Aaron Chun Hei LO

✉ aaronlo326@gmail.com
🌐 https://aaronlolo326.github.io/

SUMMARY

- PhD graduate with over 6 years of ML and NLP experiences ranging from parsing, distributional semantics to information retrieval and natural language generation
- Demonstrated abilities in publishing papers in top conferences as well as developing practical applications
- Fast and mindful independent learner and team player

EDUCATION

Ph.D. (Research Area: Natural Language Processing)   Aug 2019–Jan 2024

The Chinese University of Hong Kong (CUHK)

Thesis: *Learning Semantics from Meaning Representations: From Distributional and Graph-Grammatical Perspectives*

B.Sc. (Hons) with First Class Honors in Computer Science   Sept 2015–July 2019

The Chinese University of Hong Kong

- **Dean's List**: Awarded for academic excellence in the Faculty of Engineering   2016, 2017, 2018, 2019
- **Master's List**: Awarded to the top student of each major of each class in Wu Yee Sun College 2017, 2018, 2019
- **ELITE Stream Student Scholarship**: To recognize excellence in advanced-level ELITE courses   2017, 2019
- **Computer Science Scholarship**: Awarded to top students admitted to the computer science program   2016

RESEARCH

Word to Function: Functional Distributional Semantics (FDS)   2021–2023

- Enhanced linguistic preciseness and computational efficiency of FDS for complex sentence structures
- Developed a variational autoencoder from scratch using PyTorch sped up with distributed data parallelism
- Outperformed >20 models on semantic tasks, including BERT that uses 12× more data and 2× more parameters
- Revealed that FDS captures hypernymy if trained on a corpus that follows the Distributional Inclusion Hypothesis (DIH) and proposed an alternative training objective that improves hypernymy detection from corpora[5]
- Published a paper at **SEM 2023*[4], and was invited to present at the 19th DELPH-IN Summit

Text Generation via Semantic Graph Parsing   2019–2021

- Demonstrated the use of a synchronous graph grammar in approximating the syntax–semantics interface of English
- Developed a probabilistic graph parser from scratch that reconstructs syntactic derivations from semantic graphs, with devised adaptations that improve accuracy, efficiency, and coverage of graph parsing
- Achieved better graph-to-text translation than a neural sequence-to-sequence method under out-of-domain settings, while additionally providing derivations for syntactic disambiguation
- Published in *ACL 2023*[3], and was invited to present at the Foundations of Language Processing of Umeå University

PROJECTS

English-to-Cantonese Machine Translation System   2024

- Developing an English-to-Cantonese machine translation system by leveraging rich Mandarin–English and Mandarin–Cantonese parallel corpora using PyTorch and Huggingface on TPUs sponsored by the TPU Research Cloud (TRC) program of Google

Task-Oriented Dialogue System with Unstructured Knowledge Access   2021

- Collaboratively developed a dialogue system with the CUHK team that performs retrieval-augmented generation (RAG) of natural language response based on dialogue contexts[2]
- Coordinated a sub-team that works on ranked retrieval from knowledge base using Huggingface and PyTorch
- Participated in the Ninth Dialog System Technology Challenge (DSTC9) and our team ranked 12 out of 24

CV–JD Recommendation System   2019

- Researched on automatic taxonomy induction using distributional and pattern-based approaches
- Wrote a web crawler to scrape over 40,000 publicly available CVs and 20,000 job descriptions (JDs)
- Devised a CV–JD matching algorithm using doc2vec and Latent Dirichlet Allocation (LDA)

INVITED TALKS

Functional Distributional Semantics (FDS) at Scale and *Probing Hypernymy in FDS*, 19th DELPH-IN Summit, Language and Information Society of University of A Coruña   27 June 2023

Semantic Composition with PSHRG for Derivation Tree Reconstruction from Graph-Based Meaning Representations, Seminar at Foundations of Language Processing of Umeå University, Virtual   16 Sept 2022

WORK EXPERIENCE	**Junior Research Assistant**	July 2019

<table>
<tr><td>WORK
EXPERIENCE</td><td>Junior Research Assistant
Department of Systems Engineering and Engineering Management, CUHK
Conducted collaborative research on cross-framework meaning representations (MRs) parsing
Developed a transition-based text-to-MRs parser
Published a paper in CoNLL 2019 Shared Task[1]</td><td>July 2019</td></tr>
</table>

WORK EXPERIENCE

Junior Research Assistant   July 2019
Department of Systems Engineering and Engineering Management, CUHK
- Conducted collaborative research on cross-framework meaning representations (MRs) parsing
- Developed a transition-based text-to-MRs parser
- Published a paper in CoNLL 2019 Shared Task[1]

Software Engineer Intern   June 2018–Aug 2018
Set Sail Software
- Collaborated with digital marketing agencies in developing chatbots for multiple clients tailored to their needs
- Developed backends of chatbots using Node.js and Firebase Cloud Functions
- Created tools for automating performance analyses of chatbots

ADDITIONAL EXPERIENCE

Teaching Assistant, Faculty of Engineering, CUHK
- CSCI2100: Data Structures   2020–2023
- SEEM3550: Fundamentals in Information Systems   2021–2023

Resident Tutor, Wen Lin Tang, Chung Chi College, CUHK   2021–2023
- Provided pastoral care to undergraduate residents and served as an intermediary between them and the warden

CORE SKILLS

Natural Languages Cantonese (*native*), English (*proficient*), Mandarin (*proficient*)

Programming Languages Python, C, SQL

ML Libraries PyTorch, Tensorflow, scikit-learn

NLP Libraries PyDelphin, NLTK, Spacy, Gensim, WordNet, Hugging Face

Cloud Platform Google Cloud

PUBLICATIONS AND PREPRINTS

[5] Chun Hei Lo and Guy Emerson. 2023. Distributional Inclusion Hypothesis and Quantifications: Probing Hypernymy in Functional Distributional Semantics. arXiv:2309.08325

[4] Chun Hei Lo, Hong Cheng, Wai Lam, and Guy Emerson. 2023. Functional Distributional Semantics at Scale. In *Proceedings of the 12th Joint Conference on Lexical and Computational Semantics (*SEM 2023)*, pages 423–436, Toronto, Canada

[3] Chun Hei Lo, Wai Lam, and Hong Cheng. 2022. Semantic Composition with PSHRG for Derivation Tree Reconstruction from Graph-Based Meaning Representations. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5425–5439, Dublin, Ireland

[2] Mudit Chaudhary, Borislav Dzodzo, Sida Huang, Chun Hei Lo, Mingzhi Lyu, Lun Yiu Nie, Jinbo Xing, Tianhua Zhang, Xiaoying Zhang, Jingyan Zhou, Hong Cheng, Wai Lam, and Helen Meng. 2021. Unstructured Knowledge Access in Task-oriented Dialog Modeling using Language Inference, Knowledge Retrieval and Knowledge-Integrative Response Generation. arXiv:2101.06066

[1] Sunny Lai, Chun Hei Lo, Kwong Sak Leung, and Yee Leung. 2019. CUHK at MRP 2019: Transition-Based Parser with Cross-Framework Variable-Arity Resolve Action. In *Proceedings of the Shared Task on Cross-Framework Meaning Representation Parsing at the 2019 Conference on Natural Language Learning*, pages 104–113, Hong Kong