

Integer Linear Programming formulations in Natural Language Processing

An EACL 2017 Tutorial

Dan Roth

University of Illinois, Urbana-Champaign
danr@illinois.edu

Vivek Srikumar

University of Utah
svivek@cs.utah.edu

1 Introduction

Making decisions in natural language processing problems often involves assigning values to sets of interdependent variables where the expressive dependency structure can influence, or even dictate what assignments are possible. This setting includes a broad range of structured prediction problems such as semantic role labeling, named entity and relation recognition, co-reference resolution, dependency parsing and semantic parsing. The setting is also appropriate for cases that may require making global decisions that involve multiple components, possibly pre-designed or pre-learned, as in event recognition and analysis, summarization, paraphrasing, textual entailment and question answering. In all these cases, it is natural to formulate the decision problem as a constrained optimization problem, with an objective function that is composed of learned models, subject to domain or problem specific constraints.

Over the last few years, starting with a couple of papers written by (Roth and Yih, 2004, 2005), dozens of papers have been using the Integer linear programming (ILP) formulation developed there, including several award-winning papers, including (Martins et al., 2009; Koo et al., 2010; Berant et al., 2011)

This tutorial will present the key ingredients of ILP formulations of natural language processing problems, aiming at guiding readers through the key modeling steps, explaining the learning and inference paradigms and exemplifying these by providing examples from the literature. We will cover a range of topics, from the theoretical foundations of learning and inference with ILP models, to practical modeling guidance, to software packages and applications.

The goal of this tutorial is to introduce the computational framework to the broader ACL commu-

nity, motivate it as a generic framework for learning and inference in global NLP decision problems, present some of the key theoretical and practical issues involved and survey some of the existing applications of it as a way to promote further development of the framework and additional applications. We will also make connections with some of the “hot” topics in current NLP research and show how they can be used within the general framework proposed here. The tutorial will thus be useful for many of the senior and junior researchers that have interest in global decision problems in NLP, providing a concise overview of recent perspectives and research results.

2 Tentative Outline

After briefly motivating and introducing the general framework, the main part of the tutorial is a methodological presentation of some of the key computational issues studied in this area, that we will present by looking at case studies published in the NLP literature.

2.1 Introduction

We will begin by introducing structured prediction with various NLP examples. We will motivate the framework of Constrained Conditional Models using examples from sequential inference, sentence compression and semantic role labeling.

2.2 Applications of ILP Formulations in Natural Language Processing

Examples will be used to explain several of the key properties the framework offers. In particular, we will discuss several ways in which expressive constraints can be introduced into an application. We will discuss several applications including: Co-reference resolution, Sentence compression, Information extraction, and Parsing.

2.3 Modeling: Inference methods and Constraints

In this part we will discuss how to model problems as structured prediction problems, and focus on the use of hard and soft constraints to represent prior knowledge and the augmentation of probabilistic models with declarative constraints.

We will then discuss a few algorithmic inference frameworks that can be used for ILP formulations, including search methods, Lagrangian Relaxation and the use of existing Packages. We will also introduce the notion of amortized inference for ILP formulations.

2.4 Training Paradigms

We will present several training paradigms that are relevant in the context of ILP formulations. In particular we will discuss and distinguish between Structured learning, and Decomposed learning and study methods for Constraints Driven Learning including Semi-supervised Learning with Constraints, Constrained Expectation Maximization, and Learning with Indirect Supervision.

2.5 Developing ILP based Applications

We will provide a “template-based” approach” for developing applications with ILP formulations and describe some of the tools available.

2.6 Future work

We will discuss some open questions and address questions related to the used of ILP Formulations in the NN era.

3 Tutorial Instructors

Dan Roth

Dan Roth is the Founder Professor of Engineering in the Department of Computer Science at the University of Illinois at Urbana-Champaign and the Beckman Institute of Advanced Science and Technology (UIUC). He is a fellow of the AAAS, ACM, AAAI and the ACL. Roth has published broadly in machine learning, natural language processing, knowledge representation and reasoning and received several paper, teaching and research awards. He has developed several machine learning based natural language processing systems that are widely used in the computational linguistics community and in industry and was the program chair for a number of major conferences, including ACL and AAAI. Roth is currently the

Editor-in-Chief of the Journal of Artificial Intelligence Research (JAIR). Roth has given tutorials on these and other topics in all NLP and AI major conferences.

Vivek Srikumar

Vivek Srikumar is an Assistant Professor in the School of Computing at the University of Utah. His research interests are in the areas of machine learning and natural language processing, particularly in the context of structured learning and prediction. Recently, he has been working on problems related to predicting semantic representations of text, structured learning from limited supervision and scaling inference. He has given multiple tutorials in major conferences such as AAAI and NAACL on predicting NLP structures.

References

- Jonathan Berant, Ido Dagan, and Jacob Goldberger. 2011. Global learning of typed entailment rules. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*. Association for Computational Linguistics, Stroudsburg, PA, USA, HLT '11, pages 610–619.
- T. Koo, A. M. Rush, M. Collins, T. Jaakkola, and D. Sontag. 2010. Dual decomposition for parsing with non-projective head automata. In *Proceedings of the Conference on Empirical Methods for Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Stroudsburg, PA, USA, pages 1288–1298.
- A. F. T. Martins, N. A. Smith, and E. P. Xing. 2009. Polyhedral outer approximations with application to natural language parsing. In *Proceedings of the International Conference on Machine Learning (ICML)*. ACM, New York, NY, USA, pages 713–720. <https://doi.org/http://doi.acm.org/10.1145/1553374.1553466>.
- D. Roth and W. Yih. 2004. A linear programming formulation for global inference in natural language tasks. In Hwee Tou Ng and Ellen Riloff, editors, *Proc. of the Conference on Computational Natural Language Learning (CoNLL)*. Association for Computational Linguistics, pages 1–8. <http://cogcomp.cs.illinois.edu/papers/RothYi04.pdf>.
- D. Roth and W. Yih. 2005. Integer linear programming inference for conditional random fields. In *Proc. of the International Conference on Machine Learning (ICML)*. pages 737–744. <http://cogcomp.cs.illinois.edu/papers/RothYi05.pdf>.