**Authors:**

**Tianhao Fang**

**Mitt Shuby**

**Describe the algorithm you developed for Q-learning training**

The algorithm we use is using liner function approximation in Q-learning. We do Q-Learning on the path selector in the previous version of AI in Assignment 3. In Assignment 3, we only use the path selector in a fix method. However in this assignment we use a policy based liner function approximation in Q-learning.

**Describe the states, actions and rewards used in Q-learning. If you are using approximate Q-learning, please specify what features constitute the states.**

In our implementation, we use approximate Q-learning, the action is actual select a path, the state for Q-learning is just the path static, like power pills on road, pills on road, ghost on road and so on, and they served as features.

**What is the policy of Ms. Pac Man in the training?**

The policy of Ms.PAC MAN is choose path between many paths, it is using heuristic function for the path and choose the path with highest heuristic.

**What is the format for model storage?**

The format for model storage is just using floats convert into string with delimiter blank space. And is stored in myData/result.txt.