

Emergence of Numerals in Multi-Agent Autonomous Communication System

Shangmin Guo

Master of Science
School of Informatics
University of Edinburgh
2019

Abstract

This project aims to propose a new computational simulation method for the emergence of numerals based on multi-agent autonomous communication system following deep reinforcement learning methodology.

Acknowledgements

Any acknowledgements go here.

Table of Contents

1	Introduction	1
1.1	Evolutionary Linguistics	1
1.2	Deep Reinforcement Learning and Its Application in Grounded Language Learning	2
2	Background	3
2.1	Computer Simulation Methods in Evolutionary Linguistics	3
2.2	Multi-agent Games in Grounded Language Learning	3
3	Set Generation Game and Models	4
3.1	Game Description	4
3.1.1	Game Procedure	4
3.1.2	Numerals in the Game	4
3.2	Proposed Models	4
3.2.1	Set2Seq2Seq Models	4
3.2.2	Numeral Iterated Learning	4
3.2.3	Baseline Models	4
4	Experiment Results and Analysis	5
5	Conclusions	6
5.1	Final Reminder	6
5.2	Further Discussion	6
	Bibliography	7

Chapter 1

Introduction

Natural language processing (NLP) is an important and long-standing topic in artificial intelligence (AI), in which a core question is natural language understanding (NLU). With the rapid development of deep learning (DL), most current state-of-the-art methods in NLP, e.g. [Socher et al., 2013, Mikolov et al., 2013, Kim, 2014], are based on DL models trained on massive static textual corpora. From an information processing perspective, I illustrate the global view of NLP-oriented human-computer interaction system in Figure given as follow. As we can see in the diagram, the input of NLP systems are various kinds of textual materials generated by human beings to describe their experiences or perceptions. Under such a perspective, symbols in natural languages are actually representations of features of the original experiences or perceptions, whereas most NLP systems directly take these symbols as original features.

Therefore, grounded language learning (GLL) argues that models need a grounded environment to learn and understand language [Matuszek, 2018]. However, natural languages of the time have been developed for at least tens of thousands of years [Berwick and Chomsky, 2016] and already became very sophisticated. Thus, to verify that computational agents can truly understand and complete the tasks specified by natural languages, it is necessary to facilitate them to develop various kinds of characteristics of natural language during autonomous communication.

1.1 Evolutionary Linguistics

The emergence and evolution of natural language have always been critical questions to

Such question is a critical question in evolutionary linguistics.

However, previous works have to pre-define the basic elements of language.

1.2 Deep Reinforcement Learning and Its Application in Grounded Language Learning

With the recent development of DRL, we do not need to pre-define any linguistic element and thus can simulate the emergence of preliminary linguistic phenomenon.

With these promising progresses in GLL and characteristics of numerals, this project proposes a new simulation methods of the emergence of numeral systems and also methods to analyse them.

Chapter 2

Background

A dissertation usually contains several chapters.

- 2.1 Computer Simulation Methods in Evolutionary Linguistics**
- 2.2 Multi-agent Games in Grounded Language Learning**

Chapter 3

Set Generation Game and Models

3.1 Game Description

One hypothesis of our work is that, the linguistic hypotheses can be implied by game dynamics.

3.1.1 Game Procedure

3.1.2 Numerals in the Game

3.2 Proposed Models

3.2.1 Set2Seq2Seq Models

3.2.2 Numeral Iterated Learning

3.2.3 Baseline Models

Chapter 4

Experiment Results and Analysis

Chapter 5

Conclusions

5.1 Final Reminder

The body of your dissertation, before the references and any appendices, *must* finish by page 40. The introduction, after preliminary material, should have started on page 1.

You may not change the dissertation format (e.g., reduce the font size, change the margins, or reduce the line spacing from the default 1.5 spacing). Over length or incorrectly-formatted dissertations will not be accepted and you would have to modify your dissertation and resubmit. You cannot assume we will check your submission before the final deadline and if it requires resubmission after the deadline to conform to the page and style requirements you will be subject to the usual late penalties based on your final submission time.

5.2 Further Discussion

Bibliography

- [Berwick and Chomsky, 2016] Berwick, R. C. and Chomsky, N. (2016). *Why only us: Language and evolution*. MIT press.
- [Kim, 2014] Kim, Y. (2014). Convolutional neural networks for sentence classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*.
- [Matuszek, 2018] Matuszek, C. (2018). Grounded language learning: Where robotics and nlp meet. In *IJCAI*, pages 5687–5691.
- [Mikolov et al., 2013] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 26*, pages 3111–3119. Curran Associates, Inc.
- [Socher et al., 2013] Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A., and Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1631–1642, Seattle, Washington, USA. Association for Computational Linguistics.

Appendices