Final Project -Gomoku

18300180075 Chen Xiaohan, 19340980040 Lai Weipo

Including Algorithms:

- Alpha-beta pruning

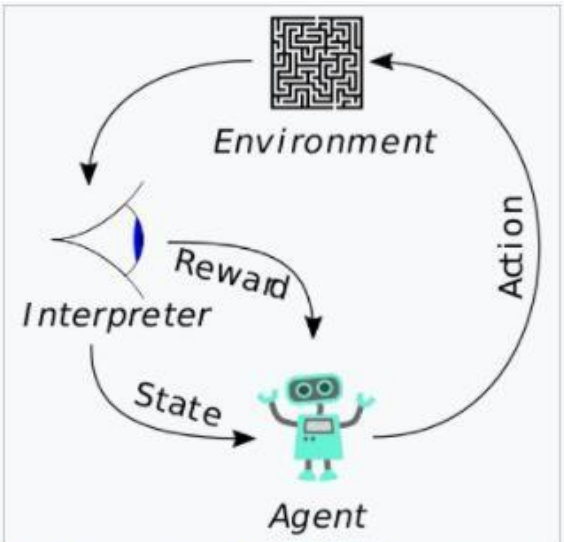
  Using MinMax implemented by Alpha-Beta Pruning, the AI searches the next move which gives itself the max score and gives its opponent the least score(if possible, reducing the opponent’s score.

- ADP (Reinforcement Learning):

  Reinforcement learning is the third fundamental machine learning method in addition to supervised and unsupervised learning. In machine learning problems, the environment is usually abstracted as a Markovian decision process.

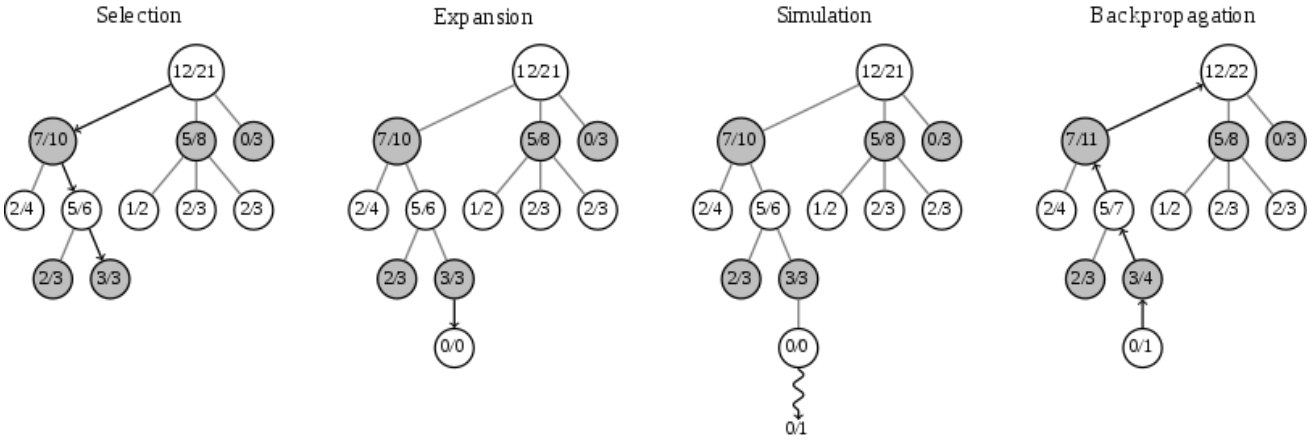
- The basic reinforcement learning is modeled as a Markov decision process:

  1. The set of environmental states S;
  2. The set of actions A;
  3. Rules for transitions between states (transfer probability matrix) P. 4. Rules specifying the "immediate reward" after a transition (reward function) R.



强化学习的典型框架：智能体在环境中采取一种行为，环境将其转换为一次回报和一种状态表示，随后反馈给智能体。

- The MCTS(Monte-Carlo Tree) algorithm includes the following 4 steps.
  1. Tree traversal.
  2. Node expansion.
  3. Rollout (random simulation).
  4. Back propagation.



In our code, a fatal error happened when using MCTS, causing the agent to not respond to the program. In order to solve this, we took a different approach, using Minmax to start our first step and replacing it with MCTS after the agent responded.