Reinforcement Learning for Temporal Patterns
using a Discrete Hierarchical Self-Organizing Map

Georgios F. Pierris and Torbjørn S. Dahl

Cognitive Robotics Research Centre, University of Wales, Newport, UK

Introduction

Reinforcement Learning Self-Organizing Maps (RLSOM) is a novel reinforcement learning approach based on a hierarchical structure, whose building blocks are augmented Kohonen Self-Organizing Maps. RLSOM reduces the input space and further encodes it in sequences, producing increasingly abstract representations towards the top of the hierarchy. Complex trajectories which even include perceptual aliasing can be learned and later reproduced using varying policy.

Motivation

Reinforcement learning algorithms have shown remarkable results in static, discrete environments, e.g., agents playing chess. However, complex problems, e.g., humanoid robot control, lie in high dimensional continuous state-action environments making a large number of classic RL algorithms ineffective.

RLSOM

RLSOM is based on a hierarchical structure, whose building blocks are augmented Kohonen Self-Organizing Maps. A SOM is a type of artificial neural network that reduces and discretizes the input space. Tracking temporal information using a single map is achieved through the introduction of decaying activation levels.

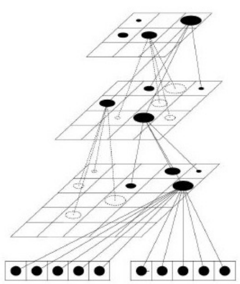

Figure 1: *A Hierarchy of augmented Kohonen Self-Organizing Maps*

The number of steps to consider a node inactive is the Short Term Memory (STM). The activation pattern is compiled to a sparse signal of zeros (no activation) and non-zero activations equal to the size of the STM. In RLSOM, the bottom layer serves as state and action space reduction algorithm whereas the higher layers encode activation patterns producing increasingly abstract representations towards the top of the hierarchy. Winning nodes are not found in the conventional way (input match), but also considers the match between STM, encoded in the node activations and long term memory, encoded in the connection weights, and the discounted future reward.

Experiments and Results

In a rather complex experiment, only a part of the Cornu Spiral is used to artificially produce a loop, i.e., a hidden state.

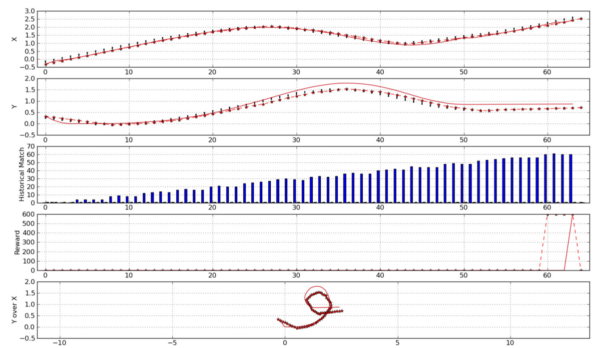

Figure 2: *The red line is the original data and the stars represent the reproduction. X, and Y, are time-series, and the bottom figure shows the Y over X to highlight the hidden state. The blue bars indicate the historical match.*

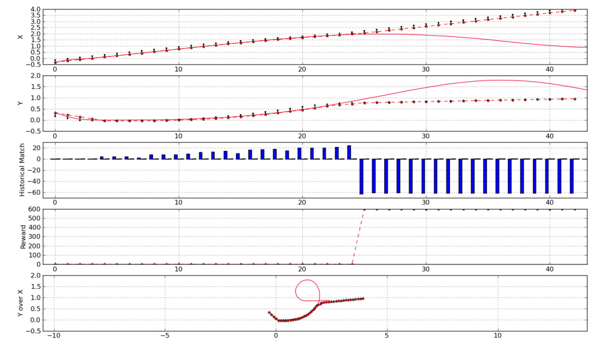

Figure 3: *Note the rapid change in the policy at step 25. The Historical Match drops and RLSOM follows the shorter path to the reward.*

In Figure 2, RLSOM better fits past experiences, thus the loop is followed. In Figure 3, a balance between history and reward, suggests using the node with the hidden state within a different context following the shorter path towards the goal.

Conclusions

- RLSOM discretizes the input space and successfully encodes temporal sequences in a hierarchical SOM.
- Problems that include hidden states can be learned and reproduced under different contexts.

Acknowledgements: The research leading to these results has received funding from the European Commission's Seventh Framework Programme (FP7/2007-2013) under grant agreement ROBOSKIN ICT-FP7-231500.

References: L. B. Cohen, H. H. Chaput, and C. H. Cashon, "A constructivist model of infant cognition", *Cognitive Development*, vol. 17, no. 3-4, 2002

A. Shimada and R. Taniguchi, "Gesture recognition using sparse code of hierarchical som" in ICPR, 2008, pp. 14.