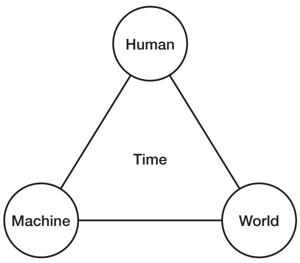
November 2020

– Master thesis exposé / summary – Christina Schinzel

THE DISAPPEARANCE OF THE PRESENT

Intelligent Machines and the Anticipation of the Future: Temporal Relations in Machine Learning Networks and Beyond



Abstract

The project centers around the question of machine learning and time and asks, what kind of temporal frameworks are at work in machine learning systems like neural networks and how the creation of and the interaction with these networks influences and creates temporal relations. Code, data, hardware and user input intertwine and engage in mutual world-making, shaping human and computational epistemology on the way. The research horizon of this project can be marked out at the triangulation of humans, machines and world and explores how these entities are bound together on a temporal axis.

Starting Point

Artificial intelligence is all over the place. From newspaper coverage to Twitter threads, from governmental strategies to economy 4.0, from everyday virtual assistants to Pepper-robots rolling happily along: AI gained center stage at the public debate. But with all this hype it is easy to forget that the breakthrough of modern-day artificial intelligence systems was not too long ago. Besides having a long history with many ups and downs that goes back to the 1940s, a turning point can be observed around 2012. At that time, an image data set called ImageNet was used to train a neural network to classify images.[1] Classifying images means, that you have an image and the network can tell what is depicted thereon: be it a dog, a flower, a bike, or one of the other classes. It turned out, that this network performed extremely well, better than any other neural network before – up to a human-like recognition rate.

1 see Krizhevsky, Alex, Ilya Sutskever, and Geoffrey Hinton. "ImageNet Classification with Deep Convolutional Neural Networks." Proceedings of the 25th International Conference on Neural Information Processing Systems 1 (2012): 1097-1105

This breakthrough was mainly reached because modern neural networks became deep with multiple layers stacked on top of each other. Together with new accelerated hardware, especially parallel graphic processing units, neural computation systems became powerful data crunching machines, that were able to process networks with millions of parameters – tasks that were previously not possible to this extent.

Fast forward eight years, and the first euphoric phase gave way for some pressing questions. With interest-based advertising, personalized recommendations, discriminatory hiring systems, predictive policing and recidivism risk level assessment, racist chat bots, humans misrecognized as apes, data-driven drone strikes and autonomous cars that should make decisions on life and death, doubts about the reliability and trustworthiness of artificial intelligence arose.

These are the socially and politically visible effects of machine learning, which are often cited and used as a general critique against artificial intelligence. And sure, these problems still persist, but during the last years, society has become more attentive and the (computer) science community is well aware of these flaws and algorithmic bias and AI ethics have become an active field of research. But being aware does not mean that there is a solution (if there could be any).

Examples like these were also the initial starting point for my interest in this topic and the starting point for this thesis. I kept asking myself: what is going on there, and why? But this thesis is not about algorithmic bias or AI applications itself. It is more a study of human and machinic relations, of which the aforementioned examples can be viewed as symptoms. In this thesis I will look deeper into the epistemological and technical structures which fostered such developments in the first place.

What all of these examples have in common is a distinct temporality and the trade-off of knowledge for data, which led to a bottom-up process of inductive decision-making and reasoning. In the early days of digital technologies, data was seen as a by-product of computation. Storing data manually in a place in memory was necessary to be able to run any program at all. Nowadays it seems to be the other way round: computation is the by-product of – or the only way to manage - all of today's overabundance of data. Computer scientist Ethem Alpaydin states, that "before, data was what the programs processed and spit out – data was passive. [With machine learning], data starts to drive the operation; it is not the programmers anymore but the data itself that defines what to do next."[2] Programming became data-driven programming.

And everything can be translated into digital data. Pictures are no longer impressions of what you see but broken up into pixel values and color channels, novels are no longer written in verbose prose but split into words and stored in numerical encoding vectors, faces don't smile anymore but consist of an ensemble of facial key points that relate to each other in a certain way. The human world is digitized. We have long arrived at what Katherine Hayles calls the 'posthuman subject.'[3] It is time to interfere as to which direction to take.

2 Alpaydin, Ethem. Machine Larning: The New AI. Cambridge (Mass.), London: The MIT Press, 2016. p. 11

3 see Hayles, Katherine. How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics. Chicago (Ill.): University of Chicago Press, 1999

Research Interest

The research horizon of this project can be marked out at the triangulation of humans, machines and world/environment and explores how these entities are bound together on a temporal axis.

First, I will outline a topology of the field of machine learning with a focus on the present situation and look at machine learning from a global (concerning the whole network) and local (concerning certain parts of a network) perspective. What has to be clarified: how do these systems work and what temporal relations exist? The subsequent question is: why are machine learning networks build like that? It is of interest here, to look back at the history of artificial intelligence and the beginnings of digital computers during the post-war period, which was accompanied by a rise of cybernetics. How has this beginning phase shaped the logic behind modern-day networks, with a focus on memory, states, and feedback?

Early AI networks were mainly of theoretical or prototypical nature, but they haven't experienced widespread usage. This changed when artificial intelligence took off with machine learning, ubiquitous devices and a never-ending stream of data. It is now possible to deploy machine learning systems at scale, which results in a widespread impact on the individual and social level. In the last part I will explore what real-life consequences machine learning has on the world outside of the machine. Focusing again on temporal relations and time constituting factors like recursion, data storage, memory and retention, I am finally going to ask: does the present really disappear? And if so, how and with what consequences?

The research hypothesis is, that the recent development in artificial intelligence intervenes on our very ability to navigate this world, make decisions and imagine a future. This shift happens as a consequence of realigning and prothesizing individual experience and memory as constituting factors of meaning through machines.

Besides following up on this argumentation it is inevitable to also look deeper inside the machine. Herman H. Goldstine and John von Neumann contended in the early days of modern programming in the 1940s, that programming is "the technique of providing a dynamic background to control the automatic evolution of a meaning."[4] So, before analyzing what kind of meaning evolves, I will examine this programmed background more thoroughly. Because machines, even very dynamic and non-linear ones like in the case of machine learning, are bound to the environment which is given to them and the meaning therein created is restricted to its provided parameters.

Methods

With regard to artificial intelligence and its technical manifestation in the form of machine learning and deep learning I will explore this topic by juxtaposing current technological developments and research findings from computer science to a media studies point of view. It is an interdisciplinary approach, which can be located in the field of critical code and software studies.[5]

The focus of this project lies on the current situation and present-day technologies. It is not a media archaeology or a history of technology. Of course, machine learning hasn't evolved in a vacuum, so selected historical sources, which could be seen as precursors of the current systems, will also be referred to.

4 Herman H. Goldstine and John von Neumann: Planning and Coding of Problems for an Electronic Computing Instrument, Report on the Mathematical and Logical Aspects of an Electronic Computing Instrument, Part II, Volume I, Princeton, N.J.: Institute for Advanced Study, 1947. p. 2

5 see for example: Marino, Mark C. Critical Code Studies, 2020 – Berry, David M. The Philosophy of Software: Code and Mediation in the Digital Age, 2011 – Chun, Wendy Hui Kyong. Programmed Visions: Software and Memory, 2011

The argumentation builds up from inside the machine and goes on to examine what happens when the discrete world of the machine is left and meets the outside world, the world of its users, the world of humans. With this approach I want to stress the importance of including technical structures and computer code into the analysis. Because code is more than a set of logical instructions and can therefore be read and interpreted beyond its pure functionality.

When analyzing the effects of artificial intelligence, a technical perspective is seldom subject to consideration. At the center of investigation lies the input and the output, or the process of interaction, but not what lies in between. But a device is always more than its input and output, not least because it reproduces only a fraction of the processes from inside the machine, which are then passed on as output and made available for human sensing. Thus, it is inevitable to not just consider the surface layer of interfaces, but also the hidden layer of computation.

The obfuscation of technical processes is very common and, as a result, responsibility is delegated to algorithms and away from humans. The delegation is twofold: the operation becomes easy and leads to greater control through interfaces, but the technical reality stays hidden and intangible inside the machine. Computers become a convenient and authoritative source of causality.

Media scholar Wendy Hui Kyong Chun argues that this invisibility leads to a fetishization of code. A fetish refers to an object worshiped by primitive peoples because of its supposed inherent magical power. It is a source of action, but in unexplainable ways. Now, replace natural sources of magical power like trees and stones with computers and you get to a similar situation. Code, because of its hidden and intangible nature, becomes the unquestioned, black-boxed power of decision making and a source of causality. Chun argues, that "the parallel to source code seems obvious: we "primitive folk" worship source code as a magical entity – as a source of causality – when in truth the power lies elsewhere, most importantly, in social and machinic relations."[6] Exploring the meaning inherent in code and algorithmic structures, which is constituent of machinic and human relations, will therefore help lifting this obfuscation.

A tech-savvy view on technology is needed to face the current technical development not in a passive way, which may on the one hand lead to an unquestioned believe in algorithms, or, on the other hand, to some kind of techno-pessimism and fear. Both scenarios result in a misunderstanding of the machine. A misunderstanding that leads – in the words of Gilbert Simondon – to alienation: "The strongest cause of alienation in the contemporary world resides in the misunderstanding of the machine, which is not an alienation caused by the machine itself, but by the lack of understanding of its nature and its essence, because of its absence from the world of significations and its omission from the canon of values and concepts which make up culture."[7]

6 Chun, Wendy Hui Kyong. Programmed Visions: Software and Memory. Cambridge Mass.: MIT Press, 2011. p. 51

7 Simondon, Gilbert. Du Mode d'Existence des Objets Techniques. Paris: Aubier, 1958. p. 9 f., "La plus forte cause de l'aliénation dans le monde contemporain réside dans cette méconnaissance de la machine, qui n'est pas une aliénation causée par la machine, mais par la non-connaissance de sa nature et de son essence, par son absence du monde des significations, et par son omission dans la table des valeurs et des concepts faisant partie de la culture."

Long Version

In the following part, I will elaborate more on the main arguments of this project in a concise summary.

Computation, Neural Networks, and Time

All computation is temporal. Computation can only exist when it unfolds in time. For a digital device to function means the constant internal flow of information, even when it seems idle. From the flow of electricity, which can be either on or off, to the constant exchange between memory and central processing unit, to basic functions that run along a structured path and return a result. Computation operates in time and is pervasive at every level of the machine.

Computational processes operate on a scale of complexity: from linear step-by-step calculations to the complex and intricate intertwinings inside neural networks. In neural networks, linearity is replaced with iteration and feedback. Negative feedback controls the network, so that the output is regulated by a margin of error.

It was in the 1940s, when the cybernetics around Norbert Wiener opened up the thought to a feedback-driven world where animals/humans and machines are actants in a system of communication and control. The concept of circular regulatory systems received widespread acclaim, so that in the following decades cybernetic thought kept percolating into many other disciplines like sociology, biology or psychology and strengthened cybernetics' claim to universality.

Still, control through feedback lies at the heart of every modern neural network. The networks are built of nodes, so-called artificial neurons, that are stacked in layers and each node is connected to each node of the following layer. Feedback is generated through backpropagation and gradient descent, meaning that the result of one pass through the network is the input to the next (backward-) pass through the same network. Neural networks are generally goal-oriented, insofar as the aim is map an input to a desired output through learning. After the learning phase, the network can generalize and recognize similar inputs the next time. Looking closer, it is this feedback-driven and at the same time goal-oriented structure, that leaves its trace on temporal relations. Because what is looping inside the network is not only numbers and data, but also reasoning, logic and knowledge.

There are different network architectures for different use cases, like classification or regression. A task that is of major interest since the beginning days is textual processing. A text is a set of words that has a linear time of narrative, meaning that you read one word after another in sequential order. For a long time, it was hard to keep data circulating inside a network and reactivate past instances. Because of its recurring nature, the signal became weaker with every iteration until it completely vanished.[8] Networks advanced significantly, when the processing of exactly this type of sequential data became possible. And it is not only text, that is sequential, but also speech or video.

When working with time-sequential data you need neural networks with memory, so that past computations are not forgotten and can be reactivated at a specific time. This is done with memory-nodes at every neuron which are build up like a mini-network with different logic gates that control what information is to be retained and what is to be forgotten.[9] Most of today's state-of-the-art networks use memory nodes. With them, tasks like handwritten character recognition, speech recognition, machine translation, text-to-speech synthesis and natural language processing can be processed with high accuracy.

8 see for example: Jordan, Michael I. Serial Order: A Parallel Distributed Processing Approach, 1986 or Elman, Jeffrey L. Finding Structure in Time, 1990

9 see Hochreiter, Sepp, and Jürgen Schmidhuber. "Long Short-Term Memory." Neural computation 9, no. 8 (1997): 1735-1780

Historical Precursors

The machine learning networks we have today and that seemingly appeared out of nothing some years ago, actually have a long history that goes back to the 1940s. The grand project of that time was the mapping of the anatomic mind of humans onto a synthesized version that can be implemented in machines.

Without going too deep into the historicity of machine learning, I want to outline some important bifurcations which lead to the machine learning systems we have today. Many of the logical foundations of machine learning – and digital computers in general – were laid out during the post-war period of the 1940s-1960s. Consequently, many ideas persist in modern-day networks. However, today's systems tend to be mainly concerned with functionality, accuracy or scalability. They don't question their very foundations anymore. Looking back at these beginning days gives important insights into the precursors of modern networks and the logic behind them. I will focus here on the work of Warren McCulloch, Walter Pitts, John von Neumann, Norbert Wiener and Ross Ashby.

Neural networks were inspired by a simplified and idealized version of the central nervous system and originated from biophysics. When referring to seminal historical articles, Warren McCulloch and Walter Pitts' work has certainly to be included. They did trailblazing work at the intersection of computation and brain. In their 1943 paper "A logical calculus of the ideas immanent in nervous activity,"[10] they outlined for the first time a computational theory of mind and brain. The project intended to bridge the gulf between mind and body, between mind anatomized and mind synthesized. It was the first use of computation to address the mind-body problem.

Many of their ideas were developed further during the following decades and became important stepping stones for the logic of computation: the idea, that the brain is a digital organ with all-or-none activity, that networks with circles could make synthetic a priori judgments, and that, with a determinate network, you can make the unknown (1) knowable and (2) predictable. McCulloch and Pitts contend, that "with determination of the net, the unknowable object of knowledge, the "thing in itself," ceases to be unknowable."[11] Although McCulloch and Pitts' theory didn't stand the test of time concerning neural processes, the techniques were developed further into the design of neural networks and led to the development of finite automata, one of the most important formalisms in the theory of computation.[12]

It was John von Neumann, who developed McCulloch and Pitts' concepts further into finite automata. Despite von Neumann's vast influences on the logical and technical construction of digital computers, I will refer here mainly to the construction and components of digital computers and on the notion of memory and what function it has inside a machine. In the 1940s, John von Neuman constructed the first digital computers and sketched out their internal logical structure,[13] which later should become known as von Neumann architecture. Still, almost all of our modern computers follow this architecture.

A computer consists of several 'organs': the arithmetic organ, the control organ and the memory organ. The memory organ becomes the central component of digital machines, through which all internal processes have to pass at one time or another. Before von Neumann's computer architecture, data was stored externally on punch cards or teletype tape. These storage media had to be manually inserted, so the machine can execute certain functions. With von Neumann's stored-program architecture, memory becomes an internal part of

10 McCulloch, Warren S., and Walter Pitts. "A Logical Calculus of the Ideas Immanent in Nervous Activity." The Bulletin of Mathematical Biophysics 5, no. 4 (1943): 115-33

11 McCulloch, Warren S., and Walter Pitts. "A Logical Calculus of the Ideas Immanent in Nervous Activity." The Bulletin of Mathematical Biophysics 5, no. 4 (1943): 115-33. p. 17

12 see Piccinini, Gualtiero. "The First Computational Theory of Mind and Brain: A Close Look at Mcculloch and Pitts's "Logical Calculus of Ideas Immanent in Nervous Activity"" Synthese 141, no. 2 (2004): 175-215

13 see for example: Neumann, John von. First Draft of a Report on the EDVAC, 1945 – Burks, Arthur W., Herman H. Goldstine, and John von Neumann. Preliminary Discussion of the Logical Design of an Electronic Computing Instrument, 1946 – Neumann, John von. The Computer and the Brain, 1958

the machine. He took the external storage and put it right inside the machine. Storage became memory – by name and by function.

Von Neumann defined different hierarchical memory organs: a primary memory, that is temporary and can be accessed quickly, and a secondary memory, that serves as a storage medium for longer periods of time. In a short paragraph, almost like a side note, he states, that there is also a third memory in the hierarchy: the outside world: "The very last stage of any memory hierarchy is necessarily the outside world, that is, the outside world as far as the machine is concerned, i.e. that part of it with which the machine can directly communicate, in other words the input and the output organs of the machine."[14]

Dead Storage, Memory Supports, and Prediction

When viewed from the perspective of the machine, the final stage of any memory hierarchy is the outside world. Or, it is always the first one, when viewed from a human perspective.

In a last step, I will leave the discrete world of the machine and explore the world of the – in von Neumann's words – 'third memory'. A characteristic of this last kind of memory is its "vast quantity of dead storage", that is "not integrated with the machine."[15]

In fact, humans live by 'dead storage'. It is called history. Bernard Stiegler describes this condition in a less technical way: he calls it epiphylogenesis. It is the bridging of the biological memory of genetics and the somatic memory of epigenetics which emerged by way of an exteriorization process. It is technical by its very nature and can only exist through objects and tools, and hence becomes transferable. Letters mark the beginning of the 'historical' age of historicality, the moment when knowledge could be exteriorized and defies the death of an individual.[16]

Technical memory supports like books, photographs and film constitute the way we conceive the past, the present and the future. But when we all rely on similar memory supports with similar inscribed knowledge, it gets easy to collectivize and industrialize them. This leads to the industrialization of memory itself, as Stiegler argues: "The 20th century is the century of the industrialization, the conservation and the transmission – that is, the selection – of memory. This industrialization becomes concretized in the generalization of the production of industrial temporal objects (phonograms, films, radio and television programs, etc.), with the consequences to be drawn concerning the fact that millions, hundreds of millions of consciousnesses are every day the consciousnesses, at the same time of the same temporal objects."[17]

Based on Edmund Husserl's work on the phenomenology of time consciousness, Stiegler develops the concept of temporal objects further. After primary retention (perception in the flux of consciousness) and secondary retention (reactivation of a previous experience), Stieger identified another kind of retention: it is realized through memory supports and can be retained with the help of technical objects. It is called tertiary retention and exists a priori and has far-reaching influences on our consciousness and individuation.

Stiegler's main topic is the industrialization of memory, however, he mainly focuses on 'old' media and their impact on consciousness and the present situation. Yuk Hui follows Stiegler's train of thought but expands it to include 'new' media – digital technologies – like algorithms and machine learning into the temporal process. Hui adds the future counterpart to Stiegler's tertiary

14 Neumann, John von. The Computer and the Brain. New Haven, London: Yale University Press, 1958. p. 36

15 Burks, Arthur W., Herman H. Goldstine, and John von Neumann. "Preliminary Discussion of the Logical Design of an Electronic Computing Instrument." (1946), In The Origins of Digital Computers: Selected Papers. Edited by Brian Randell, 399-413. Berlin, New York: Springer-Verlag, 1973. p. 405

16 see Stiegler, Bernard. La Technique et le Temps: 2. La Désorientation. Paris: Éditions Galilée, 1996. p. 12

17 Bernard Stiegler cited in: Roberts, Ben. "Cinema as Mnemotechnics: Bernard Stiegler and the "Industrialization of Memory"" Angelaki 11, no. 1 (2006): 55-63. p. 6

retention: tertiary protention. After primary protention (anticipation of the next moment) and secondary protention (anticipation based on past experience), tertiary protention is based on tertiary retention through technical objects and can thus be algorithmically analyzed in order to extract probable future scenarios. Hui contends, that "(1) tertiary protention tends to depend on tertiary retention, for example, the relations given by digital objects, those traces we have left, such as pictures, videos, or geolocations; and (2) orientation becomes more and more an algorithmic process that analyzes and produces relations to pave the way for the experience of the next now or the immediate future."[18]

Machine learning is tertiary protention in its purest form. It is based on past data in all possible forms: you take what you know and predict what you want to know. And so, pixel values predict the presence or absence of an object, words predict someone's mood, movies you liked predict more movies you may like. Hui states, that "without algorithms, digital objects would be mere retentions residing on the hard drives of computers and servers. Through the analysis of data – or, more or less, through speculation – the machines are able to produce surprises [...] by identifying a possible (and probable) "future," a specific conception of time and space that is always already ahead but that we have not yet projected."[19]

The future is linked to the past by way of the present. Through algorithmic analysis and processing, the future becomes something that can be quantified, predicted and sold. Computers become memory machines, that are always already there, and leave the 'I' thrust into the a priori collective memory of the 'we'. It is a future based on the past, where the present becomes a neglectable part of this transitional process. Of importance is only the mapping of the past to the future by means of a set of relations and patterns.

McCulloch has already foreseen this situation in a prescient statement in 1950: "The earmark of every predictive circuit is that if it has operated long uniformly it will persist in activity, or overshoot; otherwise it could not project regularities from the known past upon the unknown future. This is what, as a scientist, I dread most, for as our memories become stored, we become creatures of our yesterdays – mere hasbeens in a changing world. This leaves no room for learning."[20]

One-Way Causality

The situation gets even more intricate when looking at the treatment of internal past inside neural networks. A characteristic of neural networks, and the reason why they are often referred to as black boxes lies in the fact that they use data and knowledge from outside (externalized memory) to produce results and predictions. But when processing this knowledge, the network is 'unaware' of the ways intermediate results come about, because it is programmed to get better at approximating a desired output and on focusing at a goal, not to look back. In technical terms it means, that you cannot revers engineer a neural network's prior state. There are efforts to do exactly this within the emerging field of explainable AI (XAI).[21] Although an interesting field to look at, explainable AI should not be further discussed in this inquiry.

So, it is generally very hard to trace a network's logic backwards, because millions, if not uncountable possible combinations of neurons, weights, biases, optimizers and other hyperparameters pass through the network, creating disjunctive and transitional results at each neuron on the fly without human intervention, always being forward-directed, without leaving traces of their internal past. Causality runs only one way. Without the possibility of retrospection, past

18 Hui, Yuk. On the Existence of Digital Objects. Minneapolis (Minn.), London: University of Minnesota Press, 2016. p. 221

19 Hui, Yuk. On the Existence of Digital Objects. Minneapolis (Minn.), London: University of Minnesota Press, 2016. p. 241

20 McCulloch, Warren S. "Why the Mind Is in the Head?" dialectica 4, no. 3 (1950): 192-205. p. 10

21 A good overview of recent (technical) papers can be found in: Samek et al., eds. Explainable AI: Interpreting, Explaining and Visualizing Deep Learning, 2019

context gets eliminated. The network produces context-free circuits that lead to 'eternal ideas'. Wendy Hui Kyong Chun expands on this idea, when she refers to McCulloch and Pitts' "Logical Calculus" in the light of memory: "What is retained [in neural networks] is the memory, not all the events that led to that memory. In this sense, they threaten to become "eternal ideas," separated from context. This separation, combined with the fact that the neural nets can specify the next but not the previous state, means that "our knowledge of the world, including ourselves, is incomplete as to space and indefinite as to time."" [22]

Jacques Derrida identified writing as the disappearance of the origin.[23] With the creation of neural networks by means of writing computer code, the origin disappears as well, but what's more, also all possibilities to back trace it. What remains is empty knowledge separated from any source.

The Destruction of Memory

Geoffrey C. Bowker takes up the argumentation of knowledge deprived of its origin when he examines memory practices in cybernetics. Because much of cybernetic argumentation is based on feedback, neural networks with their recursive feedback-structure carry the same ideas in themselves.

Bowker refers to cybernetic memory as an empty archive. He argues that the cybernetic worldview – with feedback being one of its main characteristics – leads to a destruction of memory. He contends, that the new digital age is one of synchrony and form, as opposed to the old age of diachrony and materialism, where "humans become purified of matter and started to act as information."[24] In such a system you don't need memory, just rules.

At least a conscious holding of past memories is not needed. Like in neural networks, where the circulating data creates a transitory memory, from which every next step can be calculated, and that is updated with every pass through the network. Where should individual memory have its place in such a system? Bowker says, that "the theme of the destruction of memory is a complex one. It is not that past knowledge is not needed; indeed, it most certainly is in order to make sense of current actions. However, a conscious holding of the past in mind was not needed: the actant under consideration – a dog, a person, a computer – had been made sufficiently different that, first, past knowledge was by definition retained and sorted and, second, only useful past knowledge survived."[25]

This train of thought takes a very radical form in Ross Ashby's cybernetic theory.[26] He argues, that memory is an epiphenomenon that evolved because of an incomplete knowledge about a system (in Ashby's view, the human brain is also such a system): "the concept of 'memory' arises most naturally in the Investigator's mind when not all of the system is accessible to observation, so that he must use information of what happened earlier to take the place of what he cannot observe now. 'Memory,' from this point of view, is not an objective and intrinsic property of a system but a reflection of the Investigator's limited powers of observation."[27] Memory is not needed anymore, once every part of the system, the environment, its past and its future, is known. When every part is linked together you don't need memory. You just need to know all details and how to calculate their connections, for example, by means of a Markovian chain. This argument is like a Laplacian determinism, in which, at one moment in time, the universe's past and future can be calculated from a complete set of states.

22 Chun, Wendy Hui Kyong. Programmed Visions: Software and Memory. Cambridge Mass.: MIT Press, 2011. p. 155

23 see Derrida, Jacques. "La Pharmacie de Platon." In La Dissémination. Paris: Éditions du Seuil, 1988

24 Bowker, Geoffrey C. Memory Practices in the Sciences. Cambridge (Mass.): MIT Press, 2005. p. 90

25 Bowker, Geoffrey C. Memory Practices in the Sciences. Cambridge (Mass.): MIT Press, 2005. p. 100

26 see for example: Ashby, W. Ross. An Introduction to Cybernetics, 1956 and Ashby, W. Ross. Design for a Brain: The Origin of Adaptive Behaviour, 1952

27 Ashby, W. Ross. "General Systems Theory as a New Discipline." General Systems Yearbook 3 (1958): 1-6. p. 3. as cited in Bowker, Geoffrey C. Memory Practices in the Sciences. p. 100

Memory can be used (and exploited) as a way of short circuiting the synchronization process of humans and machines. It bridges across the human and the machinic. A synchronized memory assimilates different pasts and makes the – one – future calculable. Moreover, it eliminates individuation. Individuation always includes the possibility of outliers and unforeseeable conditions. Yet, every outlier is a possible source of disturbance. Just like a sigmoid function in a neural network, which brings all results between a predefined value range, a synchronized memory brings individuation between a predefined range of possible personas. As a result, "difference [is] banished as noise and distraction" and stories become "stories of sameness."[28]

Conclusion

In popular discourse artificial intelligence is still loaded with science-fiction tropes like the assumption, that machines and humanoid robots are about to take over the world. If not right now, then surely withing the next years. But as I have shown, this picture is outright wrong. We must not fear some arbitrary humanoid robots of our own making, but the inflationary use of external mnemotechnics qua machine learning, where individual experience is erased. This leads to a leveling of individual experience into collective memory and goes far beyond traditional mnemotechnics like writing or tape-recording. It results in an accelerated extrapolation of the self, which creates – feared by some and anticipated by others – a world of dataism.

A powerful antidote to this disempowerment lies in an understanding of the machine and in taking models of the brain and simulations of intelligence for what they are: models and simulations. Therefore we have to look beyond the interfaces right into the machines and see how their architecture and code is linked to and influences the world we all live in. We have to disclose the equalizations and conflations that take place in the course of an anthropomorphizing process of today's intelligent machines on various levels: the conflation of intelligence with goal-directed and problem-solving behavior, the conflation of memory with a storage medium, and ultimately the conflation of mind and brain.

28 Bowker, Geoffrey C. Memory Practices in the Sciences. Cambridge (Mass.): MIT Press, 2005. p. 112

Table of Contents

References

Agamben, Giorgio. Infancy and History: The Destruction of Experience. London: Verso, 1993

Allen, James F. "Towards a General Theory of Action and Time." Artificial Intelligence 23, no. 2 (1984): 123-154

Alpaydin, Ethem. Machine Learning: The New AI. Cambridge (Mass.), London: The MIT Press, 2016

Ashby, W. Ross. An Introduction to Cybernetics. Mansfield Centre, CT: Martino Publishing, 1956

Ashby, W. Ross. Design for a Brain: The Origin of Adaptive Behaviour. London: Chapman and Hall, 1952

Bengio, Yoshua. "Learning Deep Architectures for AI." Foundations and Trends in Machine Learning 2, no. 1 (2009): 1-127

Bengio, Yoshua, Patrice Simard, and Paolo Frasconi. "Learning Long-Term Dependencies with Gradient Descent Is Difficult." IEEE transactions on neural networks 5, no. 2 (1994): 157-166

Berry, David M. The Philosophy of Software: Code and Mediation in the Digital Age. London: Palgrave Macmillan UK, 2011

Bowker, Geoffrey C. Memory Practices in the Sciences. Cambridge (Mass.): MIT Press, 2005

Bowker, Geoffrey, and Ray-Shyng Chou. "Ashby's Notion of Memory and the Ontology of Technical Evolution." International Journal of General Systems 38, no. 2 (2009): 129-137

Brooks, Rodney A. "Intelligence Without Representation." Artificial Intelligence 47, 1-3 (1991): 139-159

Burks, Arthur W., Herman H. Goldstine, and John von Neumann. "Preliminary Discussion of the Logical Design of an Electronic Computing Instrument." (1946), In The Origins of Digital Computers: Selected Papers. Edited by Brian Randell, 399-413. Berlin, New York: Springer-Verlag, 1973

Castelvecchi, Davide. "Can We Open the Black Box of AI?" Nature 538, no. 7623 (2016): 20-23

Chun, Wendy Hui Kyong. Programmed Visions: Software and Memory. Cambridge Mass.: MIT Press, 2011

Deleuze, Gilles. Différence et Répétition. Paris: Presses Universitaires de France, 1968

Derrida, Jacques. "La Pharmacie de Platon." In La Dissémination. Paris: Éditions du Seuil, 1988

Dreyfus, Hubert L. What Computers Can't Do: The Limits of Artificial Intelligence. New York: Harper and Row, 1972

Dyke, Heather, and Adrian Bardon, eds. A Companion to the Philosophy of Time. Malden: Wiley Blackwell, 2013

Elman, Jeffrey L. "Finding Structure in Time." Cognitive Science 14, no. 2 (1990): 179-211

Goldstine, Herman H., and John von Neumann. "Planning and Coding of Problems for an Electronic Computing Instrument: Report on the Mathematical and Logical Aspects of an Electronic Computing Instrument." Part II, Volume II, The Institute for Advanced Study, Princeton (N.J.), 1948

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep Learning. Adaptive Computation and Machine Learning. Cambridge (Mass), London: The MIT Press, 2016

Graves, Alex. "Generating Sequences with Recurrent Neural Networks.", 2013

Hayles, Katherine. How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics. Chicago (Ill.): University of Chicago Press, 1999

Hinton, Geoffrey E. "Learning Distributed Representations of Concepts." Proceedings of the Eighth Annual Conference of the Cognitive Science Society 1 (1986): 1-12

Hochreiter, Sepp, and Jürgen Schmidhuber. "Long Short-Term Memory." Neural computation 9, no. 8 (1997): 1735-1780

Horkheimer, Max, and Theodor W. Adorno. Dialektik Der Aufklärung: Philosophische Fragmente. Frankfurt am Main: Fischer, 1944

Hoy, David Couzens. The Times of Our Lives: A Critical History of Temporality. Cambridge (Mass.), London: The MIT Press, 2009

Hui, Yuk. On the Existence of Digital Objects. Minneapolis (Minn.), London: University of Minnesota Press, 2016

Hui, Yuk. "Archives of the Future: Remarks on the Concept of Tertiary Protention." Inscription, 2018, 129-151

Hui, Yuk. Recursivity and Contingency. London, New York: Rowman & Littlefield, 2019

Husserl, Edmund. Vorlesungen zur Phänomenologie des inneren Zeitbewusstseins. 3. Auflage, unveränd. Nachdr. d. I. Aufl. 1928. Tübingen: Max Niemeyer Verlag, 2000

Jordan, Michael I. "Serial Order: A Parallel Distributed Processing Approach." ICS Report 8604, Institute for Cognitive Science, University of California, San Diego, La Jolla (Cal.), 1986

Kittler, Friedrich A. Draculas Vermächtnis: Technische Schriften. Leipzig: Reclam Verlag, 1993

Koselleck, Reinhart. Vergangene Zukunft: Zur Semantik Geschichtlicher Zeiten. Frankfurt am Main: Suhrkamp, 1979

Kowalski, Robert. "Algorithm = Logic + Control." Communications of the ACM 22, no. 7 (1979): 424-436

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey Hinton. “ImageNet Classification with Deep Convolutional Neural Networks.” Proceedings of the 25th International Conference on Neural Information Processing Systems 1 (2012): 1097–1105

LeCun, Yann. “Generalization and Network Design Strategies.” Technical Report CRG–TR-89-4, University of Toronto Connectionist Research Group, 1989

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. “Deep Learning.” Nature 521, no. 7553 (2015): 436–444

LeCun, Yann A., Léon Bottou, Genevieve B. Orr, and Klaus-Robert Müller. “Efficient BackProp.” (1998), In Neural Networks: Tricks of the Trade: Second Edition. Edited by Grégoire Montavon, Geneviève B. Orr and Klaus-Robert Müller, 9–48. Berlin, Heidelberg: Springer, 2012

Marino, Mark C. Critical Code Studies. Software Studies. Cambridge (Mass.): The MIT Press, 2020

McCulloch, Warren S. “Why the Mind Is in the Head?” dialectica 4, no. 3 (1950): 192–205

McCulloch, Warren S., and Walter Pitts. “A Logical Calculus of the Ideas Immanent in Nervous Activity.” The Bulletin of Mathematical Biophysics 5, no. 4 (1943): 115–33

Montavon, Grégoire, Geneviève B. Orr, and Klaus-Robert Müller, eds. Neural Networks: Tricks of the Trade: Second Edition. Berlin, Heidelberg: Springer, 2012

Neumann, John von. “The General and Logical Theory of Automata.” John von Neumann Collected Works Volume V: Design of Computers, Theory of Automata and Numerical Analysis (1948): 288–326

Neumann, John von. The Computer and the Brain. New Haven, London: Yale University Press, 1958

Neumann, John von. “First Draft of a Report on the EDVAC.” (1945), In The Origins of Digital Computers: Selected Papers. Edited by Brian Randell, 383–392. Berlin, New York: Springer-Verlag, 1973

Piccinini, Gualtiero. “The First Computational Theory of Mind and Brain: A Close Look at Mcculloch and Pitts's “Logical Calculus of Ideas Immanent in Nervous Activity”” Synthese 141, no. 2 (2004): 175–215

Randell, Brian, ed. The Origins of Digital Computers: Selected Papers. Berlin, New York: Springer-Verlag, 1973

Roberts, Ben. “Cinema as Mnemotechnics: Bernard Stiegler and the “Industrialization of Memory”” Angelaki 11, no. 1 (2006): 55–63

Rosenblueth, Arturo, Norbert Wiener, and Julian Bigelow. “Behavior, Purpose and Teleology.” Philosophy of Science 10, no. 1 (1943): 18–24

Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. “Learning Internal Representations by Error Propagation.” University of California, San Diego, Institute for Cognitive Science, La Jolla (Cal.), 1985

Russell, Stuart, and Peter Norvig. Artificial Intelligence: A Modern Approach. Third edition. Boston: Pearson, 2016

Samek, Wojciech, Grégoire Montavon, Andrea Vedaldi, Lars Kai Hansen, and Klaus-Robert Müller, eds. Explainable AI: Interpreting, Explaining and Visualizing Deep Learning. Cham, Switzerland: Springer Nature, 2019

Simondon, Gilbert. Du Mode d'Existence des Objets Techniques. Paris: Aubier, 1958

Stiegler, Bernard. La Technique et le Temps: 2. La Désorientation. Paris: Éditions Galilée, 1996

Stiegler, Bernard. La Technique et le Temps: 3. Le Temps Du Cinéma et la Question du Mal-Être. Paris: Éditions Galilée, 2001

Wiener, Norbert. Cybernetics: Or Control and Communication in the Animal and the Machine. Cambridge (Mass.): MIT Press, 1948