Learning and Sequential Decision Making

CSCI 2951-F

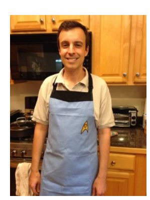
Ron Parr
Duke University Department of Computer Science
Visiting Brown University AY23-24

Who is Ron?

- Visiting faculty from Duke CS
- Worked in MDPs and reinforcement learning when it was cool the first time (1990's)
- Stuck with it when it wasn't cool until it became cool again
- (Nearly 30 years)
- Will stick with it if ever becomes uncool again


1990s


2020s


Uncool for 30 years

Who else is here to help?

Introduce David and Saket

Review class web page

What is this class about?

How an **agent** can **deduce or learn** to act **intelligently** in the presence of **uncertainty**.

Everything in **BOLD** is something we need to define more precisely to address this challenge.

Who should be interested?

- AI/ML/Robotics students
- Students who want to use AI/ML as a special sauce for making better decisions in some other area
- Econ/OR/Control students
- Psych/Neuro students, possibly to a lesser extent
  (Why? Normative vs. descriptive view of behavior)

Am I ready for this class?

- First: I'm sorry about the prerequisite issues

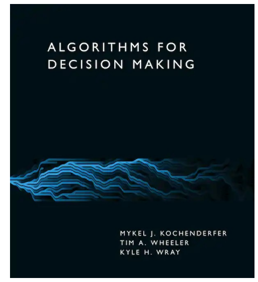
- What is needed really:
  - Basic stats, linear algebra and calculus
  - Mathematical maturity: Can do understand and produce short proofs/derivations?
  - Understanding of core computer science concepts such as computational complexity, dynamic programming, and analysis of algorithms
  - Python
  - Some previous exposure to machine learning and neural networks
  - Some exposure to AI concepts

What are the learning objectives?

RP hopes students will:

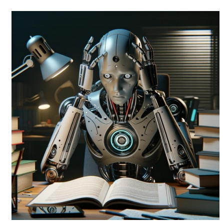
- Get a broad introduction to the theory that define optimal behavior under conditions of uncertainty
- Get a broad introduction to the algorithms that compute optimal decisions given this theoretical background
- (The above includes: Decision Theory, MDPs, Bandits, Reinforcement Learning (RL), POMDPs, and Game Theory.)
- Be able to implement RL algorithms
- Be able to understand modern RL papers

What is expected of students?



- Read the assigned textbook chapters
- Come to class(preferred) and/or review the slides
- Ask questions!
- Do ~4 homework assignments (50% of grade)
- Do a modest project with a small number of partners (last month of class)

Is this a hard class?



- It’s not intended to be!
- But we do cover a lot of material quickly at the beginning
  (Why? So, you will know enough to be able to formulate a project)
- Homework assignments:
  - Not IQ tests,
  - but designed to make you think and gain insight

- Don’t create unnecessary stress:
  - Keep up with the reading
  - Ask plenty of questions

A note about academic honesty

- It is OK to talk with classmates about homework assignments
- What you write up/code must be your own
- (Assignments that are blatantly copied will be flagged and reported.)

- It is ok to use Wikipedia or LLMs to ask basic reference questions (though you should be careful about uncritically accepting explanations from LLMs)
- It is **NOT OK** to turn in homework answers or projects that are found on the web or generated by LLMs

Defining things

What is this class about?

How an **agent** can **deduce or learn** to act **intelligently** in the presence of **uncertainty**.

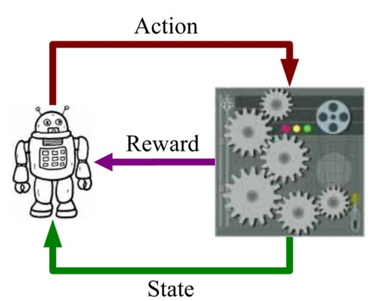
Everything in **BOLD** is something we need to define more precisely to address this challenge.

What is an agent?

- Anything/anybody that interacts with the world, and makes decisions in response to observations

- A piece of software that makes decisions
- A robot
- A person accomplishing a task
- You!

- (This is the simplest case. Games add the other participants)



From Lagoudakis 2003
(my first Ph.D. student)

What is this class about?

How an **agent** can **deduce or learn** to act **intelligently** in the presence of **uncertainty**.

Everything in **BOLD** is something we need to define more precisely to address this challenge.

What forms of uncertainty do we face?

- Uncertainty about the immediate costs/reward of actions
- Uncertainty about how our actions change the world
- Uncertainty about state of the world
- Uncertainty about the choices of other agents
- Uncertainty about the intentions/goals of other agents

What is this class about?

How an **agent** can **deduce or learn** to act **intelligently** in the presence of **uncertainty**.

Everything in **BOLD** is something we need to define more precisely to address this challenge.

What is Intelligent Behavior? (at least for this class)

- Utility theory tells us how to **map our preference to real numbers**
- Decision theory tells us how to **maximize utility**
- **Intelligent behavior maximizes utility**
- Is it really that simple???

Why “simply” maximizing utility isn’t simple

- Uncertainty
- Short vs. long term benefit (This is the *sequential* part in the class title!)
- Computational challenges
- Behavior of other agents

What is this class about?

How an **agent** can **deduce or learn** to act **intelligently** in the presence of **uncertainty**.

Everything in **BOLD** is something we need to define more precisely to address this challenge.

Deducing or learning behaviors

- If we have a mathematical model of the environment:
  - Use dynamic programming or mathematical programming
  - (Often involves non-trivial computational challenges!)
  - Machine learning sometimes used even w/full knowledge of the environment
- If we don't have a model:
  - Bandit algorithms
  - Reinforcement learning

Example Applications

- Controlling robots
- Optimizing user interactions
- Increasing energy efficiency
- Playing games
- Investment strategies
- Health management

Example: Playing a Game Show

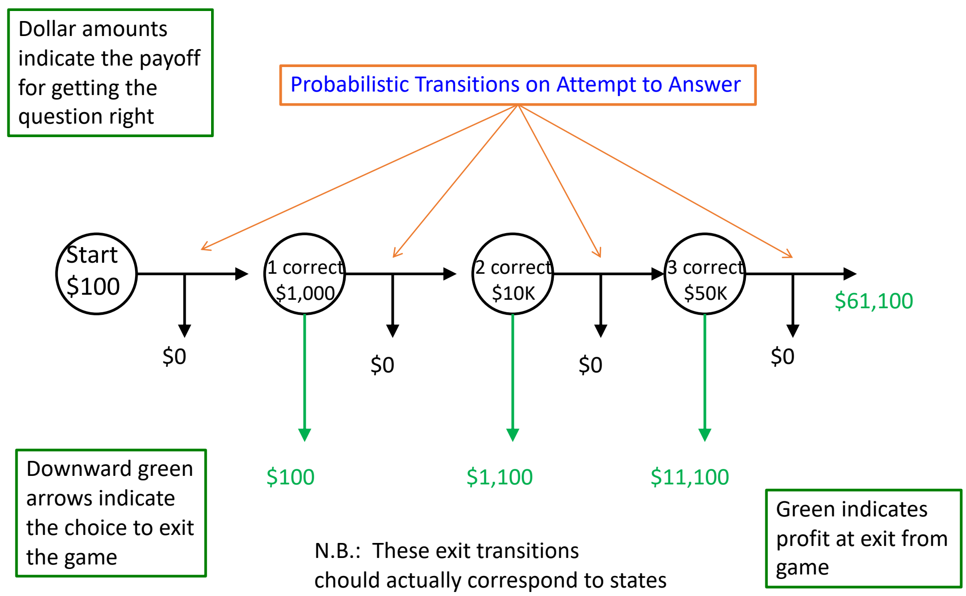
- Assume series of questions
  - Increasing difficulty
  - Increasing payoff
- Choice:
  - Accept accumulated earnings and quit
  - Continue and risk losing everything
- "Who wants to be a millionaire?"

Image source: https://en.wikipedia.org/wiki/Who_Wants_to_Be_a_Millionaire_(American_game_show)

Modeling assumptions (for today)

- Utility of money (assumed 1:1)
- Probabilities are known
- Every question answering opportunity is a "state"

State Representation



Probabilities

- Questions are of increasing difficulty

- First question: 9/10 probability of giving correct answer
- Second question: 3/4
- Third question: 1/2
- Fourth question: 1/10

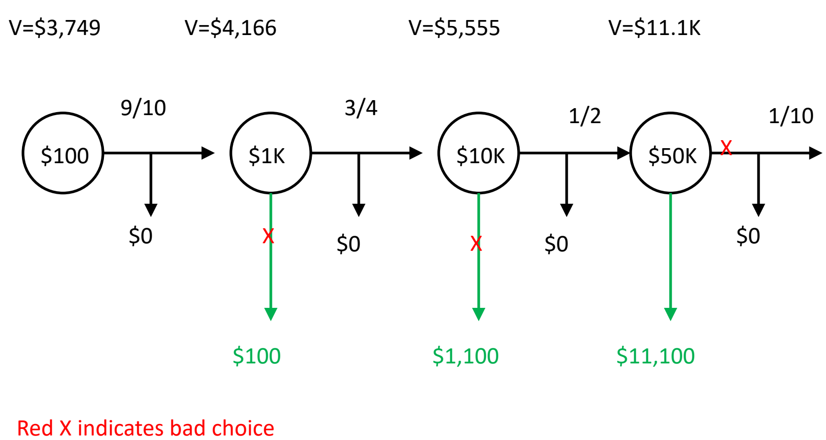
Making Optimal Decisions

- Work *backwards* from future to present – why?

- Consider $50,000 question
  - Suppose P(correct) = 1/10
  - V(stop)=$11,100
  - V(continue) = 0.9*$0 + 0.1*$61.1K = $6.11K

- Optimal decision stops

Working Backwards



Conclusions

- Intelligent behavior is "easy" to define – maximize utility

- Challenges:
  - How do we represent the problem, i.e., what is a "state"
  - Uncertainty – modeled and unmodeled
  - Sequential nature of decisions and ensuing computational challenges
  - Participation of other agents

- This class: Addresses these issues **theoretically** and **computationally**