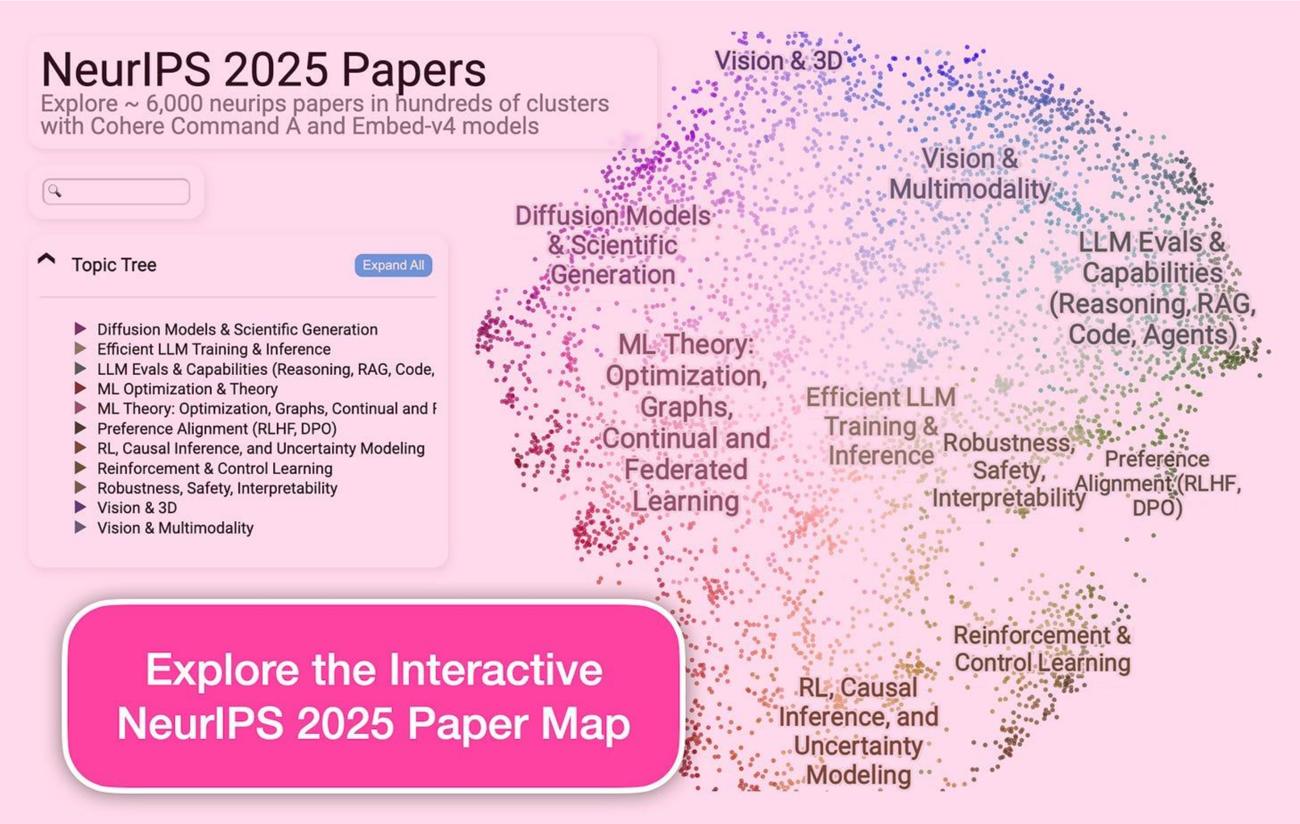
NeurIPS Observations

December 2025

NeurIPS Even Bigger

- Main track had 21,575 paper submissions, with 5,290 accepted



NeurIPS Trends

- Most popular topics
  - Multimodality
  - Lots of RL
  - Reasoning LLMs (and evaluation)
- Diffusion still heavily represented
- Also, applications and intersections of ML with other sciences
- Some faster inference papers because it’s not just about training
- Exhibit hall was large, with some, but not all big players
- Big companies outpacing research universities, maybe startups too
- Interpretability research showing some cracks

Paper Awards [1]

Outstanding Paper Awards:

- Gated Attention for Large Language Models: Non-linearity, Sparsity, and Attention-Sink-Free
- 1000 Layer Networks for Self-Supervised RL: Scaling Depth Can Enable New Goal-Reaching Capabilities
- Why Diffusion Models Don’t Memorize: The Role of Implicit Dynamical Regularization in Training
- Artificial Hivemind: The Open-Ended Homogeneity of Language Models (and Beyond)

Paper Awards [2]

Runner-Up Awards:

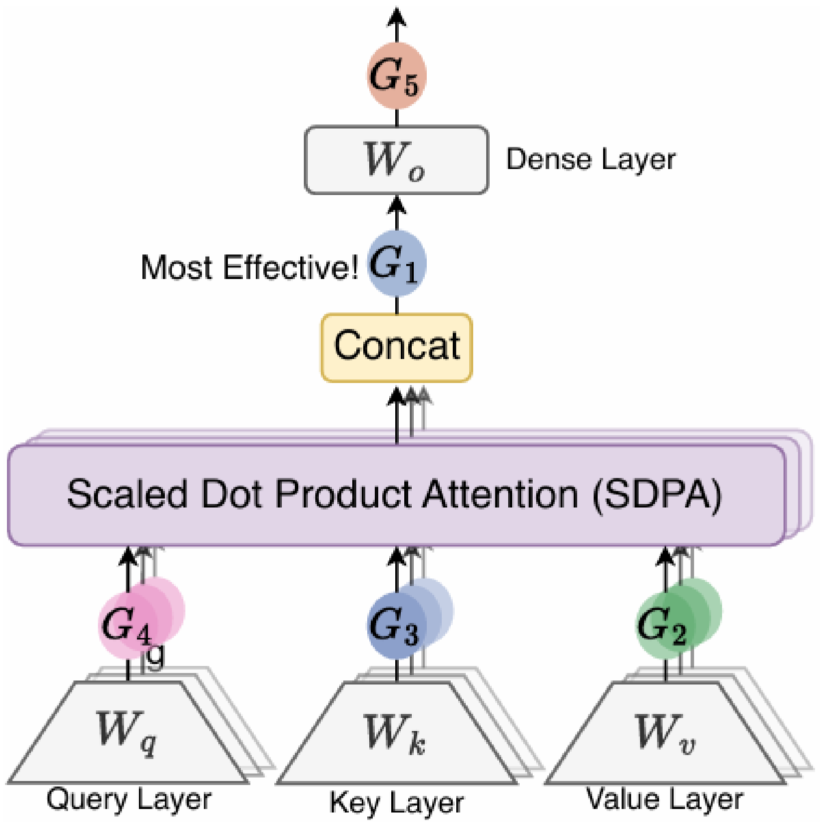
- Does Reinforcement Learning Really Incentivize Reasoning Capacity in LLMs Beyond the Base Model?
- Superposition Yields Robust Neural Scaling
- Optimal Mistake Bounds for Transductive Online Learning

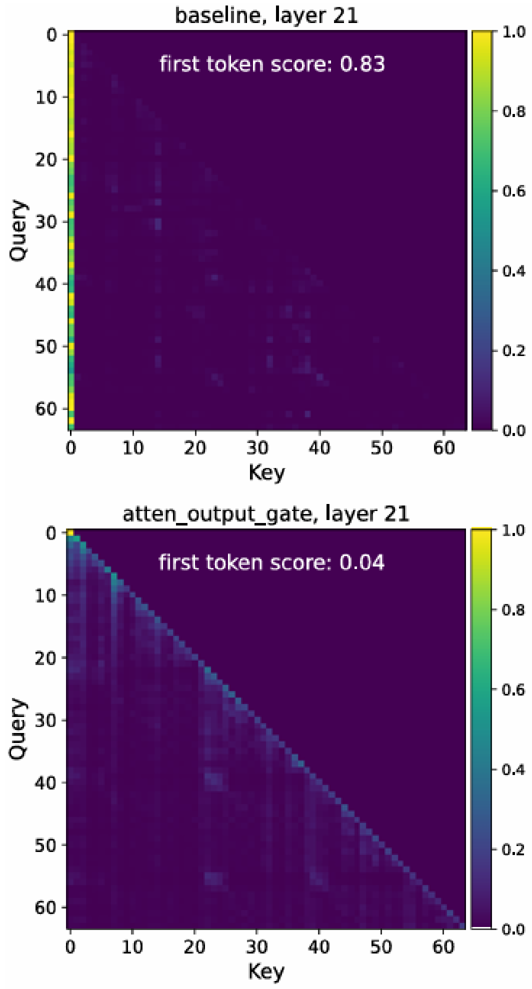
Gated Attention for Large Language Models: Non-linearity, Sparsity, and Attention-Sink-Free [1]

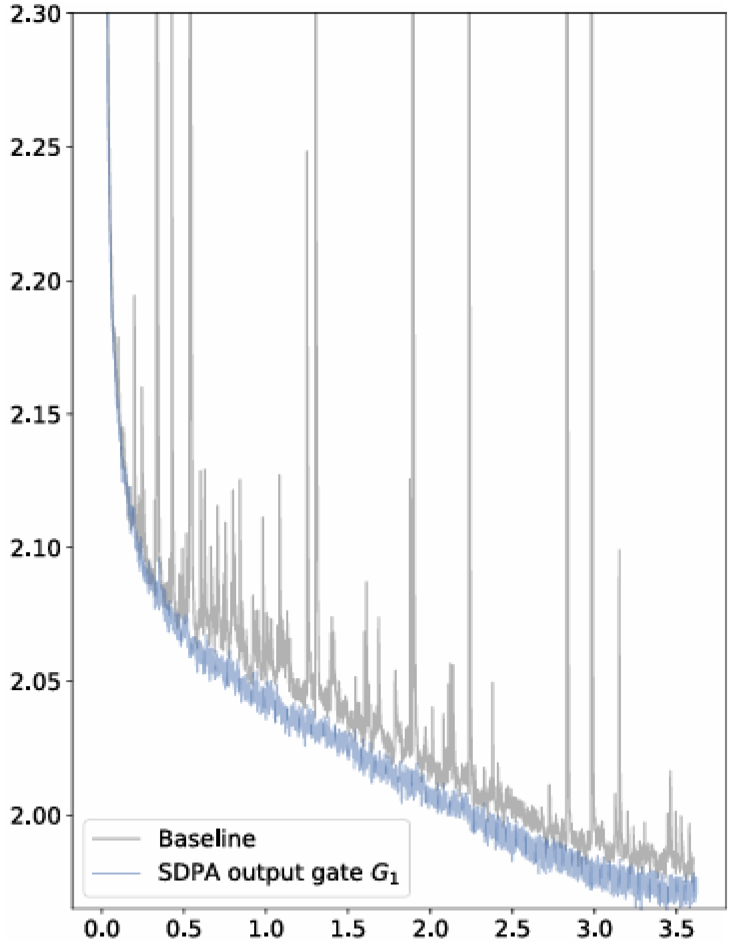
https://arxiv.org/abs/2505.06708

- Qwen team added a data-dependent sigmoid gate after attention
- This addition adds a nonlinearity, increasing expressiveness, and also increases sparsity, because many attention outputs become zero
- Additionally, this dramatically reduced the presence of attention sinks and massive activations, improving training stability

Gated Attention for Large Language Models: Non-linearity, Sparsity, and Attention-Sink-Free [2]







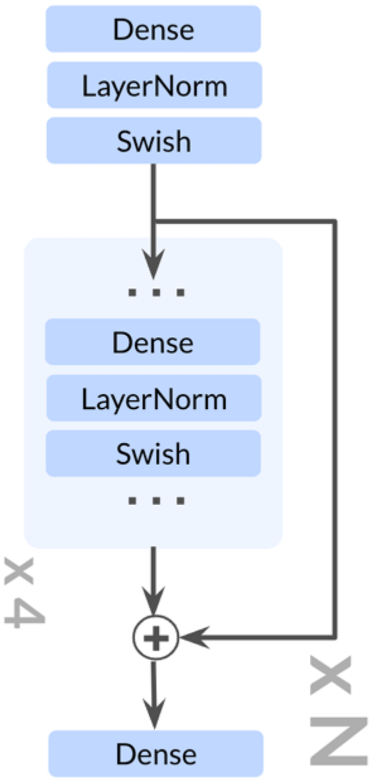
1000 Layer Networks for Self-Supervised RL: Scaling Depth Can Enable New Goal-Reaching Capabilities [1]

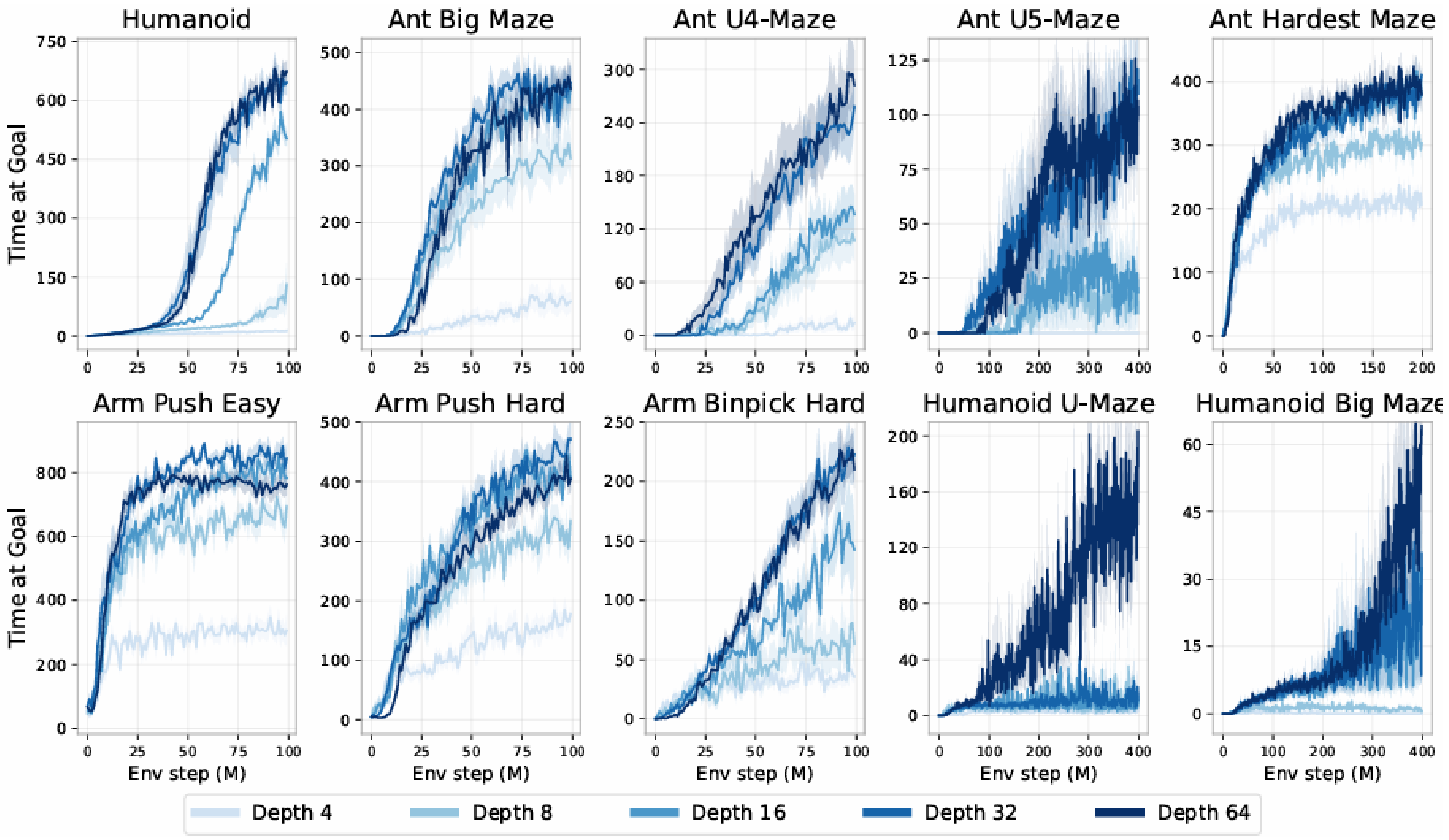
https://arxiv.org/abs/2503.14858

- Princeton-led worked to get self-supervised learning to work for RL, using contrastive RL
- Additionally, they looked to break past typically depth limits ~5 layers
- Their contrastive RL is an actor-critic method
  - The critic is set up as two NNs, one for state action and the other for goals
  - The critic outputs the L2 norm of the difference between the embedding outputs of the two NNs
  - They trained the critic with InfoNCE objective
  - The actor is trained to optimize its policy to maximize the critic's value

1000 Layer Networks for Self-Supervised RL: Scaling Depth Can Enable New Goal-Reaching Capabilities [2]

- Used residual nets with MLPs, and tried scaling up to 1024 layers





1000 Layer Networks for Self-Supervised RL: Scaling Depth Can Enable New Goal-Reaching Capabilities [2]

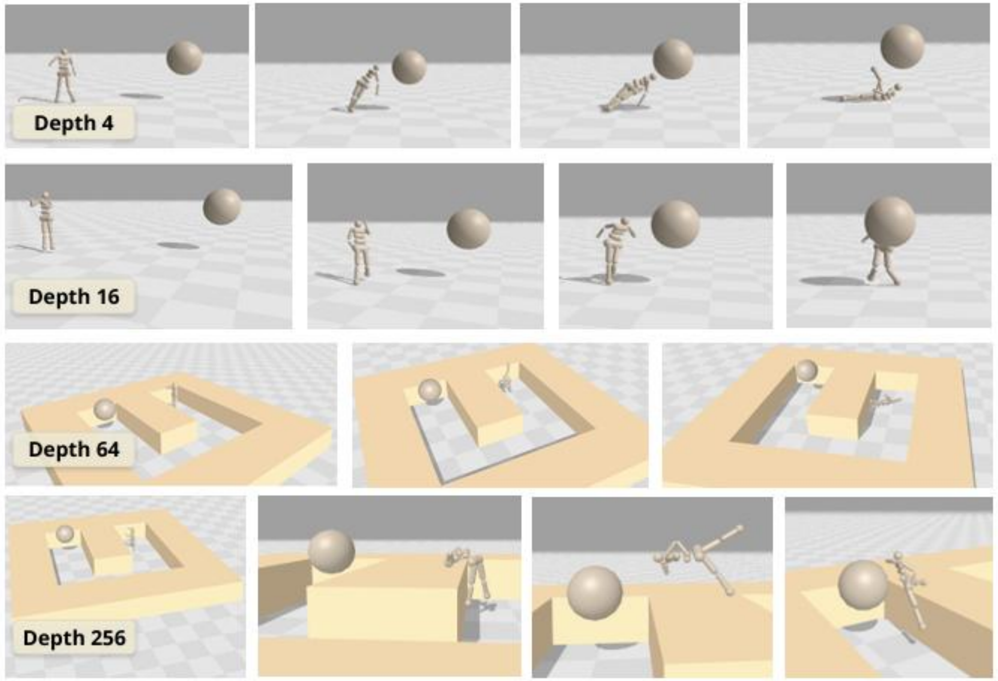


Figure 3: **Increasing depth results in new capabilities: Row 1**: A depth-4 agent collapses and throws itself toward the goal. **Row 2**: A depth-16 agent walks upright. **Row 3**: A depth-64 agent struggles and falls. **Row 4**: A depth-256 agent vaults the wall acrobatically.

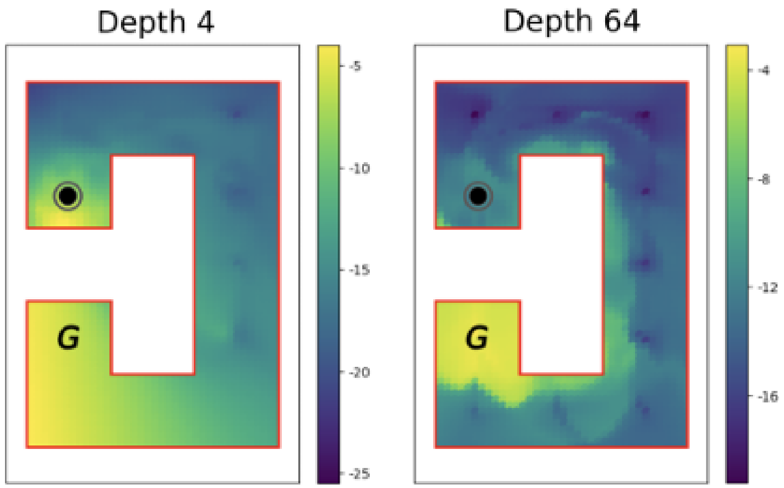


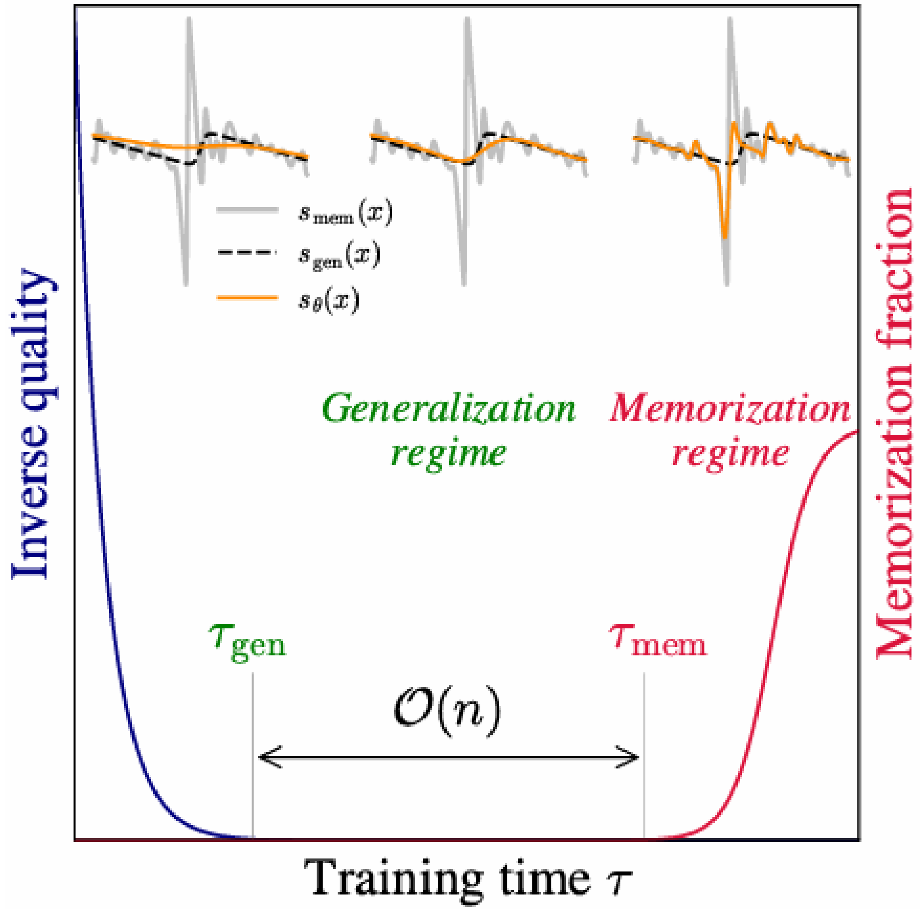
Figure 9: **Deeper Q-functions are qualitatively different.** In the U4-Maze, the start and goal positions are indicated by the ⊙ and **G** symbols respectively, and the visualized Q values are computed via the L_2 distance in the learned representation space, i.e., $Q(s, a, g) = \|\phi(s, a) - \psi(g)\|_2$. The shallow depth 4 network *(left)* naively relies on Euclidean proximity, showing high Q values near the start despite a maze wall. In contrast, the depth 64 network *(right)* clusters high Q values at the goal, gradually tapering along the interior.

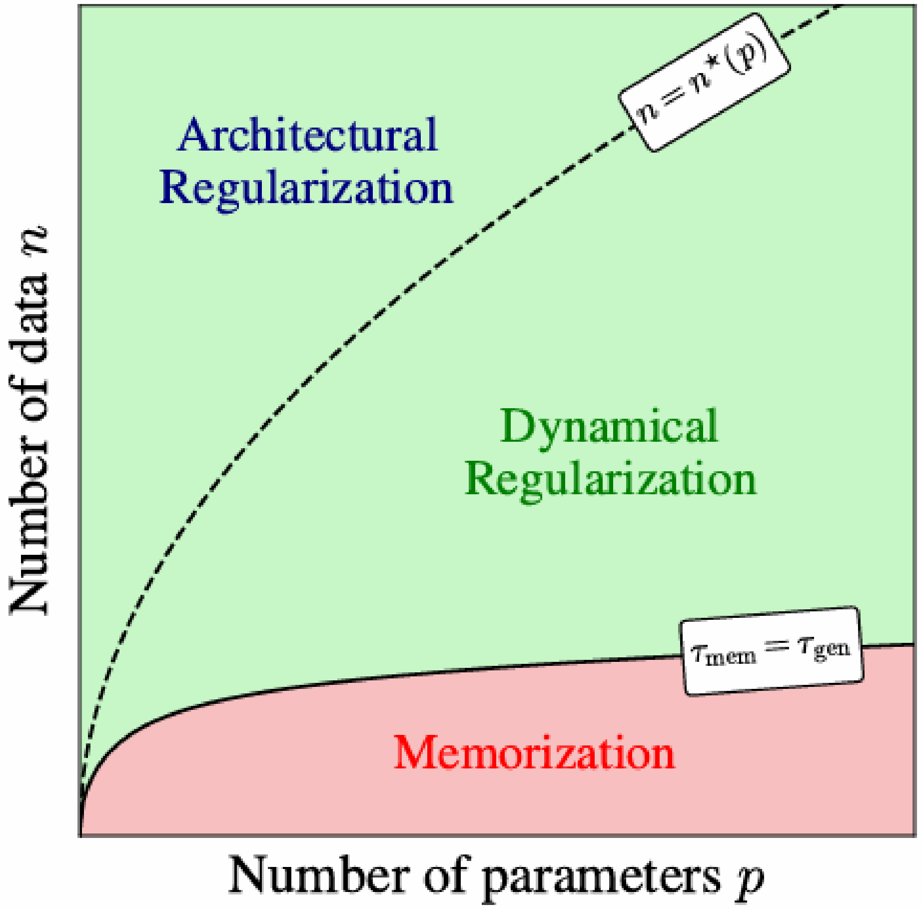
Why Diffusion Models Don’t Memorize: The Role of Implicit Dynamical Regularization in Training [1]

https://arxiv.org/abs/2505.17638

- Université PSL, Paris-led study why diffusion models generalize well
- Demonstrated that diffusion models start to generalize early at a data-independent time, and will memorize if trained long enough
  - Importantly, the memorization time grows with the training set size, creating a wide window of early stopping to get good generalization
- Pattern is shown empirically with both real data and controlled synthetic data, then proven theoretically

Why Diffusion Models Don't Memorize: The Role of Implicit Dynamical Regularization in Training [2]





Artificial Hivemind: The Open-Ended Homogeneity of Language Models (and Beyond) [1]

https://arxiv.org/abs/2510.22954

- University of Washington-led benchmark, called Infinity-Chat, for open-ended generation
  - A couple years ago, NeurIPS started rewarding benchmark creators
- Develops a taxonomy of open-ended prompts
- Studied over 70 large models and found current methods are unintentionally reducing diversity, the “artificial hivemind”

Artificial Hivemind: The Open-Ended Homogeneity of Language Models (and Beyond) [2]

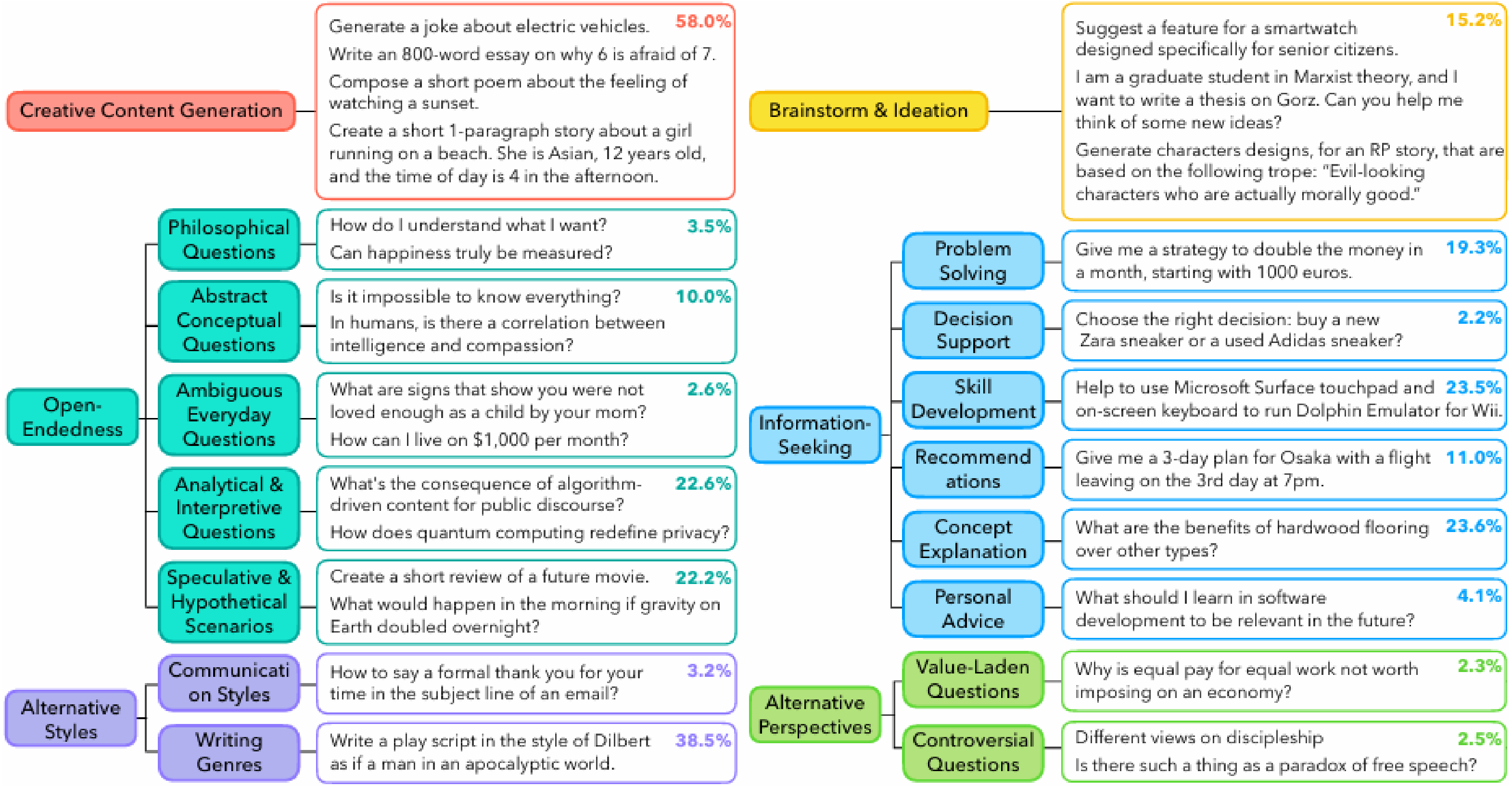


Figure 2: A taxonomy of **real-world open-ended queries that invite diverse model responses** that are mined from in-the-wild user-chatbot interactions, categorized into 6 top-level and 17 fine-grained subcategories, along with their occurrence percentages.

Artificial Hivemind: The Open-Ended Homogeneity of Language Models (and Beyond) [3]

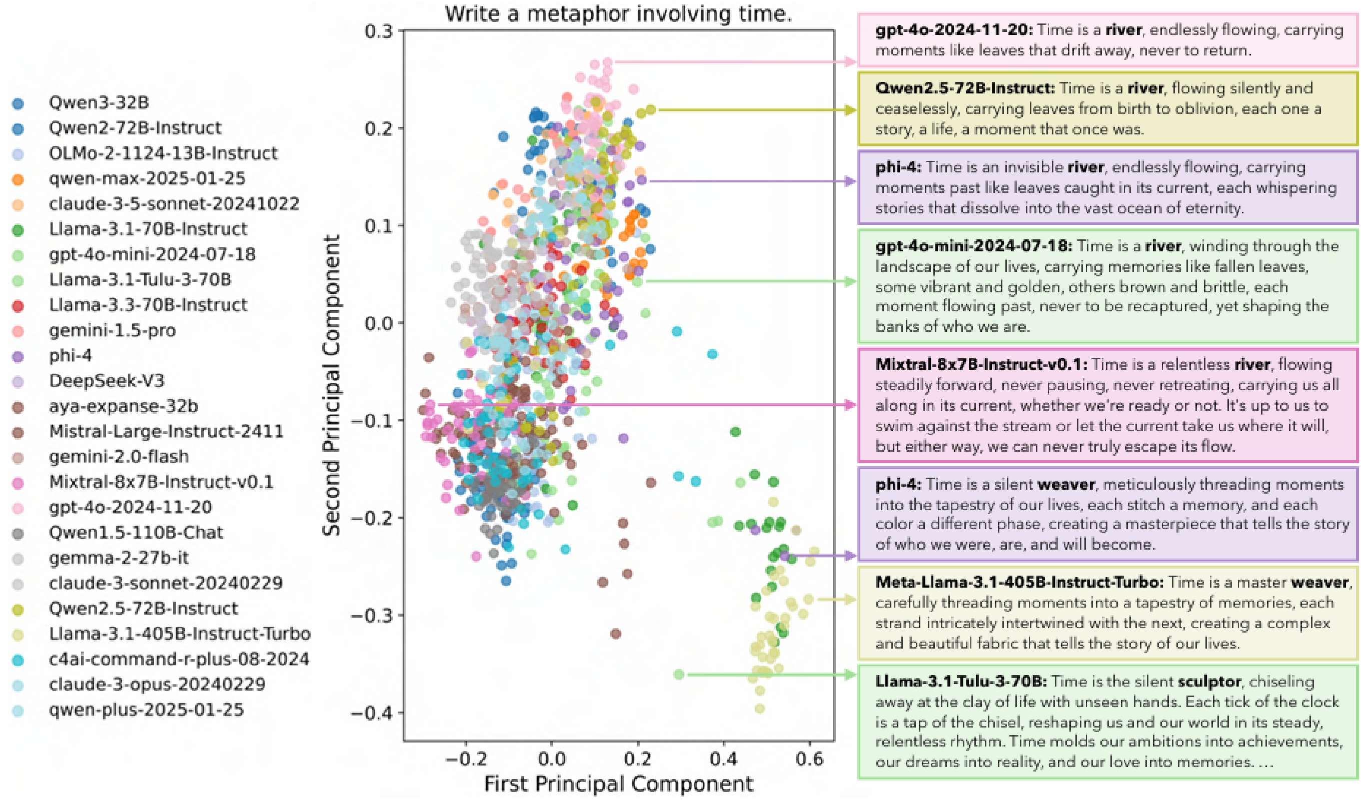


Figure 1: Responses to the query **"Write a metaphor about time"** clustered by applying PCA to reduce sentence embeddings to two dimensions. Each of the 25 models generates 50 responses using top-p sampling ($p = 0.9$) and temperature $= 1.0$. Despite the diversity of model families and sizes, the responses form just two primary clusters: a dominant cluster on the left centered on the metaphor "time is a river," and a smaller cluster on the right revolving around variations of "time is a weaver."

Does Reinforcement Learning Really Incentivize Reasoning Capacity in LLMs Beyond the Base Model? [1]

https://arxiv.org/abs/2504.13837

- Tsinghua University-led analysis of the potential of reinforcement learning with verifiable rewards (RLVR) to improve LLM reasoning
- Demonstrated that RLVR improves sampling efficiency at a cost of narrowing exploration, ultimately optimizing within the base model's capability distribution
- The key tool is using *pass@k* metric for large values of *k*

Does Reinforcement Learning Really Incentivize Reasoning Capacity in LLMs Beyond the Base Model? [2]

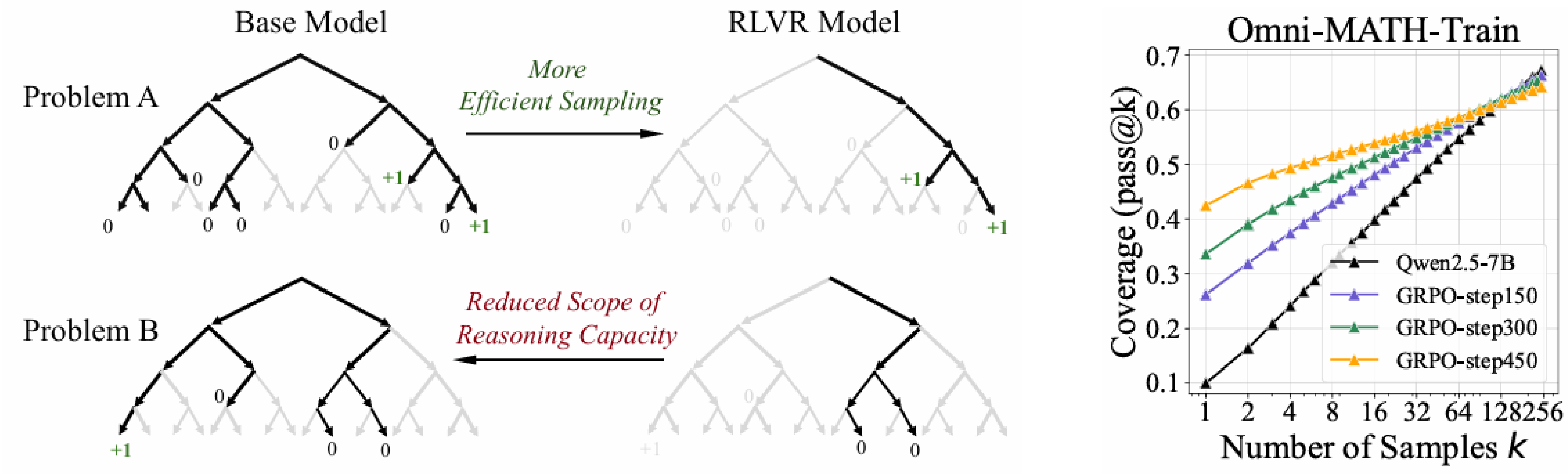


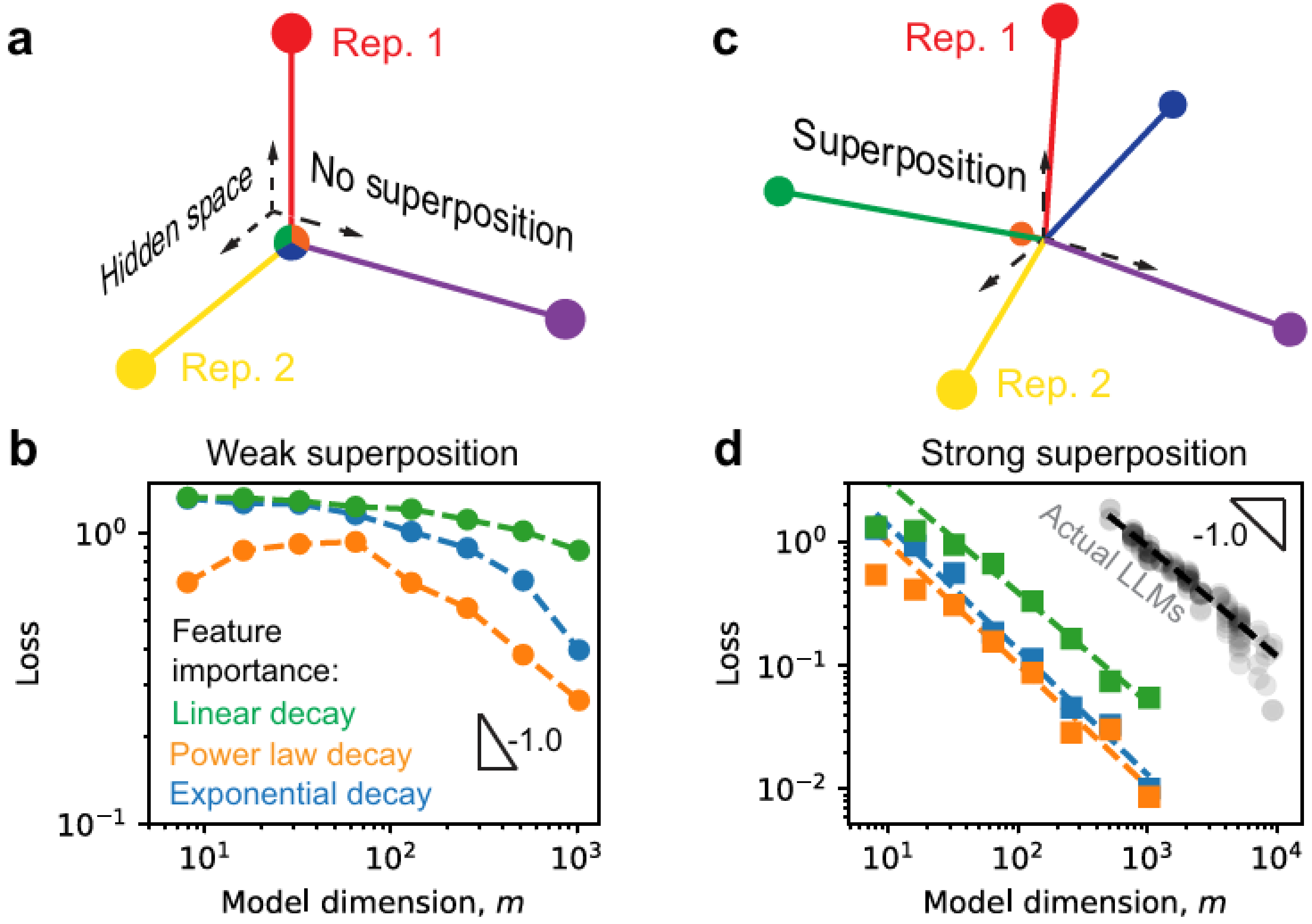
Figure 1: **(Left)** The effect of current RLVR on LLM's reasoning ability. Search trees are generated by repeated sampling from the base and RLVR-trained models for a given problem. Grey indicates paths that are unlikely to be sampled by the model, while **black** indicates paths that are likely to be sampled. Green indicates correct paths, which has positive rewards. Our key finding is that all reasoning paths in the RLVR model are already present in the base model. For certain problems like Problem A, RLVR training biases the distribution toward rewarded paths, improving sampling efficiency. However, this comes at the cost of reduced scope of reasoning capacity: For other problems like Problem B, the base model contains the correct path, whereas that of the RLVR model does not. **(Right)** As RLVR training progresses, the average performance (*i.e.*, pass@1) improves, but the coverage of solvable problems (*i.e.*, pass@256) decreases, indicating a reduction in LLM's reasoning boundary.

Superposition Yields Robust Neural Scaling [1]

https://arxiv.org/abs/2505.10465

- MIT paper shows why LLMs have Kaplan/Chinchilla power law scaling
- Using a toy model, they show that without concept superposition, model scaling pattern depends on the feature frequencies of the data
- Under strong superposition, loss scales as an inverse power law with respect to the model dimension regardless of the distribution of feature frequencies of the data
- Weight decay was the tool for controlling degree of superposition

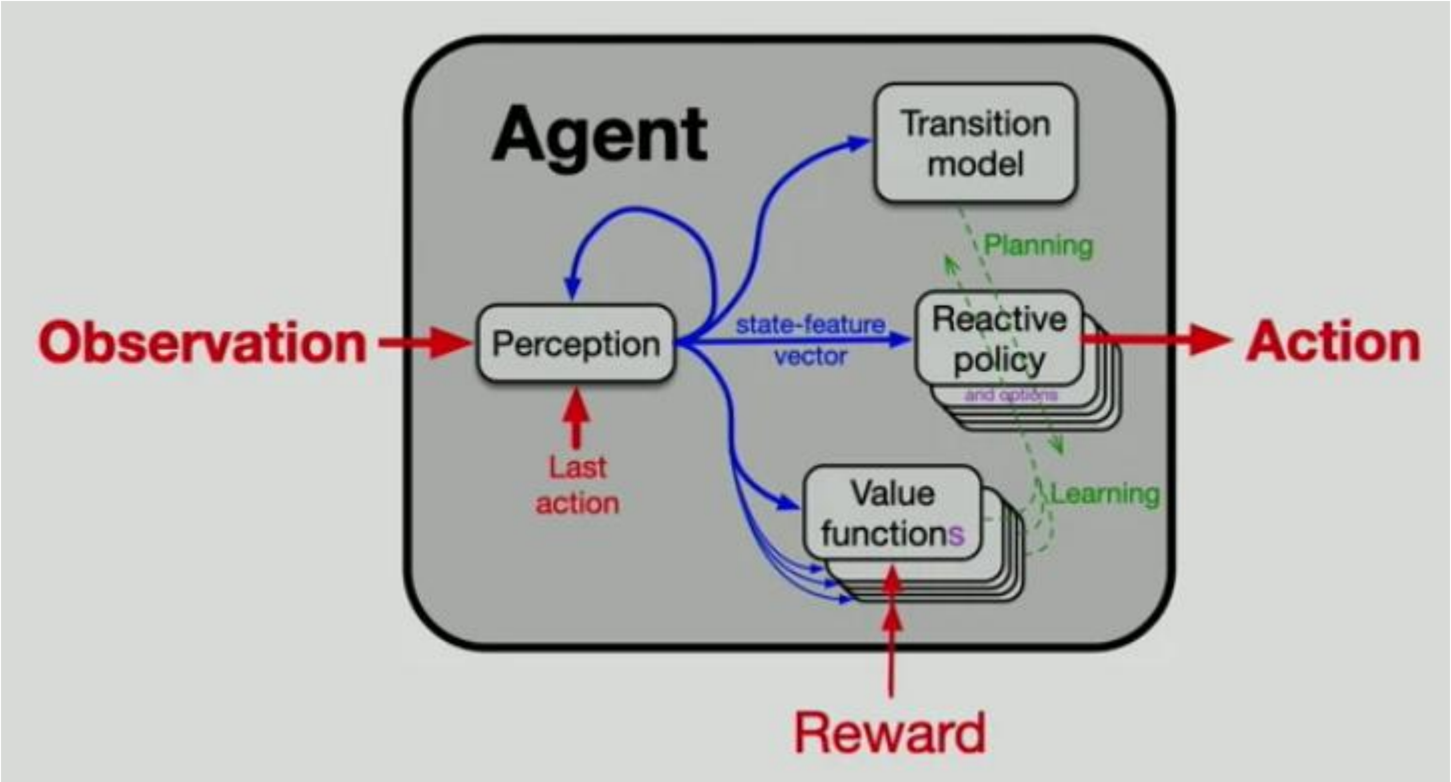
Superposition Yields Robust Neural Scaling [2]



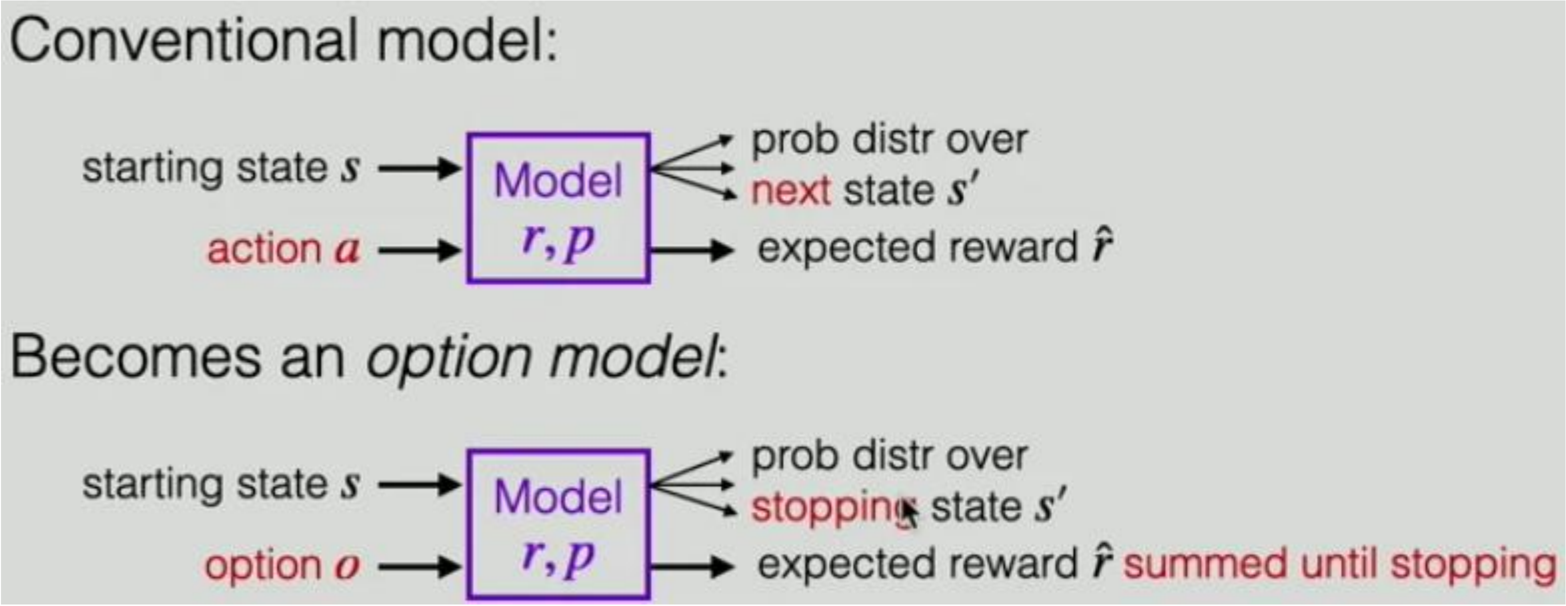
The Oak Architecture: A Vision of SuperIntelligence from Experience [1]

- Invited talk by Rich Sutton
- Extends the Bitter Lesson concept to reinforcement learning
- In a Big World, everything must be approximate and world appears non-stationary, so must be able to convert experience into knowledge
- The OaK architecture introduces the *option*, a series of actions
- Agent has many options and learns a high-level transition model enabling planning with longer jumps
- Key open problem: the agent must create its own subproblems
- Reliable continual learning seems necessary

The Oak Architecture: A Vision of SuperIntelligence from Experience [2]



The Oak Architecture: A Vision of SuperIntelligence from Experience [3]



Are We Having the Wrong Nightmares About AI? [1]

- Invited talk by Zeynep Tufekci
- Historically, people think about 1:1 replacement by new technology, and they don’t anticipate the impact of massive increase in scale
    - For example, people asked if cars were faster than horses. Didn’t consider how scaling cars would lead to living in suburbs, or pollution, or global warming
- Scarcity to abundance breaks old defenses that use difficulty signaling

Are We Having the Wrong Nightmares About AI? [2]

The Many Mechanisms Generative AI Will Break

- Proof of Effort – essays, custom job cover letters
- Proof of Authenticity – video as evidence, who is one the phone?, financial system, court system
- Proof of Accuracy – our credibility filters
- Proof of Sincerity – can you believe any feedback?
- Proof of Humanity – art as a window into humanity, the key beneficiary of copyright is the public (not the creator)

Are We Having the Wrong Nightmares About AI? [3]

- Believes the transition with AI will be extremely destabilizing
- Centralization and super surveillance will work, and they will be both demanded and offered
- Recommends restabilizing with more innovation of the right type so we can benefit from generative AI “in a manner compatible with freedom, dignity, liberty, and humanity”
  - Think of something revolutionary like Diffie-Hellman, but for the Generative AI age

SuffixDecoding: Extreme Speculative Decoding for Emerging AI Applications [1]

https://arxiv.org/abs/2411.04975

- Snowflake-led research extending speculative decoding to agentic workloads where long sequences may reoccur frequently
- Builds fast suffix trees to cache long token sequences seen
- Adaptively speculates more tokens when confident that acceptance likelihood is high and fewer when it is low
- Efficiently verifies entire speculation subtrees in one forward pass
- Achieved speedups up to 5.3x on benchmarks like SWE-Bench

SuffixDecoding: Extreme Speculative Decoding for Emerging AI Applications [2]

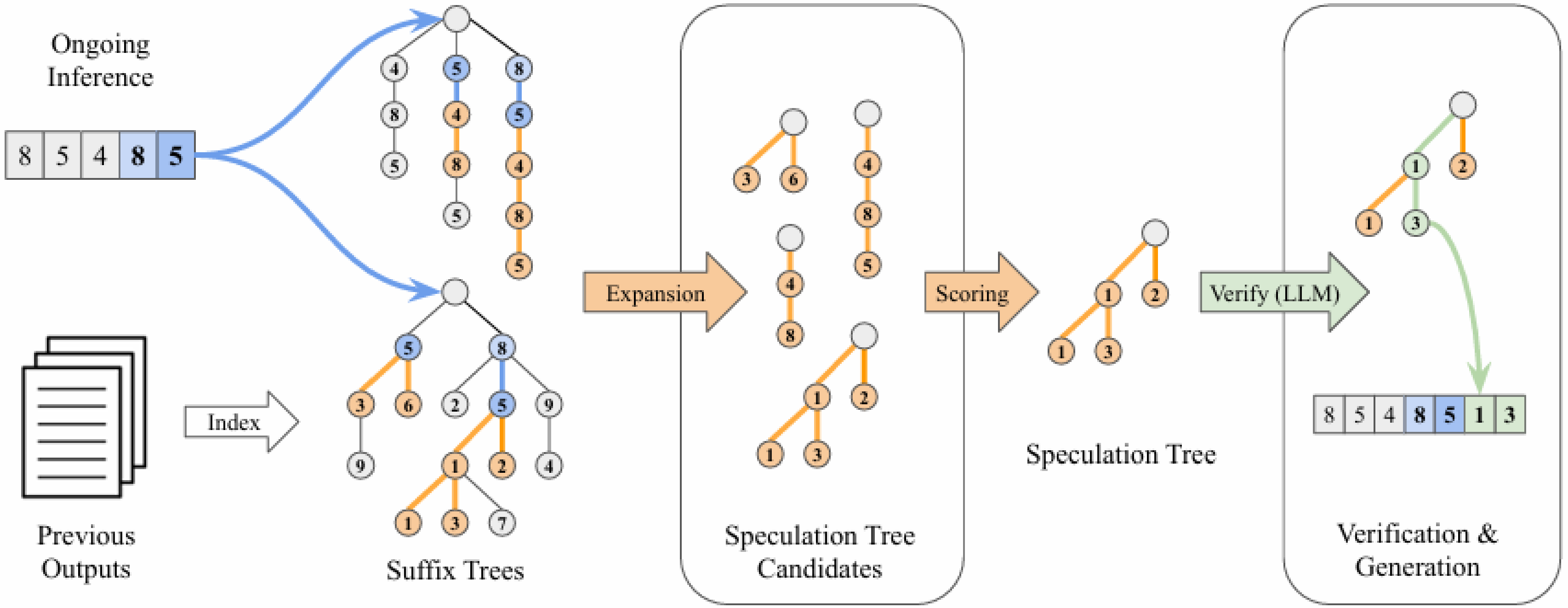


Figure 1: Overview of SuffixDecoding's algorithm. Two suffix trees track ongoing inference (top-left) and previous outputs (bottom-left). SuffixDecoding uses these trees to find matching patterns based on recently generated tokens. It constructs a speculation tree (middle) by selecting the most likely continuations, scoring them based on frequency statistics. Finally, the best candidate is verified by the LLM in a single forward pass (right), with accepted tokens (shown in green) being added to the output and used for the next round of speculation.

Other Interesting Papers [1]

- How to Scale Second-Order Optimization
- Does Thinking More always Help? Mirage of Test-Time Scaling in Reasoning Models
- Large Language Diffusion Models
- REOrdering Patches Improves Vision Models
- Differentiable Hierarchical Visual Tokenization
- Vision Transformers Don't Need Trained Registers
- Perception Encoder: The best visual embeddings are not at the output of the network
- ImageNet-trained CNNs are not biased towards texture: Revisiting feature reliance through controlled suppression

Other Interesting Papers [2]

- A is for Absorption: Studying Feature Splitting and Absorption in Sparse Autoencoders
- Finding Manifolds With Bilinear Autoencoders
- The Dual-Route Model of Induction
- Learning Linear Attention in Polynomial Time

- And model merging, such as: Train with Perturbation, Infer after Merging: A Two-Stage Framework for Continual Learning