Denis Emelin

D.Emelin@sms.ed.ac.uk | @emelin_denis | demelin.github.io | github.com/demelin

EDUCATION

Ph.D in Informatics   late 2021 / early 2022
University of Edinburgh   *Edinburgh, UK*
- Thesis: *Evaluating and Improving Natural Language Understanding in Neural Machine Translation*
- Advised by Dr. Rico Sennrich and Dr. Ivan Titov

M.Sc. in Language Science and Technology   2018
Saarland University   *Saarbrücken, Germany*
- Thesis: *Adversarially Learning to Manipulate the Cognitive Load of Sentences*
- Advised by Dr. Dietrich Klakow
- Developed IDGAN: A fully-unsupervised neural sentence simplification system that combines insights from neural machine translation, psycholinguistics, and adversarial learning
- Grade: 1.4 ($\sim$3.7 GPA)

B.A. in German Studies, Linguistics Concentration   2012
University of Tübingen   *Tübingen, Germany*
- Grade: 1.0 ($\sim$4.0 GPA)

PUBLICATIONS

Emelin, Denis, Ronan Le Bras, and Yejin Choi (2020). "Moral Stories: Situated Reasoning about Norms, Intents, Actions and their Consequences." *Pending submission to ACL 2021.* Non-archival version available at: demelin.github.io/media/moral_stories_draft.pdf.

Emelin, Denis, Ivan Titov, and Rico Sennrich (2020). "Detecting Word Sense Disambiguation Biases in Machine Translation for Model-Agnostic Adversarial Attacks." In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).*

Emelin, Denis, Ivan Titov, and Rico Sennrich (2019). "Widening the Representation Bottleneck in Neural Machine Translation with Lexical Shortcuts." In: *Proceedings of the Fourth Conference on Machine Translation (WMT).*

Alhafni, Bashar, Justin Cho, **Denis Emelin**, Mozhdeh Gheini, Jonathan May et al. (2019). "ELISA System Description for LoReHLT 2019.", In: *Proceedings of LoReHLT2019.*

Haddow, Barry, Nikolay Bogoychev, **Denis Emelin**, Kenneth Heafield, Rico Sennrich et al. (2018). "The University of Edinburgh's submissions to the WMT18 news translation task." In: *Proceedings of the Third Conference on Machine Translation (WMT).*

WORK EXPERIENCE

Research Intern   Jun 2020 - Sep 2020
Allen Institute for Artificial Intelligence (AI2)   *Seattle, WA, USA*
- Advised by Dr. Ronan Le Bras and Dr. Yejin Choi
- Explored value-aligned commonsense reasoning capabilities of natural language generation models
- Crowd-sourced a dataset of structured, annotated narratives through Amazon Mechanical Turk
- Developed novel decoding algorithms and performed extensive model analysis

Research Intern   Jun 2019 - Sep 2019
University of Southern California, Information Sciences Institute (ISI)   *Los Angeles, CA, USA*
- Advised by Dr. Jonathan May
- Evaluated the ability of neural models to translate figurative and non-compositional expressions
- Constructed a novel, wide-coverage idiom explicitation corpus from web data

Research Assistant
Nov 2013 - Apr 2014

German Research Center for Artificial Intelligence (DFKI)
Saarbrücken, Germany

- Investigated agent planning for plan verbalization in human-robot interaction
- Contributed to a natural language generation system by refining CCG grammars

Research Assistant
Dec 2012 - Sep 2013

Saarland University
Saarbrücken, Germany

- Gathered language processing data using methods such as EEG and eye tracking
- Analyzed experimental data for statistical significance

PROJECTS

Nematode | *Python, TensorFlow, numpy*

- Extendable neural machine translation toolkit built around the Transformer model
- Multi-GPU training and gradient aggregation enable large-scale experimentation
- Transformer implementation now merged into Nematus
- Available at: github.com/demelin/nematode

Paper re-implementations | *Python, PyTorch, Numpy*

- Working re-implementations of several natural language processing publications
- Noise Contrastive Estimation: github.com/demelin/Noise-Contrastive-Estimation-NCE-for-pyTorch
- Neural sentence similarity classifier: github.com/demelin/Sentence-similarity-classifier-for-pyTorch

TECHNICAL SKILLS

Python: PyTorch, TensorFlow, Numpy, Scikit-learn, Pandas, spaCy, NLTK, SciPy, Matplotlib / Seaborn
R: Limited proficiency focusing on statistical methods
Statistical analysis: Significance testing, effect size analysis
Data collection: Crowd-sourcing (e.g. via Amazon Mechanical Turk), web scraping, cleaning and curation

RESEARCH INTERESTS

Machine translation and multilingual technologies: Lexical ambiguity, co-reference resolution, low-resource NMT
Natural language generation: Narrative generation, constrained generation, decoding algorithms
Natural language understanding: Commonsense reasoning, social and moral reasoning, question answering
Data biases: Dataset artifacts, adversarial samples, sociological biases, fairness

TEACHING

Teaching Assistant, Tutor, Demonstrator
Apr 2018 - Feb 2020

University of Edinburgh
Edinburgh, UK

- Assisted with teaching undergraduate courses on neural machine translation and language understanding
- Created and evaluated coursework submitted by several hundred students

SERVICE

Reviewer: EMNLP, EACL, ACL Student Research Workshop
Mentor: ACL Student Research Workshop
Representation: Research student representative on the Informatics Staff-Student Liaison Committee,
research student representative for the Institute for Language, Cognition and Computation, University of Edinburgh

LANGUAGES

Native: Russian
Native-like: English, German
Novice: French

HOBBIES

Climbing, hiking, gardening, literature and philosophy