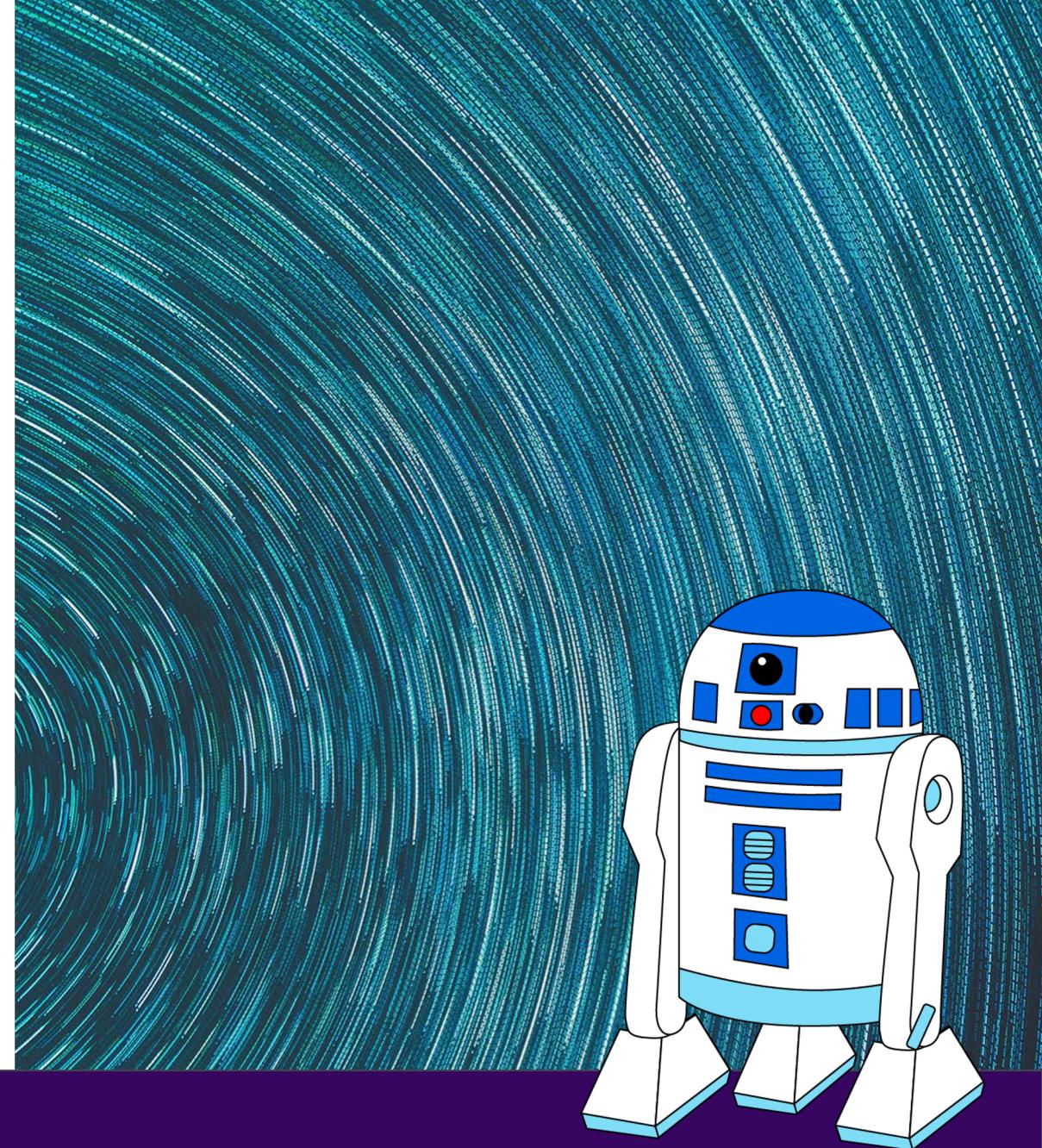


CIS 421/521:
ARTIFICIAL INTELLIGENCE

Rational Agents



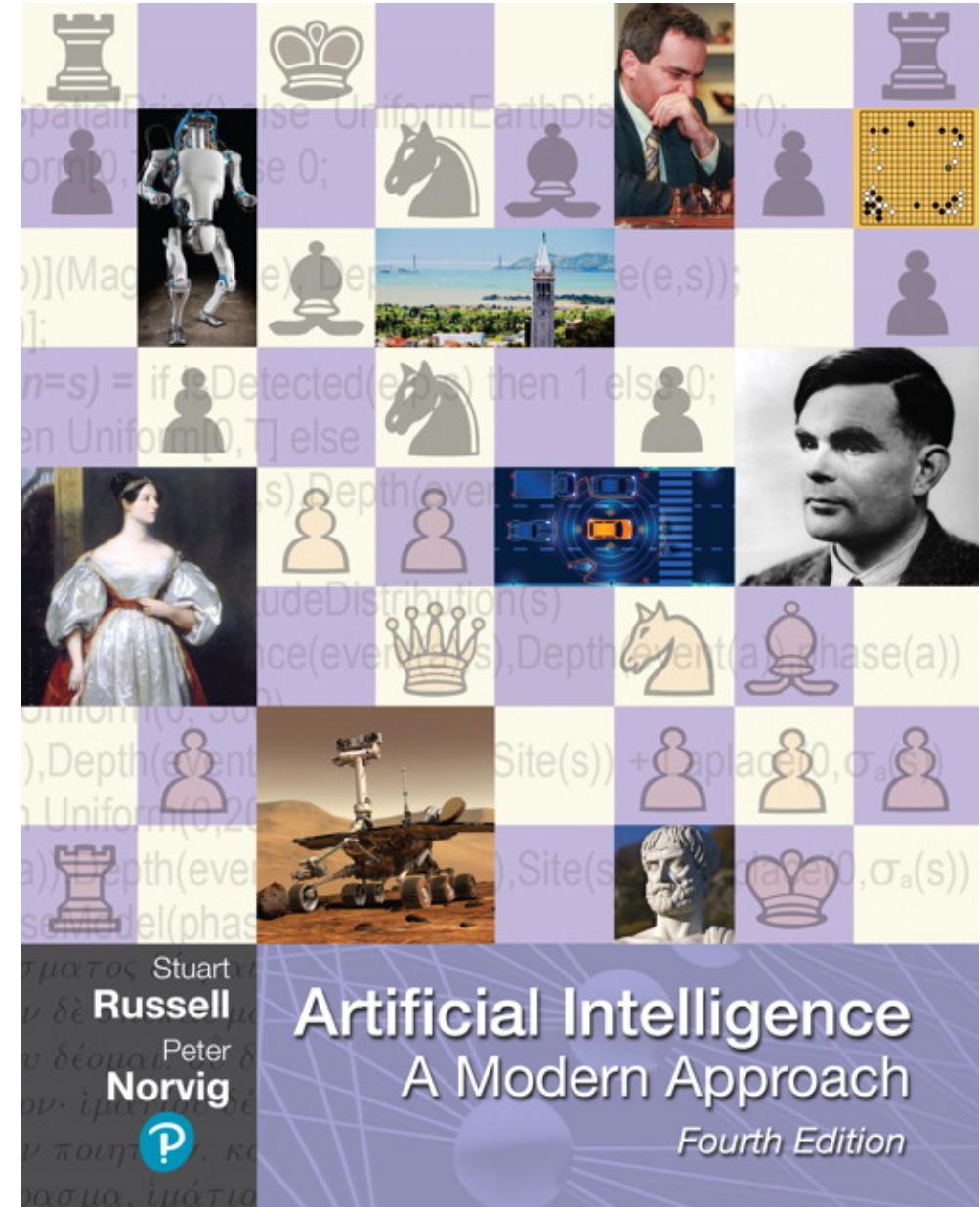
Outline for today's lecture

Intelligent Agents (AIMA 2.1-2.4)

Task Environments

Formulating Search Problems

Uninformed Search (AIMA 3.1-3.4)



**Artificial Intelligence
A Modern Approach**

Fourth Edition



Four views of Artificial Intelligence

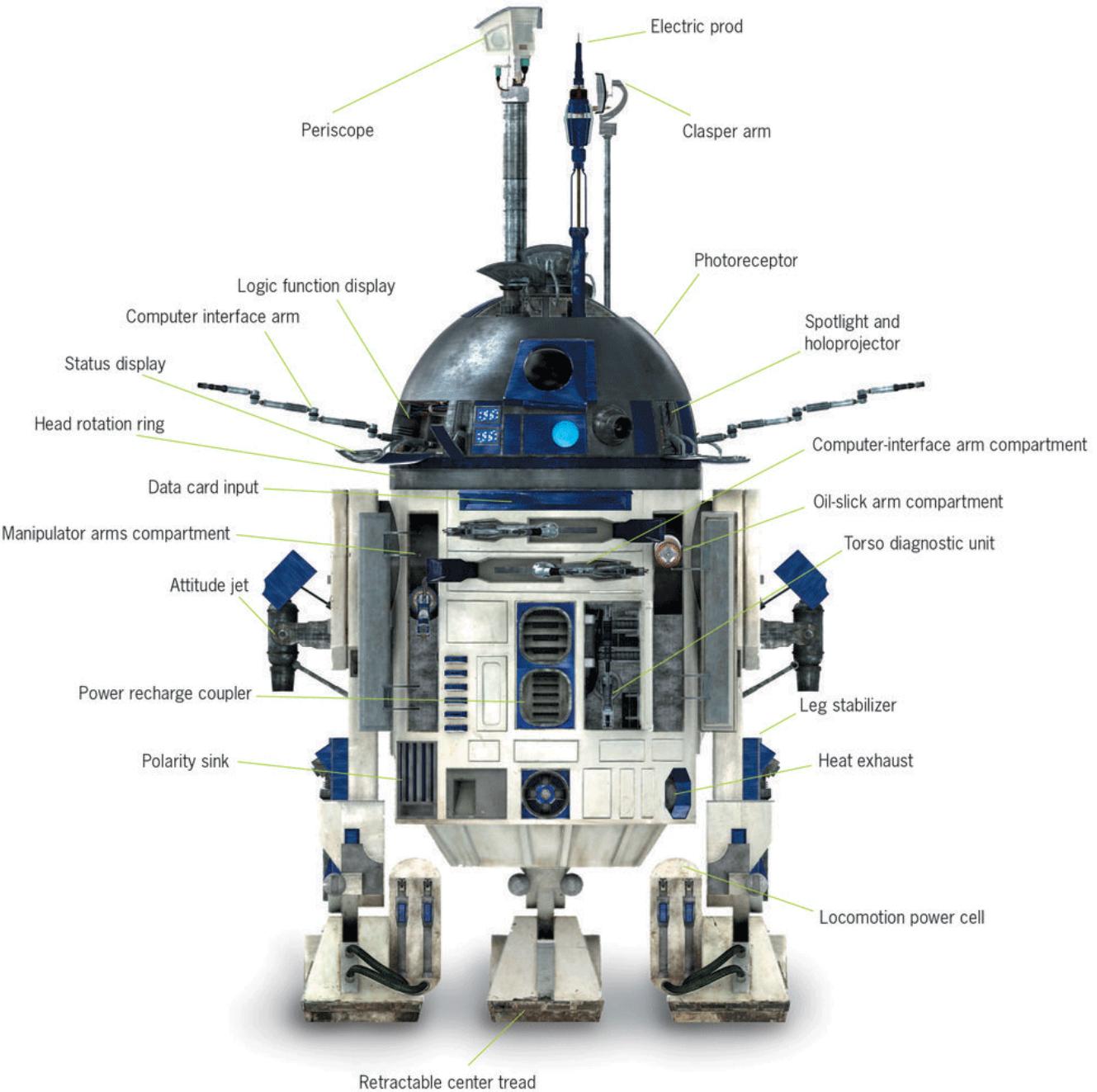
Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

This course is about effective programming techniques for designing rational agents

Agents

An **agent** is anything that **perceives** its environment through **sensors** and can **act** on its environment through **actuators**

A **percept** is the agent's perceptual inputs at any given instance.



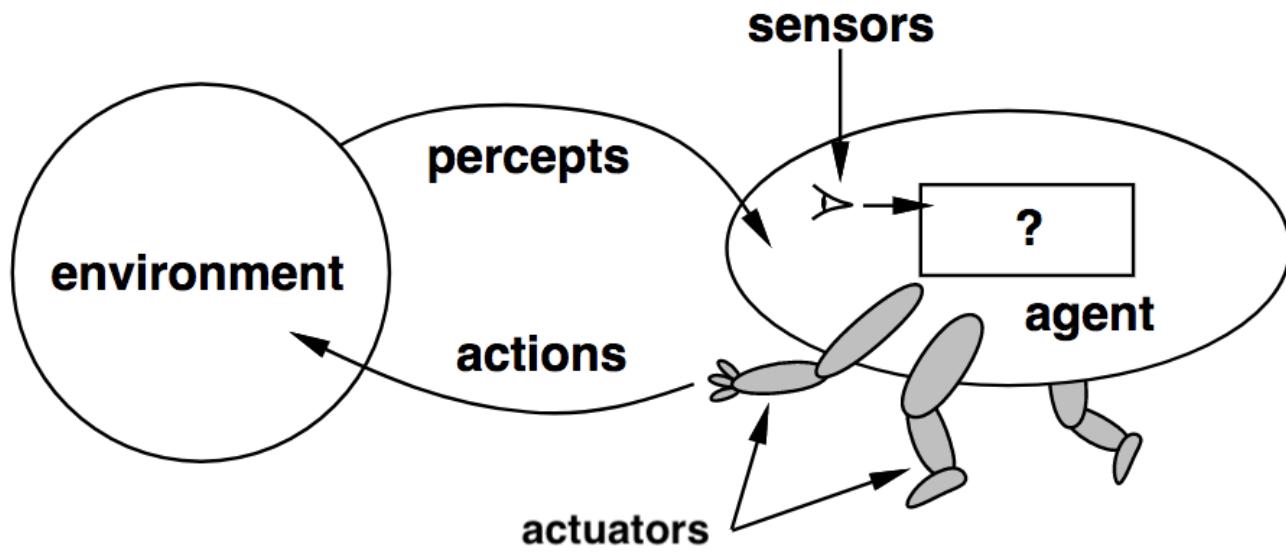
What about your robot?



What actuators does it have?

What sensors does it have?

Agents and environments



An agent is specified by an *agent function* $f:P \rightarrow A$ that maps a sequence of percept vectors P to an action a from a set A :

$$P = [p_0, p_1, \dots, p_t]$$

$$A = \{a_0, a_1, \dots, a_k\}$$

abstract
mathematical
description

Agent function & program

The *agent program* runs on the physical *architecture* to produce f

- $\text{agent} = \text{architecture} + \text{program}$

“Easy” solution: a giant table that maps every possible sequence P to an action a

- One small problem: exponential in length of P

Agents

An *agent* is anything that can be viewed as

- *perceiving* its *environment* through *sensors* and
- *acting* upon that environment through *actuators*

Human agent:

- Sensors: eyes, ears, ...
- Actuators: hands, legs, mouth, ...

Robotic agent:

- Sensors: cameras and infrared range finders
- Actuators: various motors

Agents include humans, robots, softbots, *thermostats*, ...



Rational Agent

Let's try to define "rational agent".

A **rational agent** is an agent that perceives its environment and behaves rationally

Rational behavior: doing the right thing

Obviously doing the right thing is better than doing the wrong thing, but *what does it mean to do the right thing?*

In Philosophy

Moral philosophy has developed different notions of “the right thing”.

AI is usually concerned with **Consequentialism**.

We evaluate an agent’s behavior by its consequences.



The Good Place

In Economics

A BEHAVIORAL MODEL OF RATIONAL CHOICE

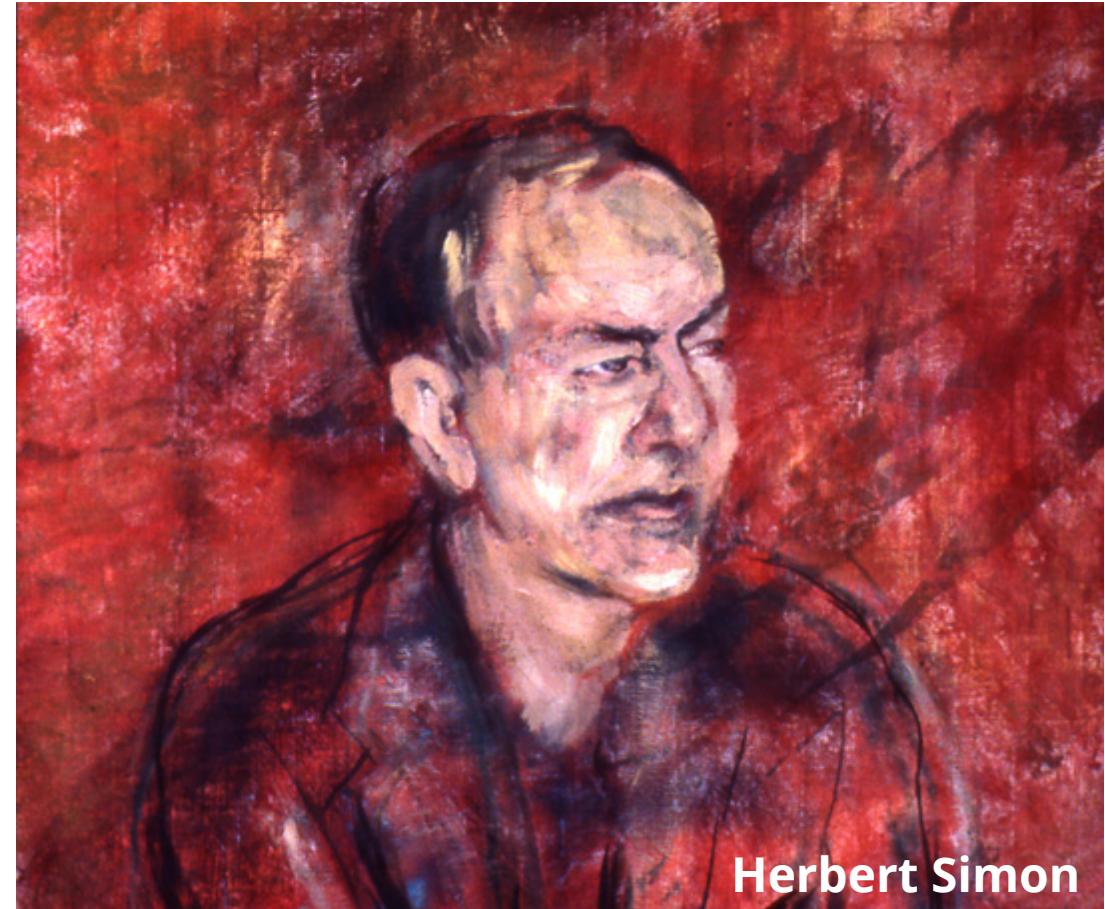
Herbert A. Simon*

Summary: A model is proposed for the description of rational choice by organisms of limited computational ability.

Rational choice theory is a framework for understanding social and economic behavior.

The basic premise is that aggregate social behavior results from the behavior of individual actors, each of whom is making their individual decisions.

It assumes that individuals have preferences and choose the alternative that they prefer.



Herbert Simon

Performance measure

How do we know if an agent is acting rationally?

- Informally, we expect that it will do the right thing in all circumstances.

How do we know if it's doing the right thing?

We define a **performance measure**:

- An objective criterion for success of an agent's behavior
- given the evidence provided by the percept sequence.

Performance measure - example

A performance measure for a vacuum-cleaner agent might include e.g. some subset of:

- +1 point for each clean square in time T
- +1 point for clean square, -1 for each move
- -1000 for more than k dirty squares



Performance measure – rule of thumb

It is better to design performance measures according to **what you want to be achieved** in the environment, **rather than how you think the agent should behave**.

For example what might happen if we said

- +1 point for each time the robot cleans a square instead of
- +1 point for each clean square in time T

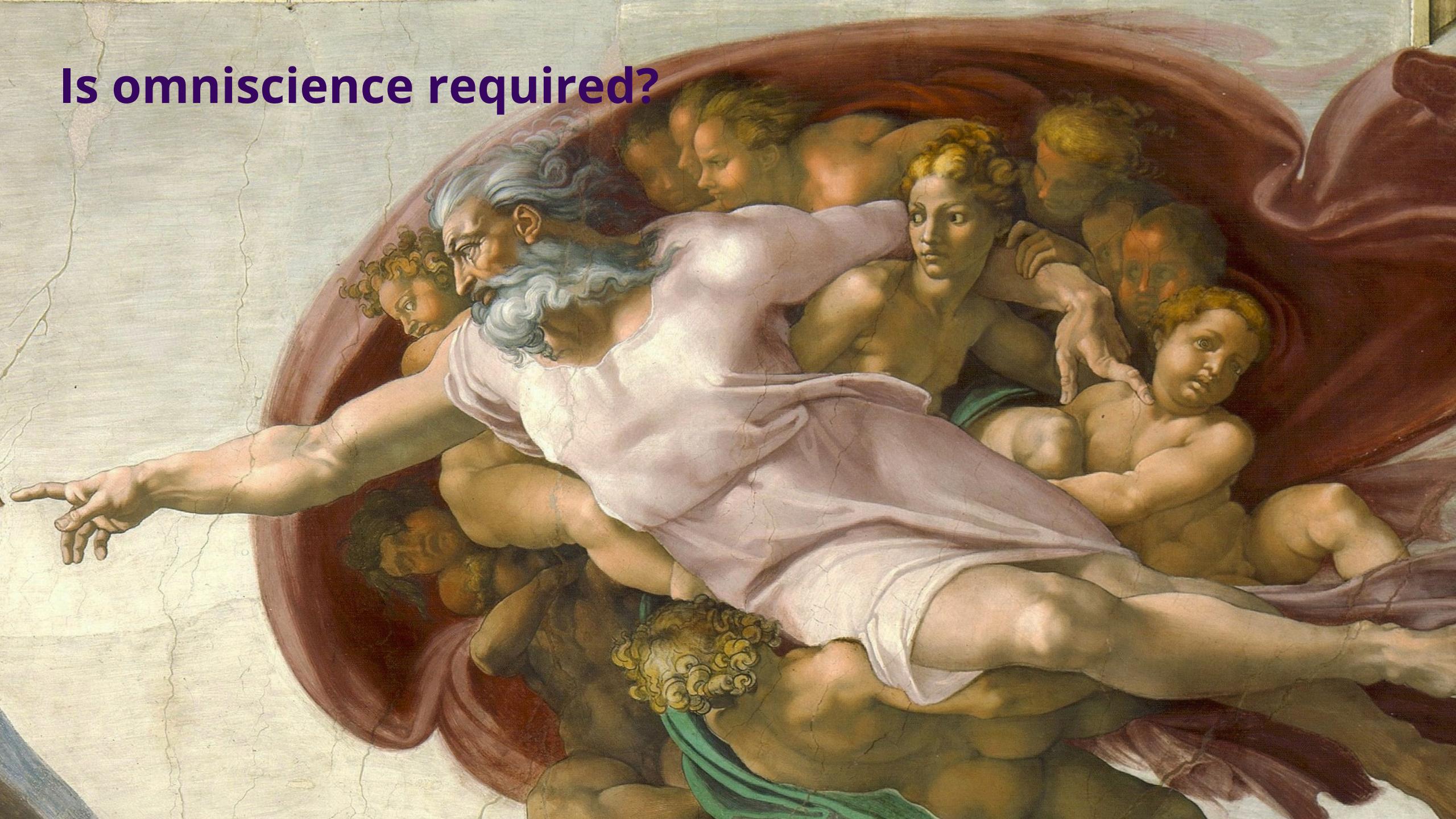


Rational agents

Rational Agent:

- For each possible percept sequence P
- a rational agent selects an action a
- to *maximize* its *performance measure*

Is omniscience required?



Expected value

Rational Agent (initial definition):

- For each possible percept sequence P,
- a rational agent selects an action a
- to maximize its performance measure

Rational Agent (revised definition):

- For each possible percept sequence P,
- a rational agent selects an action a
- that maximizes the **expected value** of its performance measure

*It doesn't have to know
what the actual outcome
will be.*

Task environments

To design a rational agent we need to specify a *task environment*

- a problem specification for which the agent is a solution

PEAS: to specify a task environment

- **P**erformance measure
- **E**nvironment
- **A**ctuators
- **S**ensors



PEAS: Specifying an automated taxi driver

Performance measure:

- ?

Environment:

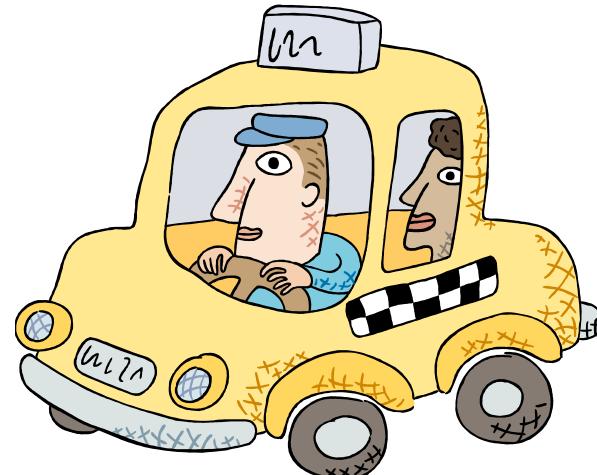
- ?

Actuators:

- ?

Sensors:

- ?



PEAS: Specifying an automated taxi driver

Performance measure:

- safe, fast, legal, comfortable, maximize profits

Environment:

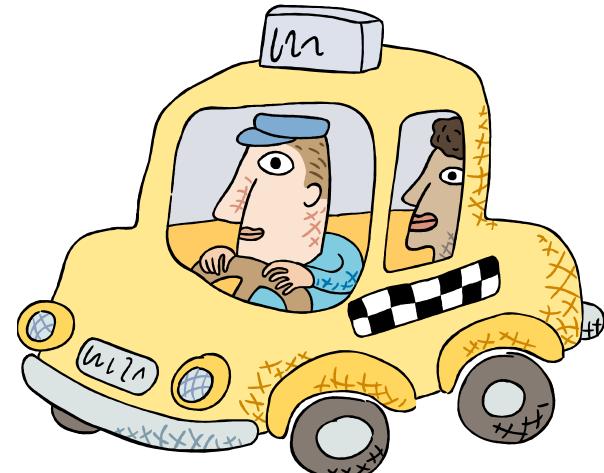
- roads, other traffic, pedestrians, customers

Actuators:

- steering, accelerator, brake, signal, horn

Sensors:

- cameras, LiDAR, speedometer, GPS



PEAS: Amazon Prime Air

*P*erformance measure:

- ?

*E*nvironment:

- ?

*A*ctuators:

- ?

*S*ensors:

- ?





<https://www.today.com/video/amazon-adebuts-new-package-delivery-drone-61414981780>

PEAS: Specifying an Amazon delivery drone

Performance measure:

- maximize profits - minimize time - obey laws governing airspace restrictions -
- deliver package to right location - keep package in good condition - avoid accidents
- reduce noise - preserve battery life

Environment:

- airspace - obstacles when airborne (other drones, birds, buildings, trees, utility poles) - obstacles when landing (pets, patio furniture, lawnmowers, people, cars) -
- weather - distances/route information between warehouse and destinations -
- position of houses, and spaces that are safe for drop-off- package weight

PEAS: Specifying an Amazon delivery drone

A^ctuators:

- Propellers and flight control system- Payload actuators: E.g. Arm/basket/claw for picking up, dropping off packages- Lights or signals - Mechanism to announce/verify delivery- Device for delivering packages to customers

S^enso^rs:

- GPS - radar/Lidar- altitude sensor- weather sensors (barometer, etc). - gyroscope- accelerometer- camera- rotor sensors- weight sensor to recognize package

The rational agent designer's goal

Goal of AI practitioner who designs rational agents:

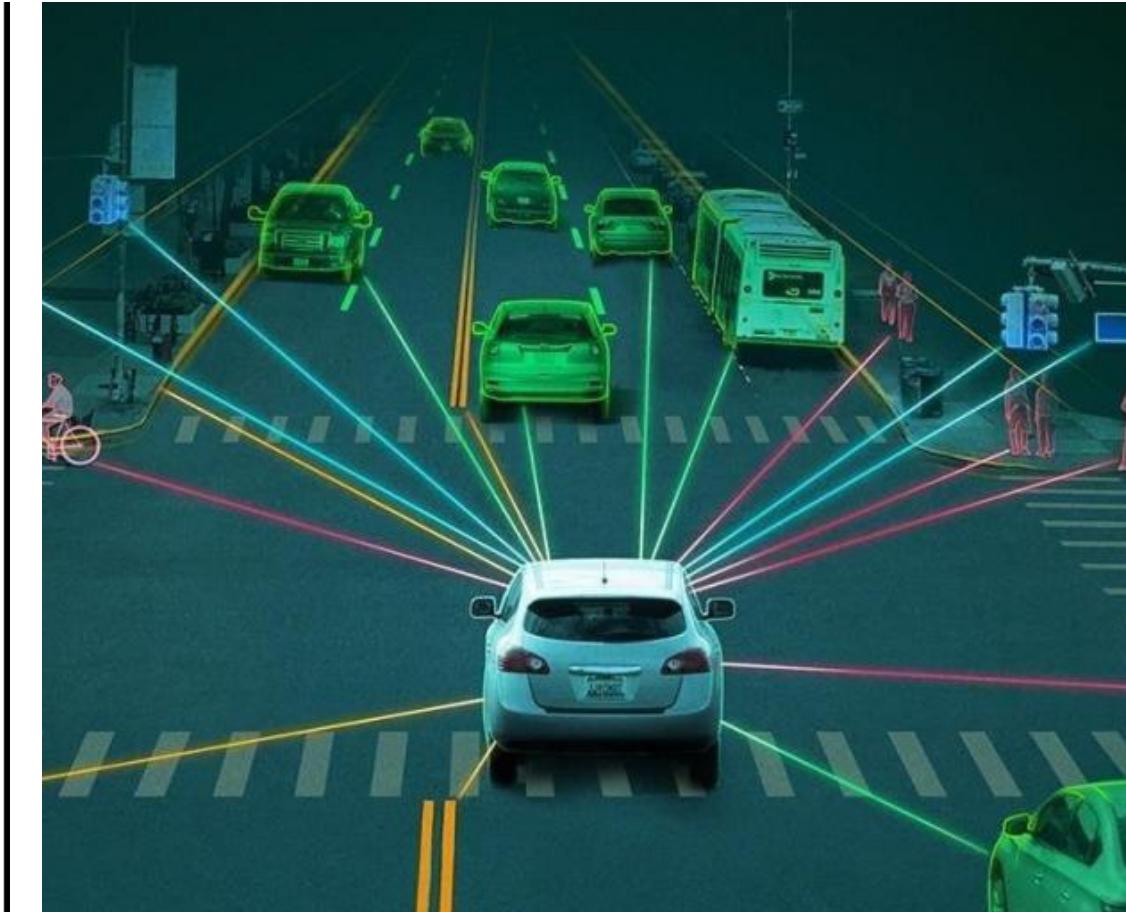
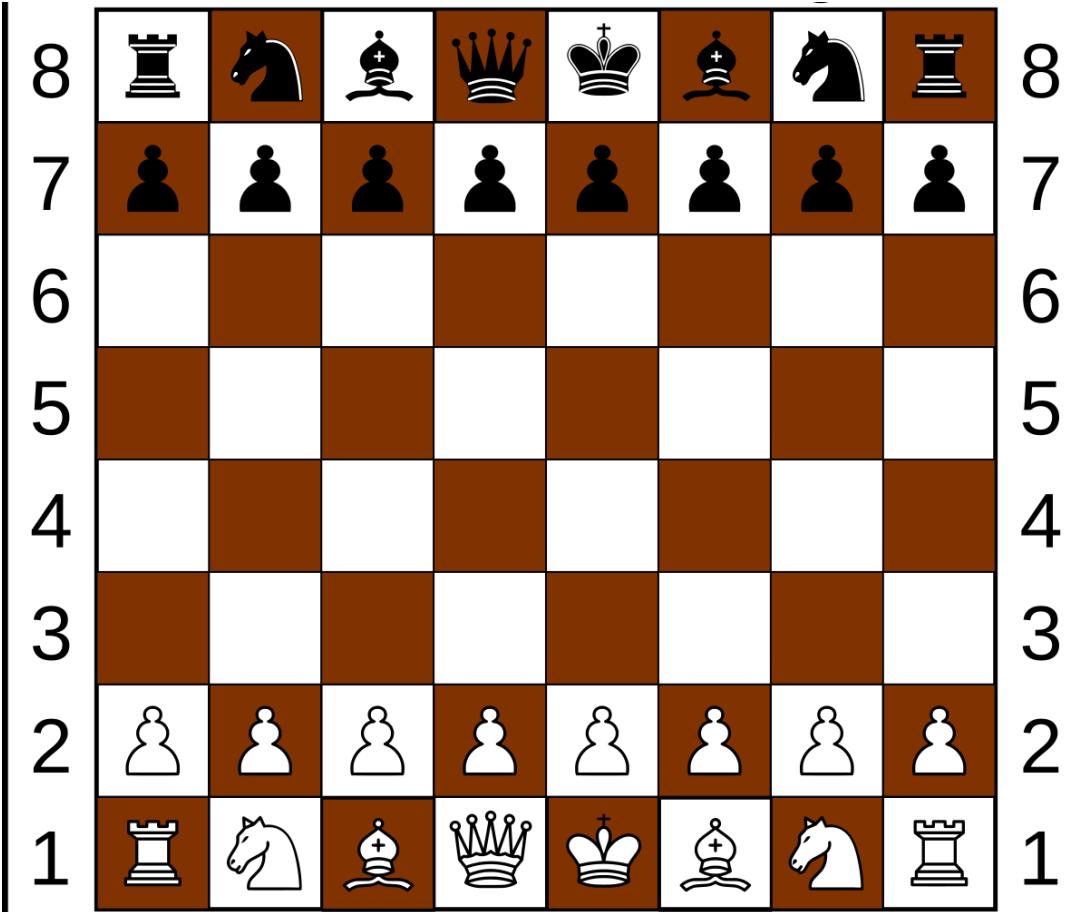
given a *PEAS* task environment,

1. Construct *agent function* f that maximizes the expected value of the performance measure,
2. Design an *agent program* that implements f on a particular architecture

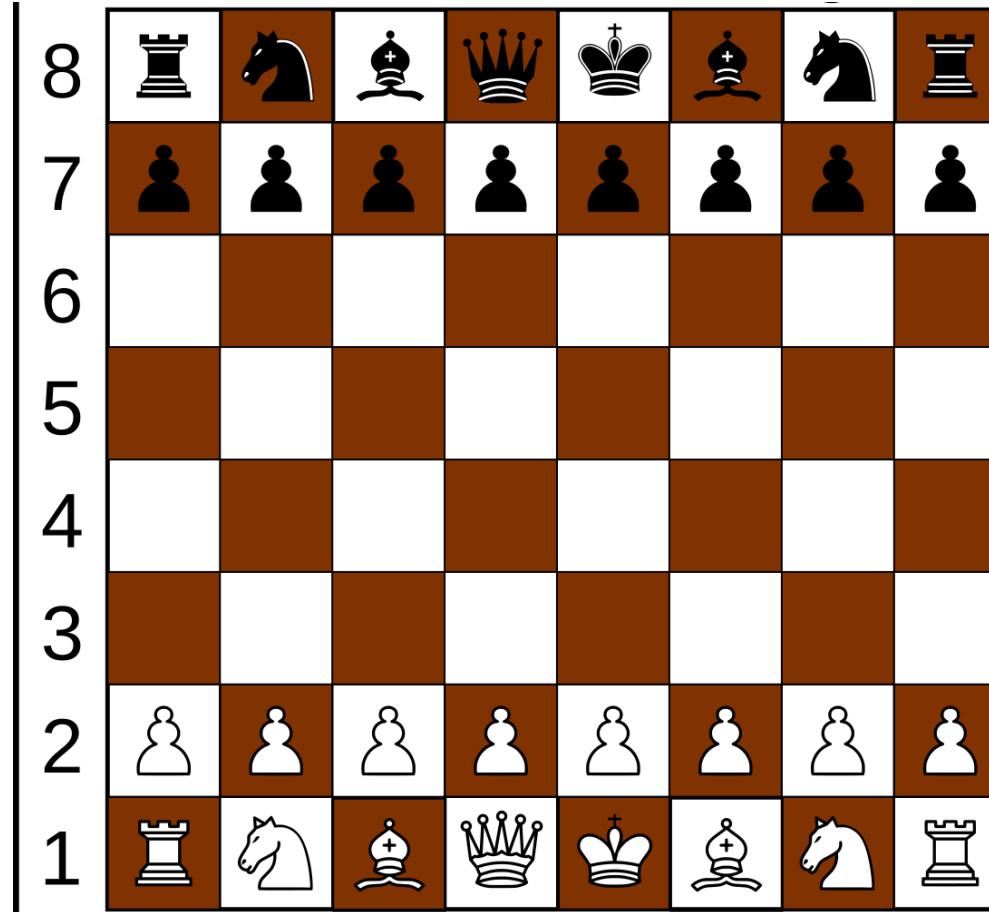
abstract
mathematical
description

concrete
implementation

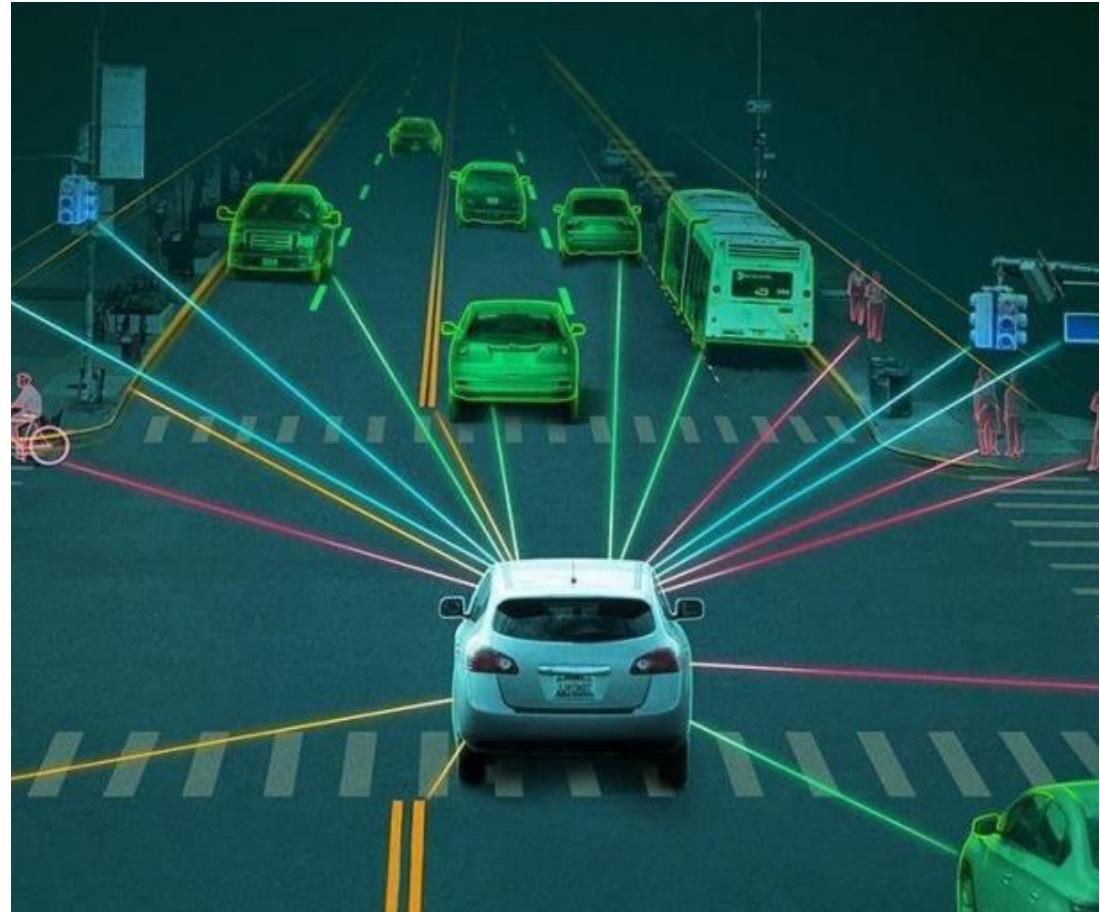
Fully Observable v. Partially Observable



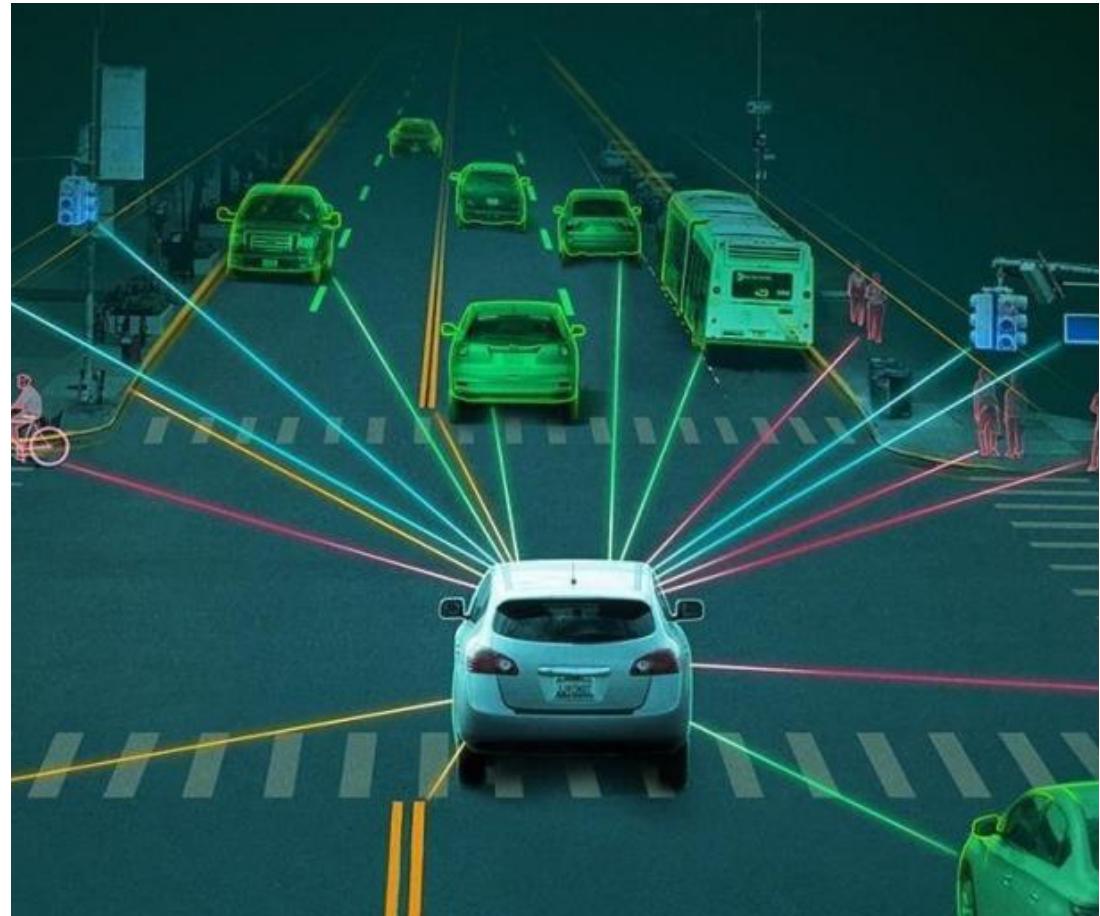
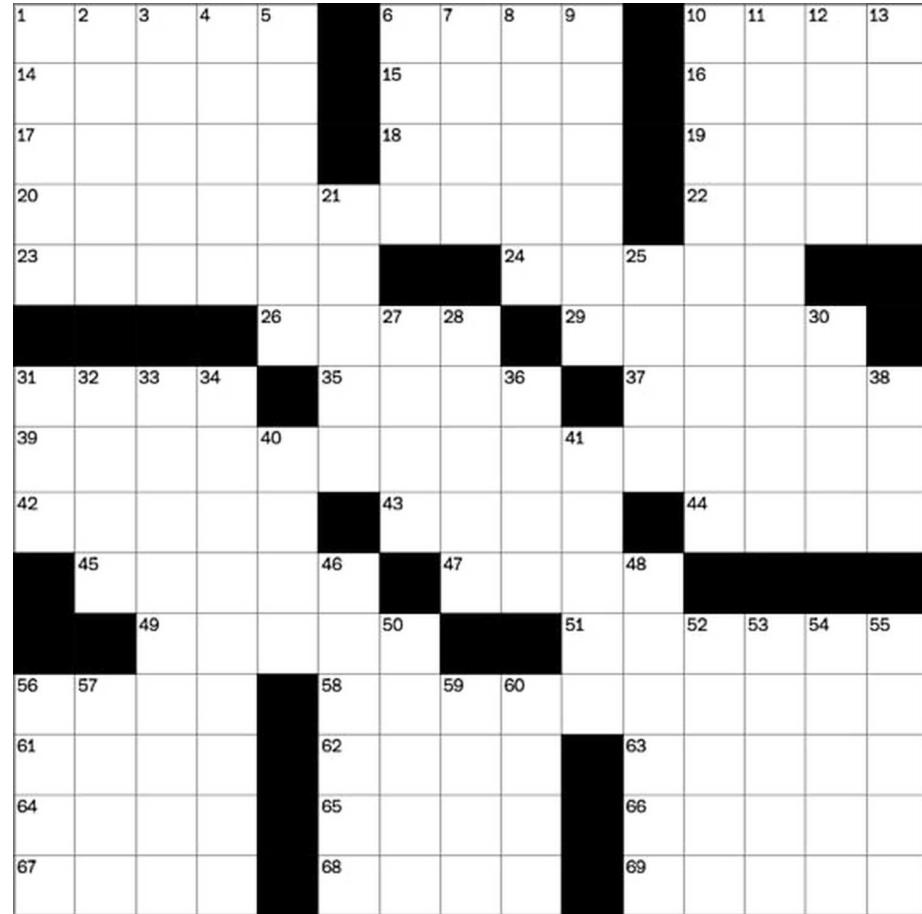
Deterministic v. Nondeterministic v. Stochastic



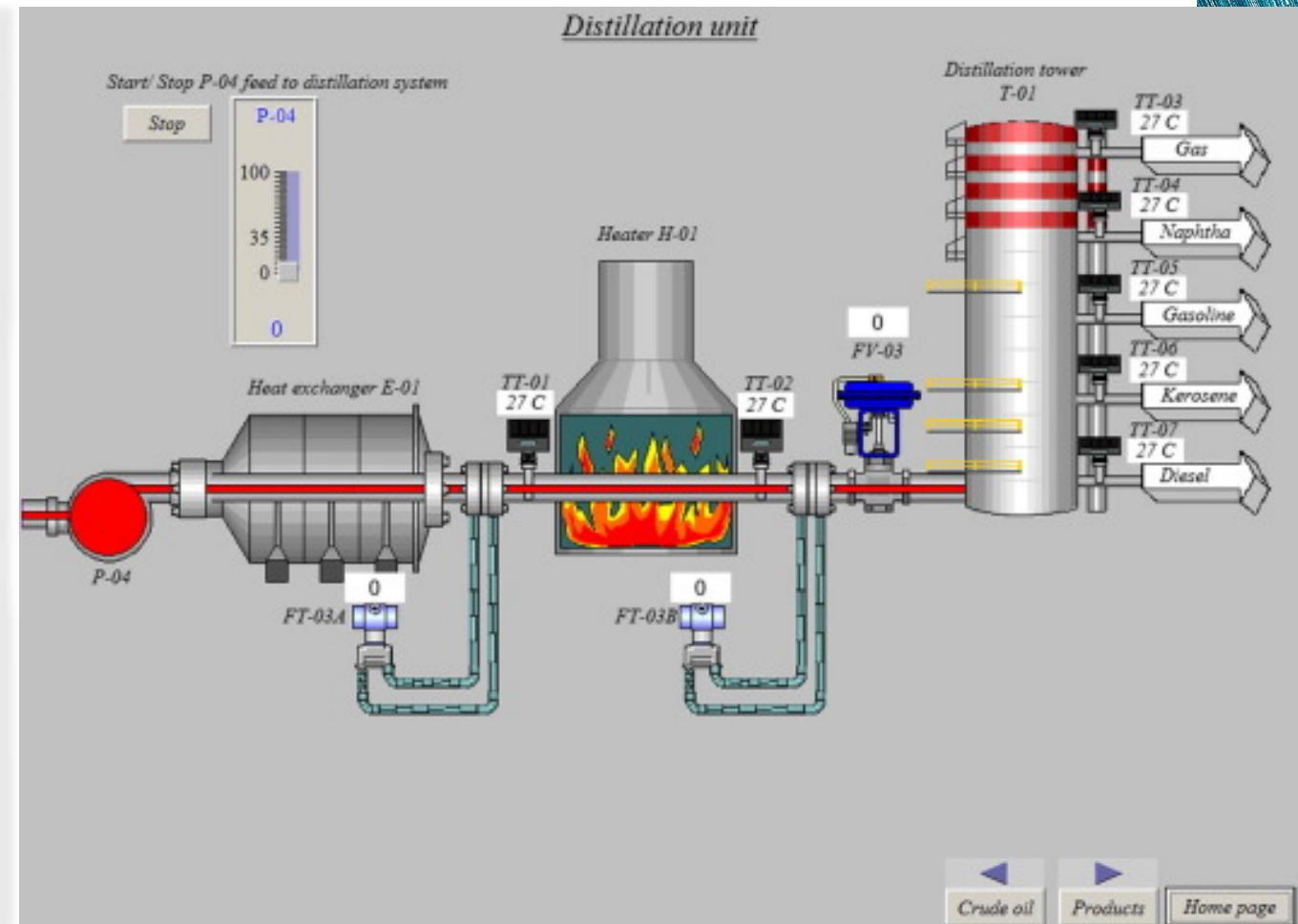
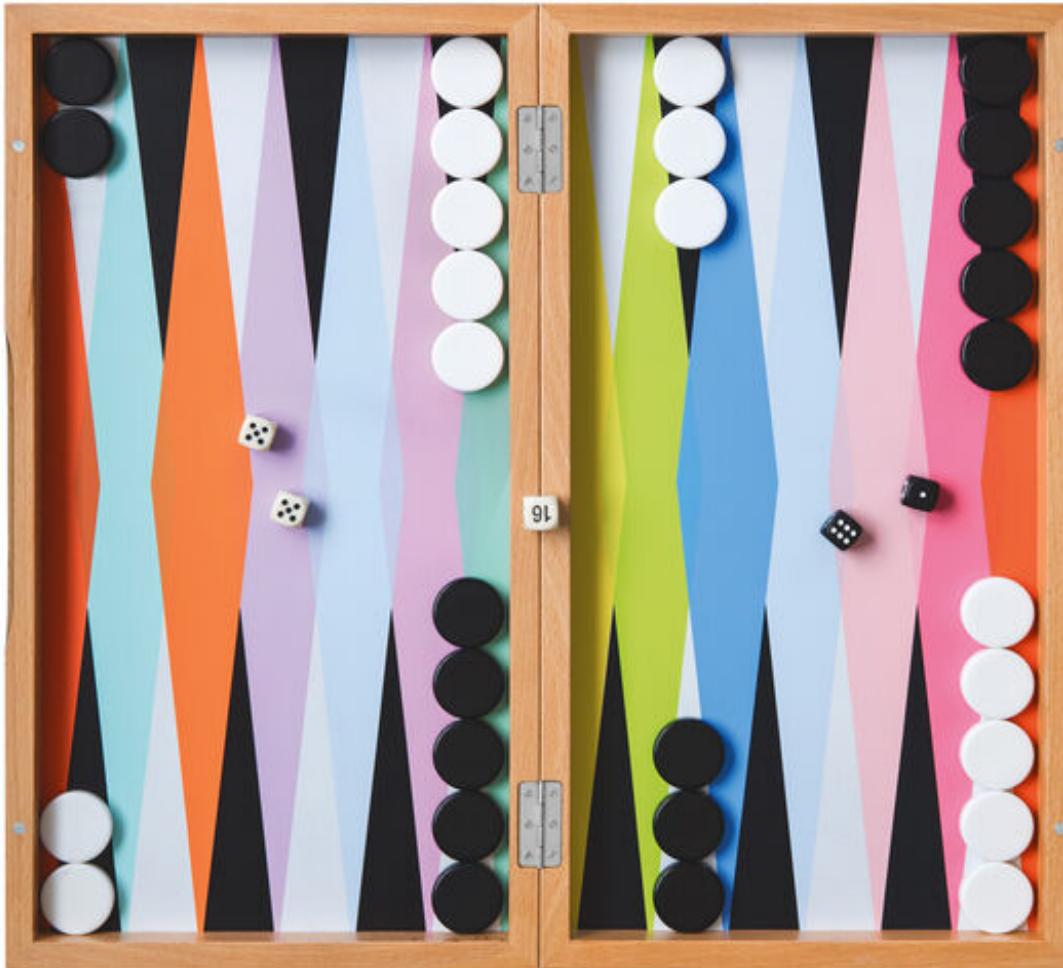
Episodic v. Sequential



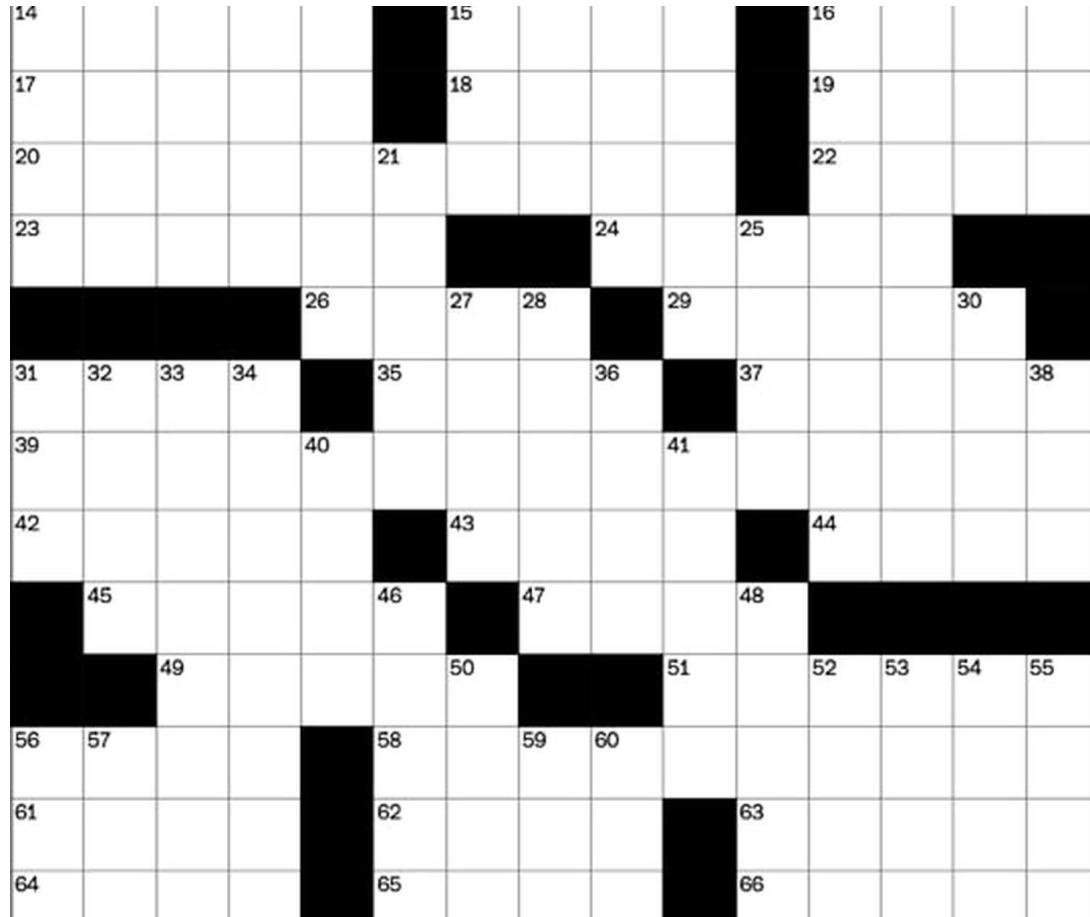
Static v. Dynamic



Discrete v. Continuous



Single Agent v. Multi Agent



When should something be considered an agent?

When should something be considered another agent?

If we're talking about a self driving taxi, when should we consider something part of the environment versus another agent?

For instance, a telephone pole is part of the environment, but a car might be another agent.

When something behavior can best be described as having its own performance measure, then we should consider it to be an agent.

Examples

Task Environment	Observable	Agents	Deterministic	Episodic	Static	Discrete
Crossword puzzle Chess with a clock	Fully Fully	Single Multi	Deterministic Deterministic	Sequential Sequential	Static Semi	Discrete Discrete
Poker Backgammon	Partially Fully	Multi Multi	Stochastic Stochastic	Sequential Sequential	Static Static	Discrete Discrete
Taxi driving Medical diagnosis	Partially Partially	Multi Single	Stochastic Stochastic	Sequential Sequential	Dynamic Dynamic	Continuous Continuous
Image analysis Part-picking robot	Fully Partially	Single Single	Deterministic Stochastic	Episodic Episodic	Semi Dynamic	Continuous Continuous
Refinery controller Interactive English tutor	Partially Partially	Single Multi	Stochastic Stochastic	Sequential Sequential	Dynamic Dynamic	Continuous Discrete

The Hardest Environment

The hardest case is

- *Continuous*
- *Partially Observable*
- *Stochastic*
- *Continuous*
- *Multiagent*
- *Unknown Outcomes*

Environment Restrictions for Now

We will assume environment is

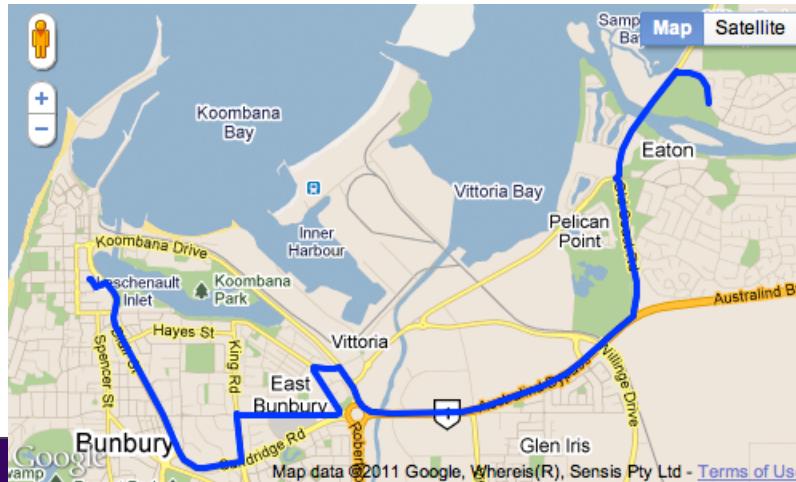
- *Static*
- *Fully Observable*
- *Deterministic*
- *Discrete*

Reflex agents v. Problem solving agents

A simple reflex agent is one that selects an action based on the current percept, and ignores the rest of the percept history.



A problem-solving agent must plan ahead. It will consider a sequence of actions that form a path to a goal state. The computational process that it undertakes is called search.



Problem Solving Agents & Problem Formulation

AIMA 3.1-3.2

Example search problem: 8-puzzle

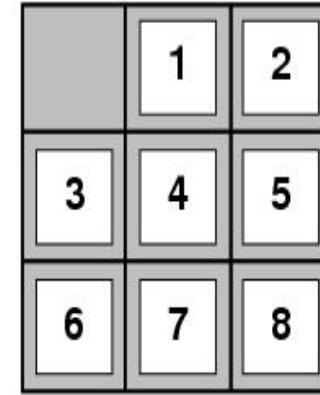


Formulate *goal*

- Pieces to end up in order as shown...



Start State



Goal State

Formulate *search problem*

- States:** configurations of the puzzle ($9!$ configurations)
- Actions:** Move one of the movable pieces (≤ 4 possible)
- Performance measure:** minimize total moves

Find *solution*

- Sequence of pieces moved: 3,1,6,3,1,...