

# SHANGMIN (SHAWN) GUO

Research Assistant, Inst. for Automated Language Teaching and Assessment, University of Cambridge

@ sg955@cam.ac.uk    +44-07523075684    shawnguo.cn    github.com/Shawn-Guo-CN  
A33, 7 Charles Babbage Road, Cambridge, United Kingdom, CB3 0FT

## PUBLICATIONS

- The Emergence of Compositional Languages for Numeric Concepts Through Iterated Learning in Neural Agents  
Shangmin Guo, Yi Ren, Serhii Havrylov, Stella Frank, Ivan Titov & Kenny Smith  
3rd NeurIPS Workshop on Emergent Communication
- Enhance the Compositionality of Emergent Language by Iterated Learning  
Yi Ren, Shangmin Guo, Matthieu Labeau, Shay B. Cohen & Simon Kirby  
3rd NeurIPS Workshop on Emergent Communication
- IJCNLP-2017 Task 5: Multi-choice Question Answering in Examinations  
Shangmin Guo, Kang Liu, Shizhu He, Zhuoyu Wei, Cao Liu & Jun Zhao  
IJCNLP-2017
- Which is the Effective Way for Gaokao: Information Retrieval or Neural Networks?  
Shangmin Guo, Xiangrong Zeng, Shizhu He, Kang Liu & Jun Zhao  
EACL-2017
- Employing External Rich Knowledge for Machine Comprehension.  
Bingning Wang, Shangmin Guo, Kang Liu & Jun Zhao  
IJCAI-2016

## PROJECTS

### SpikeInterface: A Unified Framework for Spike Sorting

📅 Sep. 2019 –    📍 Edinburgh, UK

- Perform unit testing and integrated testing, and fix the detected bugs.
- Proposed a Bayesian algorithm to combine sorting results from different spike sorters.
- This is an open-source project which can be accessed  
<https://github.com/SpikeInterface>.

### Emergence of Compositional Languages for Numeric Concepts in Multi-agent Autonomous Communication

📅 Jan. 2019 – Sep. 2019    📍 Edinburgh, UK

- This is my master dissertation project, and can be accessed at  
<https://arxiv.org/pdf/1911.01098.pdf>.
- Proposed this project to explore the emergence of compositional languages for numeric concepts in multi-agent communication protocols.
- Designed and implemented multi-agent population models based on deep learning, as well as two different language games.
- Found that iterated learning could facilitate the emergence of compositional language, and their emergence are heavily influenced by input representations.
- Found that emergent languages from different language games have different expressivity.
- This work has been submitted to EvoLang-2020 and ICLR-2020.

### Automatic Medical Record Generation based on Doctor-Patient Dialogue

📅 Mar. 2018 – Aug. 2018    📍 Beijing, China

- Designed the schema of a new medical knowledge base and implemented an information gain algorithm to do diagnosis on it.
- Led the development of a disease tag prediction system based on the utterances of doctor-patient dialog.

## WORKING EXPERIENCE

### Research Assistant

Institute for Automated Language Teaching and Assessment, University of Cambridge

📅 Nov. 2019 –    📍 Cambridge, UK

### Research Assistant

Institute for Adaptive and Neural Computation, School of Informatics

📅 Sep. 2019 – Nov. 2019    📍 Edinburgh, UK

### Research Engineer

National Laboratory of Pattern Recognition

📅 Oct. 2015 – Aug. 2018    📍 Beijing, China

## EDUCATION

### M.Sc. (Distinction) in Data Science

Avg Score: A

The University of Edinburgh

📅 Sep. 2018 – Aug. 2019    📍 Edinburgh, UK

### B.Eng. in Computer Science

Rank: 1<sup>st</sup> / 66 | GPA: 3.88 / 4.0 | Avg Score: 87.20 / 100

National University of Defense Technology

📅 Sep. 2010 – Jun. 2014    📍 Changsha, China

## RESEARCH INTERESTS

Evolutionary Linguistics

Grounded Language Learning

Natural Language Understanding

## PROGRAMMING SKILLS

Python/PyTorch/L<sup>A</sup>T<sub>E</sub>X  
C++/Java/Django/Qt5  
Distributed Computing  
CUDA/R



## HIGHLIGHT COURSES

Probabilistic Modeling and Reasoning

Computational Cognitive Science

Reinforcement Learning

Algorithmic Game Theory

Question Answering System for Gaokao History Multiple Choice Questions

Oct. 2015 – Oct. 2017 Beijing, China

- Designed and developed a novel deep neural network model that is combined with IR method, which gained best performance on real Gaokao questions in subject History.
- Implemented various kinds of baseline models including memory networks and dynamic memory networks.
- This work has been published on EACL-2017.

Employing External Rich Knowledge for Machine Comprehension

Oct. 2015 – Feb. 2016 Beijing, China

- Developed a RTE model that is trained on rich external corpus and then be applied to evaluate the confidence of a specific answer candidate in machine comprehension.
- This work has been published on IJCAI-2016.

Research and Analysis of Behavior Modeling in Simulation System

Oct. 2013 – Jun. 2014 Changsha, China

- This is my diploma project, the overall objective is to build a distributed interactive simulation system. I took the responsibility of research and analysis of the behavior modeling.
- Concluded the methods of behavior modeling and abstracted them with first-order sequential logic.
- Built up several behavior models in different scenarios and checked feasibility of these models.

REFEREES

Prof. Kenny Smith  
kenny.smith@ed.ac.uk

Centre for Language Evolution, Linguistics and English Language, University of Edinburgh

Dr. Ivan Titov  
ititov@inf.ed.ac.uk

Institute for Language, Cognition and Computation, School of Informatics, University of Edinburgh

Prof. Rico Sennrich  
sennrich@cl.uzh.ch

Institut für Computerlinguistik, Universität Zürich

Prof. Stuart Anderson  
soa@staffmail.ed.ac.uk

Schools of Informatics, University of Edinburgh

Prof. Jun Zhao  
jzhao@nlpr.ia.ac.cn

Institute of Automation, Chinese Academy of Sciences

LIFE PHILOSOPHY

“Truth can be known but it may not be the well-known truth.”

HOBBIES

Paintball  
Archery  
Racing Car

