plasticity, which, as stated by Beer, refers to the distance between the glider structure and its organization. In other words, the plasticity of the glider is constituted by the possible patterns of locations and states underpinning its distinctive organization.[1] The Game of Life model largely contributes to the illustration of Maturana and Varela's concepts, with images (see Figure 1 to 4 in Beer 2004) showing the glider structure, its organisation, its boundaries, and a series of destructive and non-destructive perturbations. The simplicity and astonishing clarity of the glider as a model of an autopoietic system makes me wonder what glider features, if any, could illustrate the concept of granularity as discussed by Füllsack. What does granularity ultimately add to the agency-constructive process? (Q1) In my view, I see granularity as another term to refer to the space of possible structures that underpin the glider's organisation.

« 7 » The simulation model in Di Paolo (2000) focuses on the issue of social coordination observed in a robotic scenario where two simulated robots are required to get and remain closer to each other. The robots are placed in a boundless arena at a certain distance from each other. Their task is to move towards and remain close to each other by exploiting acoustic signals. The robots can move forward and backwards by rotating their wheels, and they can emit and detect sound using two loudspeakers and a microphone mounted on their chassis. The robots' controller is a dynamic neural network synthesized using evolutionary computation techniques (Beer & Gallagher 1992).

« 8 » Di Paolo (2000) provides an extensive series of evaluation tests aiming to investigate the operational principles underlying the behavioural strategies used by a successful group of robots to perform the task. The author shows how the robots exploit their movements to avoid cancelling each other's sound signals, and also to distinguish between their own and the other's sound. The successful group also undergoes a series of tests in which one robot is tested with a beacon that emits pre-recorded signals. In one of these tests the beacon emits the "taped" signal of one robot recorded during a successful execution of the coordination task. The amazing result is that the robot required to coordinate its actions with respect to this taped-signal-emitting beacon fails to generate those virtuous actions that it generates when interacting with the original group mate. The comparative analysis of the behaviour of this robot, while interacting with its mate and with the beacon, shows that the failure to approach the beacon can be explained with reference to the lack of plasticity in the beacon's behaviour. That is, the beacon is unable to enter into a transient state of mutual triggering of change of structure, which is referred to as structural congruence. In the attempt to provide an operational analysis of the mechanisms underpinning the robots' actions in this social coordination task, Di Paolo emphasises the existence of operational links between social behaviour (the robots' emission of acoustic signals) and what an observer would see as purely individual (the robots' movements). For example, the analysis shows that signalling behaviour is not only used to interact with the other robot, but is also "integrated into the movement of the producing agent" (Di Paolo 2000: 1).

« 9 » The results presented in Di Paolo (2000) give me the opportunity to comment on the Füllsack claim according to which disembodied computational tools like artificial neural networks "[substantiate] the claim of constructivism" (§50) by generating a more plastic and more granular system. I certainly agree with the author that a neuro-controller significantly enlarges the distance between the agent's structure and its organisation, providing an agent with the means to retain its organisation in response to a larger set of perturbations. However, there are plenty of examples in the artificial life and adaptive behaviour literature showing how the embodiment and situatedness of a system are often sufficient to underpin complex behaviour in the absence of those neural mechanisms that a priori would be considered necessary (see Pfeifer & Bongard 2007). In view of the results of the last 20 years of research in artificial life and adaptive behaviour, I would consider it fairer to claim that artificial neural networks are one possible means among others to meet the plasticity and the granularity criteria.

1| For an illustration of a glider see https://en.wikipedia.org/wiki/Conway%27s_Game_of_Life

Elio Tuci is a Senior Lecturer in the Department of Computer Science, Middlesex University, UK. In 2003, he received a PhD in computer science and artificial intelligence from the University of Sussex, UK. In 1996, he graduated in experimental psychology from the University of Rome "Sapienza," Italy. His research interests fall into the interdisciplinary domain of bio-inspired robotics and computational intelligence. His research activity consists in generating computational instances of relatively complex behavioural phenomena, possibly using the minimal set of mechanisms sufficient to underpin the target behaviour. Elio Tuci has co-authored about 70 peer-reviewed scientific publications and 4 book chapters.

Received: 13 February 2018
Accepted: 19 February 2018

New Concepts or Just Re-Wording?

Filipo Studzinski Perotto
Normandy University, France
filipo.perotto/at/litislab.fr

> Upshot • In this commentary, I question the necessity of a recursive definition of agency, and I also examine the novelty of some proposed concepts such as plasticity, granularity, distinction, and irritation.

« 1 » In his target article, Manfred Füllsack focuses on three theoretical aspects necessary for developing artificial constructivist agents: *plasticity*, *granularity*, and *multiple contingency*. In addition, he claims that these aspects (although not explicitly) are already present in current AI technology. In my commentary, I present arguments challenging these claims.

« 2 » Füllsack proposes a recursive definition of agency: "agency can only emerge against a backdrop of a multitude of agents" (§1). In my view, this concept could be better defined by considering the notion of "proto-agents" as entities endowed with the capacity of becoming "whole-agents." After a period of interaction with other agents, such "proto-agents" can evolve into "whole-agents." In this way, an artificial constructivist agent is assumed to emerge from the interaction with other entities of the *same* kind (§7). This

makes me wonder: Why can a constructivist agent not develop itself through interaction with *other* kinds of (complex or simple) agents, objects or elements of its environment? Why is such a restriction necessary? (Q1) To illustrate such a recursive definition of agency, in reference to self-constructing systems (§2), Füllsack uses the mythological metaphor of two ouroboroses eating each other by their tails. I cannot see how this illustration helps to consolidate the necessity of such recursion: the same picture could be used to illustrate the exact opposite, i.e., a self-*destructive* system.

« 3 » Füllsack is subscribing to *epistemological solipsism* when he proposes the new notion of "irritations" (§9) in order to define something that impacts the "on-board means" of the agent, but which is not under the control of the agent. This new term is supposed to establish a distinction against terms such as "inputs," "observations" or "perceptions," which could be interpreted as representing some kind of direct "revelation" of the state of the environment. The same motivation is used to establish the distinction between actions (oriented to on-board means) and reactions (oriented to some "external world"). In philosophical terms, epistemological solipsism cannot be refuted. The assumption that nothing outside the own mind can be proved to exist is a reasonable point about the limits of our knowledge (Snider 2017: 73). Questions such as "What is the real world outside the subjective sensorimotor perceptions?" and even "Is there something beyond the perception to be called an external world?" are simply unresolvable. However, when referring to computational agency, a distinction between agent and environment must be made because (a) we program both the agent and (often) the environment, and (b) if such a distinction does not exist (even if it is complex, volatile, or dynamic), the use of these concepts is pointless. From this computational point of view, it is effectively possible to implement agents that do not have any kind of direct access to the state of the environment.

« 4 » Besides being distinguishable, agent and environment must be (by definition) systems that influence each other in some manner (coupled systems). Such a notion had been philosophically established by Ross Ashby (1952), Humberto Maturana and Francisco Varela (1980) and Mark Bickhard (2009), but also computationally by Michael Anderson (2003), Randall Beer (1995) and me (Perotto 2013). In all these references, an agent is characterized as a self-maintaining (cybernetic or autopoietic) system. So, is Füllsack not simply rewording well-known basic terms? How is his proposed concept of "*maintainable plasticity*" (§10) related to the concept of "autopoiesis" (Maturana & Varela 1980)? Does the new term introduce something new that facilitates the understanding of the learning process inside a constructivist agent? (Q2)

« 5 » Concerning "inputs" (§9), in computational models they are, in general, implemented as a filter, i.e., as a function indicating how the state of the environment (or its transformations) influences the perception that the agent has of it. There is always an "external" factor defining the agent's perception (i.e., a factor that is not directly controlled by the agent, but which is determined in some manner by the environment), even if associated to other internal (native or constructed) factors. Hence, the use of the concept of "agent" can never be dissociated from the concept of "environment," from a constructivist point of view or not, considering the existence of other complex agents in such an environment or not.

« 6 » Being aware (or not) of the existence of an "outside" (§9) is not related to being constructivist (or not). From the sceptical perspective of epistemic solipsism, an agent that evolves in order to become self-aware or conscious of the outside is just creating an illusion. Perhaps this is a useful, coherent, consistent and meaningful illusion, but it remains an illusion because the existence of an external mind-independent world can never be confirmed. The constructivist aspect of the learning process relies on the fact that the knowledge constructed by the agent does not represent any correspondent structure of some "outside," but is anchored in the sensorimotor experience, which is always subjective. For this reason, Füllsack's metaphor of the *minimal artificial constructivist agent* as a "formable something into which patterns can be iteratively inscribed" (§11) seems inappropriate as a constructivist intuition. Such an inscription represents the conformation of the "medium" to an external force rather than the composition of such a force with some already existing organization (i.e., rather than a process of construction based on the progressive equilibrium between assimilation and accommodation). The same kind of criticism can be made concerning the notion of "memorized experiences" (§8). They do not suppose any kind of generalization or abstraction, which is a necessary feature to allow an artificial agent to learn a model of its interactions (Perotto 2013).

« 7 » Yet another instance of rewording of well-known concepts from the literature concerns the two elementary aspects, which are defined by Füllsack as part of the minimal constructivist agent, namely *plasticity* and *granularity*. They can be directly related to the classical constructivist concepts of *accommodation* and *assimilation*, respectively. Plasticity describes situations in which perceptual experience does not fit in the subjective structures, triggering the transformation or the construction of new and possibly more complex structures in order to organize this experience, which in the constructivist literature is known as accommodation. Granularity is the opposite force, where the experience is integrated into existing structures, organized in different levels. This is known as assimilation. Why are these new terms needed? Could the classic concepts not be applied to Füllsack's model instead? (Q3)

« 8 » The use of a variant of the back-propagation algorithm in simple simulation scenarios is supposed to illustrate the proposed constructivist minimal necessary aspects, theoretically proposed in the article. However, the link between the theoretical concepts proposed in the first part of the article and the computational simulation presented at the end would deserve more clarity. In this regard, the most contradictory point seems to be the claim that initial conditions are irrelevant and lead to stable "eigenforms" (§4), related to the notion of multiple contingency, while the performance of the agents in the given experimental simulations (§31) seems to be completely determined by the initial conditions.

« 9 » Finally, defending the view that the working principles of artificial neural networks (ANNs) substantiate the claim of constructivism (§50) strikes me as a con-

troversial yet insufficiently demonstrated conclusion. An ANN is a multi-parametric model trained with a set of examples. Through learning, its parameters are adjusted to approximate the function or the classifier that best matches the examples. Given the non-symbolic character of ANNs, even if an ANN has layers, they cannot be said to represent levels of abstraction. From my perspective, saying that an ANN implements the fundamental aspects of constructivism in artificial intelligence is like saying that a neurobiological model is sufficient for explaining the constructivist psychology.

Filipo Studzinski Perotto is a member of the MIND team, affiliated with the LITIS lab at Normandy University, in France. He obtained a PhD in computer sciences from the University of Porto Alegre (UFRGS, Brazil) and the University of Toulouse (INPT, France) in 2010, addressing the subject of constructivist anticipatory learning algorithms. His main research interests are related to artificial intelligence, machine learning and sequential decision-making problems.

Received: 15 February 2018
Accepted: 26 February 2018

The Folly of a Normative Account of "Constructivist Agents"

Bruce Edmonds
Manchester Metropolitan University
Business School, UK
bruce/at/edmonds.name

> Upshot • The target article is criticised on four counts. It fails to make clear what is meant by the phrase "constructivist agent," and whether the author is trying to *define* "constructivist agent" or arguing what the minimal criteria for this are (there are problems with either). It does not make clear whether weak or strong emergence is intended (there are problems with either). The arguments for a minimal level of granularity are incoherent. To summarise, the whole project has a normative flavour that seems odd given the constructivist stance it intends to argue from.

What is a "constructivist agent"?

« 1 » Manfred Füllsack's target article presents an account of "constructivist agents," to establish a

> "minimum set for conceiving an artificial constructivist agent: these are (a) the multitude criterion as mentioned above, which, as I see it, necessitates (b) a minimum plasticity of the agents, and (c) a minimum granularity of the stuff that agents consist of." (§3)

« 2 » However, what is not clear is in which direction Füllsack is arguing. There seem to be two possibilities for this, namely that the article is an attempt…

- to define what "constructivist agents" *are* (i.e., they are those agents that satisfy the three criteria above);
- to show that those three criteria are the *necessary* conditions for a "constructivist agent."

So, which of the above possibilities did the author intend? (Q1)

« 3 » If the intention is the former, then the question naturally arises as to why anyone should accept Füllsack's definition, since he does not establish any criteria for the "goodness" or purpose of the definition. For example, he does not argue that this will result in agents that are more useful, nor does he argue that such agents would be better representations of anything observed. In this case, the definition floats in a kind of limbo without apparent purpose or a particular role.

« 4 » If the intention is to argue for the necessity of the three conditions, then we also need to know how to judge whether an agent is "constructivist" or not in order to be able to tell whether these conditions are, in fact, necessary. The text gives only one hint at this, namely that their agency is not programmed in, but *emerges* during a complex process of interaction with other agents. However, this merely shifts the argument to the nature of emergence and agency. Unfortunately, neither of these concepts is defined in the target article. How do we know whether an entity has agency? How do we know if something is built-in or emerges?

« 5 » Constructivism is a philosophical position. The whole idea of labelling an entity that we either build or observe as "constructivist" does not make obvious sense, because constructivism is a way of thinking about the nature of knowledge and not a property of any *thing*. I *guess* Füllsack simply means that the agency was not designed but emerged, but (a) this has not been clearly stated and (b) very simple things can emerge from such processes, many of which could not be said to have any agency at all (e.g., mathematical solutions found via a co-evolutionary genetic programming algorithm). What kind of emergence does the author intend? (Q2) This suggests the question concerning what precisely Füllsack means by a "constructivist agent" (assuming that it is not defined by his three criteria – is identical with a "minimal artificial constructivist agent" – which would make the argument circular)?

Misunderstanding emergence

« 6 » There are a number of kinds of emergence discussed in the literature, of which the most prominent are:

- *Strong emergence*, where the emergent properties are not reducible to those of the component parts (Laughlin 2005), even in principle.
- *Weak emergence*, where new properties do arise but can be determined by simulating the system (though where no short-cuts to doing the simulation are sufficient) (Bedeau 1997).

« 7 » It would seem that Füllsack is not talking about strong emergence, since he gives an example of a simulation that is, in principle, reducible to many small knowable computational steps (if the random seed of the simulation is known). If he is talking about weak emergence, then the macro-level framework for describing the new properties is very important (this being different from the framework or language that describes the micro-level properties). However, this macro-level framework is not specified – indeed it appears from the text that the author has not considered this. Given the lack of communicated purpose or role for this definition one cannot but suspect that Füllsack has effectively adopted an essentialist view of emergence, due to his wish to specify what a "constructivist agent" *should* be (see section below).