Samuel J. Wiseman

swiseman@cs.duke.edu

Current Appointment	**Assistant Professor** Duke University, Department of Computer Science	July 2021 – present
Previous Appointments	**Research Assistant Professor** Toyota Technological Institute at Chicago	Sept. 2018 – July 2021
Education	**Harvard University**, Ph.D., Computer Science Dissertation: *Structured Neural Models for Coreference and Generation* Advisors: Alexander M. Rush, Stuart M. Shieber	Sept. 2012 – May 2018
	Princeton University, A.B., Philosophy, Magna Cum Laude Certificate: Program in Applications of Computing	June 2010
Honors and Awards	**Best Short Paper**, CRAC Workshop @ EMNLP	2021
	Siebel Scholar	2018
	Honorable Mention for Best Paper, EMNLP	2016
	Harvard Bok Center Certificate of Distinction in Teaching	Spring 2014, 2016
	Phi Beta Kappa, Princeton University	June 2010

Refereed Conference Publications

Baked-in State Probing. Shubham Toshniwal, Sam Wiseman, Karen Livescu, Kevin Gimpel. *Findings of EMNLP* 2022.

SummScreen: A Dataset for Abstractive Screenplay Summarization. Mingda Chen, Zewei Chu, Sam Wiseman, Kevin Gimpel. *ACL*, 2022.

Chess as a Testbed for Language Model State Tracking. Shubham Toshniwal, Sam Wiseman, Karen Livescu, Kevin Gimpel. *AAAI*, 2022.

Data-to-text Generation by Splicing Together Nearest Neighbors. Sam Wiseman, Arturs Backurs, Karl Stratos. *EMNLP*, 2021.

WikiTableT: A Large-scale Data-to-text Dataset for Generating Wikipedia Article Sections. Mingda Chen, Sam Wiseman, Kevin Gimpel. *Findings of ACL*, 2021.

Learning to Ignore: Long Document Coreference with Bounded Memory Neural Networks. Shubham Toshniwal, Sam Wiseman, Allyson Ettinger, Karen Livescu, and Kevin Gimpel. *EMNLP*, 2020.

Learning Discrete Structured Representations by Adversarially Maximizing Mutual Information. Karl Stratos and Sam Wiseman. *ICML*, 2020.

ENGINE: Energy-Based Inference Networks for Non-Autoregressive Machine Translation. Lifu Tu, Richard Yuanzhe Pang, Sam Wiseman, and Kevin Gimpel. *ACL*, 2020.

Discrete Latent Variable Representations for Low-Resource Text Classification. Shuning Jin, Sam Wiseman, Karl Stratos, and Karen Livescu. *ACL*, 2020.

Amortized Bethe Free Energy Minimization for Learning MRFs. Sam Wiseman and Yoon Kim. *NeurIPS*, 2019.

Label-Agnostic Sequence Labeling by Copying Nearest Neighbors. Sam Wiseman and Karl Stratos. *ACL*, 2019.

Controllable Paraphrase Generation with a Syntactic Exemplar. Mingda Chen, Qingming Tang, Sam Wiseman, and Kevin Gimpel. *ACL*, 2019.

A Multi-Task Approach for Disentangling Syntax and Semantics in Sentence Representations. Mingda Chen, Qingming Tang, Sam Wiseman, and Kevin Gimpel. *NAACL*, 2019.

Learning Neural Templates for Text Generation. Sam Wiseman, Stuart M. Shieber, and Alexander M. Rush. *EMNLP*, 2018.

Entity Tracking Improves Cloze-style Reading Comprehension. Luong Hoang, Sam Wiseman, and Alexander M. Rush. *EMNLP*, 2018.

Semi-Amortized Variational Autoencoders. Yoon Kim, Sam Wiseman, Andrew C. Miller, David Sontag, Alexander M. Rush. *ICML*, 2018.

Challenges in Data-to-Document Generation. Sam Wiseman, Stuart M. Shieber, and Alexander M. Rush. *EMNLP*, 2017.

Sequence-to-Sequence Learning as Beam Search Optimization. Sam Wiseman and Alexander M. Rush. *EMNLP*, 2016. Honorable Mention for Best Paper.

Learning Global Features for Coreference Resolution. Sam Wiseman, Alexander M. Rush, and Stuart M. Shieber. *NAACL*, 2016.

Learning Anaphoricity and Antecedent Ranking Features for Coreference Resolution. Sam Wiseman, Alexander M. Rush, Stuart M. Shieber, and Jason Weston. *ACL*, 2015.

Discriminatively Reranking Abductive Proofs for Plan Recognition. Sam Wiseman and Stuart Shieber. *ICAPS*, 2014.

Workshop and Preprint Papers

Sequence Reducible Holdout Loss for Language Model Pretraining. Raghuveer Thirukovalluru, Bhuwan Dhingra, Sam Wiseman. To appear at SustaiNLP @ ACL 2023.

Exploring the Effect of Frequency Resolution in FNet. Greg Szumel, Ghazal Khalighinejad, Rickard Stureborg, Sam Wiseman. To appear at SustaiNLP @ ACL 2023.

BM25 Query Augmentation Learned End-to-end. Xiaoyin Chen, Sam Wiseman. arXiv:2305.14087. May 2023.

Approximating CKY with Transformers. Ghazal Khalighinejad, Ollie Liu, Sam Wiseman. arXiv:2305.02386. May 2023.

CREATIVESUMM: Shared Task on Automatic Summarization for Creative Writing. Divyansh Agarwal et al. (including Sam Wiseman). Workshop on Automatic Summarization for Creative Writing @ COLING 2022.

On Generalization in Coreference Resolution. Shubham Toshniwal, Patrick Xia, Sam Wiseman, Karen Livescu, Kevin Gimpel. CRAC Wokshop @ EMNLP, 2021.

Learning Deep Latent-variable MRFs with Amortized Bethe Free Energy Minimization. Sam Wiseman. DeepGenStruct @ ICLR, 2019.

A Tutorial on Deep Latent Variable Models of Natural Language. Yoon Kim*, Sam Wiseman*, Alexander M. Rush. arXiv:1812.06834. EMNLP 2018 Tutorial Document.

Training Language Models Using Target-Propagation. Sam Wiseman, Sumit Chopra, Marc'Aurelio Ranzato, Arthur Szlam, Ruoyu Sun, Soumith Chintala, Nicolas Vasilache. arXiv:1702.04770, February 2017.

Antecedent Prediction without a Pipeline. Sam Wiseman, Alexander M. Rush, and Stuart M. Shieber. CORBON Workshop @ NAACL, 2016.

Teaching Experience	**Instructor**	
	• Duke CS 590.06 (Now CS 574): Fundamentals of Deep Learning	Spring 2023
	• Duke CS 590.03: Neurosymbolic ML	Fall 2022
	• Duke CS 590.03: Advanced Natural Language Processing	Spring 2022
	– Co-taught with Bhuwan Dhingra	
	• Duke CS 590.03 (Now CS 572): Introduction to NLP	Fall 2021
	– Co-taught with Bhuwan Dhingra	
	Teaching Fellow	
	• Harvard CS 287: Statistical Natural Language Processing	Spring 2016
	• Harvard CS 187: Computational Linguistics	Fall 2014
	• Harvard CS 181: Machine Learning	Spring 2014

Service

- Publications Co-Chair: EMNLP 2022
- Organizing Committee: Creative-Summ Workshop @ COLING 2022, Midwest Speech and Language Days 2019
- Area Chair: EMNLP 2021 (Machine Learning track), ACL 2020 (Generation track), EMNLP 2020 (Generation track)
    - EMNLP 2020 Outstanding AC Award
- Reviewing: TACL, ACL ARR, EMNLP, ICML, ICLR, NeurIPS
    - NeurIPS 2021 Outstanding Reviewer Award
    - NAACL 2018 Outstanding Reviewer Award
- EMNLP 2018 Tutorial: Deep Latent Variable Models of Natural Language
- Workshop Program Committees: CORBON @ EACL 2017, CRAC @ NAACL 2018, NeuralGen @ NAACL 2019, DSNNLG @ INLG 2019, WNGT @ EMNLP 2019

Academic Internships

Facebook AI Research, New York, NY
Research Intern   Summer 2016, Summer 2017

- Research on retrieval-based text generation, with Marc'Aurelio Ranzato, Arthur Szlam, and Mike Lewis (Summer 2017)
- Research on training RNNs with target-propagation, with Sumit Chopra, Marc'Aurelio Ranzato, and Arthur Szlam (Summer 2016)