Jayden Teoh
https://jaydenteoh.github.io/

How do we build models that learn abstractions necessary for *reasoning in the real world*? And how can they learn *data-efficiently*, *self-improve*, and generalize beyond human supervision? These questions keep me up at night. In the last three years, I was fortunate enough to explore a wide breadth of research topics encompassing **language modeling and reinforcement learning (RL)**, leading to several publications in top research conferences including NeurIPS, ICML and ICLR [1, 2, 3, 4, 5].

I am currently a research intern at Microsoft Research with Dr. **John Langford** and Dr. **Alex Lamb**. I am also a research assistant at Singapore Management University (SMU) with Prof. **Pradeep Varakantham**. Throughout my undergraduate studies at SMU, I have been ranked the **top Computer Science student in my cohort every year**. I also serve as Programs Lead of "SMU AI", the university's AI club, where I help expand access to AI education by organizing and teaching workshops.

Long-term Vision. Success in my PhD would mean developing generative architectures that (a) learn structured latent representations data-efficiently, (b) leverage those representations to reason and plan, and (c) continuously self-improve with minimal human supervision. I find joy and meaning in expanding access to education, so my goal is to impart knowledge to future students as a professor. Below, I outline the directions I plan to advance in graduate study[1].

1 Parallel vs. Sequential Computation for Complex Reasoning

How can we instill an inductive bias for recurrent computation while retaining the parallelizability that made transformers so powerful? A growing body of work has hinted at the **necessity of sequential computation** for models capable of complex reasoning (e.g. latent CoT, HRMs, Chomsky Hierarchy, circuit complexity literature etc.). Although CoT prompting emulates recurrence in token space, manually curating high-quality CoT data to elicit reasoning is non-scalable. I am interested in developing learning methods and neural architectures that enable implicit discovery of such recurrent reasoning structures.

At Microsoft Research, I lead a project [6] that tackles this challenge. Transformers replaced recurrence with self-attention, removing the incentive to maintain compact latent dynamics and often leads to shortcut, non-generalizable solutions. To address this, I proposed a self-supervised **next-latent prediction** objective that instills recurrent inductive biases in transformers while preserving parallel training. Empirically, our auxiliary objective improves transformer performance in various domains, including world modeling, reasoning, planning, and language.

Much remains to be explored in this direction. During my PhD, I aim to scale recurrent reasoning models by leveraging surrogate gradients, hardware-aware optimization, and neural architectural innovations to parallelize and amortize sequential computation.

2 Data-Efficiency and Representation Learning

Large language models, despite training on trillions of tokens, still fail at simple tasks such as arithmetic. Humans, in contrast, are more data- and energy-efficient. My research addresses this by developing methods that extract richer gradient signals from datasets, allowing efficient learning from less data. In Hu et al. [7], I optimized the architecture and performance of **Belief State Transformers (BST)**, a method that harvests $O(N^2)$ gradients per token sequence to improve goal-conditioned planning in language.

However, BST and most language models, still rely on token-level predictions. The **myopic nature of token predictions** contributes to data-inefficiency as it biases learning towards exploiting n-gram regularities rather than broader abstractions, thereby hindering generalization. My subsequent work [6] addresses this limitation by extracting richer gradients through self-supervised predictions in the latent space.

I believe future progress hinges on self-supervised representation learning approaches. However, current self-supervised approaches rely on heuristic discrete tokenization or patch-level boundaries that inhibit the end-to-end learning of compact, generalizable abstractions across modalities. During my PhD, I plan to develop better representation learning methods.

[1] References to external works are provided as URLs for brevity.

3 Open-Ended Self-Improving Agents

At the current rate, we will run out of web-text data for training models by 2028. It is imperative to develop algorithms capable of open-ended learning. Yet, current training regimes still heavily rely on human annotation and dataset curation. How can models continually invent new challenges and generalize beyond human supervision?

My research approaches this via RL and unsupervised environment design (UED). In Teoh et al. [2], I developed a novel **framework for quantifying environment novelty** in UED, enabling RL agents to self-direct curricula toward not only "challenging", but also "novel" environments, encouraging exploration essential for generalization. Our method achieved state-of-the-art zero-shot generalization performance across multiple UED benchmarks and was selected for an **Oral presentation at NeurIPS 2024** [2] (top 0.4%). In Chirra et al. [5], we proposed a LLM-guided evolutionary framework for stabilizing training in Adversarial Imitation Learning. Our meta-learned algorithm generalizes across different RL domains and significantly outperforms human-designed baselines.

Recently, RL has shown promise in post-training language models. However, approaches predominantly rely on RL with verifiable rewards, which necessitates *apriori* curation of verifiable tasks. I foresee a promising future in marrying open-ended learning with post-training.

Why Massachusetts Institute of Technology?

My research interests sits at the intersection of efficient representation learning for generative models and generally-capable agents that can reason and plan. MIT's collaborative culture makes it an exceptional fit for me as several faculties' research align closely.

I am inspired by Prof. **Philip Isola's** research and his Platonic Representation Hypothesis. I believe the field still lacks efficient algorithms with the right inductive biases to learn unified abstractions across modalities in unsupervised settings, and I hope to develop such methods together.

Prof. **Leslie Kaelbling's** research on POMDPs and belief states has motivated my own [5, 6]. I really like her RLC talk on rational learning, especially her recent focus on harnessing the priors of VLMs to propose predicates and improve planning.

I am inspired by Prof. **Yoon Kim's** and Prof. **Jacob Andreas'** research on understanding the representational limitations of machine learning models, especially in language. I am excited to collaborate on innovating neural architectures and data-efficient learning algorithms, with a focus on representation learning that ensures these models capture faithful world models.

I look forward to contributing rigor, curiosity, and collaborative energy to MIT, while learning from a community that pushes the frontier of machine learning.

References

[1] **Jayden Teoh**, Wenjun Li, and Pradeep Varakantham. Unifying regret and state-action space coverage for effective unsupervised environment design. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '24. International Foundation for Autonomous Agents and Multiagent Systems, 2024.

[2] **Jayden Teoh**, Wenjun Li, and Pradeep Varakantham. Improving environment novelty quantification for effective unsupervised environment design. *Advances in Neural Information Processing Systems*, 2024.

[3] **Jayden Teoh**, Pradeep Varakantham, and Peter Vamplew. On generalization across environments in multi-objective reinforcement learning. In *The Thirteenth International Conference on Learning Representations*, 2025.

[4] Felix Hofstätter*, Teun van der Weij*, **Jayden Teoh***, Rada Djoneva, Henning Bartsch, and Francis Rhys Ward. The elicitation game: Evaluating capability elicitation techniques. In *Forty-second International Conference on Machine Learning*, 2025.

[5] Shashank Reddy Chirra, **Jayden Teoh**, Praveen Paruchuri, and Pradeep Varakantham. On discovering algorithms for adversarial imitation learning. In *The Fourteenth International Conference on Learning Representations*, 2026.

[6] **Jayden Teoh**, Manan Tomar, Kwangjun Ahn, Edward S. Hu, Pratyusha Sharma, Riashat Islam, Alex Lamb, and John Langford. Next-latent prediction transformers learn compact world models, 2025. URL https://arxiv.org/abs/2511.05963. Microsoft Research Preprint, Work in Progress.

[7] Edward S. Hu, Kwangjun Ahn, Qinghua Liu, Haoran Xu, Manan Tomar, Ada Langford, **Jayden Teoh**, Bryon Xu, David Yan, Dinesh Jayaraman, Alex Lamb, and John Langford. The belief state transformer, 2025. URL https://arxiv.org/abs/2410.23506. *Contributed post-publication at ICLR 2025.*