Human Computation and Serious Games: Papers from the 2012 AIIDE Joint Workshop
AAAI Technical Report WS-12-17

Representing the Human to the Systems that They Use
A Role for Neurocognitive Models

Joseph V. Cohn and Elizabeth B. O'Neill
Office of Naval Research
875 N Randolph Street
Arlington VA 22203-1995

Abstract

The systems with which humans interact are becoming increasingly complex; there is a corresponding need for such systems to anticipate and understand human action. Current approaches to develop this ability do not robustly represent human users to the systems with which they are interacting. Fusing neurocognitive models with existing approaches may provide an effective way to capture and represent neural action in a way that behaviors can be predicted and shared with the systems that humans are operating.

Introduction

Traditional approaches to creating effective couplings between humans and the systems they use are often based on artificial intelligence or machine learning techniques that focus on detecting statistical or probabilistic regularities in data (Cooley, 2007) or on cognitive architectures that are based on computer processing metaphors.

Because these approaches are not grounded in the core processes that drive human action, the resultant outputs – predictions of behavior, estimates of errors and the like – do not provide a robust basis for representing human users to the systems with which they are interacting. Consequently, the responses that these systems provide are oftentimes inappropriate or insufficient to the current task demands. This state of affairs is a direct result of the levels of technology available to understand and represent the processes through which the human brain transforms information into action. In the past, when one wished to create effective human system collaborations, one was forced to do so either by building predictive models based on human behaviors or by developing symbolic representations of the unseen and uncharacterized cognitive processes leading to those behaviors. Notably, the link between brain, cognition and behavior cannot be captured through these techniques.

A more effective solution is to develop methods for capturing and representing neural action in such a way that the resultant behaviors can be predicted and shared with the systems that humans are operating. These representations may be found in the neural processes leading to the actual, observed behavior. Just as understanding the equations of motion provides a much broader set of capabilities than inferring these equations from a limited set of observations (Kelso, 1995), so too understanding and modeling the dynamics of neural activity as it leads to behavior should provide a much richer and more robust set of models than those based on either observed behavior or inferred cognitive processes. Today, advances in neuroscience and engineering provide the basis for building these neurocognitive models and for using brain-based techniques to create and maintain very robust human machine interactions. The net result of this approach should be to either provide a viable alternative to classical artificial intelligence / machine learning (AI, ML) approaches or. Alternatively, to provide a more neurocognitively – inspired approach to developing these AI and ML routines.

Background

As early as the 1940s, researchers were concerned with the question of how to represent the human element in human machine systems. Bates (1947), Craik (1947/1948; 1948) and others attempted to represent human performance in control theory terms with the goal of developing engineering representations of the human that could be used to improve the effectiveness of human machine interactions.

In all cases, the properties of the human being modeled resided at the observed behavior level. For example, Fitts' speed-accuracy tradeoff (Fitts, 1954) emphasized the development of basic relationships guiding human motor and cognitive action in response to specific dimensions of a cue; and the Hick-Hyman law related decision response time to the number of possible choices (Hick, 1952; Hyman, 1953). These behavior based representations were effective for well-bounded and simple tasks but did not prove capable of predicting large scale behaviors in response to complex or open ended tasks.

Recent advances in neuroscience and related technologies have provided deeper insights into how neural activity gives rise to observed behaviors. Consequently, it has become possible to develop a more dynamic and integrative theory of brain in which different regions each contribute, in different ways depending on the task environment and user state, to the processing of information leading to behavior (Singer, 1999; Philiastides & Sajda, 2007). These regions form 'ad hoc' networks across the physical substrate of the brain via synchronization signals that briefly bind them together, leading to observed behavior (Kahana, 2006; Lisman, 1995; Singer, 1999). Importantly, human performance in this view, results from the interaction and integration of 'building blocks' like perception, attention & memory, which, in turn results from activity across multiple brain regions. Guided by these promising advances in understanding the brain's operating principles, and harnessing and extending brain imaging technologies, advanced signal processing techniques and data modeling approaches, it is now possible to begin to develop representations of human performance that are able to more effectively:

1. Adapt to new situations
2. Account for individual users' varying physiological / mental states
3. Update based on individual users' experiences

Background

These detailed representations of human performance are known as Neurocognitive models. Neurocognitive models are based on the idea that understanding how human cognition evolves in the context of neural action is key to understanding human behavior, because cognition is how the human brain transforms sensed information into behavior.

On the surface, addressing this challenge may appear to be a tall order. Interpreting actual meaning from neural activity has been a long-sought dream of the neuroscience community. In the past, the reliance of cognitive models on observed behavior was in many ways a tacit admission that, while there is a great need to 'go to the brain', in the available technologies and theories at the time simply precluded doing this – making observed actions the most readily accessible feature of human behavior that could be accessed to provide critical data to populate models. Recently, though, there has been a shift in theories of brain, made possible by advances in neuroimaging technologies, data analysis techniques and representational methods. These advances have been made possible as direct result of developments in three core domains:

1. Neural Activity Detection Technologies
2. Decoding Methodologies
3. Modeling Approaches

Neural Activity Detection Technologies

Access to the brain has been one of the key limiting steps in demonstrating coordinated activity across the brain as behavior develops.

Technologies, like functional Magnetic Resonance Imaging (fMRI), dense array Electroencephalography (dEEG), and others are at the point where they may now be applied to the challenge of capturing integrated neural action as it occurs simultaneously across multiple brain regions. Figure (1) provides an overview of some of these technologies. Detection technologies can categorized in terms of those which require penetration through the head ('Invasive') and those which do not ('Non-Invasive'). Data sources range from subcellular processes, like ion flow through individual membrane spanning channels, to individual nerve cell action, neuron-neuron synchronization, neural function, and neural structures. The processes represented by each of these data sources occur across different time scales. Lastly, different data sources provide insight into different scales of neural action (e.g. ion flow vice integrated action across brain regions).

Decoding Methodologies

Access to integrated neural data is necessary but not sufficient for representing performance. New methods for analyzing these multivariate data sets must also be established, refined and applied to data captured as users perform a range of cognitive and motor tasks. One promising emerging technique, multivariate decoding (Mitchell et al 2004), has the ability to take into account the full spatial pattern of brain activity, measured simultaneously across many regions, enabling the decoding and translation of measured brain activity. Other machine learning-based techniques (e.g. Kay et al, 2008) show similar promise.

Modeling Approaches

The final challenge is developing representation frameworks into which the detected and decoded information may be organized, to simulate and predict human performance as part of a neuroadaptive human - system control scheme. Possible solutions include extending current modeling architectures to handle neural data (Anderson, et al 2008), developing software or silicon-based representations of the detected and decoded data that produce behaviors similar to those that result from actual neural activity (Fleisher & Krichmar, 2007) or developing frameworks intended to capture neurocognitive data from the outset (Figure 2; Cohn, 2012).

Summary

As the systems with which humans interact become increasingly complex, there is a corresponding need for these systems to be able to anticipate and understand human action. Many current and anticipated technologies rely on classical Artificial Intelligence or Machine Learning approaches to provide these representations. However, such approaches fail to capture the richness of human neurocognitive processes. Blended approaches fusing neural, cognitive, and related measures with the power of Artificial Intelligence or Machine Learning may alleviate these challenges.

References

Anderson, J. R., Carter, C. S., Fincham, J. M., Qin, Y., Ravizza, S. M., & Rosenberg-Lee, M. (2008). Using fMRI to Test Models of Complex Cognition. Cognitive Science, 32, 1323-1348.

Bates, J. A. V. (1947). Some characteristics of a human operator. Journal of the Institute of Electrical Engineering, 94:298-304.

Cohn, J.V., & Wheeler, T. (2012). Neuroadaptive systems: challenges and opportunities with creating symbiotic relationships between humans and the machines they use. In: M. Fafrowicz, T. Marek,W. Karwowski, D. Schmorrow (Eds). Neuroadaptive Systems: research, theory, and applications. Boca Raton:CRC Press

Cooley, M. (2007). Cognition, communication and interaction: Transdisciplinary perspectives on interactive technology. In S.P. Gill (Ed.), On Human-Machine Symbiosis Human-Computer Interaction Series. (pp. 457-485). London: Springer.

Craik, K. J. W. (1947/1948). Theory of the human operator in control systems I: The operation of the human operator in control systems. British Journal of Psychology, 38:56-61.

Craik, K. J. W. (1948). Theory of the human operator in control systems II: Man as an element in a control system. British Journal of Psychology, 38:142-148.

Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47(6):381-391.

Fleischer, J.G. & Krichmar, J.L. (2007). Sensory integration and remapping in a medial temporal lobe model during maze navigation by a brain-based device. Journal of Integrative Neuroscience, 6(3):403-431.

Hick, W. E. (1952). On the rate of gain of information. Quarterly Journal of Experimental Psychology, 4:11-26.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. Journal of Experimental Psychology, 45:188-196.

Kahana, M. J. (2006). The cognitive correlates of human brain oscillations. The Journal of Neuroscience, 26(6):1669-1672.

Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Identifying natural images from human brain activity. Nature, 452, 352-355.

Lisman, J. E., & Idiart, M. A. P. (1995). Short-term memories in oscillatory subcycles. Science, 267(5203):1512-1515.

Mitchell, T., Hutchinson, R., Niculescu, R. S., Pereira, F., Wang, X., Just, M. A., & Newman, S. D. (2004). Learning to decode cognitive states from brain images. Machine Learning, 57:145-175.

Philiastides, M. G. & Sajda, P. (2007). EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. Journal of Neuroscience, 27(48): 13082-13091.

Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? Neuron, 24:49-65.