Abulhair Saparov

Assistant Professor of Computer Science

475 Stadium Mall Drive, West Lafayette, Indiana, 47907, United States

asaparov {at} purdue.edu | asaparov.org | asaparov | YouTube

Research interests

- Applications of machine learning to natural language processing (NLP), semantic parsing, natural language understanding (NLU)
- Reasoning in large language models (LLMs), evaluation and analysis of LLMs, mechanistic interpretability
- Symbolic and neuro-symbolic methods for reasoning, language understanding, worldbuilding
- Representations of meaning/knowledge, reasoning, especially in natural language understanding
- Statistical machine learning, interpretable machine learning, Bayesian nonparametrics, scalable inference
- Broadly interested in applications in e.g. epidemiology, linguistics, phylogenetics, biology, information security, etc.

Education

Carnegie Mellon University — *Pittsburgh, PA*

PhD in Machine Learning — *2017 - 2022*

- Advisor: Tom M. Mitchell
- Thesis: Towards General Natural Language Understanding with Probabilistic Worldbuilding
- Thesis committee: Tom M. Mitchell, William Cohen, Frank Pfenning, Vijay Saraswat

Carnegie Mellon University — *Pittsburgh, PA*

MS in Machine Learning — *2013 - 2017*

- Advisor: Tom M. Mitchell

Princeton University — *Princeton, NJ*

BSE in Computer Science — *2009 - 2013*

- Summa cum laude
- Certificate (minor) in Applied and Computational Mathematics
- Certificate (minor) in Neuroscience
- Thesis advisors: Ken A. Norman, David M. Blei

Publications

- Jimson Huang, Maxwell Nickell, Ruowang Zhang, Abulhair Saparov (2026). Curriculum Learning Teaches LLMs Long-Horizon Implicit Reasoning, Up To A Limit. CoRR, abs/2610.25612. [**Link**]
- Nischay Dhankhar, Alisher Amantay, Zhanibek Datbayev, Darkhan Akhmetov, Yash Shah, Nurmukhamed Ubaidullayev, Raman Saparkhan, Dos Baha, Abulhair Saparov (2026). Towards Foundation Decision Models: Drex. CoRR, abs/2611.07854. [**Link**]
- Amirhossein Basareh, Rishabh Sanjay, Daniel de Mello, Abulhair Saparov (2026). How Much Information Can We Fit inside a Transformer? CoRR, abs/2610.34553. [**Link**]
- Atharva Thakur, Abulhair Saparov, Bruno Ribeiro (2026). DeclineCraft: A RAG Abstention Benchmark for Jointly Measuring Combined Retrieval and Parametric Knowledge. CoRR, abs/2610.28340. [**Link**]
- Minh-Nam Tran, Abulhair Saparov (2026). GraphNIAH: A Benchmark for Long-Context Reasoning over Real-World Graphs. CoRR, abs/2610.41704. [**Link**]
- Nischay Dhankhar, Dos Baha, Abulhair Saparov (2026). Scaling Laws for Hypernetwork-Based Knowledge Injection in Large Language Models. CoRR, abs/2607.19604. [**Link**]
- Andreas Opedal, Francesco Ignazio Re, Abulhair Saparov, Mrinmaya Sachan, Bernhard Schölkopf, Ryan Cotterell (2026). Learning to Reason Efficiently with A* Post-Training. CoRR, abs/2605.24597. [**Link**]
- Tianle Wang, Zhaoyang Wang, Guangchen Lan, Xinpeng Wei, Sipeng Zhang, Guanwen Qiu, Abulhair Saparov (2026). How Deep Can LLMs Learn to Reason? Expressiveness Is Key. CoRR, abs/2605.06638. [**Link**]
- Yunxin Sun, Abulhair Saparov (2026). Do Language Models Follow Occam's Razor? An Evaluation of Parsimony in Inductive and Abductive Reasoning. Empirical Methods in Natural Language Processing, EMNLP 2026. [**Link**]

- Nathaniel Getachew, Abulhair Saparov (2026). Language Models Might Not Understand You: Evaluating Theory of Mind via Story Prompting. Third Conference on Language Modeling. [**Link**, **Oral** and **Spotlight** at Social Sim Workshop @ COLM 2025]
- Revanth Rameshkumar, Jimson Huang, Yunxin Sun, Fei Xia, Abulhair Saparov (2025). Reasoning Models Reason Well, Until They Don't. International Joint Conference on Natural Language Processing and the Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics, AACL/IJCNLP 2025. [**Link**, **Best Theme Paper**, **Oral**]
- Andreas Opedal, Yanick Zengaffinen, Haruki Shirakami, Clemente Pasti, Mrinmaya Sachan, Abulhair Saparov, Ryan Cotterell, Bernhard Schölkopf (2025). Are Language Models Efficient Reasoners? A Perspective from Logic Programming. Advances in Neural Information Processing Systems 38, NeurIPS 2025. [**Link**]
- Zhoujun Cheng, Shibo Hao, Tianyang Liu, Fan Zhou, Yutao Xie, Feng Yao, Yuexin Bian, Yonghao Zhuang, Nilabjo Dey, Yuheng Zha, Yi Gu, Kun Zhou, Yuqi Wang, Yuan Li, Richard Fan, Jianshu She, Chengqian Gao, Abulhair Saparov, Haonan Li, Taylor W. Killian, Mikhail Yurochkin, Zhengzhong Liu, Eric P. Xing, Zhiting Hu (2025). Revisiting Reinforcement Learning for LLM Reasoning from A Cross-Domain Perspective. Advances in Neural Information Processing Systems 38, NeurIPS 2025. [**Link**]
- Abulhair Saparov, Srushti Pawar, Shreyas Pimpalgaonkar, Nitish Joshi, Richard Yuanzhe Pang, Vishakh Padmakumar, Seyed Mehran Kazemi, Najoung Kim, He He (2024). Transformers Struggle to Learn to Search. International Conference on Learning Representations. [**Link**]
- Andreas Opedal, Haruki Shirakami, Bernhard Schölkopf, Abulhair Saparov, Mrinmaya Sachan (2024). MathGAP: Out-of-Distribution Evaluation on Problems with Arbitrarily Complex Proofs. International Conference on Learning Representations. [**Link**]
- Daking Rai, Yilun Zhou, Shi Feng, Abulhair Saparov, Ziyu Yao (2024). A Practical Review of Mechanistic Interpretability for Transformer-Based Language Models. To appear in ACM Trans. Intell. Syst. Technol. [**Link**]
- Nitish Joshi, Abulhair Saparov, Yixin Wang, He He (2024). LLMs Are Prone to Fallacies in Causal Inference. Empirical Methods in Natural Language Processing, EMNLP 2024. [**Link**]
- Usman Anwar, Abulhair Saparov, Javier Rando, Daniel Paleka, Miles Turpin, Peter Hase, Ekdeep Singh Lubana, Erik Jenner, Stephen Casper, Oliver Sourbut, Benjamin L. Edelman, Zhaowei Zhang, Mario Günther, Anton Korinek, Jose Hernandez-Orallo, Lewis Hammond, Eric Bigelow, Alexander Pan, Lauro Langosco, Tomasz Korbak, Heidi Zhang, Ruiqi Zhong, Seán Ó hÉigeartaigh, Gabriel Recchia, Giulio Corsi, Alan Chan, Markus Anderljung, Lilian Edwards, Aleksandar Petrov, Christian Schroeder de Witt, Sumeet Ramesh Motwan, Yoshua Bengio, Danqi Chen, Philip H.S. Torr, Samuel Albanie, Tegan Maharaj, Jakob Foerster, Florian Tramer, He He, Atoosa Kasirzadeh, Yejin Choi, David Krueger (2024). Foundational Challenges in Assuring Alignment and Safety of Large Language Models. Transactions on Machine Learning Research. [**Link**]
- Andreas Opedal, Alessandro Stolfo, Haruki Shirakami, Ying Jiao, Ryan Cotterell, Bernhard Schölkopf, Abulhair Saparov, Mrinmaya Sachan (2024). Do Language Models Exhibit the Same Cognitive Biases in Problem Solving as Human Learners? Forty-first International Conference on Machine Learning, ICML 2024. [**Link**]
- Nitish Joshi*, Javier Rando*, Abulhair Saparov, Najoung Kim, He He (2024). Personas as a Way to Model Truthfulness in Language Models. Empirical Methods in Natural Language Processing, EMNLP 2024. [**Link**]
  ***equal contribution*
- Abulhair Saparov, Richard Yuanzhe Pang, Vishakh Padmakumar, Nitish Joshi, Seyed Mehran Kazemi, Najoung Kim, He He (2023). Testing the General Deductive Reasoning Capacity of Large Language Models Using OOD Examples. Advances in Neural Information Processing Systems 36, NeurIPS 2023. [**Link**]
- Hongyi Zheng, Abulhair Saparov (2023). Noisy Exemplars Make Large Language Models More Robust: A Domain-Agnostic Behavioral Analysis. Empirical Methods in Natural Language Processing, EMNLP 2023. [**Link**]
- Vaibhav Mavi, Abulhair Saparov, Chen Zhao (2023). Retrieval-Augmented Chain-of-Thought in Semi-structured Domains. Natural Legal Language Processing Workshop @ EMNLP 2023. [**Link**]
- Andreas Opedal, Niklas Stoehr, Abulhair Saparov, and Mrinmaya Sachan (2023). World Models for Math Story Problems. Findings of the Association for Computational Linguistics: ACL 2023. Association for Computational Linguistics. [**Link**]
- Abulhair Saparov, He He (2023). Language Models Are Greedy Reasoners: A Systematic Formal Analysis of Chain-of-Thought. International Conference on Learning Representations. [**Link**]
- Abulhair Saparov, Tom M. Mitchell (2022). Towards General Natural Language Understanding with Probabilistic Worldbuilding. Transactions of the Association for Computational Linguistics (TACL), 10, 325–342. [**Link**, **Oral**]

- **Ph.D. Thesis:** Towards General Natural Language Understanding with Probabilistic Worldbuilding [**Link**]
- Abulhair Saparov (2022). A Probabilistic Generative Grammar for Semantic Parsing. CoRR, abs/1606.06361. [**Link**]
- Emmanouil A. Platanios*, Abulhair Saparov*, Tom M. Mitchell (2020). Jelly Bean World: A Testbed for Never-Ending Learning. International Conference on Learning Representations. [**Link**]
  ***equal contribution*
- Tom M. Mitchell, William Cohen, Estevam Hruschka, Partha Talukdar, Bishan Yang, Justin Betteridge, Andrew Carlson, Bhavana Dalvi, Matther Gardner, Bryan Kisiel, Jayant Krishnamurthy, Ni Lao, Kathryn Mazaitis, Thahir Mohammad, Ndapa Nakashole, Emmanouil A. Platanios, Alan Ritter, Mehdi Samadi, Burr Settles, Richard Wang, Derry Wijaya, Abhinav Gupta, Xinlei Chen, Abulhair Saparov, Malcolm Greaves, Joel Welling (2018). Never-Ending Learning. Communications of the ACM, 61(5), 103–115. [**Link**]
- Abulhair Saparov, Vijay Saraswat, Tom M. Mitchell (2017). A Probabilistic Generative Grammar for Semantic Parsing. *Proceedings of the Twenty-First Conference on Computational Natural Language Learning*. [**Link**, **Oral**]
- Abulhair Saparov, Michael A. Schwemmer (2015). Effects of passive dendritic tree properties on the firing dynamics of a leaky-integrate-and-fire neuron. *Mathematical Biosciences*, 269, 61-75.
- Tom M. Mitchell, William Cohen, Estevam Hruschka, Partha Talukdar, Justin Betteridge, Andrew Carlson, Bhavana Dalvi, Matthew Gardner, Bryan Kisiel, Jayant Krishnamurthy, Ni Lao, Kathryn Mazaitis, Thahir Mohammad, Ndapa Nakashole, Emmanouil A. Platanios, Alan Ritter, Mehdi Samadi, Burr Settles, Richard Wang, Derry Wijay, Abhinav Gupta, Xinlei Chen, Abulhair Saparov, Malcolm Greaves, Joel Welling (2015). Never-ending Learning. AAAI. [**Link**]
- Xiaobai Chen, Abulhair Saparov, Bill Pang, and Thomas Funkhouser (2012). Schelling Points on 3D Surface Meshes, *ACM Transactions on Graphics (Proc. SIGGRAPH)*. [**Link**]

Honors & Awards

2023	**Best Reviewer,** Conference on Empirical Methods in Natural Language Processing	*Singapore*
2023	**Outstanding Area Chair,** 61st Meeting of the Association for Computational Linguistics	*Toronto, Canada*
2015	**Teaching Assistant Award,** Machine Learning Department, CMU	*Pittsburgh, PA*
2013	**Honorable Mention,** NSF Graduate Research Fellowship	
2013	**Best Paper,** Program in Applied and Computational Mathematics, Princeton University	*Princeton, NJ*

Teaching

CS390AIF AI Forge
West Lafayette, IN

Instructor — *Fall 2026*

- Developed new lower-level undergraduate introductory course designed to quickly train responsible practitioners of AI.
- Enrollment: 22 students currently enrolled.

CS490 Natural Language Processing
West Lafayette, IN

Instructor — *Spring 2026*

- Developed new undergraduate version of CS577 Natural Language Processing with a focus on training practitioners of NLP methods.
- Sponsored the addition of a new permanent undergraduate course in the Purdue CS curriculum (CS477).
- Enrollment: 51 students completed course.

CS577 Natural Language Processing
West Lafayette, IN

Instructor — *Fall 2025*

- Adapted graduate course on NLP with new content in rapidly-changing field, with an updated slide deck.
- Enrollment: 71 students completed course.

CS577 Natural Language Processing
West Lafayette, IN

Instructor — *Spring 2025*

- Adapted graduate course on NLP with new content in rapidly-changing field, with an updated slide deck.
- Enrollment: 100 students completed course.

CS592-LLM Can Machines Think? Reasoning with Large Language Models
West Lafayette, IN

Instructor — *Fall 2024*

- Designed and developed syllabus for graduate seminar course on recent research on LLMs and reasoning.
- Enrollment: 35 students completed course.

10-601 Introduction to Machine Learning *Pittsburgh, PA*

Teaching Assistant *Jan. 2015 - May 2015*

- Recorded, edited, and uploaded lecture videos to YouTube.
- Led recitations, created and graded homework assignments/exams, and supervised student groups on final project work.
- Received the Teaching Assistant Award from the Machine Learning Department.

10-701/15-781 Introduction to Machine Learning *Pittsburgh, PA*

Teaching Assistant *Sep. 2014 - Dec. 2014*

- Recorded, edited, and uploaded lecture videos to YouTube.
- Led recitations, created and graded homework assignments/exams, and supervised student groups on final project work.

Experience

Purdue Department of Computer Science *West Lafayette, IN*

Assistant Professor *August 2024 -*

- Continuing research on utilizing symbolic methods to better study and understand NLU models, including LLMs.
- As well as research into improving the reasoning abilities of NLU models by incorporating symbolic methods.
- Teaching seminar and lecture courses in NLP and LLMs.

Nace.AI *Palo Alto, CA*

Chief Scientist *June 2025 -*

- Leading the research efforts of the company, and providing guidance on relevant cutting-edge NLP research.
- Advising company on the development of methods and strategies for solving reasoning-intensive tasks.

Centaur AI Institute *Lincoln, CA*

Consultant *June 2025 -*

- Advising instutute on the development of novel neuro-symbolic models to build more robust AI systems.
- Guiding research into AI systems that are hallucination-free by design by leveraging symbolic representations.

NYU Center for Data Science *New York, NY*

Post-Doctoral Associate *June 2022 - July 2024*

- Worked on reasoning in NLP and NLU; deductive reasoning, causal reasoning, truthfulness in large language models.
- Advisor: He He, Research Group: ML2, CILVR

IBM Thomas J. Watson Research Center *Yorktown Heights, NY*

Research Intern *Summer 2016*

- Developed grammar induction algorithms to train a semantic parser on a dataset of questions and corresponding logical forms.
- Advisor: Vijay Saraswat

The McGraw Center *Princeton, NJ*

Freshman Scholars Institute QuantLab Tutor *Summer 2012*

- Led group tutoring sessions for incoming Princeton University freshman students coming from schools without access to AP courses.

Princeton Department of Computer Science *Princeton, NJ*

Research Assistant *Feb. 2012 - May 2013*

- Developed and implemented novel inference algorithms to fit a probabilistic model of latent sources to fMRI brain activity data.
- Advisors: Sam Gershman, Ken Norman, David M. Blei

Princeton Program in Applied and Computational Mathematics *Princeton, NJ*

Research Assistant *Oct. 2011 - May 2013*

- Explored and analyzed novel dynamical behavior in the multi-compartment leaky integrate-and-fire model of a neuron.
- Advisors: Michael A. Schwemmer, Philip J. Holmes

Google *New York, NY*

Software Engineering Intern *Summer 2011*

- Designed and developed new features for the front-end of the DoubleClick for Publishers advertising platform, collaborating with many different software engineering and product management groups at every stage of the test-driven development process.

Re-verb Apparel Co.

Chief Web Designer *Nov. 2010 - 2012*

- Implemented and helped design the website, including online store, using HTML5 frontend technologies and a Wordpress backend.

Princeton Department of Computer Science *Princeton, NJ*

Research Assistant *Jun. 2010 - Jan. 2011*

- Designed a novel feature point detection algorithm for 3D surface meshes.
- Analyzed large data sets gathered from an online user study to find patterns in human-generated point sets on 3D meshes.
- Advisor: Thomas Funkhouser.

Talks

March 2026	Guest Lecture for CS592EAI: Appls. of ML and Explainable AI (**INVITED**)	*Purdue University*
November 2025	Can/Will Language Models Learn to Search? (**INVITED**)	*CERIAS, Purdue University*
November 2025	Can/Will Language Models Learn to Search? (**INVITED**)	*Centaur AI Institute*
March 2025	Transformers Struggle to Learn to Search (**INVITED**)	*Boston University*
November 2024	Can Language Models Reason? (**INVITED**)	*New York University*
October 2024	Guest Lecture for CS591RS: Graduate Research Seminar	*Purdue University*
October 2024	Guest Lecture for CS397: Honors Seminar	*Purdue University*
August 2023	Reasoning with Large Language Models (**INVITED**)	*Neuro-symbolic Summer School*

Professional service

NAACL 2027	Website Chair	*Organizing Committee*
AAAI 2027	Special Track on AI Alignment	*Senior Program Committee*
ICLR 2027	International Conference on Learning Representations	*Area Chair*
ACL 2026	Website and Conference App Chair	*Organizing Committee*
IJCAI-ECAI 2026	Workshop on Neuro-Symbolic Intelligence for LLMs and Autonomous Agents	*Organizing Committee*
ACL 2026	ACL Rolling Review, January 2026	*Area Chair*
EACL 2026	ACL Rolling Review, October 2025	*Area Chair*
COLM 2025	Workshop on Applications of LLM Explainability to Reasoning and Planning	*Organizing Committee*
AACL-IJCNLP 2025	ACL Rolling Review, July 2025	*Area Chair*
NSF	National Science Foundation, June 2025	*Panelist*
EMNLP 2025	ACL Rolling Review, May 2025	*Area Chair*
ACL 2025	ACL Rolling Review, February 2025	*Area Chair*
NeurIPS 2024	Workshop on Socially Responsible Language Modelling Research	*Organizing Committee*
AAAI 2024	Workshop on Neuro-Symbolic Learning and Reasoning in the Era of LLMs	*Organizing Committee*
ACL 2023	Main conference: Sentence-level Semantics Track	*Area Chair*
SDB 2019	Workshop on Machine Learning for Developmental Biology	*Organizing Committee*

Other information

Languages native in English, conversational in Kazakh, intermediate in Spanish, beginner in Japanese