Understanding the Brain's Emergent Properties

Don Miner and **Marc Pickett** and **Marie desJardins**
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
1000 Hilltop Circle, Baltimore, Maryland 21250

Abstract

In this paper, we discuss the possibility of applying rule abstraction, a method designed to understand emergent systems, to the physiology of the brain. Rule abstraction reduces complex systems into simpler subsystems, each of which are then understood in terms of their respective subsystems. This process aids in the understanding of complex systems and how behavior emerges from the low-level interactions. We believe that this technique can be applied to the brain in order to understand the mind and its essential cognitive phenomena. Once a sufficient model of the brain and mind is created, our framework could then be used to build artificial general intelligence that is based on human intelligence.

Introduction

In this paper, we propose a method of understanding human intelligence by understanding how the mind emerges from the physiology of the brain. The brain may be viewed as a complex system that produces features of human-level intelligence from the low-level physical mechanisms in the neural system. We hypothesize that we can improve our understanding of how the brain works by reducing its emergent behavior into layers of successively more complex behaviors on top of the neurological subsystem. To achieve this goal, we propose the use of *rule abstraction*, our mechanism for creating hierarchies of emergent behaviors (discussed in more detail in the next section). The purpose of this paper is to stimulate discussion about the value of such an approach for understanding the human brain and, as a result, understand intelligence.

Understanding the mind by directly studying low-level structures, such as neurons and glial cells has not proven fruitful to date. For example, biologically inspired systems such as Jeff Hawkins' Memory Prediction (Hawkins & Blakeslee 2004) and Blue Brain (Markram 2006) have not led to general models of intelligence. The leap from neurons to high-level processes, such as reasoning and language, is too great for humans or machines to decipher in a single step as of 2009. However, in smaller-scale complex systems, such as boid flocking (Reynolds 1987), we can mathematically model how simple agent-level rules produce the flocking emergent behavior (Miner, Hamilton, & desJardins 2008).

We propose trying to understand the brain by first partitioning it into hierarchical sub-processes. Each sub-process has emergent behavior that results from the emergent behavior of its lesser sub-processes. Then, we find mathematical correlations between low-level behaviors and abstract-level properties using this sub-process structure. The result will be a hierarchy of successively more complex emergent systems.

In this paper, we outline how rule abstraction and hierarchies can be used to understand emergent systems. We then discuss the challenges in applying this method to the brain and intelligence.

Rule Abstraction and Hierarchies

Rule abstraction is the process of finding a mathematical correlation between low-level rules and abstract properties in an emergent system (Miner, Hamilton, & desJardins 2008). Low-level rules are the basic actions and atomic structures in the emergent system. Abstract properties are the higher-level emergent behaviors. For example, a multicellular organism exhibits emergent behaviors (the abstract properties) that result from the individual behaviors of the cells (the low-level rules).

Applying rule abstraction may be difficult in some emergent systems. The most obvious approach to rule abstraction is manually specifying the correlation. This requires a significant level of human intuition into how the abstract level properties of the complex system emerge. The lack of human understanding of complex systems makes manual specification impossible at this time. To make this problem tractable, a number of computer science and artificial intelligence techniques can be applied to learn the mathematical correlations between the low-level rules and the abstract properties. For example, we have experimented with a sampling and regression technique in which we observed several different configurations of a complex system and then used regression to create a continuous two-way mapping of the low-level rules and the abstract properties (Miner, Hamilton, & desJardins 2008). Similarly, a theory of a single cell's behavior could be developed by observing several different cells with different internal configurations.

This methodology provides the ability to use previously mapped abstract properties as low-level rules in higher-order complex systems. For example, once we have developed a

theory of a cell, we can use its mapping to develop theories of small multicellular organisms or organs in larger organisms. The intermediate step of the cell theory enabled an understanding of multicellular organisms that may not have been previously possible. These hierarchies of emergence are clear in many cases. However, in other complex systems, such as the brain, emergent sub-processes may not be easily identifiable.

Understanding the Brain

The brain is a mysterious complex system. Rule abstraction is general enough, in theory, to handle any emergent system. However, there are three key challenges. First, a certain level of human engineering is currently required to identify the appropriate level of abstraction. Second, in the brain, the correlation models that are required may be too complex and may require more sophisticated learning methods than what we have tried with rule abstraction. Finally, it could be the case that the system we are trying to understand simply has no "midpoints." That is, emergent behavior results from the low-level rules and no meaningful reduction of abstraction layers can be found. Regardless of these hurdles, we hope that our approach will one day be able to build artificial general intelligence.

The first step to applying rule abstraction to the brain and mind, as with any complex system, is by declaring the obvious: the cognitive powers of the mind and brain result from the physiology's emergent properties. This statement represents the initial state of the hierarchy. At this point, learning a mathematical correlation would be too difficult. To overcome this problem, we break down the complex system by adding additional mid-level abstract properties to the hierarchy and learning the correlations between these, instead. For example, the top-level emergent behavior may be a combination of lesser emergent properties such as language, memory, reasoning, etc. and could each be added as nodes in the hierarchy. A similar approach can be taken from the bottom-up: the neurons, glial cells, chemical reactions, and other physiology could be parts of emergent subsystems in the brain. These subsystems may include physical structures such as cortical columns or more abstract features such as prediction in the auditory cortex. Hopefully, at some point, no correlation in the hierarchy between higher-order emergent subsystems and lower-order emergent subsystems will be too difficult to learn.

We do not believe that achieving artificial general intelligence through this type of framework is possible at this time because neuroscience, cognitive science and other related fields not yet able to explain how the mind works as a whole. However, we do believe that our framework scales to any complex system and thus increasingly accurate rule abstraction hierarchies can be built as more scientific information on the brain and mind are gathered.

Discussion

We now ask several questions to ourselves and the research community. Answers to these questions would be useful in understanding emergence, general intelligence and specifically human intelligence.

How many midpoints or layers would be in a rule abstraction hierarchy model of a brain? If there are too many layers (greater than ten), too much error may be introduced and may yield a unfavorable results. If there are too few layers (less than three), current machine learning techniques may not be powerful enough to build correlations in these massive emergent systems. Also, how deep does the model have to be? Strong enough abstract models of cortical columns may make modeling individual neurons unnecessary. On the other hand, perhaps a neural-level base in the hierarchy is not deep enough.

Unfortunately, there is much speculation and much uncertainty in defining the mind's subsystems, due to lack of scientific understanding in relevant fields. Concepts such as language, memory and reasoning are easily observable, but are there some phenomena that have not been discovered? Also, specifying the order of known phenomena in a hierarchy is difficult. Do some of these known phenomena, such as language, emerge from other essential subsystems and thus not a foundation of general intelligence?

We may be able to take lessons learned from using rule abstraction in simpler domains and apply them to the brain. Perhaps nature has used emergent systems similar to the ones in our brains in other complex systems such as ants, traffic jams, rat brains, and galaxies. Is there some overarching theme of emergence in our universe? Is there a general theory of emergence? This question may be harder to answer than understanding the brain and developing artificial general intelligence. However, any useful information that comes out of trying to answer this question may be helpful in understanding human-level intelligence.

Conclusions

We have given an overview of our method of rule abstraction, and explained how it can bring new understanding to emergent systems. We believe that rule abstraction could be applied to the brain in order to understand the mind. We hope that by introducing this new approach, we will stimulate discussion on our method's usefulness for this domain and will inspire novel views of human intelligence.

References

Hawkins, J., and Blakeslee, S. 2004. *On Intelligence*. Times Books.

Markram, H. 2006. The Blue Brain Project. *Nature Reviews Neuroscience* 7(2):153–160.

Miner, D.; Hamilton, P.; and desJardins, M. 2008. The Swarm Application Framework. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence (Student Abstract)*.

Reynolds, C. W. 1987. Flocks, herds and schools: A distributed behavioral model. *SIGGRAPH Comput. Graph.* 21(4):25–34.