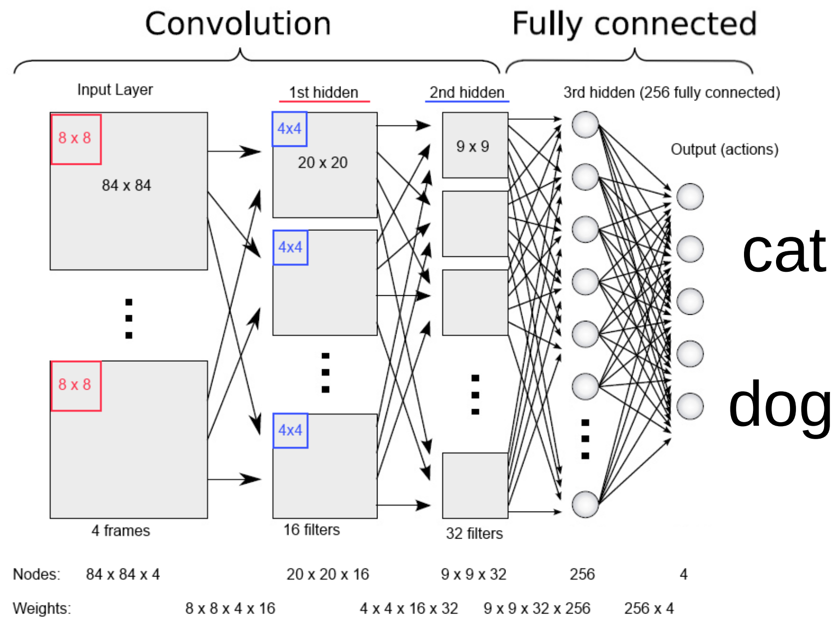
Course Logistics

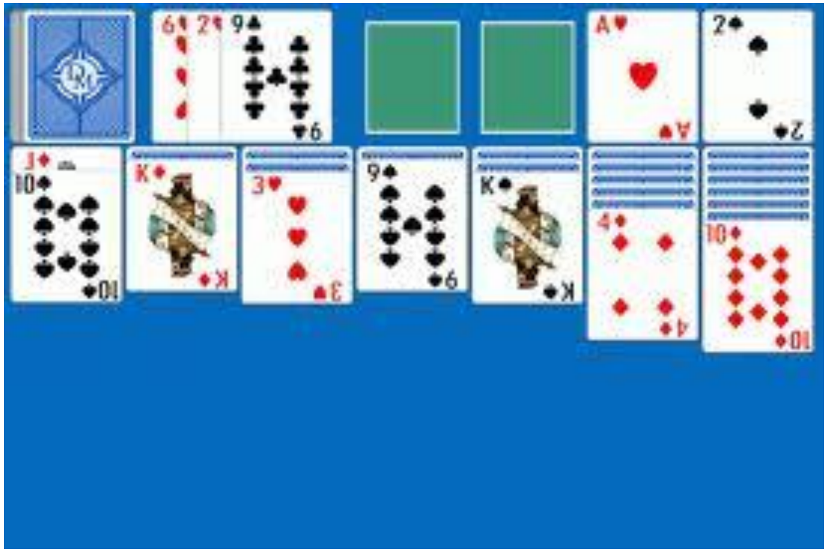
- CS533: Intelligent Agents and Decision Making
  - M, W, F: 1:00—1:50 (KEC1003)
  - Instructor: Alan Fern (KEC2071)
  - Office hours: by appointment
- Course Web Site:
  - Sign Up: http://classes.engr.oregonstate.edu/eecs/spring2016/cs533/
  - Will post lecture-schedule, notes, reading, and assignments
- Grading
  - 75% Instructor Assigned Projects (mostly implementation and evaluation)
  - 25% Student Selected Final Project (work in teams of 2-3)
- Assigned Projects (work alone)
  - Generally will be implementing and evaluating one or more algorithms
- Final Project (teams allowed)
  - Last month of class
  - You select a project related to course content

One Shot Decision Making



Each decision/classification can be made without considering future decisions making.

Sequential Decision Making

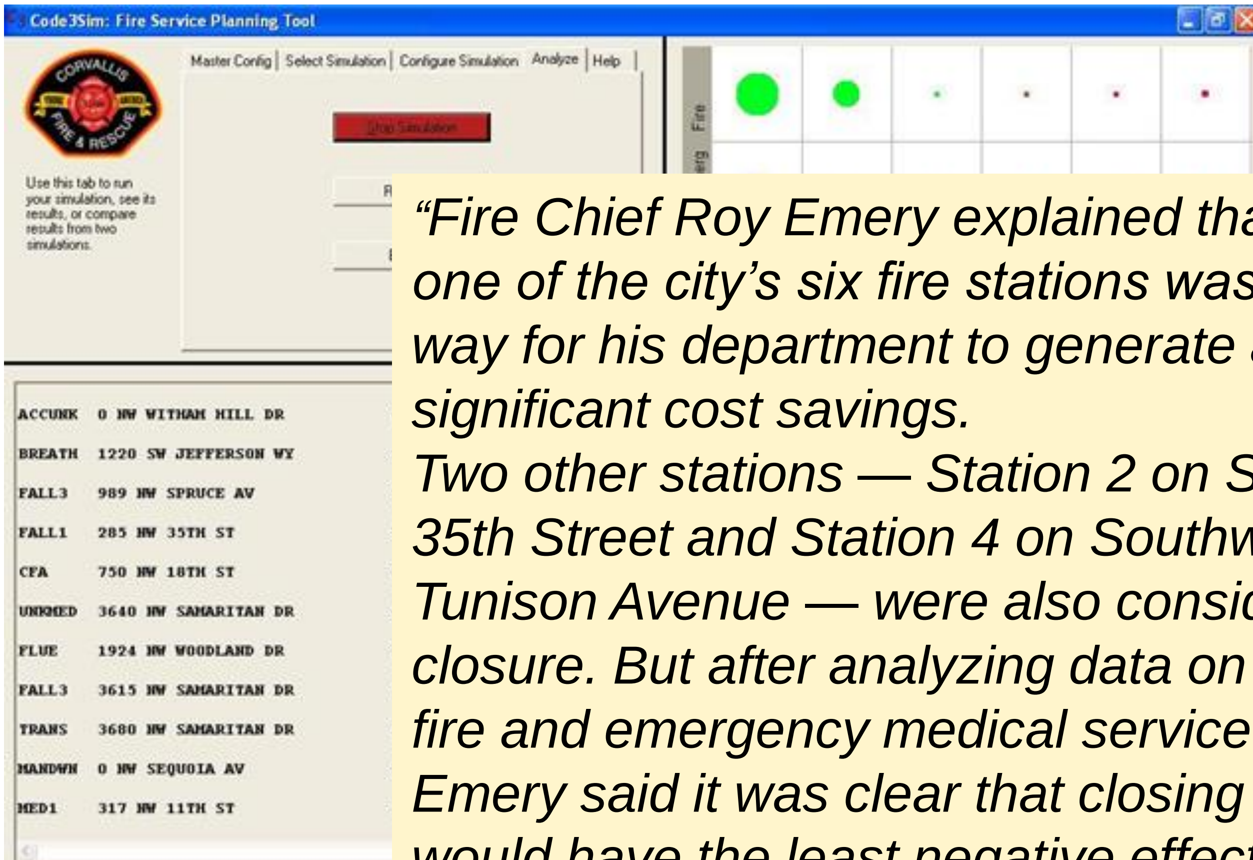


Klondike Solitaire

Real-Time Strategy Games

Sequential Decision Making

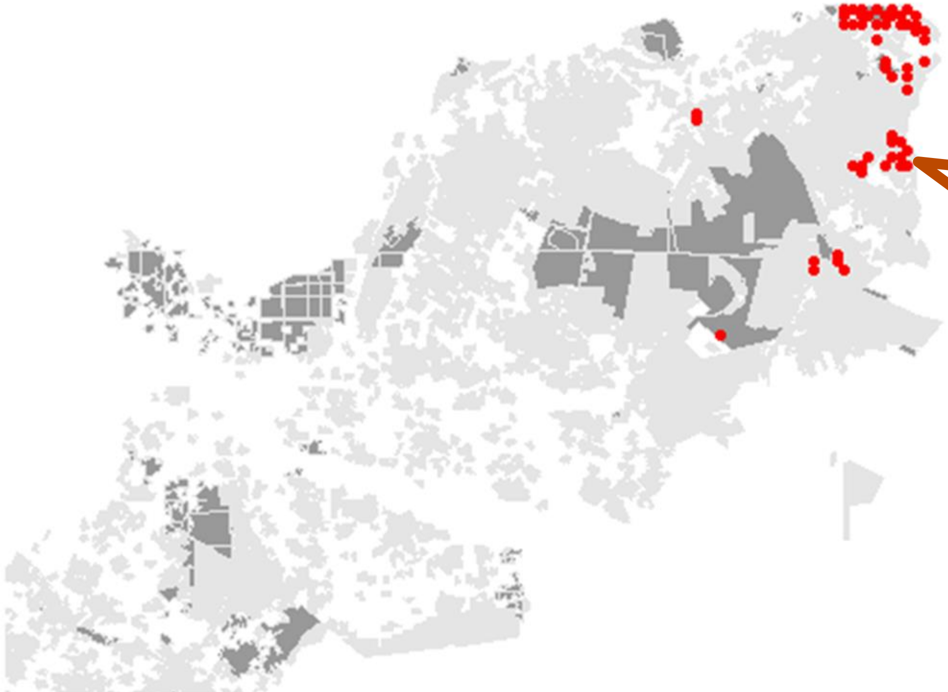
Optimizing Fire & Rescue Response Policies



"Fire Chief Roy Emery explained that closing one of the city's six fire stations was the only way for his department to generate any significant cost savings.
Two other stations — Station 2 on Southwest 35th Street and Station 4 on Southwest Tunison Avenue — were also considered for closure. But after analyzing data on calls for fire and emergency medical service citywide, Emery said it was clear that closing Station 5 would have the least negative effect."
Corvallis Gazette Times

Sequential Decision Making

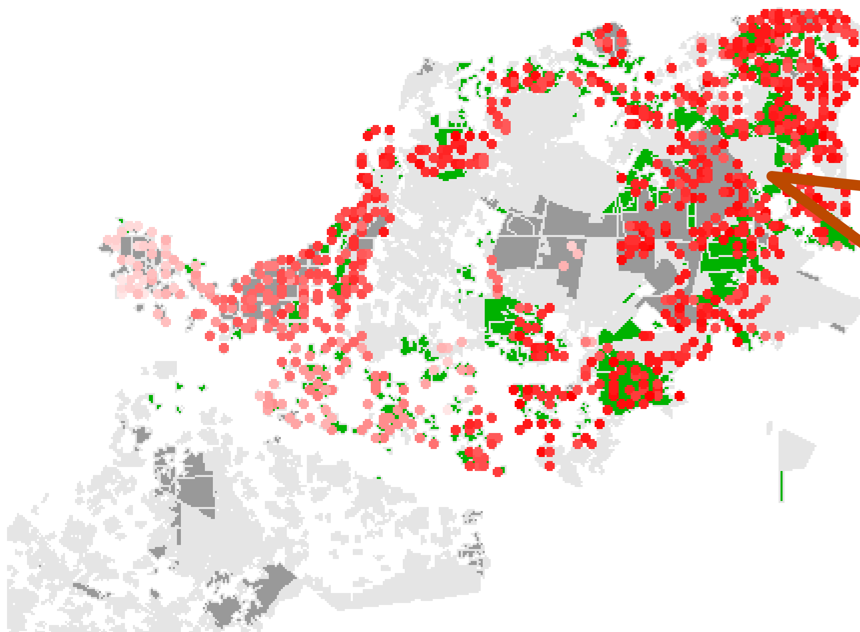
Conservation Planning:
Recovery of Red-cockaded Woodpecker



From http://www.fws.gov/
rcwrecovery/rcw.html

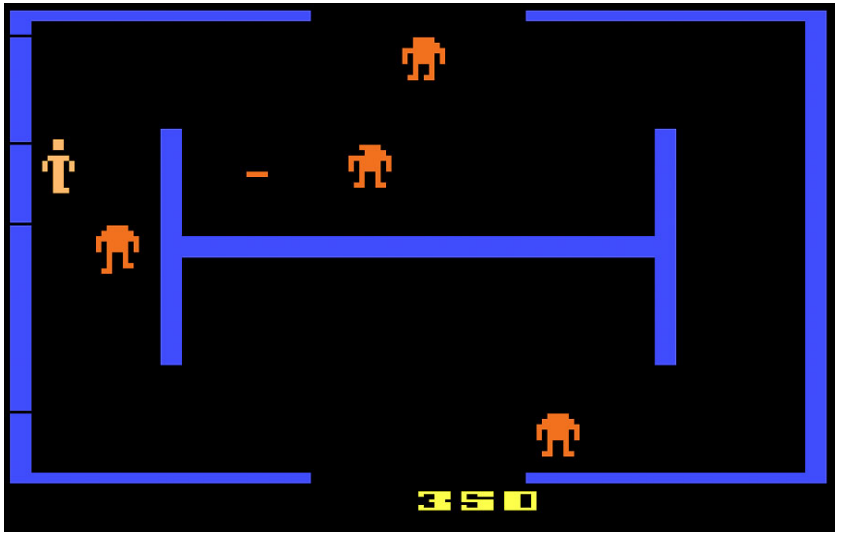
Sequential Decision Making

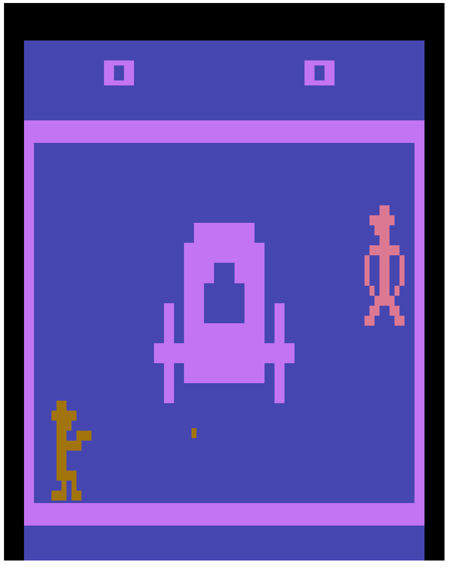
Conservation Planning:
Recovery of Red-cockaded Woodpecker

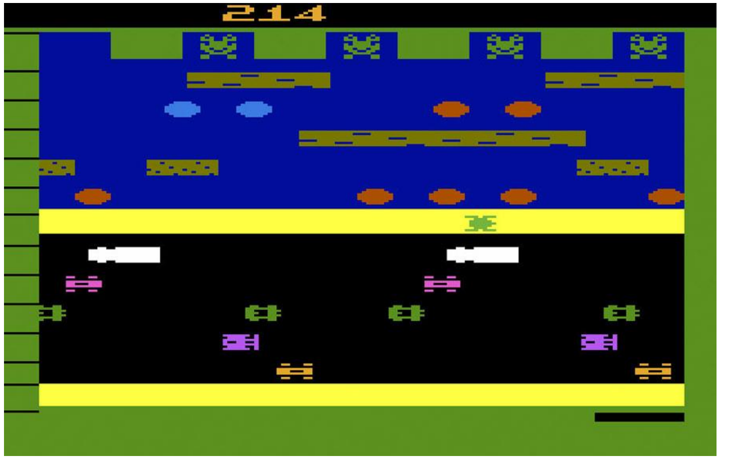


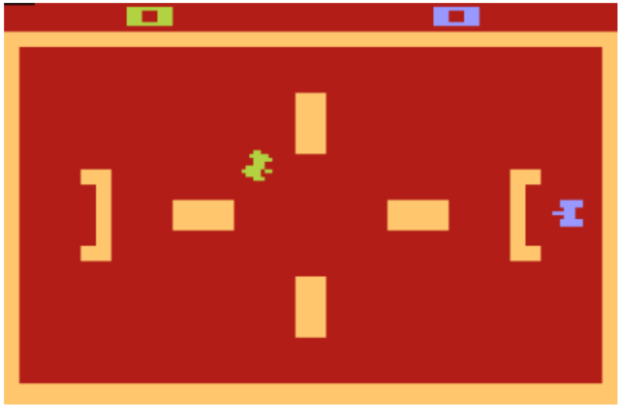
From http://www.fws.gov/
rcwrecovery/rcw.html

Sequential Decision Making





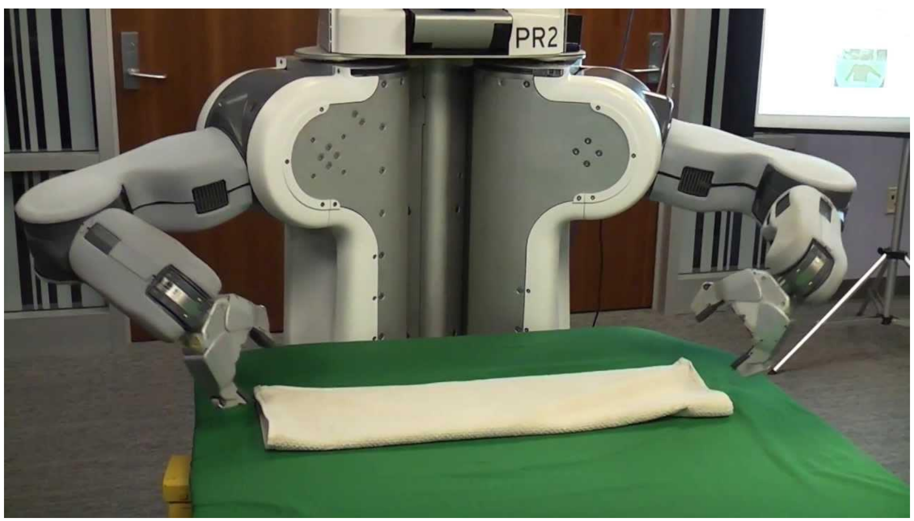




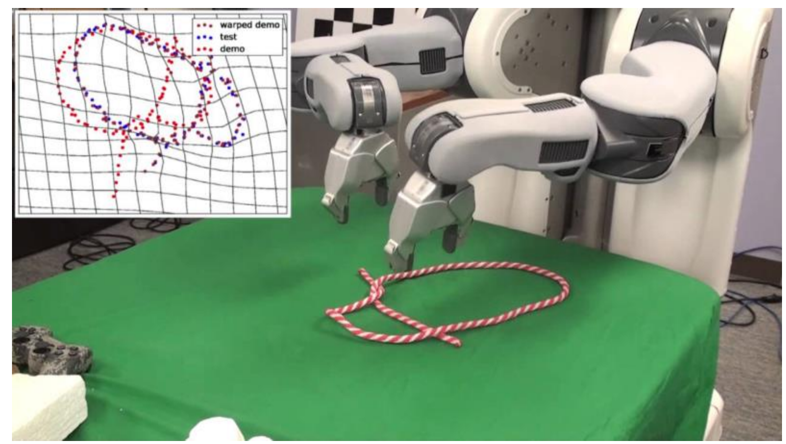
Sequential Decision Making

Helicopter Control

Legged Robot Control



Laundry



Knot Tying

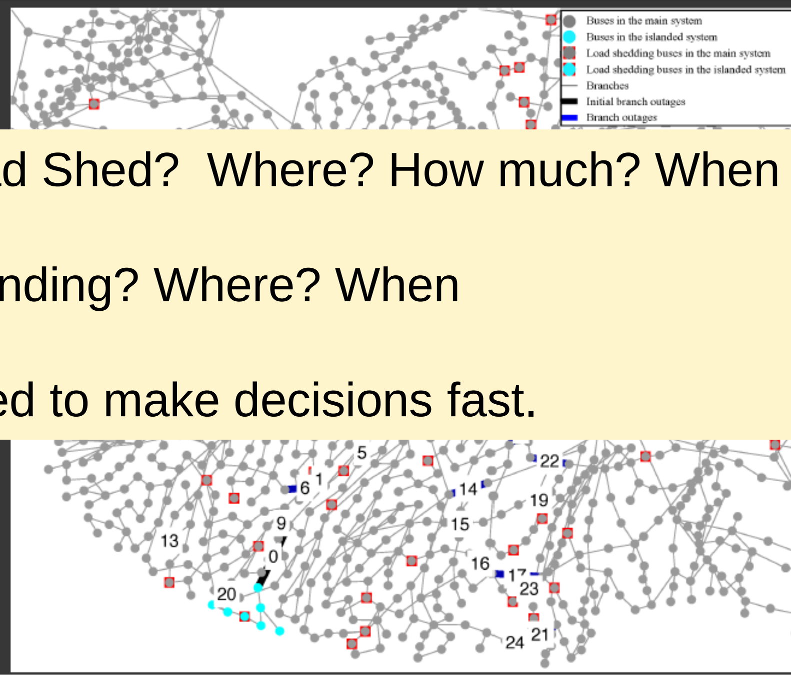
Intelligent Simulator Agents

Immersive real-time training

Intelligent Remedial Action Schemes for Power Grids w/ Eduardo Cottila-Sanchez

Goal: minimize impact of power system faults

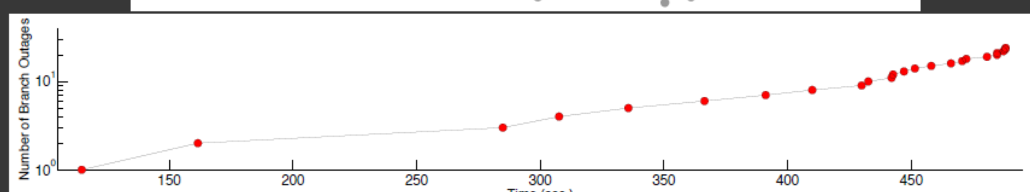
Cascading Paths



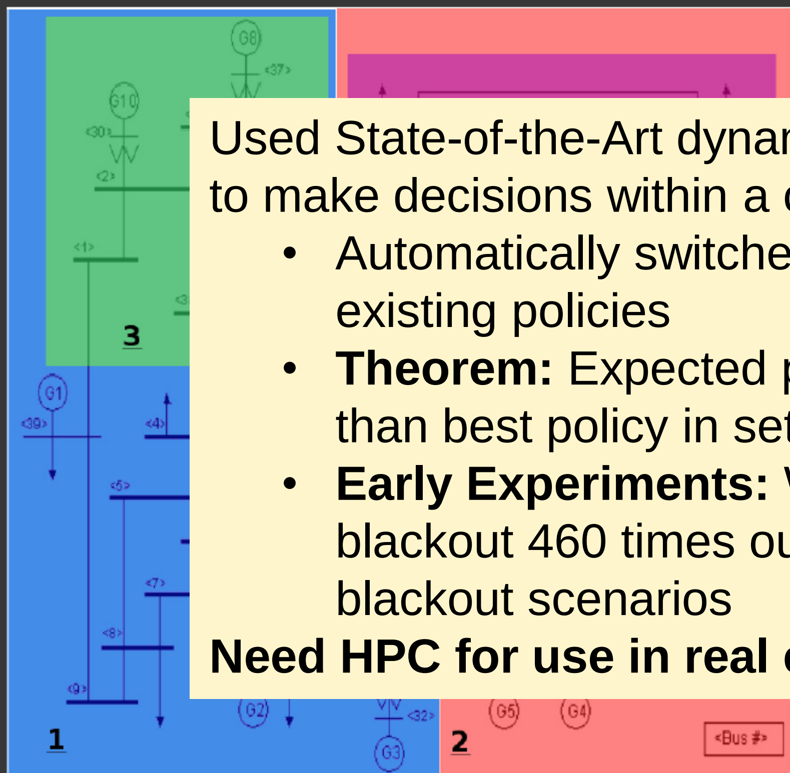
Load Shed? Where? How much? When

Islanding? Where? When

Need to make decisions fast.



Case Study: IEEE 39-Bus



Used State-of-the-Art dynamic power grid simulator to make decisions within a control loop.

- Automatically switches between library of existing policies
- **Theorem:** Expected performance is no worse than best policy in set for the current scenario.
- **Early Experiments:** Was able to prevent total blackout 460 times out of 600 simulated blackout scenarios

Need HPC for use in real environments!

AlphaGo

- Deep Learning + Monte Carlo Tree Search
- Learn from 30 million expert moves and self play
- Highly parallel search implementation
- 48 CPUs, 8 GPUs (scaling to 1,202 CPUs, 176 GPUs)

AlphaGo vs. Lee Sedol (9-dan pro w/ 18 world titles)

AlphaGo won 4 games to 1

https://deepmind.com/alpha-go.html

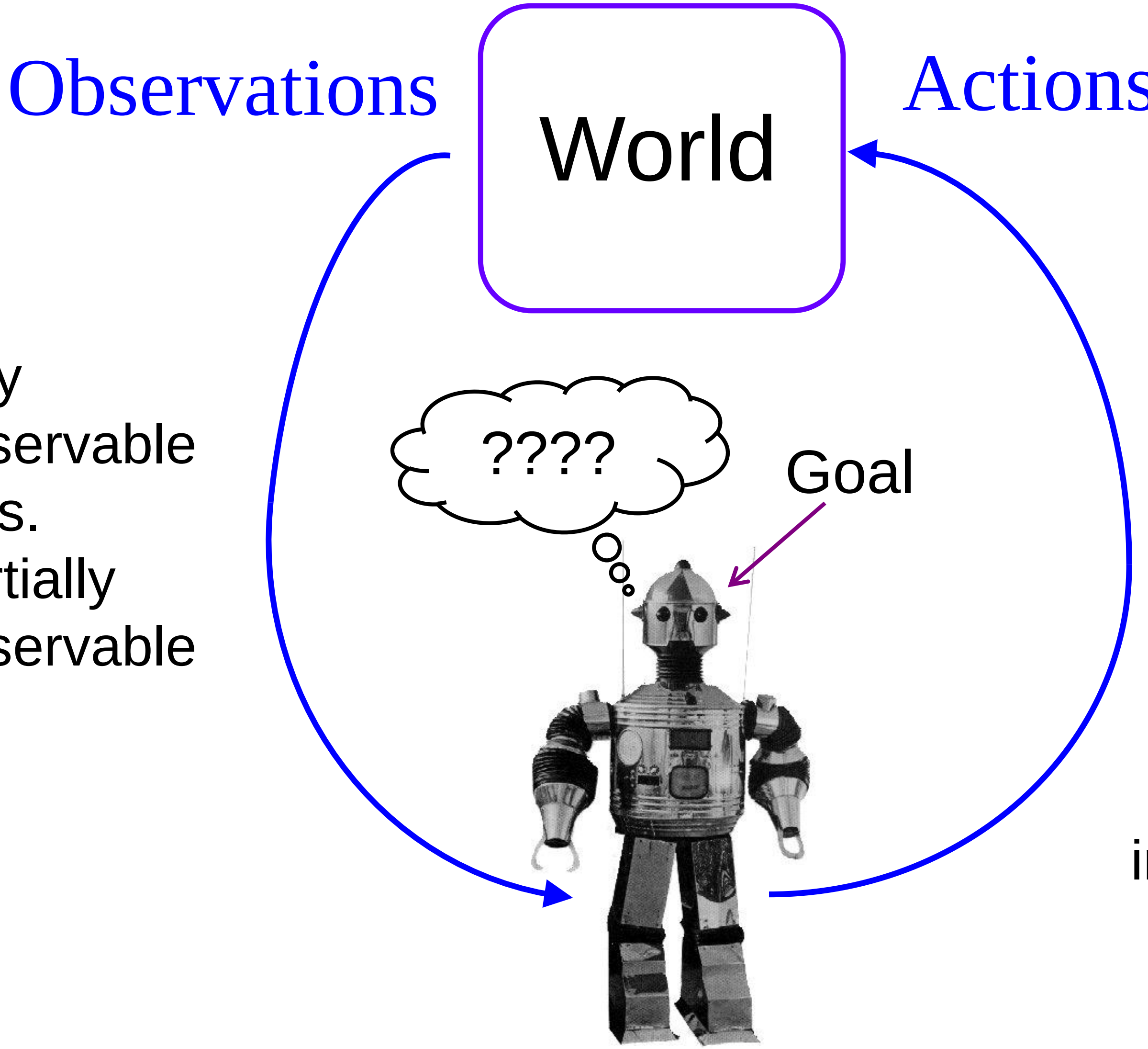
More Sequential Decision Problems

- Health Care
  - Personalized treatment planning
  - Hospital Logistics/Scheduling
- Transportation
  - Autonomous Vehicles
  - Supply Chain Logistics
  - Air traffic control
- Assistive Technologies
  - Dialog Management
  - Automated assistants for elderly/disabled
  - Household robots
  - Personal planner

Common Elements

- We have a controllable system that can change state over time (in some predictable way)
  - The state describes essential information about system (the visible card information in Solitaire)
- We have an objective that specifies which states, or state sequences, are more/less preferred
- Can (partially) control the system state transitions by taking actions
- **Problem:** At each moment must select an action to optimize the overall objective
  - Produce most preferred state sequences

Some Dimensions of AI Planning



fully
observable
vs.
partially
observable

sole source
of change
vs.
other sources

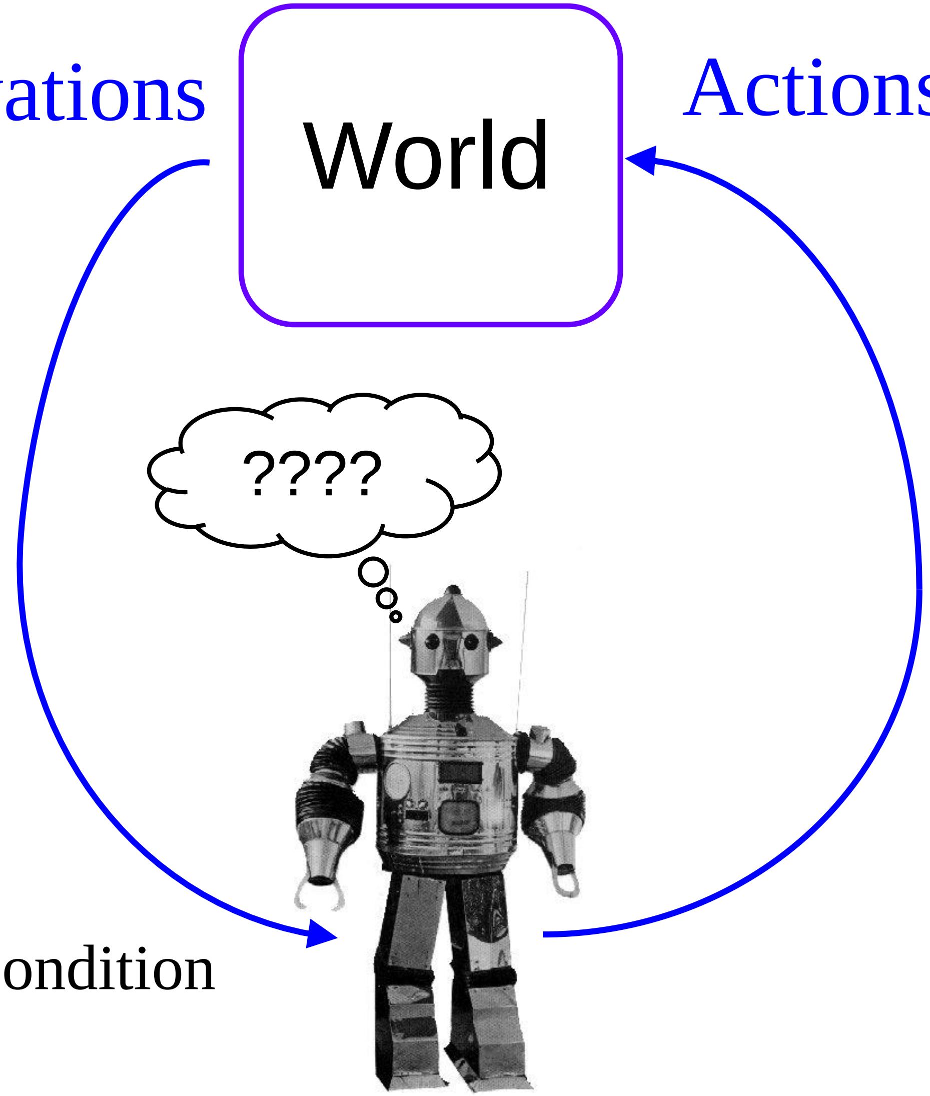
deterministic
vs.
stochastic

instantaneous
vs.
durative

Classical Planning Assumptions
(primary focus of AI planning until early 90's)

Observations

Actions



sole source
of change

fully
observable

deterministic

instantaneous

Goal

achieve goal condition

Classical Planning Assumptions
(primary focus of AI planning until early 90's)

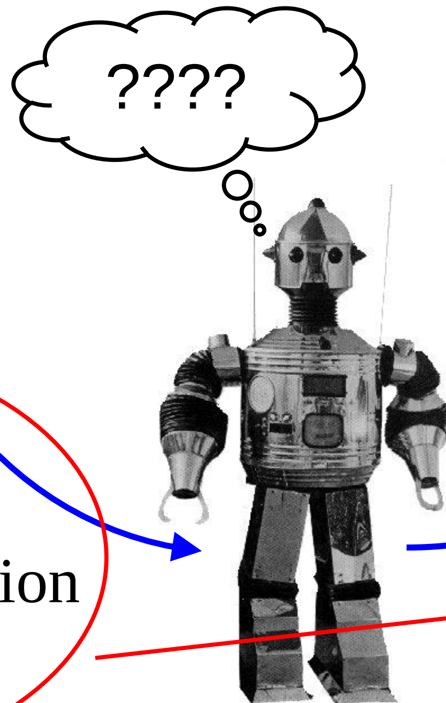
Observations

World

Actions

sole source
of change

fully
observable



deterministic

instantaneous

Goal
achieve goal condition

Greatly limits
applicability

Stochastic/Probabilistic Planning: Markov Decision Process (MDP) Model

Observations

World

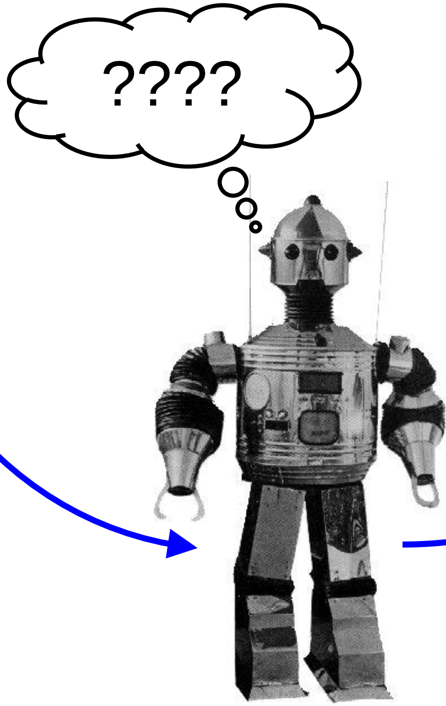
Actions

fully
observable

sole source
of change

stochastic

instantaneous



Goal

maximize expected
reward over lifetime

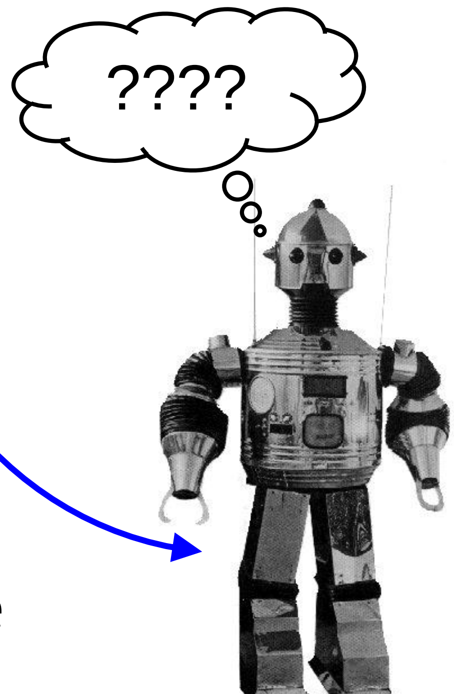
We will primarily
focus on MDPs

Stochastic/Probabilistic Planning: Markov Decision Process (MDP) Model

Probabilistic state transition (depends on action)

World State

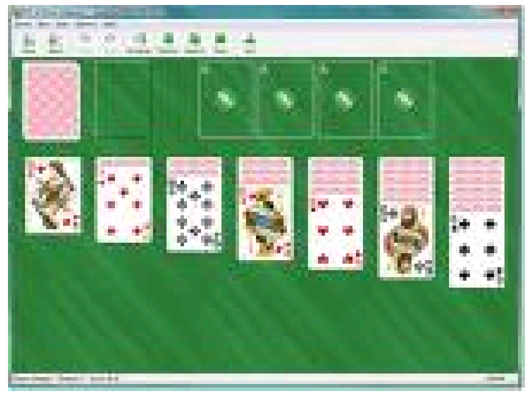
Action from finite set



Goal

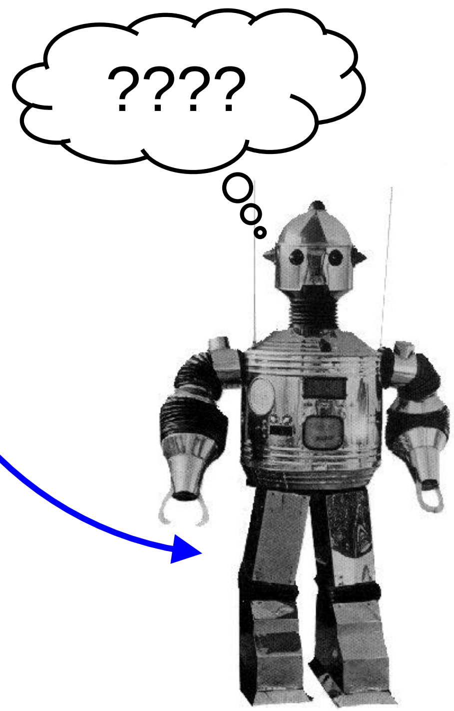
maximize expected reward over lifetime

Example MDP



State describes all visible info about cards

Action are the different legal card movements



Goal

win the game or
play max # of cards

Course Outline

Course is structured around algorithms for solving MDPs

- Different assumptions about knowledge of MDP model
- Different assumptions about prior knowledge of solution
- Different assumptions about how MDP is represented

1) Markov Decision Processes (MDPs) Basics
   - Basic definitions and solution techniques
   - Assume an exact MDP model is known
   - **Exact solutions** for **small/moderate** size MDPs

2) Monte-Carlo Planning
   - Assumes an MDP simulator is available
   - **Approximate solutions** for **large** MDPs

Course Outline

3) Reinforcement learning
   - MDP model is not known to agent
   - **Exact solutions** for **small/moderate** MDPs
   - **Approximate solutions** for **large** MDPs

4a or 4b) as time allows

- a) Planning w/ Symbolic Representations of Huge MDPs
  - Symbolic Dynamic Programming
  - Classical planning for deterministic problems
- b) Imitation Learning