CS 190 – Intro to Deep RL
4/8 – Simulated Environments and Reality

Prithviraj Ammanabrolu

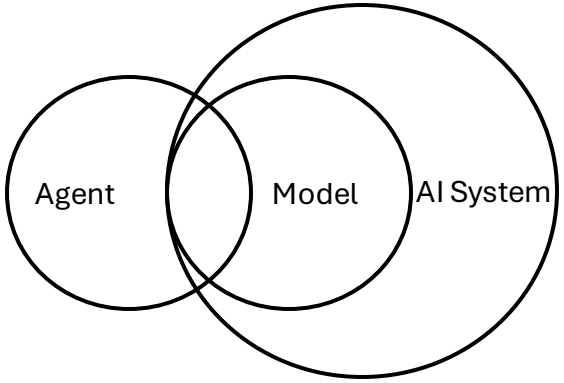
Recap: Components of an Agent

Required: Grounding, Agency (ability to act), Planning, Memory, Learning

Additional: Embodiment, Communication, World Modeling, Multimodality

Model vs AI System vs Agent: Rough Intuition

Model	AI System	Agent
GPT-4	ChatGPT	ChatGPT (computer use)
Forward passes of neural net	Mixing models together, model + scaffolding but no agency	Has agency + discussed components



Many software engineering abstractions and definitions exist.
All are roughly correct. Some are useful.

What is a simulation?

- Before you get all excited about simulating reality you need to understand verifiability and what makes a simulation useful
- Simulation : RL :: Dataset : Supervised Learning

Why do we need simulations?

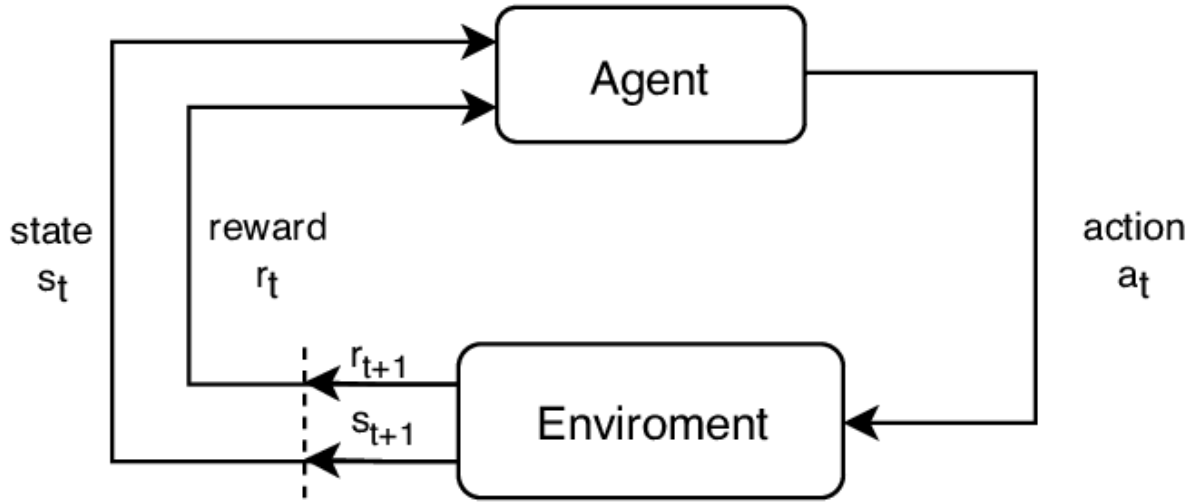
- Most tasks have many ways you can do them, e.g. “do the laundry” → how many clothes, which machine, what detergent, etc. etc.
- You usually do not know the “global” optimal solution ahead of time but usually know when you are done
- So you need to explore! Find many solutions and compare to see which is most efficient

Why do we need simulations?

- Exploration directly in the real world is expensive, wear and tear of robots, excessive compute, danger to humans, etc.
- A bunch of simple rules can compose to create very complex systems - https://robinforest.net/post/cellular-automata/

Basic components of a simulation

- From an MDP perspective, it contains at least <S, A, T>
  - S = set of all states
  - A = set of all actions
  - T = transition matrix T: (S, A)→S



Basic components of a simulation

- S = set of all states
  - propositions that are true: you are in a house, door is open, knife in drawer
- A = set of all actions
  - take knife from drawer, walk through door
- T = transition matrix T: (S, A)→S
  - (you are in a house & door is open, walk through door) → you are outside

There are pre-conditions that need to be met to perform a certain action, and post-conditions that are true after

Basic components of a simulation

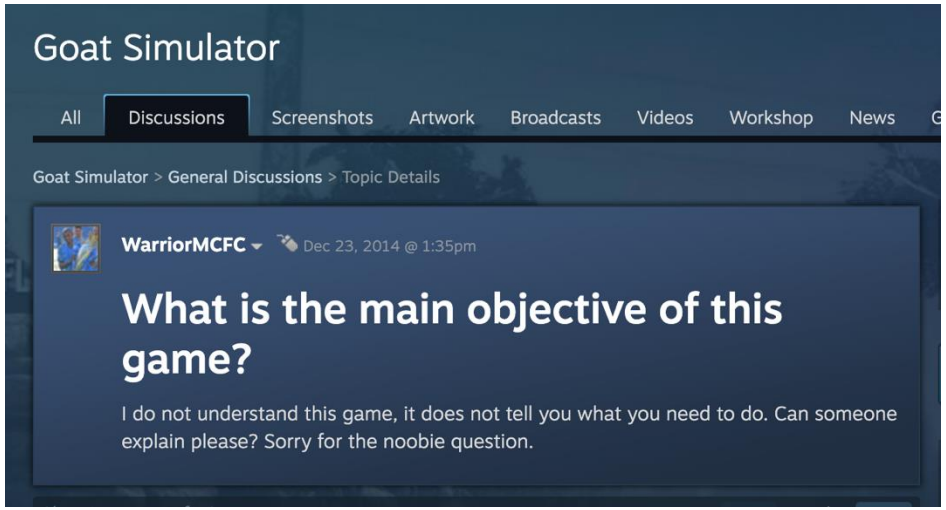
- S = set of all states
  - propositions that are true: you are in a house, door is open, knife in drawer
- A = set of all actions
  - take knife from drawer, walk through door
- T = transition matrix T: (S, A)→S
  - (you are in a house & door is open, walk through door) → you are outside

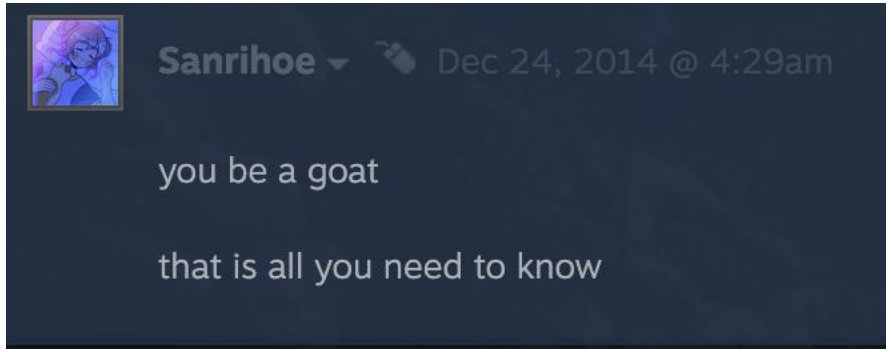
There are pre-conditions that need to be met to perform a certain action, and post-conditions that are true after

(hint: this is exactly what you learned about in intro to software eng but framed differently)

Basic components of a simulation

- From an MDP perspective, it contains at least <S, A, T>
  - S = set of all states
  - A = set of all actions
  - T = transition matrix T: (S, A)→S
- An explicit reward is technically not necessary (e.g. Goat Simulator)





Basic components of a simulation

- From an MDP perspective, it contains at least <S, A, T>
  - S = set of all states
  - A = set of all actions
  - T = transition matrix T: (S, A)→S
- A useful sim usually has an R, even just a +1 for the final goal state

Sim2real Transfer

- Does an agent trained in simulation transfer to reality (environment it is deployed in)?
- This is the same question as “does your model extrapolate out of distribution”?
- Answer is ???

Sim2real Transfer

- Does an agent trained in simulation transfer to reality (environment it is deployed in)?
- This is the same question as “does your model extrapolate out of distribution”?
- Answer is not really (for now) so rule of thumb is to make the sim as close to reality as you can

Dimensions of Complexity

- How to measure closeness to reality?
- First thing to think about is the research question you want to answer and the task you need to do.
- Many dimensions, we’ll focus on two

Cognitive Complexity

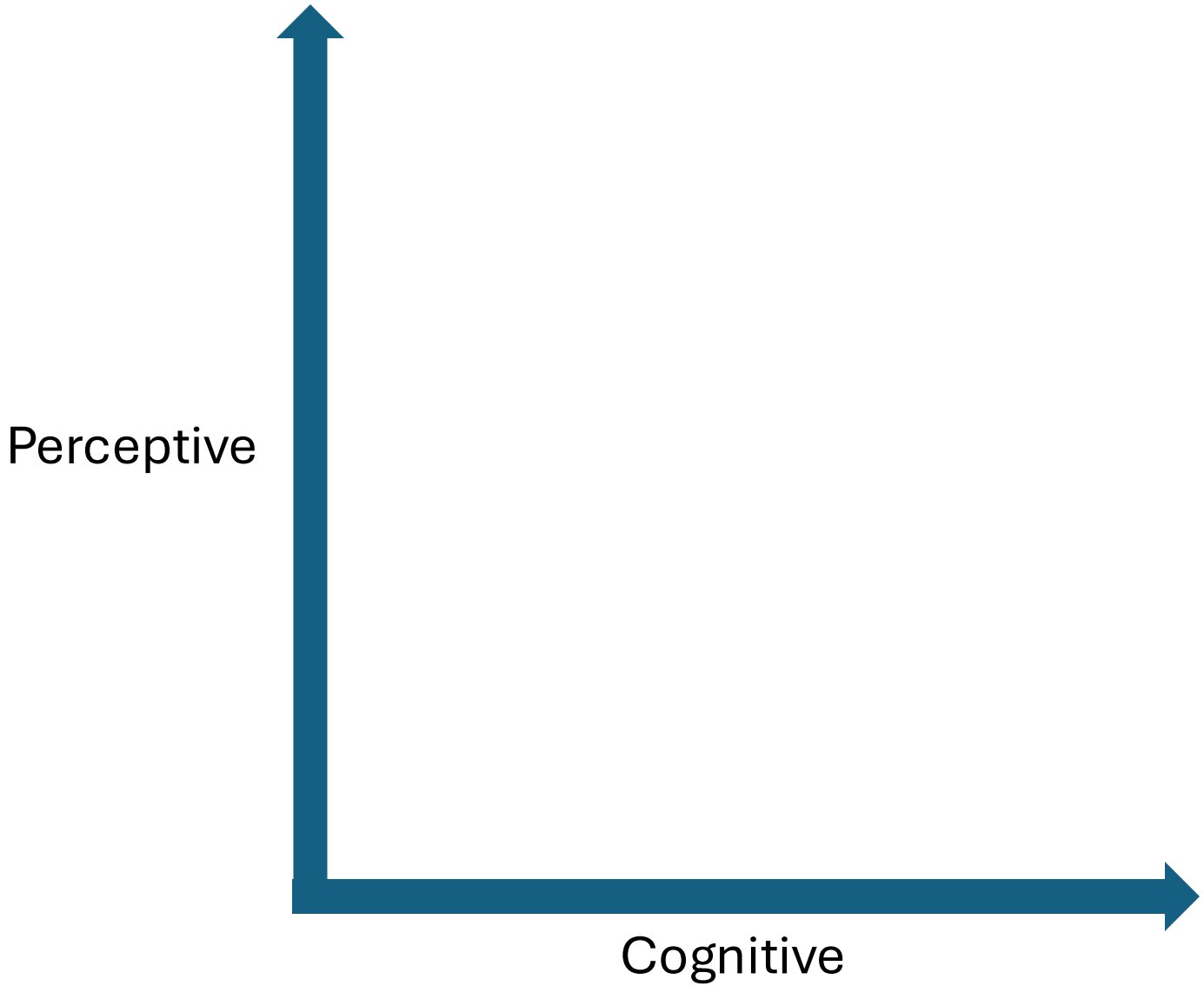
- Requires long chains of “reasoning”
- Think puzzles, math problems, moral dilemmas, etc

Perceptive Complexity

- Requires high levels of vision and/or precise motor skills
- Birdwatching, threading a needle, Where’s Waldo

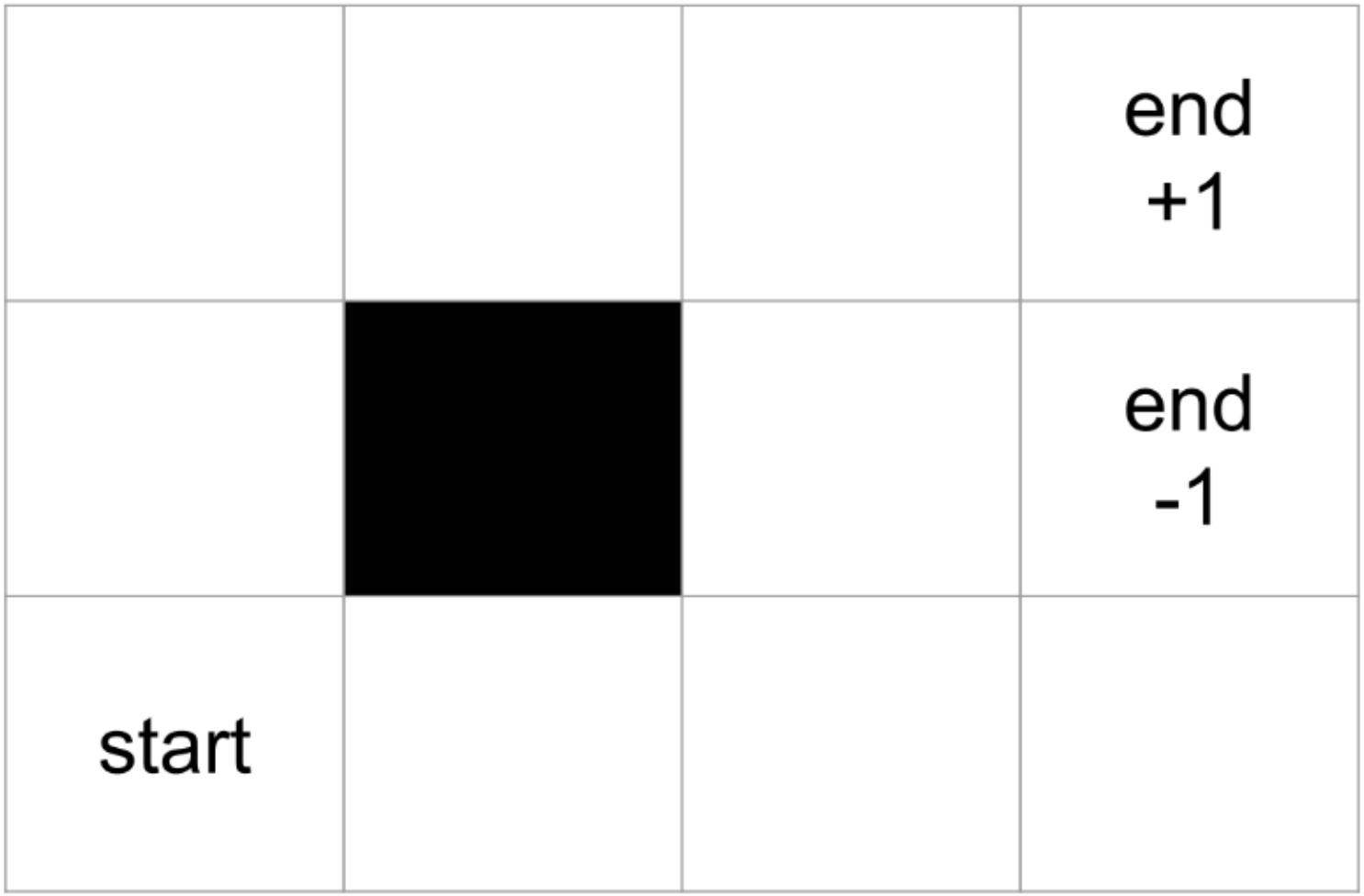
Reality

Matrix of Simulations



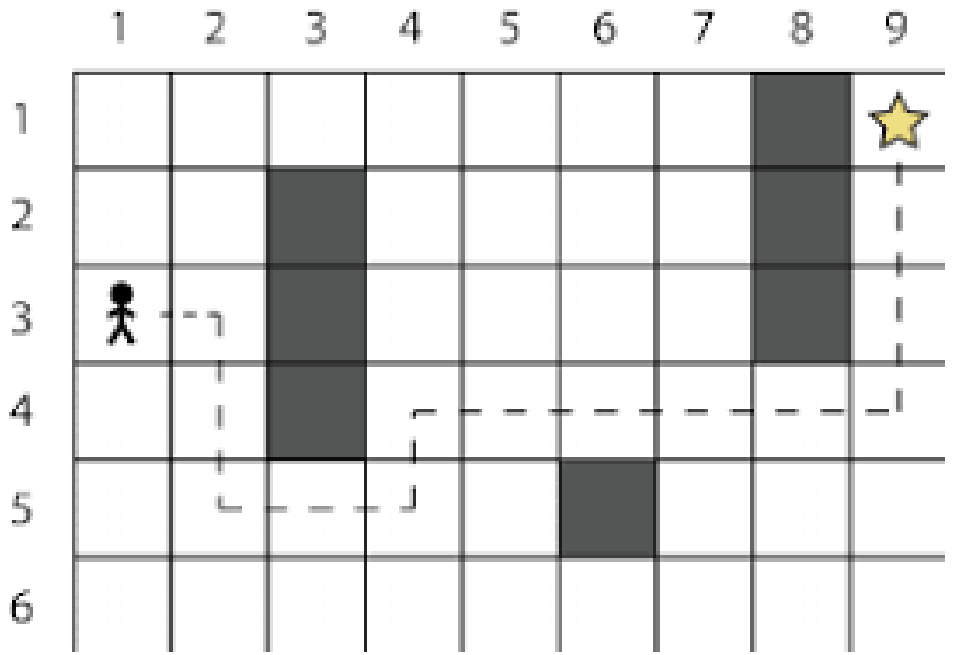
Low Perceptive, Low Cognitive

- Gridworld



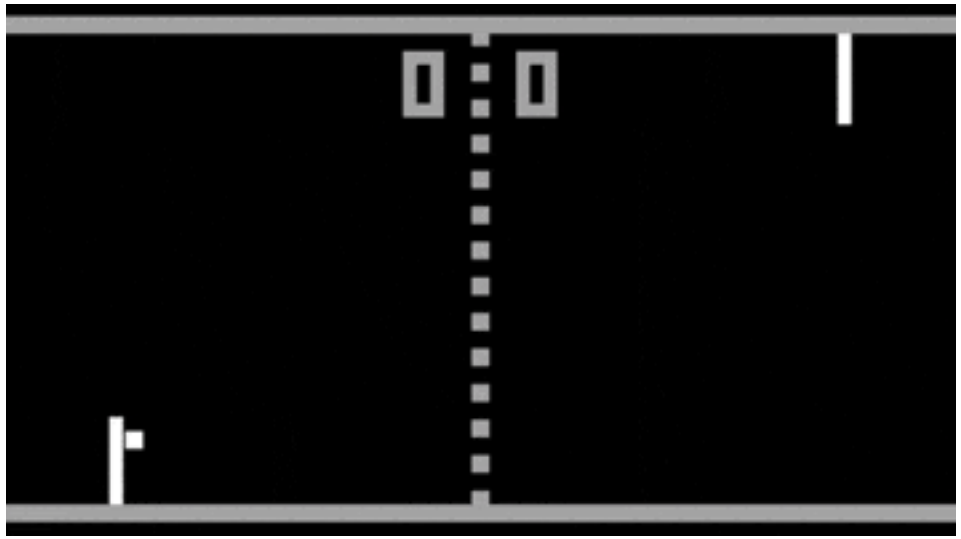
Low Perceptive, Low Cognitive

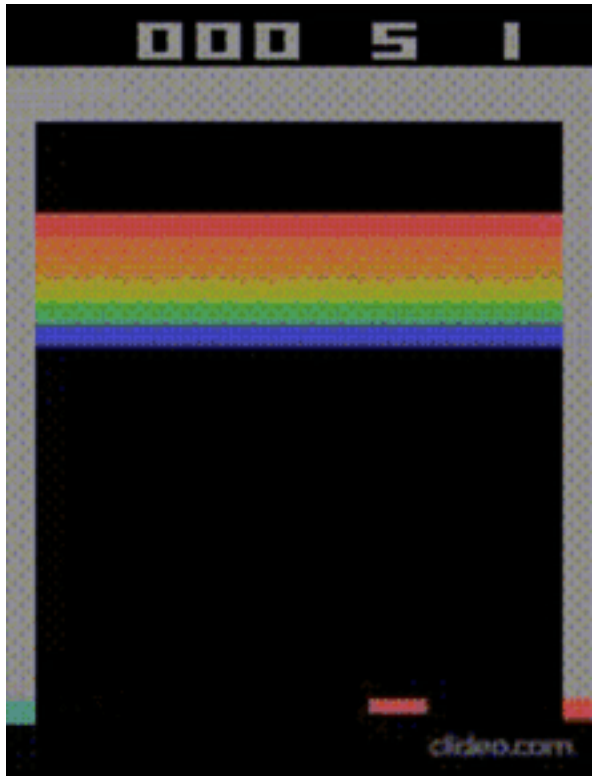
- Gridworld – seems simple but can arbitrarily scale. Can test algo generalization potential in controllable settings



Low Perceptive, Medium Cognitive

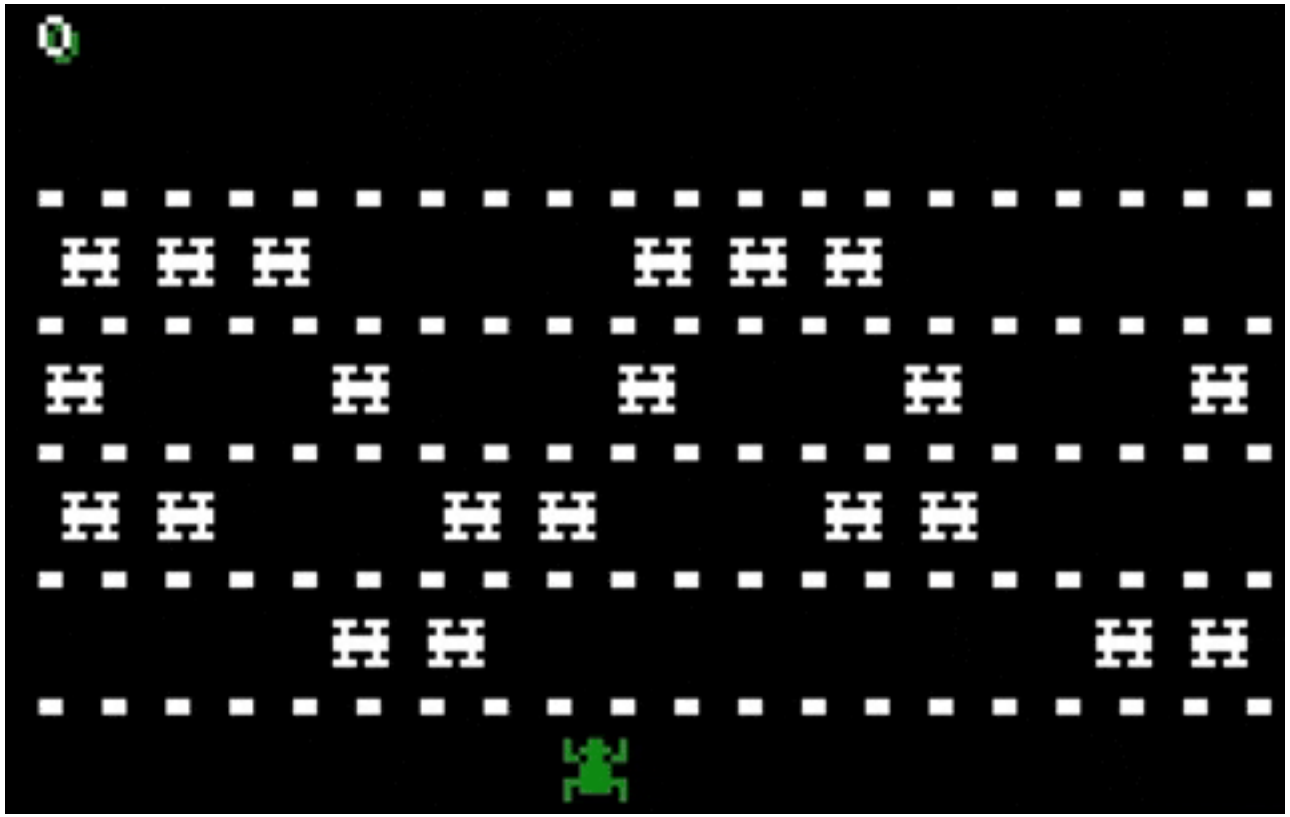
- Atari





Low Perceptive, Medium Cognitive

- Atari



Low Perceptive, High Cognitive

- Zork, NetHack

You are standing in an open field west of a white house.
There is a small mailbox here.

> *open mailbox*

Opening the small mailbox reveals a leaflet.

> *read leaflet*

Welcome to Zork!

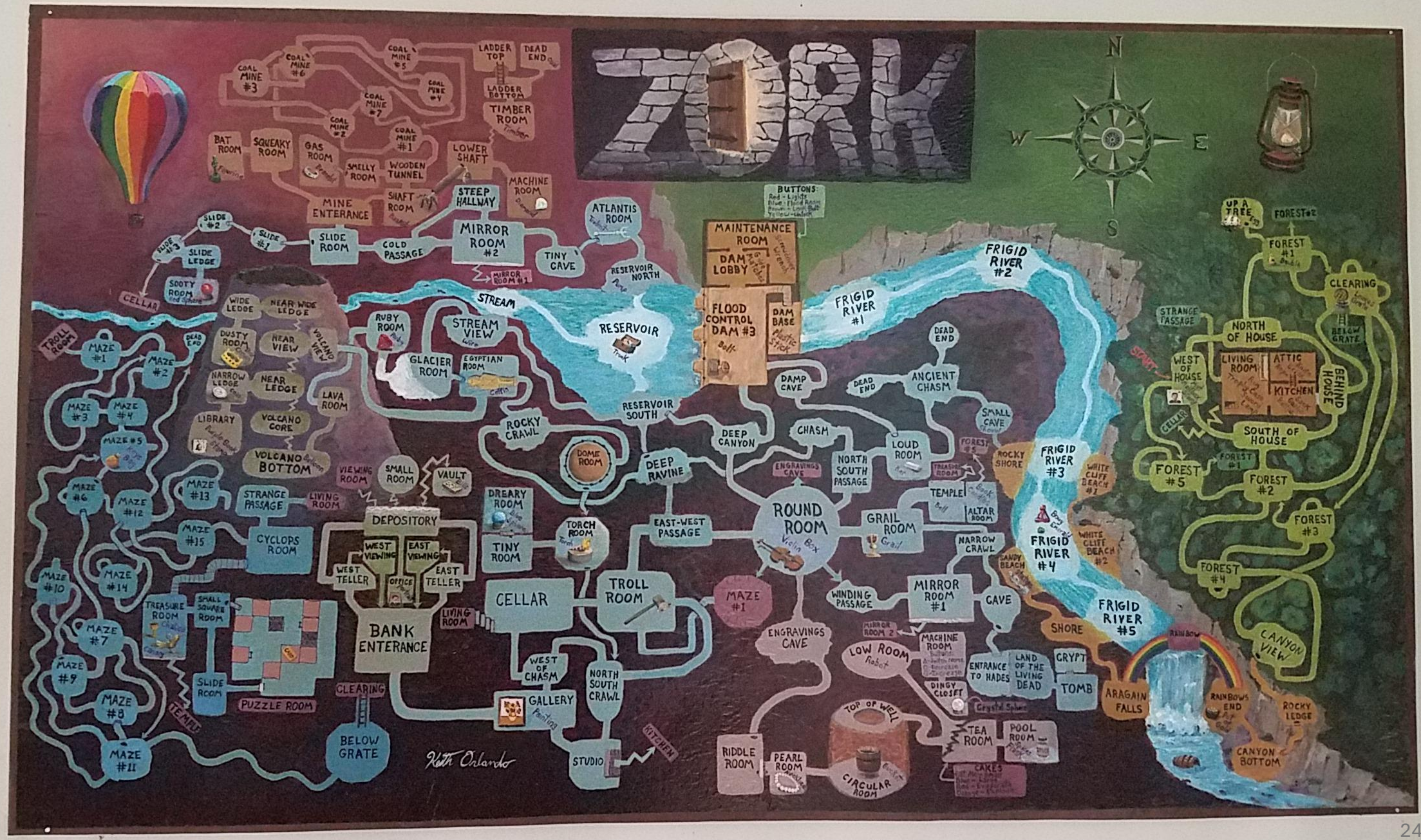
> *go north*

You are behind the white house. In one corner of the house there is a small window which is slightly ajar.

> *open window*

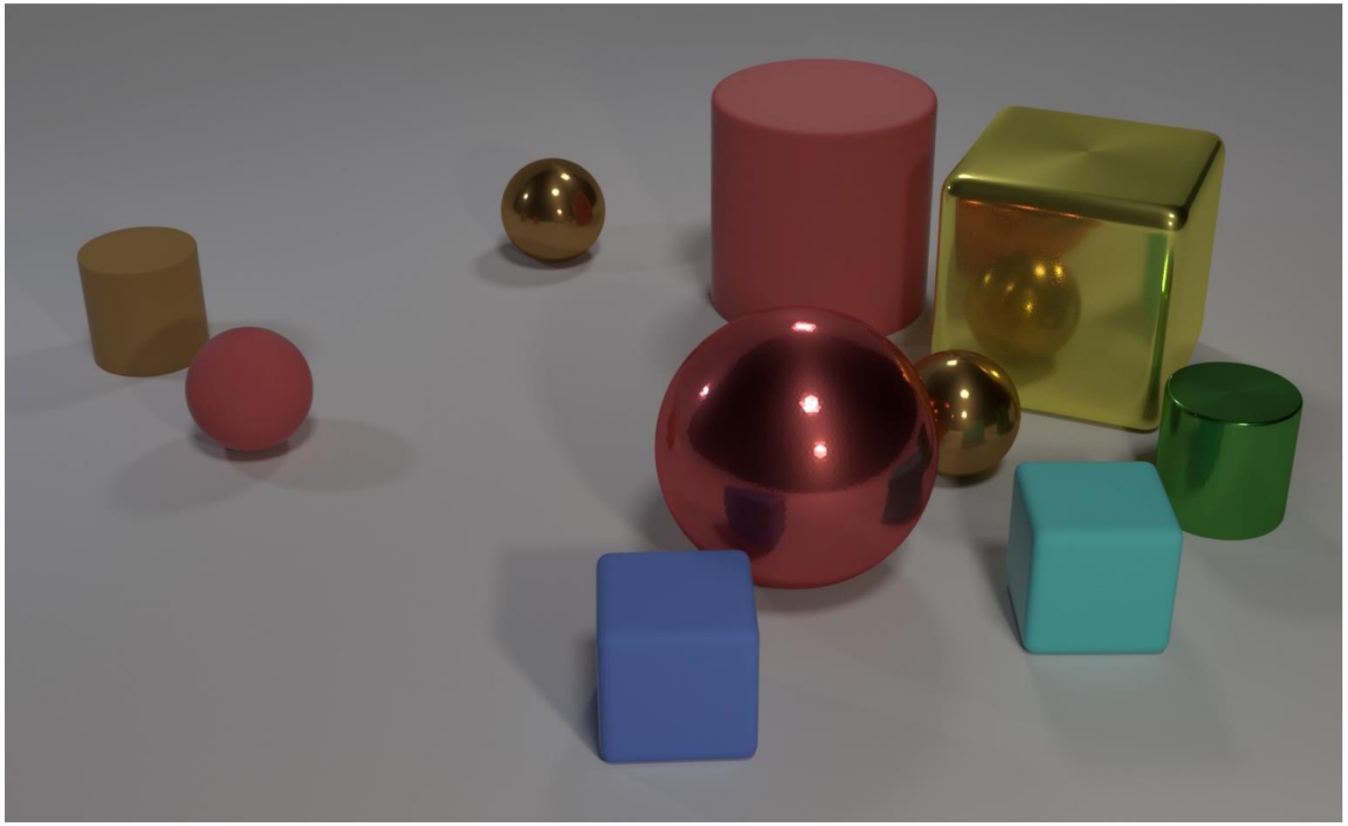
With effort you open the window far enough to allow entry.

> *enter house*



Hausknecht, Ammanabrolu, Alexandre-Cote, Yuan. *Interactive Fiction Games: A Colossal Adventure.* AAAI-20.

Medium Perceptive, Low Cognitive



Q: Are there an equal number of large things and metal spheres?

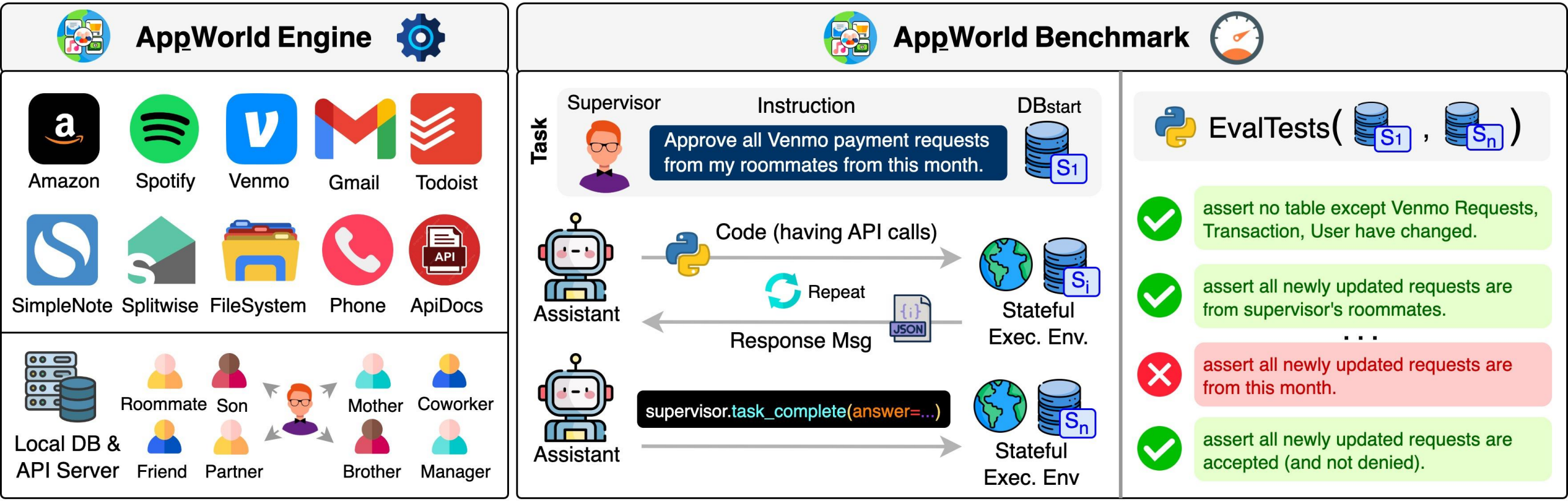
CLEVR: A Diagnostic Dataset for Compositional Language and Elementary Visual Reasoning. Johnson et al. 2017

Medium Perceptive, Medium Cogntive

- Ai2 THOR

Medium Perceptive, Medium Cogntive

- AppWorld (Trivedi et al. 2024, ACL)



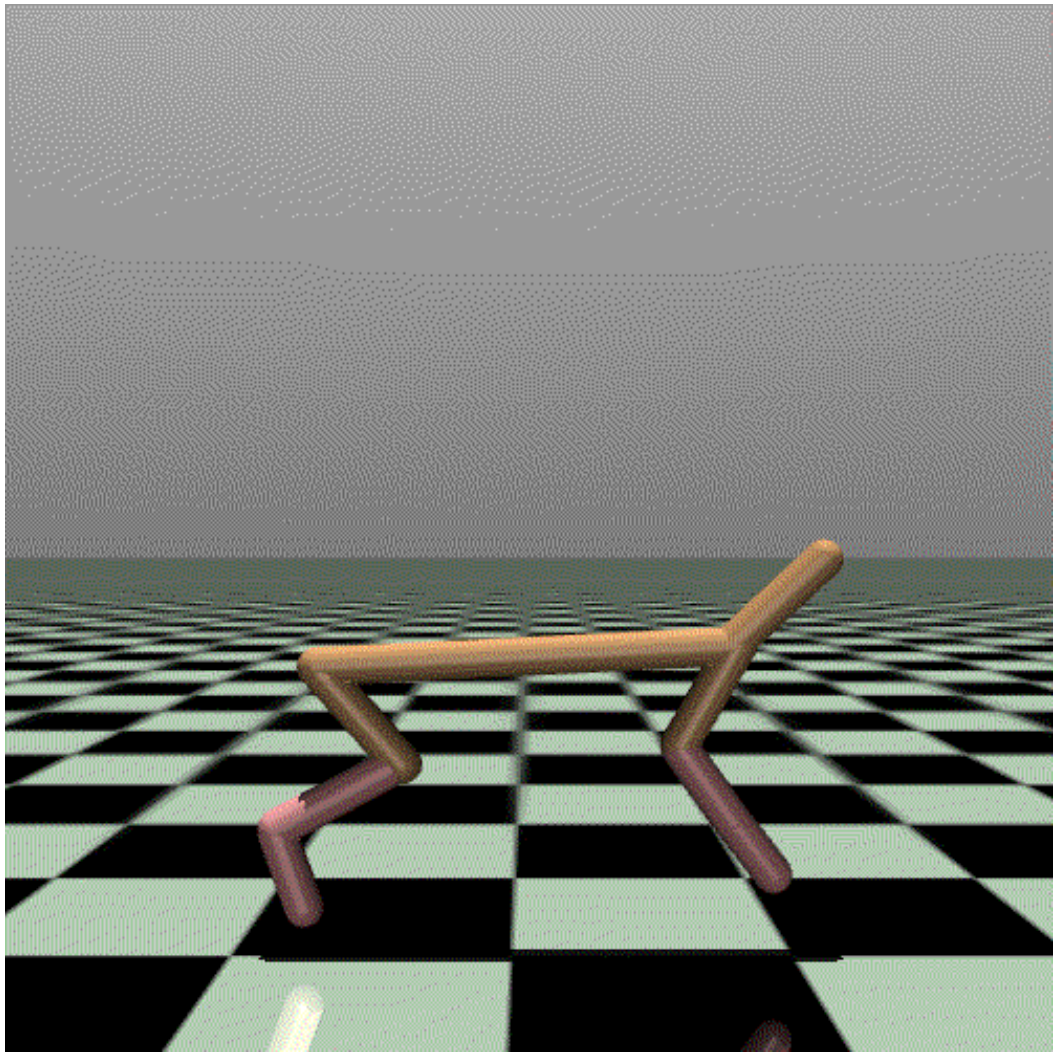
Medium Perceptive, High Cognitive

- Minecraft

Minedojo: Building Open-Ended Embodied Agents with Internet-Scale Knowledge. Fan et al. 2022.

High Perceptive, Low Cognitive

- Mujoco



High Perceptive, Medium Cognitive

- Habitat (Meta, Savva et al. 2019)

High Perceptive, High Cognitive

- Real world, whoever gets this sim first cheaply wins

Questions to think about

- For all the sims how were the dimensions of complexity related to size of state/action space?
- How were they related to how many steps you'd have to take before getting a reward? (aka reward sparsity)

Sim2real Transfer

- Which dimensions of complexity transfer more easily?
- Can you train on lower complexity and switch to a higher complexity?
- (world model holy grail) sims are very costly to make, can you just learn one?