

Abulhair Saparov

MACHINE LEARNING PHD CANDIDATE

8003 Gates Hillman Ctr, 5000 Forbes Avenue, Pittsburgh, Pennsylvania, 15213, United States

[✉ asaparov@cs.cmu.edu](mailto:asaparov@cs.cmu.edu) | [🏠 asaparov.org](http://asaparov.org) | [📘 asaparov](https://www.facebook.com/asaparov) | [📺 YouTube](https://www.youtube.com/channel/UC8v31111111111111111111)

Research interests

- Statistical machine learning, interpretable machine learning, Bayesian nonparametrics, scalable inference
- Applications of machine learning to natural language processing (NLP), semantic parsing, natural language understanding (NLU)
- Representations of meaning/knowledge, reasoning, especially in natural language understanding
- Improving NLP by incorporating prior knowledge, such as from a knowledge base (KB), or other forms of indirect supervision
- Broadly interested in applications in e.g. epidemiology, linguistics, phylogenetics, biology, information security, etc.

Education

Carnegie Mellon University

Pittsburgh, PA

PHD CANDIDATE IN MACHINE LEARNING

Sep. 2013 - PRESENT

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

Princeton University

Princeton, NJ

BSE IN COMPUTER SCIENCE

Sep. 2009 - May 2013

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

Publications

- **Abulhair Saparov**, Tom M. Mitchell (2022). Towards General Natural Language Understanding with Probabilistic Worldbuilding. Transactions of the Association for Computational Linguistics (TACL), (accepted). [\[link\]](#)
- **Ph.D. Thesis**: [Towards General Natural Language Understanding with Probabilistic Worldbuilding](#) [\[link\]](#)
- **Ph.D. Thesis Proposal**: [Never-Ending Learning for Open-Domain Semantic Parsing](#) [\[link\]](#)
- 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*. [\[link\]](#)
- **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)*. [\[link\]](#)

Experience

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 in calculus.

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, consulting 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](#).

Teaching

10-601 Introduction to Machine Learning (graduate course)

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.
- Teaching Assistant Award.

10-701/15-781 Introduction to Machine Learning (graduate course)

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.

Honors & Awards

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

Other information

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