Survey of recent progress in intelligent robotics

Jonathan Mugan
@jmugan

DeUmbra

Data Day Texas
January 27th, 2024

Slides at https://www.jonathanmugan.com/WritingAndPress/

Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence

- The Evolutionary Paths of Intelligence

Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence

- The Evolutionary Paths of Intelligence

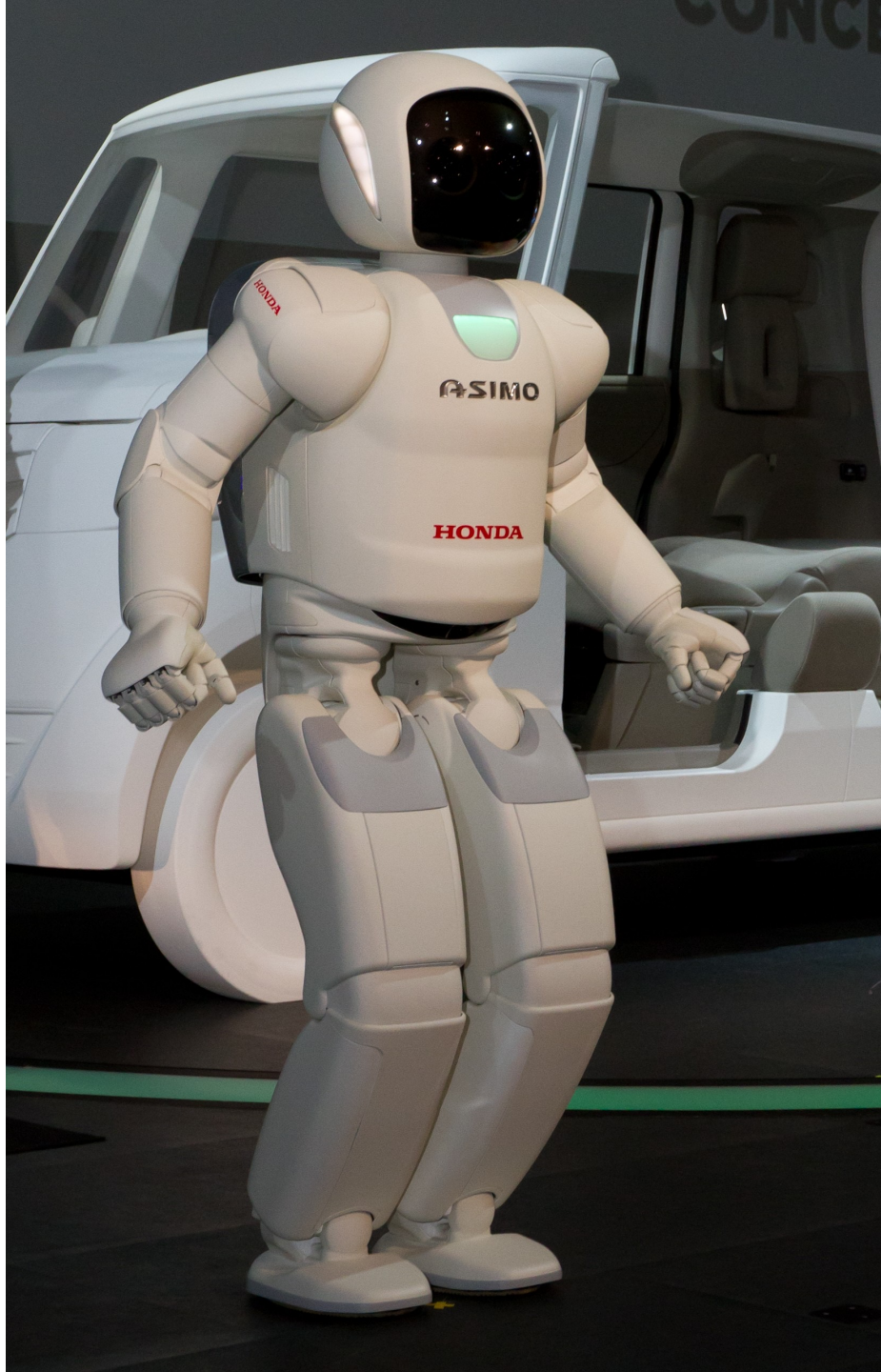
Hardware Advances: Humanoid Robots Are Arriving

The humanoid shape is important because our environment is built for humans.

I've been watching this space since ASIMO. We seemed to be at a plateau for almost 20 years, but now we may be at an inflection point.

Humanoid robots are finally starting to be produced in larger numbers.

https://asimo.honda.com/



Tesla Robot

Most human looking, and it moves its fingers!

https://twitter.com/DrJimFan/status/1705982525825503282

https://www.youtube.com/watch?v=cpraXaw7dyc

Video of it folding clothes
https://twitter.com/elonmusk/status/1746964887949934958

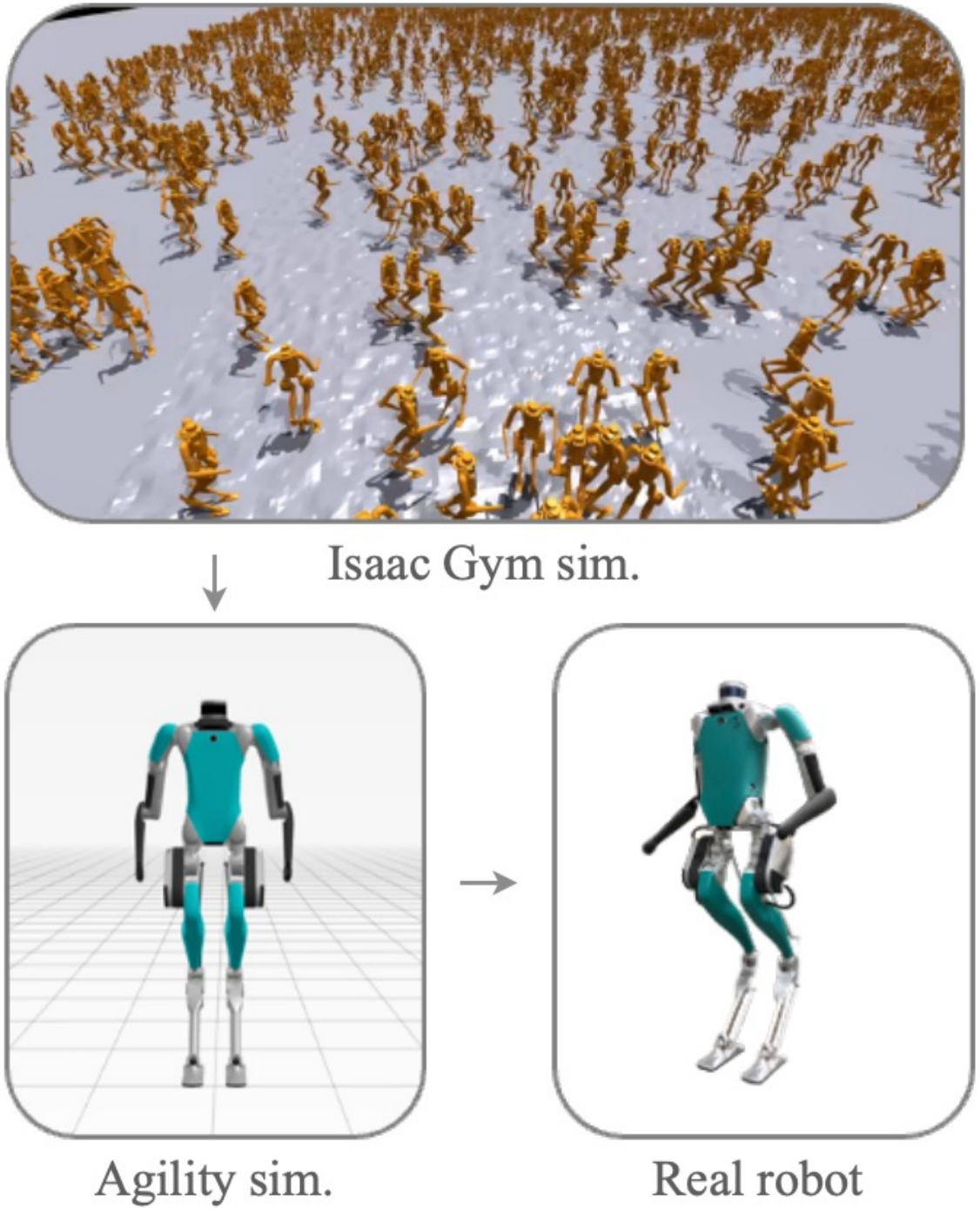
Photo by Benjamin Ceci, Public domain, via Wikimedia Commons

Berkeley

Locomotion based on reinforcement learning with causal transformer.
https://learning-humanoid-locomotion.github.io/

https://venturebeat.com/ai/uc-berkeleys-transformer-based-robot-control-system-generalizes-to-unseen-environments/

Image used with permission, thanks Ilija!
https://arxiv.org/abs/2303.03381



Agility Robots (at Amazon)

https://agilityrobotics.com/

- Cost $250k
- Still crazy expensive but getting into the realm of possibility for a medium sized business.

https://www.youtube.com/watch?v=ZWonAz7Kczs&list=WL&index=1&t=19s

video on next slide courtesy of Agility Robotics

Image by Agility Robotics, used with permission

Others

Unitree

- https://m.unitree.com/h1/
- Buy one for $150k! https://shop.unitree.com/products/unitree-h1

ALOHA

- Wheeled, but it cooks shrimp! https://twitter.com/zipengfu/status/1742602881390477771

Apptronik

- Right here in Austin!
- https://apptronik.com/
- https://apptronik.com/videos
- And maybe some day in space! https://www.reuters.com/science/humanoid-robots-space-next-frontier-2023-12-27/

Speaking of Austin,
Day Zero by C. Robert Cargill

A robot shootout on 2222, right on my way to work!

Image from ChatGPT for “robot shootout in Austin, Texas”

Others

Unitree

- https://m.unitree.com/h1/
- Buy one for $150k! https://shop.unitree.com/products/unitree-h1

ALOHA

- Wheeled, but it cooks shrimp! https://twitter.com/zipengfu/status/1742602881390477771

Apptronik

- Right here in Austin!
- https://apptronik.com/
- https://apptronik.com/videos
- And maybe some day in space! https://www.reuters.com/science/humanoid-robots-space-next-frontier-2023-12-27/

When I tell it to put the shootout on 2222 it puts it on 360, and in the 1970s?

Speaking of Austin,
Day Zero by C. Robert Cargill

A robot shootout on 2222, right on my way to work!

But autonomy is currently limited

Many of these robots are teleoperated.

In 2008, the movie Sleep Dealer depicted a future in which workers in other countries would control robots in the United States. Like Gattaca, it came out about 20 years too early for the concept to pop.

I'm not too worried about the teleoperation. Intelligence is advancing quickly with machine learning training in simulation, as we will see.

Image from Wikipedia, By Impawards.com, Fair use, https://en.wikipedia.org/w/index.php?curid=21544666

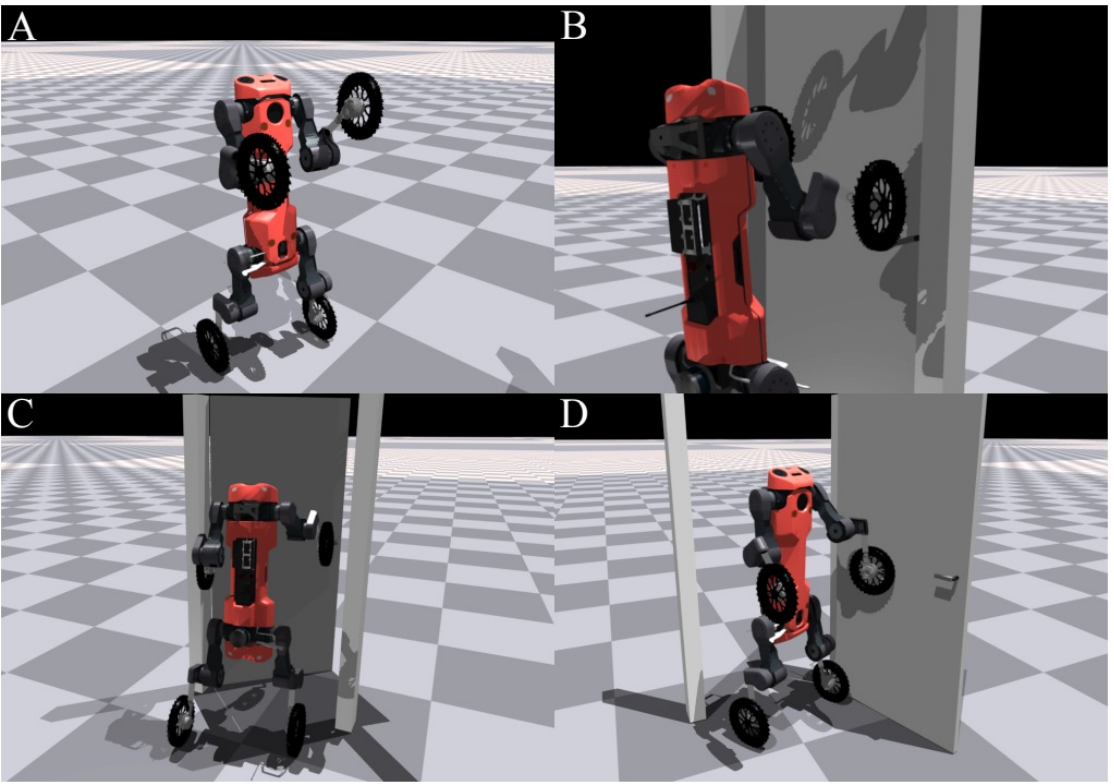
Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence

- The Evolutionary Paths of Intelligence

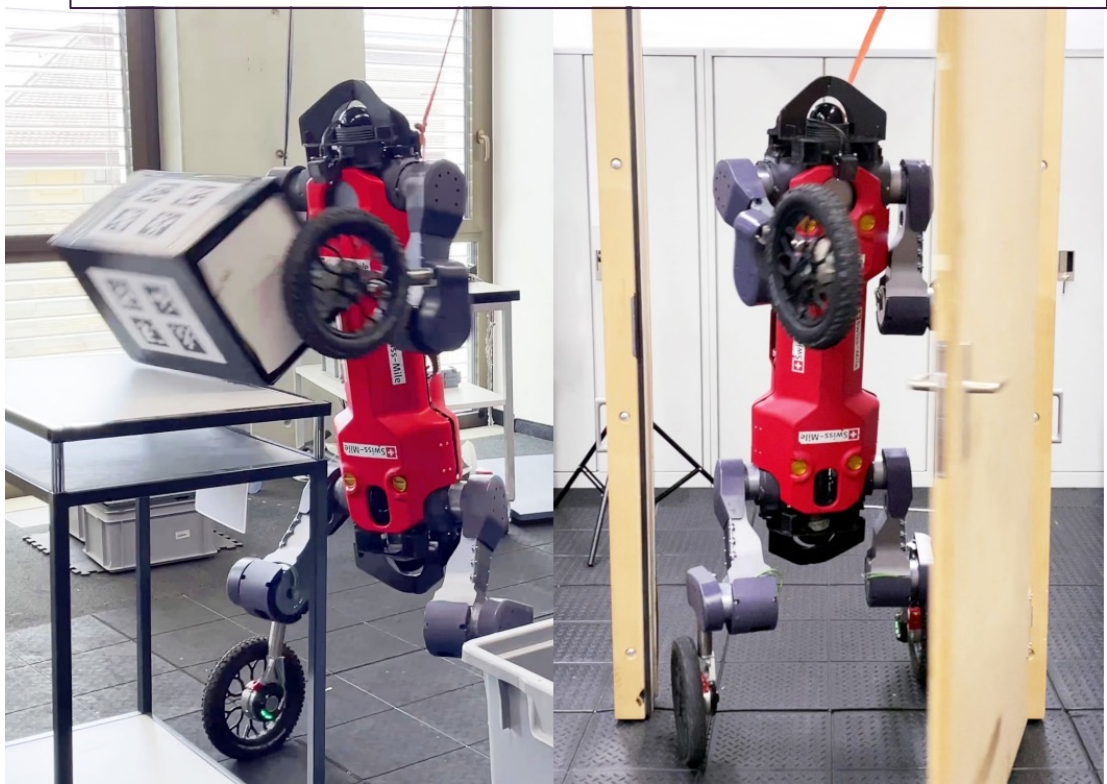
Simulation Advances: Infrastructure for Accelerated Learning

- Simulation is important because robots can learn in more diverse situations faster without wearing out their parts.
- You can train 100 robots simultaneously in different situations
- And because they are robots, these 100 robots share everything they learn
- https://openreview.net/pdf?id=QG_ERxtDAP-
- Also see, https://www.youtube.com/watch?v=Nnpm-rJfFjQ



transfer to real

Images from https://openreview.net/pdf?id=QG_ERxtDAP- used with permission, thanks Clemens!



Purpose-Built Simulations

- NVIDIA Isaac Sim, https://developer.nvidia.com/isaac-sim
- Photorealistic
- Uses a scene graph, which is how Pixar builds scenes
- Movies are becoming more like video games and vice-versa

Also

- Habitat https://aihabitat.org/habitat3/
- ai2thor https://ai2thor.allenai.org/
- ThreeDWorld https://www.threedworld.org/

Automatically Generated Simulations

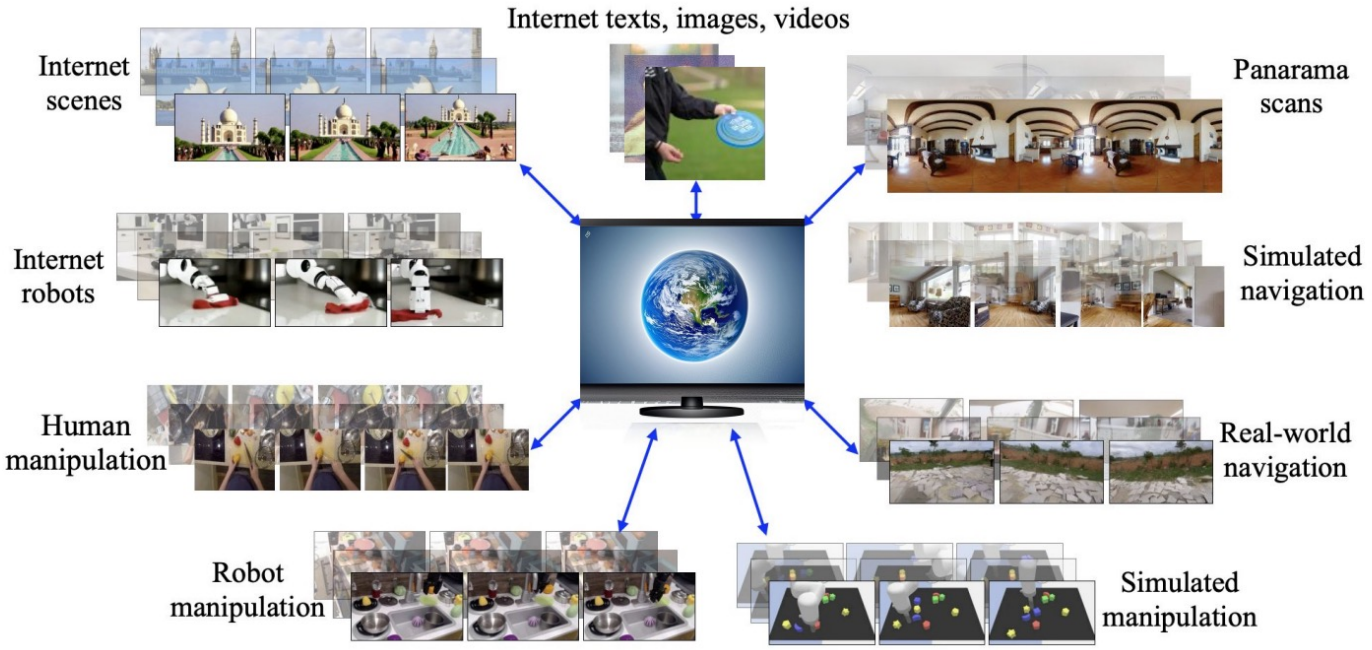
Exciting because expands the scope of training even more because humans don't have to set up the situation

The scenario is generated automatically, resulting in an unbounded amount of training data

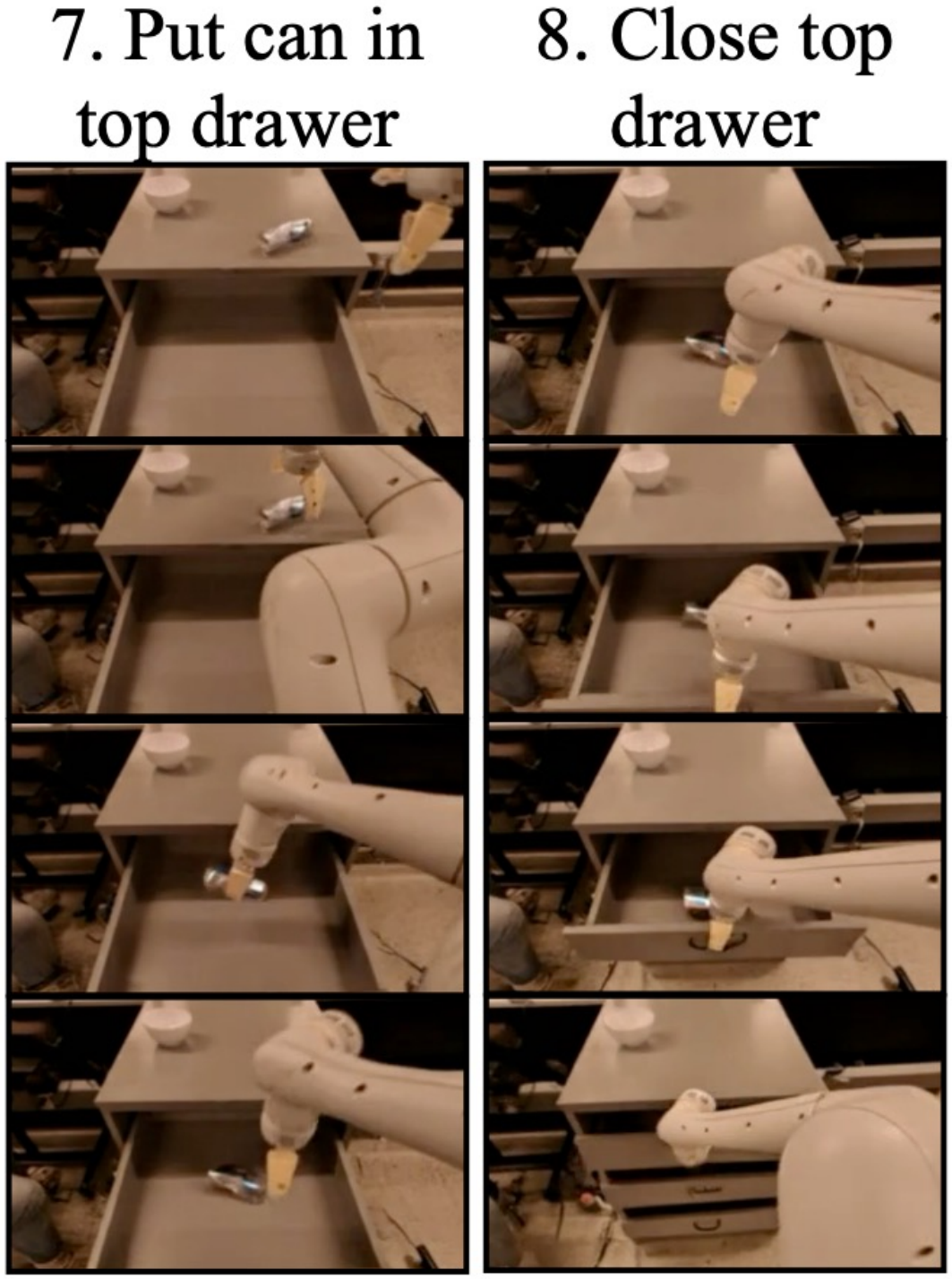
Universal Simulator

https://arxiv.org/pdf/2310.06114.pdf
https://universal-simulator.github.io/unisim/

- Combines multimodal data to create simulations using conditional video generation.
- Used to train a robot. Given a state, when the robot takes an action, it can simulate what that action will look like to the it, so the robot can learn and plan



Images used with permission from https://arxiv.org/pdf/2310.06114.pdf, thanks Sherry!



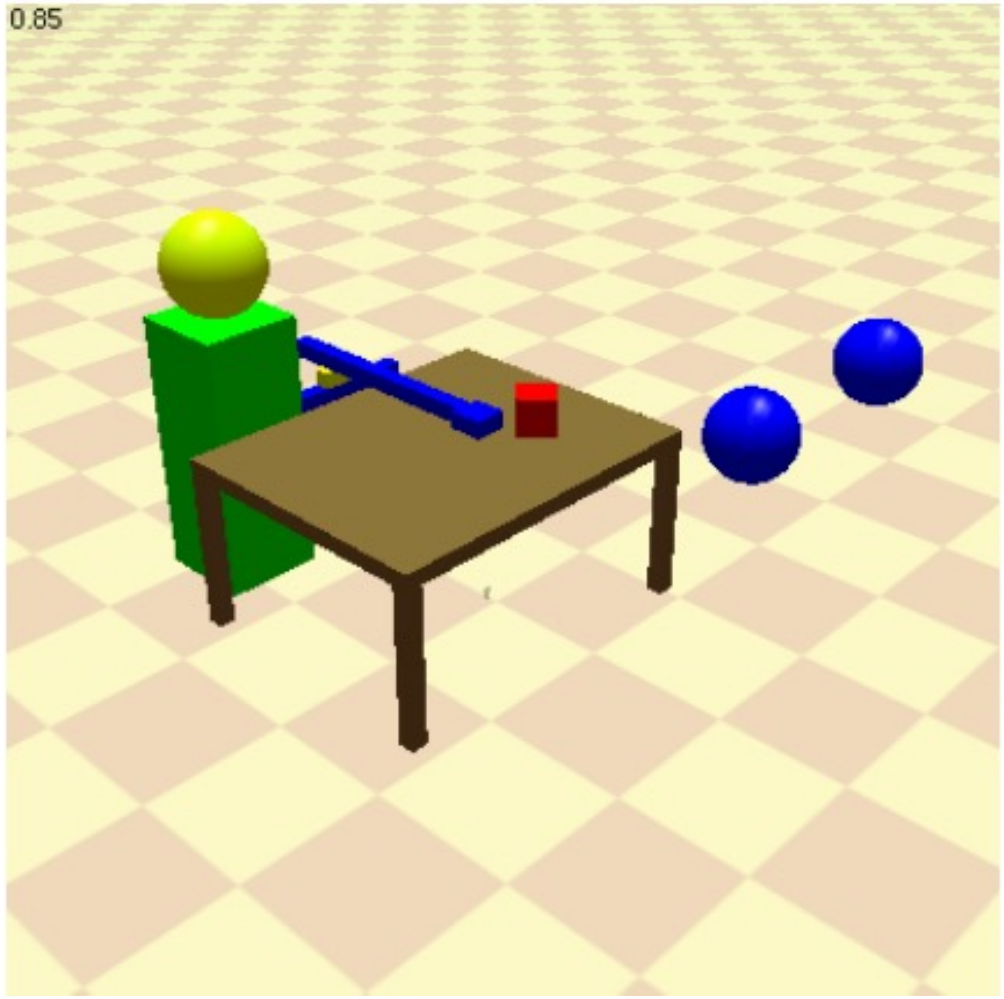
GAIA-1: generates simulations for autonomous driving

https://wayve.ai/thinking/scaling-gaia-1/
Image from https://arxiv.org/abs/2309.17080 used with permission, thanks Jamie!

Also, later, Genesis will be used in a paper we will see in a bit https://github.com/Genesis-Embodied-AI/Genesis

Expanding Rapidly

And becoming mainstream with Hugging Face tasks
https://huggingface.co/tasks/image-to-3d
https://huggingface.co/tasks/text-to-3d



compared with now

Used to train robots in my thesis in 2010
https://www.jonathanmugan.com/Publications/mugan_thesis.pdf

Image from https://arxiv.org/pdf/2311.01455.pdf, used with permission, thanks Yufei!

Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence

- The Evolutionary Paths of Intelligence

Outline

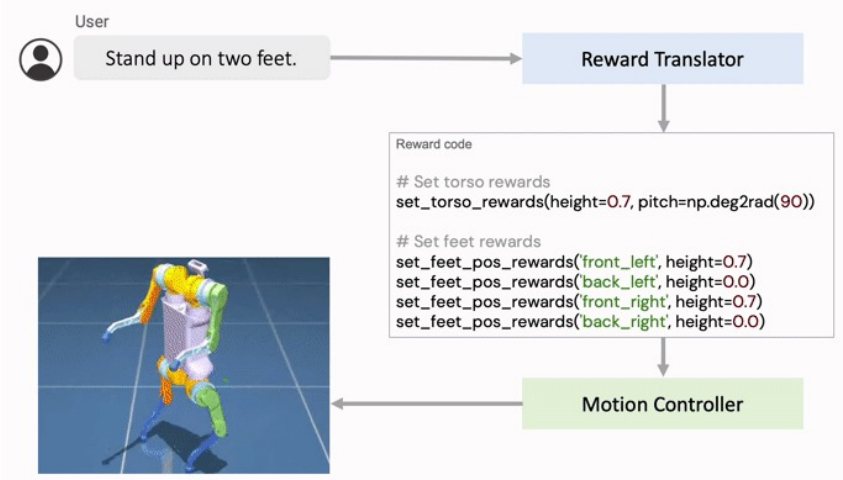
- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence
  - Skills: Moving its body to achieve its goals

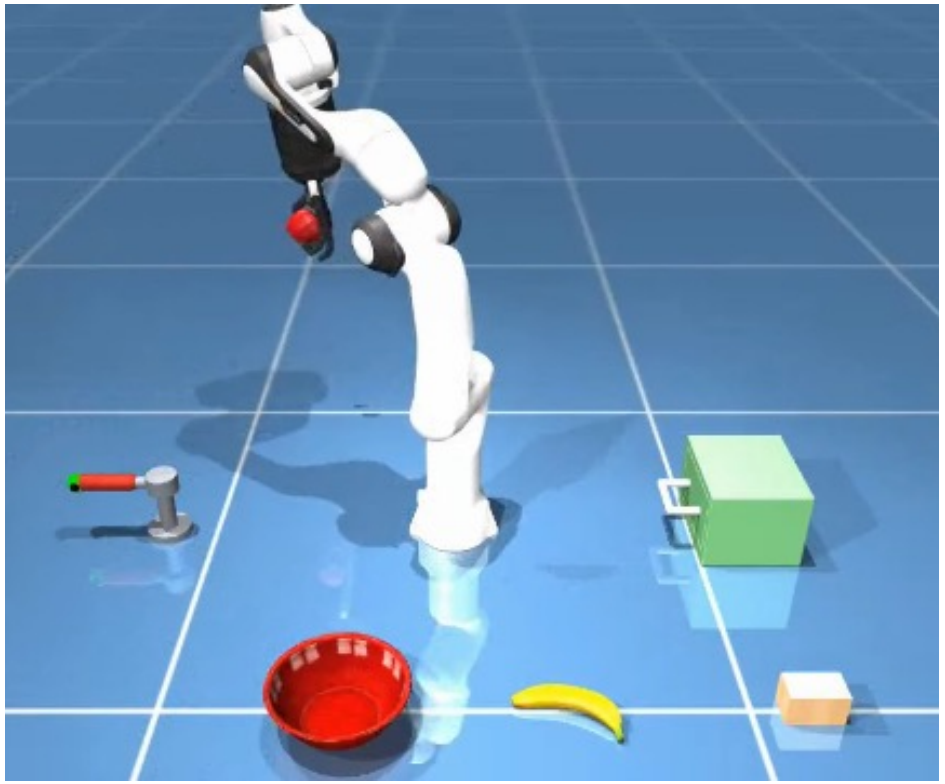
- The Evolutionary Paths of Intelligence

Google Research

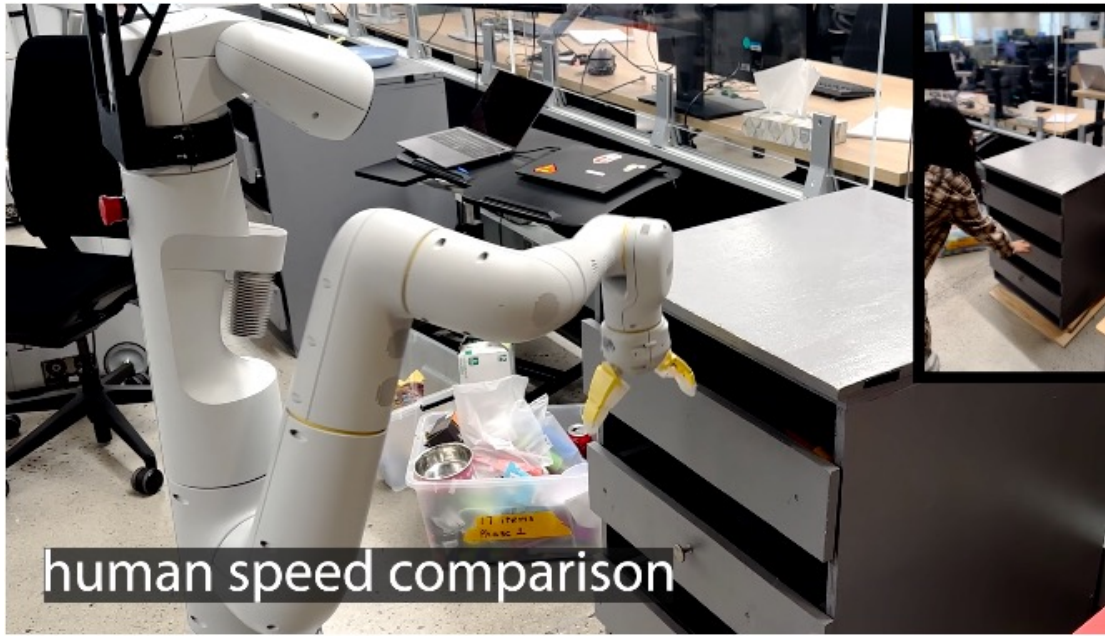
Trained through reward functions constructed by LLMs.

- LLM takes natural language and writes Python
- https://arxiv.org/abs/2306.08647





transfer to real robot

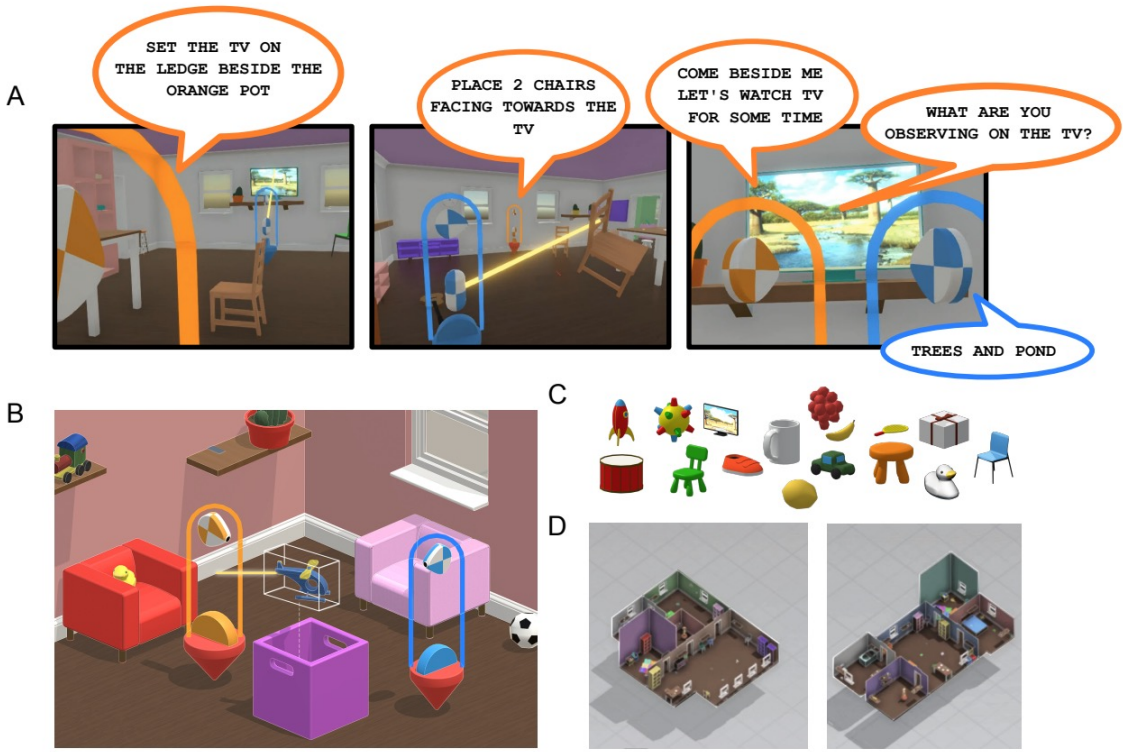


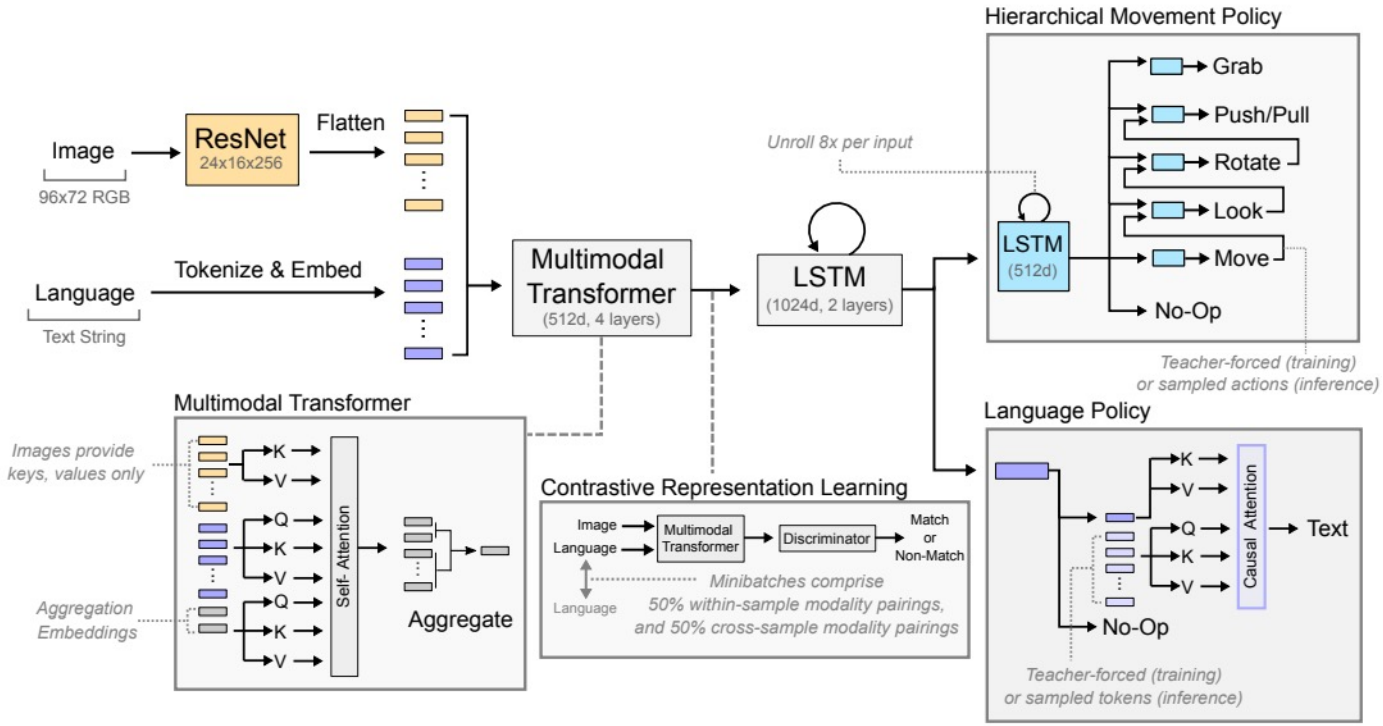
Images used with permission from blog below, thanks Wenhao!
https://blog.research.google/2023/08/language-to-rewards-for-robotic-skill.html

Deep Mind: Agents Group

Train in simulated environments with imitation learning and RL

- https://deepmind.google/discover/blog/building-interactive-agents-in-video-game-worlds/
- https://arxiv.org/pdf/2112.03763.pdf





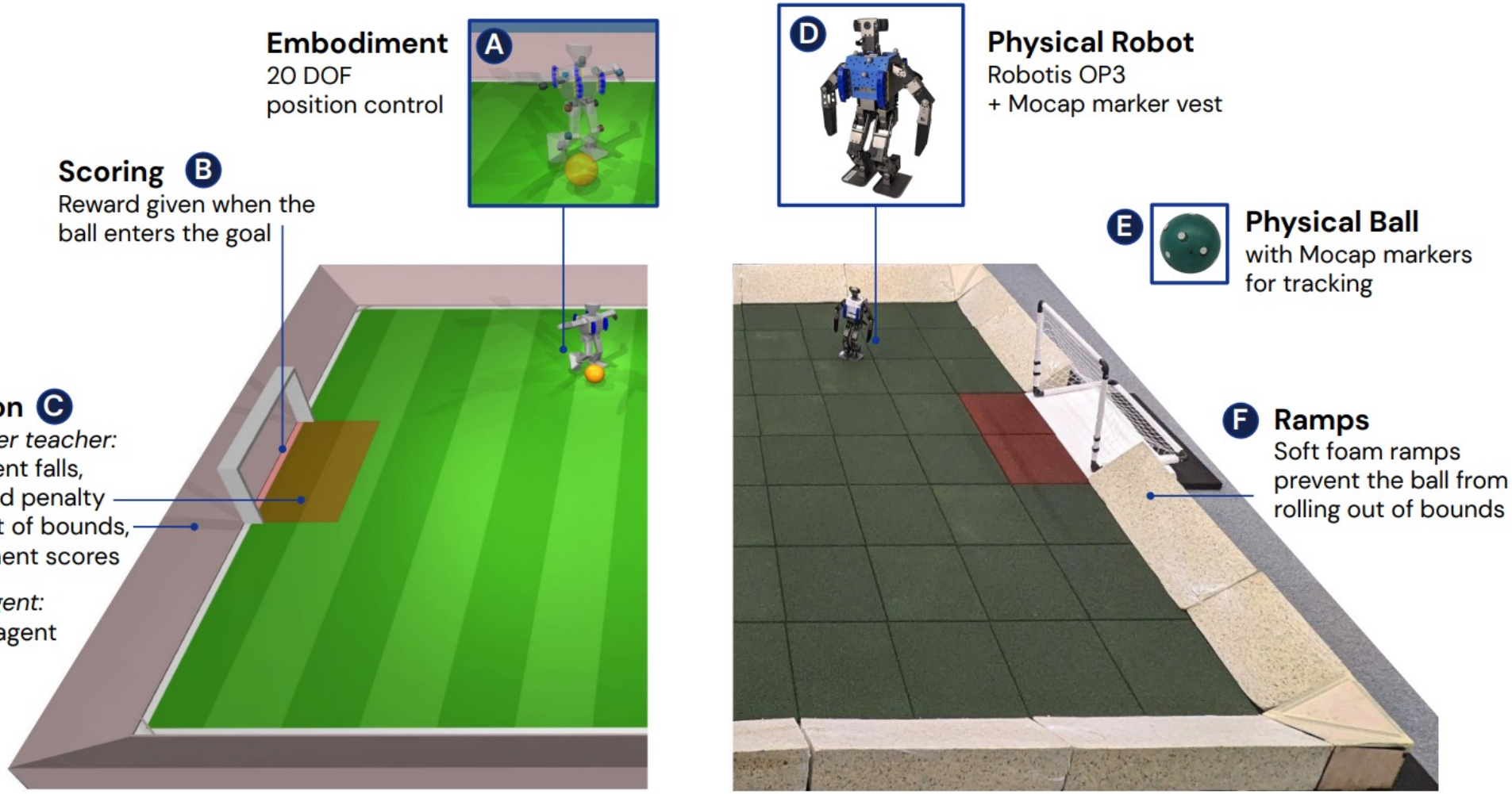
images from https://arxiv.org/pdf/2112.03763.pdf
used with permission, thanks Greg!

Soccer from DeepMind, sim to real

Learning Agile Soccer Skills for a Bipedal Robot with Deep Reinforcement Learning

- project page https://sites.google.com/view/op3-soccer
- paper https://arxiv.org/pdf/2304.13653.pdf
- earlier in simulation https://deepmind.google/discover/blog/from-motor-control-to-embodied-intelligence/

Image used with permission, thanks Tuomas! from https://arxiv.org/pdf/2304.13653.pdf

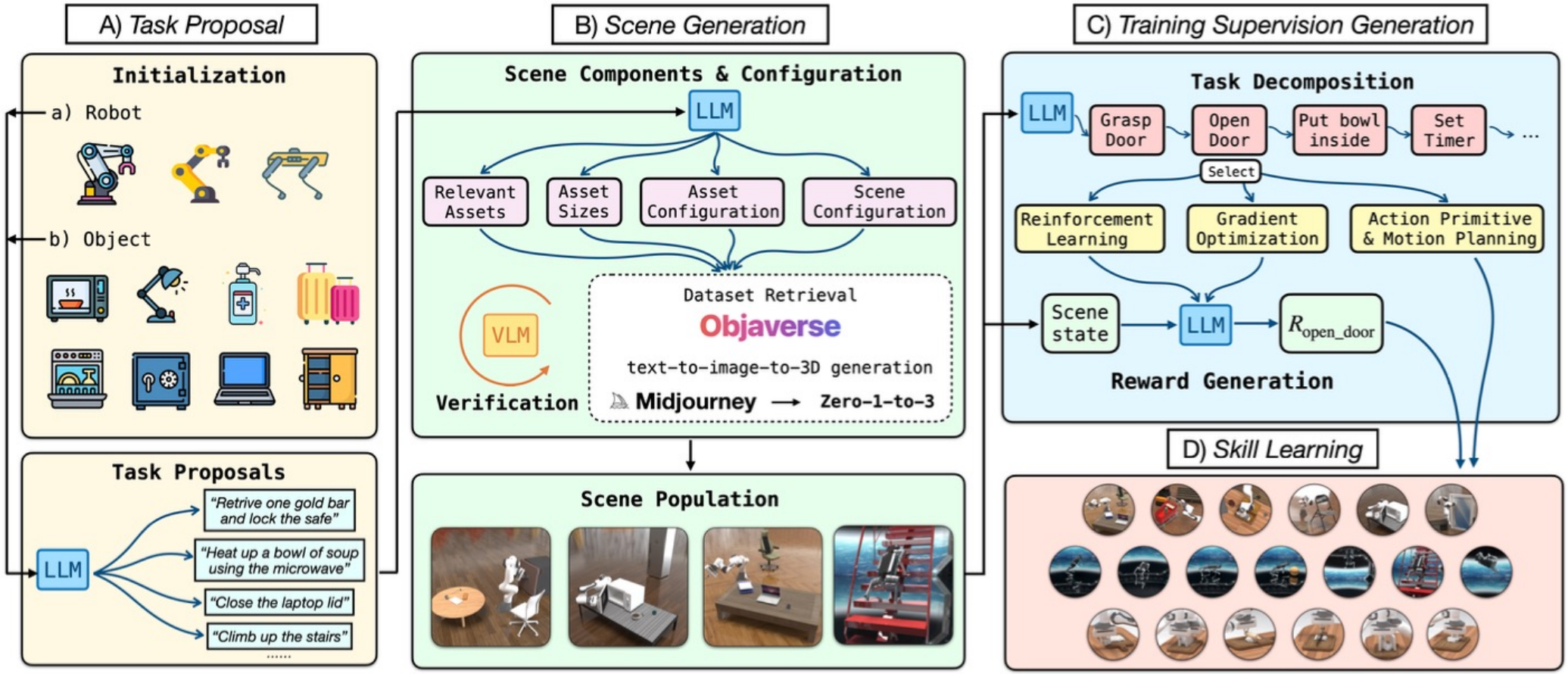


RoboGen: Training by autonomously building scenes

- https://github.com/Genesis-Embodied-AI/RoboGen
- https://robogen-ai.github.io/

Use LLM to create a scene in a simulator. Then learn skills in that scene.

Image from https://arxiv.org/pdf/2311.01455.pdf, used with permission, thanks Yufei!



Deep Mind: RT-X

RT stands for Robotics Transformer. X means using data across robot embodiments.
https://deepmind.google/discover/blog/scaling-up-learning-across-many-different-robot-types/
https://arxiv.org/pdf/2310.08864.pdf

AutoRT https://auto-rt.github.io/ Uses foundation models and can direct multiple robots
https://deepmind.google/discover/blog/shaping-the-future-of-advanced-robotics/

And Toyota!

Uses diffusion methods to denoise a policy to match a demonstration policy. Gradually transforms from demonstration to noise and then learns by going back in the opposite direction.

https://www.tri.global/news/toyota-research-institute-unveils-breakthrough-teaching-robots-new-behaviors
Uses Diffusion Policy https://arxiv.org/pdf/2303.04137.pdf.

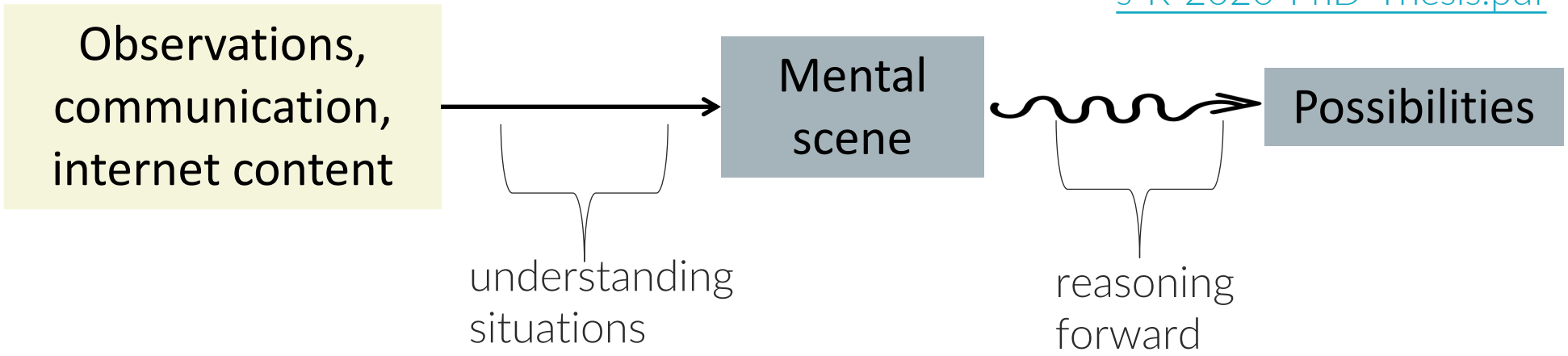
Pretty amazing stuff. A robot can spread a spread on a cracker.

Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence
  - Skills: Moving its body to achieve goals
  - Thinking: Understanding situations and forming high-level plans
- The Evolutionary Paths of Intelligence

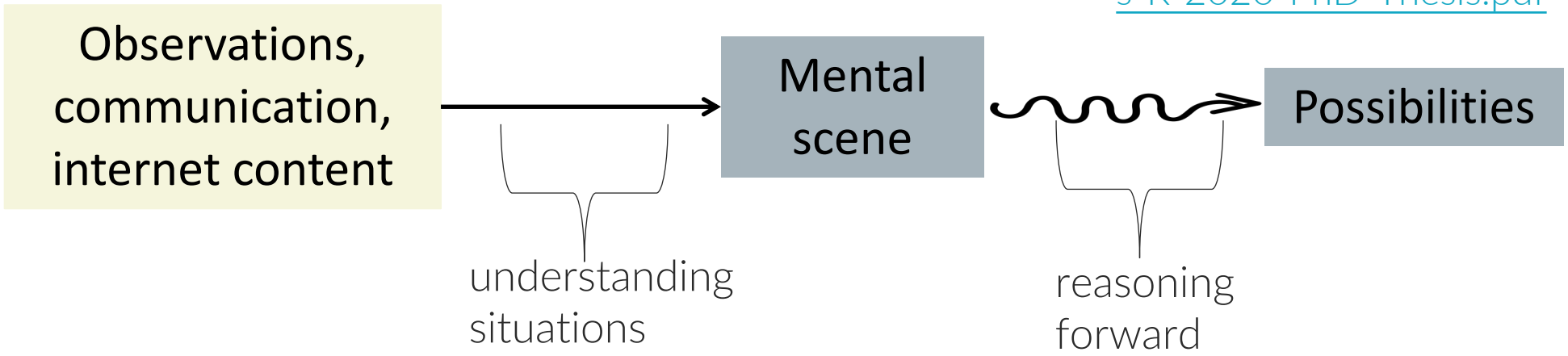
Thinking: Understanding situations and forming high-level plans

Reminiscent of Robert Brandom through Richard Evans
https://www.doc.ic.ac.uk/~re14/Evans-R-2020-PhD-Thesis.pdf



Thinking: Understanding situations and forming high-level plans

Reminiscent of Robert Brandom through Richard Evans
https://www.doc.ic.ac.uk/~re14/Evans-R-2020-PhD-Thesis.pdf



Example:

"The table is on the table" → What?!??

Thinking: Understanding situations and forming high-level plans

Reminiscent of Robert Brandom through Richard Evans
https://www.doc.ic.ac.uk/~re14/Evans-R-2020-PhD-Thesis.pdf

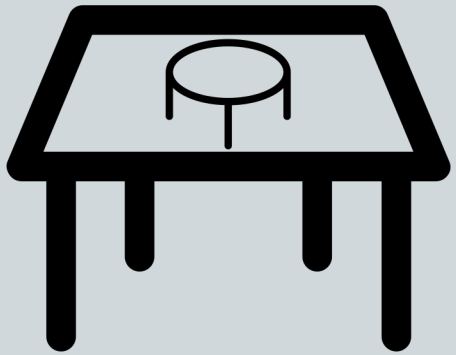
Observations, communication, internet content → Mental scene → Possibilities

understanding situations

reasoning forward

Example:

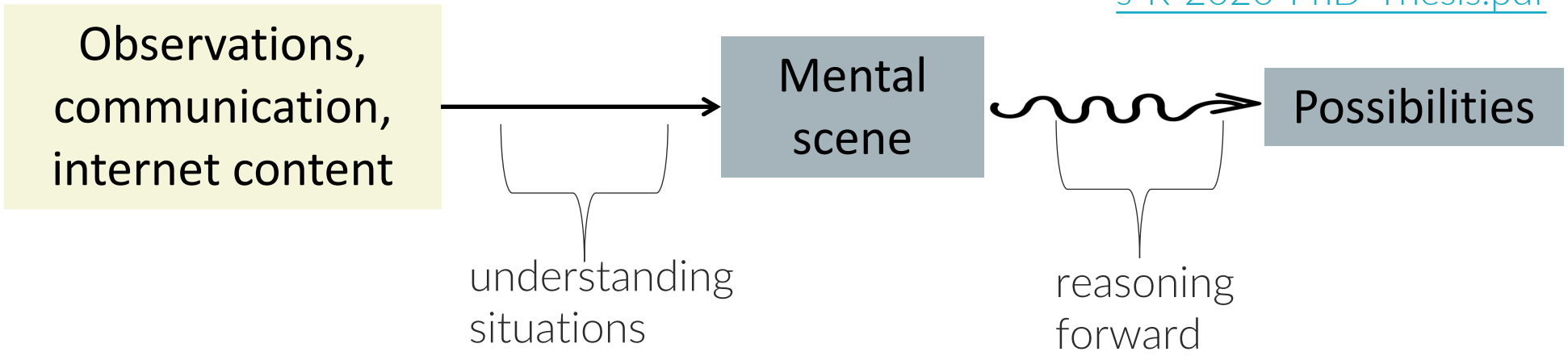
"The table is on the table"



- If you want to pick up top table, you must first walk to the table
- If you push the bottom table, the top table will fall
- Party guests will think this is weird looking.

Thinking: Understanding situations and forming high-level plans

Reminiscent of Robert Brandom through Richard Evans
https://www.doc.ic.ac.uk/~re14/Evans-R-2020-PhD-Thesis.pdf



Two ways to not understand the environment:

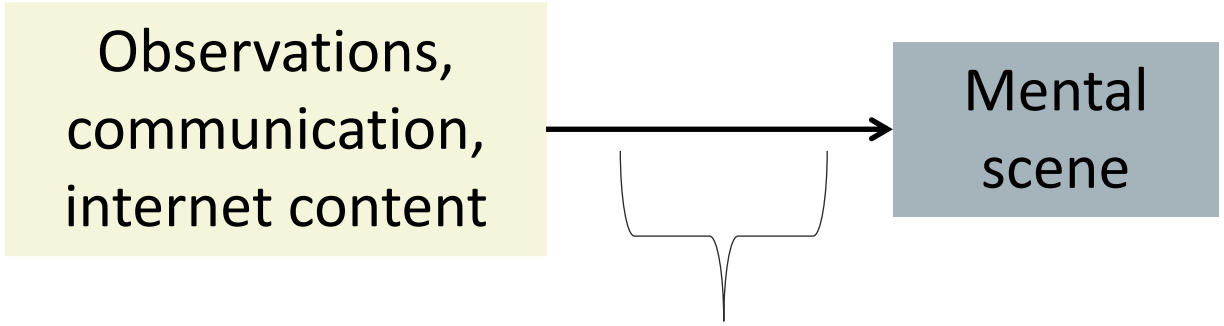
1. Wrong mental scene
2. Not knowing the possibilities

Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence
  - Skills: Moving its body to achieve goals
  - Thinking: Understanding situations and forming high-level plans
    - Grounded Representation: Map sensory input to a representation

- The Evolutionary Paths of Intelligence

Grounding representation



Through LLM to JSON,
my earlier talk covered this,
as we will see used by Tenenbaum and GOAT

LLMs to translate natural language into a machinery that allows AI to think deliberately.

Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence
  - Skills: Moving its body to achieve goals
  - Thinking: Understanding situations and forming high-level plans
    - Grounded Representation: Map sensory input to a representation
    - Reasoning: Determine possibilities from representation

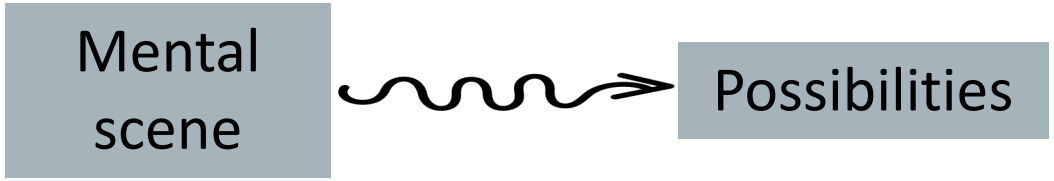
- The Evolutionary Paths of Intelligence

Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence
  - Skills: Moving its body to achieve goals
  - Thinking: Understanding situations and forming high-level plans
    - Grounded Representation: Map sensory input to a representation
    - Reasoning: Determine possibilities from representation
      - Linear Thinking: Search, Inference, and Data Processing

- The Evolutionary Paths of Intelligence

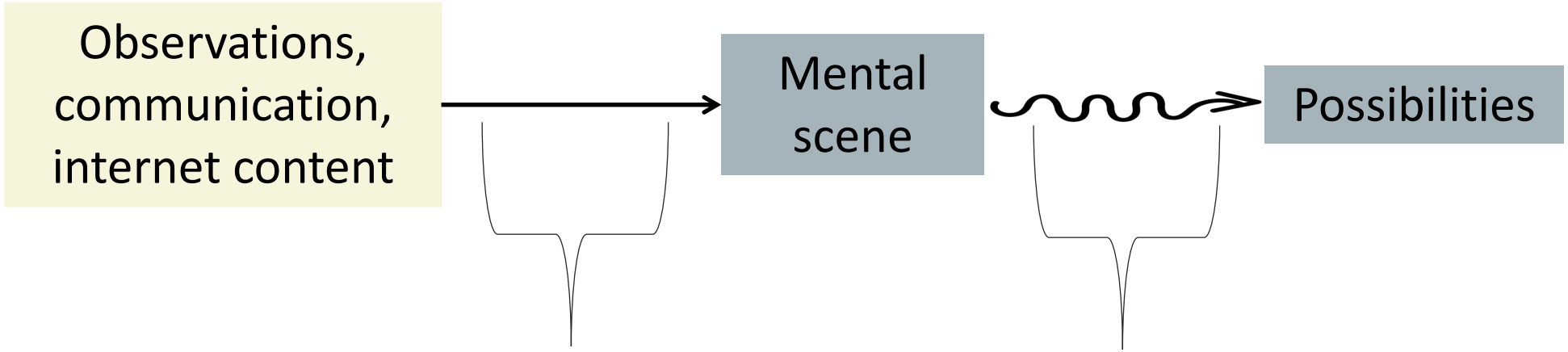
Linear Thinking: Search, Inference, and Data Processing



Search is simulation with branching.

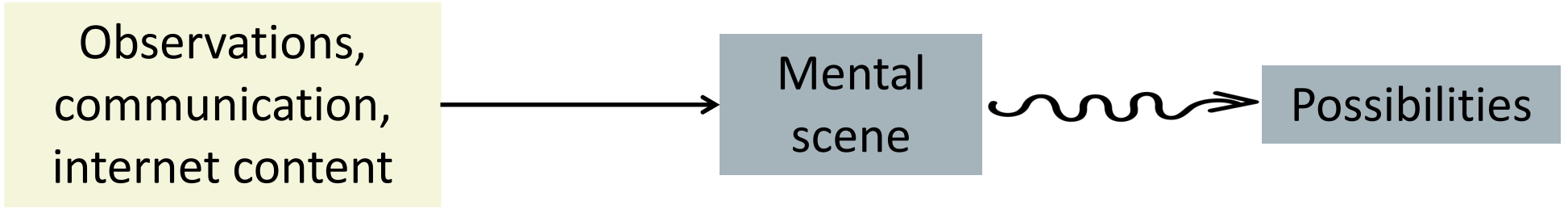
Major labs are working on combining search with LLMs
https://twitter.com/ylecun/status/1728126868342145481

Tenenbaum and friends



https://arxiv.org/abs/2306.12672

Tenenbaum and friends



Easier to understand if you mentally replace the probabilistic programming language with first-order logic.

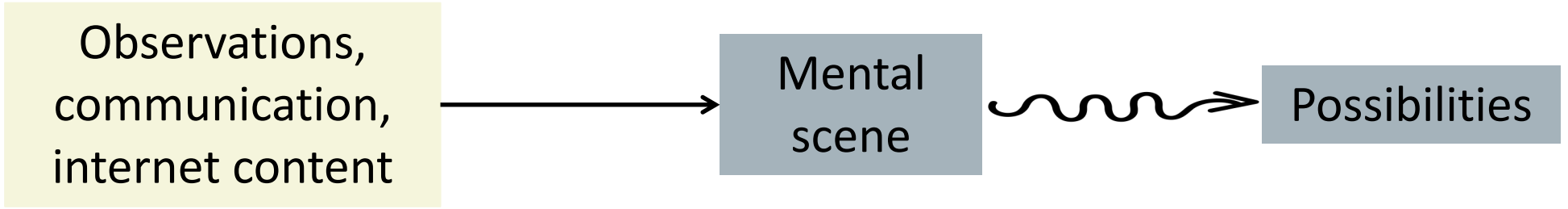
Example: They could have a domain that says if you sing to someone then they are happy. They could represent that as a logic formula: `sing-to(x,y) -> happy(y)`

Then they could use an LLM to convert “Billybob sang to Anika a wonderful song" to
`sing-to(Billybob, Anika)`

Then they run the model forward
`sing-to(Billybob, Anika) -> happy(Anika)`

And the program knows Anika was happy. It uses a world model to "think" instead of relying on interpolation.

Tenenbaum and friends



Best resource for probabilistic programming: https://probmods.org/

We need a practical probabilistic programming language in Python. We have PyMC, but to use that you have to pull out your old notes on Theano.

There's Pyro, but I wish Pyro would do a better job of hiding the implementation details. I shouldn't need to understand variational inference and such just to get the probability of a god dang hot dog. I've tried to use Pyro a few times, but every time I spend more effort trying to understand poutines and such instead of modeling my problem. /rant

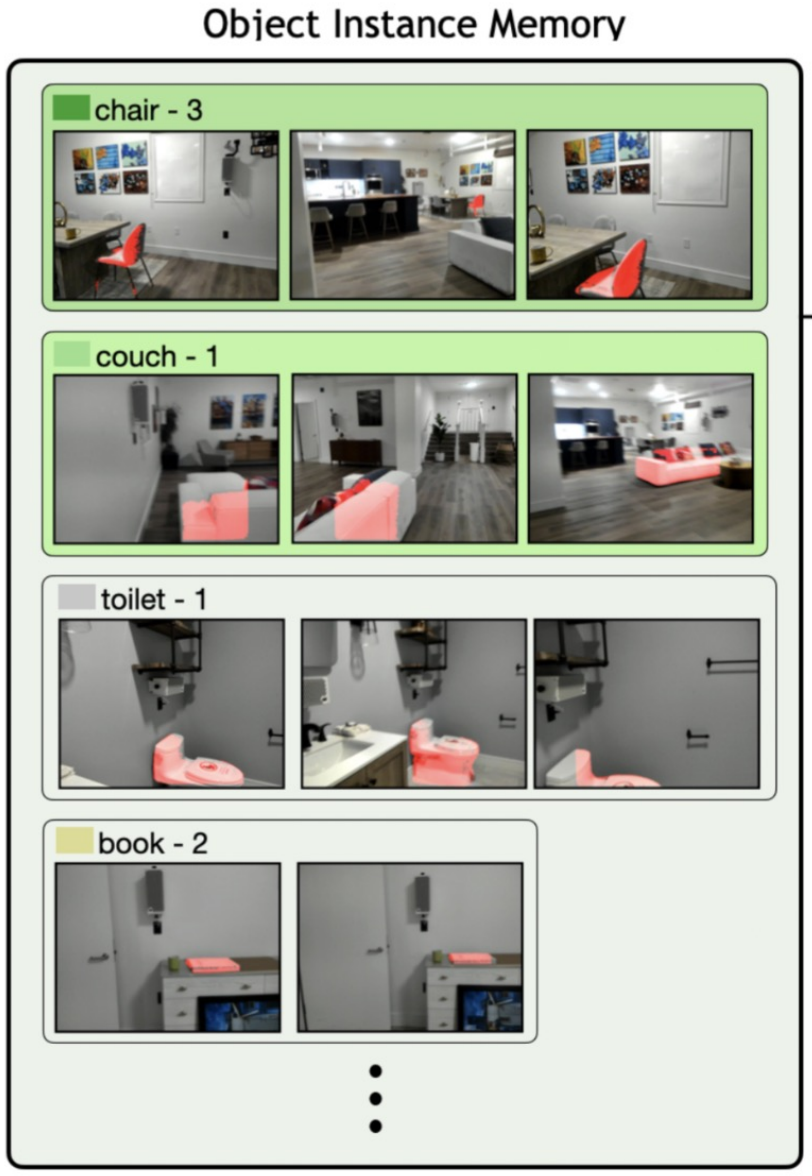
GOAT: GO to Any Thing

Robot explores and builds a map of the environment and when it sees objects it associates their location in the map. It can then navigate to those objects and place them in other locations.

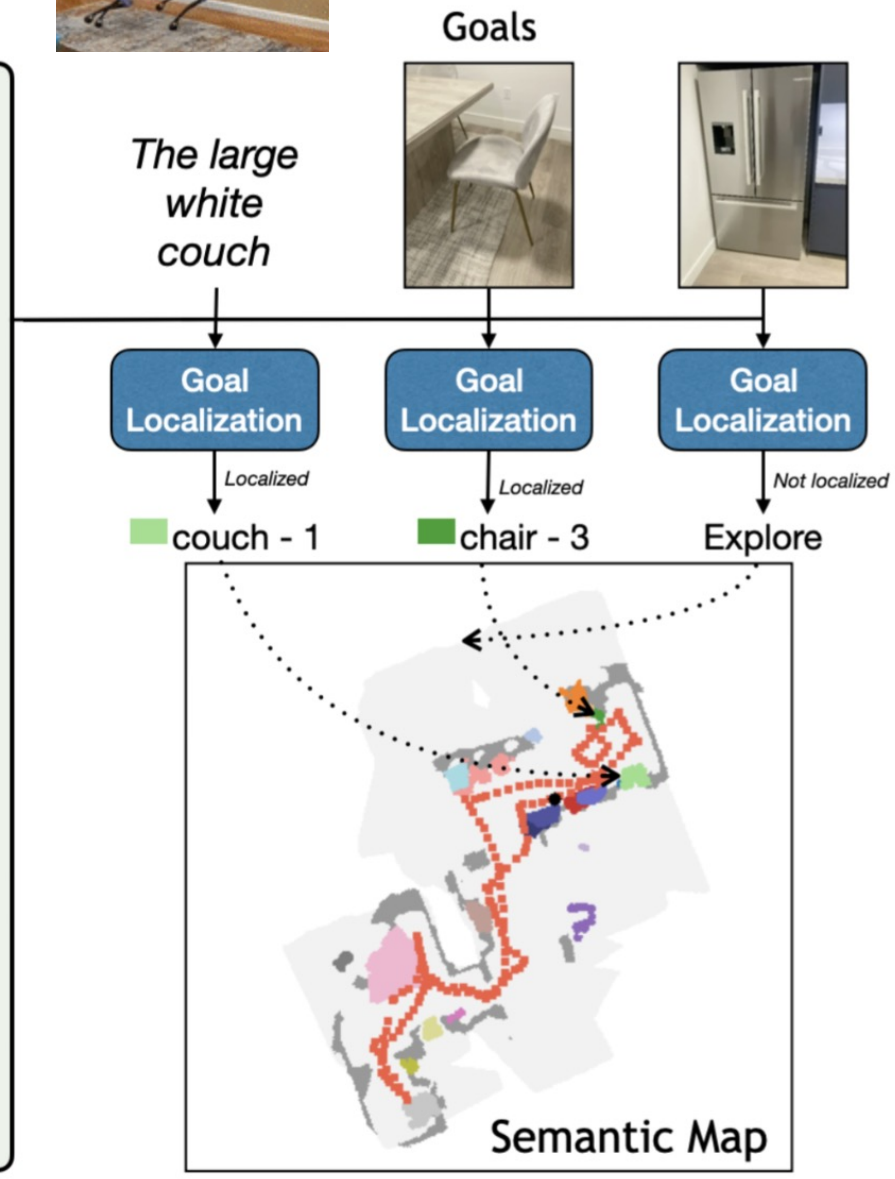
https://theophilegervet.github.io/projects/goat/

Map is the underlying data structure; objects associated with locations; images associated with objects. Then it can go to those objects.

Images used with permission, thanks Theophile!



The robot

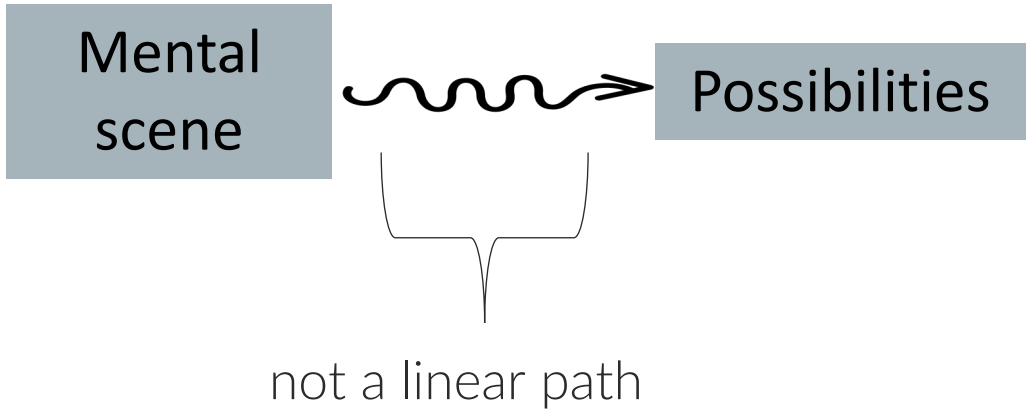


Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence
  - Skills: Enabling a robot to move its body to achieve goals
  - Thinking: Understand situations and decide what to do
    - Grounded Representation: Map sensory input to a representation
    - Reasoning: Determine possibilities from representation
      - Linear Thinking: Search, Inference, and Data Processing
      - Thought Jumping: Dynamically condensing time and space with abstraction
- The Evolutionary Paths of Intelligence

Thought Jumping

To solve wider problems, AI can't just think linearly.



Thought jumping is when you need to switch between levels of abstraction when you are trying to reason from first principles.

Thought Jumping Example 1: Halting Problem

Consider the problem of whether these algorithms will terminate. It is obvious they will.

```
10 sum = 0
20 sum = sum + 1
30 IF sum < 10 THEN GOTO 20
40 REM end of program
```

```
10 sum = 0
20 sum = sum + 1
30 IF sum != 10 THEN GOTO 20
40 REM end of program
```

`!=` is `<>`
for you purists out there

```
10 sum = 0
20 sum = sum + 1
30 IF sum != 1000000000 THEN GOTO 20
40 REM end of program
```

Thought Jumping Example 1: Halting Problem

This one will not.

```
10 sum = 0
20 sum = sum + 1
30 IF sum != 1000000000.3 THEN GOTO 20
40 REM end of program
```

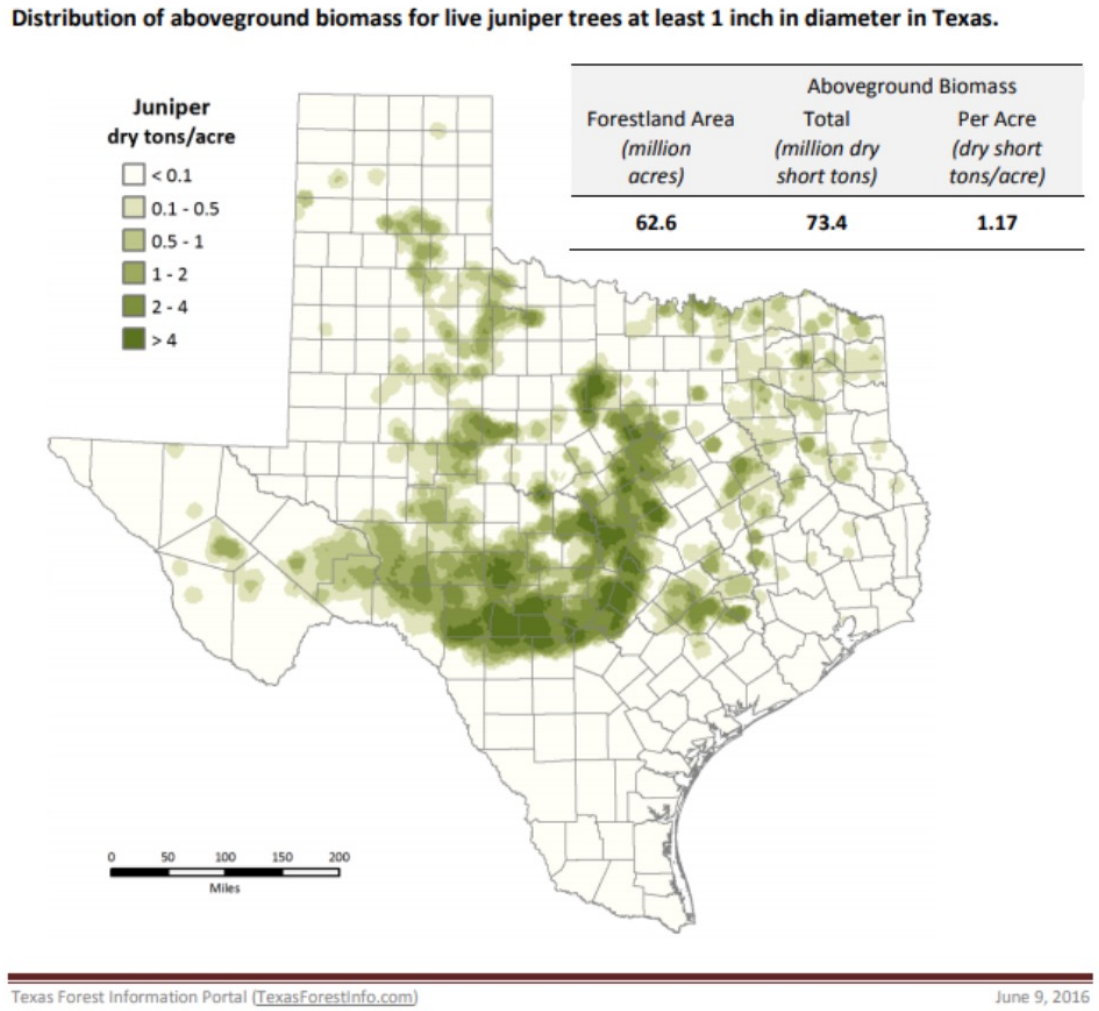
`!=` is `<>`
for you purists out there

Obvious to us. We “jump” to the point 999999999.

We need computers to do this.

- Thought jumping doesn’t “solve” the halting problem—it makes it irrelevant, just like it is for humans.
- This inability to jump out of the system is behind the Gödel Incompleteness Theorem and barbers who cut everyone’s hair who doesn’t cut their own.

Thought Jumping Example 2: Cedar Fever



https://tfsweb.tamu.edu/content/article.aspx?id=31295
used with permission, thanks Robert!

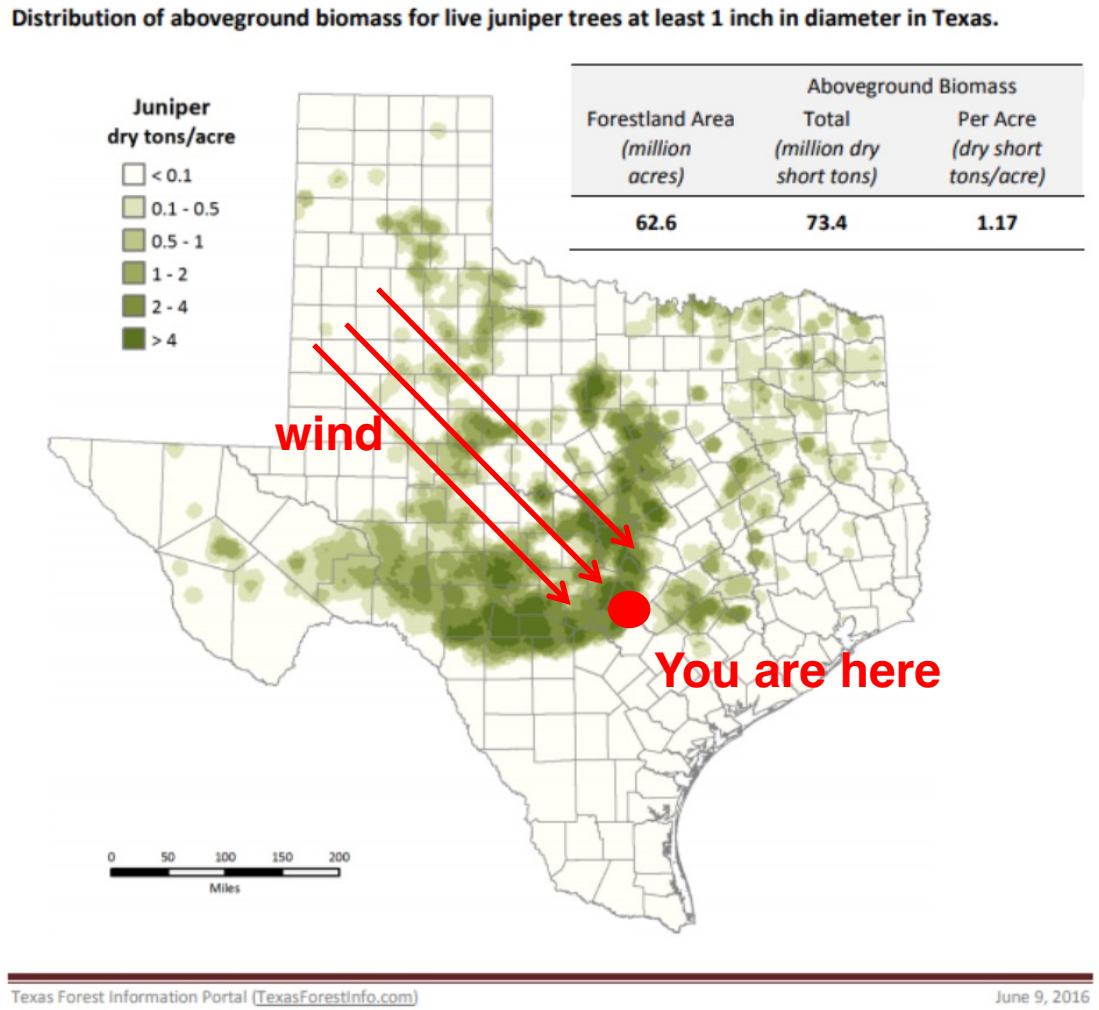
How can we predict what day cedar fever will be bad?

For those who don't live in Austin,

- January is miserable because the ash juniper trees give off a lot of pollen that makes you sneeze and want to gash your eyes out.
- It takes a few years to develop the allergy, so newcomers seemingly have a superhuman ability to walk around outside in January without a care.

We want robots to reason from first principles why dry, windy days are the worst.

Thought Jumping Example 2: Cedar Fever



https://tfsweb.tamu.edu/content/article.aspx?id=31295
used with permission, thanks Robert!

- The model is that pollen is bad when the wind blows from the northwest on cold, dry days.
- A robot can create this model by simulating the pollen coming off trees and floating in the air, but for the pollen to travel long distances, it must jump up a level and simulate it traveling at a rate of speed.
- We don't even notice we do these jumps
- A computer using simulation to think and understand must do the same.

Thought Jumping is what makes human thought so powerful

Humans are terrible at thinking too far linearly. I still remember the 44 I got on my long division test in 4th grade.

- We believe we think linearly, but we don't. We jump from incomplete piece to piece.
- Imagine the car you drove here. Got it? Did you imagine the bumper?
  - See Daniel Gilbert, *Stumbling on Happiness*, sounds like a self-help book but is about cognition

It is what makes human thinking so powerful. We don't get trapped because we can jump up above a system when we need to.

Example 3: Cyber Security

Computers get hacked because they are blind automatons. There is "nobody" watching. See Mugan, 2014
https://www.jonathanmugan.com/Publications/mugan_spie_14.pdf

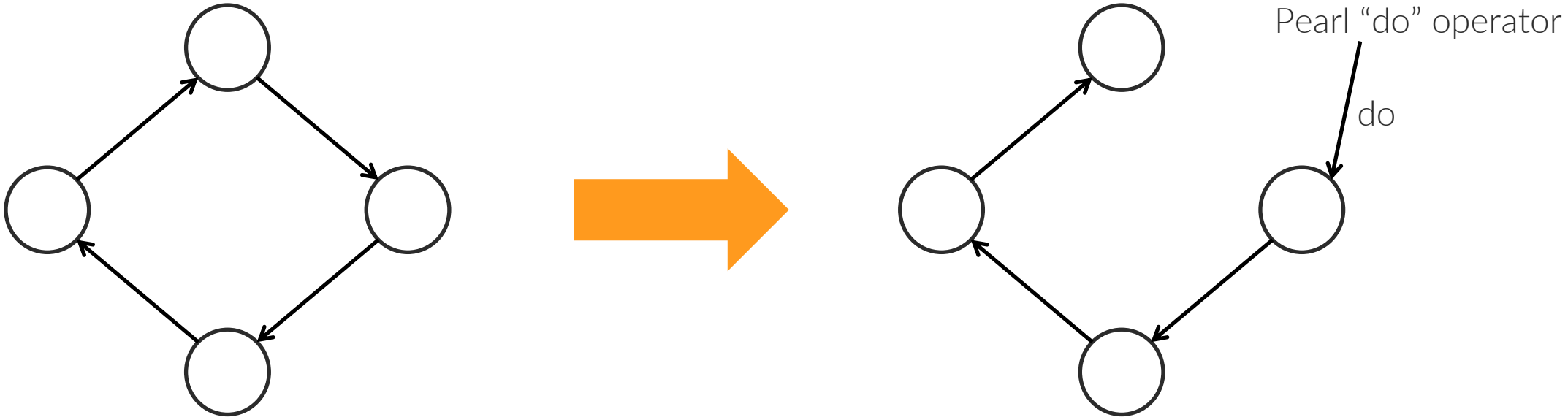
Buffer overflows and SQL injection, and now with LLMs, prompt injection

- somebody sends you an email and the computer interprets it as a prompt
  https://simonwillison.net/2023/Apr/14/worst-that-can-happen/

With LLMs as the execution platform that both reads data in natural language and takes instructions in natural language, it becomes harder to separate instructions from data.

Example 4: Understanding causality

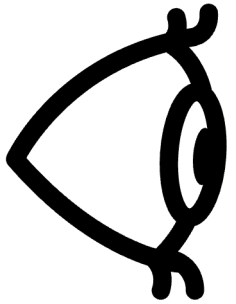
Causal thinking in general requires “jumping” out of a system and asking, “If I were to do this, what would happen?”



http://bayes.cs.ucla.edu/BOOK-2K/

Robots must have a process that determines when to jump between search spaces

$T(s, a) \rightarrow s'$
$V(s) \rightarrow [0,1]$



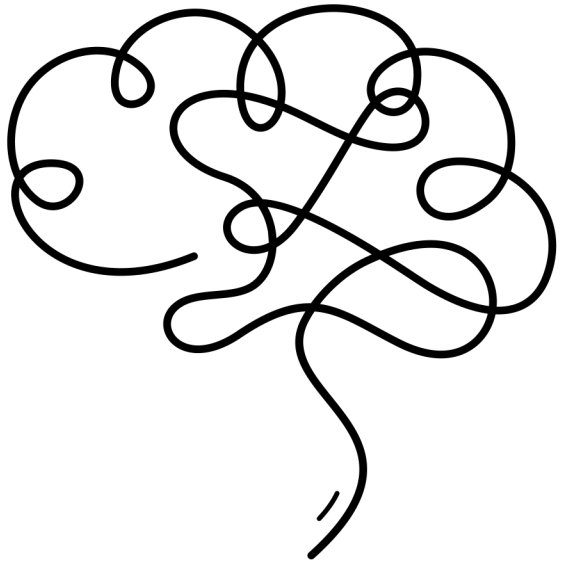
when to jump

$\mathfrak{T}(\mathfrak{s}, \mathfrak{a}) \rightarrow \mathfrak{s}'$
$\mathfrak{V}(\mathfrak{s}) \rightarrow [0,1]$

- Transition function T takes state s and action a (if any) and gives next state s'.
- Value function V says how good that state is (0 worst imaginable, 1 best imaginable)

Search space at different level of abstraction

And maybe even consciousness



- Graziano: consciousness is more than having a global workspace—it is the internal focus of attention into that workspace from the outside (reminiscent of Hofstadter's strange loops).
- You jump to different spaces. You are sitting there thinking about that time in third grade and then it's your turn to bowl.

Probably necessary but not sufficient:

- My current thinking is that it is possible to build intelligent machines that are not consciousness

Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence

- The Evolutionary Paths of Intelligence

Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence

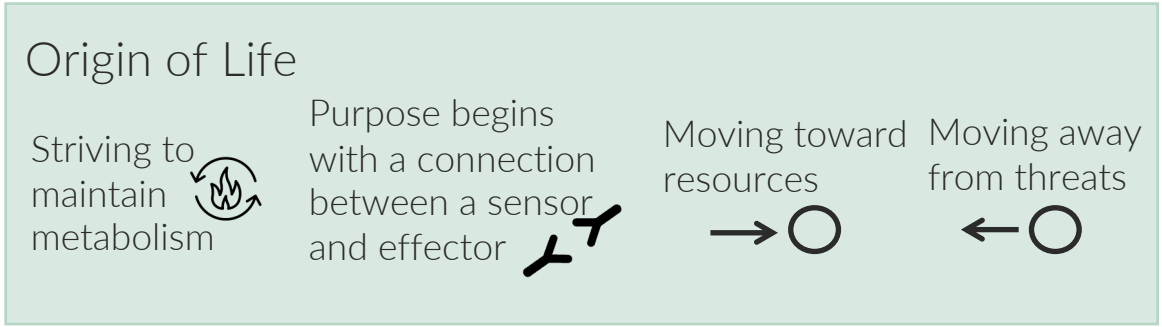
- The Evolutionary Paths of Intelligence
  - Human intelligence evolved bottom-up

Origin of life

During evolution, purpose came into being when by chance the first sensor element connected to the first effector (motor) element [1].

The purpose of life is to maintain metabolism.

[1] Gaddam, S., & Ogas, O. (2022). Journey of the Mind: How Thinking Emerged from Chaos. WW Norton.



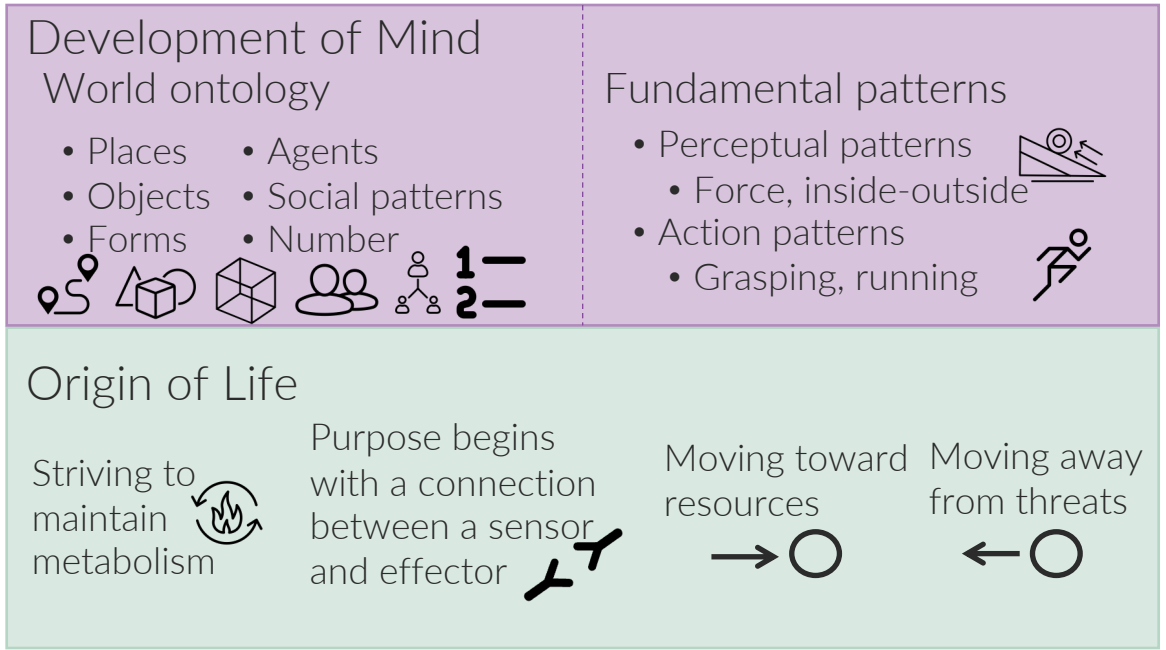
Development of mind

The developmental psychologist Elizabeth Spelke describes the ontology used by the human mind as consisting of six systems of core knowledge [1,2].

1. Spelke, E. (1994). Initial knowledge: Six suggestions. Cognition, 50(1–3), 431–45.
2. Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. Psychological Review, 99(4), 605.
3. Pinker, S. (2003). How the mind works. Penguin UK.
4. Johnson, M. (1987). The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason. University of Chicago Press.
5. Mandler, J. (2004). The Foundations of Mind, Origins of Conceptual Thought. Oxford University Press.
6. Lakoff, G., & Johnson, M. (1980). Metaphors We Live By. University of Chicago Press.

Alongside this world ontology is a set of fundamental patterns that seem to enable many of our cognitive abilities. Perceptual patterns include those such as force and inside-outside.

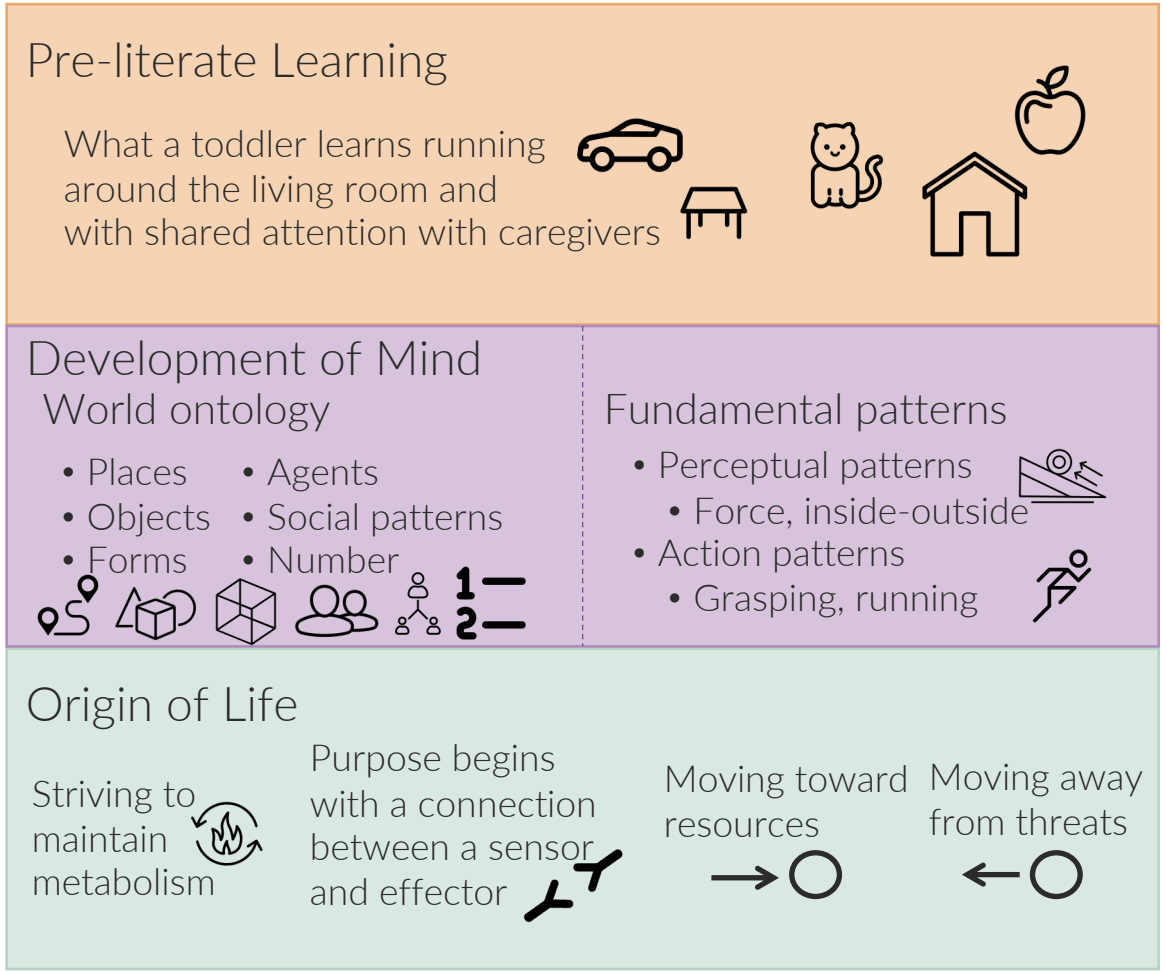
We understand the world in terms of these patterns [3,4,5]. These patterns likely evolved by being useful for one decision and were then reused by evolution for many decisions, even later becoming abstract through metaphor [6]



Pre-literate Learning

Children learn through exploration and through shared attention with caregivers [1,2,3].

At this level, learning entails properties and interactions of specific kinds of objects, especially the kinds of objects that are of interest to your domain.

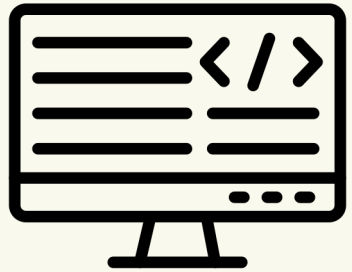


1. Mandler, J. (2004). The Foundations of Mind, Origins of Conceptual Thought. Oxford University Press.
2. Gopnik, A. (2009). The Philosophical Baby: What Children's Minds Tell Us About Truth, love, and the meaning of life. Farrar Straus & Giroux.
3. Tomasello, M. (2019). Becoming human: A theory of ontogeny. Harvard University Press.

Then we can read

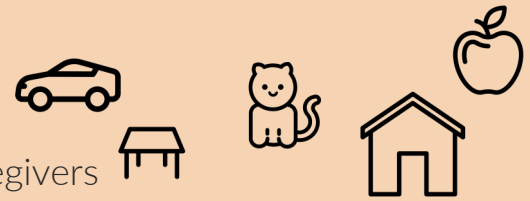
Internet content

Consuming blogs, news articles, essays, code, comics, videos, ...



Pre-literate Learning

What a toddler learns running around the living room and with shared attention with caregivers



Development of Mind

World ontology

- Places
- Objects
- Forms
- Agents
- Social patterns
- Number

Fundamental patterns

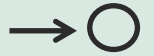
- Perceptual patterns
  - Force, inside-outside
- Action patterns
  - Grasping, running

Origin of Life

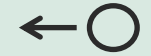
Striving to maintain metabolism

Purpose begins with a connection between a sensor and effector

Moving toward resources



Moving away from threats



Outline

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence

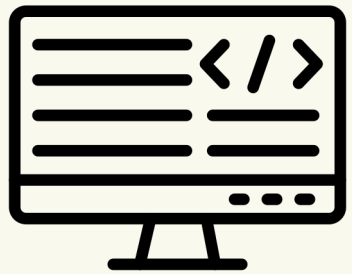
- The Evolutionary Paths Intelligence
  - Human intelligence evolved bottom-up
  - Robot intelligence is evolving top-down

ChatGPT and friends

Large Language Models (LLMs) just predict the next word, and any “understanding” must be emergent

Internet content

Consuming blogs, news articles, essays, code, comics, videos, …



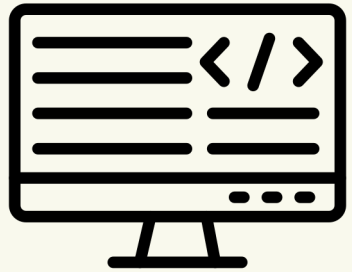
Learning from images and video

We progressed from LLMs over text to multimodal LLMs over text, images, and video,

- but at this level we still don't have a grounded foundation for learning

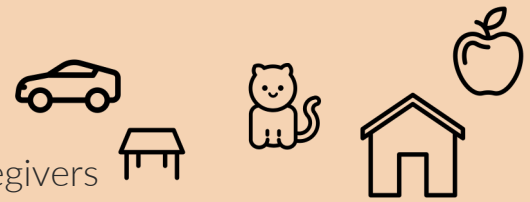
Internet content

Consuming blogs, news articles, essays, code, comics, videos, …



Pre-literate Learning

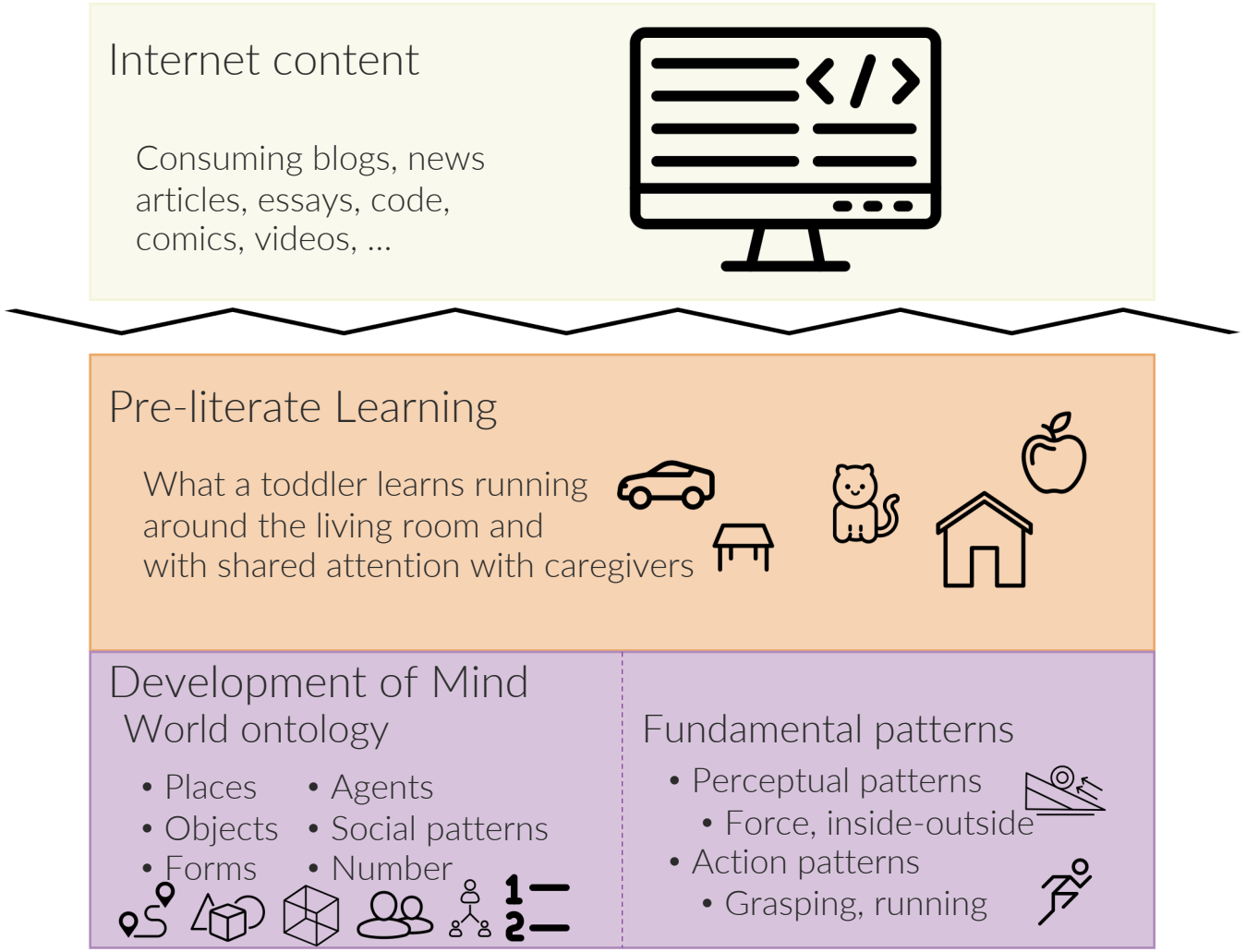
What a toddler learns running around the living room and with shared attention with caregivers



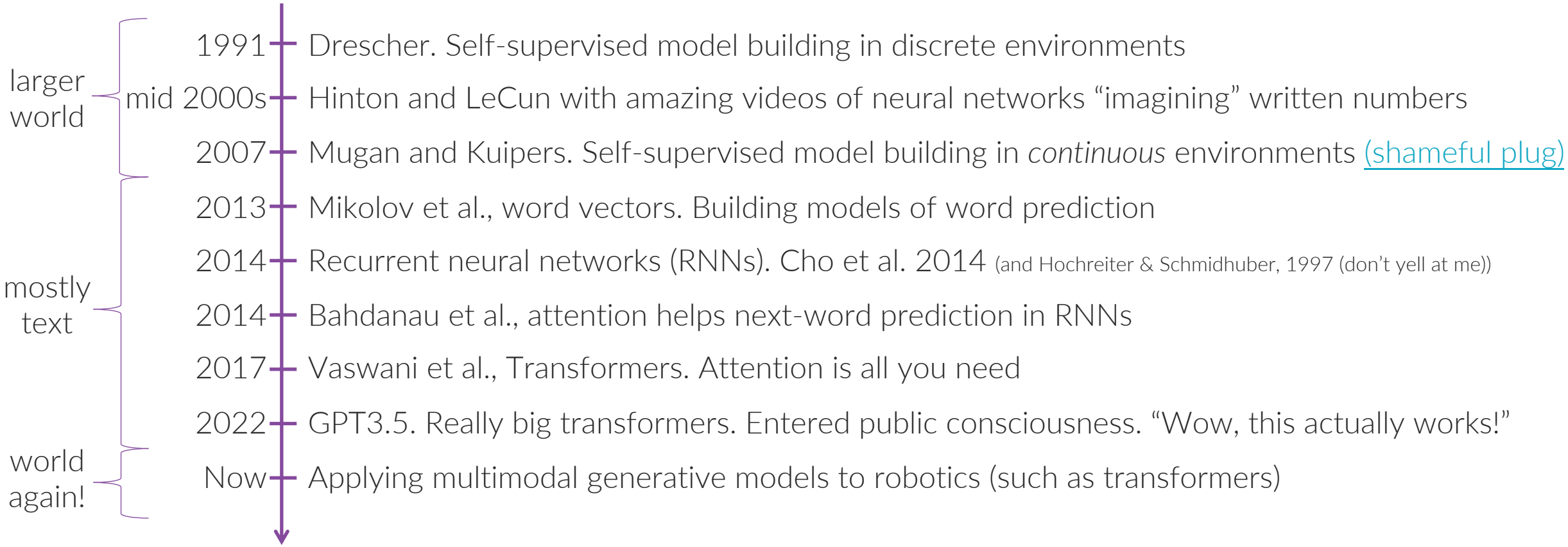
Learning in simulation

As we have seen, researchers have recently been using large neural networks to have robots actively learn in the world

- A compelling form of learning is self-supervision because the robot doesn't need external labels



A biased view of the progression of self-supervision

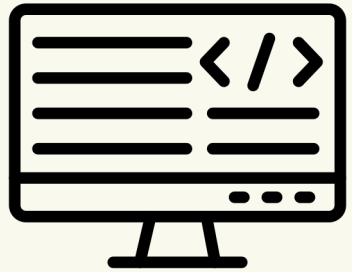


We are finally starting to move toward a theory integrating homeostasis and intelligence

Theory is called Active Inference and is the culmination of a lot of work in probabilistic modeling and inference

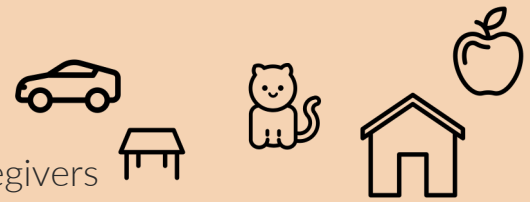
Internet content

Consuming blogs, news articles, essays, code, comics, videos, ...



Pre-literate Learning

What a toddler learns running around the living room and with shared attention with caregivers



Development of Mind

World ontology

- Places
- Objects
- Forms
- Agents
- Social patterns
- Number

Fundamental patterns

- Perceptual patterns
  - Force, inside-outside
- Action patterns
  - Grasping, running

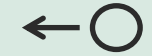
Origin of Life

Striving to maintain metabolism

Purpose begins with a connection between a sensor and effector

Moving toward resources

Moving away from threats



Active Inference

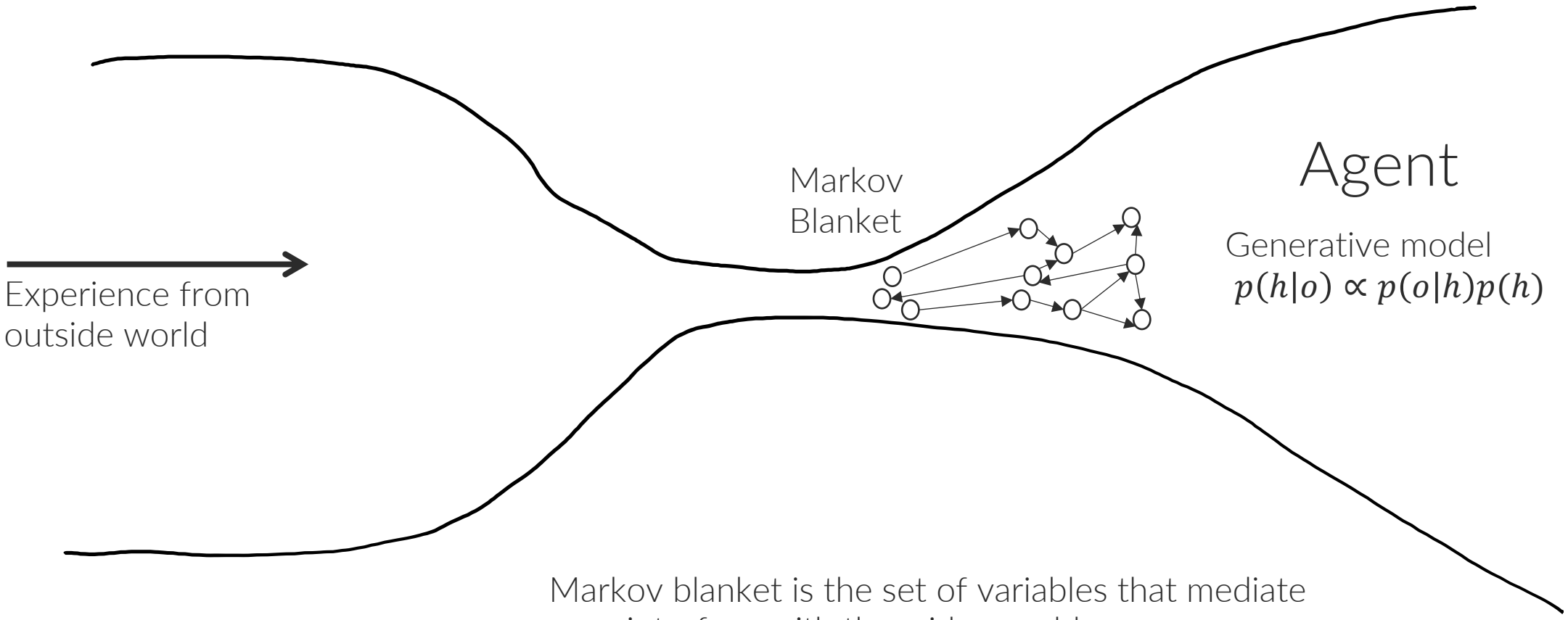
Active inference entails choosing actions that are expected to lead to observations that minimize free energy.

Free energy, in this context, is a measure that combines the surprise (or improbability) of sensory observations given the agent's model and the uncertainty (or entropy) of the agent's beliefs about the hidden states of the world.

Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active inference: The free energy principle in mind, brain, and behavior*.
Also see Namjoshi https://www.youtube.com/watch?v=3VpptCcInjU

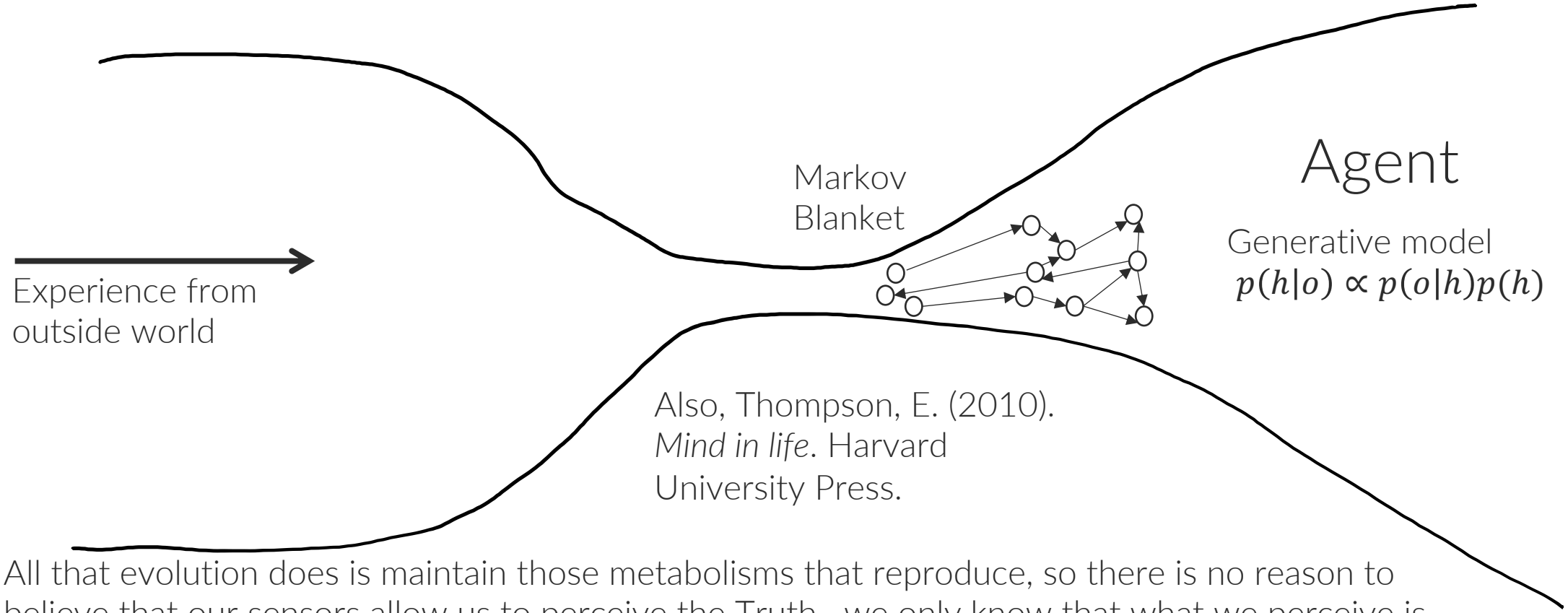
- Python library
  - paper https://arxiv.org/pdf/2201.03904.pdf
  - code https://github.com/infer-actively/pymdp
  - documentation https://pymdp-rtd.readthedocs.io/en/latest/index.html
  - start here https://pymdp-rtd.readthedocs.io/en/latest/notebooks/active_inference_from_scratch.html
- Discord https://discord.com/channels/798653042609094686/803420712445673472

I've always seen this as a "thin pipe"



Markov blanket is the set of variables that mediate your interface with the wider world

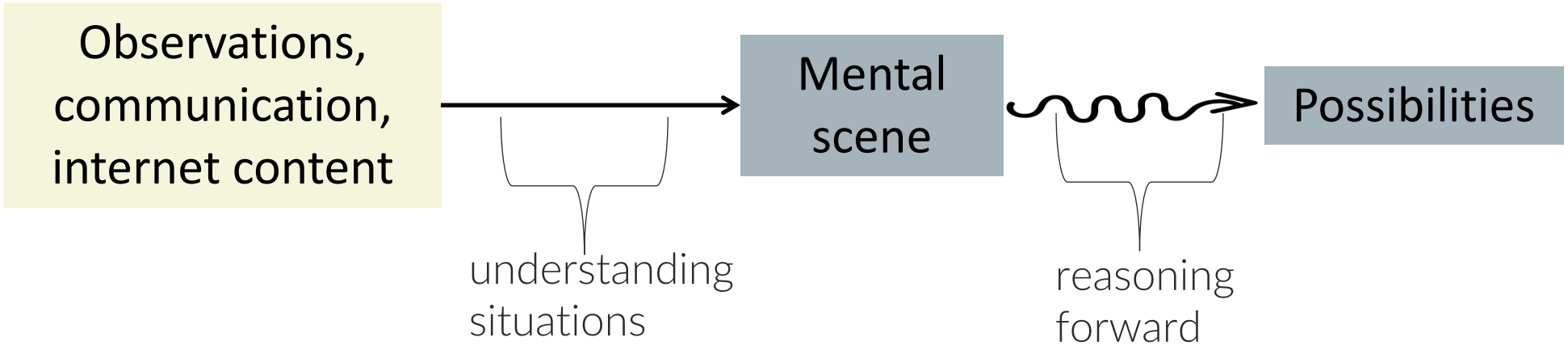
We don’t see the world as it is, only as it is useful



- All that evolution does is maintain those metabolisms that reproduce, so there is no reason to believe that our sensors allow us to perceive the Truth—we only know that what we perceive is mostly internally consistent and allows us to survive on Earth [1,2].
- To illustrate the point, there’s a kind of bird called the Common Cuckoo (Cuculus canorus) that lays its eggs in the nest of another kind of bird, often a small passerine, and the new parents raise it even though the imposter bird is six times as heavy as the real young’s parents [3].

1. Hoffman, D. (2019). *The case against reality: Why evolution hid the truth from our eyes*. WW Norton & Company.
2. Schreiter, M. L., Chmielewski, W. X., Ward, J., & Beste, C. (2019). How non-veridical perception drives actions in healthy humans: Evidence from synaesthesia. Philosophical Transactions of the Royal Society B, 374(1787), 20180574.
3. Krüger, O. (2007). Cuckoos, cowbirds and hosts: Adaptations, trade-offs and constraints. Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1486), 1873–1886.

Active Inference is a probabilistic theory



$$P(mental\ scene \mid stimuli)$$
$$\propto P(stimuli|mental\ scene)P(mental\ scene)$$

$P(next\ state|mental\ scene)$
Transition model

$P(stimuli|mental\ scene)$
State model

$P(mental\ scene)$
State prior

Example of building the mental scene

"The first dog grabbed a toy and then the second dog grabbed it and played tug of war. But, things got out of hand. The second dog won! Though, he growled harshly, and his hair rose like a mohawk. In the end they both ended in their bedrooms and their toys were taken away"

See the video at
https://pika.art/video/c96f7af6-606b-4a54-9628-8d0f95463c74

- The second video was generated using only that story of the first
- The algorithm doesn't know what as dog is as deeply as we would like, but it shows the progress of how AI can now generate internal representations of the outside world
- Of course, there isn't a little person, homunculus, inside of us watching the video. We "feel" the "pixels."

Conclusion

- Hardware Advances: Humanoid Robots Are Arriving
- Simulation Advances: Infrastructure for Accelerated Learning
- Algorithm Advances: Toward Artificial General Intelligence
  - Skills: Moving its body to achieve goals
  - Thinking: Understanding situations and forming high-level plans
    - Grounded Representation: Map sensory input to a representation
    - Reasoning: Determine possibilities from representation
      - Linear Thinking: Search, Inference, and Data Processing
      - Thought Jumping: Dynamically condensing time and space with abstraction
- The Evolutionary Paths of Intelligence: humans bottom-up, robots top-down

201 West 5th Street,
Suite 1575
Austin, TX. 78701