From Calculus to Language Game: The Challenge of Cognitive Technology

Christoph Durt[1]

Abstract

Cognitive technology is an increasingly important form of technology that can deal with meaning by either replicating or simulating human cognition. Cognitive technology can make use of information technology, but it strives to go beyond mere information processing by recognizing, changing, and creating meaning. This presents us with a two-sided challenge: On the one hand, cognitive technology is challenged to "understand" meaning in ordinary language. And on the other, it challenges us to rethink fundamental questions of human cognition and sense-making. Both challenges demand a better understanding of the difference between the technical transformation of symbols and the understanding of meaning in the ordinary sense.

After explaining the topic in relation to both the insights and the limitations of the reflections by Turing, Searle, and Heidegger, this paper primarily builds on Wittgenstein's contributions to a better understanding of the difference between two conceptions of meaning and their implications for technical replication and simulation. The paper shows that Wittgenstein developed his early calculus account of meaning into that of language games and that language games not only come in many different varieties, but are also much more flexible than calculi. Of particular interest will be the difference between rigid and creative rule-following. Creative rule-following involves an intricate interplay of very different bodily, mental, and cultural constituents, so that its simulation is not merely a technical problem but also requires clarification of a number of profound philosophical questions. It will become clear that the challenge of cognitive technology shows up at unexpected places and that is much bigger than usually assumed.

---

[1] Christoph Durt, Institute of Philosophy, University of Vienna, Austria; christoph@durt.de/www.durt.de

1. The Forgotten Meaning of Turing's Question

Prominent AI pioneers started from the premise that "every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it" (McCarthy et al. 1955, 1). The claim is not that machines are "really" intelligent, but that they can simulate any feature of intelligence. Correspondingly, Alan Turing's (Turing 1950) famous operationalization of the question "Can machines think?" allows for an affirmative answer regardless of whether the machine in question really is thinking. The machine may well pass the test by merely simulating a human participant of a dialogue in such a perfect manner that, to the conversation partner, the machine becomes indistinguishable from a human. What today is known as the "Turing test"[2] does not test whether a machine can think but whether it can simulate human thinking within given limits.

That eventually some machines will pass instances of the Turing test is likely because the test relies on a—fallible—human or jury evaluating the machine, is relative to another—imperfect—human being, and has a—limited—time frame. Thus, a simulation that is not perfect but still very good may be able to pass the Turing test. The claim that "every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it" (McCarthy et al. 1955, 1) may on a weak interpretation just mean that there can be a description of every feature of intelligence that is precise enough to enable some—more or less adequate—simulation. But the suggestion in this citation—and in many discussions of AI, information technology, and cognitive technology—is that every aspect of intelligence can be described in precise terms, i.e., terms that are unambiguous and rigid enough to allow for the construction of a machine that can exactly simulate the respective aspect of intelligence.

The idea that a machine could simulate human thought in a perfect or close to perfect manner was not only advocated by the early and later proponents of AI. Often, the affirmative answer is simply presupposed, and instead another question takes over: Does perfect simulation amount to *real* intelligence? The problem with this shift is

---

[2] Turing himself spoke simply of a "test" in "Computing Machinery and Intelligence" (Turing 1950) and on other occasions, e.g. in a BBC radio interview (Turing 2004a, 495). As discussed below, in the latter paper Turing actually suggested two versions, both of which differ from the test known under his name today (Turing 1950, 434, 446).

not only that the question what 'real' means demands a stance on precisely the hard ontological problems Turing attempted to circumvent with his test, but also that that question itself easily stirs up strong intuitions that can inhibit a clear understanding of the issue. Yet the same presupposition is also at work in a number of arguments concerning AI, including some of the very thought experiments intended to refute the claim that a digital computer could understand[3] ordinary language.

For instance, in the "Chinese Room Argument," a person in a room who doesn't understand Chinese consults a set of rules, which, in combination with "some instructions" (Searle 1980, 418), is taken to be sufficient for transforming some input (in Chinese characters) into the required output (in Chinese characters). The thought experiment builds on the assumption that this is possible, and then proceeds to the intuitive claim that the person is not really understanding Chinese. In the buzz surrounding the Chinese Room thought experiment, it is often overlooked that it doesn't address Turing's question but rather presupposes an affirmative answer. But if sets of rules and instructions like the ones assumed are to be found anywhere outside of philosophical thought experiments, then probably it is in books of fairy tales. While it is possible to somewhat imagine a book comprising such rules and instructions, it is not at all clear how such a book is possible any more than a talking teapot is possible. More concretely, Searle's argument presupposes but does not make clear how it is possible that conversations in natural language are rule-driven in such a way that they can be reduced to mechanical transformations of symbols, such as mapping input shapes in a book on the basis of rules and mechanically producing output.

In the context of Searle's argument, this may be considered unproblematic as long as it is taken to be merely a *reductio ad absurdum* of claims of "strong AI," according to which a computer would exhibit understanding that is in fact in no better position than that of the person in the Chinese Room. Since a *reductio ad absurdum* makes use of the same presuppositions as the contention it rejects, and then exhibits the absurd consequences of that presupposition, in the end

---

[3] Instead of giving a—necessarily simplistic—definition of 'understanding,' I intentionally leave the definition of this concept open-ended. The term will be introduced by example and further elucidated it in the course of the paper. Of special relevance will be the distinction between rigid and creative rule-following, the latter of which necessitates understanding in a sense that is not required for rigid rule-following.

it does not need to endorse the presupposition. But beyond showing that there is something wrong with the presupposition, rejections of this kind are prone to bring up other confusing problems rather than clarifying the problem at stake. In the specific case, Searle may expose that symbol processing according to a given set of rules does not necessarily amount to "real" understanding.[4] However, in doing so he does not refute but rather fosters the idea that it is possible to perfectly simulate understanding behavior.

Popular culture, too, often presupposes an affirmative answer and then asks whether this amounts to "real" cognition, understanding, feeling, consciousness, or personhood. How intriguing such questions can be is shown in films such as *Blade Runner*, *The Matrix*, *A.I.*, *Her*, and *Ex Machina*. Whether the simulation amounts to a "real" mind or not can be an exciting question even for popular audiences. What really stimulates the fantasy, however, is the additional question concerning the consequences of machines that look and behave like humans. All of the above movies, and many more, thus entertain different possible consequences of the fictionally presupposed fact that human intelligence can either be replicated or simulated by a machine in a nearly indistinguishable manner. Fictional presuppositions may bestow the possibility of nearly perfect simulation or replication with intuitive plausibility, but clearly they do not prove that possibility.

Since the Turing test restricts the tested interaction to symbolically mediated exchange, *prima facie* it seems plausible that a very complex program could be good enough to prove indistinguishable from a human. The idea of an exchange of text messages may appear to be a rather limited task. Notwithstanding this, in actual practice the limitations of that task can turn out to be less restricting than imagined and lead to further complications and challenges. What Turing called the "imitation game" as well as his first version of what today is called the "Turing test" involved a male participant pretending to be female (Turing 1950, 434). The female participant is supposed to be truthful about her gender, and either participant attempts to convince the jury that she or he is of the female gender. It is up to Turing's biographers to consider whether and in which way his personal history may have made him especially aware of the possible intricacies of this task.

---

[4] In other words, the Turing test only shows that the machine is, within the limits of the test, indistinguishable from humans. Turing certainly was aware of this limitation as he had explicitly designed the test this way.

We don't need to go that far to recognize that tasks like this can turn out to be quite complicated even for humans, and it is easy to underestimate the intricacy of the extended test in which a machine pretends to be a human. It is possible to imagine questions about feelings and emotions and personal history, and Turing also mentions requests to write poetry. Furthermore, the test not only allows for questions and answers but can also develop into a playful exchange between the candidates and the jury. Open-ended discussions and a free exchange of feelings and ideas are possible, and all kinds of emotions could develop between the participants and the jury. In more personal settings, one could imagine exchanges of erotic and sexual phantasies or even the development of some kind of love relation, such as portrayed in *Her* and *Ex Machina*.[5] Any technology for mastering such exchanges would need to be able to participate in language games that are potentially free, open-ended, and emotional in intricate ways.

The measure the Turing test is not only human intelligence, but also human limitations and stupidity. A human jury may sometimes be tricked into believing they are interacting with a human, as demonstrated already in the mid 1960ies by the simple computer program *ELIZA* (Weizenbaum 1966; Weizenbaum 1976). Furthermore, the "competing" human may make mistakes that suggest she or he is a machine. But tricking the jury in a Turing test is much harder than in the context of *ELIZA*, which imitates a Rogerian psychotherapist for the very reason that this is relatively easy because "much of his technique consists of drawing his patient out by reflecting the patient's statements back to him" (Weizenbaum 1976, 3). Furthermore, a machine pretending to be a human would also need to make "human" mistakes to convince the jury. That is harder for the machine than one may think since those mistakes wouldn't come naturally to it. The setup of the Turing test moreover excludes forms of intelligence that cannot be conveyed to the jury. Turing recognizes that this may represent an unfair disadvantage for the machine, which may exhibit a different kind of intelligence (Turing 1950, 435).

While text messages themselves are restricted to a rather narrow form that avoids difficulties such as imitating the human voice, in another respect the same restriction makes them even more difficult to simulate. Due to the restriction, text messages leave many things unsaid

---

[5] The latter explicitly refers to the Turing test and plays through versions of it, some of which involve feelings or simulation of feelings of bondage, love, incarceration, mistreatment, and hate.

but imply them. The participants add unspoken context—and sometimes only one of them and not the other(s). This is something especially difficult for the computer who cannot naturally add much of the context that is self-evident or seems appropriate to humans who share a culture. Even for humans, exchange via text messages and emails can be prone to misunderstanding and is in fact often avoided when it is important to avoid misunderstandings. This may give the impression that text-message exchanges are rather superficial, but the Turing test does not exclude profound exchanges. Because in text exchanges so much context needs to be inferred and for the other reasons given above, the restriction to exchanges of symbols makes the Turing only *prima facie* a straightforward target of machine simulation. It may actually be easier to build a robot that outwardly looks and behaves like a human than a machine that can engage in intricate text-message exchanges.

If one counts artificial organisms as machines, then one could "simply" replicate human physiology, and the question whether the resulting being could think would be similar to that whether a human body could think.[6] Turing thought that, for the purpose of his test, "there was little point in trying to make a "thinking machine" more human by dressing it up in such artificial flesh" (Turing 1950, 434). In *Intelligent Machinery*, however, Turing emphasizes the importance of things such as culture, community, emotion and education for thinking (Turing 2004b, 430–431). Today, many proponents of AI attempt to replicate some of the ways cognition works in nature and promote neural networks or other biological models. Cognitive technology does not have to rely on classical information technology.

Notwithstanding the above, the main question of this paper is relevant for both computational *and* biological models of thinking. The question is not a yes-no question as to whether simulation of thinking is altogether possible or not, and I will not point to any alleged instance of intelligent behavior that could in principle never be replicated or simulated. Instead, the question that will be asked is, to what extent does cognitive technology go beyond technical symbol-processing? The meaning and context of this question will be explained in the next section. Section 3 then elaborates this topic further by considering

---

[6] Similarly, but in a different context, Wittgenstein asks: "Could a machine think?—Could it be in pain?—Well, is the human body to be called such a machine? It surely comes as close as possible to being such a machine" (Wittgenstein 1998a, §359).

Wittgenstein's shift from a calculus account of meaning to that of language games, and section 4 contends that to fully participate in open-ended language games, cognitive technology would have to replicate or simulate creative rule-following. Drawing on some of Wittgenstein's thoughts on rule-following, I will explain why this is an enormous challenge.

2. Heidegger's Old New Technology, Information Technology, and the Challenge of Cognitive Technology

The development of technologies such as the telescope, steam and combustion engines, trains, cars, and planes is carried further in information technological devices such as personal computers, robots, the Internet, iPhone and self-driving cars. From a wider perspective, however, all this is only the most visible part of an ongoing and next technological revolution. Contrary to the tendency in ordinary talk and major accounts in the philosophy of technology, technology cannot be reduced to its physical manifestations. We need to look beyond technological devices to understand the nature of technology.

One way of looking at previous developments in technology is using Martin Heidegger's distinction between "old" and "new" technology (*Technik*) (Heidegger 2000). Heidegger's preferred examples for old technology are the windmill and traditional agriculture. Under the header of "new technology," he lists items such as hydroelectric plants, coal mines, the mechanized food industry, and nuclear technology. All of the above examples are not merely meant as technological things that change the world, but as expressions of a more fundamental, technological, manner in which humans relate to nature. While old technology builds on experience and tradition, new technology puts science into use. In the framework of modern science, nature is divided into forces and resources, which are understood as a standing reserve (*Bestand*) ready to be ordered (*bestellt*) by technology. Relating to nature within this framework (*Ge-stell*) is the new "challenge" (*Herausforderung*) humans find themselves in. New technology "challenges" (*fordert*) and "exploits" (*fördert*) nature in its new division in forces and resources. Nature, in turn, "reports itself in some way or other that is identifiable through calculation and […] remains orderable as a system of information" (Heidegger 1977, 23). The concept of nature itself is changed through technology and with it the concept of human existence.

Considering that Heidegger wrote that new technology and science together treat nature as a system of *information*, that he was interested in cybernetics, and that he showed foresight with respect to other technologies,[7] it is an unfortunate failure that he never put information technology in the focus of his investigation. As Heidegger himself points out, information and calculation are central players in the new modern scientific picture of the world. The world of science is represented by numerical information, which allows for the calculation of future states. The development of a technology that is explicitly focused on calculation and information allows for increasingly efficient dealing with nature as conceived by modern science and technology. Information technology is most apt to further advance science and "new" technology to new levels.

'Information technology can even serve as an ideal example for technology conceived as a system of techniques and methods of transforming entities. According to Heidegger's own account, "new" technology already essentially restricts its dealings with nature to one kind of causality, the *causa efficiens*,[8] in spite of the complexity of modern technology and its tight and symbiotic relation to disinterested knowledge or *epistēmē*. One does not need to go much beyond Heidegger to realize that the technical transformation processes are further limited in information technology, the core of which, as already Turing had shown (Turing 2004c), is reducible to simple operations on strings of binary states, such as 0 or 1. Information technology consists in its core of the simplest possible technical transformation of the simplest possible states. Information technology is the purest form of technology.

'Information' is frequently used in very different ways, which are often confused. In the current context we need to distinguish at the very least between (1) an ordinary sense of information as meaningful facts, information[o], (2) the sense of information as a numerical representation of nature, information[r], and (3) a narrow technical sense, information[t]. Speaking simplistically, information technology only

---

[7] As early as 1939, Heidegger wrote—critically—about technologically-produced (human) life, i.e., in today''s vocabulary, biotechnology (Heidegger 1976, 257). In 1975, he claimed in a TV interview that the impact of biotechnology (*Biophysik*) would surpass that of the atomic bomb (Rüdel and Wisser 1976, running time 37min., 55 sec.).

[8] Following Heidegger, "old" technology requires that an agent craftily bring together *causa materialis*, *causa formalis*, and *causa finalis*, the latter of which is now even excluded from the modern concept of causality.

deals directly or immediately with information^t. Information^t consists of distinctive states that can be transformed in rigid processes and without any understanding of what they are supposed to represent. This sense of information can be associated with "data," including "big data," but the processes at the core of information technology are neutral with respect to what they may represent. To turn information^t into information^r and information^o, the states need to be interpreted, and it must be understood what they stand for. Usually, this is done by the user, who may erroneously attribute the understanding to the information^t processing machine—a frequent cause of the attribution of "intelligence" to where there is only mechanical calculation. For instance, it may seem possible that there could be a "novel-writing machine" (Dennett 1992, 107). Such a machine would not need to be able to understand its product—the black dots on white paper that make up a printed novel. It is easily overlooked that there would be no novel if the dots could not be interpreted and understood by somebody. This somebody is, of course, not a mere body but somebody who has a perspective on the world from which they can interpret and understand the world of the novel. In this sense, Dennett presupposes precisely what he sets out to disprove: an interpreting and understanding self.[9]

The concept of nature as a system of information^r antecedes the rise of information^t technology by centuries, which continues to deal with nature in the "new" technological way. However, there is also a fundamental difference. Much of the activity of "old" and "new" technology consists in efficiently moving and transforming energy and resources to make and move physical objects. Information technology, in contrast, does not essentially concern the movement or transformation of material entities, but rather that of immaterial entities such as symbols. Information technology developed out of and makes use of what Heidegger calls "old" and "new" technology, but in contrast to these it is no longer essentially but merely contingently related to the material world. Information^r can be nearly seamlessly transformed into information^t and thus mechanically processed. This may have led Heidegger to think that information technology would play only an auxiliary role for "new" technology. Information technology doesn't have to be related to the world in that way, however. Today it is clearer

---

[9] Dennett seems to overlook the need for interpretation and understanding of the ink dots on the paper when he argues that the self could be produced by mechanical processes analogous to those that produce a protagonist in the novel written by a non-understanding novel-writing machine (see Carr 1999, 124).

that information technology can also be used for the simulation of cognitive processes that have a much more intricate relation to the world. Furthermore, there may be other technological means, such as artificial neural networks, that may have uses beyond information^r.

In spite of constituting a new kind of "immaterial" technology, information technology is really only the beginning of the next technological revolution. The protagonist of the next technological revolution is also the protagonist of this paper: cognitive technology. Cognitive technology deals with meaning as it is understood by humans. Like information technology, cognitive technology may use information^t processing, but it is not defined by it. In both cases, information^t processing is not an end in itself but only a means of processing information^o, or, in the case of cognitive technology, of dealing with all kinds of meaning. We should not assume that all kinds of meaning are reducible to information^o, let alone information^r and information^t. The relation between information^r and other kinds of meaning will be considered further in the subsequent sections.

What Heidegger calls "old" and "new" technology, including biotechnology, is interwoven with our ways of thinking about and relating to the world, but the impact of its material manifestations on cognition is mostly indirect. The printing press did not create new religious or political ideas, although it did contribute to their spreading and favored the spread of some ideas over others. Other possible changes to cognitive abilities may be due to genetic modifications, which may one day lead to superhumans who exhibit advanced cognitive abilities. Or, much less excitingly, genetic modifications may contribute to a reduction of nutrients and microorganisms that are needed for a healthy digestive system essential to the proper functioning of organs such as the brain, and consequently interfere with the cognitive abilities of part of humanity. Such influences may profoundly alter human cognition, but in indirect ways because they do not alter the content of cognition (unless memories or ideas are "implanted" into brains). Information technology by itself, too, mostly has only an indirect impact on cognition; e.g. through the interfaces and methods of use of digitally stored and modified information, by favoring certain kinds of information processing and exchange, or through adaptations by the users to the technology, apart from all the other consequences for the world we live in. Already such indirect impacts can profoundly alter human cognition and change the course of history in unpredictable ways.

Like "old," "new," and information technology, cognitive technology gives rise to numerous ethical and social concerns. For instance, ethical and social issues with regard to robots include safety issues such as those caused by error or hacking, responsibility for or of the robot, privacy concerns, as well as the social and environmental impact of robots (Lin 2012, 7–11). In addition to such ethical and social concerns, cognitive technology also gives rise to a whole new dimension of issues. These derive from the fact that it engages not only indirectly but also directly with human experience and cognition. Human experience and cognition is immediately affected by technologies such as augmented and virtual reality. Already today, online profiles allow the construction of altered identities that may lead to an "onlife personality," a "hyperconnected reality within which it is no longer sensible to ask whether one may be online or offline" (Floridi 2015, 1). We are beginning to get used to the fact that bots, robots, and other artificial systems and devices somehow autonomously interact with humans. The traction of such interaction increases exponentially when the devices or systems become able to autonomously navigate and manipulate the space of human meaning and reason. The philosophical foundations of this technology—cognitive technology—is what this paper is concerned with.

Understanding the impact of cognitive technology on human cognition requires research on fundamental philosophical questions concerning meaning, understanding, the human mind, reality, culture, and many more. The ambiguous title of this paper and section, "the challenge of cognitive technology," is to be understood in both the objective and the subjective sense: Cognitive technology is challenged by fundamental philosophical issues and it challenges us to think about fundamental philosophical issues. This paper focuses on one aspect of this two-folded challenge, namely rule-following in symbolic language. In the following sections, I will mainly draw on Ludwig Wittgenstein, who is especially suited for this endeavor. His two main general accounts of meaning, that of language as a calculus and that of language as a game, show how most of use of language goes beyond rigid transformations of information^t and information^r (and often beyond exchanges of any information of any kind of information), and thus throw light on the challenge of cognitive technology.

3. From Calculus to Language Game

While Wittgenstein did not put forward two completely antagonistic philosophies, his early and his later accounts of language and meaning differ in ways that concern key questions of this paper. On the one hand, Wittgenstein presents in the *Tractatus* a technical account of language, according to which language is strictly rule-governed. The rules are rigidly applied and obeyed. On the other hand, the later Wittgenstein offers us a multitude of considerations that demand a much more complex view, according to which language is characterized by flexible rules and creative rule-following. This does not force us to draw normative consequences for ethics and politics, and in this respect it may be true that "Wittgenstein leaves us with little more than a passive traditionalism" (Winner 2001, 16). But Wittgenstein's thinking on the above topics is radical and revolutionary. This paper attempts to show that his later account of meaning can lead to a radical reconsideration of AI and its importance for the philosophy of technology.

I will argue that Wittgenstein never completely abandoned the idea that some language games can be described as calculi. For such language games, meaningful sentences can be mapped directly on entities in the simulation environment and essential moves can be simulated straightforwardly. Even here, however, Wittgenstein emphasizes that rules alone do not determine their application. In other words, something else must come into play, which may be connected to natural or cultural features of the rule-follower, and that may be much harder to implement in technology. Moreover, according to the later Wittgenstein, most language games cannot be described as calculi in the first place. They entail flexible rules and require creative rule-following. In such language games, the essential moves cannot be directly mapped and simulated.

Wittgenstein never claimed that machine intelligence would take over human intelligence, but his early account of meaning is akin to that presupposed by those who believe that machines will soon excel humans in general intelligence. The resemblance is not coincidental since Wittgenstein's early account was, on the one hand, influenced by Bertrand Russell and other contemporaries that have also influenced thought on AI. On the other, Wittgenstein's early account itself has influenced thought on AI, either directly or mediated through other philosophers such as the members of the Vienna Circle. Wittgenstein and Turing were contemporaries at Cambridge University and had discussions in 1937 and 1939 (Floyd 2017a, 6). It has been argued that

Turing's "anthropological" approach to the foundations of logic was influenced by Wittgenstein (Floyd 2017b, 103, 110), and that in turn "Wittgenstein's later philosophy was partly shaped in response to Turing" (id., 104). Wittgenstein read Turing's *On Computable Numbers* (Turing 2004c; Hacker 1990, 163), and although there is very little explicit discussion of the universal Turing machine by Wittgenstein (Wittgenstein 1988, §1096), it is clear that Wittgenstein thought about issues arising from that idea.

As is well known, Wittgenstein became the possibly sharpest critic of his own earlier account. Of the many relevant and profound continuances and differences, this paper concentrates on his shift from viewing language as a logical calculus to viewing it a plurality of language games. Wittgenstein pursued the calculus account of language from the *Tractatus* up until the 1930s. Yet he became increasingly critical of the view of language as a calculus. Along with other central tenets, Wittgenstein came to profoundly criticize, modify and reconceive the calculative account of meaning until the end of his life. He increasingly replaced the concept of calculus with that of language game.

The *Tractatus* had put forward a notion of language according to which everything that can be said meaningfully can be expressed clearly in a rigorous calculus. A calculus is a system of rules with which one can deduct or compute propositions from other propositions in a rigid fashion. Nonetheless, the *Tractatus* doesn't claim that there is nothing more to reality. Indeed, Wittgenstein attaches great importance to that which cannot be put into the calculus form of language. He famously claims, however, that we cannot say anything about what cannot be put in that form and must hence be silent about it. The concept of calculus has two pivotal parts. The first is that it comprises everything that can be meaningfully said in language. Already the very first statement in the *Tractatus*, "The world is all that is the case" (Wittgenstein 2001, §1), suggests the calculus account of meaning. That what is the case consists of facts that can be represented by propositions. Wittgenstein writes that "[t]he general form of a proposition is: Such and such is the case." (Wittgenstein 1981, §4.5). Looking back at this phrase, Wittgenstein writes in *Philosophical Investigations* that he may as well have written "[t]his and that is true" (Wittgenstein 1998a, §136). He points out that this amounts to saying: "we call something a proposition when *in our language* we apply the calculus of truth functions to it" (ibid.).

The seemingly innocent opening statement of the *Tractatus* in fact imposes a severe restriction. The statement recognizes only those sentences as meaningful that can be either true or false. It thereby forces the logic of an ideal language that features a binary truth function—"*our language*"—upon ordinary language. The early Wittgenstein's concept of language treats nature as something that corresponds to a system of information[r] that can be manipulated by means of logical operations. If language is described in this way, it is no wonder that every meaningful proposition looks like a truth function that can be part of a calculus. As the saying goes: If all you have is a hammer, everything looks like a nail. In contrast, the later Wittgenstein doesn't just propose one single tool for describing all meaningful language, but rather speaks of a whole tool-box:

> Think of the tools in a tool-box: there is a hammer, pliers, a saw, a screw-driver, a rule, a glue-pot, glue, nails and screws.—The functions of words are as diverse as the functions of these objects. (id., §11)

This and other remarks of the later Wittgenstein are sometimes misinterpreted as dismissing the concept of meaning altogether and as an attempt to replace it with a functionalistic notion of technical usage. But Wittgenstein does not claim that language can be reduced to some set of tools, or tools in general, or any other form of technology. Rather, the above citation suggests that tools can be very different and have a variety of different functions, and in analogy that there is a variety of words and that their functions can be seen in analogy to the variety of functions of tools. It would be erroneous to deduce from superficial resemblances between words that they are all meaningful in the same way. Furthermore, the constitution of the "tools" of language matters only in relation to their use; the use affords possible applications. Accordingly, we can distinguish two claims: (1) tools are not all the same kind of object, and (2) they cannot be understood without seeing them in the context of their use. Both (1) and even more so (2) bring with them a complexity that calls into doubt any attempt for a reductive account of tools, for instance in philosophy of technology. With regard to language, the tools analogy illustrates that there are many ways of saying something meaningful, all of which cannot be forced in a calculus.

Already in the notes he dictated to his class in Cambridge in 1933–34, known as *The Blue and Brown Books*, Wittgenstein explicitly

rejects his earlier idea that all of language can satisfyingly be described as a calculus:

> In practice we very rarely use language as such a calculus. For not only do we not think of the rules of usage—of definitions, etc.—while using language, but when we are asked to give such rules, in most cases we aren't able to do so. We are unable clearly to circumscribe the concepts we use; not because we don't know their real definition, but because there is no real 'definition' to them. To suppose that there *must* be would be like supposing that whenever children play with a ball they play a game according to strict rules. (Wittgenstein 1958, 25)

While Wittgenstein sometimes continues to speak of the "calculus of language" (id., 42, 65), this citation makes it clear that his idea of the use of language fundamentally shifted. Wittgenstein's term "language game" now begins to replace that of calculus. There is an intuitive difference between calculi and games, with one being easily associated with rigorous mathematics and the other with playful behavior. But the exact nature of the difference may not be clear right away. The crucial ambiguity is already inherent in the Latin root of the word calculus. *Calculus* can either mean a stone or piece used for calculating (e.g. in an abacus), or one used for playing (e.g. in a board game). In mathematical text books, games sometimes serve as examples to illustrate a calculus. Like calculi, games are rule-driven, and both calculi and games are embedded in a wider context of human purpose and behavior.

For Wittgenstein, too, the two concepts are not necessarily contradictory, and the concept of language game inherited core features from that of calculus. A key commonality is the importance of use for meaning, which he developed in a mathematical context, most prominently in the *Remarks on the Foundations of Mathematics* (Wittgenstein 1956, II, §80). These posthumously published investigations are mostly concerned with mathematics and logic and, different from what Wittgenstein seems to have originally planned, did not make it into the *Philosophical Investigations* (id., vii). But many important ideas developed in the *Remarks on the Foundations of Mathematics* were taken over into the *Philosophical Investigations*; for instance, in §80 he describes both language games and the calculus in terms of their use (id., II, §80). In this and many other passages he emphasizes both with regard to language and calculus that meaning is constituted by the *use* of language rather than the givenness of some

mental state or given thing. Meaning is due to some form of doing, namely acts of speech (*Sprachhandlungen*) (Wittgenstein 2005, 145).

In the *Philosophical Investigations*, most uses of 'calculus' point to the shortcomings of using that concept to account for language. Wittgenstein now assigns calculus a limited place within a changed framework. He speaks of calculi and language games in the plural, and explicitly rejects the idea that ordinary language can be sufficiently described as a calculus (Wittgenstein 1998a, §81). Speaking of language games in the plural still allows for the conceiving of *some* language games as instances of a calculus. Wittgenstein now uses instances of specific language games to illustrate specific and limited ways of sense-making, such as that of construction workers passing on building materials (id., § 2).

When the use of language is conceived as a calculus, it appears as if mechanically-operating machines such as today's computers could soon master language. Those who adhere to the calculus conception of language are hence prone to believe that the time is close when the human mind will be able to be saved on a hard drive, or that AI could become better than humans in understanding human language. A different philosophical view on language, in contrast, will result in a very different idea of what would be required to match or excel the use of language by humans. The philosophical view here not only impacts the concept of AI, but also the concept of human intelligence and understanding. In this respect, Wittgenstein's plural concept of language games is much more demanding than that of calculi. To participate in language games that cannot be described as calculi, a human or a machine must be able to do much more than apply a given set of rules. The next section gives examples for the capabilities involved in language games and argues that ultimately the capability of creative rule-following is required.

4. Rigid and Creative Rule-following

Rule-following in language games is generally surprisingly demanding. It stands in the context of a shared practice that is complex and presupposes shared physiological and cultural conditions. It not only consists in the rule-following that is actually exhibited but also involves a number of abilities:

> A being can be said to be following a rule only in the context of
> a complex practice involving actual and potential activities of

justifying, noticing mistakes and correcting them by reference to the rule, criticizing deviations from the rule, and, if called upon, explaining an action as being in accordance with the rule and teaching others what counts as following a rule. (Bennett and Hacker 2008, 256)

Although such abilities are learned, they seem so natural and self-evident for humans that they are easily overlooked. Their technical replication or simulation, however, presents a huge challenge. Yet there is an even more fundamental reason for why information technology as defined above involves transformation processes that can only be described as rule-following in a very limited and usually metaphorical sense. Bennett and Hacker rightly point out that "[c]omputers cannot correctly be described as following rules any more than planets can correctly be described as complying with laws" (ibid.). Analogously, the brain, too, by itself cannot follow a rule. To think otherwise would be to commit what Bennett and Hacker call the "mereological fallacy" (Bennett and Hacker 2003, 68; Bennett et al. 2007, 22). Not just one of its parts but only a whole being that has the respective abilities and is embedded in the right context can follow a rule. Unless a machine is a being in this sense, it cannot follow a rule either, neither mechanically nor otherwise (Hacker 1990, 165).

The point here is not a distinction between rule-following by a machine and rule-following by a human but between two kinds of human rule-following that might be accomplished or simulated by either a machine or a human. In a limited sense, even today's computers may have a part in language games by contributing steps that can be computed by the mere rigid application of rules. Information technology will continue to surprise us with rigid ways of doing things that are today widely believed to require human intelligence. It seems very possible that technologies such as artificial neural networks will, like the brain today, enable "intelligent" features even when we do not understand why. By themselves, artificial systems and devices do not literally follow rules, but within a whole system they can compute steps that resemble human rule-following, i.e., they can simulate parts of human rule-following within a wider context.

Of course, even when they enact rigid rule-following, humans do not follow rules solely by mechanical means but in "analogue" ways. Nevertheless, the decisive measure in rigid rule-following is the rigid application of the strict rules of a calculus. Humans engaging in limited processes pertaining to rigid rule-following, such as in a calculation,

may with reason feel that they are only regarded for how precisely and reliably they can apply a strict rule. A worker who is solely considered in such a limited way may with reason feel dehumanized and "like a machine." Wittgenstein even writes that in such cases the human *is* the machine: "If calculating looks to us like the action of a machine, it is *the human being* doing the calculation that is the machine" (Wittgenstein 1956, III, §20). Wittgenstein does not define machines as certain kinds of physical things but in terms of the processes they do. A human can be a machine if the human is reduced to rigid rule-following processes and the wider context is disregarded.

Because of the affinity of rigid applications of the rules of a calculus with the function of computing machines, the simulation of rigid rule-following seems relatively straightforward. The concept of calculus suggests that there is little freedom in the application of rules. Some think that this holds for the concept of rule-following in general and speak of a "*coercive aspect* to following a rule as part of a practice" (Gøranzon 1998, 252). But already in rule-following within a calculus, and more so in less rigid language games, there is also a *free aspect* to following a rule. In *Philosophical Investigations* and in *Remarks on the Foundations of Mathematics*, Wittgenstein considers numerous examples of rule-following that stand somewhere in-between coercion and freedom. He contends that, even in mathematical proofs, we should not think of the rules as coercing but rather as *guiding* the rule-following:

> Do not look at the proof as a procedure that *compels* you, but as one that *guides* you.—And what it guides is your *conception* of a (particular) situation.

> But how does it come about that it guides *each one* of us in such a way that we agree in the influence it has on us? Well, how does it come that we agree in *counting*? "That is just how we are trained" one may say, "and the agreement produced in this way is carried further by the proofs." (Wittgenstein 1956, III, §30)

Letting oneself be guided by rules goes much beyond blind rule-following. Even counting is a complex practice, and although it seems rather unconditioned, in fact it relies on such things as one's grasp of a situation, given conventions, training received, shared agreement, and willingness to letting oneself be guided. A being that lets itself be guided in this sense is not a mere object in the world but also a subject for the world. Today's computers thus do not even count in the literal

sense even though, as shown in section 2, their processing of information[t] is easily interpreted as a processing of information[r] and information[o].

Yet in non-rigid rule following, letting oneself be guided by a rule is much more intricate. The comparison of speech-acts with games expressed in the concept of language game emphasizes that rule-following is usually much less rigid than suggested by the concept of calculus. One reason is that meaning in ordinary language is usually vague rather than precise (Wittgenstein 1998a, §98–102), which does not mean random or senseless (id., §71). The above citation from *The Blue and Brown Books* highlighted that, contrary to what the earlier proponents of AI and others have thought, often there is no precise definition to ordinary concepts. While sometimes more precise definitions would help, ordinary language often requires vagueness and flexibility. Furthermore, Wittgenstein's concept of language games allows, in contrast to that of the calculus, for the possibility that the rules themselves are not rigid. They even may be creatively modified in the course of the language game:

> And is there not also the case where we play and—make up the rules as we go along? And there is even one where we alter them—as we go along. (id., §83)

In such cases, there is a complex interplay between being guided by and guiding the rule-following practice, an interplay I denote by the term 'creative rule-following.' The alteration of rules may sometimes be rather random, but usually it will make sense and come naturally to the players. If it did not, other players would have a hard time following, at least for prolonged periods of time. This is a further reason for why the nature and cultural embeddedness of the rule-follower are important and neither a randomizer nor mere rigid rule-following suffice to simulate creative rule-following. Beyond rigid rule-following, creative rule-following comprises not only a free aspect of choosing which rules to apply and how to apply them, but also of how to adjust the rules and make up new rules in the course of the language game.

For the above reasons, it is by no means obvious that all creative rule-following can be precisely described and simulated through rigid rule-following in the way presupposed by proponents of AI and their opponents. The investigations of the later Wittgenstein show over and over again the intricate nature of the interplay between the different constituents of rule-following in general and creative rule-following in particular. It is not difficult to see that many interactions that are

possible in a Turing test, such as those described in section 1, require creative rule-following. Any machine that is supposed to universally simulate human intelligence would have to be able to simulate a large variety of creative rule-following behavior. This is an enormous challenge the dimension of which, I suspect, is usually not recognized by those who make sweeping claims about the possibility of replicating or simulating human intelligence.

5. Summary

This paper addressed the extent to which cognitive technology can replicate or simulate human understanding of meaning. It showed some respects in which cognitive technology is challenged by fundamental philosophical questions concerning meaning, understanding, self, reality, and culture, and in turn challenges us to rethink those questions.

The first section considered Turing's question "Can machines think?" and explained some reasons for why even "only" the simulation of rule-following in the Turing test is much more intricate than assumed by leading proponents of AI, as well as by their critics. After Turing, e.g. in Searle's "Chinese Room" thought experiment, an affirmative answer is usually presupposed, and instead it is either asked whether simulation "really" amounts to understanding, or what consequences nearly-perfect machine simulation would have. Since these questions rely in part on an affirmative answer, however, the original question remains important. Instead of making yet another argument for a yes or no answer, this paper reconsidered the concept of technology with respect to the requirements of replicating or simulating human understanding of meaning.

The second section contended that to understand the nature of technology, we need to look beyond technological devices. The paper built on Heidegger's distinction between "old" and "new" technology and his idea that the latter treats nature as a "system of information." Beyond Heidegger, the section argued that information technology is both a further development of technology and the purest form of technology. Yet, information technology is only the beginning of the next technological revolution, at the heart of which lies cognitive technology. The section distinguished different kinds of information and explained that cognitive technology is not essentially about information processing, but about dealing with all kinds of meaning.

The third section investigated the demands on cognitive technology by showing how Wittgenstein's concept of language game

goes beyond that of calculus. The section explained that the concept of calculus was influenced by and influential on thought on AI, and that Wittgenstein developed the concept of language game from that of the calculus. It was shown that Wittgenstein's analogy of the toolbox goes beyond the concept of calculus and that the concept of language game involves a new account of meaning, according to which (1) meaning can be constituted in many ways, (2) it is often vague, and (3) there do not have to be strict rules.

Furthermore, rule-following itself is often not rigid, and the fourth section contrasted rigid with creative rule-following. The concept of calculus suggests strict rules and rigid rule-following, although even here Wittgenstein asserts that rules guide rather than force rule-following. The concept of language game furthermore brings to the fore that ordinary language use often engages in creative rule-following. Creative rule-following comprises not only a free aspect of choosing which rules to apply and how to apply them, but also of how to adjust the rules and make up new rules in the course of the language game.

When language is conceived in terms of the rigid application of strict rules, it seems relatively easy to replicate or simulate human rule-following, which can give rise to the delusion that the era of universal machine-intelligence is near. The difficulties in the development of cognitive technology become clearer, however, when other kinds of rule-following are investigated. While information technology will surely continue to develop astonishing capabilities that once were believed to require human understanding, it will also continue to rely on human interpretation rather than replicating or simulating full-blown understanding. More autonomous forms of navigating the space of meaning, even when limited to the exchange of text messages, require capabilities that go beyond mere information processing. In particular, this paper argued that advanced forms of cognitive technology require the complex integration of at least some of the heterogeneous constituents of creative rule-following.

Acknowledgments

References

Bennett, M.R. et al. 2007. *Neuroscience and Philosophy: Brain, Mind, and Language*. New York: Columbia University Press.

Bennett, M.R. and P.M.S. Hacker. 2003. *Philosophical Foundations of Neuroscience*. Oxford and Malden: Blackwell.

Bennett, M.R. and P.M.S. Hacker. 2008. *History of Cognitive Neuroscience*. Oxford and Malden: Wiley-Blackwell.

Carr, D. 1999. *The Paradox of Subjectivity: The Self in the Transcendental Tradition*. Oxford University Press.

Dennett, D.C. 1992. "The Self as a Center of Narrative Gravity." In *Self and Consciousness: Multiple Perspectives*, ed. Kessel, F.S., P.M. Cole and D.L. Johnson, 103–118. Hillsdale, N.J: L. Erlbaum.

Floridi, L. 2015. "Introduction." In *The Onlife Manifesto: Being Human in a Hyperconnected Era*, e. Floridi, L., 1–3. Cham a.o.: Springer International Publishing. https://doi.org/10.1007/978-3-319-04093-6_1. Accessed September 15, 2018.

Floyd, J. 2017a. "Introduction." In *Philosophical Explorations of the Legacy of Alan Turing*, ed. Floyd, J. and A. Bokulich, 1–35. Cham a.o.: Springer International Publishing. https://doi.org/10.1007/978-3-319-53280-6. Accessed September 15, 2018.

Floyd J. 2017b. "Turing on 'Common Sense': Cambridge Resonances." In *Philosophical Explorations of the Legacy of Alan Turing*, ed. Floyd, J. and A. Bokulich, 103–149. Cham a.o.: Springer International Publishing. https://doi.org/10.1007/978-3-319-53280-6. Accessed September 15, 2018.

Gøranzon, B. 1998. "Beyond All Certainty: Wittgenstein and Turing: An Account of a Philosophical Dialogue on Skill and Technology." In *The Third Culture: Literature and Science*, ed. Shaffer, E.S. Berlin, New York: De Gruyter. https://doi.org/10.1515/9783110882575.237. Accessed September 15, 2018.

Hacker, P.M.S. 1990. *Wittgenstein: Meaning and Mind*. An Analytical Commentary on the Philosophical Investigations 3. Basil Blackwell.

Heidegger, M. 1976. *Gesamtausgabe. Abt. 1, Band 9: Veröffentlichte Schriften 1910–1976 Wegmarken*. 3. Aufl. Frankfurt a.M: Klostermann.

Heidegger, M. 1977. *The Question Concerning Technology, and Other Essays*. New York: Garland Pub.

Heidegger, M. 2000. "Die Frage Nach Der Technik." In *Vorträge Und Aufsätze (1936–1953)*, 7, 5–36. Gesamtausgabe. Frankfurt am Main: Vittorio Klostermann.

Lin, P. 2012. "Introduction to Robot Ethics." In *Robot Ethics: The Ethical and Social Implications of Robotics*, ed. Lin, P., K. Abney and G.A. Bekey, 4–15. Cambridge, Mass: MIT Press.

McCarthy, J. et al. 1955. *Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.* https://rockfound.rockarch.org/digital-library-listing/-/asset_publisher/yYxpQfeI4W8N/content/proposal-for-the-dartmouth-summer-research-project-on-artificial-intelligence. Accessed September 9, 2017.

Rüdel, W. and R. Wisser. 1976. *Im Denken Unterwegs.* https://www.youtube.com/watch?v=WxjjgGcx6o8. Accessed September 15, 2018.

Searle, J.R. 1980. "Minds, Brains, and Programs." *Behavioral and Brain Sciences*, 3 (03), 417–457.

Turing, A.M. 1950. "Computing Machinery and Intelligence." *Mind*, 59 (236), 433–460.

Turing, A.M. 2004a. "Can Automatic Calculating Machines Be Said To Think? (BBC Script)." In *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life, plus the Secrets of Enigma*, ed. Jack Copeland, B., 494–506. Oxford: Oxford University Press. Original MS: http://www.turingarchive.org/browse.php/B/6. Accessed September 15, 2018.

Turing, A.M. 2004b. "Intelligent Machinery." In *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life, plus the Secrets of Enigma*, ed. Jack Copeland, B., 410–432. Oxford: Oxford University Press.

Turing, A.M. 2004c. "On Computable Numbers, with an Application to the Enscheidungsproblem." In *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life, plus the Secrets of Enigma*, ed. Jack Copeland, B., 58–90. Oxford: Oxford University Press.

Weizenbaum, J. 1966. "ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine." *Communications of the ACM*, 9 (1), 36–45.

https://doi.org/10.1145/365153.365168. Accessed September 15, 2018.

Weizenbaum, J. 1976. *Computer Power and Human Reason: From Judgment to Calculation*. San Francisco: Freeman.

Winner, L. 2001. *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. Chicago: Univ. of Chicago Press.

Wittgenstein, L. 1956. *Bemerkungen Über Die Grundlagen Der Mathematik—Remarks on The Foundation of Mathematics*. Trans. Anscombe, G.E.M., G.H. von Wright and R. Rhees. Oxford: Basil Blackwell.

Wittgenstein, L. 1958. *Preliminary Studies for the "Philosophical Investigations": Generally Known as The Blue and Brown Books*. Oxford: Basil Blackwell.

Wittgenstein, L. 1981. *Tractatus Logico-Philosophicus*. Trans. Ogden, C.K. London and New York: Routledge.

Wittgenstein, Ludwig. 1988. *Remarks on the philosophy of psychology: Bemerkungen über die Philosophie der Psychologie. Vol. 1*. Edited by G. E. M. Anscombe. Repr. Oxford: Blackwell.

Wittgenstein, L. 1998. *Philosophical Investigations—Philosophische Untersuchungen*. 2nd ed., trans. Anscombe, G.E.M. Cambridge, Mass: Blackwell.

Wittgenstein, L. 2001. *Tractatus Logico-Philosophicus*. Ed. Pears, D. and B. McGuinness. London and New York: Routledge.

Wittgenstein, L. 2005. *The Big Typescript, TS. 213*. Ed. Luckhardt C.G. and M. Aue. German-English scholar's ed. Malden, MA: Blackwell Pub.