CSCI 3202
Lecture 28
October 31, 2025

MOTHER GOOSE & GRIMM: By Mike Peters

Mother Goose and Grimm by Mike Peters. https://www.grimmy.com/

Announcements
- Quiz #9 Today
  - Policy and value iteration
  - Active learning
- HW #3 returned on Monday
- Project Intermediate Report due Wednesday
- Homework 7 due Friday, November 7 by 11:59 pm

Project
- CSCI 3202 Course Project 2025.pdf
- Intermediate report due Wednesday
  - Turn in a Jupyter notebook or Zip file
  - Include your code
  - Write a paragraph describing your results so far
  - For 100 iterations of Random vs. Random
    - Show the percentage of wins for Player 1 vs. Player 2 and the average number of moves (both players) per game
    - We expect to see a small first player advantage
  - No need to save state for Random vs. Random
  - Should run in a few seconds for 100 iterations

- For final report, add AI player that can do Minimax and AlphaBeta at 5 and 10 plies
  - Will need to save state for AI player
  - Depending on how you save state, the time may be from 10 minutes for 5 plies to overnight
- See TAs, CAs or me with questions

Policy Iteration
- A good comparison of Value Iteration and Policy Iteration is in Geeks for Geeks https://www.geeksforgeeks.org/what-is-the-difference-between-value-iteration-and-policy-iteration/

Reinforcement Learning and Q-Learning
- Roomba and other Robot Vacuums
  - Typically vacuum a room at night when there is little traffic
  - Want to vacuum the whole room, not just the edges
  - Avoid obstacles and stairs
  - May map the room as they go
  - Vacuum until they run out of battery

- Assume that the Robot Vacuum uses a Markov Decision Process
  - What does policy mean?
  - How does it know when to stop?
  - What does optimality mean with a robot vacuum?
  - What could the reward function look like?
  - What happens when the contents of a room move?
  - What happens when someone plops something (like a coat or shoes) down in the middle of a room?
- Active vs. Passive
- Exploration vs. Exploitation
- Q-Learning doesn't require a model, but comes at a cost in terms of how many iterations it takes to create a stable utility function
- Policy Iteration Q Learning.pdf

Upcoming
- Hidden Markov Models