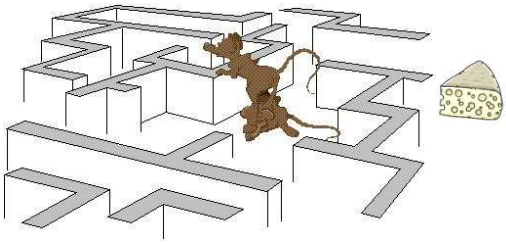
The Lorry I. Lokey Interdisciplinary Center for Life Sciences and Engineering
NETWORK BIOLOGY RESEARCH LABORATORIES

Evolutionary Algorithms in Reinforcement Learning

Introduction

Reinforcement Learning (RL) is a machine learning paradigm in which an agent tries to find the best 'behavior' in an unknown environment according to a reward that it receives for specific actions. A typical example is a mouse trying to find a piece of cheese in a maze:



Q-learning is a popular RL algorithm which maps an agent's 'best' action to each state of the world. When applying this algorithm to real world problems, the number of possible environment states is usually very large, requiring a more compact representation. A recently proposed method, NEAT+Q, uses a neural network to represent the environment, and an evolutionary algorithm to find the best network for the problem.

Aside from giving good results in real world situations, this method has specific features which make it possible to 'reverse engineer' the agents, and gain some insight as to how they behave – a question that is often left as a mystery.

Prerequisites

- 046195 - Machine learning (Recommended)
- Programming experience in Matlab/C++

Project Description

This project has 2 stages:

In the first stage the students will implement the NEAT+Q algorithm, and use it to train an agent in some simple environment.

In the second stage, which is more research oriented, the students will develop general methods to reverse engineer the trained agent, based on information from the training process. The question of how to reverse engineer an evolved system is still an open question, and is important in fields such as neuroscience and biology.

Project Duration: One semester, with an optional extension

Contact

Aviv Tamar – 04-8295084 – avivt@tx.technion.ac.il