to anticipation. In: Pezzulo G., Butz M. V., Sigaud O. & Baldassarre G. (eds.) Anticipatory behavior in adaptive learning systems: From psychological theories to artificial cognitive systems. Springer, Berlin: 132–151.

Montebelli A., Lowe R. & Ziemke T. (2013) Toward metabolic robotics: Insights from modeling embodied cognition in a biomechatronic symbiont. Artificial Life 19(3_4): 299–315.

Navarro-Guerrero N., Lowe R. J. & Wermter S. (2017) Improving robot motor learning with negatively valenced reinforcement signals. Frontiers in Neurorobotics 11: 10. https://www.frontiersin.org/articles/10.3389/fnbot.2017.00010

Schulkin J. (2011) Social allostasis: Anticipatory regulation of the internal milieu. Frontiers in Evolutionary Neuroscience 2: 111. https://www.frontiersin.org/articles/10.3389/fnevo.2010.00111

Seth A. K. (2015) The cybernetic Bayesian brain: From interoceptive inference to sensorimotor contingencies. In: Metzinger T. & Windt J. M. (eds.) Open mind. MIND Group, Frankfurt am Main: 35(T). https://open-mind.net/papers/the-cybernetic-bayesian-brain

Sterling P. (2004) Principles of allostasis: Optimal design, predictive regulation, pathophysiology, and rational therapeutics. In: Schulkin J. (ed.) Allostasis, homeostasis, and the costs of physiological adaptation. Cambridge University Press, Cambridge: 17–63.

Sterling P. (2012) Allostasis: A model of predictive regulation. Physiology & Behavior 106(1): 5–15. ▸ https://cepa.info/5330

Robert Lowe is docent in Cognitive Science, and Associate Professor, in the Department of Applied IT at the University of Gothenburg. He is the research leader of DICE Lab (https://robertlowe2.gitlab.io/publii-dicelabbers). His research interests are focused on computational modelling and gamification of cognitive science tasks focused on memory and decision making as affected by (social) interaction and emotion. He works with robots to study how such models can be used as cognitive controllers in physically embodied artificial agents. He teaches courses in Artificial Intelligence focused on introduction to Deep Learning and Deep Reinforcement Learning in the context of gaming.

Received: 30 October 2020
Accepted: 7 November 2020

The Relation between Future State Maximization and von Foerster's Ethical Imperative

Gordana Dodig-Crnkovic
Chalmers University of Technology, Sweden • dodig/at/chalmers.se

> Abstract • I review von Foerster's (bio) computational approach to cognition in relation to foresight and hindsight, and to his Ethical Imperative. For him, ethics must remain implicit and becomes manifest through metaphysics and dialogics. Only under this condition can the relevance of Future State Maximization for the system that is modelled be established.

Handling Editor • Alexander Riegler

Cognition and computation

« 1 » In "Notes on an Epistemology for Living Things" (Foerster 2003: 247) von Foerster describes cognitive processes as recursive computation, the eigenvalue of which is what we refer to as the physical world. He uses "cognition for representing something which is computing a reality" (quoted in Müller 2008: 64). However, "[o]bjects and events are not primitive experiences. Objects and events are representations of relations" (Foerster 2003: 250). That which, in everyday speech, is referred to as the "external world" manifests itself through computations of specific relations that are representations of those relations: "Reality appears only implicit as the operation of recursive descriptions" (ibid: 216). The nervous system is (self-)organized so that it computes "a stable reality for an agent" (ibid: 244).

« 2 » Von Foerster argues for recursive computation as an underlying principle of all cognitive processes, *of life itself*, as he illustrates with the evolutionary examples of independent sensorimotor units in protozoa (Foerster 2003: 217). Very much ahead of his time, von Foerster conjectured "molecular computation," arguing for the importance of molecules in the dynamics of cognition (ibid: 130). It is very important to note that von Foerster's notion of computation corresponding to biological processes denotes "any operation (not necessarily numerical) that transforms, modifies, rearranges, orders, and so on, observed physical entities ('objects') or their representations ('symbols')" (Foerster 2003: 216). About ordering, he distinguishes between two levels:

> "The one is when we wish to make a description of a given arrangement of things. The other one when we wish to re-arrange things according to certain descriptions. It will be obvious at once that these two operations constitute indeed the foundations for all that which we call 'computation.'" (ibid: 194)

« 3 » Talking about computation processes in living organisms, von Foerster (2003: 194) argues that living organisms interact with their environment in one or both of the following ways:

- developing "languages" (sensors, neural codes, motor organs, etc.) increasingly adapting to the environment; and/or
- changing the environment according to their own needs.

He adds that "whatever option they take, *it will be done by computation*" (ibid, my emphasis).

« 4 » Since biological systems have a recursive organization, programs in biological computers also can be computed on, so there are "metaprograms," "meta-metaprograms," etc. This is the kind of relationships found in the organization of neural interactions. It is, according to von Foerster (2003: 70), "to a certain extent, the genetic program that produces anatomical – read 'geometrical' – constraints which prohibit, within certain limits, arbitrary developments of conceivable structures."[1]

Modelling physical objects and properties

« 5 » In von Foerster's biocomputational definition of life, physical objects are at the bottom of computational recursion, which constitutes cognitive processes. The more evolutionarily developed (complex) organisms are, the more levels of recursive computations and programs they have. Simple unicellular organisms such as bacteria are at

1 | These insights into the nature of cognition as bio-computation make von Foerster the antecedent of contemporary natural computing and morphological computing (Dodig-Crnkovic 2017).

the bottom of the computational hierarchy. Computational building blocks are molecular computations that constitute life processes of a cell. For a bacterium, its body serves as a model, in the sense of Rodney Brooks's "the world is its own best model" (Brooks 1991). His insectoid robots, similar to Braitenberg's vehicles (Braitenberg 1989), function well without internal models of their environment. These cases are in contradiction to Hannes Hornischer et al.'s claim in §1 of their target article that "space and time are constructions emerging from the use of models." Rather, for a bacterium, space and time emerge directly from physics.

« 6 » One of the abilities of biological systems that Robert Rosen (2012) studied is anticipation. Bacteria learn directly through their adaptive morphology and by exchanging pieces of genetic code. They are an example of computation that "transforms, modifies, rearranges, orders, and so on, observed physical entities" (Foerster 2003: 216).

« 7 » In more complex organisms, responses to environmental changes are computed recursively at different levels of organization, where memory and hindsight, together with perception and logical inference help to build foresight, and anticipate and examine possible actions before executing them in the environment (Dodig-Crnkovic 2017).

Ethics as an "underground river"

« 8 » In their target article, Hornischer et al. present Future State Maximization as a computational interpretation of the Ethical Imperative, "Act always so as to increase the number of choices." As they note in Footnote 1, von Foerster later changed it to separate ethics from morality. In an interview with Bernhard Poerksen he was very explicit about it:

> "I didn't choose my words very carefully when I said that. It would have been better if I had written, 'Heinz, act always so as to increase the number of choices'." (Foerster & Poerksen 2002: 37)

« 9 » With this, von Foerster expresses the idea that ethics must be in the background, serving as a basis, where it remains implicit. He referred to it as the "underground river of ethics" (Foerster 2003: 291). Only with the help of what he calls "metaphysics" and "dialogics" can ethics become manifest without becoming explicit. (Foerster 2003: 291). Von Foerster's *metaphysics* has a very specific meaning, as it emerges in the realm of undecidable questions: "[W]e become a metaphysician any time we decide upon in principle undecidable questions" (ibid), which leads to his metaphysical postulate: "Only those questions that are in principle undecidable, we can decide." With this von Foerster draws attention to the necessity of choice, and argues that the concept of objectivity is nothing but a pretext for avoiding responsibility that comes with choice.

« 10 » In *dialogics*, language reaches out to others as the "root of conscience," such that "ethics invisibly manifests itself through dialogue" (Foerster 2003: 297). This communicative, social aspect of ethics is necessary, and it is implemented in a feedback loop where shared values are negotiated through a dialogue.

« 11 » The maximization of future states is thus a metaphysical mechanism, but on top of the metaphysical principle there must be dialogics to provide context in which shared values are embedded, and to provide criteria for which of the many choices are good.

« 12 » The connection made in the target article taking Future State Maximization to be a computational interpretation of von Foerster's Ethical Imperative prompts the question of how representative such a computational model can be for human agency. Here the historical connection can be instructive. In spite of being aware that autopoietic systems require embodiment, Ricardo Uribe, back then working in von Foerster's Biological Computer Laboratory, wrote the first computer programs that visualized autopoiesis (Varela, Maturana & Uribe 1974). The authors must have considered the model sufficiently representative of biological autopoietic mechanisms. Yet how far can one drive interpretations based on analogy between simulation and a physically embodied cognitive system when it comes to *human ethics*? If we follow von Foerster, the dialogic aspect must come in, establishing the relevance for the system that is modelled.

« 13 » To summarize, given von Foerster's understanding of life as a bio-computational process, one should acknowledge different organizational and functional levels of computational processes in living systems, from bacteria to humans. Evolutionarily, cognitively and historically, there are clear differences between the human ethics to which von Foerster refers (being the result of the dialogic process providing criteria for the maximization of choices) and the optimization processes such as those suggested in the target article.

Foresight and simulation models

« 14 » Foresight (the ability to foresee and prepare adequately for the future) is a result of envisaging possible futures so as to be prepared to act under various scenarios. This acting will necessarily alter future options, as we are intervening in a dynamical and unpredictable system, "the white water world" (Pendleton-Jullian & Seely Brown 2019). In the spirit of cybernetics, one is navigating unknown waters, which demands flexibility, agility and adaptation, thus both changing ourselves and changing the world around us. For example, in the current Covid-19 crisis, foresight is essential, as we must act, and waiting for certitude about the current and future situations can cost human lives.

« 15 » What makes simulations important, today, is not their internal maximization of space for choices for the agents in the model, but their connection to the problem they are a model of, thus providing good tools for foresight. Having maximal freedom of exploration within the model space does not guarantee the adequate solution of the problem that is modelled. In the case of Covid-19 pandemics, computational models were indeed used for foresight and they helped decision-makers to understand possible future developments (Squazzoni et al. 2020). Along with the exploration of the space of possible future states, the dialogical process of examining values of possible alternative futures is central. In that process, hindsight informs foresight, while lessons learned from acting upon foresight form the basis for future hindsight in a circular manner.

References

Braitenberg V. (1989) Vehicles: Experiments in synthetic psychology. MIT Press, Cambridge MA.

Brooks R. A. (1991) Intelligence without representation. Artificial Intelligence 47(1): 139–159.

Dodig-Crnkovic G. (2017) Nature as a network of morphological infocomputational processes for cognitive agents. The European Physical Journal 226: 181–195.

Foerster H. von (2003) Understanding understanding: Essays on cybernetics and cognition. Springer, New York.

Foerster H. von & Poerksen B. (2002) Understanding systems: Conversations on epistemology and ethics. IFSR International Series in Systems Science and Systems Engineering, Vol 17.

Müller A. (2008) Computing a reality: Heinz von Foerster's lecture at the A.U.M conference in 1973. Constructivist Foundations 4(1): 62–69. ▸ http://constructivist.info/4/1/062

Pendleton-Jullian A. M., Seely Brown J. (2019) Design unbound: Designing for emergence in a white water world. Volume 2: Ecologies of change. MIT Press, Cambridge MA.

Rosen R. (2012) Anticipatory systems: Philosophical, mathematical, and methodological foundations. Springer, New York.

Squazzoni F., Polhill J. G., Edmonds B., Ahrweiler P., Antosz P., Scholz G., Chappin É., Borit M., Verhagen H., Giardini F. & Gilbert N. (2020) Computational models that matter during a global pandemic outbreak: A call to action. Journal of Artificial Societies and Social Simulation 23(2): 10. http://jasss.soc.surrey.ac.uk/23/2/10.html

Varela F. J., Maturana H. R. & Uribe R. (1974) Autopoiesis: The organization of living systems, its characterization and a model. Biosystems 5(4): 187–196. ▸ https://cepa.info/546

Gordana Dodig-Crnkovic is Professor of Interaction Design at Chalmers University of Technology and Professor of Computer Science at Mälardalen University, Sweden. She holds PhD degrees in Physics and Computer Science. Her research focuses on the relationships between computation, information and cognition, including ethical and value aspects. She is a member of the editorial board of the Springer SAPERE series, World Scientific Series in Information Studies, and various journals. She is a member of the AI Ethics Committee at Chalmers and the Karel Capek Center for Values in Science and Technology. http://gordana.se

Received: 15 October 2020
Accepted: 12 November 2020

Future State Maximisation and Hard-Wired Structures

Bernd Porr
University of Glasgow, UK
bernd.porr/at/glasgow.ac.uk

> Abstract • Future state maximisation (FSX) offers a clear departure from both fixed reactive systems and systems that learn models based on reactive experience. In this commentary, I argue that FSX is not replacing these previous ideas but that, rather, the interplay with FSX and the older concepts create a powerful hybrid.

Handling Editor • Alexander Riegler

« 1 » In their synthetic target article, Hannes Hornischer et al. put Future State Maximisation (FSX) in the context of constructivism, in particular Heinz von Foerster's Ethical Imperative that an agent should maximise its future possible states. Crucial here is that these possible states are the states constructed by the agent from its own perspective. This forms a *closed loop* by establishing possible action–state pairs by the agent, however, not reactive by learning from experience but by the agent simulating its possible future states and maximising them. The authors successfully show, with the help of three instructional examples, that FSX complies with the central concepts of constructivism, in particular, that only quantities are processed and that behaviour can emerge from noise – again, as von Foerster postulated.

« 2 » I would like to challenge the authors on the thorny issue of "representation" in the context of constructivism. Radical constructivists reject the simple "mapping" of the physical world to a representation in the subjects' heads in terms of a distorted but perhaps "truthful" representation of the outside world. The authors make clear that they, too, reject this notion (§3). However, they then use a form of representation in Examples 1 (§§26–34) and 2 (§§35–43) when they introduce a 2D space and the ability of the agent to perceive this space in an "objective" way. This is clearly against radical constructivist thinking. However, the third example (§§44–53) is again fully compatible with constructivist thinking, because the agent's learned behaviour and the resulting collective behaviour emerges just from noise and without any qualities transmitted via its sensors. Indeed, this is the purest form, in terms of radical constructivism, because the input space codes just quantities and no qualities, which then hopefully correlate with the actions, closing the action–sensor loop. It is thus comparable to a classical instrument flight and Maturana's fictional submarine pilot who is just reacting to signals, and then turning dials – not knowing it is a submarine he is controlling. The first examples in the target article could just be dismissed as incompatible with radical constructivism, but I would argue that they also open up new avenues that might even be acceptable for radical constructivists: what do we gain by introducing hard-wired knowledge of, for example, the spatial relationships of objects in relation to the agent? This can be done in a way that will not violate the important notion in radical constructivism of not having qualities transmitted. There is no need to call these "objects," as this would imply a "mind-independent" reality. Rather, we can just stick to mathematical relationships between certain input variables, which can then be used internally to perform further computations. Elizabeth Spelke and Katherine Kinzler (2007) call hard-wired computations about the physical world "core knowledge." For example, babies have the ability to know the spatial relationship between objects and themselves. A radical constructivist might object at this point and demand that perception needs to "emerge from noise," i.e., learning this spatial relationship from scratch. However, this could have happened through evolution, which creates structures in the agent. These structures precompute certain quantities from the sensor inputs and these can then be used further downstream to make decisions about which actions should be taken. For example, visual pixels on the retina of an agent can be precomputed to egocentric coordinates of objects in the visual field of the agent. A frog will find it useful to have the pre-computed coordinates of black elongated stimuli (e.g., flies) at hand in case it becomes hungry.

« 3 » Furthermore, I argue that the first two examples in the target article, which assume knowledge of spatial relationships,