NATHANIEL WEIR

Department of Computer Science
Johns Hopkins University · Baltimore, MD
https://cs.jhu.edu/~nweir · nweir@jhu.edu

EDUCATION

2019–Present **Johns Hopkins University**
Ph.D. in Computer Science
Advisor: Benjamin Van Durme
Thesis: *Compositional Neuro-Symbolic Reasoning over Natural Language*

2019–2021 **Johns Hopkins University**
M.S.E. in Computer Science
Advisor: Benjamin Van Durme

2015–2019 **Brown University**
Sc.B. in Applied Mathematics & Computer Science, *magna cum laude* with Honors (GPA: 3.95/4)
Thesis: *Bootstrapping Generalization in Neural Text-to-SQL Semantic Parsing Models*
Advisors: Ugur Cetintemel, Carsten Binnig, and Ellie Pavlick

RESEARCH INTERESTS

Natural Language Inference, Large Language Models, Neuro-symbolic Reasoning, Knowledge Grounding, Explainable AI, Information Retrieval, Logical Reasoning, Question Answering

RESEARCH EXPERIENCE

Fall 2019 – Present **Center for Language and Speech Processing at Johns Hopkins University**
PhD Researcher
Research includes retrieval-augmented, language model-based neuro-symbolic reasoning, constrained language generation, and improving robustness of question answering models.

Summer 2023 – Winter 2023 **Allen Institute for Artificial Intelligence**
PhD Research Intern – Aristo Reasoning Team
Mentor: Peter Clark
Researched extracting and reasoning over generalizable scientific microtheories from language models. Also led a project on code-based instruction tuning and logically constrained abductive situational reasoning.

Summer 2022 **Microsoft Semantic Machines**
PhD Research Intern
Mentor: Harsh Jhamtani
Worked on knowledge-constrained dialogue generation. Published a benchmark dataset of real video game dialogue trees annotated with granular ontology constraints.

Summer 2021 **Microsoft Research – Montreal**
PhD Research Intern – Deep Learning and Language Team
Mentors: Harm Van Seijen, Xingdi Yuan and Marc-Alexandre Côté
Explored hierarchical language-based planning in embodied agents. Published a benchmark testing for learning compositional tasks in a Minecraft-like domain.

Spring 2017 – May 2019 **Brown University**
Undergraduate Researcher
Researched domain adaptation methods for text-to-SQL neural semantic parsing.

PREPRINTS

2024 **Nathaniel Weir**, Muhammad Khalifa, Linlu Qiu, Orion Weller, and Peter Clark. Learning to Reason via Program Generation, Emulation, and Search.

2024 Dongwei Jiang, Jingyu Zhang, Orion Weller, **Nathaniel Weir**, Benjamin Van Durme, Daniel Khashabi. SELF-[IN]CORRECT: LLMs Struggle with Refining Self-Generated Responses.

2024 **Nathaniel Weir**, Kate Sanders, Orion Weller, Shreya Sharma, Dongwei Jiang, Zhengping Zhang, Bhavana Dalvi Mishra, Oyvind Tafjord, Peter Jansen, Peter Clark, and Benjamin Van Durme. Enhancing Systematic Decompositional Natural Language Inference Using Informal Logic.

2024 Kate Sanders, **Nathaniel Weir**, and Benjamin Van Durme. TV-TREES: Multimodal Entailment Trees for Neuro-Symbolic Video Reasoning.

2023 **Nathaniel Weir**, Ryan Thomas, Randolph d'Amore, Kellie Hill, Benjamin Van Durme, and Harsh Jhamtani. Ontologically Faithful Generation of Non-Player Character Dialogues. Presented non-archivally at DialDoc 2023.

PUBLICATIONS

2024 **Nathaniel Weir**, Peter Clark, and Benjamin Van Durme. NELLIE: A Neuro-Symbolic Inference Engine for Grounded, Compositional, and Explainable Reasoning. IJCAI 2024.

2024 Xinrui Zou, Ming Zhang, **Nathaniel Weir**, Benjamin Van Durme, and Nils Holzenburger. Reframing Tax Law Entailment as Analogical Reasoning. Jurix Special Workshop on AI, Law and Philosophy.

2023 Orion Weller, Marc Marone, **Nathaniel Weir**, Dawn Lawrie, Daniel Khashabi, and Benjamin Van Durme. "According to ..." Prompting Language Models Improves Quoting from Pre-Training Data. EACL 2024.

2022 Orion Weller, Aleem Khan, **Nathaniel Weir**, Dawn Lawrie, and Benjamin Van Durme. Defending Against Poisoning Attacks in Open-Domain Question Answering. EACL 2024.

2022 **Nathaniel Weir**, Xingdi Yuan, Marc-Alexandre Côté, Matthew J. Hausknecht, Romain Laroche, Ida Momennejad, Harm van Seijen and Benjamin Van Durme. One-Shot Learning from a Demonstration with Hierarchical Latent Language. *AAMAS*.

2021 Jiefu Ou*, **Nathaniel Weir***, Anton Belyy*, Felix Yu, and Benjamin Van Durme. InFillmore: Frame-Guided Language Generation with Bidirectional Context. *StarSem.*

2020 **Nathaniel Weir**, João Sedoc, and Benjamin Van Durme. COD3S: Diverse Generation with Discrete Semantic Signatures. *EMNLP.* **Oral Presentation.**

2020 **Nathaniel Weir**, Adam Poliak, and Benjamin Van Durme. Probing Neural Language Models for Human Tacit Assumptions. *CogSci.* **Oral Presentation.**

2020 **Nathaniel Weir**, Prasetya Utama, Alex Galakatos, Andrew Crotty, Amir Ilkhechi, Shekar Ramaswamy, Rohin Bhusan, Nadja Geisler, Benjamin Hattasch, Steffen Eger, Ugur Cetintemel, and Carsten Binnig. DBPal: A Fully Pluggable NL2SQL Training Pipeline. *SIGMOD.*
Oral Presentation.

2018 Fuat Basik, Benjamin Hattasch, Amir Ilkhechi, Arif Usta, Shekar Ramaswamy, Prasetya Utama, **Nathaniel Weir**, Carsten Binnig and Ugur Cetintemel. DBPal: A Learned NL-Interface for Databases. *SIGMOD.* **Demo Presentation.**

2017 Prasetya Utama, **Nathaniel Weir**, Carsten Binnig, and Ugur Cetintemel. Voice-based Data Exploration: Chatting with your Database. *SCAI.*

AWARDS

2021 - 2024 **NSF Graduate Research Fellowship**

2019 **CRA Outstanding Undergraduate Researcher Award, Honorable Mention**

2019 **SIGMOD Undergraduate Research Competition, 2nd place**

2019 **Brown Computer Science Senior Award**

2019 **Sigma Xi Electee**

2019 **Brown CS Undergraduate Research Symposium, 3rd Place**

2018 **Randy F. Pausch CS Undergraduate Summer Research Award, $10,000**

2018 **NSF Travel Grant, $1,000**

2017 **Karen T. Romer Undergraduate Teaching and Research Award, $3,500**

PRESENTATIONS

Ontologically Faithful Generation of Non-Player Dialogues

07/2023 Talk DialDoc Workshop at ACL 2023

Language Models as Proposal Functions in a Neuro-Symbolic Expert System

05/2023 Talk Massachussetts Institute of Technology
05/2023 Talk Brown University

One-Shot Learning from a Demonstration with Hierarchical Latent Language

06/2023 Poster AAMAS 2023
03/2022 Talk 9th Mid-Atlantic Student Colloquium on Speech, Language and Learning

cod3s: Diverse Generation with Discrete Semantic Symbols

06/2020 Talk EMNLP 2020

Probing Neural Language Models for Human Tacit Assumptions

06/2020 Talk CogSci 2020
03/2020 Poster 8th Mid-Atlantic Student Colloquium on Speech, Language and Learning

DBPal: A Fully Pluggable Natural Language Interface to Databases

06/2020 Talk ACM SIGMOD/PODS Conference
01/2019 Talk North East Database Day @ MIT
10/2018 Talk IBM AI Systems Day @ MIT
01/2018 Demo North East Database Day @ MIT

TECHNICAL SKILLS

Languages Python (expert), R, Bash, Prolog, C/C++, SQL, Java, MATLAB, Scala

Tools PyTorch, HuggingFace Transformers, LangChain, fairseq, faiss, vLLM, DeepSpeed, MySQL, Jupyter, RStudio, NLTK, Mechanical Turk

TEACHING EXPERIENCE

Teaching Assistant

Johns Hopkins University

Fall 2022 CS 601.470/670: Artificial Agents. Instructor: Benjamin Van Durme (co-taught)

Brown University

Fall 2018 CSCI1570: Design and Analysis of Algorithms. Instructor: Paul Valiant

Spring 2018 CSCI0220: Discrete Structures and Probability. Instructor: Caroline Klivans

Spring 2017 CSCI0220: Discrete Structures and Probability. Instructor: Caroline Klivans

SERVICE

Reviewing:
Primary: ACL Rolling Review (2022, 2023, 2024), EMNLP 2023, KnowledgeNLP-AAAI'23
Secondary: ICLR 2021, ACL 2020, AKBC 2020

Committees:
JHU CS Ph.D. Admissions Committee (2021-2024)
JHU CLSP Application Support Program for underrepresented students (2022-2023)

MENTORING

- Shreya Sharma (JHU MS, 2023-2024)
- Dongwei Jiang (JHU MS, joint with Orion Weller, 2023-2024)
- Sandipan Majhi (JHU MS, 2023)
- Jingyu Zhang (JHU BS, 2022-2023)
- Xiao Ye (JHU MS, 2022)
- Chenyu Zhang (JHU BS, 2020-2021)
- Jiefu Ou (HKUST BS, 2020-2021)
- Wei Liu (UIUC BS, 2020)

COURSEWORK

Undergraduate GPA: 3.95 *Graduate GPA:* 4.0

Natural Language Processing / Artificial Intelligence: Computational Semantics, Applied Event Semantics, Deep Learning for Dialog, Natural Language Processing, Machine Learning, Causal Inference, Artificial Intelligence

Computational Cognitive Science: Computational Psycholinguistics, Logic in Language and Thought, Computational Cognitive Science

Mathematics: Computational Probability and Statistics, Information Theory, Pattern Theory, Computational Linear Algebra, Ordinary/Partial Differential Equations, Multivariate Calculus

Computer Science: Data Science, Probabilistic Algorithms, Sketching and Indexing, Compilers, Language-based Security, Prescriptive Analytics, Algorithms, Systems