How to Expand the Capabilities of Large Language Models

Jonathan Mugan
@jmugan

Large Language Models Meetup
May 17th, 2023

DeUmbra

Outline

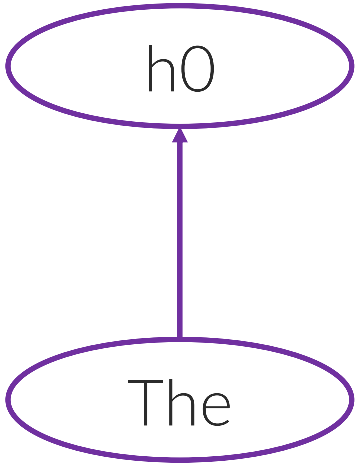
- How LLMs “think”
- Grounding LLMs with a cognitive foundation
- Expanding LLMs with GOFAI

Three technologies enabled ChatGPT

1. Language Modeling
2. Transformers
3. Instruction Tuning

It began with machine translation

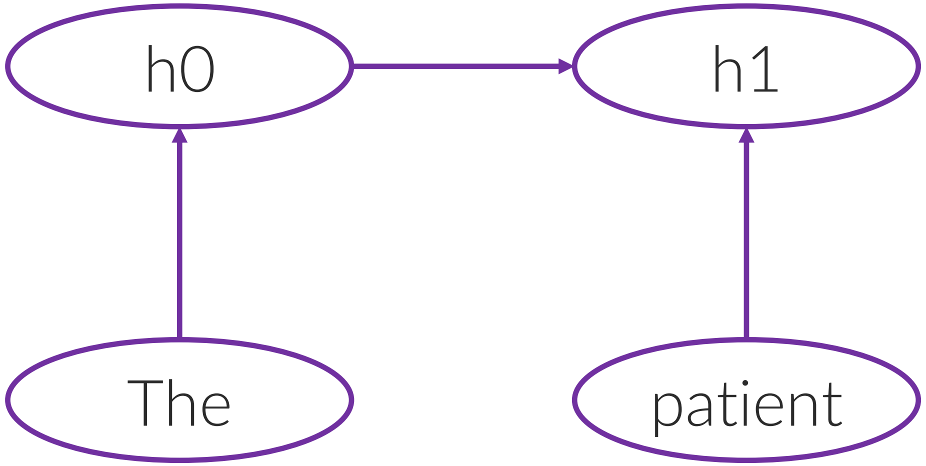
"The patient fell."



Using a recurrent neural network (RNN).

Encoding sentence meaning into a vector

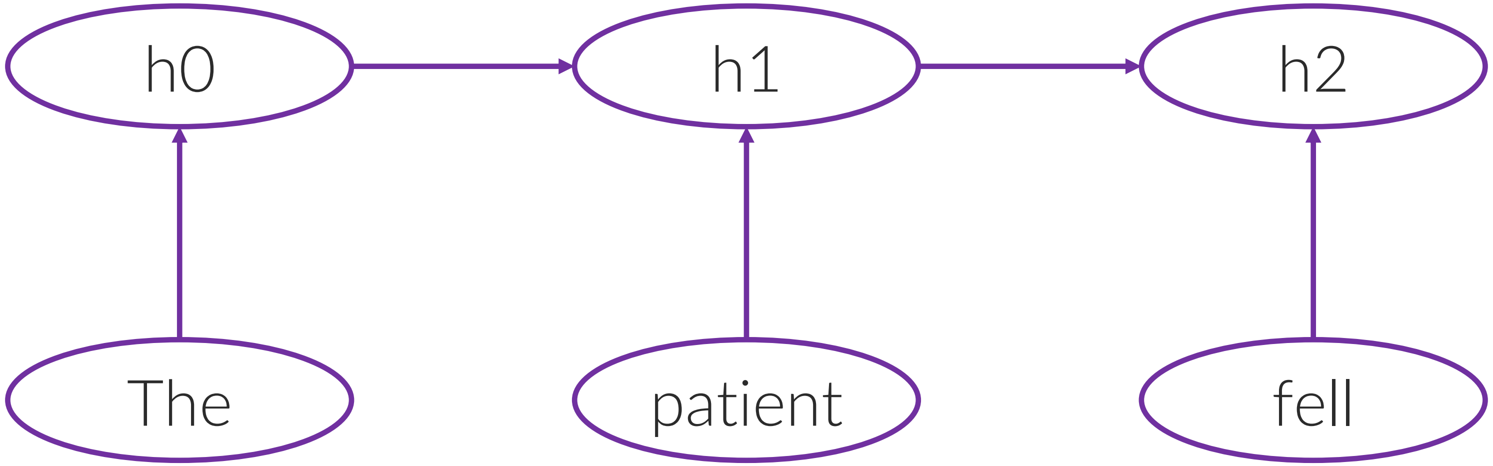
"The patient fell."



Using a recurrent neural network (RNN).

Encoding sentence meaning into a vector

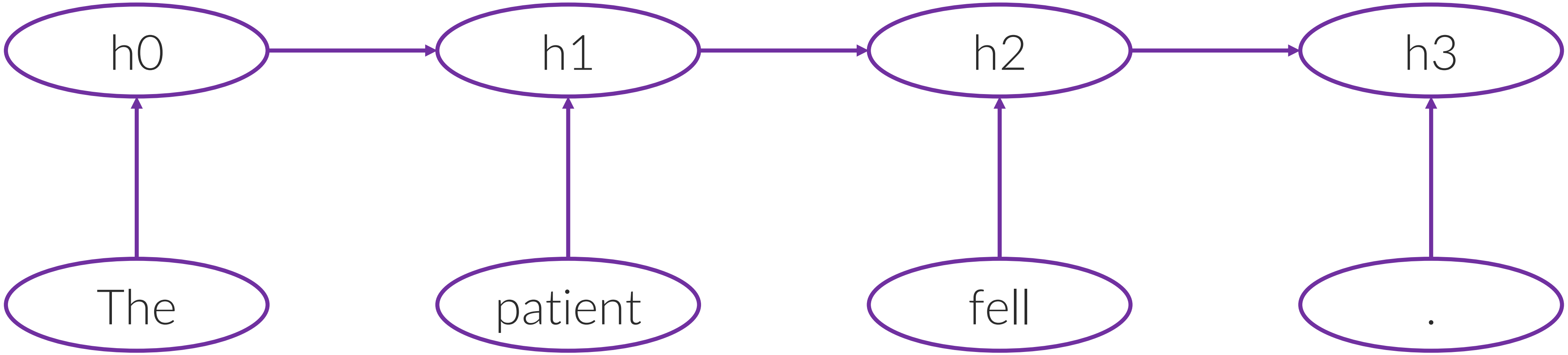
"The patient fell."



Using a recurrent neural network (RNN).

Encoding sentence meaning into a vector

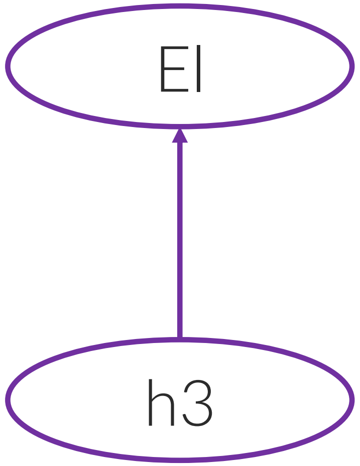
“The patient fell.”



RNN is like a hidden Markov model but doesn’t make the Markov assumption and benefits from a vector representation.

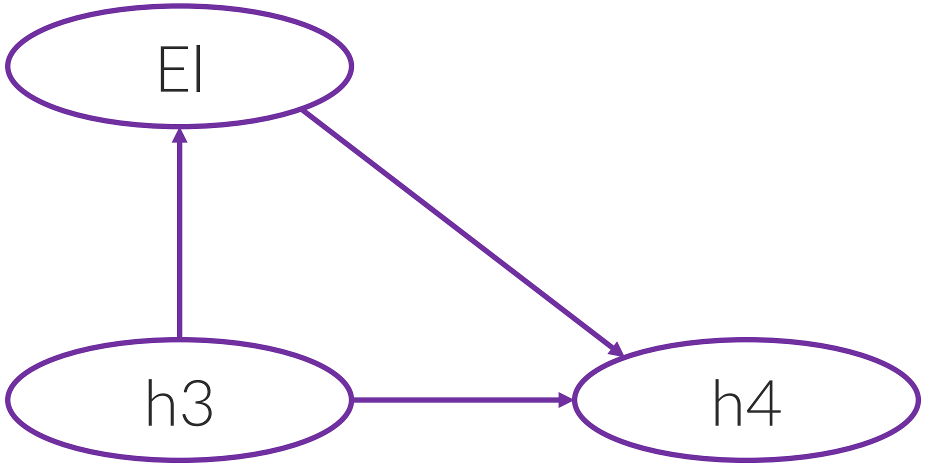
Decoding sentence meaning

Machine translation.



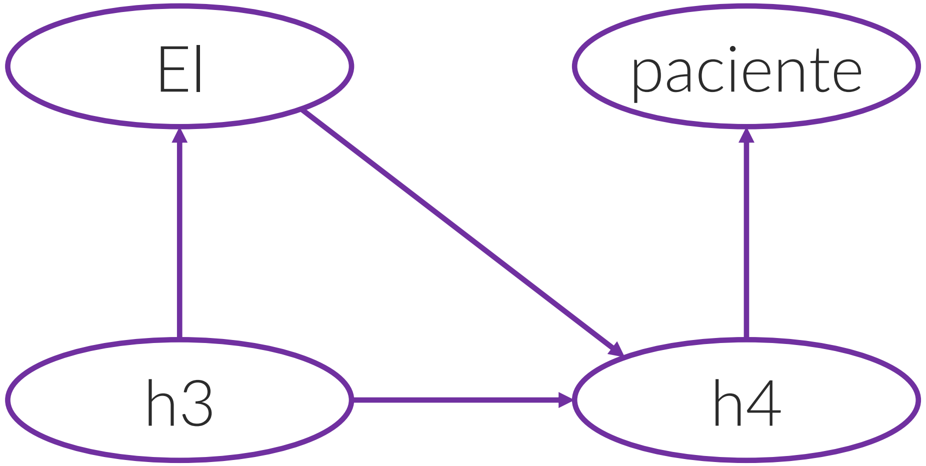
Decoding sentence meaning

Machine translation.



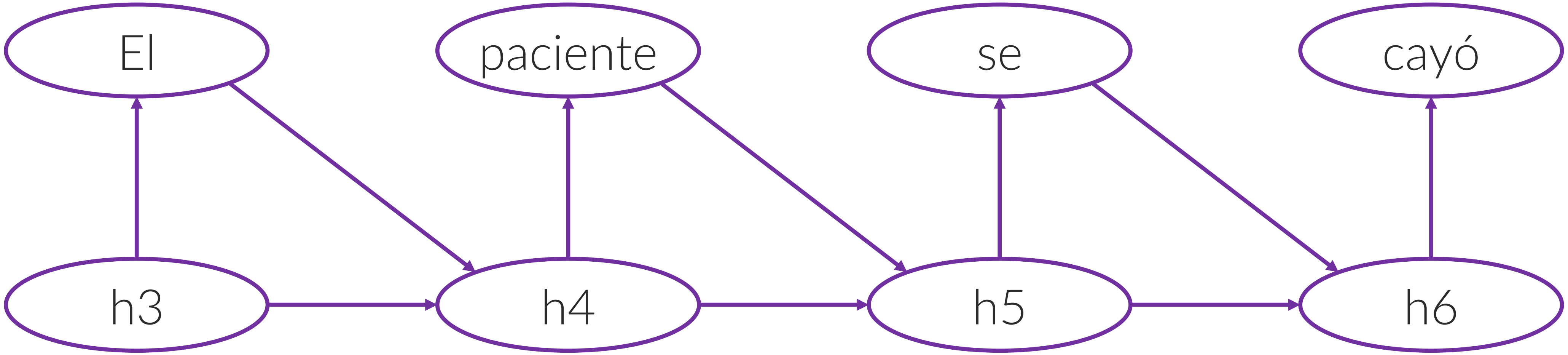
Decoding sentence meaning

Machine translation.



Decoding sentence meaning

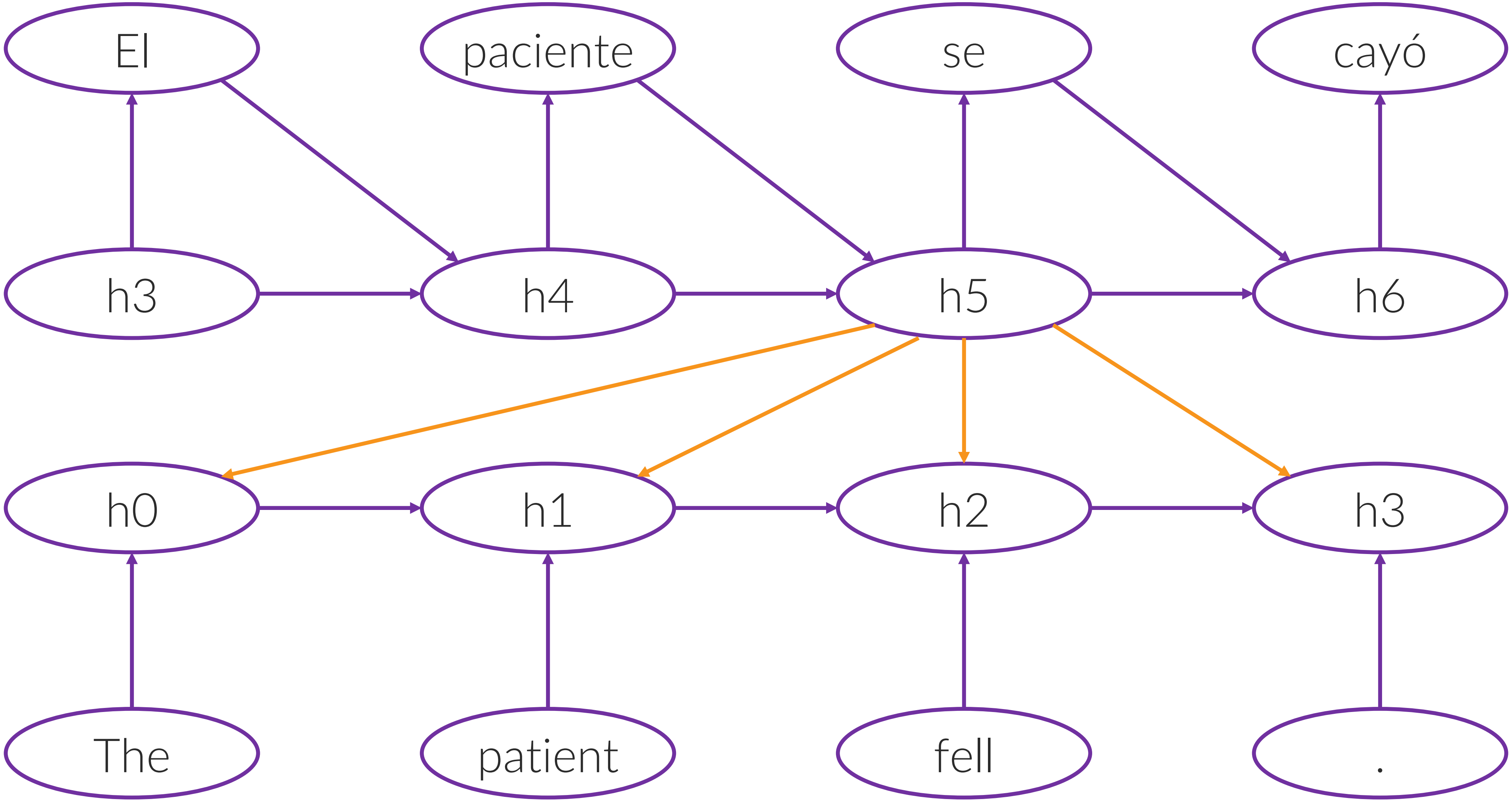
Machine translation.



[Cho et al., 2014]

- It keeps generating until it generates a stop symbol.
- It used a kind of interpolation from a huge set of training data.

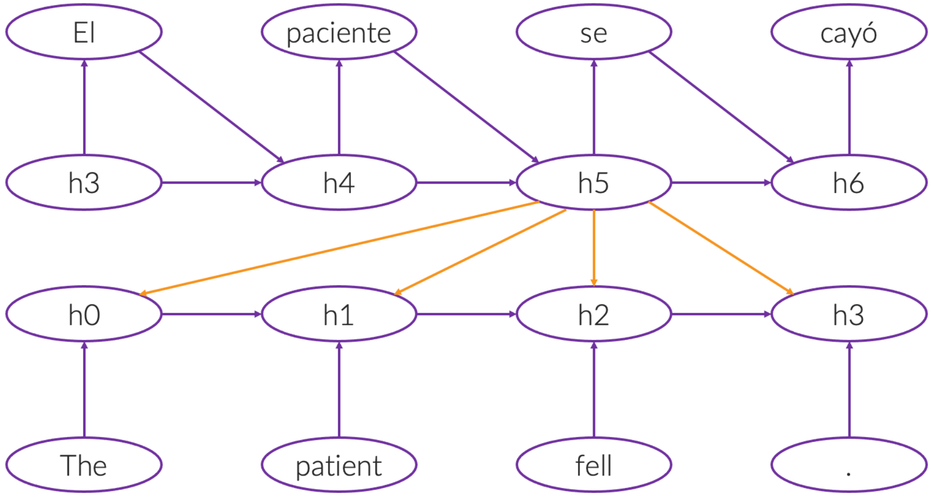
Attention [Bahdanau et al., 2014]



2

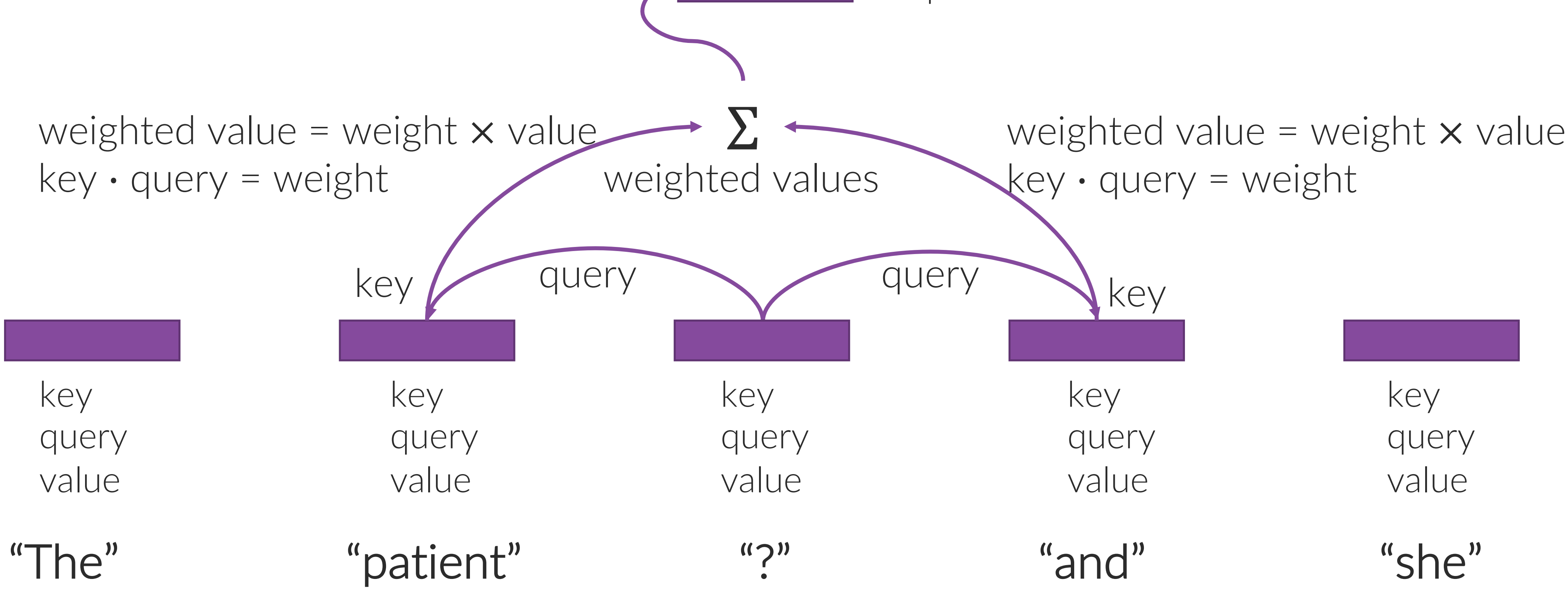
Transformers: Attention is all you need

https://arxiv.org/abs/1706.03762



How transformers think: tokens with keys, queries, and values

It computes a combined representation of itself combined with the representations of the tokens around it.



2

Recall from matrix multiplication

$X^{m \times n} Y^{n \times o} = Z^{m \times o}$

Sizes

v vocab size

d embedding size

l length of text input in tokens

Learned Weight Matrices

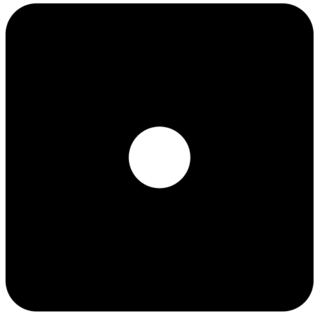
$E^{v \times d}$ embedding matrix

$W_Q^{d \times d_k}$ query matrix

$W_K^{d \times d_k}$ key matrix

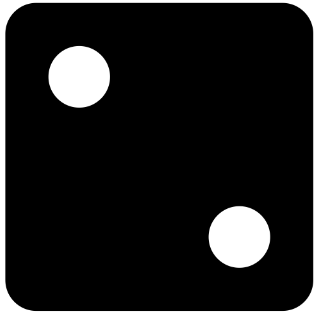
$W_V^{d \times d_v}$ value matrix

$W^{d_v \times d}$ linear



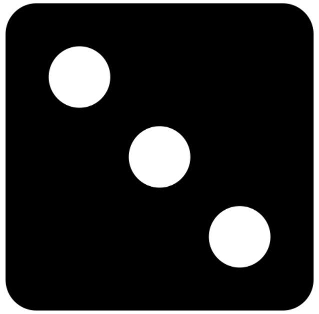
Use embedding matrix to get data

$$X^{l \times d}$$



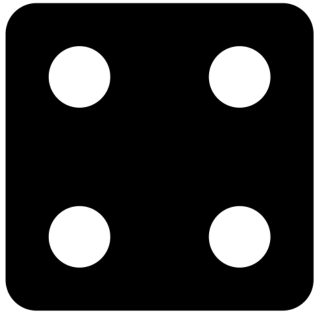
Compute queries, keys, and values

$$Q^{l \times d_k} = X^{l \times d} W_Q^{d \times d_k}$$
$$K^{l \times d_k} = X^{l \times d} W_K^{d \times d_k}$$
$$V^{l \times d_v} = X^{l \times d} W_V^{d \times d_v}$$



Compute attention weights

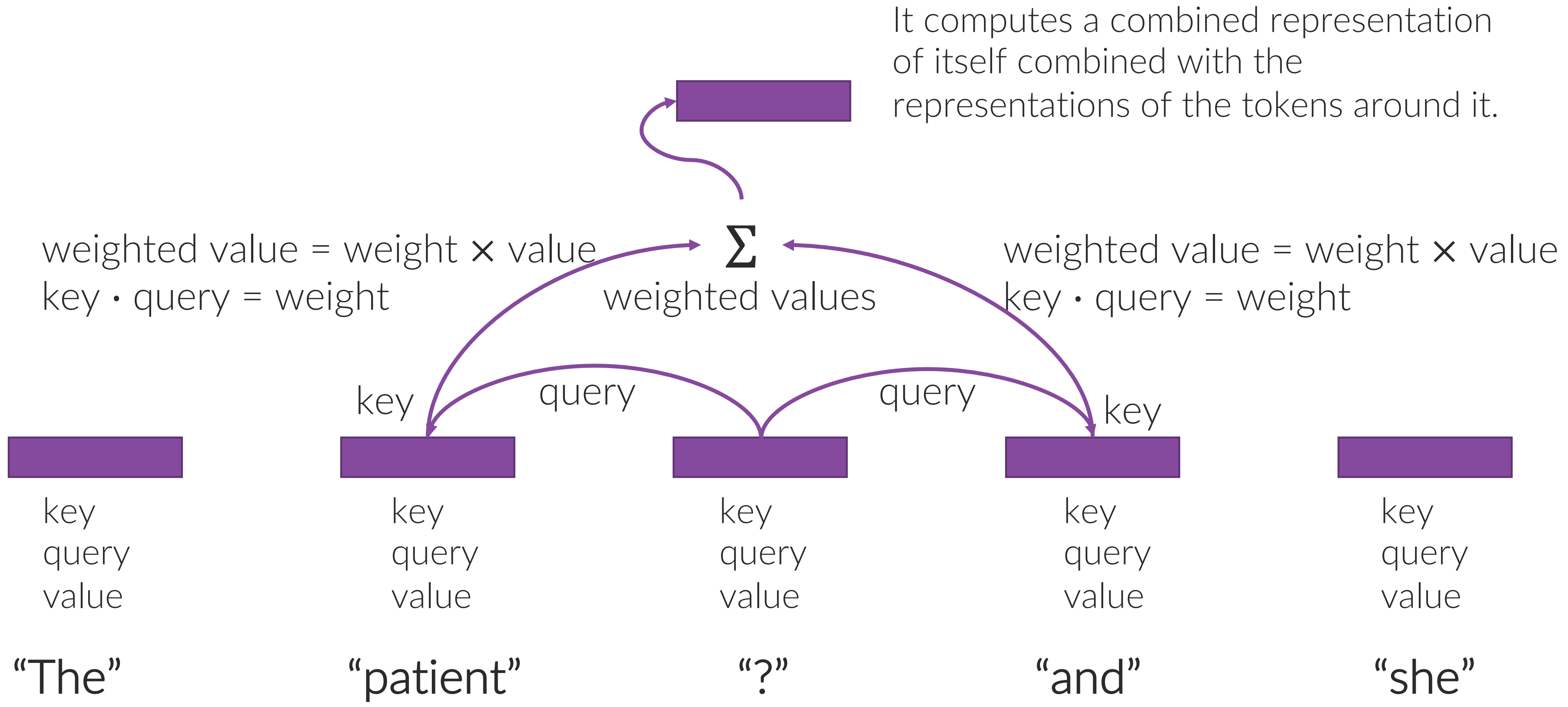
$$A^{l \times l} = QK^T$$



Use weights to get values and resize for next layer

$$\hat{X}^{l \times d_v} = A^{l \times l} V^{l \times d_v}$$
$$X_{next}^{l \times d} = \hat{X}^{l \times d_v} W^{d_v \times d}$$

How transformers think: tokens with keys, queries, and values

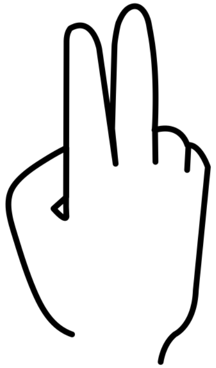


GPT-3: 96 heads, 1248 embedding size, 48 layers, plus details like positional encoding

Instruction Tuning: Reinforcement Learning with Human Feedback (RLHF)

Train an evaluation model to determine how good an output is.

1. Have humans rate outputs.
2. Train an evaluation model on those ratings. In RL, that evaluation model is called a *reward function*.



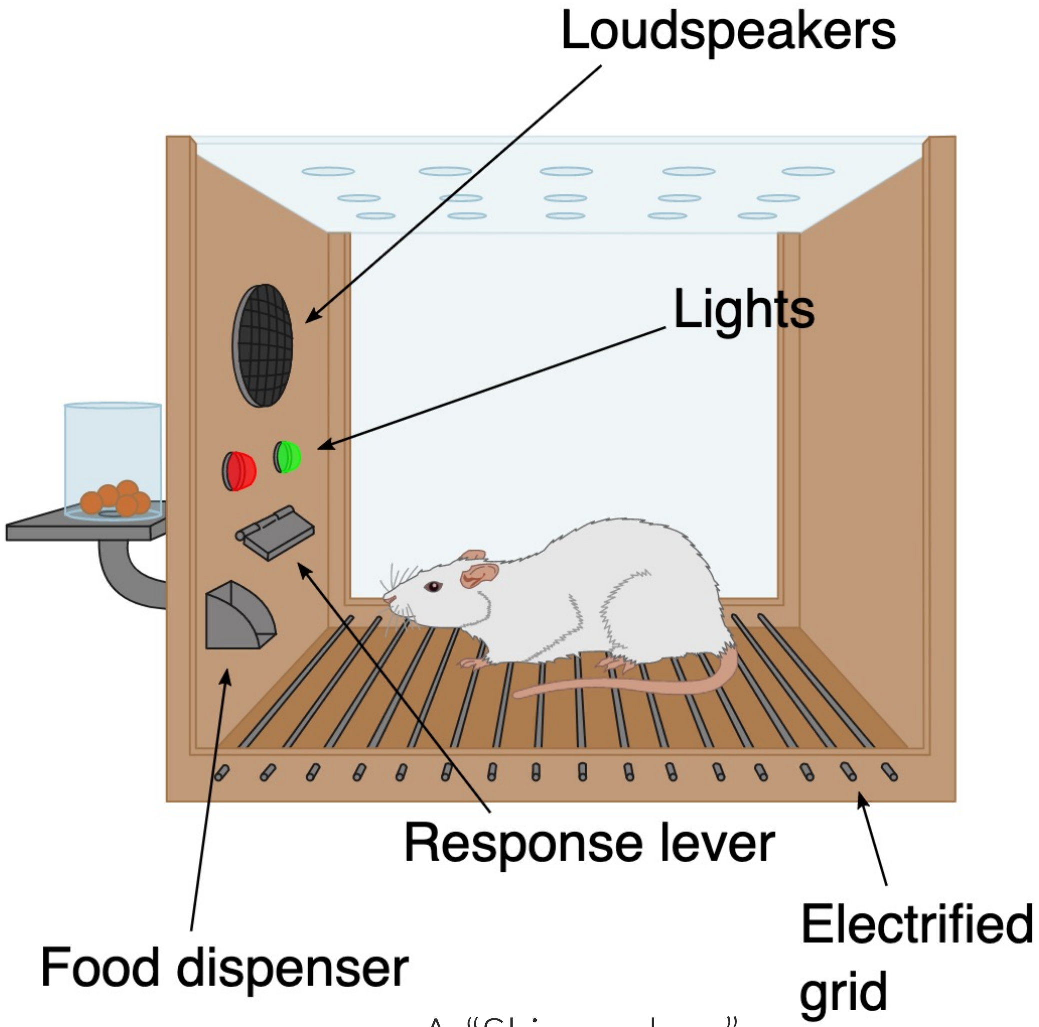
Use that evaluation model to guide autonomous learning.

- Begin with the language model trained on the internet.

RL is a gradual stamping in of behavior

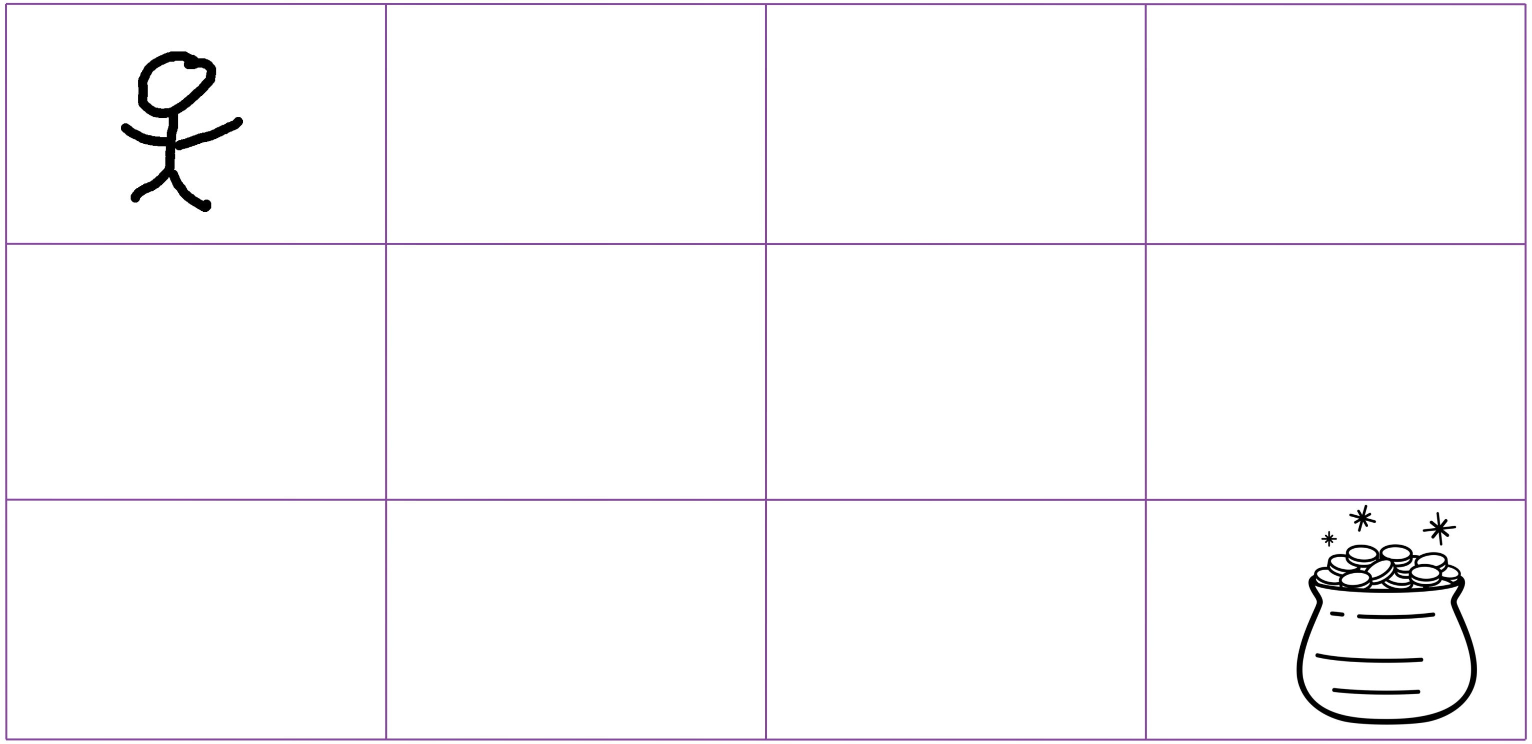
Reinforcement learning: the first 100 years

- Some behaviors arise more from a gradual stamping in [Thorndike, 1898].
- Became the study of Behaviorism [Skinner, 1953] (see Skinner box on the right).
- Formulated into artificial intelligence as Reinforcement Learning [Sutton and Barto, 1998].



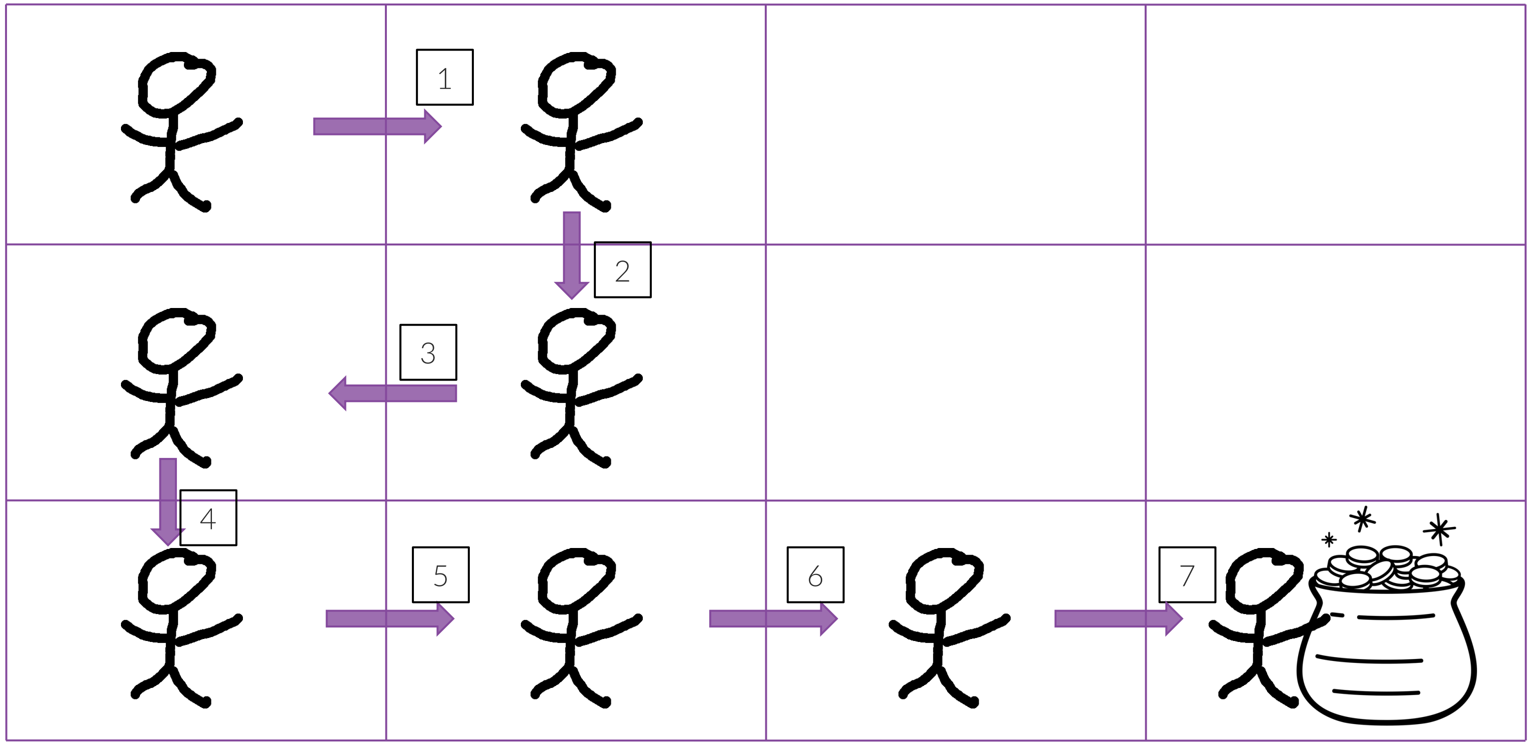
A “Skinner box”

RL in a nutshell: begin with random exploration



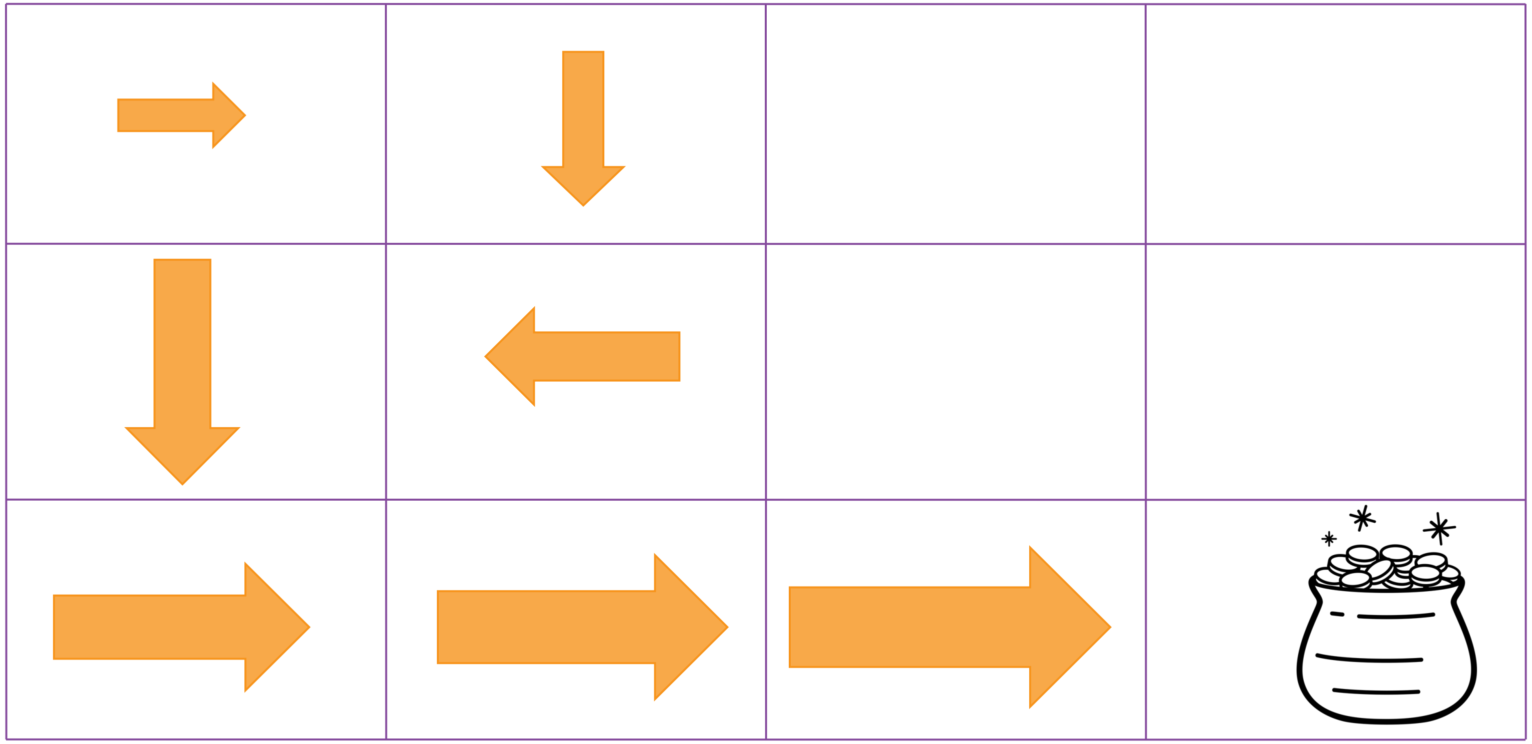
In reinforcement learning, the agent often begins by randomly exploring until it reaches its goal.

RL in a nutshell: begin with random exploration



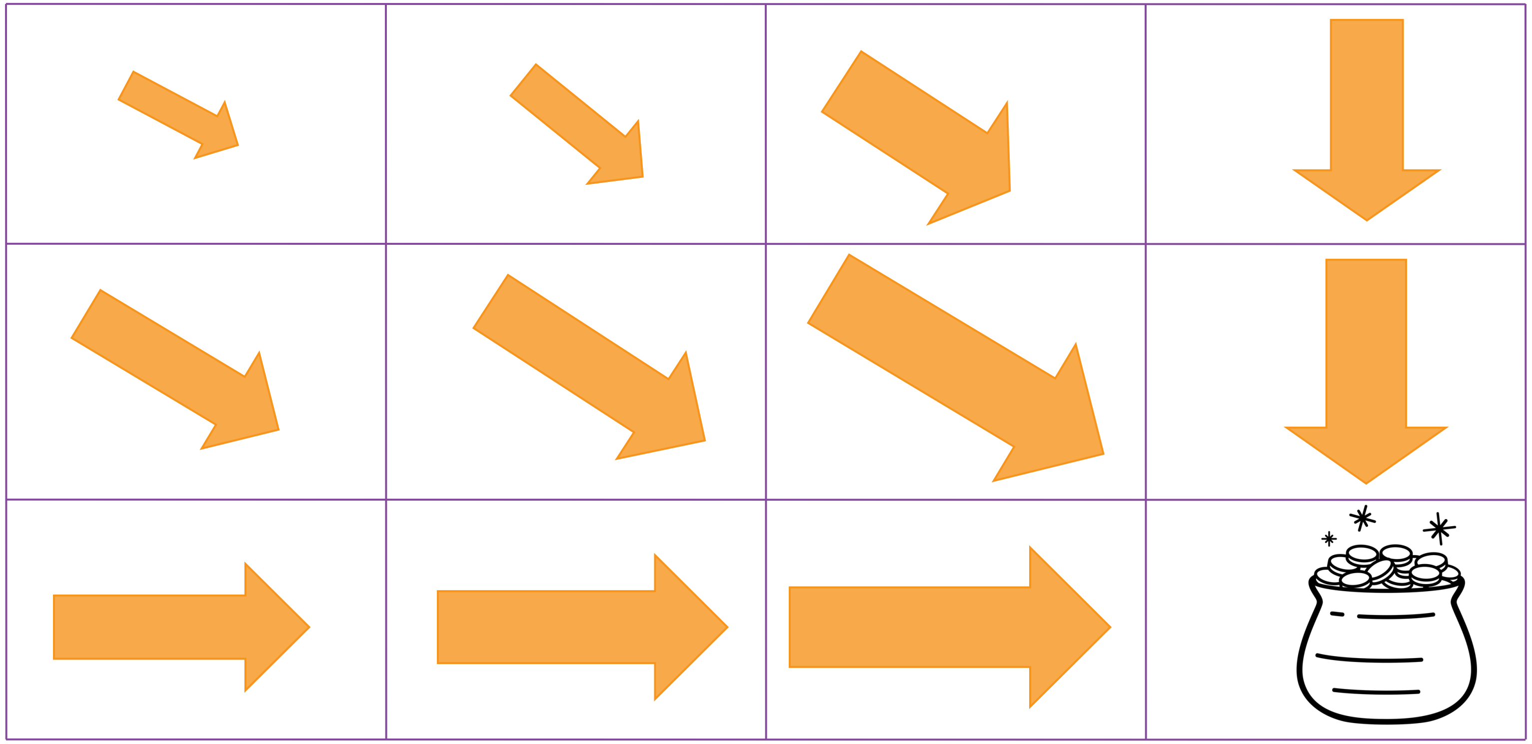
In reinforcement learning, the agent often begins by randomly exploring until it reaches its goal.

RL in a nutshell: remember what got you there



- When it reaches the goal, credit is propagated back to its previous states.
- Simplest case: the agent learns the function $Q^{\pi}(s,a)$, which gives the cumulative expected discounted reward of being in state s and taking action a and acting according to policy π thereafter. Modern uses PPO.

RL in a nutshell: learn a policy for behavior



Eventually, the agent learns the value of being in each state and taking each action and can therefore always do the best thing in each state. This behavior is then represented as a policy.

Instruction Tuning: Reinforcement Learning with Human Feedback (RLHF)

Train an evaluation model to determine how good an output is.

1. Have humans rate outputs.
2. Train an evaluation model on those ratings. In RL, that evaluation model is called a *reward function*.

Use that evaluation model to guide autonomous learning.

- Begin with the language model trained on the internet.

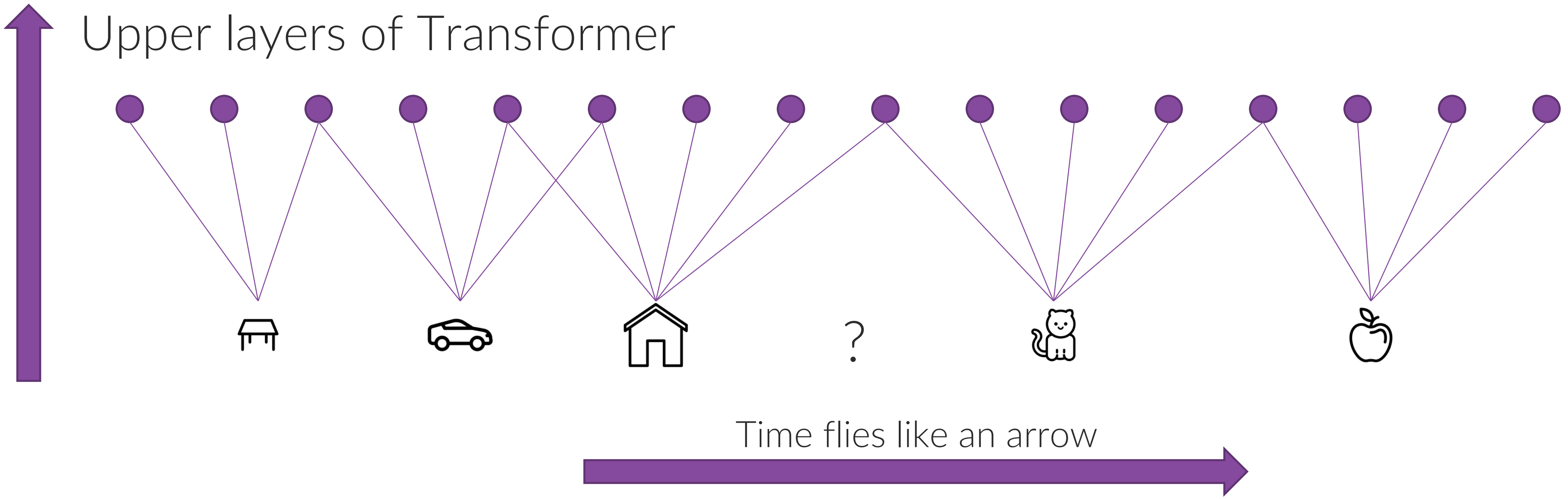
RLHF: Reinforcement Learning with Human Feedback
https://openai.com/research/instruction-following
https://arxiv.org/pdf/2203.02155.pdf

Outline

- How LLMs “think”
- Grounding LLMs with a cognitive foundation
- Expanding LLMs with GOFAI

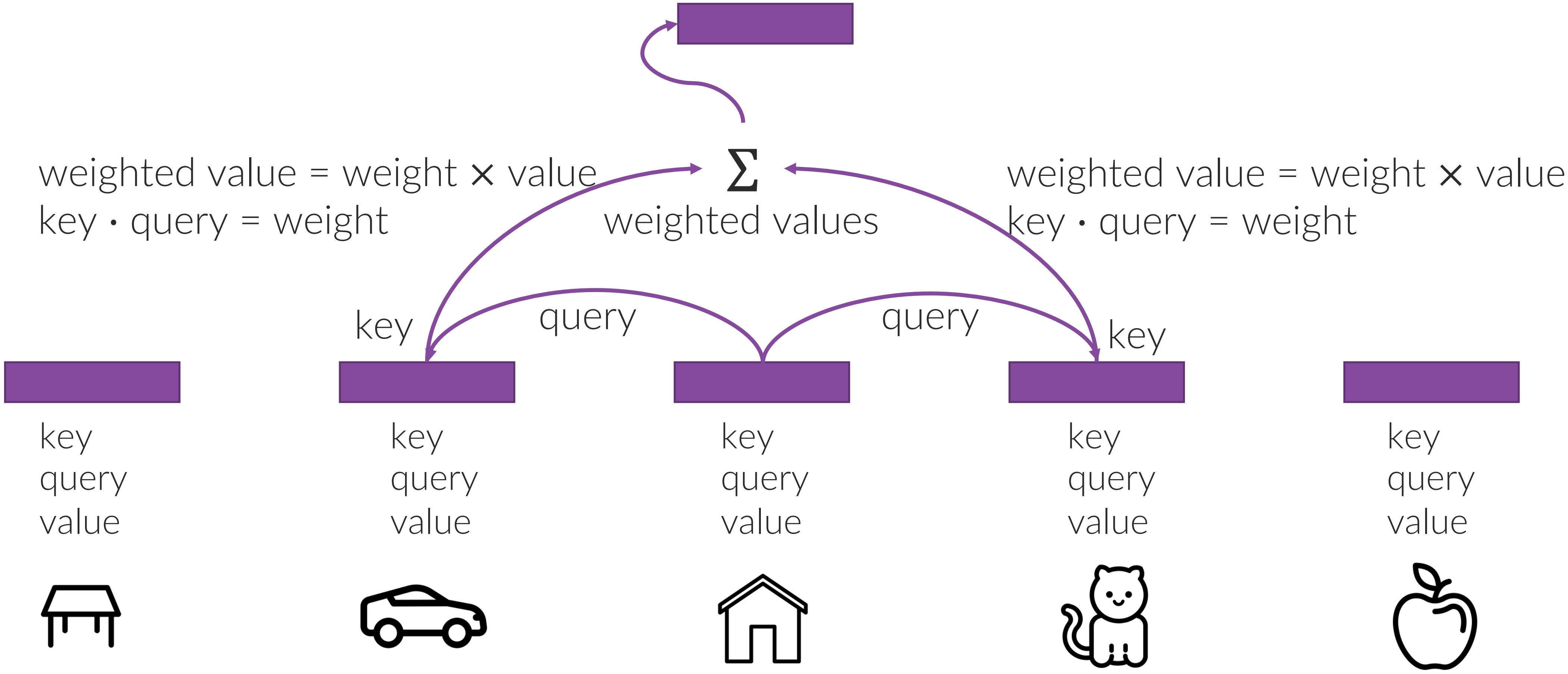
A cognitive foundation will lead to a deeper understanding

Multi-modal curriculum learning: objects instead of just words



Real-world objects

A cognitive foundation will lead to a deeper understanding



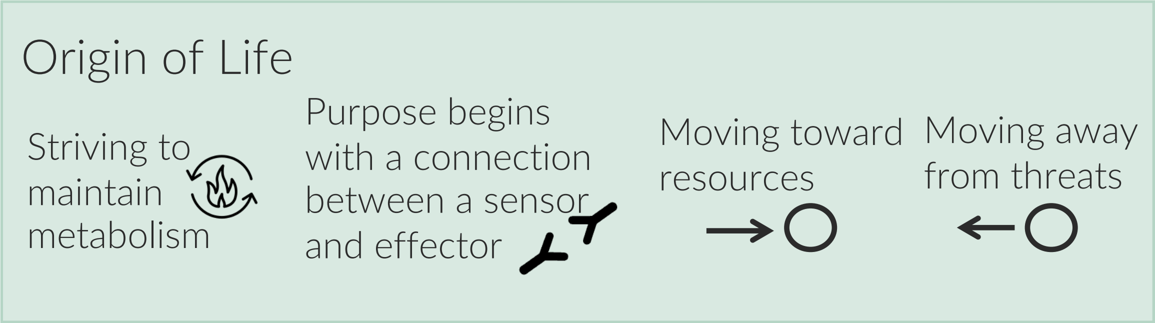
Origin of life

During evolution, purpose came into being when by chance the first sensor element connected to the first effector (motor) element [10].

The purpose of life is to maintain metabolism.

Actions allow the agent to run experiments to expose spurious correlations.

Curriculum learning entails specifying that some tokens are more important to predict than others. What is important to predict will depend on the type of robot or specialized AI you want to build.



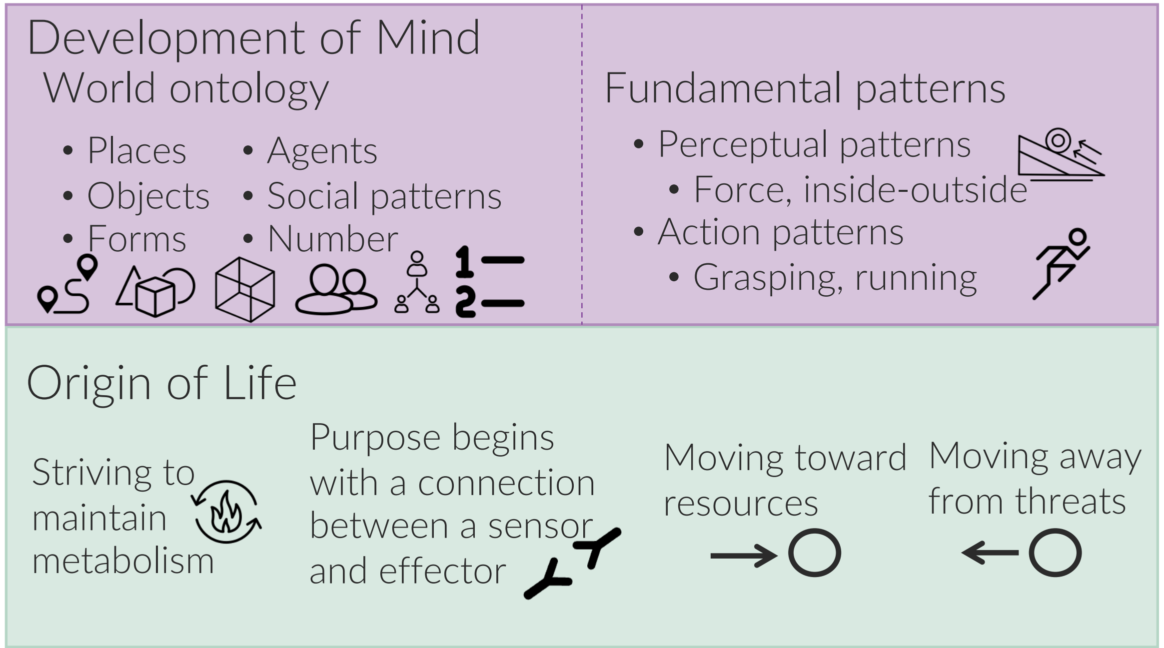
Development of mind

The developmental psychologist Elizabeth Spelke describes the ontology used by the human mind as consisting of six systems of core knowledge [11-12].

At this level,, curriculum learning entails training data that represents basic objects, relationships, and interactions

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 [3,4,7]. 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 [13]

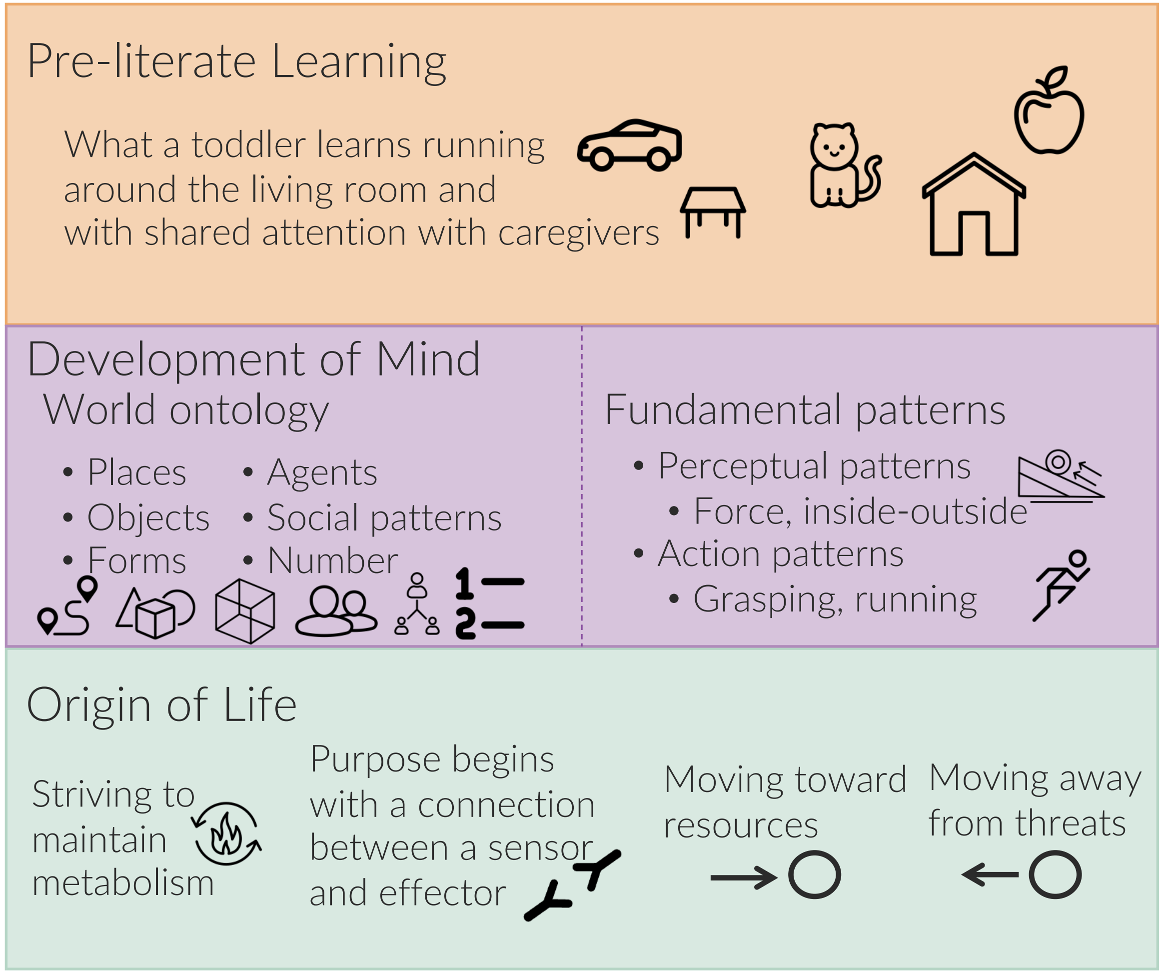


Pre-literate Learning

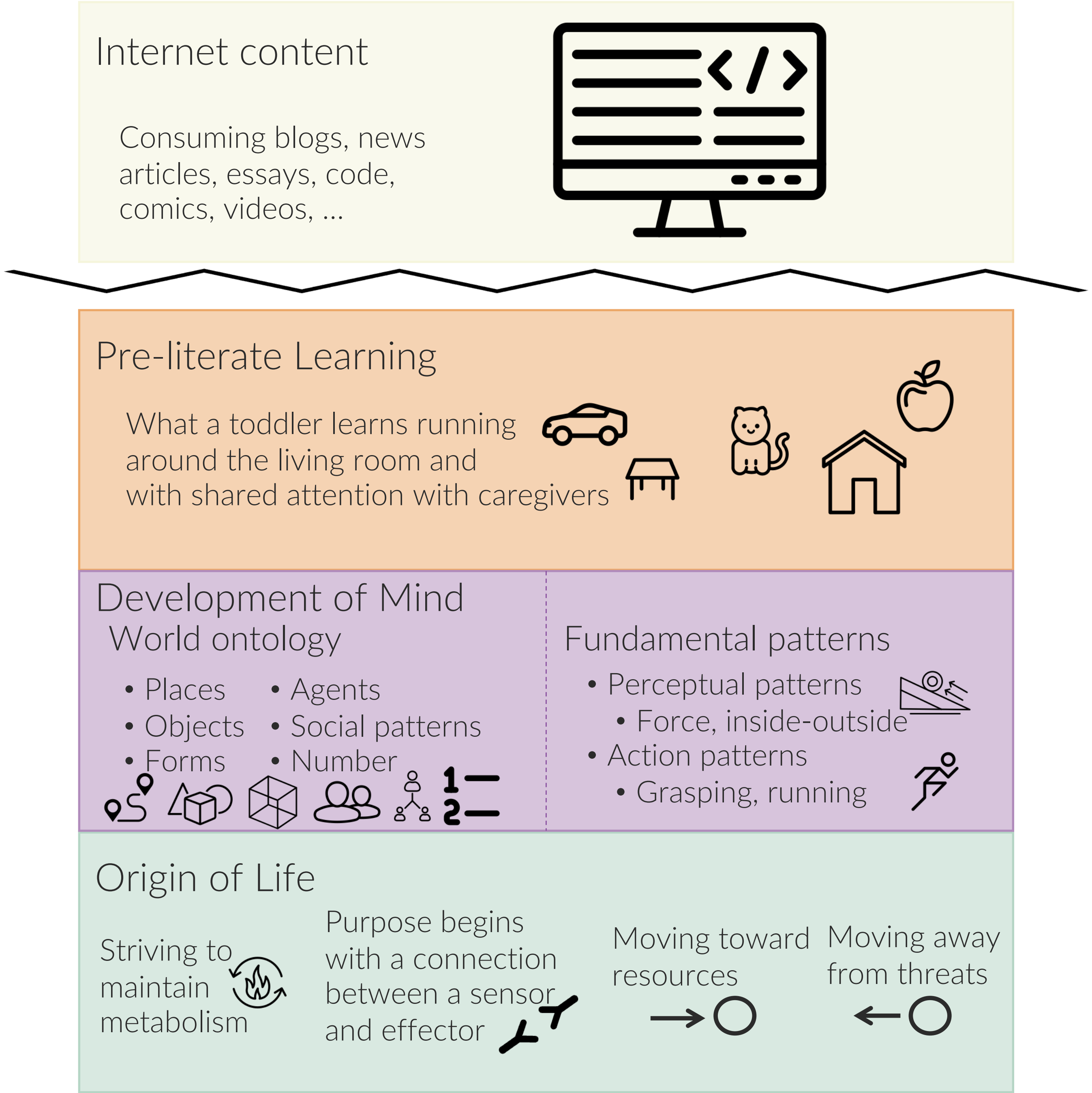
Children learn through exploration and through shared attention with caregivers [7,16,17].

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

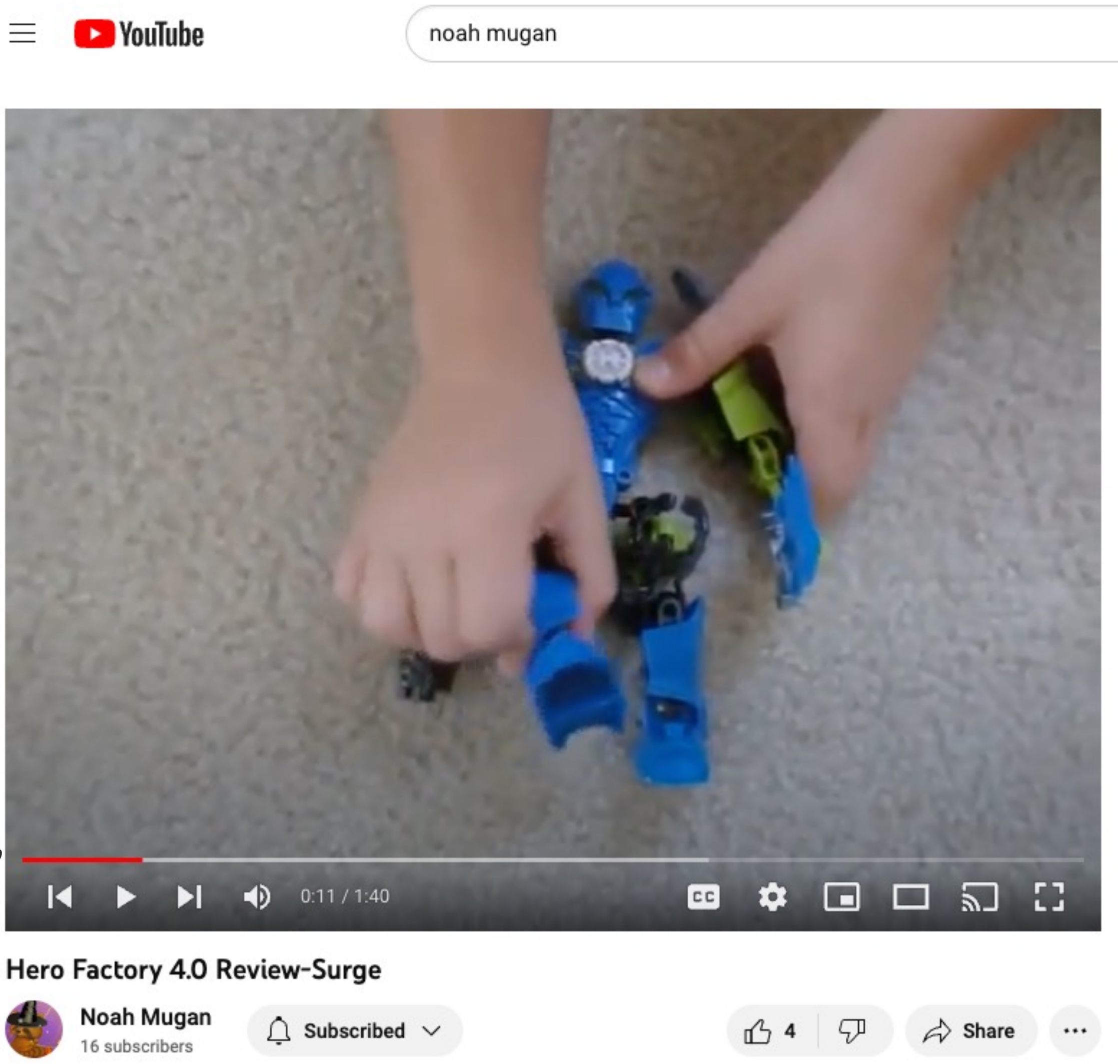
Tiny models
https://www.reddit.com/r/MachineLearning/comments/13j0spj/r_tiny_language_models_below_10m_parameters_or/



Current LLMs are trained only by what is floating on top



Imitation Learning from Video with Transformers



Robots can watch YouTube and learn to imitate, analogous to ChatGPT

Multimodal, language and object manipulation

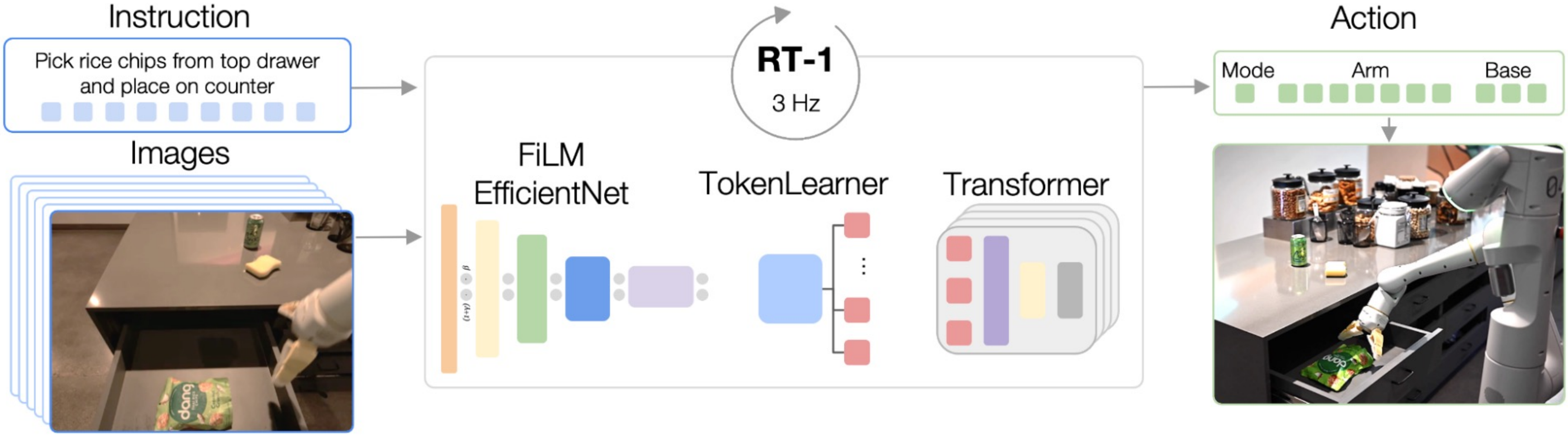
The trick is the tokenization of events in video, but Google has made some good progress

Robotics Transformer 1 (RT-1)

- Transformer model trained by copying demonstrations
- Predict the next most likely action based on what it has learned from the demonstrations

https://blog.google/technology/ai/helping-robots-learn-from-each-other/
https://ai.googleblog.com/2022/12/rt-1-robotics-transformer-for-real.html
https://robotics-transformer.github.io/assets/rt1.pdf

Imitation Learning from Video with Transformers



The trick is the tokenization of events in video, but Google has made some good progress

Image used with permission.
Thanks Keerthana Gopalakrishnan!
@keerthanpg

Robotics Transformer 1 (RT-1)

- Transformer model trained by copying demonstrations
- Predict the next most likely action based on what it has learned from the demonstrations

https://blog.google/technology/ai/helping-robots-learn-from-each-other/
https://ai.googleblog.com/2022/12/rt-1-robotics-transformer-for-real.html
https://robotics-transformer.github.io/assets/rt1.pdf

Reinforcement learning trained in simulation

AI2Thor by Allen AI
https://ai2thor.allenai.org/

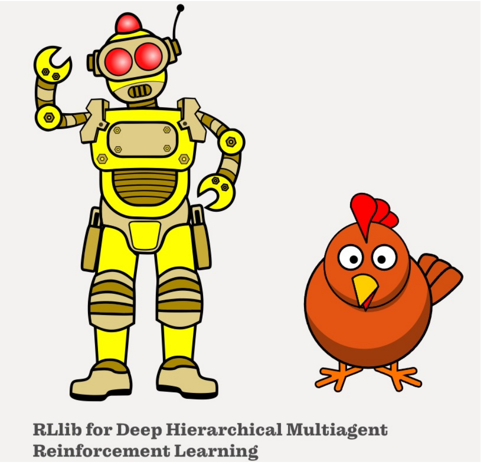
Microsoft Flight Simulator
https://www.flightsimulator.com/

Reinforcement learning trained in simulation

There has been some exciting progress in the area from DeepMind and others
https://www.deepmind.com/blog/from-motor-control-to-embodied-intelligence

Soccer!

Decision transformers
https://huggingface.co/blog/decision-transformers



https://deumbra.com/2022/08/rllib-for-deep-hierarchical-multiagent-reinforcement-learning/

Outline

- How LLMs "think"
- Grounding LLMs with a cognitive foundation
- Expanding LLMs with GOFAI

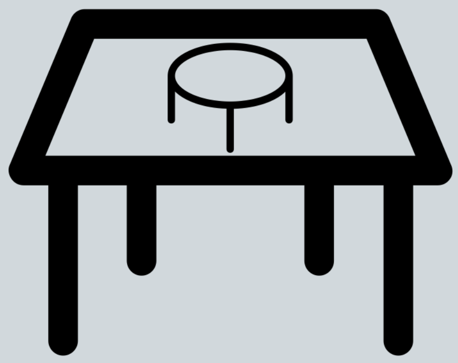
LLMs are great at what we do subconsciously

LLMs are great at this

LLMs are less good at this, especially multiple steps into novel possibilities

Words, communication, internet content → Mental scene → Possibilities

“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.

Possibilities for expanding LLMs with GOFAI

1. Expand to tools: WolframAlpha, web search, travel sites
2. Generate code to write GOFAI tools: first-order logic
3. Use simulation as a tool



Writing GOFAI Autonomously

Classic AI question, can birds fly?

$$\forall x\ bird(x) \rightarrow can\_fly(x)$$

Okay, okay

$$\forall x\ bird(x) \wedge flying\_bird(x) \rightarrow can\_fly(x)$$

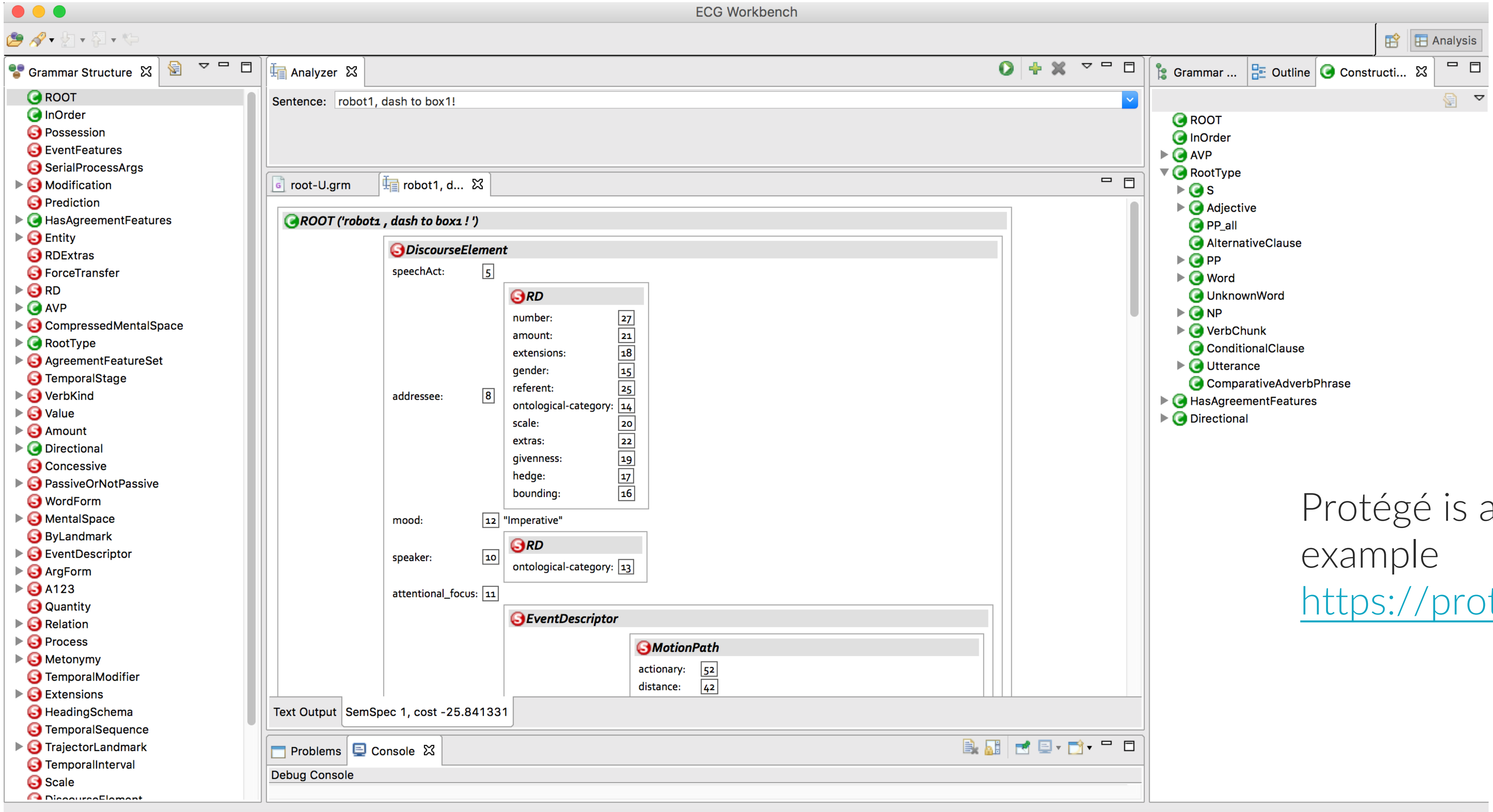
Okay, okay

$$\forall x\ bird(x) \wedge flying\_bird(x) \\ \wedge\ not\ broken\_wing(x) \rightarrow can\_fly(x)$$

WKRP "As God as my witness, I thought turkeys could fly”
https://www.youtube.com/watch?v=If3mgmEdfwg&ab_channel=EpicHouston

Okay, okay, what if it is covered in maple syrup?

Building the hypothesis space: Manual Configuration



Protégé is another familiar example https://protege.stanford.edu/

Embodied construction grammar https://github.com/icsi-berkeley

LLMs could create simulations in Unity

"The table is on the table"

By directly applying force to the bottom one in the game engine, it observes the top table fall

Recap of the talk

Three technologies enabled ChatGPT

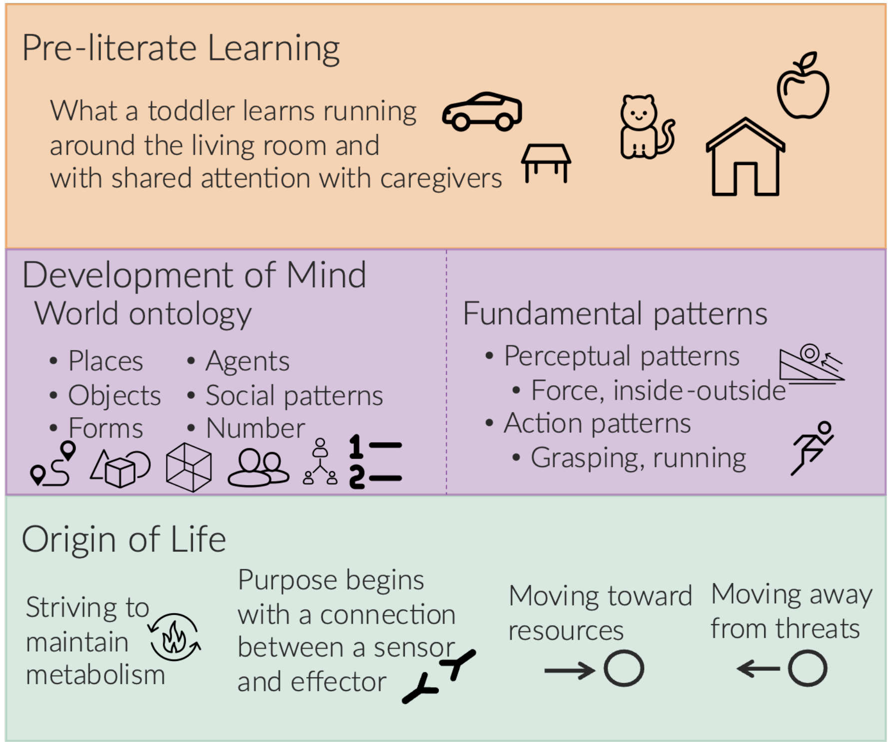
1 Language Modeling

2 Transformers

3 Instruction Tuning

We can expand LLMs by

1. Enabling tool use
2. Enabling it to write GOFAI code
3. Enabling it to use simulation as a tool



We can deepen LLM understanding by building a cognitive foundation using multi-modal curriculum learning

To learn more, see my article in *The Gradient*
https://thegradient.pub/grounding-large-language-models-in-a-cognitive-foundation/

References (1 of 2)

[1] Jackendoff, R. (2012). *A user's guide to thought and meaning*. Oxford University Press.
[2] Bergen, B. K. (2012). *Louder than words: The new science of how the mind makes meaning*. Basic Books.
[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] Feldman, J. (2006). *From molecule to metaphor: A neural theory of language*. MIT press.
[6] Shanahan, M. (2010). *Embodiment and the inner life: Cognition and Consciousness in the Space of Possible Minds*. Oxford University Press, USA.
[7] Mandler, J. (2004). *The Foundations of Mind, Origins of Conceptual Thought*. Oxford University Press.
[8] Bray, D. (2009). *Wetware: A computer in every living cell*. Yale University Press.
[9] Pross, A. (2016). *What is life?: How chemistry becomes biology*. Oxford University Press.
[10] Gaddam, S., & Ogas, O. (2022). Journey of the Mind: How Thinking Emerged from Chaos. WW Norton.
[11] Spelke, E. (1994). Initial knowledge: Six suggestions. Cognition, 50(1–3), 431–45.
[12] Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. Psychological Review, 99(4), 605.
[13] Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.
[14] Ballard, D. H. (2015). Brain computation as hierarchical abstraction. MIT Press.
[15] Kuipers, B., Browning, R., Gribble, B., Hewett, M., & Remolina, E. (2000). The spatial semantic hierarchy. Artificial Intelligence.
[16] Gopnik, A. (2009). The Philosophical Baby: What Children's Minds Tell Us About Truth, love, and the meaning of life. Farrar Straus & Giroux.
[17] Tomasello, M. (2019). *Becoming human: A theory of ontogeny*. Harvard University Press.

References (2 of 2)

[18] Vemprala, S., Bonatti, R., Bucker, A., & Kapoor, A. (2023). ChatGPT for Robotics: Design Principles and Model Abilities. Microsoft. https://www.microsoft.com/en-us/research/uploads/prod/2023/02/ChatGPT___Robotics.pdf
[19] Evans, R. (2020). *Kant's Cognitive Architecture* [PhD Thesis]. Imperial College London.
[20] Brandom, R. (2000). *Articulating reasons*. Harvard University Press.
[21] Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41–58). Brill.
[22] Goodman, N. D., Tenenbaum, J. B., & Contributors, T. P. (2016). *Probabilistic Models of Cognition* (Second). http://probmods.org/v2
[23] Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.
[24] Gilbert, D. (2006). *Stumbling on happiness*. Knopf.
[25] Hoffman, D. (2019). *The case against reality: Why evolution hid the truth from our eyes*. WW Norton & Company.
[26] 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.
[27] 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.
[28] Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., & others. (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4. ArXiv Preprint ArXiv:2303.12712.
[29] Li, K., Hopkins, A. K., Bau, D., Viégas, F., Pfister, H., & Wattenberg, M. (2022). Emergent world representations: Exploring a sequence model trained on a synthetic task. ArXiv Preprint ArXiv:2210.13382.
[30] Nanda, N., Chan, L., Liberum, T., Smith, J., & Steinhardt, J. (2023). Progress measures for grokking via mechanistic interpretability. ArXiv Preprint ArXiv:2301.05217.
[31] Schick, T., Dwivedi-Yu, J., Dessì, R., Raileanu, R., Lomeli, M., Zettlemoyer, L., Cancedda, N., & Scialom, T. (2023). Toolformer: Language models can teach themselves to use tools. ArXiv Preprint ArXiv:2302.04761.
[32] Berk, L. E. (1993). Infants, Children, and Adolescents. Allyn and Bacon.
[33] Trumper, Klaus. (2023). *Artificial Intelligence: Why AI Projects Succeed or Fail.*
[34] Chen, L., Lu, K., Rajeswaran, A., Lee, K., Grover, A., Laskin, M., Abbeel, P., Srinivas, A., & Mordatch, I. (2021). Decision transformer: Reinforcement learning via sequence modeling. Advances in Neural Information Processing Systems, 34, 15084–15097.
[35] Surís, D., Menon, S., & Vondrick, C. (2023). ViperGPT: Visual inference via python execution for reasoning. ArXiv Preprint ArXiv:2303.08128.

DeUmbra

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