Inference using a Connectionist Model

Today in cognitive modeling, there are 2 dominant paradigms: connectionist (or emergentist) approaches and Bayesian (or structured probabilistic) approaches. Connectionist models function at a representational level to explain how psychological phenomenon might emerge from the brain using neural networks as a plausibly biological implementation. Bayesian models work at a higher, computational level to explain what the brain does using probabilistic models. A common criticism of Bayesian approaches is that this probabilistic analysis is not what the brain does. Particularly, inference methods, such as MCMC, seem suspect. We might ground Bayesian approaches by finding connectionist models capable of inference as an emergent property of the network dynamics.

Currently, certain connectionist models exist that are capable of doing inference, though they are highly constrained. For example, Boltzmann machines can do Gibbs sampling [1], an idealized neuron can do basic Bayesian inference [2], and the interactive activation model can also behave in a Bayes-optimal way (unpublished work). That these exist, however, is only somewhat satisfying as the computational ability of neural networks is wide and varied. Whether we can find a similarly optimal class of models with more open, less contrived design remains an open question.

For a project, we can explore this question by picking a simple problem, such as noisy-or or cue combination [3], and attempt to generate the same distributions with a network. To do this, we need to search over a large space of networks with different sizes, activation rules, noise, and any other parameter. Regardless of its nature, the network can be run for several thousand steps, sampling the state of the relevant variables at regular intervals to generate a distribution over states. Moreover, we hope that that network would be fully generative. We can test this as well by clamping different units (as though it were observed) in the network to values, checking the resulting distribution against the corresponding conditional distributions. Afterwards, we can calculate the goodness of a network with KL-divergence between the network results against true distributions. This metric allows us to recognize promising networks, though the search itself over dynamical systems is somewhat less guided. If we can find a class of models that exhibit the right behavior, we can then analyze the specific dynamics of this network to understand how it works, inspired by previous analysis of a recurrent neural network [4].

To summarize, the goal of this project is to find a connectionist model capable of representing an arbitrary distribution from only the network dynamics. The majority of the work will be on this search process in tweaking the model design and parameters, defining an efficient search algorithm, and finding accurate measures of its performance. If successful, we can compare its behavior to existing inference methods to determine whether it is effective or innovative.

[1] Hryceij, T. (1990) Gibbs sampling in Bayesian networks. *Artificial Intelligence*, 46, 351-63.

[2] McClelland, J. L. (1998). Connectionist models and Bayesian inference. In M. Oaksford & N. Chater (Eds.), *Rational Models of Cognition.* Oxford: Oxford University Press. 21-53. Relates computational and implementational levels of analysis.

[3] Kording, KP, Beierholm, U., Ma, W., Quartz, S., Tenenbaum, J., Shams, L., (2007) Causal Inference in Cue Combination, *PLOSOne* 2(9): e943.

[4] Rodriguez, P., & Elman, J. (1999). Watching the transients: viewing a simple recurrent network as a limited counter. *Behaviormetrika*, 26, 51-74.