John Hewitt

Ph.D. Student · Natural Language Processing · Machine Learning

Stanford, CA · johnhew@stanford.edu · https://nlp.stanford.edu/~johnhew/

Education

Stanford University 2018–

Ph.D. Computer Science. *Advised by Chris Manning and Percy Liang. Funded by NSF GRF.*

University of Pennsylvania 2014–2018

B.S.E. Computer and Information Science.

Research Experience

DeepMind June 2022–

Research Scientist Intern*, with Aida Nematzadeh and Adhiguna Kuncoro*

Stanford University September 2018–

Research Assistant*, with Chris Manning, Percy Liang*

Google AI September 2020–February 2021

Research Intern*, with Vincent Zhao and Kelvin Guu*

Amazon AI May 2018–September 2018

Applied Scientist Intern*, with Katrin Kirchhoff and Arshit Gupta*

University of Pennsylvania 2016–2018

Research Assistant*, with Chris Callison-Burch*

Johns Hopkins University May 2015–May 2016

Research Assistant*, with David Yarowsky and Matt Post*

Publications

Conditional probing: measuring usable information beyond a baseline
John Hewitt, Kawin Ethayarajh, Percy Liang, and Christopher D. Manning.
In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Punta Cana, Dominican Republic. November 2021.

On the Opportunities and Risks of Foundation Models
Bommasani et. al. **John Hewitt**: co-lead, Interpretability section.
In *ArXiv*. Virtual. August 2021.

Refining Targeted Syntactic Evaluation of Language Models
Benjamin Newman, Kai-Siang Ang, Julia Gong, and **John Hewitt**.
In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics*. Virtual. June 2021.

Probing artificial neural networks: Insights from neuroscience
Anna Ivanova, **John Hewitt**, and Noga Zaslavsky.
In *Proceedings of the Brain2AI Workshop*. Virtual. May 2021.

RNNs can generate bounded hierarchical languages with optimal memory
John Hewitt, Michael Hahn, Surya Ganguli, Percy Liang, and Christopher D. Manning.
In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Virtual. November 2020.

The EOS Decision and Length Extrapolation
Benjamin Newman, **John Hewitt**, Percy Liang, and Christopher D. Manning.
In *BlackBoxNLP: Analyzing and Interpreting Neural Networks for NLP Workshop*. Virtual. November 2020
Outstanding Paper Award.

Emergent Linguistic Structure in Artificial Neural Networks Trained by Self-Supervision
Christopher D. Manning, Kevin Clark, **John Hewitt**, Urvashi Khandelwhal, and Omer Levy.
Proceedings of the National Academy of Sciences. June 2020.

Finding Universal Grammatical Relations in Multilingual BERT
Ethan Chi, **John Hewitt**, and Christopher D. Manning.
In *Proceedings of the Conference of the Association for Computational Linguistics.*. Virtual. July 2020.

Designing and Interpreting Probes with Control Tasks
John Hewitt and Percy Liang.
In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Hong Kong, China. November 2019.
Runner Up Best Paper Award.

A Structural Probe for Finding Syntax in Word Representations
John Hewitt and Christopher D. Manning.
In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Minneapolis. June 2019.

Simple, Fast, Accurate Intent Classification and Slot Labeling for Goal-Oriented Dialogue Systems
Arshit Gupta* and **John Hewitt*** and Katrin Kirchhoff.
In *Proceedings of the SIGDIAL 2019 Conference*. Stockholm, Sweden. September 2019.
*: Equal contribution; authors listed alphabetically.

A Distributional and Orthographic Aggregation Model for English Derivational Morphology
Daniel Deutsch*, **John Hewitt*** and Dan Roth.
In *Proceedings of the Conference of the Association for Computational Linguistics*. Melbourne, Australia. July 2018.
*: Equal contribution; authors listed alphabetically.

Learning Translations via Images with a Massively Multilingual Image Dataset
John Hewitt*, Daphne Ippolito*, Brendan Callahan, Reno Kriz, Derry Tanti Wijaya and Chris Callison-Burch.
In *Proceedings of the Conference of the Association for Computational Linguistics*. Melbourne, Australia. July 2018.
*: Equal contribution; authors listed alphabetically.

XNMT: The eXtensible Neural Machine Translation Toolkit
Graham Neubig, Matthias Sperber, Xinyi Wang, Matthieu Felix, Austin Matthews, Sarguna Padmanabhan, Ye Qi, Devendra Singh Sachan, Philip Arthur, Pierre Godard, **John Hewitt**, Rachid Riad, and Liming Wang.
In *Conference of the Association for Machine Translation in the Americas (AMTA) Open Source Software Showcase*. Boston. March 2018.

Learning Translations via Matrix Completion
Derry Tanti Wijaya, Brendan Callahan, **John Hewitt**, Xiao Ling, Marianna Apidianaki, and Chris Callison-Burch.
In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Copenhagen, Denmark. September 2017.

Automatic Construction of Morphologically-Motivated Translation Models for Highly Inflected Low-Resource Languages
John Hewitt, Matt Post, David Yarowsky.
In *Proceedings of the Conference of the Association for Machine Translation in the Americas*. Austin. October 2016.

Teaching & Mentoring Experience

- *Head Teaching Assistant, Co-Instructor*. Stanford. CS 224N 2021.
  Co-designed curriculum, wrote and taught lectures, managed team of 22 TAs. Awarded for being in top 5% of TAs in Department of Computer Science.
- *Teaching Assistant*. Stanford. CS 224N 2020.
  Led custom final project program; taught two lectures.
- *Mentor*. Stanford. General Advising, 2019-2022.
  Mentored Undergraduate and Master's student research projects. (Benjamin Newman, Ruth-Ann Armstrong).
- *Mentor*. Stanford. CURIS: Summer Research for Undergraduates, 2019.
  Led summer research experience for an undergraduate working on multilingual language model understanding. (Ethan Chi.)
- *Mentor*. Stanford. ROHU: Research Office Hours for Undergraduates.
  Founded weekly open office hours, mentoring undergraduates in research.
- *Mentor*. Stanford. AI Undergraduate Mentorship Program.
  Participating in 1-on-1 mentorship program (2 years running)
- *Project Mentor*. Stanford. CS 224N: Natural Language Processing with Deep Learning.
  Advised an undergraduate team working on textbook glossary extraction. (Kush Khosla, Robbie Jones, Nicholas Bowman.)
- *Teaching Assistant*. Penn. CIS 530: Computational Linguistics.
  Wrote homeworks and a lecture; advised final project teams.
- *Teaching Assistant*. Penn. CIS 121: Data Structures and Algorithms.
  Taught recitation (14 students) and held office hours.

– *Volunteer Instructor*. Old Rochester Regional High School. The Math that Runs the World.
Taught 3-lesson series and additional lectures on mathematics and algorithms.

Patents

– John Hewitt.
Capturing Rich Response Relationships with Small-Data Neural Networks.
US Patent App 15/841,963. December 2017 (granted 2019-08-13). Assigned to Qualtrics, Inc.

Professional Service

Reviewer

– **ACL Rolling Review** 2021, 2022.
– **Computational Linguistics Journal** 2021.
– **ACL** 2018 *top reviewer*, 2020 *top reviewer*.
– **NAACL** 2021.
– **EMNLP** 2018.
– **EACL** 2021.
– **AACL** 2020.
– **CoNLL** 2020.
– **BlackBoxNLP** 2020, 2021, 2022.
– **DistShift NeurIPS Workshop on Distribution Shifts** 2021, 2022.
– **DeeLIO Workshop on Deep Learning Knowledge Extraction and Integration** 2020.
– **ACL-SRW** 2019.

Departmental service

– **Stanford NLP Group Social Organizer** 2019–2020.
– **Stanford CS PhD Admissions Committee** 2020.

Invited Talks

An NLP perspective on supervised analysis of neural representations. Ev Fedorenko's EvLab (MIT), December 2020.
The Unreasonable Syntactic Expressivity of RNNs. USC ISI NLP Seminar, November 2020.
Language probes as $\mathcal{V}$-information estimators. NLP with Friends, September 2020.
Probing Neural NLP: Ideas and Problems. Berkeley NLP Seminar, November 2019.
Emergent Linguistic Structure in Neural NLP. Amazon AI, July 2019.
A Structural Probe for Finding Syntax in Word Representations. NLP Highlights Podcast of the Allen Institute for Artificial Intelligence, May 2019.
A Structural Probe for Finding Syntax in Word Representations. Stanford Human-Centered AI Initiative Symposium, March 18, 2019.

Grants & Awards

BlackBoxNLP 2020 Outstanding Paper Award.
For The EOS Decision and Length Extrapolation
Two Sigma Fellowship 4^{th} Place Prize.
2020.
NSF Graduate Research Fellowship.
(2020) In Computer Science – Natural Language Processing
EMNLP 2019 Runner-Up Best Paper Award.
For Designing and Interpreting Probes with Control Tasks
Amazon Public Datasets. 2 years' hosting of large research dataset awarded.
To provide unlimited access to the Massively Multilingual Image Dataset.
Pinkel Endowment Undergraduate Research Fund.
For the advancement of undergraduate research in cognitive science; used to attend EMNLP 2017.
Penn Engineering CIS Senior Design Fair Honorable Mention.
For excellence in year-long research project on translation of the world's languages through visual features.