

Xusheng Luo

CONTACT INFORMATION	SEIEE Bldg. 3-341 Department of Computer Science & Engineering Shanghai Jiao Tong University 800 Dongchuan Road Shanghai 200240, China	Phone: (+86) 185-1615-6174 E-mail: freefish_6174@sjtu.edu.cn URL: http://xushengl.me WeChat: Freefish_6174
RESEARCH INTERESTS	Natural Language Processing, Text Mining, Knowledge Graph Application, Machine Learning, Deep Learning, Artificial Intelligence	
EDUCATION	Shanghai Jiao Tong University , Shanghai, China M.S. Candidate, Computer Science, September 2015 (expected graduation date: March 2018) <ul style="list-style-type: none">• GPA: 2.81/3.3, Rank: Top 10%• Advisor: Prof. Kenny Q. Zhu• Research Field: Natural Language Processing, Text Mining, Knowledge Graph, Deep Learning• Thesis Topic: “<i>Text and Web Table Entity Linking</i>” B.Eng., Computer Science (IEEE Honored Class), June 2015 <ul style="list-style-type: none">• An elite program at SJTU which aims to nurture scientists in computer science, electrical and electronic technology, and information science based on MITs educational model.• GPA: 88.4/100, Rank: Top 10%• Thesis Topic: “<i>Paraphrasing of Predicates for Question Analysis in QA System</i>”	
HONORS AND AWARDS	“ Meritorious Winner ” of Interdisciplinary Contest in Modeling First Class Academic Scholarship of SJTU “ Merit Student ” of SJTU	April 2014 2011 - 2016 2012 - 2013
ACADEMIC EXPERIENCE	Shanghai Jiao Tong University , Shanghai, China <i>Graduate Student and Research Assistant at ADAPT Lab</i> September, 2014 - present http://adapt.seiee.sjtu.edu.cn/ Working at ADAPT Lab, led by Prof. Kenny Q. Zhu, collaborate with PhD candidate Kangqi Luo, work includes current Master research, coursework and developing projects. <i>Teaching Assistant</i> Various duties on undergraduate level course for Computer Science. Shared responsibility for tutorials, exams, homework assignments, and course project. Give one lecture each week for homework discussion. This is an English-only class. <ul style="list-style-type: none">• CS383 Programming Languages. 2015/16 Fall.• CS383 Programming Languages. 2016/17 Fall. Third Research Institute of Ministry of Public Security , Shanghai, China <i>Research Intern</i> July, 2014 - August, 2014 Worked in the algorithm team. Implement accurate detection and recognition of vehicles and human beings from real-time traffic videos. Won the first prize in the image detection contest of the institute.	
INDUSTRY EXPERIENCE	taobao.com @Alibaba Group , Hangzhou, China <i>Algorithm Engineer Intern</i> June, 2017 - August, 2017 Join project “Intelligent Search” for taobao.com, including query understanding, dialogue management and shopping intent question answering.	

Help construct taobao's Knowledge Graph. Use deep learning models to do semantic understanding for user query, including sequence labeling, entity linking and intention identification.

Intel Asia-Pacific Research & Development Ltd., Shanghai, China

Software Development Intern

Dec, 2014 - May, 2015

Independently developed a static multi-threads lock check tool for Intel binary translation software. The tool helped detect several lock issues for development team with good performance.

RESEARCH
PROJECTS

Understanding Natural Language Search Query in E-Commerce

As the AI technologies develop rapidly, the services provided by e-commerce companies become more and more intelligent. One inevitable tendency, different from earlier online shopping experiences, is that customers will be able to use natural language instead of key words when searching for the products they want to buy. The very first step for search engine to understand user query is to identify the query intention. There are several sub problems such as identifying central intention for query with multiple intention words, sequence labeling for query which is basically Named Entity Recognition, and entity linking which disambiguates mentions in query with knowledge bases. We attempt to solve these problem by deep neural networks. We plan to submit papers to ICSC 2018, IJCAI 2018 or ACL 2018.

Commonsense Reasoning using Word Embedding in Causality Space

Commonsense causality is the causality between actions or events that is acknowledged by human beings. Commonsense is one thing that shared by nearly all people. It is embraced in large amount of text corpus but hard to mine out because of its sparsity. Commonsense causal reasoning aims to find out the possible causality between events, measuring whether one event can lead to another. We are the first to generate the vector representations of cause/effect role for each word and those vectors can be used to not only calculate the word-level scores directly but also use for further sentence-level reasoning. For a word, we believe that it plays a cause(effect) role when it appears in a causal(effect) span in one cause-effect pair. We submitted our preliminary work to ICSC 2018.

Linking Entities in Web Tables across Languages

The very first step for machines to understand and process Web tables is to link the surface mentions of the entities in the tables to a standard lexicon or knowledge base, such as Wikipedia, which uniquely identifies entities. Existing work has focused on entity linking of Web tables in English, and we call it mono-lingual table linking. However, when we face a non-English Web table, non-English knowledge bases are often not comprehensive enough to cover all the entity mentions in the tables at hand. Thus we attempt to solve the cross-lingual table linking problem. We proposed a joint deep neural network model with some specially designed features to tackle this problem, and a translation module is used to bridge the language gap. Our paper "*Cross-Lingual Entity Linking for Web Tables*" was accepted by AAAI 2018, a top conference in AI field.

Represent Natural Language Relations with Knowledge Base Structures

This work studies the problem of discovering the structured knowledge representation of binary natural language relations. The representation, known as the schema, generalizes the traditional path of predicates to support more complex semantics. We present a search algorithm to generate schemas over a knowledge base, and propose a data-driven learning approach to discover the most suitable representations to one relation. Evaluation results show that inferred schemas are able to represent precise semantics, and can be used to enrich manually crafted knowledge bases. Our paper "*A Data-Driven Approach to Infer Knowledge Base Representation for Natural Language Relations*" was accepted by IJCAI 2017, a top conference in AI field.

Infer the Type of Relation Arguments for Open Information Extraction

This work presents a framework to model the semantic representation of binary relations produced

by open information extraction systems. For each binary relation, we infer a set of preferred types on the two arguments simultaneously, and generate a ranked list of type pairs which we call schemas. All inferred types are drawn from the Freebase type taxonomy, which are human readable. Our system collects 171,168 binary relations from ReVerb, and is able to produce top-ranking relation schemas with a mean reciprocal rank of 0.337. Our paper “*Inferring Binary Relation Schemas for Open Information Extraction*” was accepted by EMNLP 2015, a top conference in NLP field.

PUBLICATIONS

Automatic Generation of Chinese Short Product Titles for Mobile Display.

Xusheng Luo*, Yu Gong*, Wenwu Ou, and Kenny Q. Zhu

Under review of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI 2018)

Identify Central Intention of Natural Language Search Query In E-Commerce.

Xusheng Luo

In The 12th IEEE International Conference on Semantic Computing (ICSC 2018 Poster)

Word Embedding for Commonsense Causal Reasoning.

Yuchen Sha, **Xusheng Luo**, Kenny Q. Zhu.

In The 12th IEEE International Conference on Semantic Computing (ICSC 2018 Poster)

Cross-lingual Entity Linking for Web Tables.

Xusheng Luo, Kangqi Luo, Xianyang Chen, Kenny Q. Zhu.

In Proceedings of the Thirty-second AAAI Conference on Artificial Intelligence (AAAI 2018)

A Data-Driven Approach to Infer Knowledge Base Representation for Natural Language Relations.

Kangqi Luo, **Xusheng Luo**, Xianyang Chen, Kenny Q. Zhu.

In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI 2017)

Inferring Binary Relation Schemas for Open Information Extraction.

Kangqi Luo, **Xusheng Luo**, Kenny Q. Zhu.

In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)

Linking Entities in Web Tables across Languages

Xusheng Luo

Master's Thesis, 2018

Paraphrasing of predicates for question analysis in QA system

Xusheng Luo

Bachelor's Thesis, 2015

COMPUTER SKILLS

- Machine Learning and Deep Learning Packages: Scipy, Numpy, scikit-learn, TensorFlow, Keras
- Natural Language Processing Tools: Stanford CoreNLP, NLTK
- Distributed Platforms: Hadoop, Spark
- Databases: MySQL, MongoDB
- Languages: Python, Java, C/C++, PHP, HTML5, some use of Linux shell scripts, CUDA/MPI parallel processing library
- Applications: GNUPlot, L^AT_EX, common spreadsheet and presentation software
- Operating Systems: Unix/Linux, Windows