An Alternative Method of Problem Solving: The Goal-Induced Attractor

William B Levy (wbl@virginia.edu)
University of Virginia, Department of Neurosurgery, P.O. Box 800420
Charlottesville, VA 22908-0420 USA

Xiangbao Wu (xw3f@virginia.edu)
University of Virginia, Department of Neurosurgery, P.O. Box 800420
Charlottesville, VA 22908-0420 USA

One theory of problem solving posits solutions by search. That is, the generic problem has a starting point and a goal, where the goal might be precisely known or only sketchily describable. Successful problem solving entails finding a legal (i.e., biologically consistent) path from the starting state to the, or to an, acceptable goal. Influential theories of problem solving, including Newell's work, emphasize the importance of search. That is, the strategy is to try out various paths in hopes that one will lead to the goal. Such searching techniques are computationally intractable in many situations and, in our day-to-day life, we often consider a problem and then find the answer. That is, we find the logical path to the goal without much thought at all and certainly without being consciously aware of trying multiple paths. Here we present a neural network model that solves paradigmatic cognitive problems without search. The alternative to the search strategy in a recurrent neural network is the use of an attractor. An attractor affects the states of a network, and the states of a network are its representations of the world itself. When a network is designed as a sequence-learning network and when it has enough freedom to create its own solutions, i.e., to create novel paths through state space, such a network can find paths from an initial state to a goal state and can find such paths where a path has never before been experienced (Levy, 1996).

The principle of the goal-induced attractor requires or assumes that the system solving the problem has a vague knowledge of the solution. For instance, if I am hungry, I might know I want something to eat, but I might not know exactly where I want to go to eat. We propose that such notions of the goal weakly turn on certain representations. These representations heighten the probability that paths to that goal will be discovered. At the same time, because networks have activity control mechanisms, there will be a tendency not to explore or move towards other goal states. Of course, if the network is not to depend on total randomness, there must be a history of paths learned by the network that can in some way be pieced together by network dynamics.

We use a model of hippocampal region CA3 because this is a sequence-learning region that is capable of coding novel sequences. In particular, and in contrast to error correction-based models, our model is used when mammals do not know the answer and must recode the environment in order to produce simple, usable codes by other brain regions or, from our point of view, by other networks. The problems solved by the CA3 model using the goal-induced attractor are not unlimited but include the simple goal finding problem that is analogous to a rat or a human going from a starting point to a goal by piecing together small paths that have been previously learned but have taken the organism to other places. Other cognitive problems, and some that may even appear to be logical in Nature, can be cast in terms that the goal finding hippocampal model can solve. For instance, the task of transitive inference can be taught to rats, and it can be taught to people in a nonverbal mode. The hippocampal model solves this problem, and it solves the problem, in a sense, by wanting to get the right answer. That is, the goal in performing transitive inference is to get the right answer as opposed to the wrong answer. In this case, the network would have a crude version of the reinforcement "yes, you're right" turned on while it is being presented with the stimuli of the transitive inference task. The task itself is viewed as a sequence but just barely. In particular, the sequence is stimulus, decision/response followed by knowledge of whether the outcome is a success or failure (right or wrong). The model is able to make the right decision for novel, transitive pairings. Another problem that can be solved in a similar manner is the transverse patterning problem.

Our poster will discuss the critical characteristics of our CA3 model that lead to its problem-solving abilities.

References

Levy, W.B (1996). A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. Hippocampus, 6, 579-590.