Representational Dimensions

CMPUT 366: Intelligent Systems

P&M Chapter 1

Lecture Outline

1. Recap
2. Agents
3. Representations
4. Dimensions of representation

Recap:
Course Essentials

Course webpage: jrwright.info/aicourse/

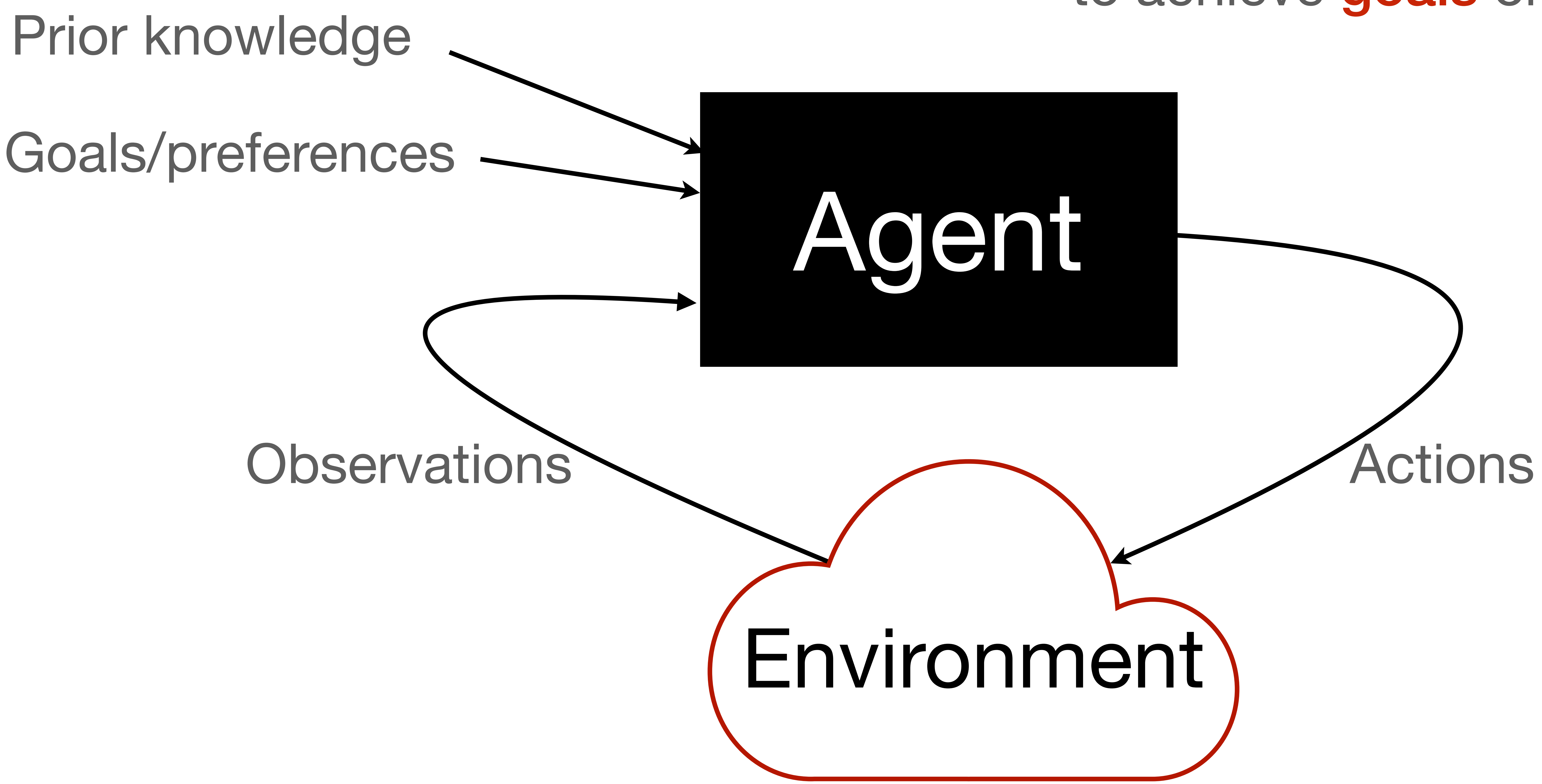
- This is the main source for information about the class
- Slides, readings, assignments, deadlines

Contacting us:

- Discussion board: piazza.com/ualberta.ca/winter2019/cmput366
  for **public** questions about assignments, lecture material, etc.
- Email: james.wright@ualberta.ca
  for **private** questions (health problems, inquiries about grades)
- Labs: Mondays 5:00pm to 7:50pm (CAB 235)
  No labs or TA office hours this week
- Office hours: After every lecture, or by appointment

Recap: Agents

An **agent** is a system that **acts** in an **environment** to achieve **goals** or optimize **preferences**.



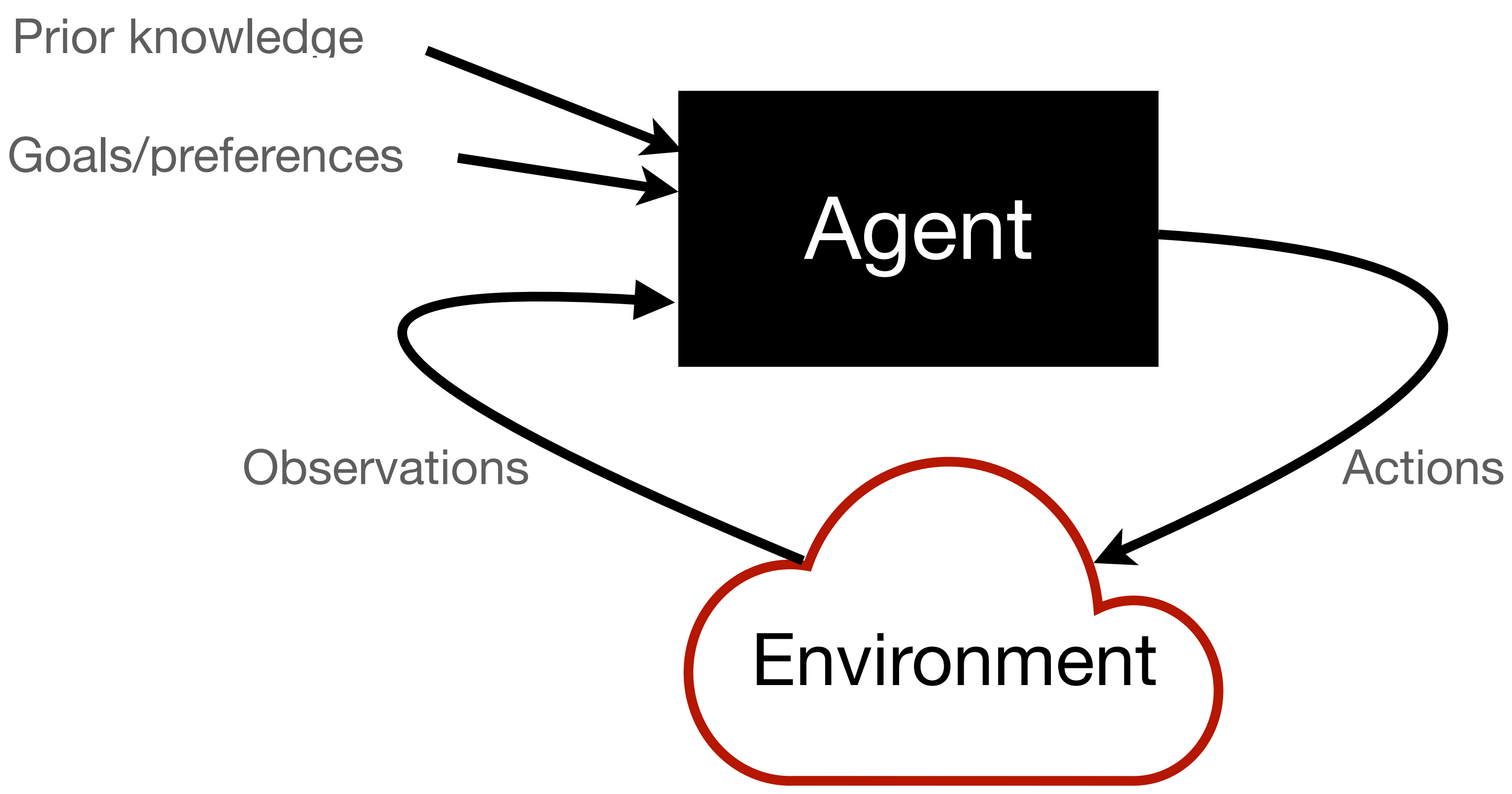
Examples

Which of these things are **agents**?

- A rock
- A tree
- A bird
- A car
- An decision assistant (previously: "expert system")
- A self-driving car
- A child
- An adult

Representations

- **Knowledge**: Information about a domain useful for solving problems in that domain
- To use knowledge, a computational agent needs to encode it into a set of data structures called a **representation**
- Representations are about the **environment**:
  - What kinds of **states** can the world be in? How should we denote them?
  - What kind of **information** do we get about the current and past states? How certain can our **beliefs** be?
  - **Dynamics**: How does the environment change based on our actions? Are the changes deterministic, or stochastic?



Representations: Wishlist

What do we want from a representation?

- **Rich enough** to express all of the knowledge necessary for solving the task
- As **close to the problem** as possible: Compact, natural, maintainable, transparent
- **Tractable**: Amenable to efficient computation
- **Learnable**: Should be able to acquire new knowledge from data, past experiences, people
- Able to **trade off** accuracy and computation time

Primary Dimensions

We will classify domains by three main dimensions:

1. **Uncertainty**: deterministic vs. stochastic settings
2. **Interaction**: Online vs. offline interaction
3. **Number of agents**: Single vs. multiple agents

1. Uncertainty

Multiple aspects of an environment may be **deterministic** (no randomness) or **stochastic** (some randomness)

1. Observations and state ("sensing uncertainty")
   - **Fully observable**: observations directly determine **state**
   - **Partially observable**: **many** possible states for same observations; or observations are **misleading**
2. Dynamics ("effect uncertainty")
   - **Deterministic dynamics**: state completely determined by action and prior state
   - **Stochastic dynamics**: probability distribution over possible states from an action

2. Interaction

When does the agent decide what to do?

- **Offline**: Agent determines what to do **before** interacting with the environment
- **Online**: Agent determines what to do **while** interacting with the environment
  - Often more computationally demanding
  - Requires timely answers

3. Number of Agents

Does the agent (need to) explicitly consider other agents?

- **Single agent:**
  - No other agents in the environment, or
  - Behaviour of other agents is fixed (part of **nature**)
- **Multiple agents:**
  - Other agents in the environment, with **distinct** goals/preferences
  - Must **reason** about other agent's behaviour *and reasoning*
  - **Other agents' actions** affect agents goals/preferences, *and vice versa*

Other Dimensions

- Static vs. sequential action
- Goals vs. complex preferences
- Episodic vs. continuing
- State representation scheme
- Perfect vs. bounded rationality

Different dimensions **interact**; you can't just set them arbitrarily

Static vs. Sequential Action

How many actions does the agent need to select?

- **Static:** the agent selects a **single** action
    - Classify an image
    - Diagnose a disease based on symptoms
    - Recommend a driving route
- **Sequential:** the agent needs to take a **sequence** of actions
    - Participate in an automated negotiation
    - Choose a series of tests to diagnose a patient
    - Navigate through an environment
- In a **deterministic** setting, this can be an arbitrary distinction

Goals vs. Preferences

How complicated a goal is the agent trying to achieve?

- **Goal**: A simple desired condition
    - **Maintenance goal**: Keep some already-true condition true in **all visited states**
    - **Achievement goal**: Condition should be true in **final state**
    - E.g.: Robot trying to deliver banana to Sam without hurting anyone or making a mess
- **Preferences**: Varying desirability of different outcomes, trade-offs
    - **Ordinal preferences**: Only the **ordering** of outcomes is important
    - **Cardinal preferences: Magnitude** of preference also matters

Knowledge Given vs. Knowledge Learned

How much does the agent know about the world in advance?

- The agent **has a model** of the world **before** it acts
- The agent must **learn** how the world works
    - from **data**
    - from **experience**
    - often still starts with *some* **prior knowledge**

State Representation

How does the agent describe the world?

- **Enumerate** every possible state of the world
- Factor each state into **features**
  - May or may not be **observable**
  - 20 binary features can represent over a million states
- **Relationships** and **objects**

Episodic vs. Continuing

Is the task ever done?

- **Episodic**: The agent eventually reaches a **final state**
- **Continuing**: The agent keeps acting **forever**
- Especially crucial distinction in reinforcement learning

Related: **Planning horizon**

Perfect Rationality vs. Bounded Rationality

Is it feasible for the agent to achieve the ideal optimum, or must it trade off solution quality against computational cost?

- **Perfect rationality:** The agent can derive the best course of action without accounting for **computational limitations**.

- **Bounded rationality:** Agent decides on best action that **it can find** within its computational limitations

  - **Anytime algorithm:** Solution quality improves with time

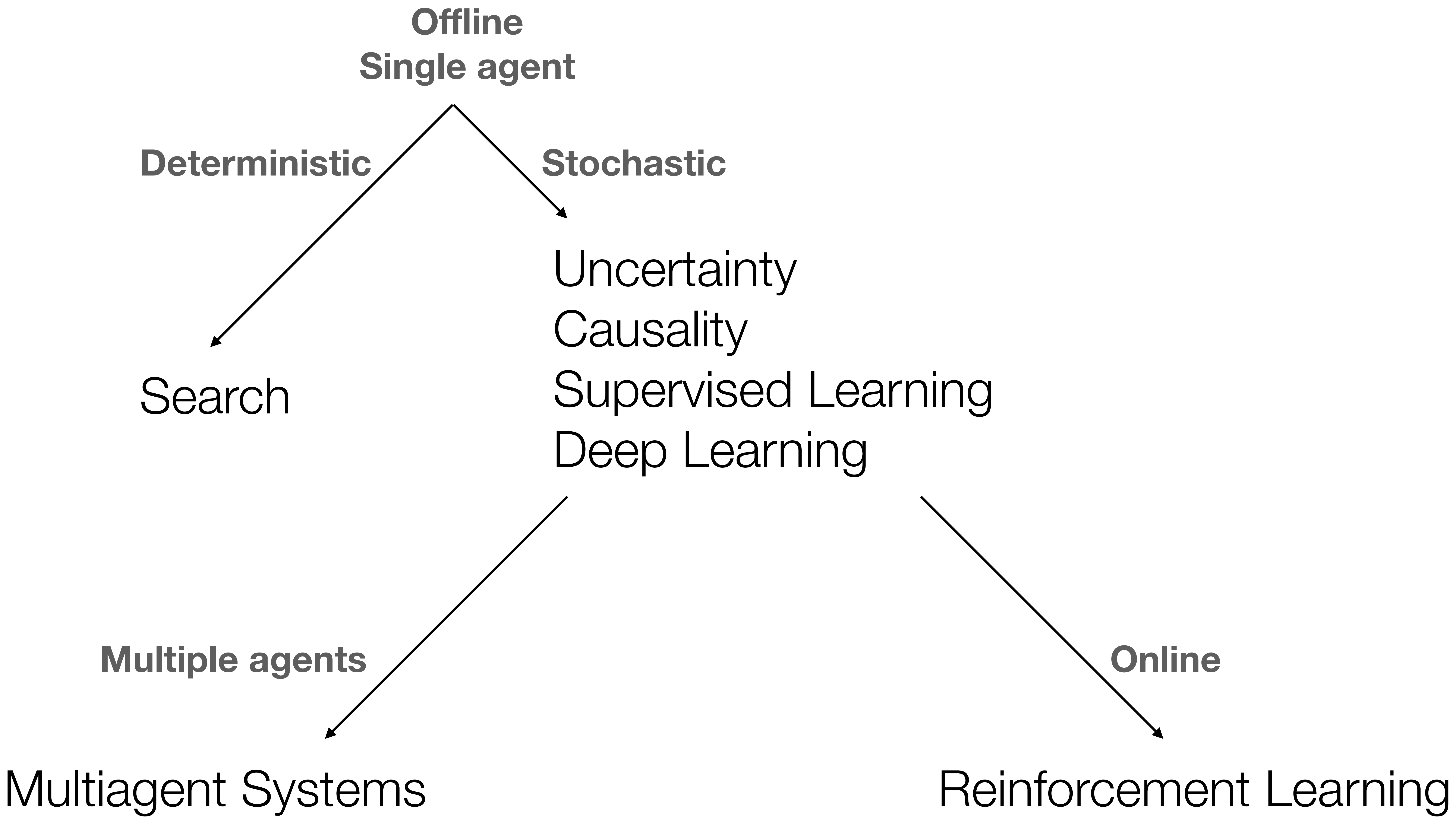
Dimensions

- Static vs. sequential action
- Goals vs. complex preferences
- Episodic vs. continuing
- State representation scheme
- Perfect vs. bounded rationality

1. Uncertainty
2. Interaction
3. Number of agents

Different dimensions **interact**; you can't just set them arbitrarily

Course Topics Breakdown



Summary

- **Course webpage:** jrwright.info/aicourse/
- Formal **representation** of an environment is essential for building agents
- We can usefully classify environments and representations according to a number of dimensions
    - These dimensions are not independent