

YU CAO

Computational Linguist

📍 New Brunswick, NJ

@ yc825@linguistics.rutgers.edu

🐙 pterosdiacos.github.io

in www.linkedin.com/in/yu-cao-pdiacos

EXPERIENCES

Ontology Linguist

Facebook via PRO Unlimited

📅 08/2019 – 10/2019

📍 Redmond, WA

- Worked at Assistant Cross-Functional Group.
- Prototyped models for cross-domain semantic parsing for question answering, using *PyTorch*.
- Automated inferences and sub-ontology extraction in OWL 2 ontology knowledge graphs, using *Python*.

Teaching Assistant

Rutgers University

📅 09/2018 – 08/2020

📍 New Brunswick, NJ

- Courses taught: *Intro to Linguistic Theory*, *Invented Languages*.

Research Assistant

The Chinese University of Hong Kong

📅 12/2014 – 07/2015

📍 Hong Kong, China

- Worked at Center for Sign Linguistics and Deaf Studies.
- Provided linguistic expertise and assisted in developing a Chinese grammatical knowledge assessment software for deaf/hard of hearing children.

EDUCATION

Rutgers University

Ph.D. in Linguistics

📅 09/2016 – Present

📍 New Brunswick, NJ

The Chinese University of Hong Kong

M.A. in Chinese Linguistics & Language Acquisition

📅 09/2013 – 06/2014

📍 Hong Kong, China

Sichuan Normal University

B.A. in Teaching Chinese as a Second Language

📅 09/2009 – 06/2013

📍 Chengdu, China

PUBLICATIONS

- [Automated fact-value distinction in court opinions](#)
European Journal of Law and Economics 2020.
👤 Yu Cao, Daniel Chen & Elliot Ash.
Automated classification of fact vs. value statements in written judicial decisions.
- [Investigating BERT's knowledge of language: Five analysis methods with NPIs](#)
Proceedings of EMNLP 2019.
👤 Alex W., Yu Cao & 14 others (equal contribution).
Supervisor: Sam Bowman. Team role: constructed and conducted the main transfer learning experiments.

PROFESSIONAL PROFILE

Linguistics

specialized in computational linguistics, semantics, syntax, discourse analysis. Strong critical-thinking and problem-solving skills.

Computation

10+ years of experiences in algorithms and data structure. The First Prize in National Olympiad in Informatics in Provinces 2007, China.

PROJECTS

Graph representation of meaning

([dissertation research](#) 📄 [repo](#))

- Designed a graph based semantic representation that encodes natural language quantification and plurality.
- Developed a CCG-based semantic parser constructing graphs from texts using *Python*.

Incremental topological sorting

([LingBuzz post](#))

- Designed algorithms for topological sorting under uncertain conditions.
- Improved on resorting time complexity with dynamic updating strategies.

SKILLS

Technical

- Machine learning algorithms, NLP/NLU
- Python, C, Scikit-learn, PyTorch, Tensorflow
- Linguistic analysis, symbolic logic, statistics

ML/NLP Coursework

- [Natural language processing](#) 📄
📅 09/2020 🏛️ HSE@Coursera
- [Bayesian methods for machine learning](#) 📄
📅 08/2020 🏛️ HSE@Coursera
- [Introduction to deep learning](#) 📄
📅 07/2020 🏛️ HSE@Coursera
- Machine learning
📅 07/2020 🏛️ Stanford@Coursera
- Seminar: Linguistic knowledge in reusable sentence encoders
📅 05/2019 🏛️ NYU
- Graph formalisms for meaning representations
📅 07/2018 🏛️ NASSLLI@CMU
- NLU & computational semantics
📅 05/2018 🏛️ NYU

Languages

ENG (fluent) CHN (native) JPN (JLPT N1)