Extending Reinforcement Learning to Incorporate Higher Order Learning

Presented By: Ashley Edwards
Mentor: Prof. William Pottenger
DIMACS REU 2010

Introduction

- **Reinforcement Learning** is a machine learning paradigm where an agent interacts with its environment and modifies its actions to maximize its reward. RL can be applied to many different domains.

Introduction

- **Reinforcement Learning** is a machine learning paradigm where an agent interacts with its environment and modifies its actions to maximize its reward. RL can be applied to many different domains.

- **Higher Order Learning**, specifically Higher Order Naives Bayes, is a classification mechanism that leveragess relationships between attributes found in *different* data instances of the same class. The data can come in many different forms – from E-Commerce to Nuclear Detection!

Related Work

- **Link Mining**

Related Work

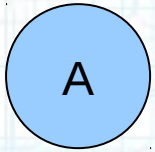
- **Link Mining**
  - Leverages the attributes and explicit links between instances

Related Work

- Link Mining
- **High-Order Paths**

Related Work

- Link Mining
- **High-Order Paths**



Related Work

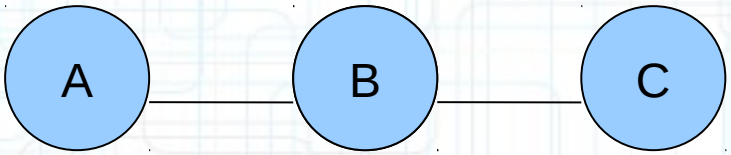
- Link Mining
- **High-Order Paths**

Related Work

- Link Mining
- **High-Order Paths**

Related Work

- Link Mining
- **High-Order Paths**



Related Work

- Link Mining
- High-Order Paths
- **LSI Algorithm**

Related Work

- Link Mining
- High-Order Paths
- **LSI Algorithm**
  - Uses higher order information within data, rather than assuming independence between instances

Related Work

- **Q-Learning**

Related Work

- **Q-Learning**
  - Learns an action value function which returns a reward for taking an action in a particular state

Related Work

- **Q-Learning**
  - Learns an action value function which returns a reward for taking an action in a particular state
  - Beneficial because a prior model of the environment is not needed

Related Work

- Q-Learning
- **Least-Squares Policy Iteration**

Related Work

- Q-Learning
- **Least-Squares Policy Iteration**
  - A batch RL algorithm based on sample experiences

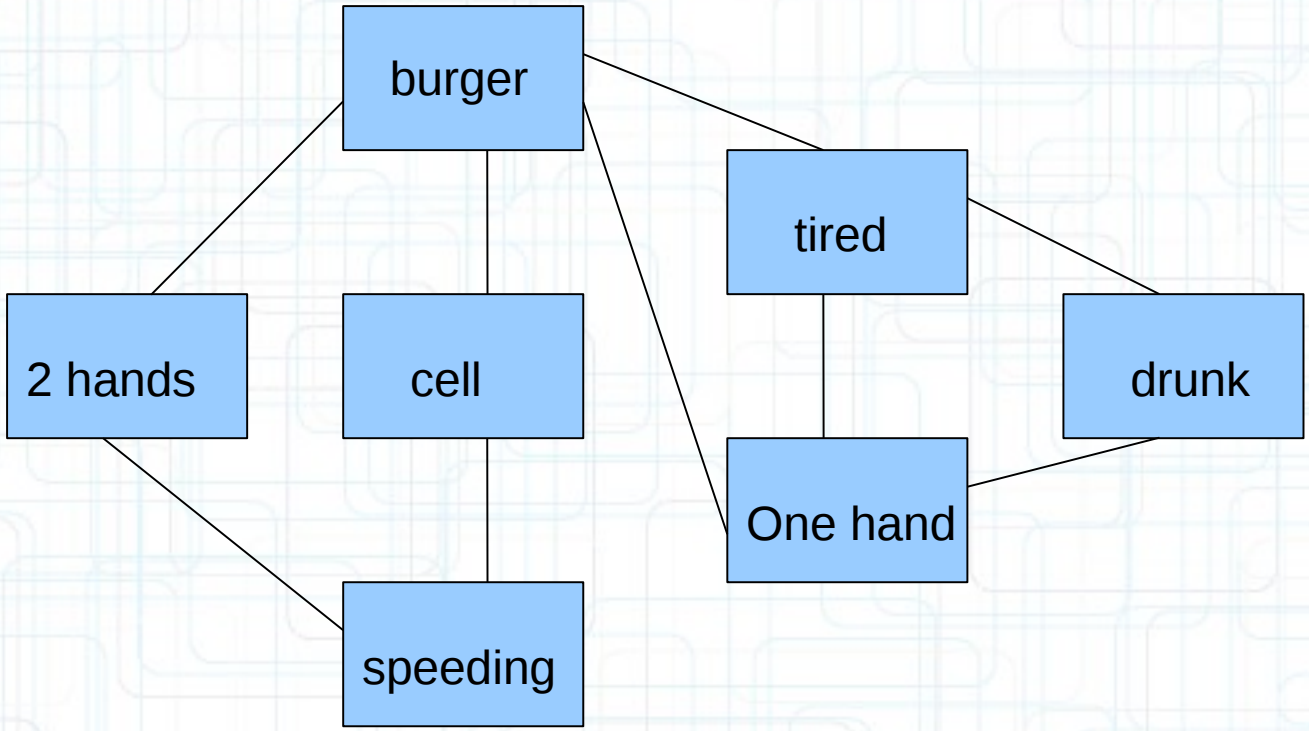
Related Work

- Q-Learning
- **Least-Squares Policy Iteration**
  - A batch RL algorithm based on sample experiences
  - A policy is iteratively approximated until it converges, then this policy is used for action selection

Approach

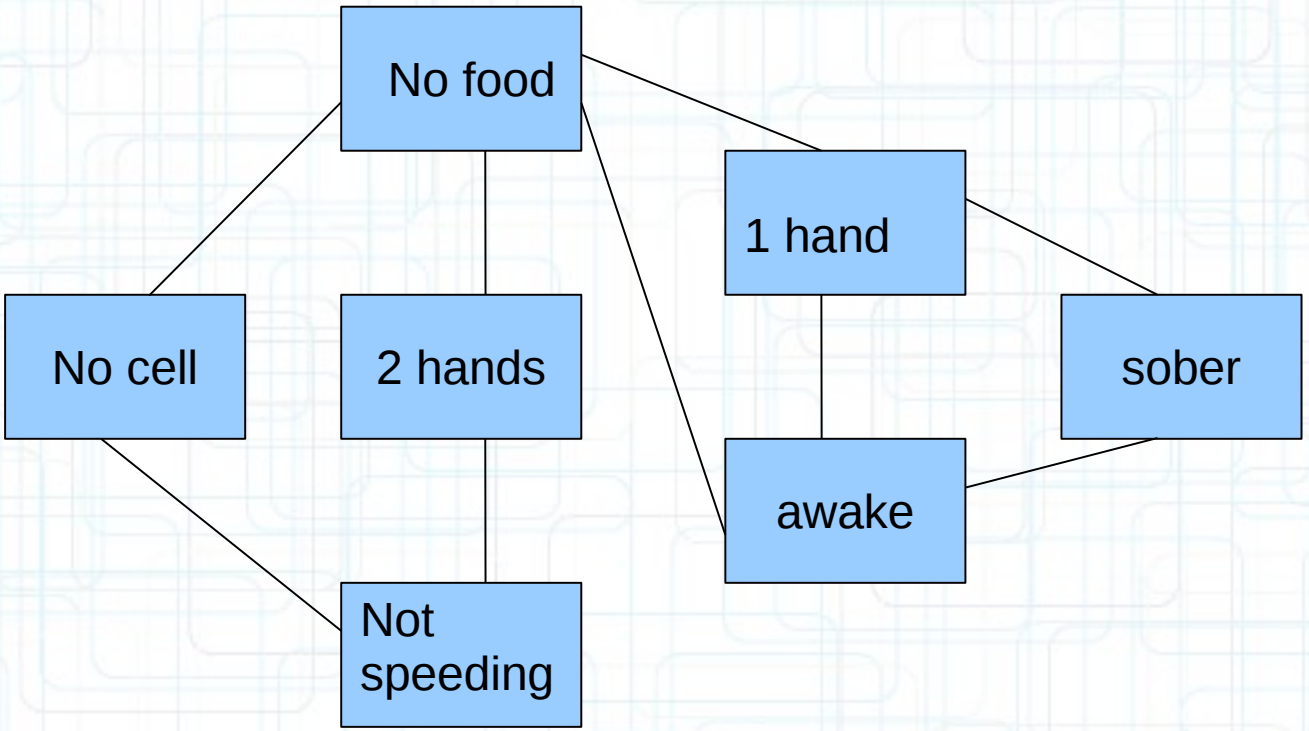
- As mentioned before, Higher Order Learning has been shown to work for many different types of data – so we will use it to classify different environmental states found in RL.
- This method can be used for action selection. If an action is classified as leading to a bad reward, a different action will be selected.

Approach



Class Low Reward

Approach



Class High Reward

Approach

- We will test our algorithm on the traditional domains used in RL.
- We will also test our algorithm on some non-traditional RL domains, such as nuclear detection

Conclusion

- Overall, we are aiming to modify Reinforcement Learning to utilize the rich features of higher-order paths found in Higher Order Naive Bayes
- We will then examine the modified algorithm's performance on Nuclear Detection Data

Questions?