Challenges for interactivist-constructivist robotics

Răzvan V. Florian

Center for Cognitive and Neural Studies (Coneural),
Cluj-Napoca, Romania

Babeş-Bolyai University,
Institute for Interdisciplinary Experimental Research,
Cluj-Napoca, Romania

Email: florian@coneural.org

Presented at the Interactivist Summer Institute 2007, Paris, France

The interactivist-constructivist (IC) approach to cognitive systems (Bickhard, 1993; Bickhard and Campbell, 1996; Indurkhya, 1992; Christensen and Hooker, 2000) offers a sound framework for understanding cognition and representation and for designing genuinely intelligent artificial systems. Motivated either by theoretical considerations, by the problems of classical artificial intelligence and robotics or by biological inspiration, there is now a substantial body of work that is guided by principles similar to the IC ones, directed towards the development of intelligent embodied agents that learn by interacting with the environment. This body of work, mostly accumulated during the last decade, is probably best represented by the biannual International Conferences on the Simulation of Adaptive Behavior (e.g., Meyer and Wilson, 1991; Nolfi et al., 2006). However, we still lack artificial systems capable of representing the world, in the IC sense, or artificial systems that feature genuine, adaptive intelligence. Why it is so? The purpose of this presentation is to pinpoint the factors that still prevent us to build such systems and to give a speculative suggestion on how we can overcome the present impasse.

At least part of the robotics community agrees that the best way to get intelligent robots is not to preprogram them, but to let them learn online, through interaction with the environment, under the direction of a value system. Training may be performed by reinforcement learning, imitation or guidance. A control system that could support this kind of learning should be a collection of parallel, heterogenous, loosely coupled processes, capable of self-organization, such as a neural network (Pfeifer and Scheier,

1999; Weng et al., 2001; Florian, 2003). The key component that is lacking is a particular learning mechanism that would allow the components of the robotic controller to learn constructively while they direct the robot's action in accordance to its value system - a kind of learning that can be identified to the microgenesis of Bickhard and Campbell (1996).

This learning mechanism is elusive because it should include several components that traditionally have been studied in isolation, and generally under different contexts: unsupervised learning, reinforcement learning, memory, and dynamic control under uncertainty. Unsupervised learning is needed because the cognitive agent should establish topologies for structuring the sensorimotor flow (Bickhard and Campbell, 1996). However, current unsupervised learning is concerned with fitting data to simple, predefined models, generally linear ones (Ghahramani, 2004). This does not seem to be appropriate for the complex topologies required by realistic models of the world. Reinforcement learning is needed because the cognitive agent should be goal directed (in order to do something useful). But current reinforcement learning theories are almost exclusively developed in the framework of a model of the world (the Markov Decision Process, sometimes a Partially Observable one) (Sutton and Barto, 1998; Kaelbing et al., 1996) that is at odds with the dynamically constructed world model that is posited by the IC framework. Also, few implementations exist that integrate unsupervised learning with reinforcement learning. The memory of neural networks was extensively studied for storing static patterns, but few models are concerned with online storage and recall, and few models will not face catastrophic forgetting in realistic situations. A robotic controller can rightfully be considered a dynamical system and analyzed as such, but dynamical systems theory is hardly of use for complex systems with more than a few variables (Strogatz, 1994). Probabilistic learning has been argued to be optimal in face of uncertainty, but it can only operate under a predefined model of the world, or it can choose between such predefined models (Jaynes, 2003). How these models can be generated online by an agent with limited memory and computing resources, while respecting the principles of optimal probabilistic inference, remains an open problem. Moreover, an integration of the above-mentioned lines of research that are relevant to IC learning has not been performed to date.

An alternative way to designing artificial intelligent systems is to mimic nature, since humans and some animals are the only systems that do feature adaptive, robust intelligence. However, there might be dangers in following this source of inspiration too closely. Biological systems are not designed optimally: through the evolutionary process, solutions were "patched" onto previously working systems. Many vestigial neurological structures, interactions, and side effects may exist in animal brains. Developmental processes needed for the growth and specialization of cells, the supportive mechanisms needed for the nutrition of neurons, and other biological constraints may also lead to side effects in the biological neural architectures. The emulation of these side effects in artificial systems may be a distraction. Moreover, most experiments in neuroscience are still dominated by representational paradigms, in the tradition of classical cognitive science, and thus it is possible that we lack data relevant to a proper understanding of the brain's function.

However, there is a learning mechanism that has been experimentally found in the brain and that has been shown, analytically and in simulations, to have an important

range of properties relevant to IC learning. Spike-timing-dependent plasticity (STDP) is a type of neural plasticity where synaptic changes depend on the relative timing of pre- and postsynaptic action potentials (Markram et al., 1997; Bi and Poo, 1998; Dan and Poo, 2004). A typical example of STDP is given by the potentiation of a synapse when the postsynaptic spike follows the presynaptic spike within a time window of a few tens of milliseconds, and the depression of the synapse when the order of the spikes is reversed. This mechanism can regulate both the rate and the variability of postsynaptic firing, and may induce competition between afferent synapses (Kempter et al., 1999, 2001; Song et al., 2000). STDP can also lead to unsupervised learning and prediction of sequences (Roberts, 1999; Rao and Sejnowski, 2001). Plasticity rules similar to STDP were derived theoretically by optimizing the mutual information between the presynaptic input and the activity of the postsynaptic neuron (Toyoizumi et al., 2005; Bell and Parrara, 2005; Chechik, 2003), by minimizing the postsynaptic neuron's variability to a given input (Bohte and Mozer, 2005), by optimizing the likelihood of postsynaptic firing at one or several desired firing times (Pfister et al., 2006), or by self-repairing a classifier network (Hopfield and Brody, 2004). It was also shown that by clamping the postsynaptic neuron to a target signal, Hebbian STDP can lead, under certain conditions, to learning a particular spike pattern (Legenstein et al., 2005). We have also shown recently that the modulation of STDP with a global reward signal leads to reinforcement learning (Florian, 2007). STDP is thus a learning mechanism that integrates both unsupervised and reinforcement learning, while working in real time, possibly contributing to memory, perception and action simultaneously, as a function of the neural network's configuration. We thus speculate that STDP may be a mechanism that supports microgenesis in the brain.

References

Bell, A. J. and Parrara, L. C. (2005), Maximising sensitivity in a spiking network, *in* L. K. Saul, Y. Weiss and L. Bottou, eds, 'Advances in Neural Information Processing Systems', Vol. 17, MIT Press, Cambridge, MA.
http://books.nips.cc/papers/files/nips17/NIPS2004_0533.pdf

Bi, G.-Q. and Poo, M.-M. (1998), 'Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type', *Journal of Neuroscience* **18**(24), 10464–10472.
http://www.neurobio.pitt.edu/bilab/papers/Bi98.pdf

Bickhard, M. H. (1993), 'Representational content in humans and machines', *Journal of Experimental and Theoretical Artificial Intelligence* **5**, 285–333.
http://www.lehigh.edu/~mhb0/repconpage.html

Bickhard, M. H. and Campbell, R. L. (1996), 'Topologies of learning and development', *New Ideas in Psychology* **14**(2), 111–156.
http://www.lehigh.edu/~mhb0/topology.html

Bohte, S. M. and Mozer, C. (2005), Reducing spike train variability: A computational theory of spike-timing dependent plasticity, *in* L. K. Saul, Y. Weiss and L. Bot-

tou, eds, 'Advances in Neural Information Processing Systems', Vol. 17, MIT Press, Cambridge, MA.
http://homepages.cwi.nl/~sbohte/publication/stdppaper.pdf

Chechik, G. (2003), 'Spike time dependent plasticity and information maximization', *Neural Computation* **15**, 1481–1510.
http://www.cs.huji.ac.il/~ggal/ps_files/chechik_stdp.pdf

Christensen, W. D. and Hooker, C. A. (2000), 'An interactivist-constructivist approach to intelligence: Self-directed anticipative learning', *Philosophical Psychology* **13**(1), 5–45.
http://www.kli.ac.at/personal/christensen/SDAL_PhilPsych.pdf

Dan, Y. and Poo, M.-M. (2004), 'Spike timing-dependent plasticity of neural circuits', *Neuron* **44**, 23–30.

Florian, R. V. (2003), Autonomous artificial intelligent agents, Technical Report Coneural-03-01, Center for Cognitive and Neural Studies, Cluj, Romania.
http://www.coneural.org/reports/Coneural-03-01.pdf

Florian, R. V. (2007), 'Reinforcement learning through modulation of spike-timing dependent plasticity', *Neural Computation* **19**(5). In press.
http://coneural.org/florian/papers/06_modulated_STDP.pdf

Ghahramani, Z. (2004), Unsupervised learning, *in* O. Bousquet, G. Raetsch and U. von Luxburg, eds, 'Advanced Lectures on Machine Learning', Vol. 3176 of *Lecture Notes in Artificial Intelligence*, Springer Verlag, Heidelberg, Germany.
http://www.gatsby.ucl.ac.uk/~zoubin/course04/ul.pdf

Hopfield, J. J. and Brody, C. D. (2004), 'Learning rules and network repair in spike-timing-based computation networks', *Proceedings of the National Academy of Sciences* **101**(1), 337–342.

Indurkhya, B. (1992), *Metaphor and cognition: An interactionist approach*, Kluwer Academic Publishers, Dordrecht, the Nederlands.

Jaynes, E. T. (2003), *Probability Theory: The Logic of Science*, Cambridge University Press.

Kaelbing, L. P., Littman, M. L. and Moore, A. W. (1996), 'Reinforcement learning: A survey', *Journal of Artificial Intelligence Research* **4**, 237–285.

Kempter, R., Gerstner, W. and van Hemmen, J. L. (1999), 'Hebbian learning and spiking neurons', *Physical Review E* **59**(4), 4498–4514.
http://itb.biologie.hu-berlin.de/~kempter/Publications/1999/Phys_Rev_E_99/kempter99c.ps

Kempter, R., Gerstner, W. and van Hemmen, J. L. (2001), 'Intrinsic stabilization of output rates by spike-based Hebbian learning', *Neural Computation* **13**, 2709–2742.
http://itb.biologie.hu-berlin.de/~kempter/Publications/2001/NeuralComput/kempter01.ps

Legenstein, R., Naeger, C. and Maass, W. (2005), 'What can a neuron learn with spike-timing-dependent plasticity?', *Neural Computation* **17**, 2337–2382.
http://www.igi.tugraz.at/maass/psfiles/154.pdf

Markram, H., Lübke, J., Frotscher, M. and Sakmann, B. (1997), 'Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs', *Science* **275**(5297), 213–215.
http://www.sciencemag.org/cgi/content/full/275/5297/213?ck=nck

Meyer, J.-A. and Wilson, S. W., eds (1991), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*, MIT Press, Cambridge, MA.

Nolfi, S., Baldassarre, G., Calabretta, R., Hallam, J., Marocco, D., Miglino, O., Meyer, J.-A. and Parisi, D., eds (2006), *From animals to animats 9: Proceedings of the Ninth International Conference on Simulation of Adaptive Behaviour*, Vol. 4095 of *Lecture Notes in Artificial Intelligence*, Springer Verlag.

Pfeifer, R. and Scheier, C. (1999), *Understanding intelligence*, MIT Press, Cambridge, MA.

Pfister, J.-P., Toyoizumi, T., Barber, D. and Gerstner, W. (2006), 'Optimal spike-timing-dependent plasticity for precise action potential firing in supervised learning', *Neural Computation* **18**(6), 1318–1348.
http://diwww.epfl.ch/~gerstner/PUBLICATIONS/Pfister06.pdf

Rao, R. P. N. and Sejnowski, T. J. (2001), 'Spike-timing-dependent Hebbian plasticity as temporal difference learning', *Neural Computation* **13**(10), 2221–2237.
http://www.cs.washington.edu/homes/rao/nc_td.pdf

Roberts, P. (1999), 'Computational consequences of temporally asymmetric learning rules: I. Differential Hebbian learning', *Journal of Computational Neuroscience* **7**, 235–246.
http://www.proberts.net/PROBERTS/LEARNPAP/JCN99.pdf

Song, S., Miller, K. D. and Abbott, L. F. (2000), 'Competitive hebbian learning through spike-timing-dependent synaptic plasticity', *Nature Neuroscience* **3**, 919–926.
http://people.brandeis.edu/~abbott/papers/STDP1.pdf

Strogatz, S. H. (1994), *Nonlinear dynamics and chaos with applications to physics, biology, chemistry, and engineering*, Perseus Books, Cambridge, MA.

Sutton, R. S. and Barto, A. G. (1998), *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA.
http://www.cs.ualberta.ca/ml

Toyoizumi, T., Pfister, J.-P., Aihara, K. and Gerstner, W. (2005), Spike-timing dependent plasticity and mutual information maximization for a spiking neuron model, *in* L. Saul, Y. Weiss and L. Bottou, eds, 'Advances in Neural Information Processing Systems', Vol. 17, MIT Press, pp. 1409–1416.
http://diwww.epfl.ch/~gerstner/PUBLICATIONS/Toyoizumi05b.pdf

Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M. and Thelen, E. (2001), 'Artificial intelligence: autonomous mental development by robots and animals', *Science* **291**, 599–600.
http://www.cse.msu.edu/dl/SciencePaper.pdf