Keiran Paster

#### Interests

I am a PhD Student at the University of Toronto supervised by **Sheila McIlraith** and **Jimmy Ba**. My research interests lie in AI reasoning and decision-making and my goal is to create AI systems that effectively process the world's information in order to fast-forward scientific progress. My work includes algorithmic work on and the scaling up of decision-making via success-conditioned sequence models, automatic prompt-engineering (language models that learn to control themselves), and the creation of a family of strong open-weights language models for reasoning which serve as a strong foundation for the next wave of reasoning research.

#### Education

University of Toronto PhD Student in Computer Science - GPA 4.0 - Projected Graduation: 2024

University of California, Berkeley B.S. In Electrical Engineering and Computer Science - GPA 3.907 Sep. 2019 – Present Toronto, ON

Sep. 2015 – May 2019 Berkeley, CA

#### Publications

### Llemma: An Open Language Model For Mathematics

Zhangir Azerbayev, Hailey Schoelkopf, **Keiran Paster**, Marco Dos Santos, Stephen McAleer, Albert Q. Jiang, Jia Deng, Stella Biderman, Sean Welleck Poster at ICLR 2024. (https://arxiv.org/abs/2310.10631)

#### **OpenWebMath:** An Open Dataset of High-Quality Mathematical Web Text

Keiran Paster, Marco Dos Santos, Zhangir Azerbayev, and Jimmy Ba Poster at ICLR 2024. (https://arxiv.org/abs/2310.06786)

### STEVE-1: A Generative Model for Text-to-Behavior in Minecraft

Shalev Lifshitz<sup>\*</sup>, **Keiran Paster**<sup>\*</sup>, Harris Chan, Jimmy Ba, Sheila McIlraith NeurIPS 2023 Spotlight. (https://openreview.net/pdf?id=YkBDJWerKg)

### Large Language Models are Human-Level Prompt Engineers

Yongchao Zhou<sup>\*</sup>, Andrei Ioan Muresanu<sup>\*</sup>, Ziwen Han<sup>\*</sup>, **Keiran Paster**, Silviu Pitis, Harris Chan, Jimmy Ba Poster at ICLR 2023. (https://arxiv.org/abs/2211.01910) Presented at the Foundation Models for Decision Making Workshop at NeurIPS 2022 (**Oral**). Presented at the Workshop on Machine Learning Safety at NeurIPS 2022.

### Return Augmentation gives Supervised RL Temporal Compositionality

Keiran Paster\*, Silviu Pitis\*, Sheila McIlraith, Jimmy Ba
Preprint. (https://openreview.net/pdf?id=q5olkWCt7nl)
Presented at the Foundation Models for Decision Making Workshop at NeurIPS 2022.
Presented at the NeurIPS 2022 Deep RL Workshop.

### You Can't Count on Luck: Why Decision Transformers Fail in Stochastic Environments

Keiran Paster, Sheila McIlraith, Jimmy BaPresented at the Decision Awareness in Reinforcement Learning Workshop at ICML 2022 (Poster).Published at NeurIPS 2022. (https://arxiv.org/abs/2205.15967)

### **BLAST: Latent Dynamics Models from Bootstrapping**

Keiran Paster<sup>\*</sup>, Lev McKinney<sup>\*</sup>, Sheila McIlraith, Jimmy Ba Presented at the NeurIPS 2021 Deep RL Workshop. (https://openreview.net/forum?id=VwA\_hKnX\_kR) Presented at the 5th Multidisciplinary Conference on Reinforcement Learning and Decision Making (Poster).

Learning Domain Invariant Representations in Goal-conditioned Block MDPs

Beining Han, Chongyi Zheng, Harris Chan, **Keiran Paster**, Michael R. Zhang, Jimmy Ba Published at NeurIPS 2021. (https://openreview.net/forum?id=oepSB9bsoCF)

# Planning from Pixels using Inverse Dynamics Models

Keiran Paster, Sheila McIlraith, Jimmy Ba Presented at the NeurIPS 2020 Deep RL Workshop (Oral). Published at ICLR 2021. (https://openreview.net/forum?id=V6BjBgku7Ro)

# Equilibrium Finding via Asymmetric Self-Play Reinforcement Learning

Jie Tang\*, Keiran Paster\*, Pieter Abbeel

Presented at the NeurIPS 2018 Deep RL Workshop. (http://goo.gl/AC3j44)

# Work Experience

# **Google Research** | Student Researcher

• Worked with the Google Deepmind Blueshift and Gemini teams on data-efficiency for LLMs.

# **University of Toronto** | TA for CSC413/2516, Neural Networks and Deep Learning

- CSC413 is a course which gives an overview of both foundational ideas and recent advances in neural net algorithms.
- Taught a tutorial, wrote a homework assignment about self-attention and normalizing flow models, and graded student projects.

# **UC Berkeley** | UGSI for CS 189, Introduction to Machine Learning

• Managed a class of 744 students, personally teaching a discussion section with 50 students. CS 189 is a proof-based class where students learn concepts like regression, classification, density estimation, dimensionality reduction, and clustering.

# **Robot Learning Lab** | Undergraduate Researcher

- Worked on deep reinforcement learning for playing real-time strategy games.
- Explored techniques to enable agents to quickly adapt to new opponent strategies.
- Created an algorithm that allows standard on-policy reinforcement learning algorithms to discover optimal policies in simple Poker games by taking gradient steps against a specialized opponent.

### **Google** | Software Engineering Intern

- Integrated Gmail Ads data into a MapReduce pipeline that generates terabytes of clean training data by combining data from multiple sources.
- Helped with the development of several internal tools, including a Gmail Ads serving stack diagnostic tool, a user model viewer and editor, and a visual user event inspector.

## Awards

## **NSERC** Postgraduate Scholarship | NSERC

• The NSERC Postgraduate Scholarship provides financial support to high-calibre students engaged in a doctoral program in natural sciences or engineering.

### Vector Research Grant | Vector Institute

• The Vector Research Grant is given to graduate students to support their research in the development of knowledge and novel applications of machine learning and AI.

### Faculty of Arts and Science Top (FAST) Doctoral Fellowship | University of Toronto

• This merit-based award is given to only one academically-excellent student in the Doctoral program each year.

### High Honors, Dean's Honors List, Honors to Date | UC Berkeley

• High honors are awarded to the top 7 percent of the graduating class by grade point average in the college.

### Member of Eta Kappa Nu (HKN) | UC Berkeley

- Eta Kappa Nu is the national Electrical and Computer Engineering honor society.
- Membership is restricted to the top quarter of the junior class.

# 2017 - 2019

2023

2021

2019

# 2017

### 2019 - 2022

2022 - 2025

# 2019

2019

2017

# **Relevant Coursework**

Itelevalit Coursework			
<ul> <li>Convex Optimization</li> <li>Natural Language Processing</li> <li>Neural Net Training Dynamics</li> <li>Data Structures and Algorithms</li> <li>Combinatorial Algorithms</li> </ul>	<ul> <li>Computer Graphics</li> <li>Computer Security</li> <li>Private Data Analysis</li> <li>Game Theory</li> <li>Linear Algebra</li> </ul>	<ul> <li>Real Analysis</li> <li>Discrete Math</li> <li>Probability and Random Processes</li> <li>Information Theory</li> <li>Statistical Learning Theory</li> </ul>	
Leadership			
<ul> <li>Mentor for Toronto Graduate Applic</li> <li>The Graduate Application Assistance someone to provide feedback on their</li> <li>Mentored and reviewed the application</li> </ul>	ation Assistance Program   Univ e Program (GAAP) aims to ensure a e application before submission. ons of two applicants.	ersity of Toronto l applicants to higher degree progra	<b>2021</b> ams have
Research Mentoring   University of Ton • Mentoring an undergraduate student	onto doing research on model-based reinf	provide the provided and the provided an	2021
<ul><li>Virtual Reality at Berkeley   Director</li><li>Worked with a team to create an interaction of the second s</li></ul>	of Resources, Augmented Reality Intuitive command system for drones in	erfaces Researcher 2 Unity using the Microsoft HoloLen	015 – 2017 s.
<ul><li>Computer Science Club   Founder, Jan</li><li>Taught the AP Computer Science cu</li></ul>	a Teacher, President rriculum to students at my high scho	ol. 2	014 - 2015
<ul> <li>FRC Robotics Team 980   Controls Te</li> <li>Lead the team responsible for design</li> <li>Designed and created an inexpensive</li> <li>Won the LA Regional Creativity Away</li> </ul>	am Leader ing the controls systems for two comp driving system using PID control an ard (2013) and was an LA Regional I	<b>2</b> betition robots. d automatic gear shifting. Finalist (2014).	012 – 2014
Other Projects			
<ul><li>Linear Programming and Policy Opt</li><li>Explored the relationship between in</li><li>Showed that some interior point met.</li></ul>	<b>imization</b>   <i>Minimizing Expectations</i> terior point LP solvers and policy gra hods can converge much faster in sim	<i>Class Project</i> adient methods for RL. aple MDPs than existing RL algorit	<b>2021</b> hms.
<ul> <li>Bayesian Cake Cutting   Algorithms for</li> <li>Created a novel problem formulation</li> <li>Proved that there always exists an er</li> <li>Introduced a new algorithm to find er</li> </ul>	r Collective Decision Making Class F in which there is a prior over valuat avy-free cut even when there is uncer nvy-free and efficient cuts for distribution	<i>roject</i> ons of a good. tainty over valuations. utions of piecewise constant valuation	<b>2020</b>
<ul> <li>Differential Privacy for Stochastic M</li> <li>Designed and tested several differents</li> <li>Found that my new algorithm, which with low time-horizons.</li> </ul>	ulti-Armed Bandits   Private Data ally private algorithms for the stocks only views new data after logarithm	a Analysis Class Project astic multi-armed bandit problem. ically-many time-steps, is empirical	<b>2019</b> lly strong
Replanning in Deep Reinforcement I	earning   Learning to Search Class	Project	2019
$\cdot$ Channel that $\cdot$ Channel that $\cdot$ Channel the that $\cdot$ Channel the		$f_{} = f_{} = f_{$	· 1 C

- Showed that a deep reinforcement learning agent can be improved by adding a few seconds of additional training before choosing an action.
- Verified the hypothesis that the performance improvement in a state is correlated with familiarity by estimating the amount of times the state was seen in training and comparing the selected action before and after additional training with that of an oracle agent.