«11» I see the previous examples as the first trials to explore the knowledge-acquiring and processing abilities of families of simple embodied computational devices in a bottom-up manner, starting with the simplest gadgets and the simplest cognitive tasks. To me, it appears that the first thing to do is to specify the set of simple cognitive tasks and scenarios out of which solutions of more complex cognitive tasks can be composed, possibly under more complex scenarios. Among such tasks, I would include mirroring of observed signals, learning of sequences of signals, reinforced learning, imitation learning, communication via simple actions, etc. The literature is full of similar approaches but I am not aware of work following the evolutionary sequence of cognitive tasks in parallel with the evolutionary sequence of increasingly more complex (embodied) computational devices. This could be a rewarding task for constructive cognitivism.

Conclusion

«12» Although I welcome the idea to build the constructive foundations of cognitivism, whose very first and timid germs I see in the commented paper, and although I have indicated that there is a tremendous amount of results and insights to begin with, it is not going to be an easy enterprise. The doors mentioned in the title of this commentary are wide open, but an immense amount of stuff has to be categorized, ordered, critically evaluated and sorted so that only the minimal body necessary to build, maintain and develop a coherent constructive theory of cognitivism remains. If properly done, cognitivism and its applications will benefit much from such an endeavor.

Jiří Wiedermann is a computer scientist specialized in computability, complexity theory and non-standard models of computations, with interests in cognitive computations and epistemology. He is a member of the Czech Learned Society and of Academia Europaea. From 2000 to 2012, he served as the director of the Institute of Computer Science of the Academy of Sciences of the Czech Republic in Prague.

Received: 8 October 2013
Accepted: 22 October 2013

Single Agents Can Be Constructivist too

Olivier L. Georgeon
Université de Lyon, France
olivier.georgeon/at/liris.cnrs.fr

Salima Hassas
Université de Lyon, France
salima.hassas/at/univ-lyon1.fr

> Upshot • We support Roesch and his co-authors' theoretical stance on constructivist artificial agents, and wish to enrich their "exploration of the functional properties of interaction" with complementary results. By revisiting their experiments with an agent that we developed previously, we explore two issues that they deliberately left aside: autonomous intentionality and dynamic reutilization of knowledge by the agent. Our results reveal an alternative pathway to constructivism that addresses the central question of intentionality in a single agent from the very beginning of its design, suggesting that the property of distributed processing proposed by Roesch et al. is not essential to constructivism.

«1» In their paper, Etienne Roesch and his coauthors formulate a constructivist approach to artificial learning in which

> "knowledge of the world, for an individual, is created from the interaction with the environment, rather than existing in an ontic reality, supposedly pre-existing or available to registration from the physical world." (§1)

They propose three models to illustrate this idea, in which a swarm of agents performs different tasks in an environment made of a string of digits. We fully agree with this theoretical stance but we feel that these models do not illustrate it as well as possible. In particular, one might argue that the swarm's knowledge does, in fact, "exist in an ontic reality" since the agents directly "perceive" the digits and apply predefined rules to process the digits for the purpose intended by the designer.

«2» Here, we present an alternative model that does not make the knowledge of the environment directly available to registration by the agent. The environment is the string of digits presented in §23, and the agent was designed to produce similar results as proposed in §25: sorting the string. Yet the agent's observations are reduced to a single bit whose significance depends on the dynamics of the agent's interactions rather than directly reflecting the state of the environment. The agent remains "unaware" that it "exists" at a particular position in a string of digits, and its own goal is not to sort this string. For the agent, the construction of knowledge consists of learning to organize its behavior to fulfill a form of intentionality defined independently of the environment.

Implementation

«3» *Initialization*: The environment is a string of 10 digits $E_0 = [6,3,5,4,7,3,5,3,9,5]$ plus an integer p in the interval $[0,9]$ that represents the agent's position. $E_t[p_t]$ denotes the digit at the agent's position at time t. At time 0, $p_0 = 0$, thus the current digit $E_0[p_0] = 6$.

«4» *Behaviors*: At time t, the agent chooses an action from amongst the set of three possible actions $A = \{step, feel, swap\}$, and then receives a binary observation from amongst the set of two possible observations $O = \{true, false\}$. The set $A \times O$ thus contains 6 possible *interactions*. The agent initially ignores the meaning of actions and observations, i.e., it implements no rule to process them specifically. However, each interaction has a predefined *valence* that plays a role in defining the agent's intentionality, as explained below. Unbeknownst to the agent, *step* consists of stepping to the next digit. If this action takes the agent to a greater or equal digit then it produces the observation *true* and has a positive valence, otherwise, it produces the observation *false* and has a strongly negative valence. *Feel* consists of testing whether the next digit is greater than or equal to the current one; if yes, it produces *true*, otherwise *false*. *Feel* interactions have a mildly negative valence. *Swap* consists of trying to swap the current digit with the next. It succeeds only if the current digit is greater than the next, producing the observation *true* and a positive valence; otherwise it does nothing and produces the observation *false* and a strongly negative valence. When the agent is at position 9, *step* returns the agent to position 0, *swap* does nothing, and the three actions produce the observation *false*. Table 1 summarizes the implementation of these possibilities of interaction.

« 5 » *Agent*: We used an agent presented previously (Georgeon & Ritter 2012), which was programmed to exhibit two forms of intentionality: the tendency to select sequences of actions that produce well-predicted observations, and the tendency to enact positive interactions while avoiding strongly negative interactions. The former type of intentionality relates to Steels's (2004) autotelic principle (the enjoyment of being *in control* of one's activity), and was implemented as a tendency to record, hierarchically organize, and appropriately re-enact sequences of interactions that capture regularities in the coupling between the agent and the environment. The latter is called interactional motivation (Georgeon, Marshall & Gay 2012), and was implemented through preferentially engaging in sequences of interactions that have the highest total valence.

Results

« 6 » Table 2 reports selected strips of behaviors, with the current digit marked in bold. The agent started by randomly picking the *step* action at times 1 and 2. Over time, the agent organized its behavior as if it had discovered that the *feel* action could be used to test the next digit. If this action resulted in the *feel_up* interaction, then the *step_up* interaction could subsequently be enacted, otherwise, the *swap* – *step_up* sequence could subsequently be enacted. This dynamics resulted in the behavior of "carrying digits to the right." This behavior is illustrated in Table 2 from time 106 to 113: the agent "carried" the "5" digit from position 2 to 4, by repeating the *feel_down* – *swap* – *step_up* sequence until the "5" digit got "blocked" by a greater or equal digit (another "5" digit at position 5). This behavior resulted in the string being entirely sorted at time 130.

« 7 » Figure 1 reports the agent's behavior until time 200 in terms of what matters to the agent: the enacted interactions, their valence, and the level of control that the agent has over its activity manifested by the length of the sequences intentionally enacted.

« 8 » In summary, this experiment helps clarify the distinction between the designer's goal (sorting the string, illustrated in Table 2) and the agent's intentionality (being in control and enacting interactions that have positive valence, illustrated in Figure 1). While the agent remained unaware of the underlying structure of the environment, it learned to master sensorimotor contingencies as if it enjoyed being able to predict its activity and to "step up," and disliked "stepping down" and failing to swap digits. The agent learned to use the *feel* action – in spite of its negative valence and the ignorance of its meaning – as an active perception of the environment to inform subsequent behaviors. This activity illustrates the property pointed out by Roesch et al. that "perception is an integral part of the process from which knowledge of the world arises" (§7) and that "Exploration of the environment provides the organism with the ability to sense and become attuned to the laws governing change" (§8). We believe that these properties, associated with the capacity of the agent to engage in incremental learning, qualify the agent as a candidate to illustrate key aspects of constructivism.

« 9 » Roesch et al. conclude by listing the properties of interactions that they judge to be paramount to the constructivist approach: "partial information, exploration, distributed processing and aggregation of information, emergence of knowledge and directedness towards relevant information" (§40). Our results support all of these except the *distributed processing* property (insofar as it applies to a swarm of agents), and suggest the additional property of *intrinsic intentionality*.

Action	Condition	Effect	Observation	Interaction	Valence
step	$p_t < 9$ and $E_t[p_t] \le E_t[p_t+1]$	$p_{t+1} = p_t + 1$	true	step_up	4
	$p_t < 9$ and $E_t[p_t] > E_t[p_t+1]$	$p_{t+1} = p_t + 1$	false	step_down	–10
	$p_t = 9$	$p_{t+1} = 0$	false	step_down	–10
feel	$p_t < 9$ and $E_t[p_t] \le E_t[p_t+1]$	–	true	feel_up	–4
	$p_t < 9$ and $E_t[p_t] > E_t[p_t+1]$	–	false	feel_down	–4
	$p_t = 9$	–	false	feel_down	–4
swap	$p_t < 9$ and $E_t[p_t] \le E_t[p_t+1]$	–	false	not_swap	–10
	$p_t < 9$ and $E_t[p_t] > E_t[p_t+1]$	$E_{t+1}[p_t+1] = E_t[p_t]$ $E_{t+1}[p_t] = E_t[p_t+1]$	true	swap	4
	$p_t = 9$	–	false	not_swap	–10

Table 1: Possibilities of interaction available to the agent.

Time	Interaction	Environment
0	–	**6** 3 5 4 7 3 5 3 9 5
1	step_down	6 **3** 5 4 7 3 5 3 9 5
2	step_up	6 3 **5** 4 7 3 5 3 9 5
…		
106	feel_down	3 4 **5** 3 3 5 5 6 7 9
107	swap	3 4 **3** 5 3 5 5 6 7 9
108	step_up	3 4 3 **5** 3 5 5 6 7 9
109	feel_down	3 4 3 **5** 3 5 5 6 7 9
110	swap	3 4 3 **3** 5 5 5 6 7 9
111	step_up	3 4 3 3 **5** 5 5 6 7 9
112	feel_up	3 4 3 3 **5** 5 5 6 7 9
113	step_up	3 4 3 3 5 **5** 5 6 7 9
…		
130	swap	3 3 **3** 4 5 5 5 6 7 9

Table 2: Behavior strips.

Olivier L. Georgeon is currently an associate researcher at the LIRIS Lab with a fellowship from the French government (ANR-RPDOC program). He received a master's in computer engineering from Ecole Centrale de Marseille in 1988, and a PhD in psychology from the Université de Lyon in 2008.

Salima Hassas leads a research activity in the area of multi-agent systems with self-*(self-adaptation, self-organization, etc.) abilities. Her research is related to complex systems, bio-inspired computing, constructivist approaches to cognition and enaction. She is an associate editor of ACM TAAS, and serves on scientific committees for major conferences in her research domain.

Received: 11 October 2013
Accepted: 18 October 2013

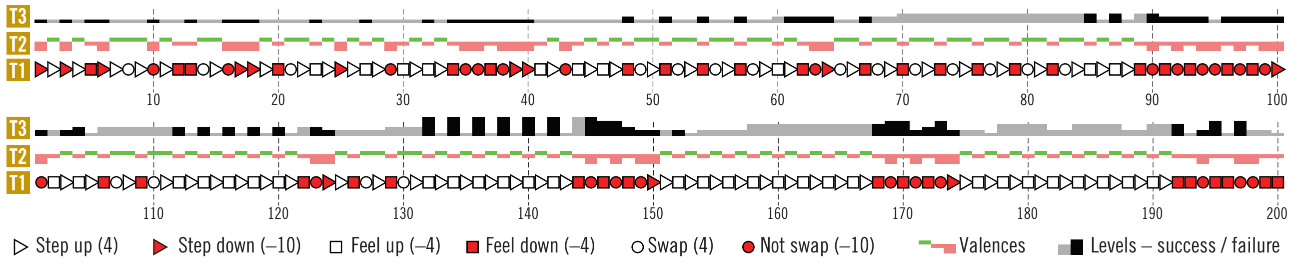

Figure 1: Analysis of the first 200 interactions enacted by the agent. Tape T1: the enacted interactions (the shape represents the action and the color the resulting observation). Tape T2: the valence of the enacted interactions displayed as a bar graph (green when positive, red when negative). Tape T3: The length of the sequences intentionally enacted, displayed as a bar graph. Higher levels of gray indicate better control over the activity; black segments indicate that an intended sequence was interrupted due to the failure to predict correctly the resulting observation. This trace shows that the behavior was unorganized approximately until time 40 (no regularities in the symbols in T1 and the presence of step_down and not_swap interactions that have strong negative valence represented by high red bars in T2). The agent intentionally enacted the second order sequence swap – step_up for the first time during time 68–69 (second level in T3), then the third-order sequence feel_down – swap – step_up during time 70–72 (third level in T3), repeating this sequence until time 85. After time 130, the digits were entirely sorted, and the agent engaged in repeating the sequence feel_up – step_up, except when reaching the end of the string, in which case it continued experimenting with other behaviors (episodes 144–150, 168–174, and 192–200). After time 310 (not shown), the agent resigned itself to merely enacting the step_down interaction when reaching the end of the string, acknowledging that it had no better possibilities.

Authors' Response: Learning, Anticipation and the Brain

Etienne B. Roesch et al.

> **Upshot** • Albeit mostly supportive of our work, the commentaries we received highlighted a few points that deserve additional explanation, with regard to the notion of learning in our model, the relationship between our model and the brain, as well as the notion of anticipation. This open discussion emphasizes the need for toy computer models, to fuel theoretical discussion and prevent business-as-usual from getting in the way of new ideas.

« 1 » Computer simulations hold great potential for charting unexplored territories. This applies especially to constructivism, which now faces the challenge of explaining the mechanisms that support knowledge acquisition. However, this potential comes at a price, i.e., that of forcing oneself to remain humble as to what can reasonably be expected of the tool. Constructivism is an "all-embracing theory," as pointed out by **Manfred Füllsack** (§2), that demands a delineation of scope that goes beyond the mere description of the organism's internal states. In this target article, we aimed to explore aspects of the interaction between the organism and the environment that support what might be called a constructivist process; aspects such as partial information, exploration, distributed processing and aggregation of information, emergence of knowledge and directedness towards relevant information.

« 2 » We do not mystify interaction – and our single-minded focus can indeed be misinterpreted (**Peter Cariani** §8) – for that would be casting a blind eye over an impressive body of data; nor do we subscribe to the mystification of internal mechanisms often found in mainstream cognitive science. In placing such a focus on aspects of interaction, we wish to highlight the contribution of these aspects to causal, powerful and goal-directed behavior, alongside more internal mechanisms, which we voluntarily did not include in our models. These passive, dynamic effects, also called "nontrivial causal spread" (Wheeler 2005: 250), contribute to cognition and our interaction with the world in a significant way (Pfeiffer & Bongard 2007), and are all too often ignored. If there is one family of theories that ought to bring credit to the interaction between the organism and the environment, it is constructivism.

« 3 » To put it differently, we concur with the notion that constructivism does not imply a particular theory of mind and brain beyond two simple axioms:

> "(1) knowledge arises from interaction with an environment, through perceiving and acting, and (2) there are internal, adaptive, self-modification processes that realize new epistemic functions that in turn modify behavior" (**Cariani** §9).

But, in the work presented, we simply wished to ascertain criteria of sufficiency and necessity for these axioms. Arguably, in the limited context we chose, our results seem to imply that the first axiom is sufficient to achieve our set goals.

« 4 » The commentaries on our target article emphasise the need for such toy models to fuel theoretical discussion, and