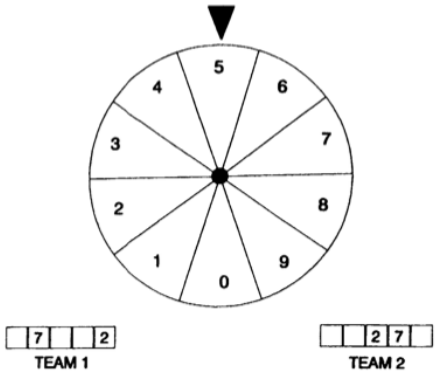
Computer Science 6804: Learning and Sequential Decision Making

Virginia Tech, Spring 2013
Instructor: Sanmay Das

So Who's Counting?



Complexities

- Future is unknown: which number will come up next? But what I choose now affects my future choices!
  - Sequential decision-making
- What if the wheel is biased? I don't know how often the different digits will show up
  - Reinforcement learning
- What is my opponent going to do?
  - Game theory

Poker

- Poker Strategy 1
- Poker Strategy 2
- Terrific testbed for AI: particularly sequential decision-making and game theory, but also learning (bet sizing, opponent strategies)

Sequential Decision-Making

- The non-sequential setting: you are in desperate need of $1000. I give you $5000 and allow you to bet any amount of it (at 1:1 odds) on a bet that you have 55% chance of winning. How much should you bet?
- The sequential setting: I give you $5000 to play Blackjack in a casino until the money runs out or you win at least $5000. How do you organize your betting?
- Not just about games: ambulance scheduling, search and rescue, logistics of disaster rescue, organ donation, all kinds of socially important issues!

What is Machine Learning?

- "Enabling computers to learn from data"
- Supervised learning: generalizing from seen data to unseen
- Unsupervised learning: Finding patterns in input data
- **Reinforcement learning**: learning how to act
  - Our focus

Where Does This Fit in AI?

- What is the goal of AI?
  - Two dichotomies: thinking/acting and humans/rationality (Russell & Norvig [RN])
- Turing test: acting like humans (restricted domain)
- Alternative goal: acting rationally?

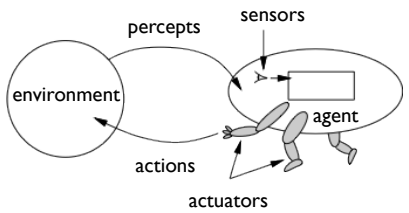
Rational Behavior

- Maximize "goal attainment"
- Easier when we specify utility functions
- Perfect rationality is problematic:
  - Computational limitations
  - Definition of optimality

Utility Theory

- Going from the real-world to sensible reward functions
- Utility theory is useful in abstracting over preferences
- Thought question: which of these two options would you prefer? $1000 with 50% prob., or $400 for sure?
- Basis for the insurance industry!

Intelligent Agents



- Agents include humans and artificial agents
- Agent function maps percept histories to actions $f : P* \rightarrow A$
- Agent program runs on the physical architecture to produce *f*

Human Beings and Humanoid Robots

- Environment: ?
- Sensors/Percepts: ?
- Actuators/Actions: ?
- Performance measure: ?

A Trading Agent

- Environment: ?
- Sensors: ?
- Percepts: ?
- Actuators: ?
- Actions: ?
- Performance measure: ?

Markov Decision Processes (Problems)

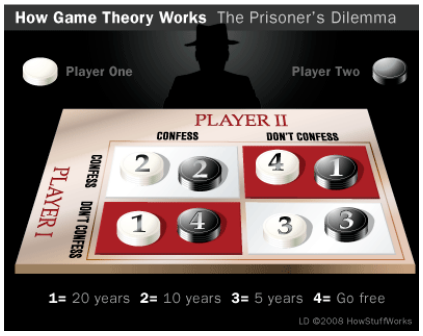
- Fundamental framework: agent interacts with the world and receives *rewards*
- State space: S
- Initial state: S_0
- Action space: A
- Transition model: $T : S \times A \rightarrow S \times [0, 1]$
- Reward function: $R : S \rightarrow \mathbb{R}$

Reinforcement Learning

- Unknown transition / reward models
- Must choose which actions to take
- There are methods for solving MDPs in many cases of interest. RL raises new major issues:
  - Delayed reward/credit assignment
  - Exploration/Exploitation

Game Theory

- Agents are assumed to be rational
- They have to reason rationally about the actions of other agents they interact with
- Game theory provides a foundation for studying rational behavior in multi-agent systems



Prerequisites

- Experience with math (reasonable level of sophistication). Comfort with notation.
- Calculus, matrix algebra.
- Probability!
- Ability to program in a mathematically flexible language (e.g. R, Matlab, Python) that enables rapid prototyping
  - Most projects will involve implementation!
- Need to force-add?
  - Come speak with me after class

Syllabus and Course Policies

- Come to class
- Participate online (Piazza) and in class
- One major project
- One presentation
- Check the website regularly!
- Let's go through the syllabus / website