Modelling the World — of Trains and Brains

Hans-Georg Stork*

Many children, young and old, love to play with toy trains. They use miniature tracks to build a railway network (in their living room or cellar), mould landscapes complete with pastures and mountains and decorate them with entire miniature villages, railway stations and other miniature replicas of objects that they find in the real railway world. Trains, tracks, the landscape and the whole paraphernalia are reproduced as realistically as possible. But of course, it is not the real thing. It is much smaller, greatly scaled down. You can watch trains moving but you can't sit in one. It is (just) a *model* of the real thing.

Replica, representation, imitation and reproduction are terms that are occasionally used as synonyms for "model". A model emphasises the characteristics of an object, a physical system or a process, that are considered important. What is important depends on what you want to achieve with your model. Aspects that are considered secondary are neglected or ignored –taken away, "*abstracted*", so to speak. In this sense, a model is an *abstraction* of a given reality.

A model can take many forms, depending on the purpose it is intended to serve. It is basically a representation of one world in the form of another world, e.g. the real railway world is represented "in the form of"a toy railway world which, like the *real world*, is also three-dimensional. But apart from appealing to children, young and old, it is not of much use.

*(h-gATcikon.de), January 2024

However, there are also other representations of real railway worlds, perhaps more useful ones, e.g. two-dimensional maps that depict the structure of the network. They take the form of graphs with nodes and edges that can be labelled and annotated. Nodes and edges represent cities and routes respectively, where the labels are the names of the stations and the annotations may indicate the time needed to travel from city A to the neighbouring cities along the connecting routes. This allows you to calculate the total time needed to get from A to a non-neighbouring city B and to find the quickest route.

In fact, we find models in almost everything that people do for a living or just for fun. And depending on the type of trade, profession or occupation in which people are engaged, there are different types of models that underlie their work. Architects, for example, draw floor plans and all sorts of diagrams, and even make small 3D replicas of the houses they want to build. Engineers, no matter what they are designing or putting together, cannot do their work without referring to models - not only of what they want to create, but also of the context in which their products will be useful.

This applies in particular to a *software engineer* who has the task of programming a computer or several interconnected computers so that it or they behave in a desired way in a specific target domain, e.g. produce an output affecting this domain at the right moment or receive an input from it.

Modelling the target domain (a bank, the administration of a trading company, the manufacture of motor vehicles, an aircraft, a city's electric power supply, a dental practice, transport logistics, etc.) of a software system is no easy task. It requires thorough analysis and a high degree of abstraction. And often also sophisticated mathematical methods and tools.

There is a plethora of models in use in physics and other sciences. Such models are not necessarily meant to serve as

blueprints for something to be built or assembled. Rather, most often they are meant to explain and understand phenomena that can be observed in Nature or in an experiment, possibly assisted by instruments (e.g., microscopes or telescopes) that extend our sensory faculties. The motion of a real pendulum, for example, can be studied to a first approximation by abstracting (i.e. ignoring) the body of the pendulum (reducing it to a point mass) as well as the mass of the rod or string connecting the pendulum to the pivot, which is assumed to be frictionless.

The resulting model is a highly simplified version of the real situation. And a very abstract one at that. We can draw it, but we can't build it. We can, however, describe its behaviour in the form of mathematical equations that allow us to predict the position of the fictitious pendulum at a given time or the duration of a single oscillation. The object of study itself is reduced to a set of parameters (e.g., weight, length of the rod, and the force of the initial push) and the relationships between them. And we should bear in mind that although our predictions may be more or less correct, they do not correspond exactly to what we would measure if we observed the real pendulum. This is due to the abstractions we have made from the real pendulum.

We call this kind of model a "*mathematical model*". Mathematical models abound in the sciences: models of an atom or its nucleus, planetary models, models of chemical reactions, of epidemics or the weather, of long term atmospheric processes, of entire ecosystems, and many, many more. Indeed, mathematics is a powerful tool for describing and understanding the natural world, and it has proven to be incredibly effective in dealing with a wide range of physical phenomena. The sky is the limit.

However, we learnt from the pendulum scenario that we should not trust these models one hundred percent. Models never correspond one hundred percent to what is modelled. Otherwise they would not be models. (Jorge Luis Borges long ago

remarked on the "*tragic uselessness of the perfectly detailed map*" in a short story entitled "*Del Rigor en la Ciencia*".)

Abstraction and identifying the relevant parameters are tricky tasks. If not done properly or based on inaccurate data, it can lead to false predictions of what is going to happen in the real world. In addition, predictive models ought to be "*stable*", i.e. their behaviour must not become chaotic with small changes in the values of one or more parameters.

By converting mathematical models into software that runs on a computer, we obtain "*computational models*" (in terms of data structures and algorithms), simulations that can be used - via sensors and actuators - to control processes in the real world, with direct effects on what goes on there. With a little imagination, we can draw parallels between what such a computer does in its world and what animals and humans do in theirs.

The world of a computer is determined by its hardware (including sensors and actuators) and the software running on it (a computational model!). It cannot do more than what it has been programmed for. It cannot "see"(hear, smell, etc.) more than its model allows it to "see". Although it is not inconceivable that a suitably primed computer could construct the model of the world it is to control and interact with, its very existence depends on human ingenuity and engineering.

The world of a sentient being (animal or human) is determined by its sensory organs and nervous system, with the brain being the key component. It controls not only its body's action in its environment (exteroceptive) but also its body and itself (interoceptive and self-referential). On the face of it, the existence of such a being depends on certain acts of its ancestors but ultimately on natural laws that govern the self-organisation of matter at different levels of granularity.

Clearly, brain-body systems cannot do more than what they have evolved and grown to do - just as a computer cannot do

what it has not been programmed to do. They cannot gather and process more data than what they have evolved and grown to gather and process - just as a computer cannot accept more data than what it has been programmed to accept.

Computer systems are designed and built. Nobody designs and builds organic sentient beings (but synthetic ones may become the next "*frontier*"). However, the fact that they evolve and grow, does not preclude using the term "model" also for what goes on in the brain of such a being. Processes in the brain somehow mirror real-world processes, just as we expect a model to do. (Nota bene: it is the brain that does the mirroring, not the "real world". But: the tree would still be "i*n the quad*" even if no brain mirrored its presence.[1]) Conditioned in many ways it will abstract from the World (at large, with a capital W) and constantly update whatever features are necessary to cope with the World (at large). Again, we should bear in mind that human and animal brains can run into all sorts of pitfalls when constructing their models of the world. Many things can go wrong. In fact, we don't know exactly how brains construct their world models –but they do it.

It should be noted that the models that are active in human brains generate the models that humans can then impose on their computing machinery or make visible and tangible –like the toy railway or the floor plan. Through suitable programming, computers can be endowed with that same faculty. Many animals are also able to "implement" the models they live by in their "environment" ("*umwelt*"): birds' nests, beaver dams, beehives and much more - presumably without realising what they are doing.

[1] An allusion to the Irish philosopher Berkeley's (1685-1753) assertion ("*esse est percipi*") that existence hinges on observation: "There was a young man who said "God / Must find it exceedingly odd / To think that the tree / Should continue to be / When there's no one about in the quad."

We are now leaving the S&T (Science and Technology) sector

When talking about the world in relation to a computer or a sentient being, the questions arise: (1) "*What is a world?*". And (2) "*What is the World?*"(Note the difference in spelling!) Questions that have kept philosophers busy since time immemorial. The answer to question (1) should be obvious by now: a world is what can be perceived by a computer or an organic brain, in terms of relationships between objects in the World, its dynamics and cause and effect structures. What can be perceived is determined by a model of the world - given either by the hardware and software of a computer or the neural networks in a brain.

What about the World? Unlike the pre-Socratic Greek sophist Gorgias (483–375 BC) , I believe that there is a World, that it is real. It exists and so do I. Surprisingly, almost two and a half thousand years later, a young German philosophy professor, a certain Markus Gabriel (1980 -), recently claimed that he had discovered that the World does not exist, that there is no World. I don't know what prompted Gorgias to make his outrageous and obviously provocative statement. Professor Gabriel, on the other hand, explains his insight in great detail in a bestseller[2]. After reading (a few chapters of) his book I still believe there is a World. He failed to convince me as he seems to get caught up in a battle of words (terms, concepts, notions, you name it) and Universal Quantifier paradoxes about the definition of "world" and "existence" that does not match my understanding of intellectual rigour and logical precision.[3] (But

[2]Warum es die Welt nicht gibt, Ullstein Buchverlage GmbH, Berlin, 2013; Why the world does not exist, Polity Press, 2015 (no wonder such a title sells well!)

[3]Time could provide another sophistical argument for the non-existence of the world: The past no longer exists, the future does not yet exist and the present is an infinitesimal value.

the story is entertaining and even edutaining, stimulating ideas, both consenting and dissenting.)

So, what do I understand by "*World*"? The Austrian philospher Ludwig Wittgenstein (1889 - 1951) begins his *Tractatus* with the statement: "*The world is everything that is the case*". Wondering what it might mean for something to be the case I say quite simply and perhaps naively: everything that physically exists and is going on somewhere, be it around me, near me, far away (outside me), or inside me, including the organs and chemical reactions that enable me to perceive the world and act in it, including what is said or written, everything. A body with its organs, especially its brain, constitutes the physical substrate of a "*Self*" –an "inner world"- and allows one (in this case: me) to speak in first person terms ("*I*", "*me*", "*my*", etc.) of its *being-in-the-world*.

Unlike Descartes (1596 - 1650), who claimed that there is a categorical difference between a body and something he called a mind, I have no doubts about my existence and the nature of my Self. My Self is everything that goes on in my body in the form of electrochemical processes - perceived by my brain as thoughts, pain, pleasure and other emotions. I do not believe that I and the World around me exist as –for instance –silicon based processes in a giant machine (as in a fictitious book that Stanislaw Lem reviews under the title: "*Non Serviam*"[4]).

I assume that everything that happens in and around me can, in principle, also happen in and around everyone else. As already mentioned, it can even happen without anyone (humans or other sentient beings) being present. Therefore, the events in the world outside of me can be further divided into those that depend on my actions or the actions of other sentient beings and those that do not depend on such actions but could in principle

[4] *A Perfect Vacuum*, Stanisław Lem, 1971, Harcourt Publishers Ltd, a collection of fictitious reviews.

be observed and investigated by me or other sentient beings, i.e. beings that have sense organs and a nervous system. (Note: Observing and examining something in my external world - even from a great distance - can be considered a disturbance. However, it is unlikely that I can move a star in the Andromeda galaxy or calm a storm.)

What happens inside me can only be felt by me. It can have a bearing on the outside world, but it is not part of it, although I can be part of the outside world of another sentient being. Another being can –to a certain extent –examine my inner world but it will never feel what I feel and never have the same sensory experiences.

The outer world is huge, multidimensional and probably limitless, both in terms of its expanse and its subtlety (complexity?). It is so much more than I can perceive. In contrast, my inner world is at least spatially limited, limited by my body.

> *Definition: World (capital W) equals outer world plus inner world plus the relationships between the two.*

My ability to cope with this one World, depends on the capacity and capabilities of my brain-*cum*-body to access and communicate with the World. How does a limited brain-body system do this?

The answer is straightforward. We need only repeat what we have already stated above: The brain, in a kind of bootstrapping[5] process that abstracts from the World 'useful'(for coping) features, inevitably creates and updates the world model, that guides the behaviour of its body, its *being-in-the-world*. Since the World in which I exist includes my inner world, this model

[5]a term used in many contexts: https://www.wikiwand.com/en/Bootstrapping Here it refers to generating tools for whatever purpose. The process starts with basic tools which are then used to produce better tools, and so on, without losing sight of the actual purpose of the tool. In this case, the "tools" are the world models whose ultimate purpose is to make their "owner" fit for the World.

includes a "*self-model*". The overall model is contingent on many factors: phylogenesis (the history of my species), ancestry, onto-aand epigenesis (physical development after and before birth) and social environment (upbringing, education , etc.). Physically, it consists of the neural activity of my brain, also known as "*mental content*".[6] What can my "*mental content*"tell me (or: itself) about me and the outside world?

···and enter neo++(?)-Kantian terrirory.

"*What can I know*?" is one of the three basic questions[7], attributed to the German philosopher Immanuel Kant (1724 - 1804). He devoted a long and detailed treatise, his "*Critique of Pure Reason*", to answering this question. What troubled him was the state of metaphysics, a branch of philosophy concerned with problems such as "*What is the meaning of existence?*" (not of the word, but of the concept) or "*Why does anything exist at all?*". Questions, to which experience alone could not provide answers. He wrote about metaphysics:

> "*It has not yet had the good fortune to attain the certain scientific metho*d" (preface to the 2nd edition of his *Critique*).

Of the scientists who since the days of Copernicus had more or less successfully established natural laws he wrote:

> "*They learned that reason only perceives that which it produces after its own design; that it must not be content to follow, as it were, in the leading-strings of nature, but must proceed in advance with principles*

[6] As an aside, what we call "*consciousness*" may be described as my (Self's) ability to communicate with my Self insofar as it is represented by my self-model. My *unconscious* may not be represented in my self-model although it is in my World.

[7] "What can I know?", "What must I do?" and "What may I hope for?"

> *of judgement according to unvarying laws, and compel nature to reply its questions. For accidental observations, made according to no preconceived plan, cannot be united under a necessary law. But it is this that reason seeks for and requires. It is only the principles of reason which can give to concordant phenomena the validity of laws, and it is only when experiment is directed by these rational principles that it can have any real utility.*"(ibid.)

Kant did not know how we use the term "model" today in different contexts of science and engineering. But of course, he could have formulated the above description of "*scientific method*" as "*modelling*" the phenomenon in which a scientist is interested, and then letting Nature herself, either by observation or by experiment, confirm or disprove "*what reason produces according to its own design*". (Read: *design = model*)

He regarded the natural sciences as exemplary for principled judgement and wanted to place metaphysics - '*a purely speculative science, which occupies a completely isolated position and is entirely independent of the teachings of experience*'- on an equally solid foundation by clarifying the very nature of principled judgement.

In a nutshell: Kant believed that all we can know about the World is either through experience - *a posteriori* - or through pure thought (reason!) - *a priori* - independent of any experience. What he calls "*judgements*"(or "*propositions*") are in fact his "*atoms*"of knowledge (reminiscent of Wittgenstein's "*what is the case*"). He differentiates between "*analytic*"judgements where the predicate says no more than the subject already implies ("*a quadrilateral has four sides*"), and "*synthetic*"judgements where the predicate adds facts or insights not already implicit in the subject ("*a quadrilateral is topologically equivalent*

to a circle"). Judgements can be either true or false, *a priori* or *a posteriori*. Metaphysics is the domain of *synthetic judgements a priori* - judgements that are not grounded in experience. In his "*Prolegomena to Any Future Metaphysics*", a sort of "popular science"exposition of the *Critique*, he poses the key question: *How are synthetic propositions a priori possible*? and explains why this question is important:

> *Metaphysics stands or falls with the solution of this problem: its very existence depends upon it. Let any one make metaphysical assertions with ever so much plausibility, let him overwhelm us with conclusions, if he has not previously proved able to answer this question satisfactorily, I have a right to say this is all vain baseless philosophy and false wisdom.*

Kant does concede '*that all our knowledge begins with experience there can be no doubt*'(*Critique*, Introduction) but continues: '*it by no means follows that all arises out of experience*'. The next section is therefore entitled:: *The human intellect, even in an unphilosophical state, is in possession of certain cognitions 'a priori.'*

Of course, Kant could not have been familiar with a science that began to flourish almost two hundred years after his death: neuroscience. He knew nothing of Darwin's discoveries, of the eons of evolution of life on earth, nothing of genetics, nothing of robotics.

One wonders whether, had his mind been active two hundred years after his time, he would have come to the same conclusions regarding the distinction between *a priori* and *a posteriori* knowledge or even, to the same definitions. (And it would be interesting to learn what he would say about modern attempts to impart "knowledge" to robots and other computerised devices. What are a robot's *a prioris* and *a posterioris*?)

Is what he sees as *a priori*, independent of any experience, not in fact the result of millions of years of experience of our non-human ancestors, which has been reflected in the structures of the brain and its auxiliary organs? Must everything be understood as a posteriori? We are not born as blank slates, an assertion that seems to be supported by Kant's claim '*that even in an unphilosophical state, the human intellect is in possession of certain cognitions "a priori"*. The newborn human brain is already equipped with a basic model of the World and so is the brain of any newborn sentient being.

These remarks in no way diminish Kant's merits as the first philosopher (to the best of my knowledge) to question metaphysics for its lack of rigour, essentially ending a long scholastic tradition in which speculation and dogmatism reigned supreme. Metaphysics, as it had been known until then, was no longer seen as useful for explaining the World. (This may not have been wholeheartedly affirmed by many philosophers to this day.)

In fact, Kant recognised that modelling, as an interplay of (pure) reason and experience, has not only greatly advanced the natural sciences (where mathematics, replete with *synthetic judgements a priori*, is predominant), but also leads to great progress in understanding the World (in all its dimensions):

> *It has hitherto been assumed that our cognition must conform to the objects; but all attempts to ascertain anything about these objects a priori, by means of conceptions, and thus to extend the range of our knowledge, have been rendered abortive by this assumption.*
>
> *Let us then make the experiment whether we may not be more successful in metaphysics, if we assume that the objects must conform to our cognition. This appears, at all events, to accord better with the possibility of our gaining the end we have in view, that is*

> *to say, of arriving at the cognition of objects a priori, of determining something with respect to these objects, before they are given to us.*
>
> *We here propose to do just what Copernicus did in attempting to explain the celestial movements. When he found that he could make no progress by assuming that all the heavenly bodies revolved round the spectator, he reversed the process, and tried the experiment of assuming that the spectator revolved, while the stars remained at rest.*
>
> *We may make the same experiment with regard to the intuition of objects. If the intuition must conform to the nature of the objects, I do not see how we can know anything of them a priori. If, on the other hand, the object conforms to the nature of our faculty of intuition, I can then easily conceive the possibility of such an a priori knowledge.* (preface to the 2nd edition of his *Critique*)

This is what has become known as Kant's "*Copernican turn*": our cognition must not try to adapt to the objects in the World, but conversely, we must accept that the objects in the World conform to our faculty of intuition. A bit further on he makes '*the new method of thought which we have adopted*'even clearer as '*based on the principle that we only cognize in things a priori that which we ourselves place in them*'. In other words, it is our world model, which is located in the brain, the organ of cognition, that determines what we can know about the World. But note that what '*we place into things ourselves*' may or may not be conducive to a better life. Our models of the World may or may not be viable, depending on what we *put into things* - as we have already noted above when discussing our brain's construction of world models.

Incidentally, scientists (or "*natural philosophers*" as they were called at the time) had long recognised, as Kant also acknowledged, that - in the words of the Scottish philosopher David Hume (1711-1776) - "*Sounds, colours, heat and cold, according to modern philosophy are not qualities in objects, but perceptions in the mind.*" (Hume, *A Treatise of Human Nature.*) The things we see are not the way we see them. The things we see are what our brain makes of them. The "*thing-in-itself*", as Kant called it, is and remains hidden from us. We all live in our own model of the World, our own "*virtual reality*". We cannot know what "*reality*" really is. We can penetrate deeper and deeper into it (e.g. with high-tech help), but we will never reach the bottom.

Hans-Georg Stork, 1/2024