Deep Q-Learning With Recurrent Neural Networks

Clare Chen, Vincent Ying, Dillon Laird

Stanford, 2016

Abstract

Deep reinforcement learning models have proven to be successful at learning control policies image inputs. They have, however, struggled with learning policies that require longer term information. Recurrent neural network architectures have be used in tasks dealing with longer term dependencies between data points. We investigate these architectures to overcome the difficulties arising from learning policies with long term dependencies.

Introduction

- Recent advances in reinforcement Learning have led to human-level or greater performance on a wide variety of games (e.g. Atari 2600 Games).
- Deep Q-networks are limited in that they learn a mapping from a single previous state which constist of a small number of game screens.
- We explore the concept of a deep recurrent Q-network (DRQN), a combination of a recurrent neural network (RNN) and a deep Q-network (DQN)
- In addition to vanilla RNN architectures we also examine augmented RNN architectures such as attention RNNs.

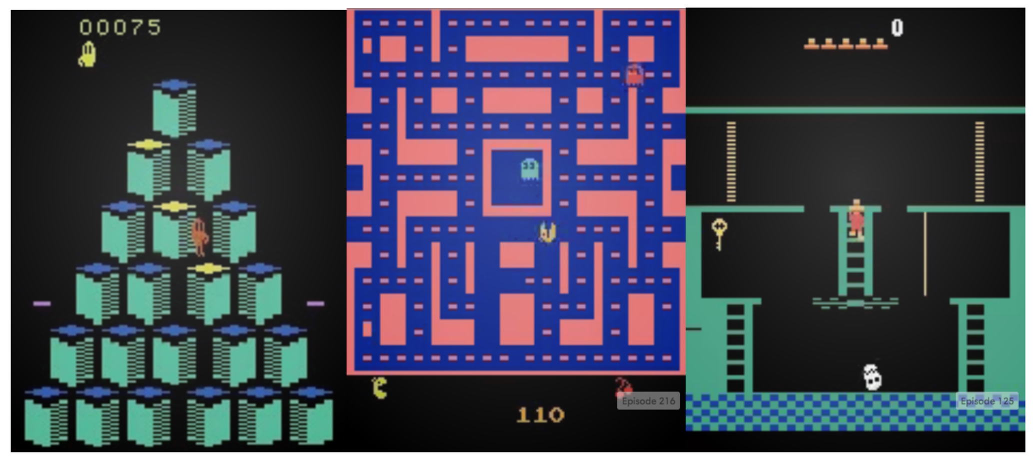

Figure 1: Q*bert, Ms. Pac-Man and Montezuma's Revenge

Background

Attention Deep Recurrent Q-Learning

Box

Text or image?

Deep Recurrent Q-Learning

- The architecture of DRQN augments DQN's fully connected layer with a LSTM.
- We accomplish this by looking at the last L states:

$$\{s_{t-(L-1)}, \ldots, s_t\}$$

- We feed these into a convultion neural network (CNN) to get intermediate outputs and finally send those through the RNN:

$$\text{CNN}(s_{t-i}) = x_{t-i}$$
$$\text{RNN}(x_{t-i}, h_{t-i-1}) = h_{t-i}$$

- The final output is used to predict the Q value.

Title

Title

References

Box

- item