implies that there are variations of some internal states. One could object to the claim that the propagation of the gradient during the stabilization of the network could be seen as a variation of some internal states. To this I would reply that gradient propagation aims at stabilizing the system in a non-autonomous dynamic state, because after the convergence of the propagation algorithm, if the inputs do not change, the system will not vary or evolve. Yet, an autonomous system is changing in time because these changes allow for an evaluation of its internal states and particularly its level of degradation. The dynamic of the system is always changing: it is an organization that can disappear because it evolves. This is why some researchers interested in the simulation of autonomous systems are using models such as recurrent and temporal neural networks, which have their own internal dynamic (Manetti & Caiani 2011) and which can adopt different operating regimes, i.e., different sensorimotor behavioral regularities. Inputs become perturbations of these regularities, as do other agents and the entire environment of the system.

« 12 » My arguments seem to be in line with Füllsack, who prefers talking about "irritation" (§§9f) rather than "external inputs." However, for the reasons mentioned above, I think that the proposition of minimal artificial constructivist agents as systems based on feed-forward neural networks, even if they are linked together, does not implement the autonomous dynamic as presented by Varela with the notion of organization as agency.

In an enactive perspective, how to deal with phenomenal domain separation and reflexivity?

« 13 » The most difficult question is certainly that of phenomenal domains. While this notion could be seen in various ways I prefer referring to a kind of "universe of contents" that has a meaning for a certain point of view. In the same way that an agent is seen as an agent by an external observer (§20), there are different levels of phenomenal domains in the agent, each with a specific nature and without shared sense. For instance, if I write on a sheet with a pen, I not only use my hand and my fingers for writing but also my gaze for checking whether the shape of my writing respects handwriting rules. I feel the pen in my hand and the marks on the sheet confirm my action. It is in a phenomenal domain that has no relation to the sense of the sentence that I write.

« 14 » In the same way, my feeling of the pen in my hand is not in the same domain as the neural activities of my brain that generate this feeling. The first domain (my hand feeling my pen) needs the second (my nervous system with neurons and electrochemical reactions), as is the case for the relations between different abstract levels, previously addressed. The point is that if we try to simulate what happens in a phenomenal domain, we certainly have to address how to simulate the interfaces that exist between two different phenomenal domains in order to make their causal links possible. I wonder whether the use of neural networks can cover this point? (Q3)

« 15 » As my final point, let me emphasize that I consider a more challenging program that of reflexivity. The capacity of an agent to observe itself internally, i.e., in a phenomenal domain, to move not only from a low abstract level to a high abstract one but also vice versa, from a high abstract level to a low abstract one. In my opinion, it is the more challenging issue in the process of obtaining a genuine constructivist agent, i.e., a genuinely autonomous system that governs itself without the need to predefine its goals by external causes. It is able to determine its behavior not only based on the behaviors in lower phenomenal domains that help to maintain its existence but also based on the behaviors in the phenomenal domain that reflect on its existence as a system. For that, it needs to know that it exists.

Pierre De Loor is Professor in Computer Sciences at Brest National Engineering School (ENIB). He develops his research at the Lab-STICC (French CNRS Lab) and is the Co-Head of the team "Human System Interaction" (IHSEV). His research focuses on the links between artificial intelligence, interactive systems and cognitive sciences. More precisely, he works on the self-adaptation of virtual environments to users, non-verbal interactions between humans and virtual agents, modeling and simulation of autonomy, as well as all the properties linked with the enactive field. Web page: http://www.enib.fr/~deloor.

Received: 7 February 2018
Accepted: 25 February 2018

Mastering the Laws of Feedback Contingencies Is Essential to Constructivist Artificial Agents

Olivier L. Georgeon
Université de Lyon, France
ogeorgeon/at/univ-catholyon.fr

Mathieu Guillermin
LBG UMRS 449, Université Catholique de Lyon, France
meguillermin/at/univ-catholyon.fr

> Upshot • We support Füllsack's claim that contingency is essential in the conception of artificial constructivist agents. Linking this claim to O'Regan and Noë's theory of sensorimotor contingencies, we argue that artificial constructivist agents should master the laws of *feedback contingencies*. In particular, artificial constructivist agents should process input data as feedback from their actions rather than as percepts representing the environment.

« 1 » Manfred Füllsack argues that *multiple contingency* counts among the essentials for conceiving artificial constructivist agents. Yet, is it not the purpose of all artificially intelligent (AI) agents to adapt to contingencies during their interaction with their environment? Indeed, all AI agents should be able to cope with some uncertainty, that is, to handle the contingent nature of the data that they receive from their environment. Many AI agents may also develop contingent behaviors depending on their contingent experience interacting with multiple other agents. Yet, we would not argue that all AI agents are constructivist. So, what kind of contingency exactly makes an artificial agent constructivist, and why?

« 2 » We argue that there is something special in how Füllsack's agents handle contingencies that makes them constructivist, namely their ability to handle a particular kind of contingency we call *feedback contingency*. Indeed, as implemented by Füllsack, the agent's input data (the *payoffs*, §27) come as feedback resulting from the agent's actions (*cooperate* or *defect*, §28). The agent develops contingent behaviors (i.e., behaviors that could have developed differently) depending on contingent feedback received at an early stage.

« 3 » Feedback contingencies contrast with the kind of contingency that most AI agents deal with. For example, Peter Russell and Stuart Norvig consider the agent's input data as *percepts* rather than feedback: "the problem of AI is to build agents that receive percepts from the environment and perform actions" (Russell & Norvig 2003: iv). By considering input data as percepts, Russell and Norvig designed algorithms that rest upon the assumption that the agent's input data are a direct function of the state of the environment. This assumption, however, is not satisfied when the agent's input data are feedback from actions. For example, in Füllsack's model, the payoff not only depends on the state of the environment (here, the other agents) but also on the agent's own previous actions. Russell and Norvig did not demonstrate that their agents could handle feedback data appropriately. We expect that such demonstration would require a significant modification of their agents and a redesign of their experiments.

« 4 » In computer science, the paternity of the term *feedback* is usually attributed to Norbert Wiener's cybernetics (1948). Figure 1 illustrates Wiener's conceptualization of a control system using a perturbation signal provided as feedback from the system's output.

« 5 » In his paper "Why I Consider Myself a Cybernetician," Ernst von Glasersfeld (1992) explained the link between constructivism and cybernetics by comparing Wiener's *feedback systems of control* with an organism that manages to "get by the constraints set by its environment" (Glasersfeld 1992: 22), thereby demonstrating "adaptedness" and "equilibration." Von Glasersfeld's arguments also apply when considering Füllsack's agents to be constructivist: the agent's input data (the feedback) "has an adaptive function, not a representational one" (ibid: 23). This differentiates Füllsack's agents from Russell and Norvig's whose input data (percepts considered as a function of the state of the environment), although partial and noisy, do have a representational function.

« 6 » In short, Füllsack seems to commit to an ‹effector→sensor› paradigm whereas most AI scientists commit to the ‹sensor→effector› paradigm. While effector data and sensor data are sent and received alternately in an infinite loop, these paradigms differ in the way sensor data is generated by the experimental settings and processed by the agent's algorithm. Figure 4 of Füllsack's article, as often in the AI literature, does not highlight which paradigm is implemented. Our Figure 2 suggests a manner of highlighting it.

« 7 » Füllsack's claim that contingency is essential to artificial constructivist agents is in line with Kevin O'Regan and Alva Noë's (2001) idea that mastering sensorimotor contingencies is important to cognition. O'Regan and Noë, however, did not specify in what sense they used the term "sensorimotor." Their statement "[…] the structures of the rules governing the sensory changes produced by various motor actions, that is, what we call the sensorimotor contingencies […]" (O'Regan & Noë 2001: 941) evokes the ‹effector→sensor› paradigm. On page 1013, however, the sentence "We agree with the functionalists, though, that it is the input/output relations that matter" evokes the ‹sensor→effector› paradigm. As an experimental psychologist and a philosopher of perception (as they present themselves on page 939), they may have conceived "sensorimotor" in its Piagetian sense that merges sensation and motion in an indivisible chunk of phenomenological experience. Yet, for a computer scientist, the distinction is important since we do not see how we could merge artificial agents' sensors and effectors. Some of O'Regan and Noë's commentators pointed out the need for clarification. In his commentary entitled "On the Distinction between 'Sensorimotor' and 'Motorsensory' Contingencies," Donald Laming noted that "the distinction is essential" because "these two sets of contingencies are entirely disjoint" (Laming 2001: 992).

« 8 » When it comes to designing constructivist artificial argents, we suggest interpreting O'Regan and Noë's theory according to the ‹effector→sensor› paradigm. To avoid ambiguity, we propose the expression *feedback contingencies*. Füllsack's model appears to us to be a valuable example of artificial agents that learn to master simple laws of feedback contingencies. On the basis of the arguments developed above, along with O'Regan and Noë's, we believe that designing artificial agents capable of learning to master complex laws of feedback contingencies constitutes an interesting challenge.

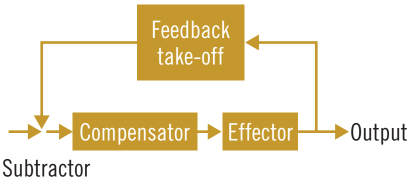

Figure 1 • Feedback system of control (adapted from Wiener 1948: 112).

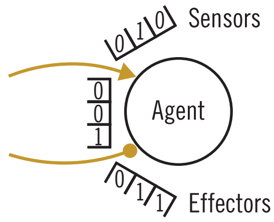

Figure 2 • Modification of target article's Figure 4. The bullet represents the premise of the interaction, i.e., the investment; the arrow represents the outcome, i.e,: the payoff..

Olivier L. Georgeon is an associate professor at the Catholic University of Lyon and a researcher in computer science at the LIRIS Laboratory (CNRS - UMR 5205). He has an interdisciplinary background with a master of engineering in computer science from the Ecole Centrale de Marseille (1988) and a PhD in cognitive psychology from the Université Lumière Lyon 2 (2008). His research focuses on the implementation of developmental learning mechanisms and intrinsic motivation in artificial agents.

Mathieu Guillermin is an associate professor at the Catholic University of Lyon. He holds a PhD in Physics (laser-matter interaction) and a PhD in Philosophy of Science (pragmatist epistemology for combining incommensurability and context-sensitivity with scientific realism). His research focuses upon philosophy and ethics of science and technology, especially regarding digital technologies.

Received: 26 February 2018
Accepted: 28 February 2018