A Standard Model of Mind:
AAAI Technical Report FS-17-05

Continuous and Parallel: Challenges for a Standard Model of the Mind

Terrence C. Stewart
Chris Eliasmith
Centre for Theoretical Neuroscience, University of Waterloo, Waterloo, Canada
tcstewar@uwaterloo.ca
celiasmith@uwaterloo.ca

Abstract

We believe that a Standard Model of the Mind should take into account continuous state representations, continuous timing, continuous actions, continuous learning, and parallel control loops. For each of these, we describe initial models that we have made exploring these directions. While we have demonstrated that it is possible to construct high-level cognitive models with these features (which are uncommon in most cognitive modeling approaches), there are many theoretical challenges still to be faced to allow these features to interact in useful ways and to characterize what may be gained by including these features.

Introduction

Minds are extraordinarily complex and building models of them is difficult. To create such models, we have to approximate and abstract from the real system. Importantly, there is no one "correct" level of abstraction for the mind (c.f. Eliasmith and Trujillo 2013). As with modeling any system, the correct degree of abstraction depends on the sorts of questions being asked and the kinds of behavior we are trying to replicate or understand. As a result, when considering a Standard Model of the Mind, the question arises as to which abstractions are the right ones.

In this paper, we argue that two common abstractions found in cognitive architectures should be avoided. In particular, we are interested in characterizing the mind as being continuous (as opposed to discrete) and parallel (as opposed to serial). These are not new ideas; indeed it has long been clear that the mind is not purely discrete or serial. However, we believe that the metaphor of the discrete and serial Turing Machine as a theoretical source of understanding for minds is still prevalent, and worthy of explicitly challenging (Eliasmith 2003). Furthermore, we believe that modern computational advances give us new possibilities for investigating such models that would have been infeasible even a decade ago.

We believe that continuity and parallelism are essential ingredients for a compelling theory of cognition, and hence should be a central part of a Standard Model of the Mind. As well, both ideas are under-represented in the cognitive architecture literature, and taking them into consideration may lead to the discovery of new capabilities, algorithms, and explanations that are not typically available to current models of the mind.

Continuous State

Neural activity (i.e. spikes) are generally considered as discrete events, and there are a finite number of neurons in the brain. Consequently, there is a technical sense in which neural states are discretely characterizable (Eliasmith 2001). Indeed, mathematically speaking, continuous states carry an infinite amount of information which is generally *not* considered a reasonable means of characterizing neural states (Rieke et al. 1997). In the literature on cognitive architectures, it is also often assumed that *cognitive* states are usefully characterizable as discrete, typically using conceptual symbolic representations. However, mathematical discreteness at the neural level should not be confused with conceptual discreteness at the cognitive level.

Rather, in theoretical neuroscience it is generally agreed that neurons are best characterized as forming *distributed* representations (where multiple neurons redundantly code for multiple values). Furthermore, the spikes from neurons release neurotransmitters which are gradually reabsorbed. As a result, there is a long tradition of considering neurons as representing continuous values with finite precision (Rieke et al. 1997; Eliasmith and Anderson 2003; Salinas and Abbott 1994). These finitely accurate values are thus most naturally described in terms of metric spaces, which lack the kind of discreteness typical of symbolic characterizations of conceptual spaces.

In fact, this kind of mapping of spiking neurons to metric spaces also holds for most non-spiking neuron models as well. In this context, the main difference between spiking and non-spiking networks is the kind of noise that is typical of the system. Unsurprisingly then, theoretical frameworks, like the Neural Engineering Framework (Eliasmith & Anderson, 2003), can be used to characterize distributed spiking or non-spiking neuron populations as generating a (noisy) representation of continuous real-valued variables.

Of course, using real-valued representations raises the question of how aspects of cognition that are heavily conceptual, such as language (which seems to be very discrete), may be represented. For this, we turn to the small (but growing) field of Vector Symbolic Architectures (Gayler 2003; Plate 2003; Eliasmith 2013), which uses high-dimensional vectors as symbol-like representations in syntactic structures. In VSAs, many of the compositional aspects of symbol systems are combined with the smooth generalization of nonsymbolic systems. This has led to novel explanations of the gradual decay of working memory (Choo and Eliasmith 2010), the number of terms and size of sentences in human language (Eliasmith 2013), and how the brain might perform inductive reasoning on general intelligence tests (Rasmussen and Eliasmith 2014). A Standard Model of the Mind should allow for continuous symbol-like representations, and VSAs are one good candidate for how that can be realized.

Continuous Time and Continuous Actions

The mind is a physical system, and physical systems do not change state instantaneously. That is, cognitive models should not jump instantly from one state to another. Rather, state transitions (from one continuous-valued state to another) take place over time, and during that time the continuous-valued representations gradually change. A Standard Model of the Mind should allow for continuous, smooth state transitions.

In some situations, the time course of this temporal process leads to predictions about the timing of cognitive models. For example, we have previously shown that the dynamics of changing values in a model of the cortex-basal ganglia-thalamus loop can explain the 50ms cognitive cycle time (Stewart, Choo, and Eliasmith 2010).

However, the introduction of these dynamics into models of the mind poses new challenges. We previously adapted an existing production-system model of expert behaviour in the Tower of Hanoi task into a continuous-time spiking-neuron-based model (Stewart and Eliasmith 2011). While this was successful, it also involved careful balancing of timing by being concerned with the inherent dynamics of the system. While one action is being performed (e.g., shifting attention from one disc to the next), the system must maintain the signal to perform that action, and it must do so *while the effects of that action are being implemented*. In this example, the representation of which disc is currently being attended to will gradually change (over a few tens of milliseconds) from the neural representation of one disc to that of another. The problem is that this change results in the very action that is being performed to become gradually less appropriate while the action is occurring. Ideally, this leads to a natural way of smoothly stopping an action as it completes. However, if there are other potential actions competing with the current one, balancing these behaviors becomes complex because the action we are attempting to complete becomes less important relative to actions we may not yet want to perform. This kind of interference is a result of smoothly changing actions, something avoided by discrete models, but inherent to real brains.

While we do not have a panacea to alleviate these concerns, we do believe that this is more than a "mere implementation" issue. If internal representations change gradually over time, the dynamics of these changes should be exploitable by a cognitive system. Neurons can very efficiently implement dynamical systems and attractor networks (Eliasmith 2005) and it seems likely to us that these timing effects and continuously changing representations would be used by a cognitive system to guide its behavior.

Continuous Learning

We believe that any learning processes within a Standard Model of the Mind must be continuous in another sense: the model must continue to learn over its entire lifetime. While this fits well with many cognitive architectures (for example, ACT-R's declarative and procedural learning mechanisms), it is concerning to us that many large connectionist learning systems, including the majority of deep neural networks, make use of a gradually diminishing learning rate. This common practice of starting with a high learning rate and slowly reducing it (i.e. simulated annealing) implies a distinct beginning and end to the learning process, and seems suited only for environments whose statistics do not change.

One way to address this issue without removing this highly useful adjustment of the learning rate would be to explicitly have the learning rate under the control of the model itself. This would require having the model determine for itself the useful value of this learning rate parameter, and that itself may be difficult, given that currently many learning rate trajectories in neural networks are hand-tuned for different circumstances by the researcher.

Furthermore, it should be noted that deep neural network learning is frequently discontinuous in another way: deep learning often makes use of *batch* learning, where learning signals are collected for a period of time, and then applied all at once. While this is often described as being similar to the sleep/wake cycle in living creatures, it seems to us that a more specific theory of the details of this relationship would be needed if such a system were to be used in a Standard Model of the Mind.

Our work in this direction has mostly focused on adaptation in the motor system and reinforcement learning (RL). In the motor system, we have shown that a continuously learning system based on mammalian motor cortices and the cerebellum can adjust to changes in the motor system and the environment (DeWolf et al 2016). Importantly, this learning is stable when given a fixed environment, but adapts quickly when and if the environment changes (such as dealing with the unexpected forces that occur when picking up an object of unknown mass). The core underlying learning algorithm has been proved to be stable (Slotine and Li 1987), and also works for other systems such as quadcopter control (Komer 2015).

However, adaptive motor control is not enough. A Standard Model of the Mind will also need to be continuously learning in many other ways as well. Classical and operant conditioning need to occur, including standard effects such as second order conditioning (not learning new associations if existing associations sufficiently explain observed effects) and spontaneous recovery (associations than have been extinguished suddenly re-appearing when an agent is placed in a new environment). While we have developed a neural models of these effects (Kolbeck, Bekolay, and Eliasmith 2013), more sophisticated forms of RL are also critical. For instance, we have developed methods for RL in continuous time and continuous state settings, that can operate in sMDP environments (where the time between an event and related rewards is continuous and unknown) and use hierarchical methods (Rasmussen, Voelker, and Eliasmith 2017). We have used these hierarchical methods to demonstrate transfer learning (Rasmussen, Voelker, and Eliasmith 2017). It is also worth noting that the PRIMs model (Taatgen 2013) makes extensive use of continuous learning to deal with transfer of skills from one task to another.

In addition, we have shown that you can use continuous learning methods to learn the operators needed to implement the symbol-like structures mentioned above (Bekolay, Kolbeck, and Eliasmith 2013).

A Standard Model of the Mind is likely to require many different learning systems, all continuously running during the lifetime of the model. These learning systems will interact in possibly unexpected ways. We believe that making this interaction stable over long time scales will be an important challenge for large-scale cognitive modeling.

Parallel Control

As more and more systems are added to a Standard Model of the Mind, a new challenge also arises. Many cognitive modeling architectures, including our own large-scale modeling efforts (Eliasmith et al. 2012), make use of a single serial bottleneck for controlling cognition. While there is extensive behavioral evidence supporting such a bottleneck, it is likely not the only control loop that exists.

In particular, it is clear that complex automatic reflex actions are exhibited by cognitive systems, and that these reflexes can be both built-in and learned. These are often thought of as purely stimulus-response systems, and may be treated as direct connections from sensory systems to motor systems (bypassing any central cognitive control). However, we believe that similar systems also exist internally to the agent, leading to complex interactions with any central serial bottleneck for cognition.

Indeed, neural evidence suggests that even in the basal ganglia (often considered to be the neurological correlate for the serial bottleneck), there are multiple parallel control loops. The selection process appears to be separated into different types of actions, including limbic, associative, sensory, and motor actions (Alexander, DeLong, and Strick 1986). Furthermore, there is evidence that other structures also perform similar control functions as the basal ganglia. For example, the amygdala is sometimes seen as a coordination hub, receiving inputs from a wide range of cortical structures, and providing specific output for thalamic routing and control of other regions (Bickart, Dickerson, and Barrett 2014) and there is extensive evidence for other loops as well (e.g. Solari and Stoner 2011).

Conclusions

We have identified a variety of challenges for a Standard Model of Mind that have become salient as we have developed cognitive models that step away from the common approach of assuming events are discrete, representations are discrete, and control between modules is achieved with a single serial bottleneck. We do not yet have solutions that comprehensively deal with these challenges. However, we have shown that it is possible to build large-scale integrated continuous models (e.g. Eliasmith et al. 2012), and we have shown that it is possible to take advantage of continuous representations to provide novel explanations of cognitive behaviors (Rasmussen and Eliasmith 2014) and timing (Stewart, Choo, and Eliasmith 2010). We have also shown that continuous learning is useful for motor control (DeWolf et al 2016) and extending reinforcement learning (Rasmussen, Voelker, and Eliasmith 2017). Importantly, we have shown that these continuous systems can also exhibit behavior that seems symbolic and discrete, allowing us to investigate continuous neural systems while considering the sorts of high-level cognitive behavior that inspires most cognitive architecture research.

Our goal is to work towards cognitive architectures that directly incorporate continuous representations, continuous timing, continuous actions, and continuous learning systems, while combining multiple control systems of different types. We believe this pushes the field of cognitive architectures in novel and scientifically useful directions. In particular, parallel and continuous characterizations of cognition may be critical for explaining aspects of the mind that have been difficult to characterize, such as how motor and perceptual systems relate to higher-level cognition. In summary, we strongly believe that a Standard Model of the Mind should support continuous and parallel characterizations of cognition that are often avoided by contemporary cognitive architectures.

References

Alexander, G. E.; DeLong, M. R.; and Strick, P. L. 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review Neuroscience* 9, 357-381.

Bekolay, T.; Kolbeck, C., and Eliasmith, C. Simultaneous unsupervised and supervised learning of cognitive functions in biologically plausible spiking neural networks. In Proceedings of the International Conference of the Cognitive Science Society, 169–174. Berlin, Germany: Cognitive Science Society.

Bickart, K. C.; Dickerson, B. C.; and Barrett, L. F. 2014. The amygdala as a hub in brain networks that support social life. *Neuropsychologia* 63, 235-248.

Choo, X., and Eliasmith C. 2010. A Spiking Neuron Model of Serial-Order Recall. In Proceedings of the International Conference of the Cognitive Science Society, 2188-2193. Portland, OR, Cognitive Science Society.

DeWolf, T.; Stewart, T.C.; Slotine, J.-J.; and Eliasmith, C. 2016. A spiking neural model of adaptive arm control. *Proceedings of the Royal Society B* 283(1843).

Eliasmith, C. 2001. Attractive and in-discrete: a critique of two putative virtues of the dynamicist theory of mind. *Minds and Machines* 11:417-426.

Eliasmith, C. 2003. Moving beyond metaphors: understanding the mind for what it is. *Journal of Philosophy* 100:493-520.

Eliasmith, C. 2005. A unified approach to building and controlling spiking attractor networks. *Neural computation* 7(6):1276-1314.

Eliasmith, C. 2013. *How to build a brain: A neural architecture for biological cognition*. New York, NY: Oxford University Press.

Eliasmith, C., and Anderson, C. H. 2003. *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. Cambridge, MA: MIT Press.

Eliasmith, C.; Stewart, T. C.; Choo, X.; Bekolay, T.; DeWolf, T.; Tang, Y; and Rasmussen, D. 2012. A large-scale model of the functioning brain. *Science* 338:1202-1205.

Eliasmith, C., and Trujillo, O. 2013. The use and abuse of large-scale brain models. *Current Opinion in Neurobiology* 25:1–6.

Gayler, R. 2003. Vector symbolic architectures answer Jackendoff's challenges for cognitive neuroscience. In Proceedings of the ICCS/ASCS International Conference on Cognitive Science, 133-138. Sydney, Australia: University of New South Wales.

Kolbeck, C.; Bekolay, T., and Eliasmith, C. 2013. A biologically plausible spiking neuron model of fear conditioning. In Proceedings of the International Conference on Cognitive Modeling, 53–58. Ottawa, Canada: Carleton University.

Komer, B. 2015. Biologically inspired adaptive control of quadcopter flight. M.A.Sc. diss. Computer Science, University of Waterloo, Waterloo, ON.

Plate, T. 2003. *Holographic reduced representations*. Stanford, CA: CSLI Publication.

Rasmussen, D., and Eliasmith, C. 2014. A spiking neural model applied to the study of human performance and cognitive decline on raven's advanced progressive matrices. *Intelligence* 42:53–82.

Rasmussen, D.; Voelker, A.; and Eliasmith, C. 2017. A neural model of hierarchical reinforcement learning. *PLOS ONE* 12(7).

Rieke, F.; Warland, D.; de Ruyter van Steveninick, R.; and Bialek, W. 1997. *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Salinas, E., and Abbott, L. F. 1994. Vector reconstruction from firing rates. *Journal of Computational Neuroscience* 1, 89–107.

Slotine J.-J., and Li W. 1987. On the adaptive control of robot manipulators. *International Journal of Robotics Research* 6, 49–59.

Solari, S., and Stoner, R. 2011. Cognitive consilience: Primate non-primary neuroanatomical circuits underlying cognition. *Frontiers in Neuroanatomy* 5:65.

Stewart, T.C.; Choo, X; and Eliasmith C. 2010. Dynamic Behaviour of a Spiking Model of Action Selection in the Basal Ganglia. In Proceedings of the International Conference on Cognitive Modeling, 235-240. Philadelphia, PA: Drexel University.

Stewart, T.C., and Eliasmith C. 2011. Neural Cognitive Modelling: A Biologically Constrained Spiking Neuron Model of the Tower of Hanoi Task. In Proceedings of the International Conference of the Cognitive Science Society, 656-661. Austin, TX: Cognitive Science Society.

Taatgen, N.A. 2013. The nature and transfer of cognitive skills. *Psychological Review* 120(3), 439-471.