Language Learning and Processing in People and Machines

Aida Nematzadeh
DeepMind
nematzadeh@google.com

Richard Futrell
UC Irvine
rfutrell@uci.edu

Roger Levy
MIT
rplevy@mit.edu

1 A Brief Description

The ambitious goal of computational linguistics (CL) is to develop systems that process, understand, and produce natural languages. To achieve this goal, most work in CL has focused on developing models for different linguistic tasks such as semantic role labeling and natural language inference. However, recent research in CL has started investigating the missing ingredients required to move towards building systems with general linguistic intelligence. For example, one area of focus is multitask learning – building models that perform well on a number of linguistic tasks (*e.g.*, Devlin et al., 2018). Other research has investigated the importance of introducing commonsense into natural language processing models (*e.g.*, Rashkin et al., 2018). Despite the recent advances in the field, we are still far from systems that exhibit human-level linguistic intelligence: great performance on a set of predefined linguistics tasks does not result in systems that can understand and produce natural language in general settings.

An alternative research direction is to build systems that mimic language acquisition and processing as it is performed by humans. Such a system might achieve the linguistic efficacy required for understanding and producing human languages. But we first need to understand how children so effortlessly learn their language. A line of research aims to reverse-engineer child language acquisition: the idea is to shed light on the cognitive processes that might be responsible for language acquisition; we can in turn use the learned lessons in designing computational (cognitive) models that exhibit human-like language performance. Understanding language acquisition is also beneficial to natural language processing (NLP) applications: we can explore how the mechanisms (such as attention) and inductive biases that facilitate human learning can be explicitly incorporated into our algorithms. Moreover, we can evaluate our NLP systems with respect to human behavior which helps us understand the limitations of these systems.

The goal of this tutorial is to bring the fields of computational linguistics and computational cognitive science closer: we will introduce different stages of language acquisition and their parallel problems in NLP. As an example, one of the early challenges children face is mapping the meaning of word labels (such as "cat") to their referents (the furry animal in the living room). Word learning is similar to the word alignment problem in machine translation. We explain the current computational models of human language acquisition, their limitations, and how the insights from these models can be incorporated into NLP applications. Moreover, we discuss how we can take advantage of the cognitive science of language in computational linguistics: for example, by designing cognitively-motivated evaluation tasks or building language-learning inductive biases into our models.

We believe now is a great time for this tutorial. Using end-to-end and deep neural approaches has resulted in significant improvements in various NLP tasks in the past years. But in 2018, we observed a shift in the field from building models to creating datasets; this mainly happened because given the current compute power and access to vast amount of data, the existing NLP tasks were not challenging enough for our models. Revisiting challenges in language acquisition will spark interest in the community in two ways: Some will be inspired to design more challenging problems, and others may work on developing models of language acquisition.

Proceedings of NAACL HLT 2019: Tutorial Abstracts, pages 19–21
Minneapolis, Minnesota, June 2, 2019.

2 Type of the Tutorial

The tutorial is mostly introductory—we will explain the literature on computational models of language acquisition and processing. However, we will also introduce some of the recent research directions in this domain.

3 Outline of the Tutorial

- Introduction: the goal of the tutorial (5 minutes)
- Language acquisition (60 minutes)
  - Views/debates on language acquisition
  - The role of categorization, memory, and attention in language acquisition
  - Segmenting speech to words
  - Learning the meaning of words
  - Unraveling the structure of the words
  - Developing theory of mind
  - Understanding the pragmatics
- Language processing (60 minutes)
  - Methods and sources of data on human language processing
  - Expectation-based syntactic processing
  - Effects of working memory and computational models of them
  - Can RNNs explain patterns of human language processing?
  - Can incremental parsers explain patterns of human language processing?
- Cognitively-informed NLP (25 minutes)
  - Evaluating language models using psycholinguistic tests
  - Evaluating sequence-to-sequence models
  - Evaluating semantic representations and vector spaces
  - Evaluating question-answering models
- Language evolution (25 minutes)
  - Emergence of linguistic symbols
  - From symbols to linguistic structure
  - Recent agent-based modeling results
- Conclusion: main take-aways and future research (5 minutes)

4 The Breadth of the Tutorial

We will cover a broad range of topics in the computational cognitive science of language: the tutorial will mostly cover work by other researchers. The presenters will discuss their research when it is the most relevant work on the topic; this would cover less than 30% of the tutorial.

5 Prerequisites

To fully take advantage of the modeling part of the tutorial, the attendees need to have introductory-level knowledge of statistics, probability theory, and machine learning.

6 Instructors

- Aida Nematzadeh, DeepMind.
  - Email: nematzadeh@google.com
  - Website: http://aidanematzadeh.me
  - Research interests: I draw on the intersection of machine learning, cognitive science, and computational linguistics to investigate how humans learn language and use the insights to improve artificial intelligence systems. My recent work has focused on statistical approaches to semantic representations and theory-of-mind reasoning. During my PhD, I used computational modeling to study how children learn, represent, and search for semantic information.
- Richard Futrell, UC Irvine.
  - Email: rfutrell@uci.edu
  - Website: http://socsci.uci.edu/~rfutrell
  - Research interest: I study language processing in humans and machines. My hypothesis is that the distinctive properties of natural language, including its syntactic structure, can be explained in terms of efficient communication given human cognitive constraints. I explore this hypothesis in large-scale corpus studies, behavioral experiments, and cognitive modeling work using information theory and neural networks.
- Roger Levy, MIT.
  - Email: rplevy@mit.edu

- Website: http://www.mit.edu/~rplevy/
- Research interest: My research focuses on theoretical and applied questions in the processing and acquisition of natural language. Linguistic communication involves the resolution of uncertainty over a potentially unbounded set of possible signals and meanings. How can a fixed set of knowledge and resources be deployed to manage this uncertainty? And how is this knowledge acquired? To address these questions I combine computational modeling, psycholinguistic experimentation, and analysis of large naturalistic language datasets. This work furthers our understanding of the cognitive underpinning of language processing and acquisition, and helps us design models and algorithms that will allow machines to process human language.

7 The Audience Size

We expect an audience size of around 100. Roger gave a similar tutorial a few years ago which was attended by around 50 people. However, this was before the rapid growth of the ACL conferences.

8 Special Requirements

We require Internet access in the tutorial room.

9 Venue

We strongly prefer NAACL for logistical reasons; if NAACL is not possible then we would be open to EMNLP instead. ACL is not possible due to overlap with the Annual Conference of the Cognitive Science Society this year.

References

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding.

Hannah Rashkin, Maarten Sap, Emily Allaway, Noah A. Smith, and Yejin Choi. 2018. Event2mind: Commonsense inference on events, intents, and reactions.