occur: due to cancelation, condition numbers can be arbitrary large then.

Two final cases: for an exponential dependency, $f(x) = e^{\lambda x}$, the condition number is $|\lambda x|$, so it monotonically increases with x, both for increasing and decreasing exponentials. For a logarithmic dependency $f(x) = \log \lambda x$ it is $\frac{1}{\log \lambda x}$.

124◁There are also situations where an expression is replaced by another expression. For instance: replace $\sin(x)$ by x. This replacement is OK for small $|x|$, but it gets increasingly inaccurate if $|x|$ increases. Analyzing the effect on resulting uncertainty of expression replacements is subtle, because the effect depends on the values of quantities in the expressions. We will limit our work to the situation where an expression is to be replaced by a constant

125◁Source: Mike Wallace, Gotham: A History of New York City to 1898, New York: Oxford University Press, 1999.

126◁When there is no risk for confusion, we drop the extension '*for-the-modeling-process-as-a-whole*'.

127◁In Chapter 6, we gave quantitative means to assess the reliability of models. We can therefore use the methods of Chapter 6 to compare outcomes of models. To compare *entire solutions*, however, statements about model outcomes are insufficient. We need to take the interpretation stage into account as well. For this aim, we introduced the eight criteria discussed above.

We have earlier needed to compare things using criteria. In Chapter 5 we studied models for supporting design decisions. There, the cat.-II quantities were a means to compare ATBD's. We introduced 4 categories of quantities to streamline constructing a design model.

Now we lift this idea to the modeling process *as a whole*. Just as we can talk about the stakeholder's value for an ATBD, expressed in terms of cat.-II quantities, we can talk about a purpose for a model and how this purpose is fulfilled. In the same way as we represent a designer's choices for the ATBD by cat.-I quantities, we can represent the modeler's choices for the type of model, and the modeling techniques, using the dimensions from in Section 1.3.

Now we make the correspondence between the 4-categories approach and assessing a model's merits more concrete. We want to discuss the way a model fulfills its purpose, and we re-use the idea of the four categories to do so.

We apply the four categories to generic modeling is as follows.

- A cat.-II quantity for a design model is the extent to which an ATBD fulfills a requirement, desire, or wish of a stakeholder. A criterion for modeling *in general*, such as the eight criteria we introduced here, is *the extent to which a model fulfills its purpose*.
- A cat.-I quantity for a design model is a free choice for the designer. A cat.-I quantity for modeling *in general* is *the choice for a modeling strategy for the modeler*.
- A cat.-III quantity for a design model is a factual circumstance from the context of the design. A cat.-III quantity for modeling *in general* is *a factual circumstance from the context of the model* (that is, the modeled system, the stakeholders or the model's purpose).
- A cat.-IV quantity for a design model relates to mechanisms, internal to the way designer's decisions affect the eventual merits of the ATBD. A cat.-IV quantity for modeling *in general* relates to *model-internal considerations*.

The interface between an ATBD and its context consists of quantities of categories I, II and III; we adopt the same convention for the interface between models and their context.

To assess if and how a model fulfills its purpose, we use the eight criteria introduced above. These are the counterparts of cat.-II quantities. To choose how to set up models (e.g., to choose among the distinctions in Section 1.3) we encounter the counterparts of cat.-I quantities, and to fix modeling contexts or constraints we work with the analogues of category III.

Similar to the connection between an ATBD and its context, the connection of a model and its context works in terms of quantities of categories I, II and III.

Analyzing the various aspects to a model as an ATBD goes a bit further. We can also say something about 'inputs' and 'outputs' of a model. Indeed, the connection of a model and its context regards an INPUT - and an

OUTPUT side. This use of 'input' and 'output', however, does not exactly coincide with the use of 'input' and 'output' for models for design. Indeed, in design, the entire ATBD was represented by a single function: the design model was a functional model, and the arguments of the function, being its input, correspond to categories I and III whereas the output quantities were, by definition, in category II.

In the context of modeling in general, the terms 'input' and 'output' relate to the modeling process. 'Input' relates to stage 1, were the problem is defined (see Section 1.4.1); 'output' relates to stage 5, where the results are presented and an interpretation of the model outcome is given (see Section 1.4.5). This means that, when we want to compare the merits of two modeling processes, we postulate cat.-II quantities-*for-the-modeling-process-as-a-whole* to do this comparison both from the 'input' and from the 'output': some cat.-II quantities-*for-the-modeling-process-as-a-whole* will relate to stage 1, others to stage 5. Stages 2, 3, and 4 could be seen as 'model-internal', and according to our analogy, they therefore constitute cat.-IV. Also in stages 2, 3 and 4, we will find room for cat.-I and cat.-III quantities.

[128]◁It has not been formally proven that an exact solution in polynomial time is impossible. A famous open problem in theoretical computing science is whether or not P=NP, that is: if the class of problems that cannot solved in polynomial time in the problem size is or is not disjunct from the class of problems that allow polynomial time solution. If P would be NP, this would be a break-through, since many important problems then could be transformed to problems with polynomial solution time.

[129]◁The Buys Ballot law states that, at the Northern hemisphere, wind circulation round a region of low pressure is counterclockwise. For high pressure, and for the Southern hemisphere, this is just opposite. This means that the knowledge of the position and distribution of depressions and regions of high pressure is enough to know the wind direction.

[130]◁In such circumstances, a so-called T-TEST can give reliability intervals for each of the two possible conclusions.

[131]◁We don't say that we *must* increase the model's convincingness: as we will see later in this chapter, convincingness is not a primary criterion for all modeling purposes; furthermore, convincingness is to a large extent a comparative, relative notion. What John finds convincing in context 1 may be found highly suspect by Mary in context 2, and they both may be perfectly right.

[132]◁A trivial compression, such that $c_A = A$ and $c'_A = A'$ has maximal distinctive power, but it is an extremely poor compression since the size of c_A (or c'_A) is not at all smaller than the size of A (or A')

Index

a priori◁ (a priori, van tevoren)
from the start, before doing anything;
Example: even before weighing a glass of water, you know that the found weight will be larger than the weight of the glass when empty , 28

a-cyclic◁ (niet-cyclisch)
of a graph: containing no cycles;
Example: a family tree is a non-cyclic graph, since nobody is his own ancestor , 137

abstraction◁ (abstractie)
- as a purpose of models: the process of leaving out details that are unnecessary for some given purpose;
Example: by leaving out the size and shape, both gas molecules and cars can be abstracted to moving points. This enables treating the flow of gas and the flow of traffic in the same way , 15

accuracy◁ (nauwkeurigheid)
quality of the outcome of a model that relates to the deviation between the model and the modeled system. Compare precision;
Example: to prosecute a motorist with an alcohol promillage of 0.3 for drunken driving (in the Netherlands, the maximally tolerated alcohol contents = 0.2 per thousand), an alcohol detector needs to have an accuracy that is better than 0.1 per thousand , 213

adaptive◁ (aanpassend, zelflerend)
the property of some quantity that it takes an adequate value, given the actual circumstance;
Example: an adaptive step size in sampling means, that it gets smaller if calculated or measured values in subsequent samples differ more; if these values differ little, the step size gets larger , 276

additive◁ (additief)
in general: related to adding or accumulating;
Example: for condition number calculation: condition numbers are less than 1 in additive expressions with all terms with the same sign , 293

aggregate◁ (samenvoegen)
grouping together;
Example: the bundle of properties with their values in a concept are aggregated , 60

aggregation◁ (samenvatting)
obtaining data that are characteristic for an entire ensemble from data associated to the individual members of that ensemble;
Example: consider a population with individuals that can be assumed similar, that each have a salary. Taking the average of all salaries produces the average salary, which is a property of the population. Averaging is an example of aggregating , 31

algorithm◁ (algoritme, rekenrecept)
programmed sequence of mathematical instructions, typically performed by a computer;
Example: long division (Dutch: 'staartdeling') is an algorithm to obtain the quotient of two numbers , 23

aliasing◁ (aliasing)
the aspect of sampling that, for a given sampling constant (e.g., a time lapse for sampling in time) there is an upper limit to the amount of detail in the

sampled signal that can be reproduced; Example: audio pitches of close to 44100 Hz can not be reproduced on an audio CD; rotating spoke wheels in movies may seem to turn backwards; near horizontal lines on a raster image show jaggies , 284

analysis◁ (analyse)
- as a purpose for models: attempting to find out about certain properties of a system, not by studying the system proper, but a model of the system instead; Example: after a traffic accident, it is common to draw a schematic sketch of the situation. This is a model of the accident, used to analyse which party was responsible for the accident , 16

analytic◁ (analytisch)
- of a modeling strategy: making use of formal mathematical elaborations, operating on symbols rather than on numbers. Opposite to numerical; Example: finding the maximum of a differentiable function by setting the first derivative equal to 0 , 23

AND◁ (en)
logical operator: P AND Q is true only if P and Q are true; Example: a transaction is a purchase if both an item of value has been transferred from A to B and a sum of money has been transferred from B to A , 28

angle◁ (hoek, boog)
the distance between directions; Example: a perpendicular angle is 90 degrees or pi/2 radians , 24

approximation order◁ (benaderingsorde)
a numerical approximation has order n if halving the sampling step size gives a reduction of the error of 1/2 to the power n; Example: estimating the area underneath a function using the rectangle rule, approximating the function as piecewise constant, has an error that is proportional to the width of the piecewise constant segments to the 3rd power: it is order 3 , 116

arc◁ (boog, kant, zijde)
an element in a graph, next to node; Example: in a diagram consisting of boxes and arrows, the arrows are the (directed) arcs; the boxes are the nodes. Indeed, arcs can be directed or undirected: a directed arc is an arrow , 30

argument◁ (argument)
-of a function: the value, taken from the domain of the function, that serves as input; Example: for the function 'color', an argument could be anything colored , 62

arrow◁ (pijl)
directed arc in a graph; Example: if a relation, say `isA` is represented by an arrow, the opposite arrow represents `specializesTo` , 30

artifact◁ (kunstproduct, maaksel)
something made by man, as opposed to a natural object; Example: an artefact needs not to be material: symphonies, laws and organizations are artifacts , 16

aspect ratio◁ (aspect(verhouding))
the ratio between height and width of a shape (e.g., a rectangle); Example: a square has aspect ratio 1, same as a circle , 73

assumption◁ (veronderstelling, aanname)
a non-proven, perhaps even false proposition about the state of affairs that is taken to be true in order for further propositions to be deducible; Example: in a system involving geometric optics, a mirror is assumed to be ideally planar. Material objects can never be mathematically planar, but in some contexts it can be reasonable to ignore deviations from planarity , 32

asymptotic behavior◁ (asymptotisch gedrag, langetermijngedrag)
behavior in the long run; Example: for a pendulum with length l,

the oscillation period is proportional to the square root of l, provided that the amplitude is small enough. But due to friction, any free swinging pendulum, in the long run, will reach a state where this condition is fulfilled. So the asymptotic value of the oscillation period of a free swinging pendulum is constant , 112

asynchronous◁ (ongerelateerd in tijd, uit de maat)
of an event with respect to some process P: the event can occur at any time during the sequence of actions in P. Opposite of synchronous;
Example: an incoming telephone call will typically be asynchronous with whatever we are doing , 102

ATBD◁ (-)
artifact to be designed;
Example: a novel type of mobile phone that should double as electric razor , 16

atomic◁ (atomair, ondeelbaar)
-of a value or a type: does not consist of a (bundle) of multiple properties;
Example: a number, a boolean, or a string are examples of values that need no further information to be fully known , 65

audience◁ (omvang gebruikersgroep)
-as a criterion for modeling: abbreviation of 'number of intended stake holders', with the interpretation that a model, in respect of this criterion, is better if it can serve a larger group of problem owners;
(no example) , 253

average◁ (gemiddelde)
measure for central tendency of a set of data. See also mean;
Example: the average number of children per woman in the Netherlands is 1.76. This does not mean that most women have 1.76 child. 'Average' is obtained by adding the quantities in a set and dividing by the number of quantities in that set. 'Mean' is a property of a distribution: the set could, for instance, be a sample from a larger collection. The calculated average is then an estimation of the (arithmetic) mean. Indeed, there are different mathematical definitions of 'mean': arithmetic, geometric and harmonic means , 220

averaging-out◁ (uitmiddelen)
the effect that, in a sufficiently large ensemble of similar entities, individual variations can sometimes be ignored in comparison with average values;
Example: if we throw a fair die, the number of times we throw each of the possible outcomes 1,2,...,6 will all approach to 16.66...percent - even though they never will get exactly the same. Their variations get smaller with an increasing size of the ensemble (=the total number of repetitions of the experiment of throwing the dice) , 26

axis intercept◁ (asafsnijding)
-of the graph of a function: either the x-value(s) where y=f(x)=0, or the y-value that is attained when x = 0;
Example: the axis-intercepts of y = 2x-1 are x = 1/2, y = -1 , 157

behavior◁ (gedraging)
a route through a state space;
Example: Hamlet's part in Shakespeare's play is the only known route through this Danish prince's state space , 97

bias◁ (systematische fout)
the systematic deviation between measured or computed values and a correct value. See also systematic error, offset;
Example: somebody measuring her body mass by stepping on a scales with snoes on , 213

binary decision◁ (binaire beslissing)
decision with two possible outcomes;
Example: 'is this pathological anomaly yes or no malign?' Or: 'which of two materials performs better in this circumstance?' , 217

binding◁ (binding)

the relation between a value and a property, assuming that value;
Example: at the time of writing, the value '56' is bound to the property 'age' of the concept 'author'. In a few months, the bound value will be '57' , 95

binding◁ (binding)
the association of a value to a quantity (say, a property of a concept);
Example: the value '54', at the time of writing, is bound to the property age of the concept `authorOfThisText` , 31

black box◁ (zwarte doos)
a form of modeling where no claims about causal mechanisms are made; the model is obtained by compressing the observable information of some system;
Example: many models in biology, psychology, medicine and economics are black box models because the inner working of the modeled systems are too complex to represent , 29

Boolean◁ (Boolean)
the type with values `TRUE,FALSE`;
Example: the value of the expression 'it is currently raining' is `TRUE` or `FALSE`; therefore this expression (and every other proposition) has type Boolean , 68

boundary points◁ (randpunten)
a point (a, b) is a boundary point of a region R if *every* open disk centered at (a, b) contains points in R and points outside R;
(no example) , 281

bounded◁ (begrensd)
a region $R \subset \mathbb{R}^2$ is bounded if there is a disk that completely contains R;
(no example) , 281

bulk quantity◁ (bulk-eigenschappen)
a quantity of a system, consisting of many similar entities, where some form of aggregation applies;
Example: pressure and temperature of a given amount of gas , 26

calculate◁ (rekenen, berekenen)
obtain the resulting value from a formal expression by applying rules from arithmetic or calculus; opposed to reasoning;
Example: obtaining the volume of a rectangular box by multiplying its height, width and depth , 28

calculus◁ (calculus)
the part of mathematics involving functions, limits, differentials and integrals;
Example: a Taylor expansion of a sufficiently differentiable function is a device from calculus , 33

camel casing◁ (geen Nederlandse vertaling)
the habit of writing multiple-word names without spaces, where every word except the first starts with a capital;
Example: `myLongIdentifier` instead of `mylongidentifier` , 136

capacity◁ (capaciteit)
a quantitative measure for available effort, proportional to the amount of resources and the amount of time;
Example: capacity assumes homogenous tasks, where resources can be freely added, redistributed and exchanged. An example is capacity of human labor, expressed in man hours , 151

cartesian product◁ (cartesisch product)
collection of all combinations;
Example: for the sets with elements 1,2,3 and with elements a,b, the cartesian product is the set with elements [1,a], [1,b], [2,a], [2,b], [3,a], [3,b] , 179

category I◁ (categorie I)
quantities under control of the modeler; decision quantities;
Example: when designing a holiday: what destination to visit (assuming the destination is to be chosen)? , 178

category II◁ (categorie II)
quantities that are properties of the output of the functional model, also called (stakeholders') happiness quantities or objective quantities. Together they determine the merits of an ATBD in a design model;
Example: when designing a holiday: how

relaxing will the holiday be (assuming that we prefer a relaxing holiday)? , 178

category III◁ (categorie III)

quantities that are necessary to evaluate the functional model, but that are not influenced by any of the decisions of the modeler;

Example: when designing a holiday: between which dates will the holiday have to take place (assuming that we have to comply with school terms)? , 178

category IV◁ (categorie IV)

quantities that are necessary to evaluate the cat.-II quantities, and that are influenced by cat.-I, cat.-III and cat.-IV quantities;

Example: when designing a holiday: what will the weather be (assuming that the weather depends on the place we visit, so we influence the experienced weather by our choice of the destination; further, the weather will influence our pleasure)? , 178

causal loop◁ (circulaire afhankelijkheid)

circular dependency - a condition where something depends on its own present value; a condition that is forbidden in modeling functional behavior;

Example: in a bicyle trip: claiming that the duration depends on the distance we travel, which depends on the speed, whereas the speed we can maintain depends on the duration , 181

central limit theorem◁ (centrale limietstelling)

the sum of sufficiently many uncorrelated numbers with given center value and standard deviation has a normal distribution;

(no example) , 216

characteristic time◁ (karakteristieke tijd)

in a dynamical process, an amount of time needed to perform a typical (part) of the behavior;

Example: for a period dynamical process, the period is a characteristic time. For a behavior of exponential increase or decrease, the characteristic time is the amount of time needed for doubling or halving , 124

closed interval or disk◁ (gesloten interval of gebied)

an interval (or disk in two dimensions) is closed if it contains all its boundary points;

(no example) , 80

closed◁ (gesloten)

- of a set of values: it is possible to enumerate all values, either directly by listing them all, or indirectly by giving a finite recipe to generate them. Opposite to open;

Example: enumerating all values: 'the taste of this candy can be sweet, sour, salt or bitter'; generating the values: 'the shape of a cog wheel is a circle with two or more equal shaped indentations, placed at regular intervals on the perimeter' , 17

cluster◁ (klont)

a subset from a data set that can be described with a single black box model, whereas the entire data set cannot be described very well. Clustering or cluster analysis is the attempt to subdivide a data set into clusters such that much better black box models per cluster result;

Example: in physics and chemistry, various types of spectra are measured. A spectrum is essentially a histogram, the quantity plotted vertically being the occurrence of events with a certain energy, a certain wavelength, etc. Peaks in a spectrum correspond to clusters in the distribution of events , 227

coefficient◁ (coëfficiënt)

a quantity, often occurring as a factor that is multiplied with a variable;

Example: in the function $z = ax+by+c$, a is the coefficient for x, and b is the coefficient for y , 68

communication◁ (communicatie)

- as a purpose of models: a way to inform

some intended audience about what is modeled;
Example: a list of numbers, representing the outcome of an experiment can be a means to communicate this outcome to an interested reader , 17

complete◁ (volledig)
- of a set of options: including all possible outcomes;
Example: earth, water, fire, air is the complete set of concepts that can be obtained by combining values hot, cold for property temperature and values wet, dry for property humidity , 276
-of a taxonomy: every concept can be put in its unique location in the taxonomy. A taxonomy with only independent, enumerable properties with disjoint value sets is complete;
Example: the Greek periodic system (earth, fire, water, air) is complete with respect to a taxonomy with properties humidity withe values dry, wet and temperature with values hot, cold , 245

compound◁ (samengesteld)
- of a value or a type: consisting of (a bundle of) multiple properties, each with their own value;
Example: a vector, consisting of 2 (in 2D) or 3 (in 3D) coordinates is a compound value. A compound value is a concept in its own right. , 65

concept graph◁ (concept(en) graaf)
entity relation graph;
(no example) , 67

concept◁ (idee, voorstelling)
defined as a bundle of properties;
(no example) , 60

concept◁ (idee, voorstelling)
a mentally conceived or imagined entity, used in a model and representing some entity in the modeled system;
(no example) , 58

conceptualization◁ (conceptualisatie)
stage in the modeling process, comprising of building the conceptual model and choosing quantities;
(no example) , 30

conclusion◁ (conclusie, afronding)
stage in the modeling process, comprising of presenting and interpreting the result;
(no example) , 33

condition number◁ (conditiegetal)
for a function with respect to an argument: a measure for the relative sensitivity of that function for the argument;
Example: a large condition number value means that a small relative variation in the argument gives a large relative change in the output , 234

congruent◁ (gelijkvormig)
of two geometric figures: have the same shape, that is: one geometric figure can be mapped onto the other one using just rotation, scaling and translation;
Example: all circles are congruent; all equilateral triangles are congruent , 69

conservative◁ (behoudend, terughoudend)
with respect to the interpretation of a model outcome: tuning down the claimed distinctive power of a model outcome if the chance exists that problem owners might use the results unwisely;
Example: business models sometimes predict a return-on-investment. Such predictions can be dangerous, since optimistic estimates may cause recklessness in the (prospect) entrepreneur , 252

consistent◁ (samenhangend, kloppend)
such that no contradiction results ;
Example: a collection of statements (propositions) is contradictive if it is possible to deduce both a statement and its negation. See contradiction , 69

constant◁ (constante)
a quantity with a value that does not change;
Example: many so-called physical constants (e.g., the speed of light, the mass of an electron) are assumed to have an invariant value , 46

constraint◁ (beperking)
a limitation that applies to values for properties of a concept. See contingent;
Example: for the area and perimeter of a rectangle the constraint holds that the square of the perimeter is at least 8 times the area. Constraints can be (logically or mathematically) necessary, as in this example; they can also be contingent, for instance the constraint that something should not be heavier than X kg because otherwise it falls through the floor , 70

construct◁ (construct)
mental artifact, an abstract notion invented by man;
Example: mathematical objects, but also social notions such as 'marriage', 'possession', 'justice', ... , 277

context (of a problem)◁ (probleemcontext)
set of circumstances, events and conditions, not immediately part of a problem statement, that partially determine the success of the solution;
Example: for a model concerning the illumination of a motor way, the circumstance that the motor way is in an area with frequent fog is part of the problem context , 33

contingent◁ (contingent)
not logically or mathematically necessary. See constraint;
Example: as a mathematical necessity, for positive a, 3a is larger than 2a. In a shop, however, the fact that three products are more expensive than two is contingent: there may be a discount programme that sells 'three for the price of two' , 70

continuity◁ (continuïteit)
the property of a quantity to be able to assume all values between a minimum and maximum, without skipping even the tiniest hole;
Example: the speed of a material object can take a continuous set of values; the price of a good, to be payed in currency with a smallest coin can not take a continuous set of values , 21

continuous◁ (continu, aaneengesloten)
having the property of continuity;
(no example) , 21

continuous◁ (continu)
-of functions: $f(x)$ is continuous in $x = c$ if, for any $\epsilon > 0$, however small, we can find a $\delta > 0$ such that there is an x, $c - \delta < x < c + \delta$ with $f(c) - \epsilon < f(x) < f(c) + \epsilon$;
Example: $f(x) = 3x + 5$ is continuous in $x = 4$. Indeed, choose $\delta = \epsilon/3$, then $4 - \delta < x < 4 + \delta$ implies that $17 - \epsilon < 3x + 5 < 17 + \epsilon$, 80

contour plot◁ (contourenkaart)
visualisation of a function of two variables by means of one or more level curves;
Example: weather maps often show contour plots to indicate locations of equal temperature (isotherms) or locations of equal pressure (isobars) , 47

contour◁ (contour)
level curve;
(no example) , 47

contradiction◁ (tegenspraak)
see consistent;
Example: the statements `greaterThan (a,b)`, `greaterThan(b,c)`, and `greaterThan(c,a)` together are contradictive: the latter two can be combined to deduce `greaterThan(b,a)`, which conflicts the first statement , 69

convincingness◁ (overtuigendheid)
-as a criterion for modeling: the extent to which people tend to accept a model. A model is more convincing if is uses fewer non-obvious assumptions;
(no example) , 254

coordinate◁ (coördinaat)
quantity, used to distinguish spatial locations;
Example: in a cartesian system, coordinates are length, width, and height; in a polar system, coordinates are radial distance and angle; in a spherical system, coordinates are radial distance,

azimuth and elevation , 23

coplanar◁ (co-planair)
being confined to the same plane;
Example: the left and right rail in a railroad track in flat terrain are coplanar , 24

correlated◁ (gecorreleerd, verwant)
two sets of quantities are correlated if one can help predict the other;
Example: there is a correlation between the amount of alcohol a woman consumes during pregnancy and the birth weight of the baby , 26

correlation◁ (samenhang)
measure for the amount of dependency or independency between two quantities;
Example: the correlation between the amount of sunlight received by a plant and its height can be investigated by measuring the heights for a large number of plants that have been subject to varying amounts of sunlight , 220

correspondence◁ (correspondentie, overeenkomst)
-a relation between two things;
Example: an entity (something in the modeled system) corresponds to a concept (something in the model). This form of correspondence is also called representation , 58

counting◁ (tellen)
establishing a correspondence between distinct entities in a set and numbers 1,2,3 ... The number of entities is the largest number encountered;
Example: our number system is base ten because our ancestors used their fingers to make a correspondence between amounts of things and numbers. In French, 'quatre-vingt' for 80 reminds of a tradition where toes were used as well , 22

critical number, point◁ (kritieke waarde, punt)
an element of the domain of f is a critical number (one-dimensional domain) or a critical point (two or more-dimensional domain) if all derivatives are zero, or if some derivative is undefined;
Example: Both for the function $f(x,y) = x^2 + y^2$ and for the function $g(x,y) = 1/(x^2 + y^2)$, (0,0) is a critical point. For f, both partial derivatives in (0,0) are 0; for g, both partial derivatives in (0,0) are undefined , 80

DAG◁ (geen Nederlandse vertaling)
directed, a-cyclic graph;
(no example) , 137

data compression◁ (datacompressie)
- as a purpose of models: the representation of a body of information such that less space is needed;
Example: the information '0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30' can be compressed to 'positive even numbers less than 31' , 15

data mining◁ (data mining)
a kind of black box modeling where the form of the model typically follows from the data mined;
Example: using devices such as the Albert Heijn (AH) Bonuskaart (a discount card, issued by a large supermarket chain in the Netherlands), AH has access to a valuable amount of customer data, helping them to predict the demands for victuals and devising effective marketing campaigns - using advanced data mining technology , 224

data visualization◁ (datavisualisatie, informatievisualisatie)
a subdiscipline from communication science, focussing on the effective design of visuals to convey the meaning of data;
Example: pie charts, bar charts, 3D plots and more advanced forms of mapping data to images are part of data visualization , 250

database◁ (databank, informatieopslag)
structure to hold information, typically in the form of tables of mathematical objects, suitable for representation in a computer;
Example: the list of ingredients in a meal, together with a list of prices for

ingredients allows the calculation of the price for a meal , 28

deadlock◁ (impasse, patstelling)
in a dynamic process: a state that cannot lead to a following state;
Example: a computer system that is said to 'hang' often is in a state of deadlock , 101

decision◁ (beslissen)
- as a purpose of models: aid in taking a decision, either by optimization or by constraint satisfaction;
Example: the crucial steps in a design are the design decisions , 17

deduction◁ (afleiding)
logical inference in which the conclusion is of no greater generality than the premisse. See induction;
Example: if we observe 20 swans that are all white, we can deduce that it is possible that swans are white , 16

definition◁ (definitie)
- in modeling: first stage of a modeling process;
Example: the problem of 'how to illuminate a motor way' is defined more precisely by asking: 'is it possible to obtain sufficient illumination conditions with LEDs for this particular motor way' , 29

dependency by anticipation◁ (afhankelijkheid door vooruitdenken)
the situation where a value is assigned to a property x by a reasoning agent, such as a designer, to ensure that some other property y (depending on x) will assume a desired value. Then x depends on y by dependency-by- anticipation, whereas y depends on x by causal dependency;
Example: if an artifact should be portable, the choice for plastic instead of steel may be appropriate. It could be, however, that a model reveals that a steel solution exists that is lighter than a plastic one; dependency-by- anticipation would be an obstacle to finding such eventualities , 180

design space◁ (ontwerpruimte)
the cartesian product of all collections of values for cat.-I quantities;
Example: in a sandwich shop: the combination of all choices for the type of bread, butter or mayo, the added vegetables, and the toppings , 179

design, evolutionary◁ (evolutionair ontwerp)
a design methodology inspired by, or making use of, principles from biological evolution theory;
(no example) , 194

design, integral◁ (totaalontwerp)
(1) design where both functional, financial, ecological and user-centered considerations are taken into account in cat.-II; (2) design where simultaneously decisions are taken that involve many different aspects of the ATBD;
Example: some Apple products are considered examples of integral design because they address both technological innovation, new business models, aesthetical considerations and manufacturing policy , 198

design◁ ((technologisch) ontwerp)
(as opposed to research) the process of systematically taking decisions with the primary purpose of creating value for stake holders;
Example: the decisions leading to the realization of a machine, a material, an organization etc. , 13

desire◁ (voorkeur, preferentie)
desired property of an ATBD, a predicate over the ATBD or over a cat.-II quantity that is preferred to be `TRUE`;
Example: `doublesAsWalkingCane()` or `foldable()` for an umbrella , 183

deterministic◁ (bepaald)
involving only known steps and dependencies;
Example: the outcome of a die throw is not deterministic, whereas the outcome of throwing a quarter in a functioning coffee machine is deterministic , 25

differential equation, linear◁ (lineaire differentiaalvergelijking)
differential equation with the property that if f1 and f2 are solutions, any linear combination alpha1 f1 + alpha2 f2 is also a solution;
Example: equations for mechanical and electrical systems, Maxwell's equations for electromagnetism, and the Schrödinger equation from quantum mechanics , 287

differential equation◁ (differentiaalvergelijking)
a mathematical equation where the unknowns are functions rather than quantities, and where derivatives of the unknown functions occur;
Example: Newton's motion equation, F=ma is a differential equation since acceleration a is the second derivative of the location with respect to time , 118

dimension◁ (dimensie)
-in organizing information: property, or aspect that can help distinguishing individual items;
Example: 'gender', 'age' and 'educational level' are three possible dimensions in demography to distinguish individuals in a population , 18
an equivalence class, belonging to the equivalence relation 'has a constant ratio with' between units;
Example: length, time, force, energy, etc. , 77

direction◁ (richting)
that which two different parallel lines have in common;
Example: North, South, East and West are distinct directions, defined everywhere on the globe except on the North- and Southpole , 24

discrete◁ (discreet, telbaar)
-of a quantity: distinct; there is no smooth route to go from one value to another. Values of discrete quantities result from counting;
Example: states of a game of chess form a discrete set , 22

disjoint◁ (losstaand)
separated from something else, standing alone;
Example: the legs of a table, although connected by the table top, are disjoint entities , 57

dissipation◁ ((wrijvings)verlies)
loss of energy in a dynamic system, typically as a result of friction or damping;
Example: a dashpot in a mass-spring system increases the rate of energy loss in the system, converting kinetic energy into heat , 285

distance◁ (afstand)
a measure for the proximity of two items;
Example: in spatial coordinates, it can e.g. be expressed as the square root of the sum of squares of the difference between the coordinates of the respective items. In a broader context, it can be applied to non-spatial quantities as well , 24

distinctiveness◁ ((onder)scheidend vermogen)
-as a criterion for modeling: the extent to which the model is capable to tell two things apart;
Example: a model, say a function y=f(x) used to predict the value of a quantity y in dependence of x has higher distinctiveness if it is more accurate, that is: if it also can tell the difference between x being inside or outside the interval (11.99,12.01) rather than merely for the interval (11.5,12.5) , 257

distribution◁ (verdeling)
range of values, where the probability of finding a particular value from the range may differ for each value;
Example: a normal distribution and a uniform distribution , 216

domain convention◁ (domein conventie)
for a given function $f(x, y)$ the largest set of pairs (x, y) for which this function $f(x, y)$ can be evaluated, unless the domain is explicitly given by a smaller set;
(no example) , 45

domain◁ (domein)

-of a function: the set of values the argument of a function can be taken from;
Example: for the function 'age', everything that was ever born or created , 62

dominance◁ (dominantie)
one ATBD, say x, dominates another ATBD, say y, if, for all cat.-II quantities qi, x.qi is better than y.qi;
Example: for cat.-II quantities speed and comfort, a Rolls Royce dominates a wheelbarrow. For the triple (speed, comfort, price), however, this is not true , 191

dynamic◁ (dynamisch, tijdsafhankelijk)
involving time; opposite to static;
Example: the balance of forces that keep a dike from collapsing under wind and surf load , 21

edge◁ (boog, kant, zijde)
see arc;
(no example) , 30

emergent◁ ((onverwacht) verschijnend)
-of a phenomenon: something that appears as the result of some internal process;
Example: the behavior of an individual ant is limited and well-understood. Yet a colony of ants is capable of building structures like anthills, which can not readily be seen to result from combining individual ants' behaviors , 26

empirical◁ (empirisch)
based on observation - as opposed to 'found by reasoning' or 'resulting from a definition';
Example: results following from a laboratory experiment, a questionnaire, etc. , 14

EMS◁ (-)
see: empirical model system;
(no example) , 256

ensemble◁ (ensemble, collectief)
a collection of many entities that each behave stochastically, but similarly, so that the law of large numbers helps obtaining meaningful expectation values;
Example: the molecules of a confined amount of gas in thermal equilibrium , 25

entity-relation graph◁ (entiteit-relatiegraaf)
a graph where nodes are concepts, referring to entities, and arcs are relations;
Example: a city map, an electronic circuit , 67

entity◁ (entiteit)
anything, represented by a concept, of which information is represented in the model;
Example: entities can be material or immaterial, real or virtual , 30

entity◁ (entiteit)
something in the modeled system (and not in the model) that can be referred to; something that can be distinguished from another entity;
Example: pointing to something can be a way to distinguish it from other things, even if these have no names: 'this' as opposed to 'that' , 58

enumerable◁ (opsombaar, aftelbaar)
-of a property: the situation where all admitted values of a property can be enumerated;
Example: a property `aggregationPhase` is enumerable. It can only take one of the values `the aggregation states of matter (=gas, liquid, and solid) form an enumerable set of values` , 244

equality◁ (gelijkheid)
relation between two quantities stating that they have equal values;
Example: x=y and x=5 implies that y=5 , 83

equidistant◁ (gelijk verdeeld)
-of a series of values: having the same distance between any two subsequent values;
Example: the pearls in a pearl necklace , 23

equilibrium◁ (evenwicht)
- state: the state of a model such that, when slightly perturbed, it will try to get back to the initial state;

Example: a spring, when gently pulled and released, after a short while will assume its initial length , 37
equivalence class◃ (equivalentieklasse)
collection of things that are pairwise connected by an equivalence relation;
Example: 'has the same color as' is an equivalence relation; an equivalence class with this relation is the class of red things; another equivalence class is the class of green things , 280
error propagation◃ (foutenvoortplanting)
the way uncertainties in input quantities give rise to uncertainties in output of a functional model;
Example: in calculating properties of a rectangle with sides a and b, a relative error of 1 percent in one of the sides (say, a) causes a relative error of 1 percent in the area; the relative error in the perimeter is a / (a+b) percent , 232
error, model◃ (modelfout)
quality of the outcome of a model that relates to the deviation between a first model and a more complete model for the modeled system. To argue about error, there has to be a reference (glass box) model that is known to be more complete in terms of its incorporated mechanisms;
Example: ignoring the moon's gravitation is a permissable error for calibrating a pendulum clock; it is unacceptable when predicting high tide , 290
error, random◃ (willekeurige fout)
deviation between a measured or computed value and the correct value, due to nonsystematic effects. See also noise;
Example: repeated measurements, even with a bias-free instrument, used to its maximum precision, will produce nonidentical outcomes as every system warmer than 0 Kelvin at least possesses thermal noise , 213
error, systematic◃ (systematische fout)
shortcoming of a model or a measurement that causes results to be consistently too large or too small. See bias, offset;
(no example) , 213
event◃ (gebeurtenis)
the external cause for a transition in a state chart;
Example: the telephone rings, somebody insert a coin in a coffee machine, or a billiard ball collides with another billiard ball , 104
executable◃ (uitvoerbaar)
of a model: the state of developing a formal model where all unknown quantities have been expressed into either known or freely selectable quantities;
Example: all occurring values can be computed , 141
execution◃ (uitvoering, executie)
stage in the modeling process, comprising of operating the model, obtaining a result plus an estimate of the accuracy of the result;
Example: after composing a set of equations to represent the behavior of a modeled system, the execution stage amounts to the solution of these equations by mathematical or numerical means , 32
exhaustive◃ (volledig, uitputtend)
complete, taking all possibilities or all options into account;
Example: simple board games like tic-tac-toe can be exhaustively analyzed: all possible states can easily be enumerated. This is not feasible for chess, checkers and most card games , 36
expectation value◃ (verwachtingswaarde)
-of a distribution: the average value of a sufficient large number of samples taken from this distribution;
Example: the expectation value for the outcome of rolling a die is (1+2+3+4+5+6)/6=3.5 , 233
expert system◃ (expert systeem)
a system to represent expert knowledge in the form of a set of rules;
Example: a system to aid medical diagnosis could represent, for a number of pathological conditions, the observ-

able symptoms. Since symptoms do not match one-to-one to pathological conditions, an expert system needs the ability to reason with logical operations , 28

explanation◁ (uitleg)

- as a purpose of models: the association between two domains of knowledge, where this association to some communities may provide a sufficient answer to a 'why' question;
Example: Q: 'why does the temperature of a cold object increase when it is brought into contact with a hot object?' A: 'because there is a substance, called phlogiston, flowing from warm to cold bodies'. One domain is the (by now abandoned) Medieval theory of phlogiston, the other domain is an experience or observation from daily practice , 14

exploration◁ (verkenning)

- as a purpose of models: imposing a structure on an open domain to facilitate producing the elements of the domain;
Example: suppose the domain is 'planar shapes'. This domain is infinite (it contains circles, squares, star shapes, letters, ...), and there is no a priori manner to classify them. Exploring could be done by proposing properties on the domain (such as 'symmetry', 'size', 'curved or straight') that help the classification , 17

extremal◁ (extreem)

-of an element of the domain of a function: the property of being an extremum;
Example: For the function $f(x) = x^2$, $x = 0$ is extremal , 81

extremum◁ (extreem (punt))

an element in the domain of a function f where f assumes a local or global maximum or minimum;
Example: The function $f(x) = x^2$ assumes an extremum in $x = 0$ - which is a global minimum , 81

factor◁ (factor)

a quantity, often occurring in a product expression. See also term;
Example: 2, 3, and 5 are the prime factors of the number 60 because 2 x 2 x 3 x 5 = 60 , 68

false negative◁ (geen Nederlandse vertaling)

the situation where we should take action and omit to do so;
Example: overlooked medical symptoms , 260

false positive◁ (geen Nederlandse vertaling)

the situation where we take an action that should not have been taken;
Example: unnecessary medical treatment , 260

feasible region◁ (geen vertaling)

part of the domain of a function, that is to be optimized, where constraints are fulfilled;
Example: for the inequality constraints x¿0, y¿0 and x+y¡3, the feasible region is an isosceles triangle, aligned with the coordinate axes of the x-y plane , 84

Fermi, Enrico◁ (Enrico Fermi)

Italian / American Physicist (1901-1954), pioneer of nuclear physics and much acclaimed champion of modeling-by-estimating;
(no example) , 149

fitting◁ (aanpassen)

obtaining the value of a quantity in a formula by demanding that this formula adequately compresses a set of data, for instance in black box modeling;
Example: by fitting an exponential curve through a collection of radioactivity intensities, measured at regular time intervals, we can deduce the halftime of that radioactive material , 28

fitting◁ (aanpassen)

constructing a function y=f(x) that approximates a number of (xi, yi)-pairs;
Example: with the linear least squares method we look for a linear function f(x) = ax+b that minimizes the sum of distance between yi and a xi+b , 162

FMS◁ (-)
see: formal model system;
(no example) , 255

formal◁ (formeel)
expressed in terms of mathematical or logical formulas, or in terms of an algorithm;
(no example) , 27
not relying on human interpretation;
(no example) , 31
that which is defined within a logically consistent system, and does not require interpretation by human intelligence in order to be operated. Arithmetic is an example of a formal system;
Example: 'when John was five years younger than Suzy, Suzy was twice as old as John' can be formally expressed as y+5 = x and x = 2y; it follows that y = 5 and x = 10 which can be interpreted as the ages of John and Suzy at the time the riddle refers to , 31

formalization◁ (formalisering)
stage in the modeling process, comprising of obtaining values for quantities, and introducing mathematical relations between quantities;
Example: the translation from a wiring scheme (= an entity-relationship graph, hence a conceptual model) of two parallel resistors with values R1 and R2 to the formula
R = R1 R2 /(R1+R2) , 31

formula◁ (formule)
formal relation between quantities;
Example: F=ma for the relationship between a force F, and the acceleration a caused by that force on a point with mass m , 31

function of two variables◁ (functie van twee variabelen)
the function $f(x, y)$ is a rule that assigns a real number $f(x, y)$ to each ordered pair of real numbers (x, y) in the domain $D \subset \mathbb{R}^2$ of the function;
(no example) , 45

Function Selector◁ (functiekiezer)
interactive tool for finding a mathematical function with a graph that has certain desired shape features, such as asymptotes, convexity or concavity, etc.;
(no example) , 142

function, logistic◁ (logistische functie)
a continuous function that approximates a linear behavior in a limited part of the domain, but that saturates for larger and smaller argument values;
Example: the simplest form is parameterized as f(x) = 1 / (1+exp(-x)). It takes the form of a stretched s-shape , 159

function, penalty◁ (boetefunctie)
a function that expresses how good something is by returning a non-negative value which is 0 if the evaluated thing is optimal; the further it deviates from optimal, the larger its value;
Example: for approximating a sequence of data points [[x0,y0], [x1,y1], ...] by a function f, a penalty function could be the residue, that is the sum, over all i, of squares of yi-f(xi) , 188

function, ramp◁ (helling)
a continuous function consisting of two piecewise constant segments connected by a linear segment;
Example: used to model price elasticity , 159

function, rational◁ (rationale functie)
a function that can be written as the quotient of two polynomials. Every function that can be computed with no more a finite number of additions, subtractions, multiplications and divisions of the argument value can be written as rational function;
Example: the function that gives the focal length f for a lens such that a point as distance v is sharply projected at a screen at distance b from the lens: f = b v / (b+v); notice that this function is usually written as the relation 1/f = 1/b + 1/v , 158

function, recursive◁ (recursieve functie)

a function where the calculation of the return value needs the evaluation of the same function on another argument value;
Example: a function that computes the sum of elements from a list, f(a1,a2,a3,...) may be defined as the first element of the list, a1, plus the same function applied to the remainder of the list; the sum over an empty list is zero. Recursive functions are convenient to represent quantities that are time dependent. For instance f(x,current)= x,previous+d express that the current value of quantity x depends on the previous value of the same quantity. If d is constant, this function describes a uniformly incrementing behavior , 107

function, separable◃ (separeerbare functie)
a function of multiple quantities that can be written as a product of functions, each of one of the quantities;
Example: the Boyle-Gay Lussac-Avogadro law, P=nRT/V is separable , 288

function, sigmoid◃ (sigmoide functie)
other name for logistic function;
(no example) , 159

function◃ (functie)
a prescription to produce a uniquely determined value, given a value;
Example: mathematical functions such as square, square root, etc., are familiar examples. But properties are also functions of the concept they belong to: the age of a person is a function of that person , 61

Galton, Francis◃ (Francis Galton)
English Scientist and polymath (1822-1911), pioneer in statistical methods and the inventor of the method of the wisdom of the crowds;
(no example) , 143

genericity◃ (algemeenheid, brede inzetbaarheid)
-as a criterion for modeling: the different kinds of modeled systems for which a model can be used;
Example: an evolutionary algorithm is a generic approach to optimization, as it can deal with continuous and discrete, linear and non-linear problems, and single- and multiple objective functions , 246

genotype◃ (genotype)
term from biological evolution: the collection of all inherited features in the form of 'building instructions' (DNA encoded genes);
Example: the genome for an individual human , 195

geometry◃ (meetkunde)
the part of mathematics studying the relations between mathematical objects, stemming from formalizing intuitions related to our perception of space;
Example: Pythagoras theorem is a result that can be proven by geometric means , 24

glass box◃ (glazen doos)
a form of modeling based on (assumed) causal mechanisms; the model is obtained by representing the causal relations by mathematical expressions;
Example: many models in chemistry and mechanical engineering are glass box models involving reaction mechanisms or laws of physics , 29

global maximum, absolute◃ (globaal maximum)
a value $f(c)$ (or $f(a,b)$ in two dimensions) is a global maximum on the region R if it is larger than (or equal to) all the other function values in the region R;
(no example) , 82

global minimum, absolute◃ (globaal minimum)
a value $f(c)$ (or $f(a,b)$ in two dimensions) is a global minimum on the region R if it is smaller than (or equal to) all the other function values in the region R;
(no example) , 82

graph (entity-relation)◃ (entiteiten-en-relaties graaf)
a diagram where nodes correspond to entities and (directed) arcs correspond to named relations;

Example: an electronic circuit, a structural formula in chemistry, or an annotated city map , 30

graph◁ (graaf, netwerk)
a drawing consisting of nodes and arcs;
Example: the London Underground map and the maps used by the Dutch Railways (NS) are graph representation of hundreds of kilometers of rail connections, stations and junctions , 30

grey box◁ (grijze doos)
- of a model: mix between a black box and a white (glass) box model;
Example: a model for propelling a ship using the theory of fluid flow (glass box), where coefficients for the friction between water and the ship hull are taken from measurements in a water tank (black box) , 29

ground truth◁ (vaststaand feit)
data that can be used to verify if the behavior of a model is consistent with our knowledge of the modeled system;
Example: a model for weather predictions can be run to predict yesterday's weather, using earlier weather data as input. The actual observations of yesterday's weather form ground truth, to compare the 'predictions' with , 36

heuristic◁ (heuristiek)
a non-proven method to obtain something;
Example: to make most effective use of the loading space in a van, it is a good idea to start with loading the bigger objects , 135

hypothesis◁ (hypothese, veronderstelling, vermoeden)
a postulated proposition or relation, that is assumed to be provisionally true, but that is subject to sceptic testing;
Example: a hypothesis in social science could be: 'frustration causes aggression'. In material science, a hypothesis could be that adding material X to a substance Y increases the melting temperature, etc. , 32

hypothetical◁ (hypothetisch, verondersteld)
imaginary;
Example: free electrons are hypothetical particles, transporting electricity through a conductor , 22

identity◁ (identiteit, eigenheid)
that which allows distinguishing one thing from other things;
Example: the identity of a Dutch citizen is reflected in his or her (unique) passport number; the chemical identity of an element from the periodic table is reflected in its atomic number , 58

impact◁ (invloed, effect)
-as a criterion for modeling: the consequence, expressed in quantities belonging to the model context, of using the model;
Example: macro-economical models as consulted by politicians, supporting (or warning against) interventions in transnational financial issues can have long-lasting impact on a global scale , 262

IMPLIES◁ (impliceert, heeft als gevolg)
logical operator: P IMPLIES Q is true if Q is true or P is false;
Example: rain IMPLIES a wet ground; this means that either the ground is wet OR it is not raining , 28

independent◁ (onafhankelijk)
-of two quantities: one cannot be deduced from, or correlated with, the other;
Example: the yearly number of sunny days in Tokio and the yearly sales of bikinis in London are two independent functions of time , 69

index◁ (index)
number to be written between [and], used to single out one element from an aggregation;
Example: For the aggregation p=[2,5,7], p[2] denotes 7, and the index is 2 , 61

index◁ (index)
additional information, added to data to speed up queries on the data. Modern digital search systems such as Google heavily relay on indexing;

Example: an alphabetically sorted list of keywords , 248

individual◃ (individu, ondeelbare entiteit)
literally: that which cannot be divided;
Example: a human being in a population, or a molecule in a gas or liquid , 58

induction◃ (generalisatie)
a form of reasoning where an attempt is made to arrive at general conclusions from limited premisses. See deduction;
Example: if we observe 20 swans that are all white, we may be tempted to state that all swans are white - which is not necessarily true , 16

inequality◃ (ongelijkheid)
relation between two ordinal quantities stating that one is larger than the other;
Example: for two ranks in the army, p and q, p¿q and p=captain, q has a lower rank than captain. An inequality can be expressed with the relation ¡ or with the relation ¿: a¡b is equivalent to b¿a , 83

inference◃ (afleiding)
a formal operation with logic quantities and logic rules. See deduction;
Example: let P represent the proposition 'all balls are spherical', and Q the proposition 'a football is a ball', then the conclusion 'a football is spherical' can be drawn purely by using the logical structure of the propositions. If propositions are written down with sufficient precision, dedicated computer languages are capable to perform such deductive inference , 33

initial value◃ (beginwaarde)
a value to be provided for a quantity in a model for a dynamic system in order to calculate subsequent states;
Example: for a billiard shot, the velocity of the bal that was hit by the cue. For a system representing financial transactions: the initial amounts on the involved accounts , 109

injective◃ (injectief, 'in' (voor functies))
of a function: two different argument values yield two different result values;
Example: mapping a telephone number to a subscriber should be an injection: no two subscribers should have the same telephone number. Mapping keys to locks does not have to be an injection, as no one will exhaustively try all locks in the world with a found key , 191

input◃ (invoer)
-related to a functional model: referring to the designer's choices (cat.-I) and context quantities (cat.-III);
(no example) , 294

input◃ (invoer)
-related to the modeling process: referring to stage 1 (model definition);
(no example) , 294

integer◃ (geheel (getal))
zero, or the successor of an integer;
Example: 3, 17, 888895, -4 , 22

interpolation◃ (interpolatie)
obtaining a value y, depending on a quantity t, from values y0 and y1, occurring for t-values t0 and t1, where t is between t0 and t1, as y = y0 + (y1-y0)(t-t0)/(t1-t0);
Example: John's length at age 12 was approximately the average of his length at ages 10 and 14, assuming growth rate to be constant in the period between 10 and 14 years , 22

interpretation◃ (duiding)
-as stage in the modeling process: the formulation of an answer to the initial problem in terms of the problem domain, rather than in terms of the model domain;
Example: imagine a model for the concentration of some medicine in the blood flow. The outcome of the model could be a table consisting of concentrations as a function of time. An interpretation could be: 'take two pills before breakfast and another pill just after lunch to have the fewest unwanted side

effects' , 33
invariant◁ (invariant)
something that stays the same when circumstances, such as measurements, change;
Example: in measuring: any physical quantity stays the same if the laboratory moves with uniform velocity. Also: when measuring length l with a unit of length u, the value (l/u)u is invariant. Example not related to measuring: the ratio between perimeter and diameter of a circle stays the same if the circle is enlarged or reduced , 75
iso-(value) curve◁ (curve van gelijke waarden van een functie)
level curve;
(no example) , 47
iterate◁ (herhalen)
-of the modeling process: repeatedly go through the subsequent process stages, e.g. because of increasing understanding of the model's purpose;
Example: to obtain an increasingly better approximation of the square root of some S, one should start with an arbitrary positive number x and repeatedly replace x by (x+S/x)/2 , 35
iterative◁ (herhaaldelijk)
based on iteration;
Example: counting beans is an iterative process, where repeatedly we set apart one bean and increase a number N by one. When there are no more beans, N, started with zero, equals the number of beans , 23

knowledge base◁ (kennissysteem)
a database, especially designed to hold mathematical objects that represent propositions;
Example: a database with medical symptoms together with a database containing rules to link symptoms to diseases , 28

label◁ (label, naam, aanduiding)
- in a graph: a name or identity given to a node or an arc. In particular if arcs in a graph can have different meanings, these should be labeled;
Example: in a geographical map, the city names printed near dots representing cities are labels , 30
leaf◁ (blad)
in a directed a-cyclic graph: a node without outgoing arrows;
Example: in the four categories-model, cat.-I and cat.-III quantities are the leaves with respect to the `dependsOn`-relation; with respect to the `isArgumentTo`-relation, the cat.-II quantities are leaves , 137
length◁ (lengte)
- of a curve: the number of sufficiently short line segments of unit length, needed to approximate the shape of that curve;
Example: the length of a curve is found by having a hypothetical rope follow the curve; next stretching the rope and measuring the distance between its two end points , 24
level curve◁ (hoogtelijn)
a collection of points where a function of two variables takes the same values;
Example: if the function represents a temperature distribution over an area, the level curves are isotherms. Other examples are isobars (equal pressure) or iso-potential curves (equal electric potential) , 47
lifelock◁ (geen Nederlandse vertaling)
in a dynamic process, lifelock is a limited collection of states where no transitions exist that lead out of this collection;
Example: two polite people in front of a narrow passage, wanting to grant each other precedence ('after you-after you blocking') , 101
line◁ (lijn)
straight curve;
Example: the trajectory of a beam of light in a space with constant index of

refraction is a straight line , 24
linear approximation◁ (lineaire benadering)
approximation of a function, in the neigborhood of a given point of the domain, by a linear function;
Example: the linear approximation of the function $f(x, y)$ at the point (a, b) is defined as $L(x, y) = f_x(a, b)(x-a) + f_y(a, b)(y - b) + f(a, b)$, 49
linear least squares◁ ((lineaire) kleinste kwadratenmethode)
a method for finding the linear function f that best approximates a collection of data points (xi,yi) by minimizing the sum of the squares of yi-f(xi);
Example: the build-in trend lines offered by MS Excel graphs are computed using linear least squares , 224
linear◁ (lineair)
(1) to be described or approximated by an expression y = ax+b where x is an independent quantity and y is a dependent quantity; (2) the property, of a dependency y = f(x), that f(x1+x2) = f(x1)+f(x2) and f(sx) = sf(x), for all x, x1, x2 and real s;
Example: example of 1: the position of a point with constant velocity is a linear function of time; example of 2: the amount of heat, produced by burning gas, is linearly dependent of the amount of burned gas. See also superposition , 22
linearization◁ (linearisatie)
a technique to approximately solve non-linear differential equations where the non-linear term is sufficiently small compared to linear terms;
(no example) , 285
local maximum, relative◁ (lokaal maximum)
a value $f(c)$ (or $f(a, b)$ in two dimensions) is a local maximum of f if it is larger than (or equal to) all the other function values in some open interval (or disk in two dimensions) containing c (or (a, b)) ;
(no example) , 80
local minimum, relative◁ (lokaal minimum)
a value $f(c)$ (or $f(a, b)$ in two dimensions) is a local minimum of f if it is smaller than (or equal to) all the other function values in some open interval (or disk in two dimensions) containing c (or (a, b)) ;
(no example) , 80
logistic growth◁ (logistische groei)
growth in the presence of limited resources;
Example: If bacteria would divide every minute, their amount would double every minute - growing beyond any bound. In practice, their growth rate decreases, among other things because of limited food, leading to a constant colony size , 125
lumping◁ (samenvoegen, klonteren)
the replacement of numerous quantities, each related to an individual item in a system by few quantities that apply to the system as a whole;
Example: molecules in a gas, citizen in a community, cars in a flow of traffic , 26
lumping◁ (samenballen)
-of cat.-II quantities: combining the effect of several cat.-II quantities into fewer ones by adding penalty functions, and postulating their relative weights;
Example: expressing the impact on environment of an individual or an organization using a so called 'ecological footprint': an equivalent to an amount of earth surface, needed to sustain the individual or the organization in terms of biomass, minerals and fossil energy , 190

machine precision◁ (machinenauwkeurigheid)
on a computer, a number is only represented with finite precision. Numbers that differ less than machine precision cannot be distinguished;
Example: on a 64 bit machine, 53 bits are typically used for the mantissa of floating point numbers, giving a ma-

chine precision of 2 to the power -52, approximately 2.2 x 10 to the power -16 , 293

macro-irreversible◁ (macro onomkeerbaar)
the property of physical transformations that, when many degrees of freedom are involved, time order is not symmetric;
Example: friction always slows moving things down, turning motion energy into heat; a melting ice cube will turn in a little puddle of water, but a freezing, unconstrained puddle of water does not assume a cube shape , 93

mathematical object◁ (wiskundig object)
a concept used in mathematical reasoning;
Example: numbers, vectors, functions, matrices, geometric shapes , 24

mean, arithmetic◁ (rekenkundig gemiddelde)
sum of the values in a set divided by the number of values, i.e. the average of the elements in a set;
Example: as a result of recent reforms of pension funds in the Netherlands, pensions will be based on the arithmetic average of the salary over the working period, rather than the salary at the time of retiring , 221

mean, geometric◁ (meetkundig gemiddelde)
n th-root of the product of the values in a set where n is the number of values;
Example: to calculate the volume of a truncated pyramid with height h and bottom- and top areas of B1 and B2, respectively, we need the geometric average of the bottom and top surfaces, the square root of B1 B2, as follows: V = h(B1+B2+sqrt(B1 B2))/3 , 221

mean, harmonic◁ (harmonisch gemiddelde)
the inverse of the sum of the inverses of the values in a set divided by the number of values;
Example: used to calculate the mean speed over the sum of a number of equal distances, each with given speed , 292

median◁ (mediaan)
for an ordinal set: the largest value x such that the number of occurring values less than x is not larger than the number of occurring values larger than x;
Example: consider 5 sticks of lengths 1, 2, 5, 10, and 99 cm. Then 5 is the median value , 72

message◁ (boodschap)
finite amount of information, to be conveyed from a sender to a receiver;
Example: 'if you can read this you are driving too near by' , 275

micro-reversible◁ (micro omkeerbaar)
the property of physical interactions that, when few degrees of freedom are involved, time order is symmetric;
Example: the collision between two ideal rigid spheres, the compression of an ideal spring , 93

model refinement◁ (model verfijning)
replacing (part of) a model by a more sophisticated version in order to better fulfill a purpose;
Example: in a geometric system, a spherical mirror can be first be assumed to reflect a parallel beam of light to one that converges through the focal point. This is only true for sufficiently narrow beams; model refinement may involve the more elaborate calculation of the actual shape of the reflected beam , 32

model system, empirical◁ (empirisch model, (w.o. schaalmodel, fysiek model))
a second 'real' (that is: not conceptual) system, e.g. a physical system or a social system, that is believed to behave sufficiently similar to a system to be modeled, so that conclusions about the system to be modeled can be derived from this second system;
Example: scale models, such as mock-ups, wind tunnel models, or historical data , 256

model system, formal◁ (formeel modelsysteem)
model, where elements of the system to be modeled are replaced by simpler elements, chosen such that (physical, economical, ...) laws exists that apply to

these elements;
Example: for meteorological purposes, our planet may be replaced by a perfect sphere with constant density equal to the average density of the Earth , 255

model, conceptual◁ (conceptueel model)
version of the model that comprises of concepts, their properties and relations between them, but not yet formal mathematical constructs. See model, formal;
Example: a block scheme for a chemical reactor is an example of an entity relationship model for that reactor , 30

model, first principles◁ (geen Nederlandse vertaling)
a model that is based on the most fundamental relations, known in a discipline;
Example: in physical chemistry, a quantum mechanical calculation of the binding energy in a molecule is a first-principles approach, whereas a quasi-classical calculation is not , 255

model, formal◁ (formeel model)
version of the model that comprises of quantities and formal relations between them. See model, conceptual;
Example: the unknown currents flowing through a network of resistors, written as a set of linear equations , 31

model, formal◁ (formeel model)
model where mathematics, logic, or computer science plays a crucial role;
Example: next to performing experiments and the development of theories, doing simulations begins to be a third direction in various scientific disciplines. Simulations are an example of formal models , 20

model, immaterial◁ (model, immaterieel)
model, only involving information and information carriers (e.g., paper and ink, computer memory), and no other material objects;
Example: a mathematical model (e.g., functions and equations), a logical model (e.g., propositions and rules for deduction), or a software model (e.g., a simulation) , 20

model, informal◁ (model, informeel)
model where arguments and reasoning are mainly stated in natural language;
Example: Einstein's famous thought experiment that lead to the idea of special relativity started with the question 'what would happen if someone could travel on the front of a light ray'. Although some elementary mathematics is required to find the basic formula of special relativity, most part of the reasoning is informal , 276

model, material◁ (model, materieel)
model involving a material object;
Example: a scaled-down aircraft used in a wind tunnel to estimate aerodynamical properties of the original, full-scale version; guinea pig used in testing medicine to predict the reactions in humans to that medicine , 20

model, Monte Carlo◁ (Monte Carlo model)
kind of model where properties of a complex system are approximated by repeating stochastic calculations;
Example: the pressure of a gas, in dependence of temperature T, could be found from a simulation using sufficiently many elastically colliding point masses with well-defined initial kinetic energy, proportional to T. The name comes from the association between Monte Carlo and casino, hence stochastic processes , 26

model, purposes◁ (model)
-in relation to purpose. See purpose;
(no example) , 13

model◁ (model)
-definition: the (mental) construct resulting from going through the modeling process, to help fulfilling a purpose;
(no example) , 29

model◁ (model)
-steps common to all models;
Example: collecting data, performing mathematical operation(s), interpreting

the outcome of the mathematical operation(s) , 12

-various dimensions of models;

Example: static vs. dynamic, continuous vs. sampled vs. discrete, numerical vs. symbolic, geometric vs. non-geometric, deterministic vs. stochastic, calculating vs. reasoning, glass box vs. black box , 18

model◁ (model)

-as the result of a process;

Example: the stages in the modeling process are: definition, conceptualization, formalization, execution, and conclusion , 29

model◁ (model)

- as a means to achieve a purpose;

Example: doing a computer calculation on atmospheric data to predict the chance that tomorrow will bring rain , 11

modeled system◁ (gemodelleerd systeem)

the existing or hypothetical system to which a model relates;

Example: if the model is a geographic map, the modeled system is the depicted area on the surface of the Earth. If the model is a drawing of an electric circuit, consisting of rectangles and circles connected by lines, the modeled system is the physical assembly of resistors, capacitors and transistors to which this drawing relates. A modeled system is either a part of existing reality for models with a research purpose; it is part of not-yet existing, possible future reality for models with a design purpose , 29

monotonic◁ (monotoon)

see monotonous;

(no example) , 156

monotonous◁ (monotoon)

-of a function: literally, 'everywhere the same (behavior)';

Example: a monotonously decreasing function, such as f(x) = exp(-x) is decreasing everywhere in the domain of the function , 156

morphological box◁ (geen Nederlandse vertaling)

example of a taxonomy, used to give a systematic exploration of a number of varieties of some item;

Example: morphological boxes are often used in shape studies, e.g. in architecture or geometric design , 258

MTF◁ ()

Mean Time Between Failures: the average amount of time that is expected to pass between two subsequent instance of failure;

Example: If the MTF of some engineered component is sufficiently long compared to its economic lifetime, it may be considered to be engineered well , 229

multiple regression◁ (meervoudige regressie)

a branch of statistics seeking to describe data sets with functions of multiple arguments;

Example: the life expectancy of a population may be correlated to smoking, but also to diet patterns and social circumstances , 164

multiplicative◁ (multiplicatief)

in general: regarding multiplication or proportionality;

Example: for condition number calculation: condition numbers are constant for multiplicative expressions , 293

node◁ (knoop, punt)

an element in a graph, next to arc;

Example: in a diagram consisting of boxes and arrows, the boxes are the nodes and the arrows are the arcs , 30

noise◁ (ruis, (willekeurige) spreiding)

a random contribution to a signal or a measured value which causes repeated samples to differ. See also random error, statistical spreading;

Example: the larger the signal-to-noise ratio, the smaller the relative differences between samples compared to their av-

erage , 213
nominal◁ (nominaal)
of a set: a set that has no ordering;
Example: a set of countries, a set of plant species, a set of car brands, ... , 71
of quantities: the property that they cannot be ordered;
Example: materials, nationalities and tastes are nominal , 71
non-rational◁ (niet-rationaal)
-of a function: the function cannot be written as a quotient of two polynomial functions;
Example: f(x) = sin(x) , 156
normal distribution◁ (Gaussische verdeling, normale verdeling)
Gaussian distribution;
Example: the statistical distribution of a quantity that is the sum of sufficiently many, independent quantities. For instance, the total score obtained by throwing with N fair dice approximates a Gaussian distribution when N increases , 216
normal◁ (normaal, Gaussisch)
of a distribution: a distribution, described by a Gaussian, featuring a center value and a quantity to indicate the width;
(no example) , 216
numerical◁ (numeriek)
operating on numbers instead of non-numerical symbols. Opposite to 'symbolic' or 'analytic';
Example: estimating the maximum of a function y = f(x) on an x-interval by repeated evaluation in a series of closely spaced x-values, recording when the largest y value is obtained , 23

objective◁ (objectief)
does not depend on an individual observer, or on an individual opinion. Opposite of subjective;
Example: the viscosity of marshmallows is lower than the viscosity of caramel candy , 15
offset◁ (systematische fout)
See bias, systematic error;
(no example) , 213
open disk◁ (open cirkelschijf)
the interior of a circle (i.e. all points inside in the circle but not on the circle;
(no example) , 80
open◁ (open)
- of a set of values: it is not possible to enumerate all values. Opposite to closed;
Example: 'the shape of this hole can be round or square or heart-shaped or something else ...' , 17
operation - mathematical◁ (bewerking, wiskundige)
processing of mathematical objects, such as numbers or functions, using mathematical or logical operators (such as add, subtract, differentiate, ...);
Example: a(b+c) = ab+ac holds for arbitrary numbers a, b and c , 12
operation◁ (operatie, bewerking)
activity, such as evaluation, optimization, solving equations, numerical approximation, etc., to be performed with a model;
Example: a weather model is run with a set of empirical weather data as input to predict tomorrow's weather , 32
optimization◁ (optimalisatie)
- as purpose of a model: finding an answer to the question 'for which value(s) of quantity x is the resulting value of quantity y, depending on x, as good as possible', where 'good' needs to be specified ('large', 'small', ...);
Example: finding the shape of a tank such that the volume is as large as possible with a given wall surface area , 17
OR◁ (of)
logical operator: P OR Q is true if at least one of the two is true;
Example: precipitation is the weather condition where either rain OR snow falls from the sky , 28
ordening - partial◁ (partiële ordening)
a relation between elements from a set that is transitive, and that introduces no cy-

cles;
Example: the relation `ancestorOf` for the collection of human beings , 71
ordening - total◁ (totale ordening, totale volgorde)
a relation between elements from a set that is transitive, anti-symmetric and total. 'Total' means that for any two different elements one exceeds the other;
Example: the relation `greaterThan()` for the collection of numbers , 71
order◁ (orde)
for the model of a dynamical system. This is the number of earlier states needed to evaluate the current state;
Example: for a system to represent financial transactions, the order is 1: to obtain a new state for a budget, we need the previous value and the amount of money transferred , 109
ordinal◁ (ordinaal, ordenbaar)
of quantities: the property that they can be ordered;
Example: numbers are ordinal , 71
orthogonal◁ (onafhankelijk, loodrecht)
for vectors or directions: being perpendicular to. In general: independent;
Example: in state charts: the behaviors of non-communicating systems, where every state in one system can occur in every state of the other system and vice versa , 101
output◁ (uitvoer)
-related to a functional model: referring to the stakeholders' values (cat.-II);
(no example) , 295
output◁ (uitvoer)
-related to the modeling process: referring to stage 5 (conclusion);
(no example) , 295

parallel◁ (evenwijdig)
the property that a pair of two coplanar straight lines do not intersect;
Example: the velocity vectors in two points of a moving, but non-rotating rigid object are parallel , 24
parameter◁ (parameter)
a quantity, sometimes known and sometimes unknown, that occurs in a function or other expression;
Example: the parameter representation of a line between points p1 and p2 is p1 + lambda(p2-p1), lambda is a parameter , 68
Pareto optimisation◁ (Pareto optimalisatie)
a technique for approximating the theoretical Pareto front for a given functional model;
Example: finding the non-dominated taxi companies using the techniques of Chapter 5 , 193
Pareto, Vilfredo◁ (Vilfredo Pareto)
Italian economist and philosopher (1848-1923);
(no example) , 193
Pareto-front◁ (Pareto-front)
the collection of non-dominated solutions in cat.-II space;
Example: the collection of taxi companies, from the model of Appendix **??**, for which no company exists that is both more profitable and more fun , 193
partial derivative◁ (partiële afgeleide)
Informal: the partial derivative of the function $f(x, y)$ with respect to x is the ordinary derivative, while treating y as a constant;
(no example) , 48
penalty function◁ (boetefunctie)
see function, penalty;
(no example) , 188
performance◁ (prestatie)
a measure for the quantitative effectiveness of a procedure. Performance can relate to time (faster is better) or memory requirements (requiring less space is better);
(no example) , 247
periodic◁ (periodiek, regelmatig herhalend)
of a process: repeating itself after a constant time lapse;
Example: a pendulum, or a bookkeeping system where each 1st of Jan-

uary the previous book-year is closed , 110
perpendicular◃ (loodrecht)
the maximal difference between two directions;
Example: North and East are perpendicular directions everywhere on the globe (except on the North pole and South pole) , 24
phenotype◃ (phenotype)
term from biological evolution: the collection of all manifest features of an individual, together determining its fitness;
Example: the long neck of the giraffe is part of its phenotype: it is a feature that makes it fit in an environment where leaves are nutritious but trees are tall , 194
polynomial◃ (polynoom, veelterm)
-of a function: the function can be written as a sum of terms, where every term is a constant coefficient multiplied with an positive integer power of the argument;
Example: f(x) = (x-3)(x+4) , 156
population◃ (populatie)
collection of individuals, in the context of SPEA: the collection of candidate solutions;
Example: when designing an X: the collection of candidate X-s considered in the design space , 195
POset◃ (partieel geordende verzameling)
partially ordered set;
Example: the collection of intervals is partially ordered under the relation `greaterThan()` , 71
possible, logically◃ (logisch mogelijk)
of a concept: the values of its properties do not contradict;
Example: Let us suppose that `big` and `small` are mutually exclusive values of the property `size`, and that `open` and `closed` are mutually exclusive values of the property `top`. Therefore, a big open box, a small closed box etc. are logically possible, whereas a big small box or a big open closed box are logically impossible , 69
postulate◃ (postuleren)
verb: formulate as a working hypothesis; substantive: a working hypothesis;
Example: the truth of the postulate can not be proven, but it is assumed true until evidently shown false. Unlike (normal) hypotheses, postulates are sometimes not subject to deliberate attempts of falsification. E.g., any formula to describe physical phenomena should be independent of the speed of the laboratorium where the phenomenon occurs , 29
precision◃ (precisie)
quality of the outcome of a model that does not relate to the deviation between the model and the modeled system. Compare accuracy;
Example: repeating a measurement 1000 times instead of 10 times gives a factor of 10 smaller standard deviation (assuming no systematic errors), that is: one more digit of precision , 213
predicate◃ (predikaat)
a proposition depending on some argument, or: a function with {`TRUE`,`FALSE`} as range;
Example: `isGreen(x)` is a predicate, which is `TRUE` if we substitute for `x` the concept `cucumber`, `FALSE` if we substitute `canary` , 182
prediction time◃ (voorspellingstijd)
the time period for which we want to, (or: are able to) obtain a valid prediction;
Example: for the weather, there are no known methods to obtain a prediction time longer than 5 or 6 days , 124
prediction, 1st kind◃ (voorspelling van de 1e soort)
unconditional prediction;
(no example) , 15
prediction, 2nd kind◃ (voorspelling van de 2e soort)
conditional prediction;
(no example) , 15
prediction, conditional◃ (voorwaardelijke voor-

spelling)
a statement about something that is going to happen provided that some condition is fulfilled, where this condition may or not may be under somebody's control;
Example: If I work hard enough, I will get I high grade for my exam on Modeling , 15

prediction, unconditional◁ (onvoorwaardelijke voorspelling)
a statement about something that is going to happen without the possibility to influence on the course of events;
Example: the weather forecast, as long as there is no technology to influence the weather, is an unconditional prediction. Also, under fair and legal circumstances, predictions of stock exchange rates are also unconditional , 15

prediction◁ (voorspelling)
- as a purpose of models: (1) a statement that at a given time point in the future something will happen, or (2) a statement that when something (not seen before) will happen, a certain property will be observed;
Example: (1) the next lunar eclipse will occur April 15, 2014 (this text is written December, 2012); (2) if we continue to consume fossile fuel in the current rate, sea levels will rise , 14

preposition◁ (voorzetsel)
word used to indicate mainly spatial or temporal relations;
Example: 'near', 'above', 'behind', 'before', 'during', ... 'Notwithstanding', although not referring to a spatial or temporal relation, is also a preposition , 62

presentation◁ (vertolking)
casting the result of formal operation with a model in a form that can be more easily understood in the context of the initial purpose;
Example: a model for predicting the weather produces a table with numerical values for the temperatures in a certain region as a function of time. These data are incomprehensible for most most stake holders; therefore they could be presented in the form of a map with small thermometers drawn in , 33

price elasticity◁ (prijselasticiteit)
one of the mechanisms economists use to describe the relation between price, demand and supply in a free market;
Example: the number of purchases as a function of price: typically, purchases decrease if price increases , 158

problem owner◁ (probleemeigenaar)
a person or group of people who benefit from the solution of the problem, or: who take the initiative for the problem being solved;
Example: for a model to predict the risk of aircraft failure, the problem owner could be the aircraft manufacturer, having primary interest in accurate estimates of this risk , 33

process◁ (proces, voortgang)
a behavior, or a collection of behaviors, of a dynamical system. The word 'process' in physics typically refers to one particular behavior; in computer science, it refers to a running program, which can display a variety of behaviors, perhaps depending on its input;
(no example) , 281
the development of some system over time;
Example: boiling an egg, doing a long division, performing a billiard shot , 21

process◁ (proces, voortgang)
the changes that a dynamical system undergoes when it develops over time, involving causes and effects;
Example: from physical processes (compression, expansion of gasses, propagation of waves or moving material objects) to social processes (the occurrence and resolution of conflicts, organization and reorganization of institutions , 94

projection◁ (projectie)

limiting the number of properties, or the number of values of properties of a system to reduce the state space;
Example: for a sock, in most cases we are not interested in exactly how dirty it is. The only values we want to distinguish are 'clean' or 'not clean'. , 99

properties, exposed -◁ (zichtbare eigenschappen)
properties in a conceptual model for which a value change in the modeled system is visible;
Example: in a clock with only an hour and a minute hand, the number of minutes since midnight is exposed , 100

properties, hidden -◁ (onzichtbare eigenschappen)
properties in a conceptual model for which a value change in the modeled system is not visible;
Example: in a clock with only an hour hand and a minute hand, the number of seconds since midnight is hidden , 100

property◁ (eigenschap)
a means for distinguishing concepts;
Example: `water` and `ice` are two very much related concepts. To distinguish them, we can use the property `aggregationPhase`. For water, the value of this property is `liquid`, for ice it is `solid` , 59
a pair (name, set of values);
Example: the property `height` of the concept `lantern` could be 6.0, stating the height of a lantern is the floating point number 6.0 (=a number of meters), that is: a set with only one element. It could e.g. also be the range of numbers between 5.0 and 8.0, stating that the lantern is anywhere between 5.0 and 8.0 meters high , 59
aspect of a concept that carries information;
Example: material things have properties such as `size`, `mass` and `aggregationPhase`. These are meaningless for a concept such as `pianoSonata`. Conversely, properties such as `duration` or `loudness`, applicable to `pianoSonata`, have no meaning for a concept such as `sandwich` or `briefcase` , 59

proposition◁ (bewering)
sentence that is either true or false;
Example: 'It is raining' is a proposition that is here and now true if at this moment and this location, droplets of water fall from the sky. 'please close the window' or 'is the window closed?' are not propositions , 182

purpose◁ (doel, doelstelling)
-of a model: what the modeler wants from the model;
Example: optimization, decision support, verification are possible purposes of a model , 13

quantities, hidden◁ (verborgen grootheid)
non-exposed quantities: quantities that may or may not be present in a model, and that may help explain seemingly non-causal behavior in a system;
Example: the number of bacteria in the body of an infected, but not yet diagnosed patient during incubation time , 281

quantity◁ (grootheid)
a mathematical object that can assume a value;
Example: in a mechanical model, g, the gravity acceleration is an essential quantity , 67

random experiment◁ (door toeval bepaald, niet voorspelbaar)
A random experiment is an experiment that can result in different outcomes, even though it is repeated in the same manner every time;
Example: throwing of dice , 25

range◁ (bereik)
-of a function: the set of values a function can return;
Example: for the function 'square', the

range consists of all non-negative reals , 62
range◁ (reeks, serie)
a set, where all elements of the set are known by knowing just a minimum and a maximum;
Example: the range of integers between 3 and 6 is the set with elements 3, 4, 5, 6 , 60
ranking◁ (ordenen in rang)
assigning integers to an ordered collection of items, such that the order of the integers matches with the order of the items;
Example: Olympic medal winners are ranked 1, 2 and 3 , 72
rational◁ (rationaal)
-of a function: the function can be written as a quotient of two polynomial functions;
Example: f(x) = (x+1)/(x-1) , 156
raw◁ (ruw, onbewerkt)
of data: data that results from observation or measurement, without any further processing;
Example: reading a thermometer, we see that the mercury level is halfway between the 12th and the 13th marks. '12.5' is a raw reading. Only by further processing, e.g., using the numeric labels near the marks, we can deduce that the temperature measured is, say, 21.5 centigrades , 31
reachable◁ (bereikbaar)
of a state S: if there is a path trough a state space, consisting of admitted transitions, leading from a reachable state to S, S is reachable;
Example: if a bulky sofa could be placed in an attic that has only a narrow spiral staircase as access, there must be a route for this sofa over the staircase such that it never gets stuck , 99
reason◁ (redeneren)
(verb): obtain the resulting value for a formal expression by applying rules of logic; opposed to calculate;
Example: 'all cars must have licence plates; a police car is a car, hence a police car must have a licence plate' is a valid reasoning. If both premises are true ('all cars must have licence plates', and 'police car is a car'), the conclusion ('a police car must have licence plate') is also true , 28
reconstruction◁ (reconstructie)
-in the context of sampling: the process to recover information about a continuous quantity from a set of sampled values of this quantity;
Example: the audible sound, produced from reading digital information from a CD , 22
recursion◁ (recursief)
the property that something is defined in terms of itself, or perhaps of earlier versions of itself;
Example: 'current age' could be defined as 'last year's age plus one year'. N factorial is defined as (N-1) factorial multiplied with N, where 0 factorial is defined as 1 , 111
reflection◁ (reflectie, bespiegeling)
the mental process of looking back to some achievement with the purpose to improve one's understanding and skills for future occasions;
Example: a modeler could conclude, that, despite her mathematical skills, she has difficulty in explaining the model outcomes in terms that make sense to the problem owner, and take corrective actions , 35
reflexive◁ (reflexief)
-of a relation: a reflexive relation applies between an item and itself;
Example: `hasSameFatherAs` , 280
regime◁ (regime, bereik)
the range of values for input quantities in a functional model where the model is supposed to work properly;
Example: for the taxi example from Appendix **??**: kilometer price somewhere between 10 cents and 10 Euro per kilo-

meter , 231
regime◃ (regime, bereik)
- in models: a range of values for the quantities in a model such that a set of assumptions holds; or a range of values for the quantities in a model such that the behavior of the model is similar but different for another regime;
Example: consider a model for the physical properties of water. For temperatures and pressures in a certain range, water is solid; in another range it is liquid and in yet another range it is a gas. The behavior in one regime is similar over the entire regime (for instance, for a gas, volume and pressure are inversely proportional for a constant temperature. This is not true for liquid or solid.) , 36
Relation Wizard◃ (relatieassistent)
interactive tool to help translate non mathematical properties of a relation into the appropriate mathematical terminology;
(no example) , 142
relation, equivalence◃ (equivalentierelatie)
a relation that is reflexive, symmetric and transitive;
Example: `hasSameFatherAs` , 77
relation◃ (verband)
way to connect two or more concepts or their properties;
Example: `isMarriedTo()`, `greaterThan()`, but also formulas such as V=IR, relating V, I and R in Ohm's law , 62
representation◃ (vertegenwoordiging)
the relation between a concept in the model and its corresponding entity in the modeled system;
Example: p represents the pressure in the vessel , 58
reproducibility◃ (reproduceerbaarheid)
the extent to which repeated executions of the same measurement yield results that are similar;
Example: if the noise in an experiment is larger, there will be a larger variation among the outcomes, and the reproducibility will be less , 214
requirement◃ (eis, noodzaak)
necessary property of an ATBD, a predicate over the ATBD or over a cat.-II quantity that needs to be true;
Example: `operatedByHand()`, `waterproof()` and `lessThan3kg()` for an umbrella , 183
research◃ ((wetenschappelijk) onderzoek)
(as opposed to design) the systematic investigation of some object or phenomenon with the primary intention to gain understanding for the benefit of a scientific community;
Example: assessment of the value of a physical constant, checking of a hypothesis regarding the behavior of a system , 13
residual error◃ (residu, overblijfsel)
for a black box model: a measure of the differences between the raw data and the corresponding values as they follow from model calculation;
Example: let y = f(x) be model for data fitting; the residual error is the average of the squares of f(xi)-yi , 226
residue◃ (residu, overblijfsel)
see 'residual error';
(no example) , 226
return value◃ (resultaatwaarde)
-of a function: the value, part of the range of the function, that is obtained by applying the function to its argument;
Example: for the function 'color', the return value is e.g. red, green or purple , 62
Richards curve◃ (Richards curve)
a generalized version of the logistic function, with quantities that allow tuning the behavior to be applicable in many practical contexts;
Example: used to describe growth processes , 160
root◃ (wortel)
in a directed a-cyclic graph: a node with

only outgoing arrows;
Example: in the four categories model, cat.-II quantities are root nodes with respect to the `dependsOn`-relation; with respect to the `isArgumentTo`-relation, the cat.-I and cat.-III quantities are roots , 137

saddle point◁ (zadelpunt)
a point $(a, b, f(a, b))$ of $z = f(x, y)$ is called a saddle point if (a, b) is a critical point but $f(a, b)$ is not an extremum;
(no example) , 81

sample◁ (monster, steekproef)
(verb) represent the behavior of a large set by knowing relatively few values of that set;
Example: estimating the quality of a batch of oranges by testing the quality a handful (stochastic sampling), or storing a continuous signal (music) in the form of a large number of values, each 1/44100 second apart on a CD (digital audio sampling) , 112

sampling◁ (bemonsteren, een steekproef nemen)
representing a continuous quantity by a discrete one that comes in steps, small enough so that for practical purposes no information is lost;
Example: a motion pixture is a way to sample the visual impression of some scene by taking 24 snapshots per second , 22

sanity check◁ ((geen Nederlandse vertaling))
a simple test to see if a construct could make sense;
Example: looking at consistency, order of magnitude, or dimensional correctness can reveal if a formula or a numerical outcome could make sense , 41

saturate◁ (verzadigen)
the behavior of a functions that with increasing or decreasing argument, the value will not exceed some threshold value;
Example: the output voltage of an amplifier may increase with the voltage of the input signal, but if it can never get higher than the voltage of the power supply, the increasing behavior must flatten for sufficiently high input voltages , 157

scalability◁ (schaalbaar)
-as a criterion for modeling: the extent to which the size or the scale of a modeled system can vary in order to be tractable by a given model;
Example: algorithms to find an optimal solution for the traveling salesman problem are not scalable, as adding 1 to a problem size N causes a multiplication in the effort to solve the problem of a factor N , 247

scalar◁ (scalair getal)
a real number, in contrast with a vector or a matrix;
Example: a vector can be multiplied with a scalar to yield another vector , 283

scale model◁ (schaalmodel)
example of an empirical model system, consisting of physical objects that are studied to obtain information on other physical objects;
Example: scale models are used in wind tunnels or hydrological laboratories to study floating and streaming behavior of vessels, dykes, and aircraft , 257

scale, interval◁ (intervalschaal)
scale that allows addition and subtraction;
Example: temperature scale in Centigrade , 73

scale, Mohs◁ (schaal van Mohs)
example of an ordinal scale that is not an interval scale: if for two minerals, A and B, A receives scratches and B does not when rubbed against each other, then B's Mohs number is higher than that of A;
Example: diamond is harder than steel and steel is harder than chalk , 72

scale, ordinal◁ (ordinale schaal)
a scale that allows ordering and median computation;

Example: Mohs scale , 72
scale, ratio◁ (ratioschaal)
a scale that allows the calculation of ratio's;
Example: the Kelvin scale and many other scales for physical quantities , 73
scale-free◁ (schaalvrij)
property of a network: a network is scale-free if, irrespective of the number of nodes in the network, the average number of edges needed to reach one node from an arbitrary other node is constant;
Example: the population of the world has the so-called 6 handshakes property: there are (on average) no more than six handshakes between any pair of people, knowing each other, in the world , 249
scale◁ (schaal(grootte))
of a problem: a characteristic quantity (n) that determines the size of the problem;
Example: the precise value is less important; modelers are usually interested in the increase of the amount of time or memory space, needed to solve the problem, when n increases. A distinction between sub linear, linear, super linear and exponential is often made , 247
scope◁ (toepassingsgebied)
-of a model: range of modeled situations to which a model should apply, or for which a model should be useful;
Example: water could be modeled as an ideal gas, provided that the temperature is not too low and/or the pressure is not too high , 39
segmentation◁ (opdeling)
separating a unity into meaningful segments;
Example: a city map is segmented into street blocks; a year schedule may be segment in semesters, trimesters or quartiles, and an living organism may be segmented into digestive, reproductive, respirational and other functional segments , 57
semantic network◁ (semantisch netwerk)
entity relation graph;
(no example) , 67
sensitivity analysis◁ (gevoeligheidsanalyse)
for a function: the calculation of all its condition numbers around a certain point of the domain;
Example: the sensitivity of a proportional model, where a cat.-II quantity is proportional or inversely proportional to all cat.-I and cat.-III quantities is 1: a change of 1 percent in an input causes a 1 percent change in the output irrespective of the values of the quantities , 234
servicing◁ (afhandelen)
- of an external event: the act of responding to that event from within a process;
Example: somebody is cooking, and the phone rings; in responding to the phone, the stove has to be put off first , 104
similarity◁ (overeenkomstigheid, overeenstemming, nabijheid)
a measure to express the extent to which two things resemble each other. Opposite of distance;
Example: two shapes can be said to be more similar if their aspect ratios are closer; two distributions can said to be more similar if their means and standard deviations are closer , 219
simplex method◁ (simplex methode)
a method to solve optimization problems with inequality constraints that are linear expressions of the unknowns;
(no example) , 281
simulation◁ (nabootsing)
for some proces P1, a simulation P2 is a second process that aims to replicate certain aspects of P1;
Example: the game of chess simulates traditional warfare , 276
simulation◁ (naspelen)
the evaluation of subsequent states of a dynamic model by unrolling the recursive definition of quantities in state i+1 in

terms of their values in state i, beginning at some starting state i=0;
Example: using the bank transcripts over a period of time to reconstruct and analyse the time-behavior of a bank account , 111

singleton◁ (singleton)
a set that contains only one element;
Example: the set of monuments in Paris being taller than 300 meters , 60

smooth◁ (glad)
-of a function: there are no abrupt changes, neither in value nor in derivative, when the argument changes little. Sometimes smoothness is indicated with a quantity, say B, indicating that for any two arguments x1 and x2, the difference f'(x1) - f'(x2) is in absolute value not larger than B times abs(x1 - x2); this condition is called Lipschitz-smoothness;
Example: a sine function is smooth, whereas a saw-tooth function is non smooth , 156

solution, closed form◁ (oplossing in gesloten vorm)
the solution of a problem, given in terms of a finite set of arithmetical operations, so that generic properties of the solution can be stated without having to rely on numerical estimates;
Example: the solution for x of the equation ax+b = y for any a, b, y, a different from 0, is x=(b-y)/a , 14

spatial◁ (ruimtelijk, ruimteachtig)
regarding space; see also temporal;
Example: length, width, and height are spatial dimensions, as well as 'per length' (spatial frequency, as in 'this necklace has 3 beads per centimeter') , 98

SPEA◁ (geen Nederlandse vertaling)
algorithm for genetic optimisation of multi-objective problems. Acronym for Strength Pareto Evolutionary Algorithm;
(no example) , 194

specialization◁ (specialisatie)
-as a criterion for modeling: abbreviation of 'level of specialization of the intended problem owner', with the interpretation that a model, in respect of this criterion, is better if the required level of specialization on the side of the user can be lower;
Example: a model used to help specialists in medical diagnosis may assume the users to be knowledgeable in the principles of medical imaging , 249

specification◁ (specificatie)
- as a purpose for models: make sure that something in the modeled system will occur of will be realized, or give a description of some artefact that is sufficiently complete so that the artefact can be realized (purely) on the basis of the specification, perhaps still allowing for open choices;
Example: a blue print of a piece of furniture (say, a chair) is a model of that chair - made at at a time where the chair doesn't yet exist, but (perhaps) leading to the actual existence of the chair once it is build , 17

spreading◁ (statistische spreiding)
see noise, random error;
(no example) , 213

stability◁ (stabiliteit)
the property of a numerical result that it does not vary wildly with small perturbations of the input;
Example: in a model that predicts the profit of a company, we don't want a large change in predicted profit if one of the assumed input quantities (being inevitably uncertain) undergoes a small change , 234

stable solution◁ (stabiele oplossing)
a solution (e.g., of a differential equation) such tat, if (initial) conditions are slightly different, the solution is not different;
Example: For the logistic equation, y'=ky(M-y), y(t)=M is a stable solution , 125

stake holder◁ (betrokkene)
a person or group of people who are, purposely or involuntary, affected by the solution of some problem or the failure thereof;

Example: an insurance company is a stakeholder for a model to predict the effectiveness of a pharmaceutic drug , 33

standard deviation◁ (standaardafwijking)
measure for the spreading of values in a set of data;
Example: when throwing a fair dice, the arithmetic mean is 3.5. Every outcome, from 1 to 6, has equal probability. The standard deviation therefore is the square root of 1/6 of the sum of squares of each of the outcomes minus 3.5, or approximately 1.71 , 220

state chart◁ (toestandsdiagram)
a process model where a process is a graph, the nodes being states and the edges transitions between states;
Example: an illustrated manual, showing step-by-step the assembly of a piece of furniture , 95

state space explosion◁ (explosie van de toestandsruimte)
the phenomenon that the number of states, and the number of possible processes, of a system, grow exponentially, both in the number of properties of the system, and the number of steps of the process;
Example: the game of chess continues to fascinate players because its state space is intractably large , 98

state space◁ (toestandsruimte)
the collection of all possible states of a system, that is: all possible bindings of the properties of the system to values;
Example: all possible configurations of 54 colored squares that can be obtained by rotating any of the six faces of Rubick's cube form the state space of this cube , 97

state transition◁ (toestandsovergang)
the replacement of one binding of values to the properties in a system to another binding;
Example: somebody celebrating his birthday, where the property age, originally bound to N, now gets bound to N+1 , 95

state, initial◁ (begintoestand)
in a dynamic process: the state that is occupied by the process, when no transitions have yet taken place;
Example: the initial state of all symphonies is a conductor raising his baton, the orchestra being silent , 99

state◁ (toestand)
the snapshot of a system, containing all its quantities and their current values;
Example: Rembrandt's Night watch depicts one state in the history of the 17th Century Amsterdam police force , 95

static◁ (statisch, stationair, tijdsonafhankelijk)
not involving time; opposite to dynamic;
Example: the balance of forces that keep a building from collapsing under its own weight , 21

stationary point◁ (stationair punt)
an argument value x of a function f(x) such that sufficiently small change of x causes the value of f to stay the same;
Example: a local maximum or minimum of a function , 23

stationary◁ (stationair, onveranderlijk)
not varying as a function of time or another argument, constant;
Example: the value of physical constants such as the speed of light and the mass of a proton , 21

steepest descent◁ (steilste daling)
a numerical technique to approximate a local minimum of a function by repeatedly evaluating the function in a point x, estimating the change of the function in the vicinity of x, and updating x by a move in the direction of steepest descent;
(no example) , 187

steer◁ (besturen)
- as a purpose of models: based on a representation of X, and measurements on some aspects of X or its environment, perform actions that influence X;
Example: a thermostat influences the

working of a heater, thereby influencing the temperature in a room , 18
step size◁ (stapgrootte)
-in sampling: the increment of a sampled quantity;
Example: in film, 1/24th of a second; in CD's: 1/44100th of a second , 276
stochastic◁ (stochastisch, door toeval bepaald)
involving randomness;
Example: the motion of molecules in a gas , 25
straight◁ (recht)
the property of a curve passing through two given points, that its length is minimal;
Example: the trajectory of a falling point mass in vacuum with zero initial speed is a straight line , 24
strength◁ (geen Nederlandse vertaling)
a quantity in the SPEA algorithm, representing the number of individuals by whom an individual is dominated;
Example: the strength of individuals on the Pareto front is 0, as they are dominated by 0 others , 195
strong◁ (sterk)
- of an assumption: a stronger assumption implies a weaker assumption;
Example: `greaterThan(x,1)` implies `greaterThan(x,0)`, so `greaterThan(x,1)` is stronger than `greaterThan(x,0)` , 212
subjective◁ (subjectief)
depending on an individual observer, or on an individual opinion. Opposite of objective;
Example: John likes marshmallows , 15
superposition principle◁ (superpositieprincipe)
the principle, in physics, by which the combined effect of two actions or interactions is described by a mathematical sum of quantities, each quantity representing one of the two actions or interactions. ;
Example: combined forces have the same effect as the sum of the forces; electromagnetic fields of two interfering waves may be added to produce the electromagnetic field of the resulting wave , 287
surjective◁ (surjectief, 'op' (voor functies))
of a function: every value in the range is reached;
Example: mapping sufficiently many throws of a fair die to the set of numbers 1, 2, 3, 4, 5, 6 is surjective, as in the long run every outcome will occur. If the oldest living person in the world is, say, 110 years old, however, the mapping from people to ages, at any time, does not have to be surjective with respect to the integers between 0 and 110 (e.g., it might be that there is nobody with age 109) , 191
surprise◁ (verbazing)
-as a criterion for modeling: the potential of a model to produce a surprising result. This is particularly relevant for models used for exploration, inspiration, abstraction or unification;
(no example) , 261
symbolic◁ (symbolisch)
- of a modeling strategy: performing operations upon mathematical or logical expressions, not dealing with values only. Opposite to numeric;
Example: a+b = b+a is a symbolic expression, denoting the property that addition is commutative, as in 3+4=4+3 or 17+1=1+17 , 23
symmetric◁ (symmetrisch)
-of a relation: a symmetric relation between A and B also holds between B and A;
Example: `marriedTo` , 280
symmetry◁ (symmetrie)
the condition that only part of a system needs to be known in order to know something about the entire system;
Example: if only the left hand part of the shape of a mirror-symmetric piece of clothing is drawn, a capable tailor can make the entire piece. Symmetry can be spatial, but temporal and other symmetries occur as well , 98
synchronous◁ (gerelateerd in tijd, in de maat)

of an event with respect to some process P: occurring between two predefined state transitions in P. Opposite to asynchronous; Example: when preparing a sandwich, the butter should be applied in between slicing the bread and putting on the topping. Applying butter is to be synchronized with the other two stages of the process , 102

T-test◁ (T-test)
a statistical method to assess if two collections of values did likely come from different distributions; Example: A common application of statistics is, to assess if a claimed effect is real. For instance, if taking a particular medicine is effective in treating a particular symptom. The T-test (in one of its many versions) may then be a useful tool. , 295

tangent plane◁ (raakvlak)
the tangent plane to the graph $z = f(x, y)$ at (a, b) is given by $f_x(a, b)(x - a) + f_y(a, b)(y - b) - (z - f(a, b)) = 0$; (no example) , 49

temporal◁ (tijdachtig)
regarding time; see also spatial; Example: duration and (temporal) frequency are temporal dimensions , 98

term◁ (term)
a quantity, often occurring in a sum expression. See also factor; Example: in (a+b)(a-b)=a x a - b x b, a, b, a x a and b x b are terms whereas (a+b) and (a-b) are factors. In a x a are two factors a , 68

theory◁ (theorie)
A body of knowledge, accepted by a community, used to explain empirical phenomena; Example: the theory of electromagnetism in physics, the theory of consumers' behavior in economy , 14

thought experiment◁ (gedachte-experiment)
attempt to find some contingent fact by mere reasoning, without any empirical observation; Example: Galilei was interested in the speed of falling objects: would speed depend on weight? Prior to the famous leaning tower experiment with the two unequal canon balls, he postulated the opinion that both would fall with equal speeds, based on a thought experiment: suppose that a canon ball would be cut in two equal halves, where the two stay in close contact during falling, would this influence the falling behavior? , 276

time lapse◁ (verstreken tijd)
the amount of time between two subsequent transitions in a model with full time ordering; Example: a day is the time lapse between rising and setting sun , 106

time reversibility◁ ((tijd)omkeerbaarheid)
the aspect of time in physical processes that, at microscopic scale, past and future can be interchanged; Example: a hypothetical film of a microscopic process can be viewed in reverse without showing any non-physical behavior , 93

to-do-list◁ (geen Nederlandse vertaling)
a list of quantities introduced during the development of a functional model, that have not yet been defined in terms of quantities of categories I, III or IV; Example: once the to-do-list is empty, a model is fully defined and, in principle, executable , 141

tolerance◁ (marge)
range of values that are acceptable for some purpose; Example: a measurement with 38 centigrade with a tolerance of plus or minus 2 is too coarse to assess if a patient has fever , 212

trace◁ (spoor)
a route through a state space; Example: the footprints left by an animal walking through the wood represent a trace in its state space , 97

trade-off◁ (afweging, dilemma)
the situation, in cat.-II space, where for two non-dominated solutions, one is better with respect to one cat.-II quantity, whereas the other one is better with respect to another cat.-II quantity;
Example: for many people, taste and health of food products form a trade-off , 193

train◁ (trainen, instrueren)
- as a purpose of models: assist trainees to get familiar with some system X, avoiding the risks if untrained personnel would work with the actual X;
Example: flight simulators to train pilots; simulators of industrial plants to train operators; anatomical simulators to train surgeons , 18

transition, internal◁ (interne overgang)
transition that is not caused by an external event;
Example: when a balloon is gradually inflated, there is a moment where it explodes. This happens when the stress in the balloon exceeds the strength of its skin. It does not require any external trigger , 104

transition◁ (overgang)
short for state transition;
(no example) , 95

transitive◁ (transitief, overdraagbaar)
the property of a relation R(A,B), implying: if R(A,B) and R(B,C) then R(A,C);
Example: `greaterThan` , 71

traveling salesman problem◁ (handelsreiziger probleem)
prototypical problem where the needed time for optimal solution is believed to grow faster than polynomially in the problem size. The problem asks for the order to visit a set of given geographic locations so as to travel the smallest total distance;
Example: many discrete search problems, such as the best next move in a given chess position, can be transformed to traveling salesman problems , 249

trend line◁ (trendlijn)
result of a linear least squares fit;
Example: stock market analysts try to spot global tendencies in stochastic data, such as Dow Jones indices, by fitting smooth curves (such as straight lines) through segments of data, sometimes to predict future market behavior , 228

triangle inequality◁ (driehoeksongelijkheid)
requirement on distance functions, that the distance from A to B via C is not shorter than the (shortest) distance from A to B;
Example: when traveling from A to B, visiting C underway can never make the route shorter , 219

triple store◁ (geen vertaling)
a method of sotring information, not in terms of tables, but rather in terms of triples;
Example: We want to store the information that a tomato is red and round, and a cucumber is green and oblong. In a tables-based approach we have two columns, one for color and one for shape, with values (red, green) and (round, oblong); there are two rows, one for tomato and one for cucumber. In a triple store, there are 4 triples, namely (tomato-color-red), (tomato-shape-round), (cucumber-color-green), (cucumber-shape-oblong) , 246

type, compound◁ (samengesteld type)
an element of a compound type is a concept that has one or more properties that themselves have a type;
Example: the type `circle` has properties `radius` and `midpoint`. These have a type: `radius` is a real number, and `midpoint` has type `point` , 68

type, elementary◁ (elementair type)
types that are not compound, a.k.a. atomic types;
Example: the type Boolean is the set with elements `TRUE,FALSE`. The values `TRUE` and `FALSE` are not compound , 68

type◁ (type, soort, specimen, verzameling van

waarden)
-of a property is the set of values that can be associated to that property;
Example: a property with type `integer` can take any integer value , 59
-of a quantity: the set of values that can be assumed by that quantity;
Example: the set with elements `red, green, blue, ...` is the type of a quantity named `color` , 67

uncertainty◁ (onzekerheid)
in measured data: repeatedly measuring a real quantity does not produce a single, unique and exact value;
Example: if a real-valued cat.-II quantity is found to have an uncertainty range of sigma, whereas it was known, prior to the model, to be within a range rho, the model has reduced the uncertainty with -log(sigma/rho) bits of information , 215

unification◁ (unificatie)
- as a purpose of models: providing a representation that allows explanations and perhaps even predictions in two or more domains that initially were thought to be unrelated;
Example: the flow of gas through a pipe and the flow of traffic on a motor way can, under some circumstances, be unified , 15

uniform◁ (uniform)
of a distribution: a distribution, described by a lower and an upper bound, where every value in between has equal chance of occurring;
Example: outcomes of a fair die throw are uniformly distributed over the set with element 1, 2, 3, 4, 5, 6 , 217

unit, pseudo◁ (pseudoeenheid)
a term that does not count or measure some quantity, but that nevertheless can appear in the definition of quantities to help analyzing whether a quantity should be multiplied or divided by another quantity, and which safeguards against adding, subtracting or comparing incompatible quantities;
Example: when we are computing with a number of chimney sweepers in Eindhoven, the unit of this quantity could be Sw/A. 'Sw' is a number of chimney sweepers; 'E' is a pseudo unit that informs us about the condition 'in Eindhoven' , 150

unstable◁ (instabiel)
the behavior of numerically evaluated recursive simulations that, typically due to too coarse sampling, approximated solutions get further and further off;
Example: a mass spring system, approximated by discrete sampling, where the sampling time step is not small compared to the period, may get instable , 285

validation◁ (vaststelling, beoordeling (van bruikbaarheid voor een doel))
assessing if a model is valid or invalid, in other words; if the model outcome is suitable to fulfill the purpose. Validation amounts to answering the question: 'Is X the right model?', as opposed to verification: 'Is model X right?';
Example: Validation includes: showing that a model behaves consistently with our best knowledge of the modeled system, and produces output that is valuable for the model's purpose. Assessing if uncertainties are sufficiently small is an example of validation. This can be done by running the model on inputs with known (observed) outputs, and see if these are reproduced within sufficient accuracy. E.g., a model for weather predictions can be validated by having it predict yesterday's weather from data of the day before yesterday , 209

value-free◁ (waardevrij)
of a model: the property that all assumptions and approximations are neutral with respect to benefits or disadvantages for the problem owner;

Example: a quantitative risk analysis, commanded by an airline carrier, when revealing unexpected hazards, may have fatal consequences for the patron - rendering the model anything but value-free , 252

value◁ (waarde)

- in design: that which is to be produced in a design process. Examples of value could be: profit, safety, health, amusement, ...;
  Example: in designing a vehicle, values could be safety, speed, and comfort , 16
- of a property: in case the set of values of a property is limited to a singleton, the (single) element of this singleton is called 'the value of the property';
  Example: '40000 km' is the value of the property 'perimeter' of the concept 'earth' , 60

variable◁ (variabele)

a quantity that can take several values, or an unknown in an equation;
Example: most text would not call pi = 3.1415... a variable, whereas in the equation 2x = pi, x would be called a variable , 68

variable◁ (variabele)

quantity that occurs as argument of a function;
Example: In y=f(x), x is called a variable. If f is a non-trivial function, y will assume different values as well (if not, the notation y=f(x) is misleading); therefore it is common to call y also a variable. Sometimes x is called an independent variable, whereas y is called dependent variabloe, as it depends on x , 45

variance◁ (variantie)

square of the standard deviation;
Example: for uncorrelated, normally distributed stochastic quantities, the variance of the sum is the sum of the variances , 223

verification◁ (verifiëren, zekerstellen)

- of a model: the assessment if some predicate over the model is true. Verification amounts to answering the question: 'Is the model X right?', as opposed to validation: 'Is X the right model?';
  Example: verification includes dimensional consistency, the correspondence between the conceptual model and the formal model, and the check that found value(set)s of quantities are part of the types defined during conceptualization , 209

verification◁ (verificatie, zekerstelling)

- as a purpose of models: assessing if something in the modeled system is true;
  Example: in a control system for railway signaling, the software should be such that it is impossible that two trains are allowed in the same block. To verify if this property holds, a model of the software is constructed. Verification is a particular instance of analysis , 17

verisimilitude◁ (verisimilitude)

the problem in philosophy related to assessing the truth of models;
Example: elementary particles are mathematically defined concepts, used to explain the outcomes of physical experiments. Whether or not they refer to entities beyond the mathematical realm is a question regarding the verisimilitude of the theoretical framework of particle physics , 211

weak◁ (zwak)

- of an assumption: a weaker assumption is implied by a stronger assumption;
  Example: `greaterThan(x,0)` is implied by `greaterThan(x,1)`, so `greaterThan(x,0)` is weaker than `greaterThan(x,1)` , 212

white box◁ (witte doos)

other term for glass box;
(no example) , 29

wicked problem◁ (geen Nederlandse vertaling)

a problem that has the property that the solution of the problem causes the orig-

inal problem to change;
Example: the introduction of DDT into agriculture, intended to increase crops, after a few years gave rise to a reduction of crops , 253

wisdom of the crowds◁ (wijsheid van de menigte)
a process of guessing the value of some quantity by averaging the guesses of individuals in some group, assuming that individual guesses are independent and not biased by systematic error;
(no example) , 143

wish◁ (wens)
the property of an ATBD that a given cat.-II quantity should be lower (or higher) than the value for any competing ATBD that we compare it to, formalising the intution 'the lower (higher), the better';
Example: an umbrella should catch as little wind as possible , 184