Proc. Of the 10th International Conference on Advances in Engineering, Technology and Management - ICETM 2024
Copyright © Institute of Research Engineers and Doctors. All rights reserved
ISBN: 978-1-63248-197-9 DOI : 10.15224/978-1-63248-197-9-01

# A Computational View at Software Production

Roland Kaschek

Abstract—*Computers are machines for the approximate execution of mathematical operations. By using appropriate software, computers can be used to help solve practical problems. The process of producing software is a quite complicated one and surprisingly is usually not discussed in terms of figuring out the Mathematical model that finally is going to be evaluated by the computer. In this paper we identify the software process stage, in which the Mathematical model is produced. We also discuss the Mathematical concept of quantity, how to measure quantities and Mathematical model. We moreover discuss the indispensability of models when it comes to acquiring knowledge about the world. We conclude the paper with a brief review of quality aspects of Mathematical models and first ideas about how to engineer Mathematical models.*

Keywords—model, conceptual model, Mathematical model, quantity, amount, software process, model engineering

## Introduction

Creating software in many cases is a difficult thing to do. This is, why it requires careful preparation and proceeding. It has turned out, that before the ultimate task of coding the desired computer programs can be conducted, a lot of preparation is required. One of the required tasks is the creation of, what nowadays is called, conceptual model. Surprisingly, however, there is neither a widely held consensus about what a conceptual model is nor is there such consensus on what its key task is. This is the more surprising as computers are machines that implement approximations of Mathematical operations and therefore, on the way to the actual coding, at some stage it needs to be specified which mathematical operations with which accuracy, speed and security precautions should be carried out under which conditions and with what input data. We are going to argue that the Mathematical stuff has to be dealt with during what in the software process setting is called design. We identify the design with what in more data oriented discussions is called conceptual model and essentially is a Mathematical model. We are going to base our discussion here on [7]. The key concept to be used for our purpose is, what in Mathematics is called “quantity”. In particular we are going to propose basic ideas about engineering Mathematical models.

## Software Production

Software, in a very narrow (and limited) understanding of the term, is a number of signs that can be interpreted by a computer and then executed. One way of understanding what is happening when these signs are executed is to understand these signs as mathematical terms and to understand the sign execution as calculating the mentioned mathematical terms. However, it seems that a mathematical understanding of what happens during software production is not very common in the literature on software production. In this paper I am going to show how Mathematics can be used to understand conceptual models as mathematical models.

When computers started to get popular in the 1950es the software used to operate these machines was part of the package purchased. Actually, looking back at that time one might feel compelled to consider that software (the operating system as well as the programming language) as rather simple. Some time later a view took hold according to which software production and use is a problem in its own right and better should not be connected too tightly with producing and selling of computers.

Software after some time turned into a business. Any software producer, who wants to survive the competition, must produce efficiently, effective software and mus be able to deploy and maintain their software as well as replace any legacy software. Therefore attempts were made to standardize software production. Finally a software process [2] was popularized hat consisted of several clearly identifiable steps that would not overlap one another.[1] That software process then was understood to consist of iterations of the following steps: requirements elicitation, design, implementation, test, deployment and maintenance. The meaning of these terms is fairly obvious. Nevertheless we are going to discuss the meaning of the first three steps.

During the first step the software engineers figure out the functionality to be implemented, the quality of that functionality, the circumstances and time at which that functionality is going to be used and the personal that is going to use that functionality. They thus create a comprehensive model of the intended software use. Obviously, during implementation the software engineers encode that functionality in a language the computer to be used can interpret. A model thus is created of the enactment of the software by the computer. In this regard it is important to note, that the implementation language usually differs significantly

1 For sake of simplicity I use here, what I think are still key activities in software production.

from the Mathematical language. If one accepts the thesis that a computer essentially is a machine for doing mathematical calculations, then the conclusion to be drawn is that the calculations to be carried out have to be specified during design. As in this paper we are dealing with a mathematical angle on software production, we are going to ignore any other tasks to be carried out during design.

A consequence from our observations is that practitioners are not able to understand and validate the conceptual model. Thus techniques should have to be worked out that allow to show that the stated requirements actually follow from the conceptual model and that this model does not have implications contradicting the stated requirements.

## Models

I cannot go into a discussion on literature on models or modeling. However, I think that a good starting point with regard to that is [9, 6].

Humans, more than other animals, live on experience of life and on knowledge. Often knowledge is hard to come by. Therefore, rhetorical devices were developed to transfer knowledge, that applies to one item, to another. Well known such devices are analogy [3], metaphor [8] and model. Amongst these, metaphor and model have become important to informatics. In the tradition of Stachowiak [11], one says, that an an agent A uses an item M, called model, for item O, called original, if A under specified circumstances and at a specified point in time satisfies A's information need concerning O by means of M. This typically applies if M is not identical to O. Note, that M needs not to be an artifact and may predate, postdate or may have come into existence at the same time as O has. The mentioned information need typically consists either in getting an (approximate) information about O or in specifying details of O. If the former applies the model is called descriptive while it is called prescriptive if the latter applies.

Quality aspects of items used as models are of interest if and when models are being used in practice. Given an agent, arbitrary usage circumstances and time of use, any item may be used as a model for any other item. Consequently there cannot be a generally accepted structural definition of the concept of model. Rather, a functional or operational explication of the term model is required and already was given by Stachowiak and was reported about above. To my knowledge, Thalheim was the first one to draw from that the consequence to understand models as instruments. This, of course, is an attempt to explicate model by a metaphor. The advantage of this, over Stachowiaks's logic based approach, is, that a larger latitude is available to extend the conceptual framework as needed for understanding the concept model. Clearly, he purpose of using the instrument is to get or provide information about the work piece, i. e. the original. Note, that if quality aspects of models are being considered, then any model, depending on the quality aspects chosen, might turn out to be a rather poor one. Even a poor model, however, may have its merits, in particular if it is the best one aid available.

Many different kinds of model exist and may be used. In Informatics the use of notional models is predominant. By a notional model we mean a model that consists of notions. The latter, however, are not understood as being the actual content of some human's brain. Rather, a notion, here, is considered as a sign in some human language that has a conventional meaning in that language, such as found in an encyclopedia, a conversational lexicon or similar. Physical models, e. g. animals or structures found or created in the world, have been used in many disciplines. They continue to be indispensable e. g. in Medicine. Analog computers are an example of physical models as used in Informatics. Understanding the Antikythera mechanism as an analog computer for representing astronomical knowledge shows, that the history of these analog computers is a rather long one [3].

## Mathematical Problem solving

A quantity [10] is a notion for which an additive rule is known, attaching to it a non negative number, called amount of that quantity.[2] A quantity's amount may be regarded as either constant or variable and is said to measure that quantity. Mathematical terms such as an equation or inequality thus maybe classified as either involving a quantity or not. For example $2 = 5$ and $2 > 6$ are Mathematical terms not involving quantities. They may be considered as either true or false. They, however, do not involve an assertion about the world. To make a Mathematical assertion about the world, a quantity has to be part of a Mathematical term. For example[3], Newton's famous, $F = M \cdot A$, involves the three quantities force F, mass M and acceleration A. The notional interpretation of it is, that a given force applied to a given mass results in an acceleration of that mass. The quantitative interpretation is, that if a force of amount f is applied to a mass of amount m then the mass is accelerated by an acceleration the amount of which is equal to f : m. The appearance of the duality of a physics equations maybe even more impressive with Einstein's equation Energy = Mass · Speed · Speed, where

2 Note, that we simplify a bit and ignore that a part of that notion is a so called dimension, i. e. a qualifier telling how to interpret the amount.

3 We are going to ignore here, that force actually is a vector.

Speed is the speed of light.[4] As is well known, it often is interpreted as saying, that energy and mass are equivalent to each other. Clearly that is not an assertion about the amounts of the energy and mass involved.

A quantity may be measured from different angels and thus may have attached to it several amounts or even classes of amounts. In Informatics the amounts may be considered as code for more semantically meaningful signs and be replaced by these. They also might be used immediately as object identity. Standard Mathematical operations may be used to define quantities from quantities. We are going to call a quantity a known or an unknown, respectively, if its measure either is known or unknown. A Mathematical or calculation model[5] is a collection of quantities some of which are connected to others by functions, relations, equations, inequalities, formula or algorithms. The purpose of the Mathematical model is to specify calculations for obtaining the amounts of its unknowns. The conceptual model (possibly additionally to other purposes) specifies calculations to be carried out and thus is regarded as a Mathematical model. If one does not use the concept of quantity when discussing calculation, then one easily can arrive at the wrong conclusion that calculations and formula like Force equals Mass times Acceleration are not models.

To solve a Mathematical problem one writes down a suitably chosen set of knowns and unknowns and then plunders the Mathematical Universe for quantities, relationships amongst quantities and formula or algorithms in such way that the amounts of the unknowns can be approximated or calculated. Obviously, to figure out the quantities to start with and to find a way to compute the unknowns from the knowns, is a creative process and we are not going to say anything about how to do it. Informatics operates the same way as Mathematics does, but rather exploits the Informatics Universe, which is a super set of the Mathematics Universe. In particular, the Informatics Universe may include signs that, strictly interpreted in the sense of Mathematics, may have no or at best a doubtful meaning.

In Mathematics, traditionally, the creator of a model evaluates it, i. e. measures the unknowns, themselves. In Informatics, however, a machine is going to do so. Therefore, usually an algorithm exists that translates the conceptual model (via a logical model) into a physical model, that can be evaluated by the available hardware.[6] In Mathematics, the model user often needs to have a relatively deep and comprehensive understanding of the model. This often is not true in Informatics. Rather, often proficiency in using the hardware, some critical software and an overall understanding of the universe of discourse suffices.

## IV. A Quantity Measuring Example7

To illustrate the mathematical way of measuring quantities I am going to follow the explanation given in chapter 1, p. 28 of [1]. Since a very long time, humans have attempted to measure elapsed time.[8] They have, for this purpose, used the perceived periodicity of change of astronomical, weather and life

4 By the way, for that equation to be true Speed squared would have to be measured in Joule : kg, which appears a little odd.

5 In [3], when discussing Mathematical models, they do not use the concept of quantity and thus ignore the Janus face of Mathematical models, i. e. The intricate interplay between the meaning aspect and the quantitative aspect of applied Mathematical formula.

6 This procedure, involving a logical model as distinguished from a physical model, is more frequently found in data oriented cases. In more algorithm oriented cases frequently the physical model is derived immediately from the conceptual one.

7 What follows is not intended to be an introduction into the modern discipline of measure theory. We merely look into the very early beginnings of it by using a modern example.

8 Note, that we are not attempting to measure an entity frequently referred to as “time” [3]. What we actually measure, when we deal with elapsed time, are past numbers of instances of occurring patterns of elementary astronomy, weather phenomena and, derived from these, patterns in the observable flora and fauna. Time, as understood as elapsed time, thus is not an “objective” phenomenon but rather is a purposefully constructed human abstraction. It is an amount attached to an interval [BUE , OM], where BUE and OM stand for “begin of used epoch” and “occasion to be measured”, respectively. However, elapsed time, by definition, has an implicit parameter, i. e. the trajectory and speed of the earth’s motion through space. This in turn might have implications with regard to the twin paradox [3] and comparable riddles. Finally, obviously, the sequence in which we have used the yardsticks for measuring elapsed time, is not the sequence in which these yardsticks have been invented.

phenomena. Thus elapsed time was measured in years passed since a mathematically arbitrary but culturally or politically important event. Examples of this were e. g. the begin of the rein of a new emperor and the alleged birth of Jesus Christ. However, in many cases the measurement of elapsed time in years was not sufficiently accurate. Therefore, ultimately the year was subdivided into months so as to be able to make required distinctions within a given year. While not each month was considered equal in that regard, months ultimately were subdivided into days, which allowed for an even more accurate measuring of elapsed time. Days then were subdivided into hours and the specification of elapsed time potentially became even more accurate. Elapsed time then could be measured as a quadruple: (y, m, d, h), when y, m, d, h are, what we now call natural numbers, that are subject to the limitations $0 \leq y$, $1 \leq m \leq 12$, $1 \leq d \leq 31$, $1 \leq h \leq 24$.

The example illustrates the basic idea of quantity measuring. The quantity measure is conceptualized as, what we now would call an interval [0,m] on the axis of real numbers, where, however m is an integer. The measure of the quantity would then be considered to be m. Essentially that number m would, moreover, be considered as the m-fould of a yardstick a (of length 1) and that yardstick would recursively be applied to the interval [i , m], where i is the number of times the yardstick has already been applied to get from 0 to m. However, if a measure x of the quantity turns out to be between m an m+1, then a new yardstick b would be used, that is shorter than a and would be used the already explained way to measure $x - m \cdot a$. Clearly, any desired accuracy of quantity measuring can be achieved that way. Obviously that basic idea of measuring quantities applies to many elementary cases. Note, finally, that in case quantities were to be measured that correspond to a curved line rather than a straight line, then appropriate yardsticks, to start the measuring with, might have to be reduced in length, so as to account for the bending of the curve.

## V. Quality Aspects of Mathematical Models

We are going to define the quality aspects repeatability, accuracy, efficiency, speed, availability, scalability and restart ability. This list most likely is not complete and might also be contested. Regardless of that, we feel that it is worth mentioning some of the quality aspects of Mathematical models.By **repeatability** we mean the degree to which any repeated calculation, on unchanged data and mathematical model, reproduces the result obtained first. By **accuracy** we mean the degree to which the difference between the correct and the obtained result is acceptable in the case at hand. By **efficiency** we mean the average consumption for any calculations of any resources needed for these calculations, such as energy, storage and processor. By **speed** we mean the average waiting time until result provision. By **availability** we mean the degree to which the services provided by the application are ready for use even if the conditions of the computing environment vary (temperature, humidity, power outage, power oscillation, maloperation, maintenance, attacks, natural disaster). By **scalability** we mean the degree to which the application can be adapted to revised user numbers per hour, data to be stored or operating ours per day. By **restart ability** we mean the degree to which the application can restart after a shut down due to maloperation, accident, attack or natural disaster.

To some extent, one obviously is free to consider some of these quality aspects as part of the explication of the term model. For example, Thalheim has incorporated repeatability into the explication of model. From an informatics point of view, this is a quite reasonable approach. However, at the time Stachowiak wrote his book about modeling, that sort of use of models was quite rare and that might have been his reason not to mention that aspect of models.

## VI. The Indispensability of Models

Here I aim at contributing to the discussion by Mayr and Thalheim in their section “Opportunities for Progress” in [9], where they say, that “… it is impossible to imagine engineering without models …”. In my view, they however, do not really get to the fundamental reason for their claim.

It is well known that Human perception is limited both with respect to quality, i. e. what can be perceived at all, and with respect to accuracy, i. e. how well things can be perceived. It is therefore obvious, that humans never can be certain about having perceived the world, or parts of it, the way these “really are”. While at his time this information might not have been generally available Nietzsche [3] added to this the insight, that perception is knowledge and information infused. His perspectivism actually got quite popular in his time but seems not so anymore. His alleged sympathies to the German Hitler fascism might have contributed to that. Nietzsche went even further by saying that not only perception always is infused by a certain vantage point. He also said that no all-inclusive vantage point exists.

If one takes for granted the utterance ascribed to Socrates “I know, that I do not know”, then one, would consider human knowledge as principally limited and hence subject to permanent change. Fleck [4] has pointed out that scientific facts have a genealogy and thus depend on a scientific community, to acknowledge them. Therefore scientific truth has a hidden parameter, i. e. the scientific community providing it with the status of truth.[9]

9 The deplorable mistreatment of Ignaz Semmelweis [3] bares witness of the cost, potentially imposed on those,

Long before Fleck, Hegel[10], in his preamble to [5], has said that the notion finally obtained of the whole is not the whole itself. That certainly is universal truth about human attempts to acquire knowledge. His insight can even be extended: the notion used to understand an item is different from that item itself. Therefore, if any knowledge is acquired in the way of thinking and in particular by scientific or engineering methods, then a model is used. That model then consists of the actual brain content used to get the desired information, together with any written, drawn or spoken signs used for that purpose. If the knowledge in question is supposed to be the common possession of several different human individuals, then, by current state of knowledge about these things, the brain content mentioned before, needs to be substituted by signs of the kind mentioned before. Clearly, perhaps together with some tutoring, that would enable each of the individuals mentioned before to rationally construct the knowledge in question. The use of models, when humans acquire knowledge about the world, is indispensable. This applies regardless of whether one knows or wants it.

By the way, later, in the same text, Hegel says two rather deep and things, that I share: 1. knowledge only is real as science or as a system and only can be represented as such; and 2. it must be claimed, that the truth is not like a stamped coin, that finally can be delivered and being put into ones purse. Since one of the key characteristics of science is its paradigm of adherence to the better argument, science never can be certain of having uncovered a final or objective truth.[11]

Remain a skeptic concerning any claim about science's ability to uncover an objective reality, actually is quite in line with Occam's razor [3]. Standard scientific principle is to accept a hypothesis only if it sufficiently simple and complies with the relevant observations that can be made. With regard to whether or not there is an objective reality there are two hypothesis that particularly need to be noted. The first is, "everything is connected to everything". The second is "certain items we observe are not related to other items at all". The first hypothesis is extremely simple and is compatible with what can be observed, since certain connections between items might be of the kind, that cannot at all, or at least not easily or currently, be observed. It cannot be decided whether or not the second hypothesis is compatible with the observations that could be made because at any time new ways to observe items might come into use, that show that items we thought were not related to each other, in fact are related to each other. Since the extremely simple hypothesis "everything is related to everything" is compatible with what can be observed, it is the scientific stance to be taken when it comes to analyzing the perceived world. The consequence of this is, that whatever investigation is carried out concerning our world, when we single out a particular item and focus the related investigation on that item, then in fact we use it as a model. As we have discussed above it is a bit more complicated if several human individuals are involved.

Looking back at the software process stages as discussed above, then we conclude, that each stage corresponds to a model that has to be created. Using the terms descriptive and prescriptive, that already were used by Stachowiak [11], we conclude that the implementation is descriptive, i. e. describes how the original actually is. The other process stages are prescriptive, i. e. they describe how the original is supposed to be. This again shows the enormous amount of planning involved in current software production.

However convincing our arguments are or appear to be, they only are valid for software production that follows the software process paradigm. An alternative to that paradigm might be under development right now. Artificial Intelligence (AI) in the long run might have the potential to change software production in a way that abandons that paradigm. It seems that in a distant future, software production might perhaps be based on machines learning, in an AI fashion. The respective learning software "simply" would have to come up with statistics backed ways of how to respond to certain inputs. Potentially the whole software process then might be abandoned and one only could argue that machines would be creating statistical models not being understood by humans and only created and used by machines. Some, considering this potential future, might object to that on grounds of humans losing control over the software development. One, who thinks that way, should, however, not lose sight of the fact that the benefit of using tools lies exactly in this kind of complexity reduction in practice. The average human does not

who undertake to change the scientific truth of their time. The most recent case of Jan Hendrik Schön that science even nowadays is not immune against attempts to corrupt the process of creating scientific facts.

10 Please note that I am neither a philosopher nor a Hegel scholar. The translations used here are mine as well as the interpretations of his text fragments. Note in particular, that much of what he writes in his preamble according to modern standards of German language are very hard to understand at all.

11 Obviously, this does not mean, that a human individual cannot arrive at their final truth regarding some matter. In so far as it is final it simply is not scientific.

understand most of the tools they use and has no problem with that as long as the tools most of the time are really helpful.

## VII. Basic Ideas Concerning Model Engineering

As models and modeling are of vital importance for Informatics, Thalheim [12, 13], as it seems, was the first one to talk about modelology, a new discipline about creating, improving and working with models. We contribute to that by looking into an approach to engineer conceptual models. With regard to that, two tasks are key. Firstly, one has to find an appropriate model. Secondly, one has to find out how to best modify that model such that it fits current needs.

With regard to the first step, we propose the creation of libraries of appropriately labeled mathematical models. These labels, on the one hand, include handling information such as the authors, the date of creation, deployment, commissioning, retirement, error reports, enhancement reports, version number and experience reports focusing on the quality of service achieved with the model and in particular the gains to effort ratio and potentially a model using fee. On the other hand, the quantities occurring in the model, including their status as either being knowns or unknowns, the connectors used to connect quantities (functions, relations, equations, inequalities, algorithms, …) are provided as labels.

Clearly, on attempting to get or create a Mathematical model fitting any given requirements, one would look at models in the library whose knowns are similar to the inputs requested for the model to be created or whose unknowns are similar to the outputs of the model to be created. If a number of related library models would be found, a rational approach to choosing one of these as the starting point for development would be to assess the cost for 1. getting the input data; 2. adapting the knowns; 3. adapting the unkowns; 4. implementing new functionality as well as update already existing ones. Finally the time to project completion and the gains to efforts ratio of the development project should be estimated and made known to the deciders. Once these have finally given green light a choice can be made based on cost, gain or development time considerations.

## *References*

[1] Aleksandrov A., Kolmogorov A., Lavrent'ev M. Mathematics, Its Content, Methods and Meaning. Dover Publications, Inc. New York, 1999.

[2] Boehm B. Software Engineering Economics. Englewood Cliffs, NJ. Prentice-Hall, 1981.

[3] Encyclopedia Britannica. Articles on: analogy, Antikythera mechanism, Mathematical model, Nietzsche, Occam's razor, Semmelweis, time and twin paradox.

[4] Fleck L. Entstehung und Entwicklung einer wisschenschaftlichen Tatsache. Benno Schwabe & Co. 1935. New editions with Surkamp.

[5] Hegel G. Phänomenologie des Geistes. Bamberg, Würzburg, 1807.

[6] Kaschek R., Konzeptionelle Modellierung. Habilitation Thesis Klagenfurt. 2003.

[7] Kaschek R., A Calculation View at Conceptual Models. In: Mayr H., Thalheim B. (eds.). Fundamentals of Conceptual Modeling, Mini-Dagstuhlseminar at CBI 2024. Vol. xx LNBIP, Springer, 2024.

[8] Lakoff G., Johnson M. Metaphors we Live By. University of Chicago Press, 1980.

[9] Mayr H., Thalheim B. Conceptual Modeling: A Still Unfinished Saga. In: Strecker S., Jung J. (eds.): Informing Possible Future Worlds. Logos Verlag Berlin, 2024.

[10] Smirnow W. Lehrbuch der höheren Mathematik, Teil 1. Europa Lehrmittel. Haan- Gruiten, 2017.

[11] Stachowiak H. Allgemeine Modelltheorie. Springer, 1973.

[12] Thalheim B. Modellkunde (Modelology) I: Sytematik. In: Proceedings of Modellierung 2022, pp. 11 - 23. Gesellschaft für Informatik. Bonn, 2023. LNI, P-324.

[13] *Thalheim B. Modellkunde: kurz & knapp. In: Loeben C. (ed.). Modellkunde und Ägyptologie im Dialog. Kadmaos, 2023.*

About Author (s):

Image

[Type a quote from the document or the summary of an interesting point. You can position the text box anywhere in the document. Use the Drawing Tools tab to change the formatting of the pull quote text box.]