Elias Stengel-Eskin

Website: esteng.github.io
Email: esteng@cs.unc.edu
LinkedIn: elias-stengel-eskin
GitHub: github.com/esteng

Education

University of North Carolina, Chapel Hill — Chapel Hill, USA
Postdoctoral Research Associate – *Advisor: Mohit Bansal* — 2023 –current

Johns Hopkins University — Baltimore, USA
Ph.D. in Computer Science – *Advisor: Benjamin Van Durme* — 2018–2023

- Ph.D. thesis: "Modeling Meaning for Description and Interaction"
- supported by NSF Graduate Research Fellowship

Johns Hopkins University — Baltimore, USA
MSE in Computer Science – *Advisor: Benjamin Van Durme* — 2018–2021

McGill University — Montreal, Canada
Bachelor of Arts and Sciences in Cognitive Science — 2014–2018

- Minor: Linguistics
- First Class Honours
- Honours thesis: "Variational Bayesian Inference for Unsupervised Lexicon Discovery" – *Advisor: Timothy O'Donnell*

Experience

Microsoft Research — Montreal, Canada
PhD Research Intern – *Advisors: Marc-Alexandre Côté, Eric Yuan, Pierre-Yves Oudeyer* — March 2022-March 2023

Microsoft Research - Semantic Machines — Remote
PhD Research Intern – *Advisor: Yu Su* — Summer 2021

Montreal Computational and Quantitative Linguistics Lab — Montreal, Canada
Research Assistant – *Advisor: Morgan Sonderegger* — 2016-2018

Publications (peer-reviewed)

[1] **E. Stengel-Eskin**, A. Prasad, and M. Bansal, "ReGAL: Refactoring programs to discover generalizable abstractions", *The 41st International Conference on Machine Learning (ICML)*, 2024.

[2] D. Wan, J. Cho, **E. Stengel-Eskin**, and M. Bansal, "Contrastive region guidance: Improving grounding in vision-language models without training", *The 18th European Conference on Computer Vision (ECCV)*, 2024.

[3] J. C.-Y. Chen, S. Saha, **E. Stengel-Eskin**, and M. Bansal, "MAGDi: Structured distillation of multi-agent interaction graphs improves reasoning in smaller language models", *The 41st International Conference on Machine Learning (ICML)*, 2024.

[4] H. Fu, P. Sharma, **E. Stengel-Eskin**, G. Konidaris, N. L. Roux, M.-A. Côté, and X. Yuan, "Language-guided skill learning with temporal variational inference", *The 41st International Conference on Machine Learning (ICML)*, 2024.

[5] H. Wang*, A. Prasad*, E. **Stengel-Eskin***, and M. Bansal, "Soft self-consistency improves language model agents", *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*, 2024.

[6] M. Hashemzadeh, **E. Stengel-Eskin**, S. Chandar, and M.-A. Cote, "Sub-goal distillation: A method to improve small language agents", *The 3rd Conference on Lifelong Learning Agents (CoLLAs)*, 2024.

[7] **E. Stengel-Eskin**, K. Rawlins, and B. Van Durme, "Zero and few-shot semantic parsing with ambiguous inputs", in *The Twelfth International Conference on Learning Representations (ICLR)*, 2024.

[8] A. Prasad, **E. Stengel-Eskin**, and M. Bansal, "Rephrase, augment, reason: Visual grounding of questions for vision-language models", in *The Twelfth International Conference on Learning Representations (ICLR)*, 2024.

[9] **E. Stengel-Eskin** and B. Van Durme, "Did You Mean...? Confidence-based Trade-offs in Semantic Parsing", *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2023.

[10] **E. Stengel-Eskin** and B. Van Durme, "Calibrated Interpretation: Confidence Estimation in Semantic parsing", *Transactions of the Association for Computational Linguistics (TACL)*, 2023.

[11] **E. Stengel-Eskin**, J. Guallar-Blasco, Y. Zhou, and B. Van Durme, "Why Did the Chicken Cross the Road? Rephrasing and Analyzing Ambiguous Questions in VQA", *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.

[12] Z. Li, X. Wang, **E. Stengel-Eskin**, A. Kortylewski, W. Ma, B. Van Durme, and A. Yuille, "Super-CLEVR: A Virtual Benchmark to Diagnose Domain Robustness in Visual Reasoning", *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[13] **E. Stengel-Eskin** and B. Van Durme, "The Curious Case of Control", *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2022.

[14] **E. Stengel-Eskin**, E. A. Platanios, A. Pauls, S. Thomson, H. Fang, B. Van Durme, J. Eisner, and Y. Su, "When More Data Hurts: A Troubling Quirk in Developing Broad-Coverage Natural Language Understanding Systems", *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2022.

[15] C. Zhang, B. Van Durme, Z. Li, and **E. Stengel-Eskin**, "Visual Commonsense in Pretrained Unimodal and Multimodal Models", in *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, Seattle, Washington: Association for Computational Linguistics, 2022.

[16] **E. Stengel-Eskin**, A. Hundt, Z. He, A. Murali, N. Gopalan, M. Gombolay, and G. D. Hager, "Guiding Multi-Step Rearrangement Tasks with Natural Language Instructions", in *5th Annual Conference on Robot Learning (CoRL)*, 2021.

[17] Z. Li, **E. Stengel-Eskin**, Y. Zhang, C. Xie, Q. Tran, B. Van Durme, and A. Yuille, "Calibrating Concepts and Operations: Towards Symbolic Reasoning on Real Images", in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

[18] **E. Stengel-Eskin**, J. Guallar-Blasco, and B. Van Durme, "Human-Model Divergence in the Handling of Vagueness", in *Proceedings of the 1st Workshop on Understanding Implicit and Underspecified Language (UnImplicit)*, Online: Association for Computational Linguistics, 2021.

[19] **E. Stengel-Eskin**, K. Murray, S. Zhang, A. S. White, and B. Van Durme, "Joint Universal Syntactic and Semantic Parsing", *Transactions of the Association for Computational Linguistics (TACL)*, 2021.

[20] **E. Stengel-Eskin**, J. Guallar-Blasco, and B. Van Durme, "Exploring Human-Model Divergence Through Vagueness", *Proceedings of the Society for Computation in Linguistics (SCiL)*, 2021, ***Abstract**.

[21] R. Culkin, J. E. Hu, **E. Stengel-Eskin**, G. Qin, and B. V. Durme, "Iterative Paraphrastic Augmentation with Discriminative Span Alignment", *Transactions of the Association for Computational Linguistics (TACL)*, vol. 9, 2021.

[22] **E. Stengel-Eskin**, A. S. White, S. Zhang, and B. Van Durme, "Universal Decompositional Semantic Parsing", in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2020.

[23] A. S. White, **E. Stengel-Eskin**, S. Vashishtha, V. S. Govindarajan, D. A. Reisinger, T. Vieira, K. Sakaguchi, S. Zhang, F. Ferraro, R. Rudinger, *et al.*, "The Universal Decompositional Semantics Dataset and Decomp Toolkit", in *Proceedings of The 12th Language Resources and Evaluation Conference*, 2020.

[24] **E. Stengel-Eskin**, T.-R. Su, M. Post, and B. Van Durme, "A Discriminative Neural Model for Cross-Lingual Word Alignment", in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.

[25] M. McAuliffe, **E. Stengel-Eskin**, M. Socolof, and M. Sonderegger, "Polyglot and Speech Corpus Tools: A System for Representing, Integrating, and Querying Speech Corpora." In *INTERSPEECH*, 2017.

Publications (other)

1. P. Hase, T. Hofweber, X. Zhou, **E. Stengel-Eskin**, M. Bansal, "Fundamental Problems With Model Editing: How Should Rational Belief Revision Work in LLMs?" (2024), *Preprint*
2. T. Hofweber, P. Hase, **E. Stengel-Eskin**, M. Bansal, "Are language models rational? The case of coherence norms and belief revision" (2024), *Preprint*
3. **E. Stengel-Eskin**, P. Hase, M. Bansal, "LACIE: Listener-Aware Finetuning for Confidence Calibration in Large Language Models" (2024), *Under Review*
4. Z. Wang*, S. Yu*, **E. Stengel-Eskin***, J. Yoon, F. Cheng, G. Bertasius, M. Bansal, "VideoTree: Adaptive Tree-based Video Representation for LLM Reasoning on Long Videos" (2024), *Under Review*
5. J. Duan, R. Zhang, J. Diffenderfer, B. Kailkhura, L. Sun, **E. Stengel-Eskin**, M. Bansal, T. Chen, K. Xu, "GTBench: Uncovering the strategic reasoning limitations of LLMs via game-theoretic evaluations" (2024), *Under Review*
6. **E. Stengel-Eskin**, "Modeling Meaning for Description and Interaction", (2023) **Ph.D. Thesis**
7. J. Guallar-Blasco, **E. Stengel-Eskin**, B. Van Durme, "Analyzing Question Ambiguity in *Why*-Questions" (2023), *Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL 2023)* ***Abstract**
8. S. Vaidya, **E. Stengel-Eskin**, J. Sedoc, "Automatic Evaluation of Chit-chat via Semantic Parsing" (2022), *Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL 2022)* ***Abstract**
9. Y. Chen, S. Ebner, T. Chen, P. Xia, **E. Stengel-Eskin**, T. Su, J. E. Hu, N. Holzenberger, R. Culkin, C. Harman, M. Thomas, T. Lippincott, A. S. White, K. Rawlins, B. Van Durme, "NIST TAC SM-KBP 2019 System Description: JHU/UR Framework", (2019)
10. **E. Stengel-Eskin**, "Variational Bayesian Inference for Unsupervised Lexicon Discovery" (2017), **Undergraduate Honours Thesis**

News Coverage

1. *The importance of ambiguity*, Johns Hopkins University CS News, 04/15/2024
2. *Speaking easy: A simple method to clarify voice commands*, Johns Hopkins University CS News, 01/08/2024

Invited Talks

- **Confidence-based Rephrasing, Refinement, and Selection** March 22nd, 2024
  Keynote *– UncertaiNLP Workshop, EACL 2024*
- **Ambiguity in NLP** October 26th, 2023
  UNC Chapel Hill – Linguistics Department
- **Uncertainty and Ambiguity in Semantic Structures** October 6th, 2023
  Emory University
- **Language, Structure, and Beyond** May 9th, 2023
  UNC Chapel Hill – MURGe Group
- **Language, Structure, and Beyond** April 4th, 2023
  Scaled Cognition
- **Language, Structure, and Beyond** March 13th, 2023
  Brandeis University – CS Department
- **Language, Structure, and Beyond** March 2nd, 2023
  University of Cambridge – CS Department
- **Language, Structure, and Beyond** February 17th, 2023
  George Mason University – George Mason NLP Group
- **Language, Structure, and Beyond** February 1st, 2023
  Georgetown University – NERT Lab
- **Joint Universal Syntactic and Semantic Parsing** April 1st, 2022
  Cornell University – Workshop on Meaning in Language

Skills

- **Programming (expert):** Python
- **Programming (proficient):** Bash, Java, Javascript
- **Libraries/Frameworks:** PyTorch, AllenNLP, Transformers, NLTK, numpy, MechanicalTurk, networkx, MXNet, React

Languages

- **Native:** English, German
- **Fluent:** French
- **Other:** Latin (reading/translation), Spanish (intermediate)

Teaching

- **Teaching Assistant** at Johns Hopkins University Fall 2019
  Artificial Intelligence (EN.601.464/664)

Mentoring

- Jimena Guallar-Blasco Summer 2020-Present
  BS expected 2024
- Maryam Hashemzadeh Summer 2023
  Research Assistant, MILA
- Zhuohong (Zooey) He Winter 2021-Spring 2021
  MSE, Spring 2021 (Current: Intuitive)
- Chenyu (Heidi) Zhang Fall 2021-Fall 2022
  BS 2022, (Current: MS at Stanford)
- Shalaka Vaidya Fall 2022
  MS (NYU) 2023

- Yi Zhou — Winter 2022-Fall 2022
  MS 2022
- Vicky Zeng — Winter 2021-Winter 2022
  PhD student

Honors, Awards, and Funding

Funding

- Center for AI Safety Compute Grant (Co-PI) — 2024
- NSF Graduate Research Fellowship — 2018–2023

Honors and Awards

- First Class Honours in Cognitive Science — 2018
- Dean's Honor List (top 10% of faculty) — 2014–2015, 2016–2018
- Arts Undergraduate Research Internship Award — 2016

Graduate Courses

- Natural Language Processing, Neural Networks for NLP, Deep Learning, Applied Machine Learning, Computational Linguistics, Causal Inference, Vision as Bayesian Inference, Parallel Programming, Software Testing and Debugging, Deep Learning for Automated Discourse, Nonlinear Optimization, Human-Computer Interaction

Service

Reviewing

- **Area Chair:** ACL ARR (2024)
- **Primary Reviewer:** ACL (2021, 2022), NAACL 2022, EACL 2023, CoRL 2023, EMNLP 2023, NeurIPS 2024, WiNLP 2024
- **Secondary Reviewer:** NeurIPS 2020, ACL 2020, NAACL 2019, TACL
- **Program Committee**: UnImplicit 2022, SouthNLP 2024

Organizing

- **Workshop Organizer:** UnImplicit 2024
- **Conference Organizer:** EMNLP 2024 (Publicity Chair)

Committees

- **CLSP Visit Weekend Committee**, 2018-2022
- **CLSP Admissions Committee**, 2022-2023
- **Application Mentoring Group**, 2022-2023