Automated Societal Option Maximization

Manfred Füllsack
University of Graz, Austria
manfred.fuellsack/at/uni-graz.at

> Abstract • In this commentary I take up thoughts about the sociology of time and Luhmann's theory to consider whether automated future option maximization can be extended with a social dimension.

Handling Editor • Alexander Riegler

« 1 » Elena Esposito shares some nice biographical anecdotes about her encounters with Niklas Luhmann's sociology, spiced with some scattered remarks about her current research on changing time perceptions in the context of digital transformation. Since my own interests are also strongly influenced by Luhmann's systems theory, and my current research revolves around predictive capabilities of digital machines, I find many of Esposito's thoughts stimulating. However, there are a few aspects I find difficult to subscribe to. In particular, I have doubts about Esposito's proposition of a substantial difference between "algorithmic forms of prediction" and probabilistic ones (§20). If anything, I would agree with a quantitative difference. Distributed computing, the speed of GPUs and the like, allow the building of predictions on different amounts of data, and in some cases, this data may be processable more thoroughly than with probabilistic techniques. Even so, the core of most machine-learning tools is statistics, after all, which provides probabilistic predictions that – yes, of course – are misinterpreted at times. This seems a point that Luhmann might have made similarly: the techniques themselves and their perception should not be confused. Today, even lay persons are used to dealing with statistical statements and, by and large, do not attribute them to individual fates. We have learned that an average life expectancy of 80 years is no death sentence for every 80-year-old. We know that the underlying uncertainty is to be shared. By comparison, the predictions of a neural network with, say, 20 hidden layers of thousands of neurons each appear as opaque, today, as the methods of Thomas Bayes or Carl Pearson may have seemed in their times. Indeed, even for specialists, the self-learning capabilities of deep neural networks seem eerie and many details are not understood (Castelvecchi 2016). It indeed seems "divinatory" and maybe even "divine" when the Go-playing program AlphaZero makes AlphaGo, which beat the human champion in 2017, look old after a few hours of training with nothing but a copy of itself (Silver et al. 2018). However, I disagree with Esposito that this is a fundamentally different kind of divination from, say, that of Bayesian inference as perceived in the 19th century. In view of current efforts to make deep learning transparent (Adadi & Berrada 2018; Barredo et al. 2020; Mahbooba et al. 2021), it seems likely that in less than the amount of time that lies between Pearson and us, the capabilities of today's AI will appear much more mundane than divinatory.

« 2 » However, this remains to be seen. In the rest of this commentary, I would like to focus on an aspect that Esposito hints at in §18 by drawing a line from sociology of time to Luhmann's theory and that is central to my own current research. Last year, we published a small compilation of reports (Hornischer et al. 2020) on experiments with a method, Future State Maximization, that is quite well explained by a principle that Heinz von Foerster called the "Ethical Imperative:" Always act so as to increase the number of your options. Interestingly, this imperative can be relatively easily implemented in a computer model to be used for the optimization of tasks, which are currently popular with machine-learning methods. Picture, for example, a small robot with just a handful of sensors and actuators. It seems conceivable that the immediate (i.e., the next-step-) consequences of all the actions this robot can take can be compared to one another using a simple brute-force algorithm. The robot then is made to take the action that offers the greatest number of further options. In this way, it will, for example, avoid moving into corners since corners impede movement. This method works fine, to some extent and yields surprising effects in many quite complex tasks. For example, Sergio Cerezo, Guillem Ballester, and Spiros Baxevanakis (2018) show that the method can beat many other advanced machine-learning tools in most of the games in the "OpenAI-gym" (a simulated programming environment for developing and comparing reinforcement-learning algorithms). However, it seems clear that the number of options even for a simple robot grows exponentially if the principle is to be computed for more than just the next time step, i.e., if the robot is made to maximize options for five, ten or even more time steps into the future. Think of the moves in the game of chess, where the "future present" quickly becomes too complex to compute straight away.

« 3 » As in chess programs, the complexity problem can be remedied with an intuitive solution, i.e., by statistically sampling the option space. The option maximization principle is made to consider not all future paths, but a selection that is assumed to be representative (Hornischer et al. 2020). However, this solution entails uncertainty, which grows rapidly with the time horizon and eventually brings the principle to its limit. The trade-off between sample sizes large enough to minimize uncertainty and small enough to accept computational costs caps option maximization. Beyond this limit, the future remains vague and unpredictable, after all.

« 4 » The question, thus, is how can this limit be extended? Esposito's tentative comments on Luhmann's social theory and on her current project brought up some ideas. The core problem of the principle to "always act so as to maximize options" is its rapidly opening future horizon, or in other words, its huge possibility space, which is far too vast to render a single option sufficiently likely to be chosen in the first place.

« 5 » By itself the possibility space for social interactions is simply too large: The problem is not that some specific actions are impossible but, rather, that there are too many actions possible for one of them to be selected for initiating social interaction. In order to make something happen on which further interactions can build, the space must be reduced, i.e., the number of options constrained. Initially, this constraint can take place by chance, as in the well-known example of the train compartment (Parsons 1951; Füllsack 2018). If the interaction resulting from the constraint turns out to be viable and thus provides a base for further interactions, the subsequent iteration of the interaction finds an already reduced possibility space in which a new interaction can further reduce options and thus increase the likelihood of next-order interactions being based on it, which in turn further reduce options, and so on. Social interaction becomes likely because the possibility space is iteratively narrowed down with each subsequent viable and connection-generating interaction. Our daily social interactions can happen because they find these reduced possibility spaces as a pre-given that results from the preceding interactions of all past human beings who, by interacting with each other, contributed to developing our language, culture, and knowledge. The probability of answering a friendly "Have a good weekend" with a nice "Same to you" is sufficiently high because the pre-constrained option space of this interaction makes some actions much more likely than others.

« 6 » I suggest that the option-maximizing handling of futures could work in similar ways. According to an estimation by Stuart Russell (2019), the number of motor-control choices a human makes in a lifetime runs up to about twenty trillion, opening up an option space that is literally immeasurable. We handle this inconceivable vastness not with a brute-force search of options, but on the groundwork of a deeply nested hierarchy of "subroutines," which equip us with temporarily relevant sub-goals that lead our actions. Similarly, chess programs handle enormous numbers of move branches, which depend on whatever move their opponent might choose, not by calculating the consequences of each and every possible move, but by pre-estimating the reasonableness of moves and distributing calculation time with regard to these estimates. In early chess programs up to Deep Blue, these estimates were gained from human expertise, in particular from the experience of grandmasters. In recent Go programs such as AlphaZero, the estimates are gained through thousands of test games against the program's copy. In a similar way, I propose that it might be possible to use "digital test pilots" to pre-structure (and thus pre-constrain) vast future option spaces. Currently, we just deploy one agent to sample an option space (or to be more precise, one agent with a high number of so-called "walkers," which represent the sample instances; see Hornischer et al. 2020). Alternatively, a society of option-maximizing agents (with each of them using high numbers of walkers, in turn) could perhaps be made to build structures into the future's horizon, which would constrain the temporal cone of uncertainty down to an extent at which the hit rate of predictions improves. These structures would then provide ground on which further explorations for maximizing option could build. In some way, this seems to be similar to what Luhmann (1997) considered as functionally differentiated societal sub-systems, which in their own particular logic supply modern society with essential functionalities. For example, while the procurement of food may have been difficult for early humans, making survival highly uncertain, an economic system such as that of modern society renders the supply of large parts of the world's population highly predictable. However, the difference from the achievements of past human social interactions so far, i.e., culture, language, and knowledge, is that it is not we who would have to run sightlessly into uncertain (and possibly fatal) futures to build structures from which our horizon would extend, but digital substitutes doing the job for us.

« 7 » No doubt, humanity disposes of structures on the basis of which it is able to predict futures with sometimes impressive accuracy. The daily weather forecast may be just one simple example. With an automated structuration mechanism, however, implemented along the line of option maximization or some other self-learning principle, I might re-consider my initially raised doubts about the difference between digital and probabilistic predictions. The automated pre-constraint of option spaces, based on the contingency of emerging cooperation among option-maximizing digital agents, could result in a difference that makes a difference to algorithmic forms of prediction – among other things, simply because of the relative effortlessness with which futures then could be screened.

« 8 » We are not there yet, though. Future State Maximization works well in the OpenAI gym and some rather limited domains of out-of-the-lab applications. Whether a *societal* implementation is feasible and would make a significant difference remains to be seen. For the moment, therefore, I would be cautious and rather not read too much into a perceived opaqueness of digital tools. However, I agree with Esposito that sociologically, the question of why, today, we seem ready to assign divinity to machines is enthralling.

References

Adadi A. & Berrada M. (2018) Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access 6: 52138–60. https://doi.org/10.1109/ACCESS.2018.2870052

Barredo A., Díaz-Rodríguez N. A., Del Ser J., Bennetot A., Tabik S., Barbado A. & Garcia S. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion 58: 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Castelvecchi D. (2016) Can we open the black box of AI? Nature 538(7623): 20–23. https://doi.org/10.1038/538020a

Cerezo S. H., Ballester G. D. & Baxevanakis S. (2018) Solving Atari games using fractals and entropy. ArXiv: 1807.01081 [Cs]. http://arxiv.org/abs/1807.01081

Füllsack M. (2018) Plasticity, granularity and multiple contingency: Essentials for conceiving an artificial constructivist agent. Constructivist Foundations 13(2): 282–291. ▸ https://constructivist.info/13/2/282

Hornischer H., Plakolb S., Jäger G. & Füllsack M. (2020) Foresight rather than hindsight? future state maximization as a computational interpretation of Heinz von Foerster's ethical imperative. Constructivist Foundations 16(1): 36–49. ▸ https://constructivist.info/16/1/036

Luhmann N. (1995) Social systems. Stanford University Press, Stanford CA. German original published in 1984.

Luhmann N. (1997) Die Gesellschaft der Gesellschaft. Suhrkamp, Frankfurt am Main.

Mahbooba B., Timilsina M., Sahal R. & Serrano M. (2021) Explainable artificial intelligence (XAI) to enhance trust management in intrusion detection systems using decision tree model. Complexity 2021: e6634811. https://doi.org/10.1155/2021/6634811

Parsons T. (1951) The social system. Free Press, Glencoe IL.
Russell S. (2019) Human compatible: AI and the problem of control. Penguin, London.
Silver D., Hubert T., Schrittwieser J., Antonoglou I., Lai M., Guez A., Lanctot M., Sifre L., Kumaran D., Graepel T. Lillicrap T., Simonyan K. & Hassabis D. (2018) A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science 362(6419): 1140–1144. https://doi.org/10.1126/science.aar6404

Manfred Füllsack is Professor of Systems Sciences at the University of Graz. His research includes: systems, complexity, networks, games and computational theory, work – its history, its sociology, its economy, and its philosophy, and computer-based simulations.

Funding: This contribution is not subject to any funding.
Competing interests: The author declares that he has no competing interests.

Received: 25 May 2021
Revised: 1 June 2021
Revised: 25 June 2021
Accepted: 25 June 2021

How Can Algorithms Participate in Communication?

Bruce Clarke
Texas Tech University, USA
brunoclarke/at/gmail.com

> Abstract • Esposito's theoretical approach indicates the fertility, first, of transplanting social systems theory into other fields, and next, of bringing classical cybernetic topics such as computation by algorithms back into Luhmann's multi-modal constructivist framework of differentiated system operations.

Handling Editor • Alexander Riegler

« 1 » In my commentary, I consider some key statements of Elena Esposito's interview from my vantage as an Anglophone and interdisciplinary non-specialist reader of Niklas Luhmann's text. The ultimate issue that I draw out is the one intimated in my title's adaptation of Luhmann's essay "How Can the Mind Participate in Communication?" (Luhmann 2002a). The short answer Esposito gives regarding algorithms may be summarized like so: algorithms participate in communication in the same way natural minds do, by manipulating the medium of meaning, but unlike minds, they do so without needing to construct an understanding of what they are doing.

"[W]hat I and a couple of friends found very interesting in the Luhmann–Habermas debate was the Luhmann side of it" (§2)

« 2 » Esposito observes Luhmann from within his dedicated academic discipline of sociology. Hers is a lucid insider's report on the theoretical workings and the current status of Luhmann's sociological theories. Esposito clarifies the differences in the circumstances of my own reception of Luhmann's work in English translation and provokes some thoughts on the status of Anglophone Luhmann scholarship.

« 3 » For instance, the published record of the debate between Luhmann and Jürgen Habermas was not a formative event for me, and its interest followed only after my initiation into social systems theory's cybernetic extraction. Having studied that connection at some length, I would summarize Luhmann's conceptual innovation in this way: Luhmann made an effective selection from the epistemological and otherwise philosophical versions of cybernetics coming out of the 1970s systems counterculture in the works of Gregory Bateson, Heinz von Foerster, George Spencer Brown, Humberto Maturana, and Francisco Varela (Clarke 2020a: 101–156), and set that synthesis quite neatly down within post-war sociology's disciplinary formations around Talcott Parsons's *The Social System* (1977).

« 4 » This major theoretical move foregrounds communication and meaning as the phenomena to be accounted for. One does not have to speak cybernetics (although it helps) to operate Luhmann's social systems theory at the level of its sociological themes. For instance, the term *autopoiesis* does not come up in Esposito's discussion. However, at the same time, one does have to speak constructivism. Luhmann's work enacts what is still the most far-reaching development of second-order cybernetics' epistemological themes, rivalling if not surpassing the conceptual evolutions of the second-order cyberneticists themselves, Varela included. If one could get beyond the factionalism that besets much of contemporary cybernetic discourse, one would find in Luhmann the full sociological and epistemological counterpart of the cognitive, neurological, and phenomenological work that Maturana and Varela mediated through von Foerster's Biological Computer Laboratory. In any event, this is the thesis at large in the essay collection I coedited a decade ago (Clarke & Hansen 2009).

"Luhmann has only a narrow group of followers in the English-speaking world" (§4)

« 5 » Matters appear somewhat more favorable in the British sociological academy, but for the normal run of American sociology departments, as near as I can tell, Luhmann's work still appears to be too heterodox. Lacking (like cybernetics more broadly) an institutionally solid disciplinary base of operations, Luhmann's Anglophone "following" is relatively dispersed. Rather, Luhmann's texts have circulated in literature, philosophy, and media studies departments, buffeted amidst other concurrent theoretical movements or fads, such as object-oriented ontology, speculative realism, the new materialism, Anthropocene studies, posthumanism, and the like.

« 6 » In the international Anglophone academy, sociologist Dirk Baecker (2021) and philosopher Hans-Georg Moeller (2012) anchor the field. Luhmann's theory is being ably developed for the environmental humanities by Hannes Bergthaller (2014), for critical theory by Anders la Cour and Andreas Philippopoulos-Mihalopoulos (2013), for media studies by Richard Grusin (2010), for new materialism and posthumanism by Anna Henkel (2016), for economic theory by Brian Massumi (2014), for politics and international relations by Hannah Richter (2020), and for literary studies by Mark Seltzer (2016). Beginning with *Critical Environments*, Cary Wolfe's (1998) work has deeply developed social systems theory alongside deconstruction, animal studies, and posthumanist discourse. My own work (Clarke 2008, 2014, 2020a,