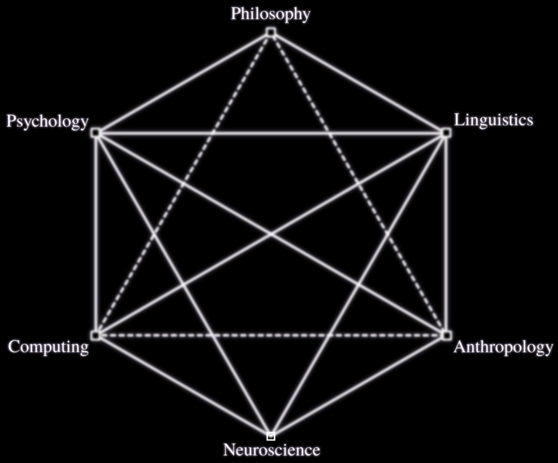
What Would it Take to Teach a Computer to Think?

Eilene Tomkins-Flanagan
Department of Cognitive Science, Carleton University

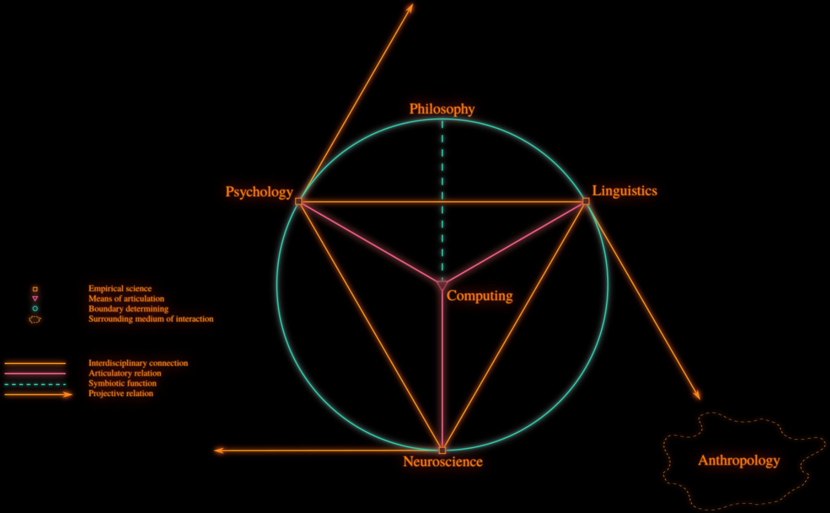
ANIMUS

Carleton University

What Is Cognitive Science?



What Is Cognitive Science?

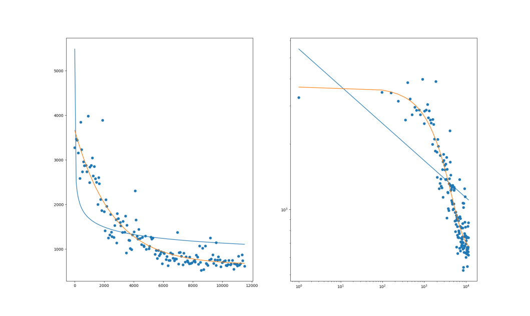


Mind as Machine

- Computation is a descriptive language for physical processes
- It describes how information-bearing states causally influence one another to change over time
- We can precisely test theories by making computational models that predict the causal structure mediating the states we measure
- So, artificial intelligence in cognitive science is a set of methodological tools we use to make our theories precise and testable

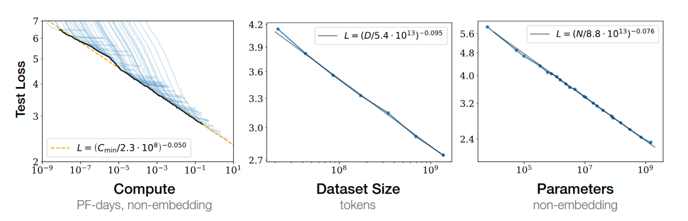
Learning's Kinda Like Searching Innit?

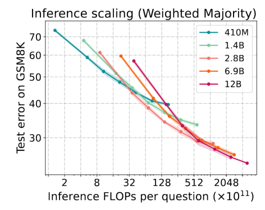
Here's an example:



What do learning curves tell us about memory and skill acquisition?

What Doesn't It Work Like





Reproduced from Kaplan et al. (2020) and Wu et al. (2025)

Search and Decision

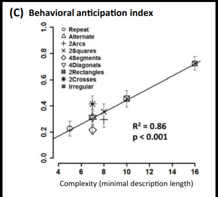
- Heathcote et al.[1] find that learning often doesn't follow a power law, but an exponential law
- Evans et al. elaborate that decision times are log-normally distributed about the main effect curve distribution of decision times is explained by drift diffusion decision models (Myers et al., 2022)
  - describe the behaviour of an attractor neural network with small basins of attraction, and
  - an evidence accumulation process driving decision
- You explore through representations until you find the correct decision, and learn by finding better exploration strategies

[1] Source of the graphic on slide 4

Characteristic Functions

- The main effect results from decreasing uncertainty about the correct response to a task as it becomes more familiar
  - We can think of this improvement like a search for more efficient procedures for reaching a correct decision
  - A power decay is characteristic of a memorization-like heuristic search strategy, with diminishing returns
  - An exponential decay is characteristic of a divide-and-conquer search strategy, that stays efficient
- So how is it that brains don't need to read the whole internet to learn anything?

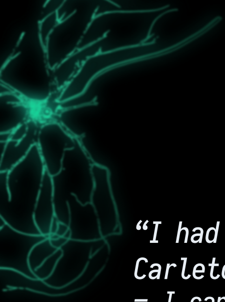
Well, Maybe They Learn Programs



- Dehaene et al. (2022) theorize that mental representations might be **programs that efficiently represent** the world would make sense, as Hutter (2005, Introduction) shows that with an efficiently representative program, agents can make **optimal decisions**
- So, then, **how can mental states be programs?** Can we learn with them?

How to Learn a Program

- First, we need to show that mental states can encode programs; in Tomkins-Flanagan and Kelly (2024), we showed that this is possible
- Second, we need to find efficient and natural encodings; this is an ongoing project, with work appearing in Hanley et al. (2025) and Tomkins-Flanagan et al. (2025)
- Third, we need to show how the right representations can constrain searching to enable it to be efficient; we began working on this problem in Tomkins-Flanagan et al. (2025)
- Then, we need to build brainlike networks that are able to take advantage of our representation scheme to learn efficiently



“I had no idea this kind of cutting edge research was happening at Carleton!”
– I can’t remember who said this

References I

- Dehaene, S., Al Roumi, F., Lakretz, Y., Planton, S., & Sablé-Meyer, M. (2022). Symbols and mental programs: A hypothesis about human singularity. *Trends in Cognitive Sciences*, *26*(9), 751–766. https://doi.org/10.1016/j.tics.2022.06.010
- Evans, N. J., Brown, S. D., Mewhort, D. J. K., & Heathcote, A. (2018). Refining the law of practice. *Psychological review*. https://doi.org/doi.org/10.1037/rev0000105
- Hanley, C., Tomkins-Flanagan, E., & Kelly, M. A. (2025). Hey pentti, we did (more of) it!: A vector-symbolic lisp with residue arithmetic. *2024 International Joint Conference on Neural Networks (IJCNN)*.
- Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review*, *7*(2), 185–207. https://doi.org/10.3758/BF03212979

References II

- Hutter, M. (2005). *Universal artificial intelligence: Sequential decisions based on algorithmic probability*. Springer.
- Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., & Amodei, D. (2020). Scaling laws for neural language models. https://arxiv.org/abs/2001.08361
- Myers, C. E., Interian, A., & Moustafa, A. A. (2022).A practical introduction to using the drift diffusion model of decision-making in cognitive psychology, neuroscience, and health sciences. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2022.1039172

References III

- Tomkins-Flanagan, E., Hanley, C., & Kelly, M. A. (2025). Hey pentti, we did it again!: Differentiable vector-symbolic types that prove polynomial termination. *Proceedings of MathPsych / ICCM 2025*. https://mathpsych.org/presentation/1997
- Tomkins-Flanagan, E., & Kelly, M. A. (2024). Hey Pentti, We Did It!: A Fully Vector-Symbolic Lisp (C. Sibert, Ed.). *Proceedings of the 22nd International Conference on Cognitive Modeling*, 65–72. https://github.com/eilene-ftf/holis
- Wu, Y., Sun, Z., Li, S., Welleck, S., & Yang, Y. (2025). Inference scaling laws: An empirical analysis of compute-optimal inference for problem-solving with language models. https://arxiv.org/abs/2408.00724

ANIMUS

https://carleton.ca/animus

Viewers Like You

Thank You