

Check autograder
visibility settings

CS 232: Artificial Intelligence

Fall 2026

Prof. Carolyn Anderson
Wellesley College

Reminders

- ♦ HW 1 due
- ♦ HW 2 will be released Tuesday
- ♦ Set up Python 3.14 virtual environment
- ♦ Reading for Tuesday: ANY Shadow Self
- ♦ I have help hours today from 3-4 in W405 nook
- ♦ Starting next week:
 - Mon 3-4 in W323
 - Fri 3-4 in W423
- ♦ Missing Semester shell workshop Tuesday 7-8pm in H402

What your CS classes didn't teach
you but everyone expects you to
know!

CS 001: MISSING
SEMESTER

X

Workshop #1: The Shell

Tuesday 9/8 @ 7-8 PM
SCI H402

Join us to learn more
about the shell, a
program that lets you
interact with your
computer's operating
system!



RSVP Here

Snacks will be provided!

Questions: rp104
accessibility@wellesley.edu

Jordan (2019)

Jordan (2019)

- ♦ Intelligent Infrastructure (II) - "a web of computers, data, & physical entities"
- ♦ Intelligence Augmentation (IA) "augment human intelligence & creativity."
- ♦ Human-Imitative Artificial Intelligence
"an entity possessing human-level intelligence."

Jordan (2019)

- ◆ Intelligent Infrastructure (II): *"a web of computation, data, and physical entities [...] that makes human environments more supportive, interesting, and safe"*
- ◆ Intelligence Augmentation (IA): *"computation and data are used to create services that augment human intelligence and creativity"*
- ◆ Human-Imitative Artificial Intelligence: *"realizing in software and hardware an entity possessing human-level intelligence"*

Jordan (2019)

Which of Jordan's categories do you think chatbots like ChatGPT fall into, and why?

Depends on how the user uses it

- marketed as human-imitative
- could be used as IA

IA : you're still the driver

Human-imitative : mimics human speech patterns
people use it as therapist etc.

Jordan (2019)

What kinds of course do you think would need to be added to a Computer Science major in order to prepare students for work on the types of systems Jordan discusses?

Ethics
Engineering
Neuroscience
Philosophy — basis of knowledge
Accessibility
History of Science
How to interpret uncertainty, probability

Task-based AI

Task-based AI

Our goal is to write programs that can solve tasks. This is sort of the goal of all computer science.

In AI, though, the tasks we focus on are ones that seem to require human intelligence. This is a moving standard- what seems impossible for a computer to solve one day may eventually become very easy.

AI Tasks

Almost all AI tasks can be grouped into one of three main categories:

- ✦ Search
- ✦ Classification
- ✦ Generation

Learning

How do people learn?

- ♦ Some tasks have a **critical acquisition period**:
 - Language acquisition
 - Vision
 - Music training
- ♦ Others can be consciously acquired

Example: Child Language Acquisition

Example 1



cj and ember manning liked



Gareth Roberts @garicgymro · 45m

Just overheard from two of my kids:

Osian (5;1): Look how I caught Mickey!

Eirwen (8;2): Do you mean caught?

Osian: ... yeah.

Eirwen: But you can keep saying caught!

Osian: Look how I caught him!

Example 2

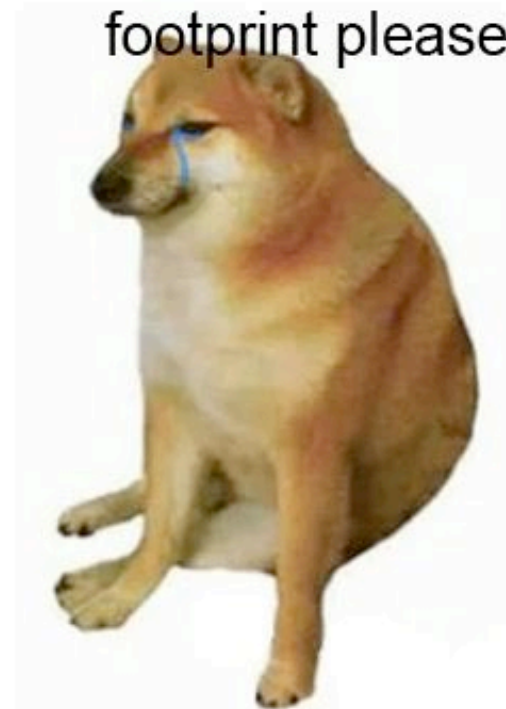


How do people learn?



human infants

with fast mapping, i
can learn the meaning
of a word in 3
exposures in my
human brain powered
by food



large language models

1 trillion
parameters and
a carbon
footprint please

How does AI learn?

- ♦ Symbolic AI
 - Program **rules** for the model
- ♦ Machine learning
 - ♦ Supervised learning
 - Give model **input/output pairs** to learn from
 - ♦ Reinforcement learning
 - Give model a **reward** function
 - ♦ Unsupervised learning
 - Model tries to **separate** data

Search

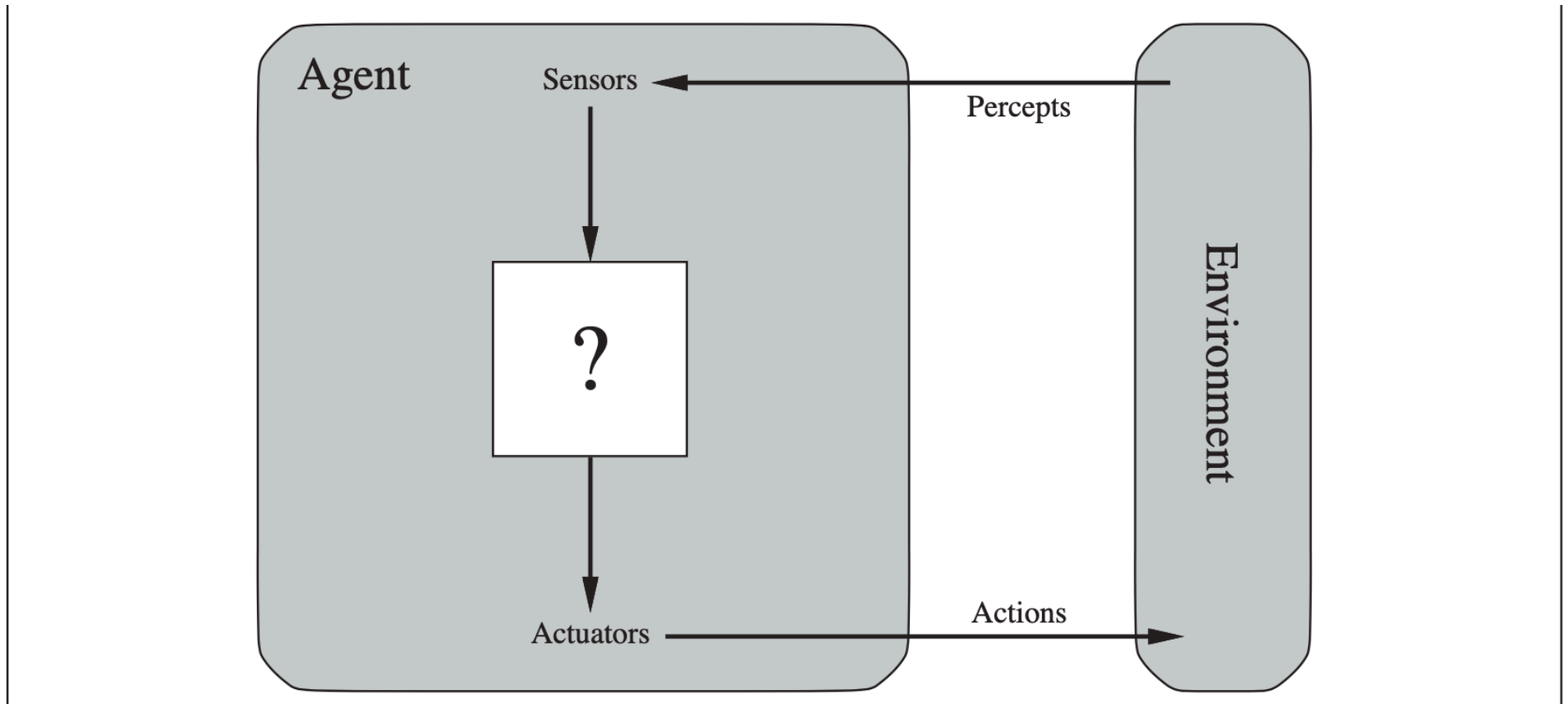
Agents

An **agent** is anything that can be viewed as perceiving its **environment** through **sensors** and acting upon that environment through **actuators**.

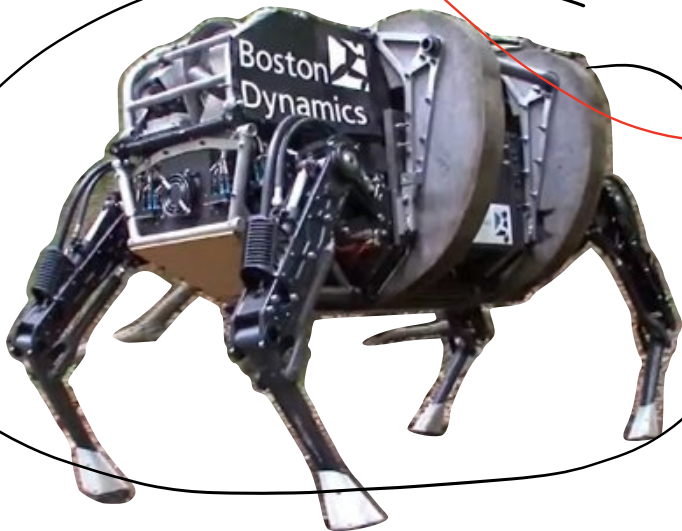
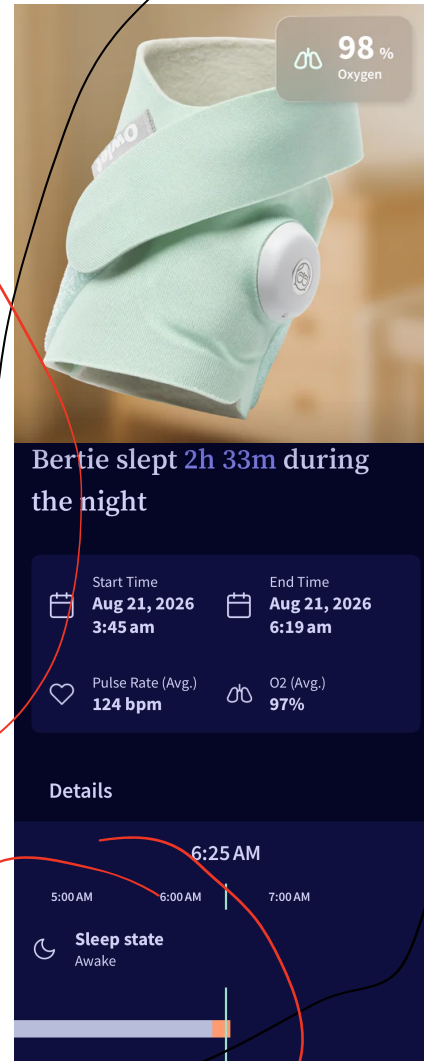
This is **extremely broad!**

Agents

An **agent** is anything that can be viewed as perceiving its **environment** through **sensors** and acting upon that environment through **actuators**.



Which of these are agents?



Defining an environment

Making problems out of the real world

If our goal is to tackle hard real-world problems, we will need a way of describing them.

Agents exist in **environments**. How do we define environments?

Defining the environment

Consider my cats Millie and Mango as agents.

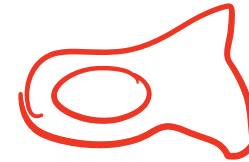


Defining the environment

We could define their environment very simply.

Here are two possible states:

State 1: `foodInBowl = true`



State 2: `foodInBowl = false`



Defining the environment

Here are their actions in each state:

State

Actions

1: foodInBowl = true

eat

2: foodInBowl = false

wail

Defining the environment

If we represent more information about the environment, we can make more nuanced choices:

State 1: humansHome = true, foodInBowl = true

State 2: humansHome = true, foodInBowl = false

State 3: humansHome = false, foodInBowl = true

State 4: humansHome = false, foodInBowl = false

Defining the environment

Here are their actions in each state:

State

- 1: humansHome = true,
foodInBowl = true
- 2: humansHome = true,
foodInBowl = false
- 3: humansHome = false,
foodInBowl = true
- 4: humansHome = false,
foodInBowl = false

Actions

eat

wail

eat

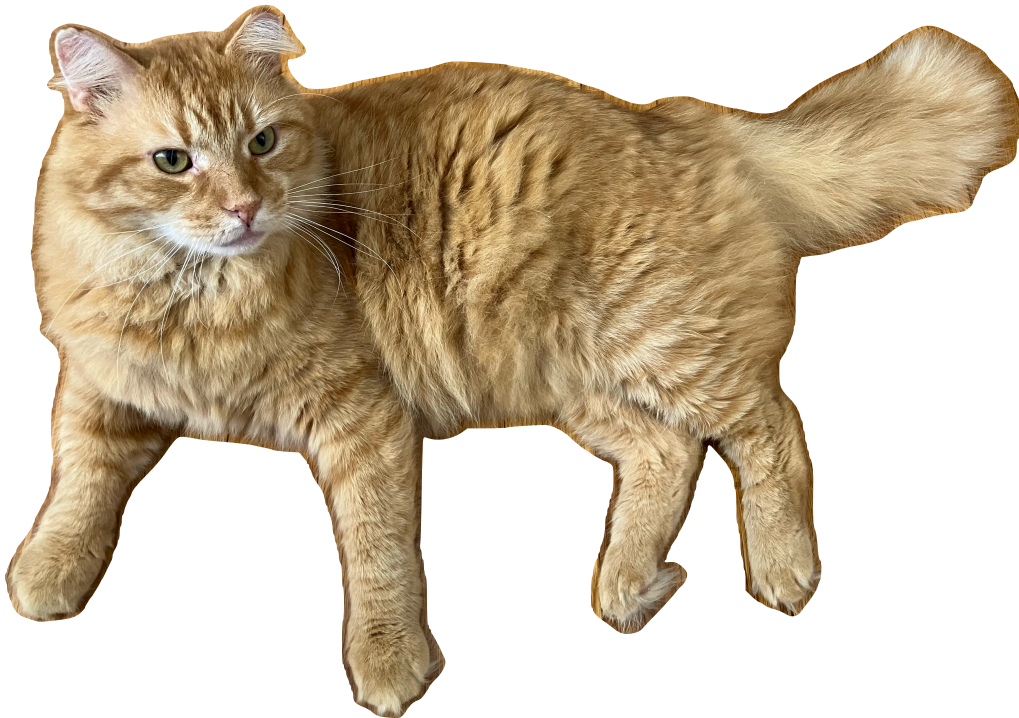
sleep

Rationality

Rational agents

A rational **agent** is an entity that acts to optimize a desired outcome (accomplish a task).

Mango: rational agent



Millie: sweet goofball



Rational agents

Given a goal, an AI agent must decide what the best action to take is in order to reach this goal.

For complex tasks, this can mean:

- ♦ gathering information
- ♦ coming up a set of possible actions
- ♦ weighing the best action
- ♦ acting
- ♦ updating and adapting based on changes to the environment

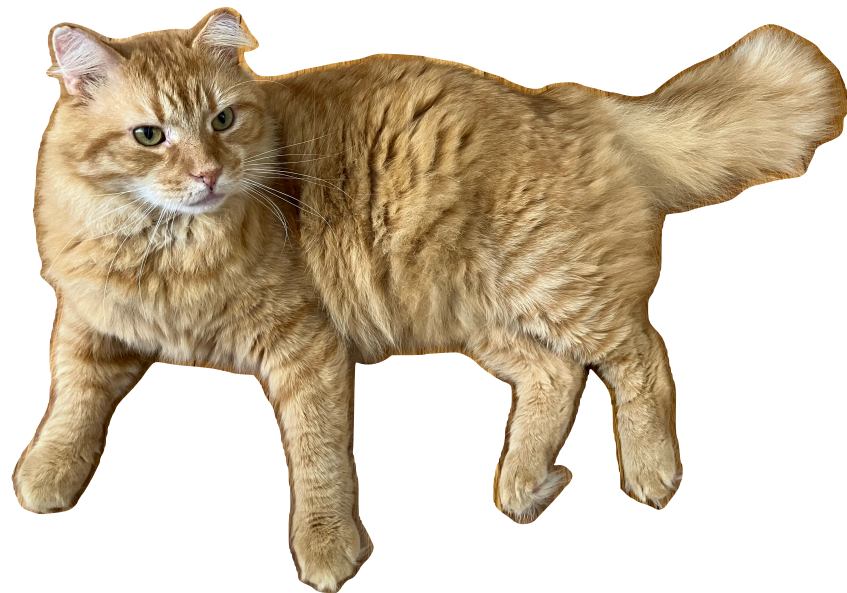
Defining Goals

In a search problem, our first step is to **define a goal**.

Mango has two goals in life:

1) eat

2) sleep



Actions

Given a goal, an AI agent must decide which actions to take.

Mango has several available actions:

- eat
- sleep
- wait
- store of birds

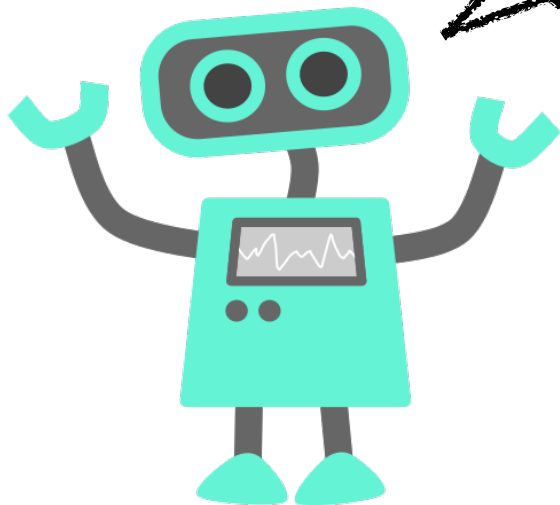
Actions

Not all of these are possible in every environment:

- ♦ eat - successful only if food available
- ♦ sleep - always successful
- ♦ wail - always successful, but most effective if an audience present
- ♦ stare at birds - successful only if birds available

Assessing Consequences

Given a goal and a set of actions, an AI agent must decide which action to take by weighing their **consequences**.



Assessing Consequences

Some actions have costs as well as consequences!

eat:

—

sleep:

wail:

stare at bird:

Assessing Consequences

Some actions have costs as well as consequences!

eat:

- achieves goal of devouring food
- updates environment by removing food

sleep:

- achieves goal of sleeping
- replenishes energy

wail:

- chance of human adding food to the bowl, if human is available and sufficiently convinced
- cost: lack of dignity

stare at bird:

- absolutely no consequences
- cost: some energy

Solutions

- ✦ A **solution** is a sequence of actions from the initial state to a goal state.
- ✦ A solution is **optimal** if no other solution has a lower cost.

Transitions

When an agent takes an action, they may transition to a new state (the environment may change).

Agent Complexity

Reflex agent: decides what to do based only on current state.



is there food? then I
will devour it!

Defining the environment

Here are their actions in each state:

State

Actions

1: foodInBowl = true

eat

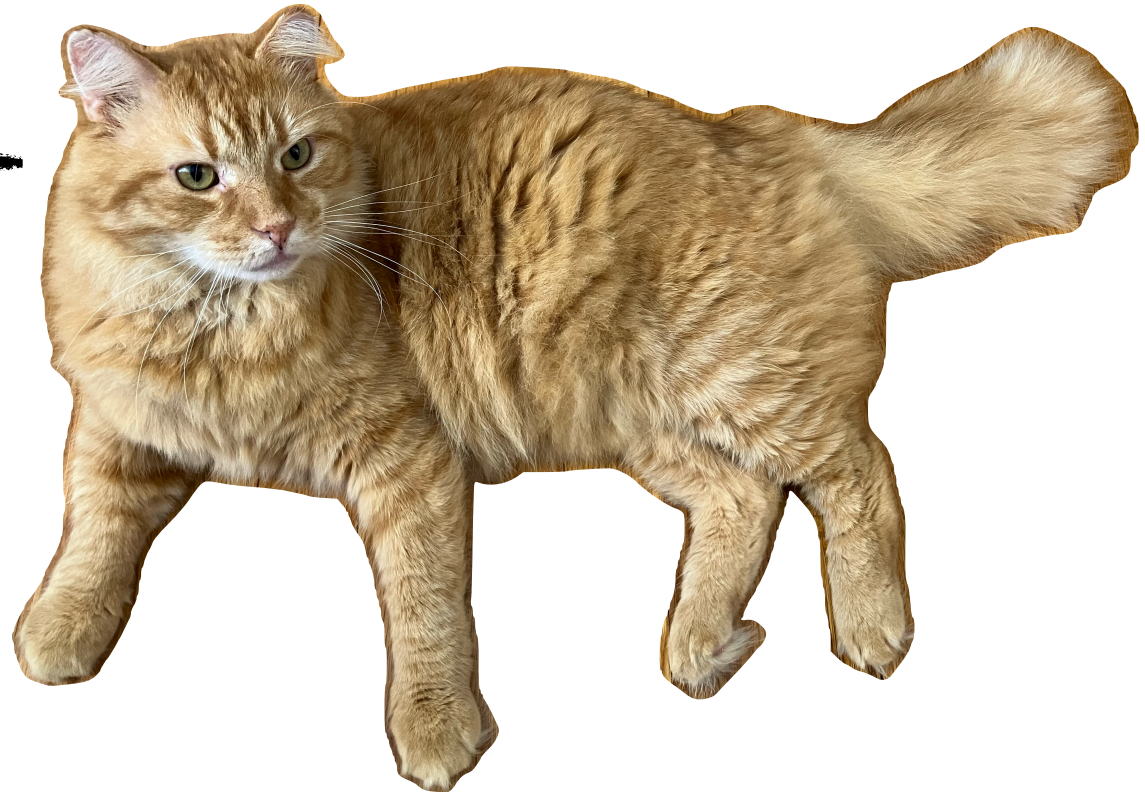
2: foodInBowl = false

wail

Agent Complexity

Problem-solving agent: capable of considering a sequence of actions that form a path to a goal state (planning ahead to find a solution).

if I wail, the human
might refill my bowl.
then I can eat more.



Defining Search Problems

Designing a State Space

The real world is absurdly complex.

To formulate a search problem, we need to **abstract** the world into a simpler representation: a *state space*.

Designing a State Space

The real world is absurdly complex.

To formulate a search problem, we need to **abstract** the world into a simpler representation: a *state space*.

We define a state space in terms of *states* and *actions*, which typically lead to transitions between states.

Formulating a Search Problem

Decide:

- Which properties of the world matter and how to represent them

initial state goal state(s) intermediate states

- Which actions are possible and how to represent them

set of actions transition model

- Which action is next

path cost function

These decisions greatly affect the size of the search space and therefore the speed of the search.

Formally Defining a Search Problem

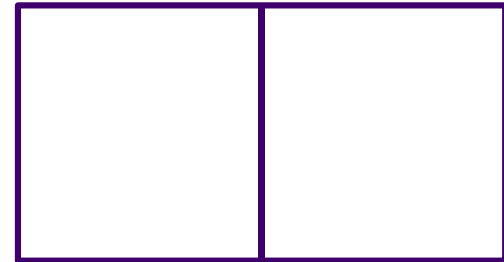
1. States : a set S
2. An initial state $s_i \in S$
3. Actions: a set A
 $\forall s \text{ Actions}(s) = \text{the set of actions that can be executed in } s$
4. Transition Model: $\forall s \forall a \text{ Result}(s, a) \rightarrow s_r$
 s_r is called a successor of s
5. Path cost (performance measure): must be additive.
6. Goal test : $\text{Goal}(s)$

Catworld

States: A state of the world describes which objects are in which locations.

In a simple two bowl version:

- ♦ the cat can be at one bowl at a time
- ♦ each bowl can be full or empty



Catworld

States: A state of the world describes which objects are in which locations.

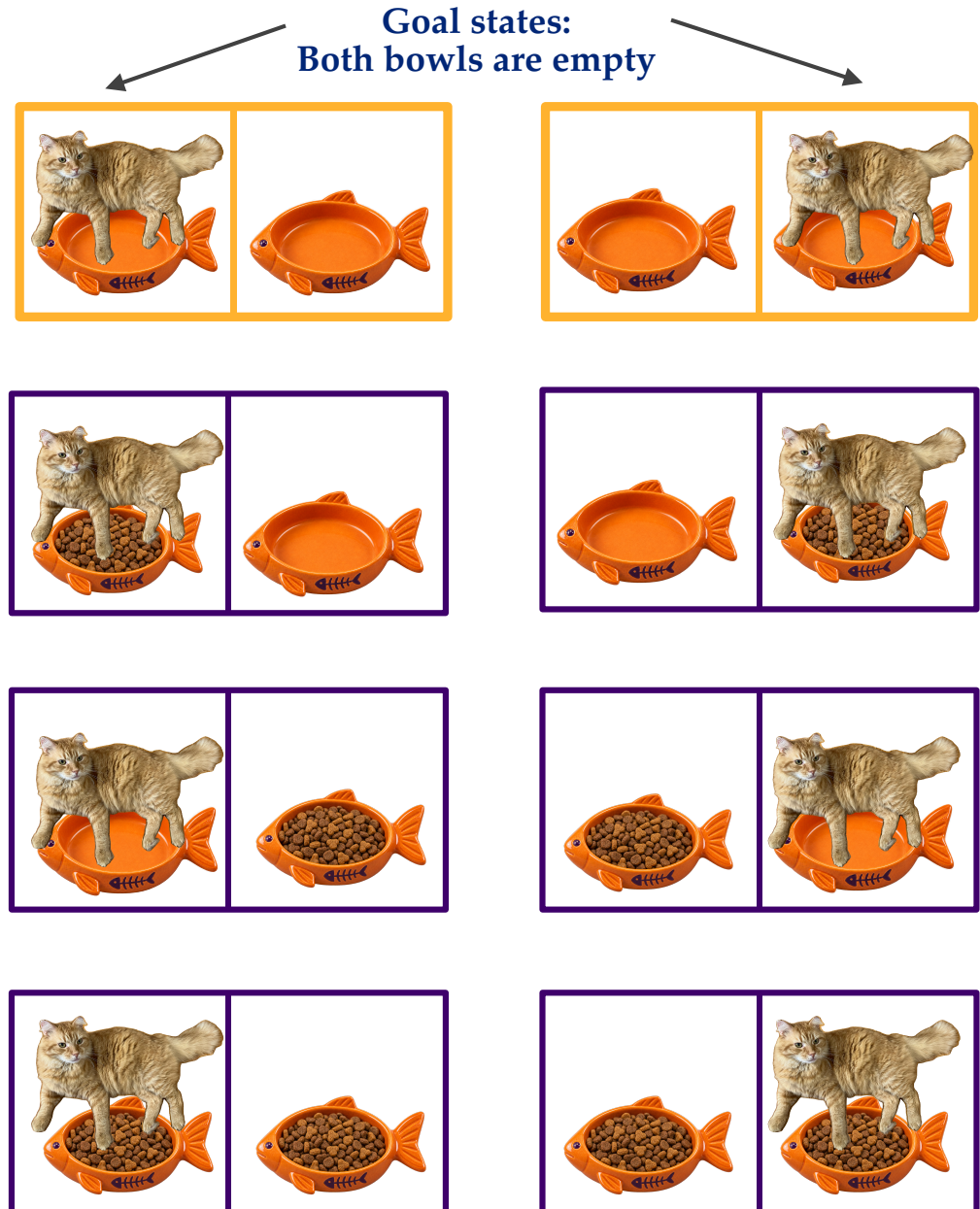
In a simple two bowl version:

- ♦ the cat can be at one bowl at a time
- ♦ each bowl can be full or empty

With n bowls, there are $n \times 2^n$ states.

2 cat positions $\times$ 2^2 bowl configurations = 8 states.

One state is designated the start state.



Catworld



Actions:

Transition:

Catworld



Actions:

Eat

Move Left

Move Right

(Move Up)

(Move Down)

Transition:

Eat — empties the current bowl

Move — moves in that direction, unless the agent hits a wall, in which case it stays put.

Catworld

Initial state



Actions:

Eat

Move Left

Move Right

(Move Up)

(Move Down)

Catworld

Initial state



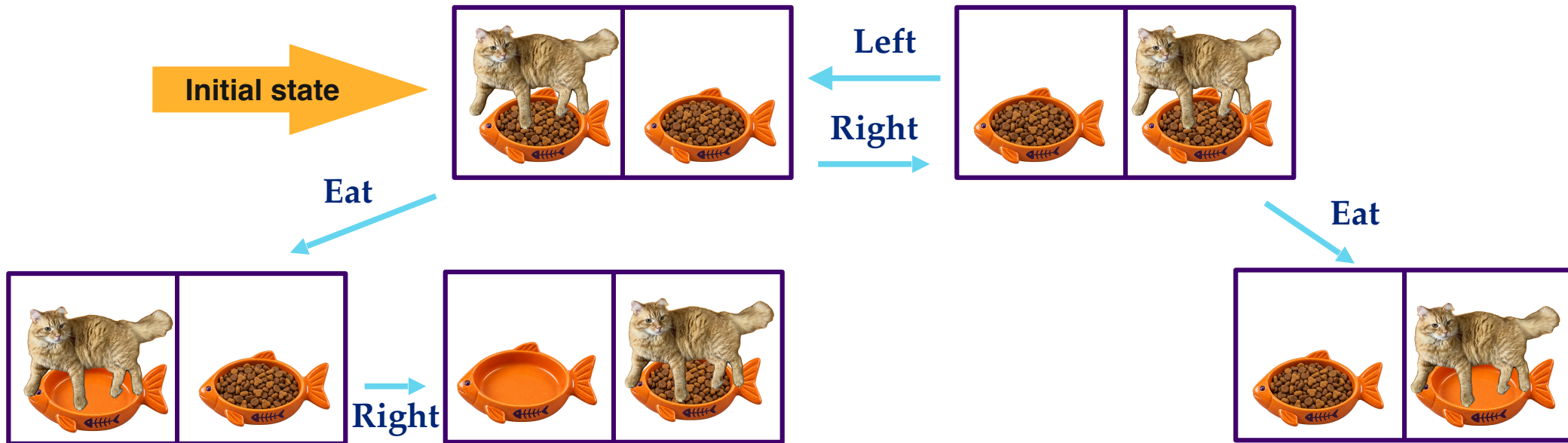
Right



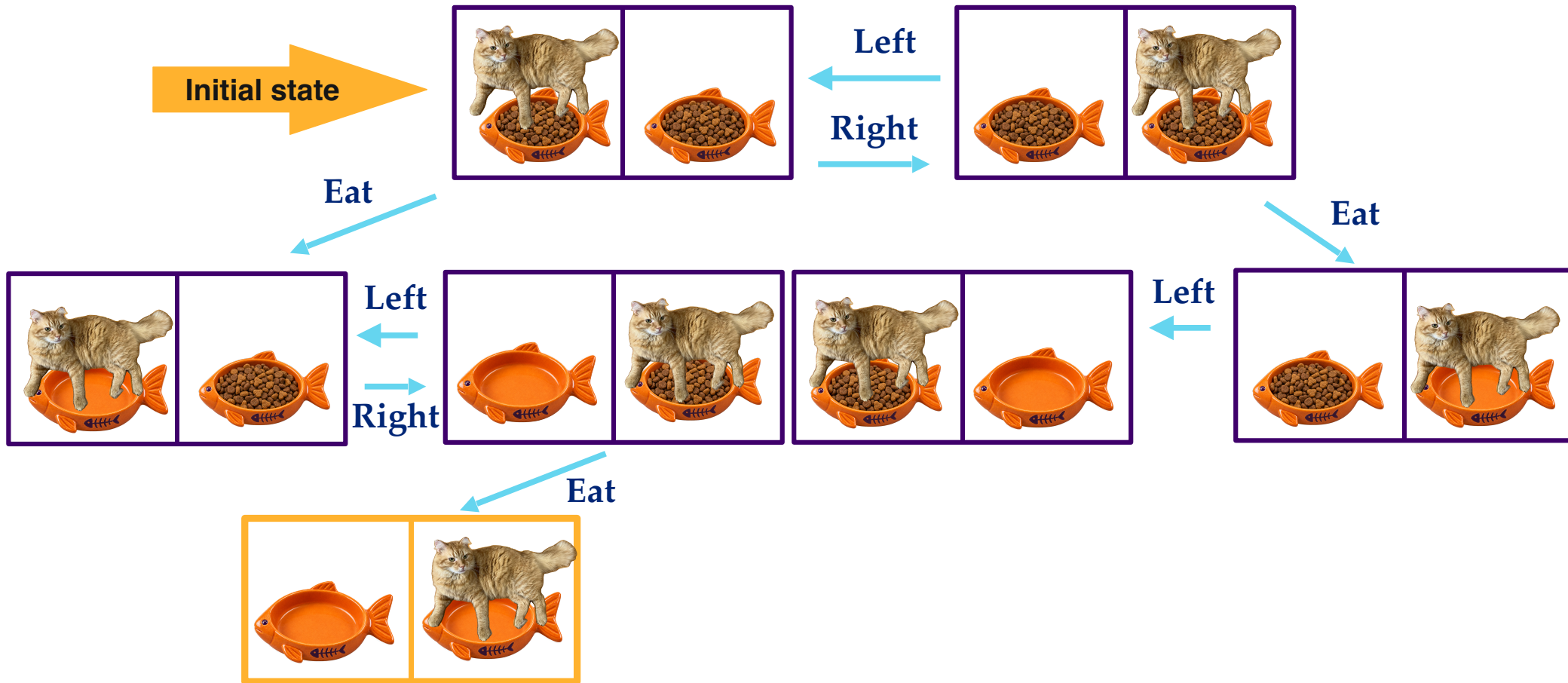
Eat

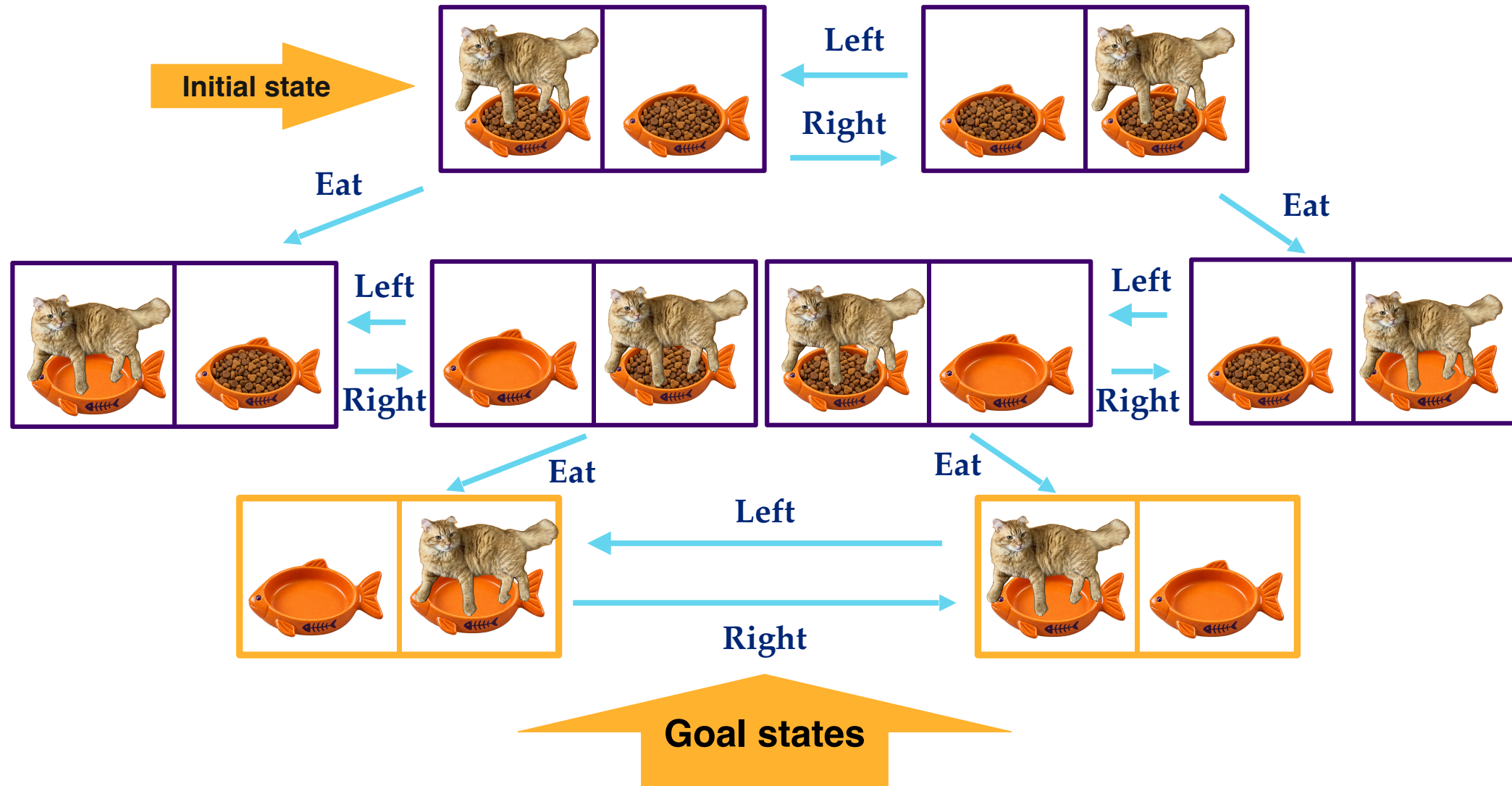


Catworld

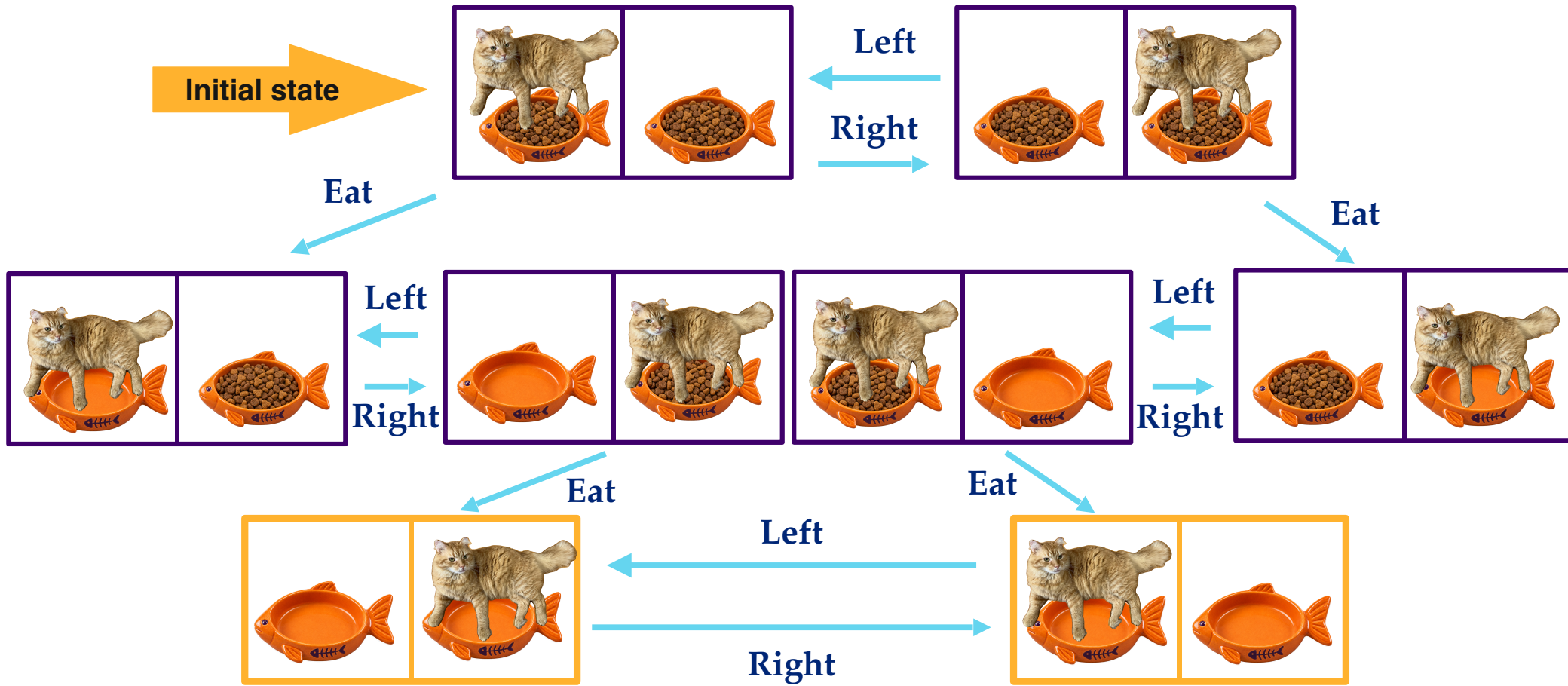


Catworld





Catworld



Solution:

A path from the initial
state to a goal state:
Eat → Right → Eat

Goal states

Useful Concepts

Evaluation

Evaluating Solvers
