In Proceedings of the Sixth International Conference on Cognitive Modeling, 408-409.
Mahwah, NJ: Lawrence Earlbaum.

Emergence of Bayesian Structures from Recurrent Networks

Sébastien Hélie (helie.sebastien@courrier.uqam.ca)
Department of Computer Science, C.P. 8888 Succ. Centre-Ville
Montréal, PQ H3C 3P8 CANADA

Abstract

The problem of representational form has always limited the applicability of cognitive models: where symbolic representations have succeeded, distributed representations have failed, and vice-versa. Hybrid modeling is thus a promising venue, which however brings its share of new problems. For instance, it doubles the number of necessary assumptions. To counter this problem, we believe that one network should generate the other. This would require specific assumptions for only one network. In the present project, we plan to use a recurrent network to generate a Bayesian network. The former will be used to model low-level cognition while the latter will represent higher-level cognition. Moreover, both models will be active in every task and will need to communicate in order to generate a unique answer.

General Problem

In cognitive science, the problem of representational form is crucial. During the cognitive revolution, the computer metaphor was used to model human intelligence, which was thus seen as a set of symbol-manipulating syntactic processes (Turing, 1936). These processes were modeled as a series of conjunctive conditions and consequential actions (known as "IF-THEN" rules). This modeling approach is referred to as the classical view (Russel & Norvig, 1995).

In the late seventies, another metaphor became increasingly popular for modeling cognitive processes, namely: the brain. The connectionist (or "neural") networks proposed during this period were mostly unsupervised networks, either competitive (Grossberg, 1976; Kohonen, 1982) or recurrent (Anderson, Silverstein, Ritz & Jones, 1977). Another class of neural networks, first proposed by Rosenblatt (1958) and further developed by Rumelhart and McClelland (1986), was based on error backpropagation. These supervised networks have since been proven to succeed in fitting complex data but are often criticized on the grounds of their biological implausibility.

Connectionist networks, in general, differ from the classical view in that they are based on a conception of intelligence as a set of parallel processes acting on distributed representations. Moreover, while the distinction between processes and representations is quite straightforward in classical models, here this difference is a bit fuzzier.

Finally, in the early nineties, pioneers of the second wave of artificial intelligence started using symbol-manipulating probabilistic models which were either operating in a serial (e.g. Hidden Markov Models, see Rabiner, 1989) or parallel manner (e.g. Bayesian Networks, see Pearl, 1988).

Each of the previous approaches has important shortcomings. For example, while symbolic models can easily be used to model tasks involving language and reasoning (e.g. the use of recursive rules, Marcus, 2001), they are considered less useful for modeling low-level similarity-based perceptual processes. This has been shown to be better achieved through the use of connectionist models (Anderson & al., 1977; Grossberg, 1977).

Two important points are of interest here. First, approaches can be classified using their representational format and this classification is informative about what kind of task it will do best. Second, and more importantly, we see that where one approach succeeds, the other falls short of explanation.

Hypothesis

Since humans are able to use both recursive rules and similarity-based judgments, they must be able to use both kinds of representations. First, we will argue that low-level implicit knowledge is best modeled by an unsupervised neural network (using distributed representations), while higher-level cognitive operations are best modeled by symbolic networks. Second, those two kinds of representational format must be able to coexist and work together, sharing the same goals.

Proposed Thesis

In the proposed project, we will use an unsupervised neural network (either competitive or recurrent) to model low-level cognition. For modeling higher-order processes, because we believe that all levels of cognition share the same parallel architecture, we will use a parallel probabilistic architecture, namely a Bayesian network (Pearl, 1988). Moreover, the use of Bayesian networks is consistent with Anderson's rational analysis of cognition (Anderson & Milson, 1989; Oaksford & Chater, 1998). The main problem is: How can an unsupervised neural network build, from scratch, a fully functional Bayesian network?

First, the unsupervised neural network will have to generate nodes (symbols) and connections for a Bayesian network. Second, each network will eventually be individually functional, and they will each be capable of learning independently. Finally, both networks will have to communicate in an interactive way in order to produce a single answer from separate processes.

What has already been done

The first part of this thesis consists in identifying which properties are needed by a connectionist network to construct a Bayesian network. First, the network must be able to identify what is important in the environment and generate symbols according to those findings. Second, it must be able to feed the Bayesian network with the probability distribution of it environment.

This second property was used to compare a hard competitive neural network[1] with the Brain-State-in-a-Box (BSB, Anderson et al., 1977), which was chosen because it is the simplest form of recurrent networks (our findings will thus generalize to other, more recent recurrent networks). We tested these networks with four environmental distributions: a bimodal distribution consisting of two Gaussian distributions, a step distribution, the exponential distribution, and the uniform distribution (Hélie & Proulx, 2003). The choice of these distributions was motivated by the fact that they could either vary the frequency of the exemplars, the frequency of the categories, both, or neither. Results show that the competitive network is not sensitive to the frequency of the exemplars. Moreover, this network was unable to learn categories that were relatively rare. The recurrent network, on the other hand, was sensitive to both frequency of exemplars and frequency of categories. The former affected the position of the eigenvectors of the weights matrix, while the latter affected the magnitude of associated eigenvalues.

The results of this research thus suggest the use of recurrent networks to generate a Bayesian network.

What is left to do

Once the lower-level network has been chosen, the next step will be to find out how this network can generate useful symbols from the environment. Recurrent networks are particularly interesting because of their dynamic properties: these dynamic properties may thus be used in symbol generation. For example, we can see activation from the environment as a trajectory in a multidimensional space and try to find useful trajectories to create symbols.

Another important unresolved issue is the nature of the interaction between both models (once they will be fully functional). Should this interaction be just a crosstalk or should it involve weights that can be trained? In the later case, should these weights affect the learning of the networks (e.g.. BAM, see Kosko, 1988)? Should networks compete to give the answer (Logan, 1988) or, work together, adding evidence until a criterion is reached (Cousineau, Lacroix & Hélie, 2003)? Clearly, a lot is left to explore.

[1] Hard competitive networks update the weights of only one winning unit. They must be distinguished from soft competitive networks were the weights of several winning units are updated simultaneously.

Acknowledgments

This research is supported by scholarships from *Le fonds québecois de la recherche sur la nature et les technologies* and the *Natural Sciences and Engineering Research Council of Canada*.

References

Anderson, J. A., Silverstein, J. W., Ritz, S. A., Jones, R. S. (1977). Distinctive features, categorical perception, and probability learning: Applications of a neural model. *Psychological Review*, *84*, 413-451.

Anderson, J. R., Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*, 703-719.

Cousineau, D., Lacroix, G. L., Hélie, S. (2003). Redefining the rules: Providing race models with a connectionist learning rule. *Connection Science*, *15*, 27-43.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics*, *23*, 121-134.

Hélie, S., Proulx, R. (2003). Effet de la distribution environnementale sur l'apprentissage dans un réseau autoassociatif. *26ième Congrès Annuel de la Société Québécoise pour la Recherche en Psychologie*. Montréal, Qc.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*, 59-69.

Kosko, B. (1988). Bidirectional Associative Memories. *IEEE Transactions on Systems, Man, and Cybernetics*, *18*, 49-60.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492-527.

Marcus, G. F. (2001). *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. The MIT Press, Cambridge, MA.

Oaksford, M., Chater, N. (1998). *Rational Models of Cognition*. Oxford University Press, Oxford.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, California.

Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*, 257-285.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*, 386-408.

Rumelhart, D. E., McClelland, J. L. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.

Russell, S., Norvig, P. (1995). *Artificial Intelligence : A Modern Approach*. Prentice-Hall, Upper Saddle River, NJ.

Turing, A. M. (1936). On computable numbers with an application to the Entscheidugsproblem. *Proceedings of the Mathematical Society, série 2*, *42*, 230-265.