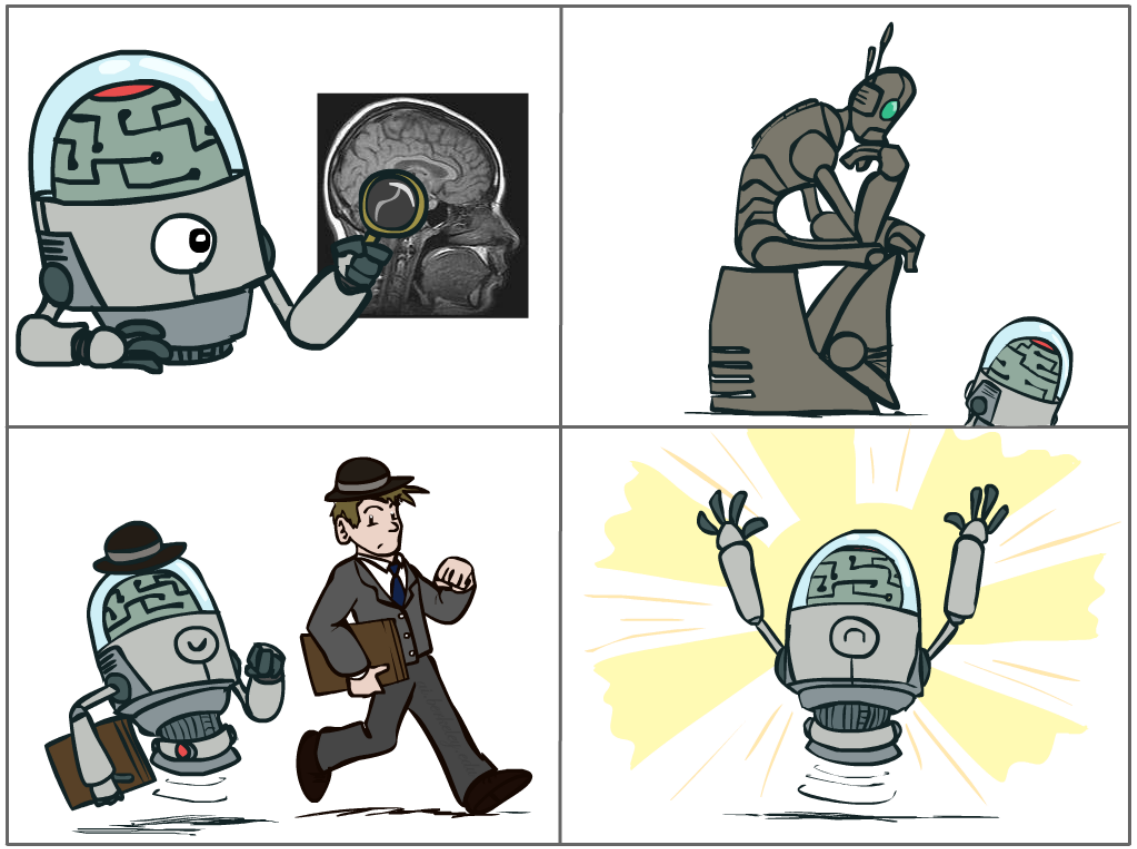
What is AI?

The science of making machines that:

Think like people



Think rationally

Act like people

Act rationally

Fundamental question for this lecture
(and really this whole class!):

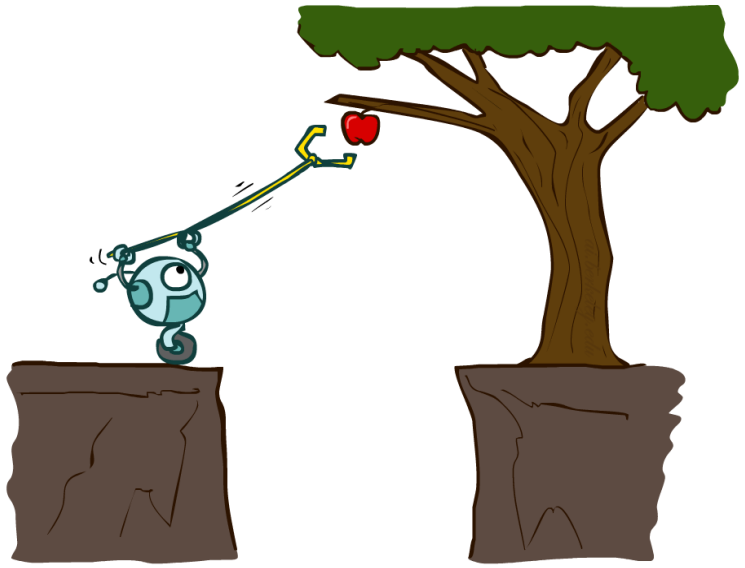
How do you turn a real-world problem into an AI solution?

AI – Agents and Environments

Much (though not all!) of AI is concerned with **agents** operating in **environments.**

Agent – an entity that *perceives* and *acts*

Environment – the problem setting



Fleshing it out

Fleshing it out

Performance – measuring desired outcomes

Environment – what populates the task's world?

Actuators – what can the agent act with?

Sensors – how can the agent perceive the world?

Arnold’s PEAS?

PEAS in a taxi

Automated taxi driver

Performance – Safe, fast, legal, comfortable trip, maximize profits

Environment – Roads, other traffic, pedestrians, customers

Actuators – Steering, accelerator, brake, signals, horn, display

Sensors - Cameras, sonar, speedometer, GPS, odometer, accelerometer, engine sensors, microphone/keyboard

Exercise - pick some PEAS

Group 1 Grocery store shelf-stocking robot	**Group 2** Twitter bot to flag hate speech
Group 3 Medical diagnosis system	**Group 4** Interactive English tutor

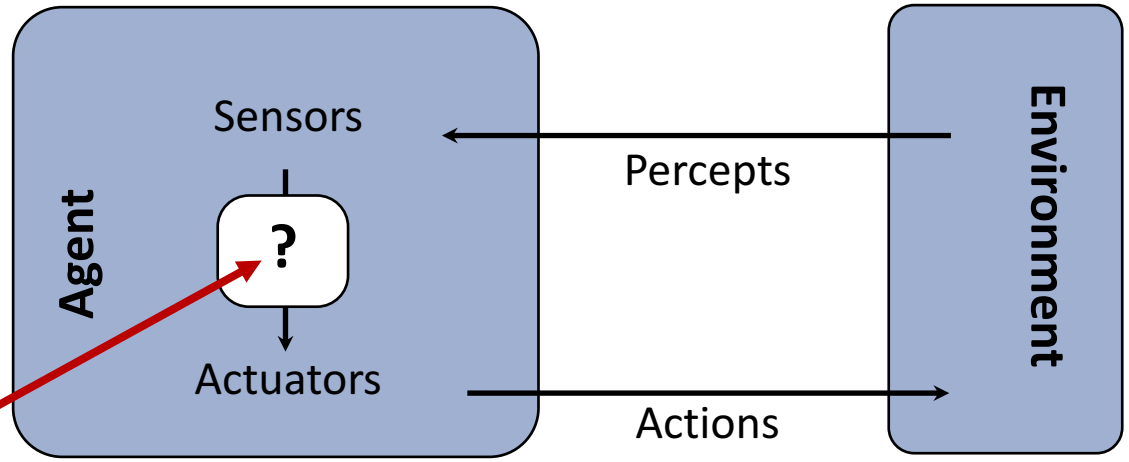
Performance
Environment
Actuators
Sensors

What makes an Agent?

Agent – an entity that perceives its environment through sensors, and acts on it with actuators.

Percepts are constrained by
Sensors + Environment

Actions are constrained by
Actuators + Environment



Agent Function – how does it choose the action?

What makes one rational?

Actually pretty simple:

A **rational agent** always acts to **maximize its expected performance measure**, given current state/percept

Not the same as **omniscience** (knowing the *actual* outcome of an action, instead of the *expected* outcome); can't account for a piano falling from the sky.

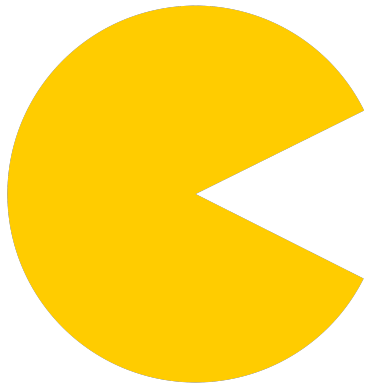
Our sample agents

Pacman

Percepts – squares around Pacman

Actions – move U/D/L/R

Environment – map with walls, dots, and ghosts



Spam detector

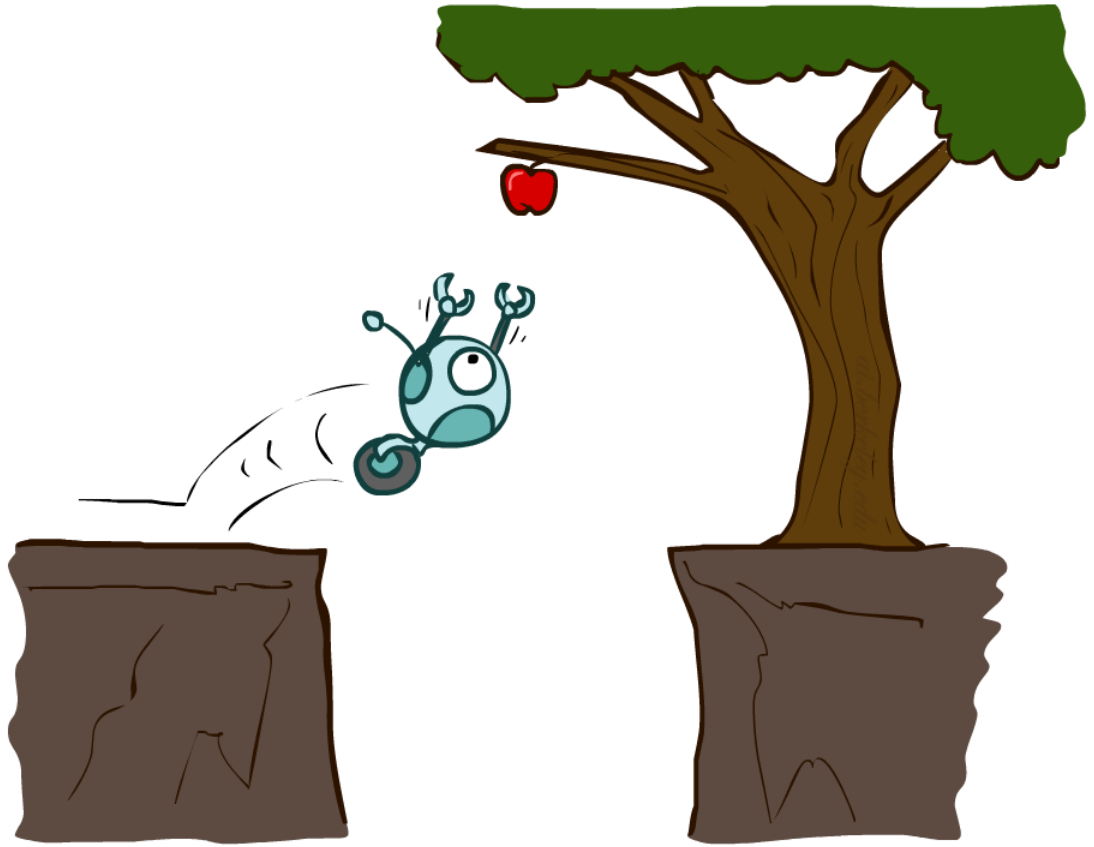
Percepts – sender, subject line, body of current email

Actions – mark Spam/Not Spam

Environment – your email inbox

Reflex Agents

- Reflex agents:
  - Choose action based on current percept (and maybe memory)
  - May have memory or a model of the world's current state
  - Do not consider the future consequences of their actions
  - Consider how the world IS



Simple Reflex Agent

Chooses action based on current percept alone

Pacman

Percepts – squares around Pacman

Actions – move U/D/L/R

Environment – map with walls, dots, and ghosts

[DEMOS]

Spam detector

Percepts – sender, subject line, body of current email

Actions – mark Spam/Not Spam

Environment – your email inbox

Agent function:

```
if hasNigerianPrince(body): return SPAM
else: return NOTSPAM
```

Model Reflex Agent

Chooses action based on current percept and internal model of environment

Pacman

Percepts – squares around Pacman

Actions – move U/D/L/R

Environment – map with walls, dots, and ghosts

Model – which squares have I eaten dots in?

Spam detector

Percepts – sender, subject line, body of current email

Actions – mark Spam/Not Spam

Environment – your email inbox

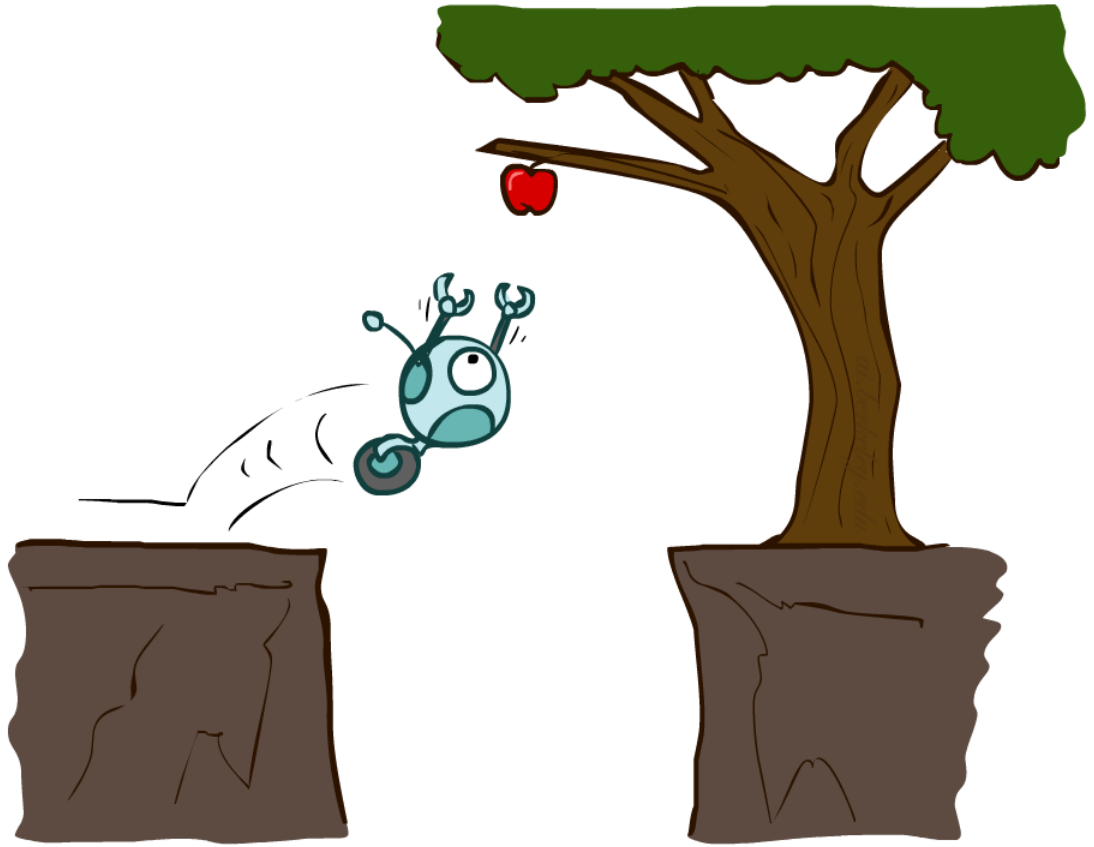
Model – per-sender message history

Agent function:

```
if haveExchangedEmails(sender): NOTSPAM
elif hasNigerianPrince(body): return SPAM
else: return NOTSPAM
```

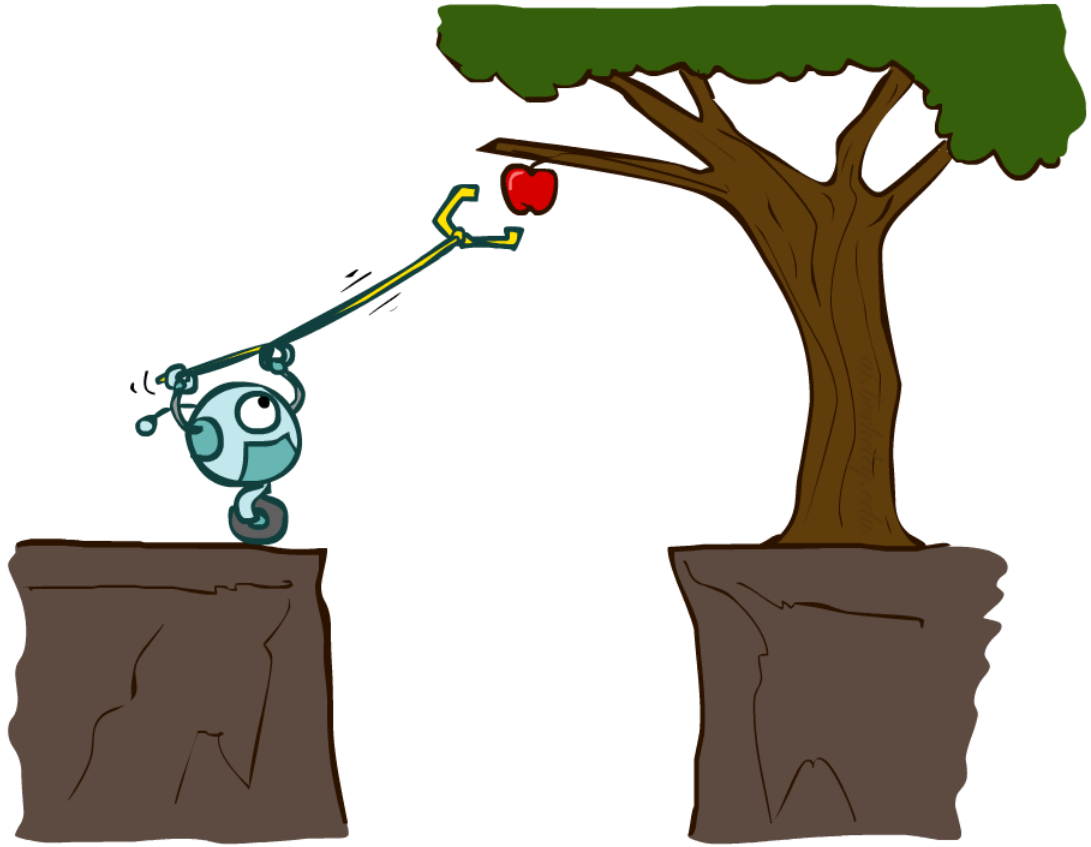
Reflex Agents

- Reflex agents:
  - Choose action based on current percept (and maybe memory)
  - May have memory or a model of the world's current state
  - Do not consider the future consequences of their actions
  - Consider how the world IS
- Can a reflex agent be rational?



Planning Agents

- Planning agents:
  - Ask “what if”
  - Decisions based on (hypothesized) consequences of actions
  - Must have a model of how the world evolves in response to actions
  - Must formulate a goal (test)
  - Consider how the world WOULD BE



Goal-based Agents

Chooses action (sequence) to get from current state to some goal

Pacman

Percepts – squares around Pacman

Actions – move U/D/L/R

Environment – map with walls, dots, and ghosts

Goal:

Spam detector

Percepts – sender, subject line, body of current email

Actions – mark Spam/Not Spam

Environment – your email inbox

Goal:

???

Utility-based Agents

Chooses action (sequence) to get from current state to some goal
with maximum utility along the way

Pacman

Percepts – squares around Pacman

Actions – move U/D/L/R

Environment – map with walls, dots, and ghosts

Goal:

…in as short a path as possible!

Spam detector

Percepts – sender, subject line, body of current email

Actions – mark Spam/Not Spam

Environment – your email inbox

Goal:

Summary

Reflex agents

Act on current state (and maybe past)

Simple – current p

Model – current p

of rest of

Goal-based agents

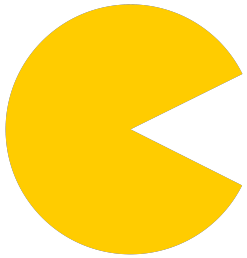
From current state to desired future

Goal only – find *any* action(s) to

ch the goal

st action(s) to

reach the goal



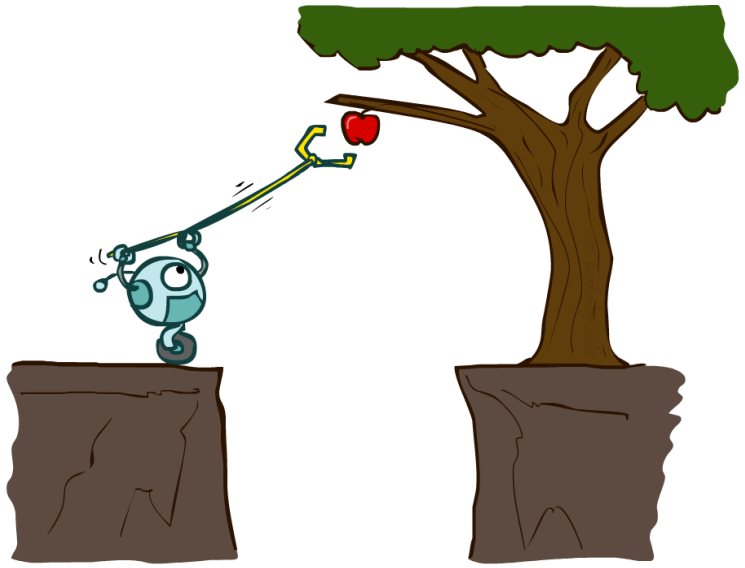
Can also have a **Learning Agent** – we'll talk about these later in the course!

AI – Agents and Environments

Much (though not all!) of AI is concerned with **agents** operating in **environments.**

Agent – an entity that *perceives* and *acts*

Environment – the problem setting



Kinds of task environments

6 common properties to distinguish tasks (not exhaustive)

- **Fully observable** vs **Partially observable**
- **Single agent** vs **Multiagent**
- **Deterministic** vs **Stochastic**
- **Episodic** vs **Sequential**
- **Static** vs **Dynamic**
- **Discrete** vs **Continuous**

Fully observable vs partially observable

Fully observable – agent is able to sense everything in the environment

ACROSS

1 See 24-Across

6 They radiate outward from an earthquake's epicenter

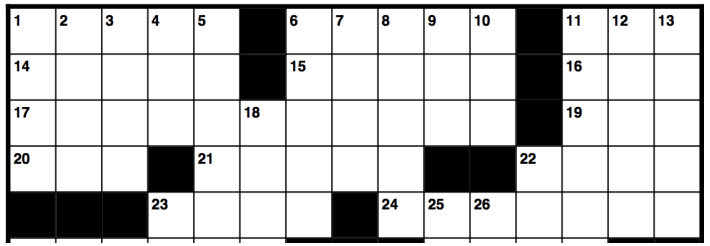
11 The "F" of "T.G.I.F.!": Abbr.

45 ___ fire under (urged to take action): 2 wds.

47 Daniel Defoe's "Robinson ___"

49 Vibrations caused by earthquakes

52 Low in fat

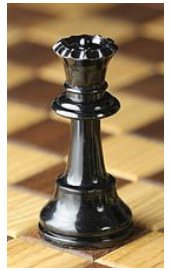


Partially observable – noisy, inaccurate, or incomplete sensors

Single agent vs Multiagent

Single agent – self-explanatory

Multiagent – task involves more than one agent, each with its own performance measure



May be **competitive** (measures are opposed)

or **cooperative** (measures align)

Deterministic vs Stochastic

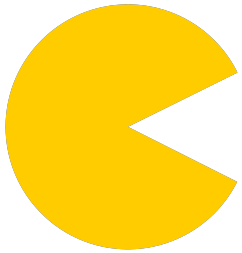
Deterministic – next state of the world is fully determined by current state + agent action

Stochastic – it's not deterministic

Episodic vs Sequential

Episodic – Each step/decision is independent of the previous ones

Sequential – Each step/decision affects later ones



Static vs Dynamic

Static – world doesn't change while agent is choosing an action

Dynamic – decision time matters!

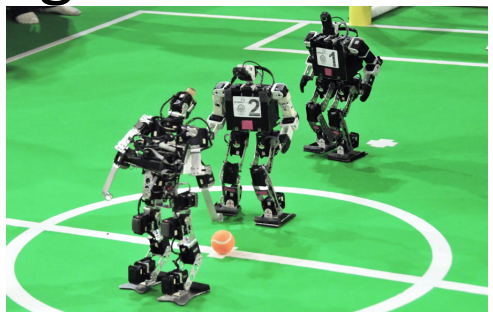
Discrete vs Continuous

Discrete – possible states/actions are distinct; world changes discretely

Continuous – states/actions take on continuous values

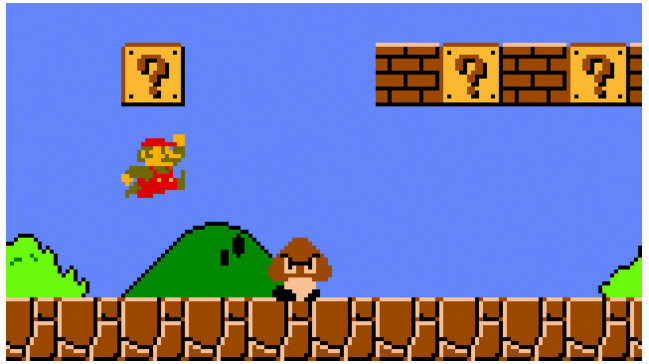
Bonus: level of agent knowledge

Environment is **known** – agent knows the “rules” of the world



Environment is **unknown** – agent has partial knowledge of the rules

???



Kinds of task environments

6 common properties to distinguish tasks (not exhaustive)

- **Fully observable** vs **Partially observable**
- **Single agent** vs **Multiagent**
- **Deterministic** vs **Stochastic**
- **Episodic** vs **Sequential**
- **Static** vs **Dynamic**
- **Discrete** vs **Continuous**

Exercise: classify some real tasks!

These help determine how to approach problems

Static -> can focus on getting really high accuracy/utility

Dynamic -> trade some utility for higher efficiency (speed!)

Episodic -> reflex agent with a great model

Sequential -> need a goal-oriented agent

Stochastic -> need robustness to uncertainty/failure (robots!)

Deterministic -> can focus on efficiency and exactness (Internet crawler)

Next up

Defining search problems – how to choose the right action sequence?

Uninformed search approaches – simple reflex agents for searching