

Wittgenstein and Artificial Intelligence: Towards an update

Skjolden Hotel, Skjolden, June 1-3 2022

Organizers: Alois Pichler and Simo Säätelä

June 1:

- Meeting time at **12:45pm in front of Johanneskirken** (5007 Bergen, Sydneplassen 5)
- Departure by bus to Skjolden from Johanneskirken at **13:00pm sharp**
- Arrival at Skjolden hotel after 5-6 hours in the early evening
- Dinner
- Opening talk by Arild Utaker:
Wittgenstein and our intelligent machines - The digital and the grammatical

June 2:

- 09:15 Alois Pichler and Simo Säätelä: Welcome and introduction

Chair: Simo Säätelä

- 09:25-10:20 Martin Pilch: Tractarian logic and the combinatorics of elementary propositions
 - Discussant: Luciano Bazzocchi
- 10:20-11:15 Joseph Wang-Kathrein: Wittgenstein's shifting from *Tractatus* to *Philosophical Investigations* and its implication for understanding AI
 - Discussant:

11:15-11:30 Coffee break

Chair: Alois Pichler

- 11:30-12:25 Radek Schuster: AI language models in the perspective of Wittgenstein's philosophy: The neural resemblance techniques and the word2picture theory of language
 - Discussant: V.L. Botelho

12:30-13:30 Lunch break

Chair: Alois Pichler

- 13:30-14:25 Herbert Hrachovec: Organisms and calculi: A Wittgensteinian take on AI
 - Discussant: Jakub Macha

Chair: Nivedita Gangopadhyay

- 14:25-15:20 Juliet Floyd: Wittgenstein, Turing, and AI
 - Discussant: Friedrich Stadler
- 15:20-16:15 S. Sunday Greve: Turing's philosophy of AI
 - Discussant: Volker Munz

16:15-16:30 Coffee break

Chair: H. Wilhelm Krüger

- 16:30-17:25 Nivedita Gangopadhyay: The duck, the rabbit, the robot and the philosopher
 - Discussant: Florian Franken Figueiredo
- 17:25-18:20 Ian Ground: Black boxes, beetles and beasts
 - Discussant: Jakub Gomułka

20:00 Dinner

This event is made possible by funding awarded by the Norwegian Research Council, Bergen Universitetsfond and the UiB Humanities Faculty to the Philosophy and Text Technology Research Group at the UiB Department of Philosophy.

An earlier preparatory workshop "Wittgenstein, Semantic technologies and Artificial Intelligence" took place in Øystese, September 4.-9, 2020.

Abstracts

Arild Utaker: Wittgenstein and our intelligent machines: The digital and the grammatical

Philosophy as a struggle against the fascination of technologies makes Wittgenstein particularly important. The point is not that of being against AI but to understand why such a technology cannot at the same time be a general model of language or intelligence. Thus the questions of my paper: What in language cannot be digitalized? And: How can the grammatical (in distinction to the logical) in Wittgenstein give us an alternative to what may be termed a “digitalized culture”?

Martin Pilch: Tractarian logic and the combinatorics of elementary propositions

In my talk I want to distinguish two ways of the construction of truth functions out of elementary propositions according to the *Tractatus*. The first one, the “operational way” consists of a successive application of the N-operator and this is the core of the “General Form of Proposition” given in TLP 6. However, there is a second way, which could be called the “combinatorial way”, also present in the *Tractatus* but less well known. All truth-functions can be achieved by a two-step procedure, which uses the specific Tractarian architecture of truth-arguments, truth-possibilities and truth-conditions. For a given number of n elementary propositions (serving as the truth-arguments) in a first step all possible conjunctions of these n elementary propositions and its negations are formed. For e.g. $n=2$ with p and q being elementary this gives 4 possible combinations $p.q$, $\sim p.q$, $p.\sim q$ and $\sim p.\sim q$ (truth-possibilities). In a second step now all possible subsets these possibilities are constructed and the elements of each subset combined by disjunction. In this way all truth-function can be constructed and this method is equivalent to the construction via N-operator. From a mathematical point of view, this procedure is equivalent to a “free Boolean Algebra” of n generators, generating 2^n so-called “atoms” of the algebra and finally 2^{2^n} elements of the algebra. This free Boolean Algebra in turn is isomorphic to the Lindenbaum-Tarski Algebra of propositional logic. In my talk I want to present an interpretation of (the finite propositional logic part of) Tractarian Logic by discussing the properties of this structure and demonstrate some connections to Hertz’ configuration space (and Boltzmann’s phase space) which can be used for a better understanding of Wittgenstein’s logical space. In the end, I want to show that based on this view there can be given ostensive examples of elementary propositions.

Joseph Wang-Kathrein: Wittgenstein's shifting from *Tractatus* to *Philosophical Investigations* and its implication for understanding Artificial Intelligence

It is said by scholars that one can distinguish between an early and a late period of Ludwig Wittgenstein’s philosophy: The early philosophy is that of *Tractatus Logico-Philosophicus* (TLP); and the late philosophy is that of the *Philosophical Investigations* (PI). One interpretation of Wittgenstein’s oeuvre is that TLP and PI are different answers to the same question: “How is it possible that words and sentences can have meaning?” In TLP Wittgenstein was convinced that words have meaning because they refer to something in the world. Later in PI he changed his mind and claimed that “the use of the word is its meaning”. Starting in the middle of the 20th century, computer scientists have approached the field of artificial intelligence using insights put forward by TLP. This kind of AI – sometimes called Good Old-Fashioned Artificial Intelligence (GOFAI) – is based on the assumption that we can store human knowledge in databases. And by applying inference rules to the data in these systems, the computer would be able to perform certain tasks that – until now – only human beings can do. In certain fields, GOFAI had great success. Medical diagnoses and treatments have benefitted largely from GOFAI. In the last few decades however, computer scientists have shifted their focus towards other machine learning methods (like neuronal networks); thus GOFAI seems to have gone out of fashion. In this presentation, I want to use Wittgenstein’s shift to explain, what are the limits of a system based on GOFAI. Furthermore, by referring to Wittgenstein, I also want to speculate on the limitation of certain AI systems: As long as these systems obey some basic principles that are widely shared among us, they will never be able to “understand” language as human beings do.

Radek Schuster: AI language models in the perspective of Wittgenstein's philosophy: The neural resemblance techniques and the word2picture theory of language

The aim of the talk is to reflect the current neural network techniques of natural language processing (NLP), particularly those based on the word2vec methods, in the context of the development of Wittgenstein philosophy. The main thesis is that although the AI-driven NLP methods can be seen as the analogy to *Philosophical Investigations* of language games, many of our expectations and evaluations of their performance and results have been still held in the Tractarian picture theory of language. While technologies that encode words as numbers in multidimensional semantic spaces make it possible to capture plausibly the diverse uses of expressions as well as their miscellaneous family resemblances across a plethora of language games, the mindset of both the designers and users of these technologies remains largely caught in the

Augustinian view of language. Moreover, while drawing on picture theory of language and factual ontology to justify the parameterization, training, and performance of language models, the distinction between saying and showing is often forgotten, and many programmers and users expect neural networks to tell them something meaningful about the ethical or mystical. In regard to practical examples to support the arguments of the talk, special attention will be paid to systems that use text2image technologies and generate images based on verbal prompts (e.g. VQGAN+CLIP or DALL-E).

Herbert Hrachovec: Organisms and calculi: A Wittgensteinian take on Artificial Intelligence

Philosophically minded computer scientists have recognized an affinity between the general form of a proposition put forward in the *Tractatus* and the construction of programs to explore and manipulate data. Hidden, however, in a few side comments, Wittgenstein regards language as an (unruly) organism. This tension can be shown to directly lead to a standoff between syntax and semantics which is the argumentative core of the books "grand finale". The Wittgenstein of the *Philosophical Remarks* and the *Big Typescript* attempts to resolve the difficulties by an analysis of a calculus' relationship to its application. Since computer programs are built to be applied "in the real world", this can be regarded as a second (presumable) approach towards the philosophical issues raised by artificial intelligence. But Wittgenstein's notion of a calculus remains antithetical to any organic development. He arrives at a second installment of the standoff mentioned. Dropping this attitude prepares the ground for him to regard language as a form of life.

Juliet Floyd: Wittgenstein, Turing, and AI

Turing drew from Wittgenstein's 1939 Cambridge lectures the idea that everyday typings of concepts, our evolving "phraseology", plays a fundamental role in the application of logic. After going to Bletchley Park, Turing continued to think about the importance of notations, and in "The Reform of Mathematical Notation" (1944/45) he suggested that symbolic logic opens itself up to a plurality of systems, attending to the specific uses to which notations are put, and Turing argued that we should take into account everyday language when constructing logical notations. This Wittgensteinian aspect of Turing's philosophy of logic culminated in his 1948 report to the National Physical Laboratory, "Intelligent Machinery", the founding document of AI. Here Turing envisioned, with great prescience, a future with machinery that would involve, not only the search for proof systems and notations, but also the use of "intelligent machinery" in biological searching and, in the end, what Turing called "the cultural search", a search involving humanity as a whole (and no machines). "Intelligence" Turing defined, tentatively, as the capacity to appreciate the importance of different kinds of searching – an echo of an idea to be found in *Philosophical Investigations*. Later in his "Computing Machinery and Intelligence" (1950) Turing devised the "imitation game", a "Turing Test" that has been much discussed in philosophy of mind and popular culture. Too often this Test has been wrongly conceived as furthering mechanism or behaviorism about the notions of thought and mentality: it is read in a dualist, Cartesian vein. On this (mis)reading, the social dimensions of Turing's game have been ignored. However, the Turing Test is actually a social experiment in "phraseology", a quest to elicit criteria from us – in brief, a language-game in Wittgenstein's sense. I will lay out this reading of the Test and discuss some implications for AI in our world today.

S. Sunday Greve: Turing's philosophy of AI

The value of Turing's work on artificial intelligence has traditionally been reduced to what is now known as the Turing test, but it is more nuanced and compelling than previously assumed. Turing's thinking on this topic was far ahead of everyone else's, partly because he had discovered the fundamental principle of modern computing machinery, the stored-program design, as early as 1936 (a full twelve years before the first modern computer was actually engineered). Careful historical reconstruction of Turing's philosophy of AI shows that the heart of this work consists of logical investigations that proceed on the basis of what the later Wittgenstein called 'language games', which Turing employs in precisely the kind of function that Wittgenstein described as their being used as 'objects of comparison' (PI 130).

Nivedita Gangopadhyay: The duck, the rabbit, the robot and the philosopher

Ambiguous figures present us with one of the biggest perceptual challenges. How does the perceiver make sense of ambiguous figures? How does the perceiver, in the process of giving meaning to the ambiguous figure, pick one aspect or interpretation of the figure over another equally probable one? A philosophical question is what do ambiguous figures tell us about the content of experience? A psychological question is how does the mind deal with uncertainty when confronted with ambiguous figures? The problems of ambiguous perception greatly occupied Wittgenstein, for example in his writings between 1946 and 1949. He discussed it under the topic of aspect perception or seeing-as. In this presentation I would like to bring in insights from Wittgenstein's discussions of aspect perception into a field of perception that did not even exist when Wittgenstein composed his works, namely, machine vision. Given the ubiquity of machine vision in our daily

life, from our smartphones to some of the most complex technologies currently found, the question is not merely how does machine vision tackle ambiguous figures (if it does so at all) but rather how should machine vision approach the perceptual challenges of ambiguous figures?

Ian Ground: Black boxes, beetles and beasts

A common ethical objection to certain classes of A.I system – the Black Box issue – depends on the realisation that answers to the question of why (at least for some senses of “why”), the system made a particular “decision” are, logically, unavailable. This objection is, or should be, independent of the objection that an A.I system may be trained on data which is some way unjustly skewed. The implications of the Black Box issue for the philosophy of mind and in particular our conception of normative rationality are less frequently explored. Can a decision be regarded as rationally based if, in principle, it is not possible to track that decision back through a series of deductions, inferences, and principles? Is it legitimate to compare such cases to our everyday reliance on testimony? What are the implications for the conceptual possibility of minded machines? In these debates, Wittgensteinians may find themselves facing some dilemmas. Many are, at least temperamentally, inclined to scepticism regarding claims about the putative intelligence of machines, rejecting a range of assumptions upon which those claims depend. However, for classes of A.I to which the black box problem applies, many of those assumptions are not in play. Moreover, Wittgenstein’s remarks on rule-following reject the idea that rational cognition is possible only via encoded representations of rules. The same rejection is central to the positive case for machine intelligence in the A.I case. How then should the Wittgensteinian respond to these issues? In this discussion, I offer reasons for thinking that Wittgensteinians should be intensely relaxed about such A.I systems. It remains for the Wittgensteinian to resist the conceptual possibility of machine intelligence by insisting on the foundational role of the concepts of biological life and (en-)action perhaps in tandem with a quietist attitude towards explanation. The discussion concludes by raising some concerns about this strategy via a comparison with the case of non-linguistic animals.

Participants:

- 1) Alois Pichler (Univ. of Bergen)
- 2) Simo Säätelä (Univ. of Bergen)
- 3) Rune Falch (Univ. of Bergen)
- 4) Nivedita Gangopadhyay (Univ. of Bergen)
- 5) H. Wilhelm Krüger (Univ. of Bergen)
- 6) Victor Lacerda Botelho (Univ. of Bergen)
- 7) Arild Utaker (Univ. of Bergen)
- 8) Filipe Campello (Univ. of Pernambuco)
- 9) Luciano Bazzocchi (Univ. of Siena)
- 10) Hanoch Ben-Yami (Central European University)
- 11) J. Matthew Fielding (Takin Solutions Berlin)
- 12) Florian Franken Figueiredo (Univ. of Lisbon)
- 13) Juliet Floyd (Boston Univ.)
- 14) Arthur Gibson (Univ. of Cambridge)
- 15) Jakub Gomulka (Pedagogical Univ. of Cracow)
- 16) S. Sunday Greve (Peking University)
- 17) Ian Ground (British Wittgenstein Society)
- 18) Herbert Hrachovec (Univ. of Vienna)
- 19) Lassi J. Jakola (Univ. of Helsinki)
- 20) Jakub Macha (Univ. of Brno)
- 21) Volker Munz (Univ. of Klagenfurt)
- 22) Martin Pilch (Vienna)
- 23) Radek Schuster (Univ. of West Bohemia)
- 24) Ludek Sekyra (Prague)
- 25) Ilse Somavilla (Univ. of Innsbruck)
- 26) Friedrich Stadler (Univ. of Vienna)
- 27) Joseph Wang-Kathrein (Univ. of Innsbruck)