Abstract

Natural organisms can memorize and process sequential information over long time lags. Chimpanzees and orangutans can recall events which occurred more than a year ago (Martin-Ordas et. al, 2013). Long term social memory can provide significant survival benefits. For example, bottlenose dolphins can recognize each others whistle sounds even after decades (Bruck, 2013). Such a capability allows the dolphin to identify adversaries as well as potential teammates for hunting. First step towards adaptive behavior is to memorize past events and utilize them for future decision making (Stanley et. al, 2003). Memory is a key cognitive component and incorporating this capability in artificial agents can make them more realistic.

New methods are presented in this paper that can evolve sequence processing networks to solve reinforcement-learning (RL) memory tasks with long time dependencies. Tasks requiring memory can be formally be described as POMDP problems. Traditionally, recurrent neural network (RNN) has been the preferred choice for this purpose. However, RNNs leak information and are unable to discover long term dependencies. Long Short Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) has been found to be successful in overcoming the limitations of RNN. It consists of memory cells with linear activations. The inflow and outflow of information to and from these cells is controlled by associated input/output gated units. While LSTM networks have been used to achieve strong results in the supervised sequence learning problems such as in speech recognition and machine translation, applying them to solve POMDP tasks has resulted in limited success (Bayer et. al, 2009; Bakker et. al, 2003). This is probably because it is difficult to train LSTM units (including its associated control logic) with weak reward/fitness signal. Also, the number of LSTM units in a network is a parameter that is often manually selected. This approach turns out to be inefficient especially in new problems where the memory depth requirements are not clear.

In this work, NEAT (neuroevolution of augmenting topologies) (Stanley and Miikkulainen, 2004) algorithm is extended to incorporate LSTM cells (NEAT-LSTM). Since NEAT algorithm can evolve network topologies, it can discover the correct amount of memory units which are suitable for the task. NEAT-LSTM outperform RNN in two distinct memory tasks. However, NEAT-LSTM solutions do not scale as the memory requirement of the task increases. Evolving LSTMs presents two challenges: (1) the fitness landscape is deceptive (2) large number of associated parameters need to be optimized. To overcome these challenges, a new secondary optimization objective is introduced that maximizes the information (Info-max) stored in the LSTM network. The network training is split into two phases. In the first phase (unsupervised phase), independent memory modules are evolved by optimizing for the info-max objective. In the second phase, the networks are trained by optimizing the task fitness. Results on two different memory tasks indicate that neuroevolution can discover powerful LSTM-based memory solution that outperform traditional RNNs.

The killer application of this work is to build a computational model of memory-based biological behaviors. In particular, this model will explain how a group of hyenas, during lion-hyena interactions, modulate their behavior over a period of time through memory-based emotions -transitioning from being fearful initially to becoming risk-taking later. Understanding the underpinnings of such hyena behavior will provide insights into building realistic artificial agents.

References


