AFRL-RI-RS-TR-2018-135

A BROAD RANGE, PURPOSEFUL, TEXTUAL INFERENCE

UNIVERSITY OF ILLINOIS CHAMPAIGN URBANA (UIUC)

JUNE 2018

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE

■ **AIR FORCE MATERIEL COMMAND** ■ **UNITED STATES AIR FORCE** ■ **ROME, NY 13441**

NOTICE AND SIGNATURE PAGE

AFRL-RI-RS-TR-2018-135 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE CHIEF ENGINEER:

/ S /
JAMES M. NAGY
Work Unit Manager

/ S /
JON S. JONES
Technical Advisor, Information Intelligence
Systems & Analysis Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
JUN 2018

2. REPORT TYPE
FINAL TECHNICAL REPORT

3. DATES COVERED *(From - To)*
OCT 2012 – NOV 2017

4. TITLE AND SUBTITLE
A BROAD RANGE, PURPOSEFUL, TEXTUAL INFERENCE

5a. CONTRACT NUMBER
FA8750-13-2-0008

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER
DEFT

5e. TASK NUMBER
12

5f. WORK UNIT NUMBER
2

6. AUTHOR(S)
Dan Roth

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Illinois at Urbana-Champaign
506 S. Wright ST 364 Henry Admin Bldg
Urbana IL 61801-3620

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/RIED
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RI

11. SPONSOR/MONITOR'S REPORT NUMBER
AFRL-RI-RS-TR-2018-135

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The objective of DARPA's DEFT program is to create capabilities for deep natural language understanding and use them to aid analysts in identifying information sources that contain new developments of interest. The goal of the Cognitive Computation Group team has been to combine Natural Language Processing (NLP), Machine Learning (ML), and Knowledge Representation and Reasoning (KRR) techniques into new technologies that support the DEFT mission. Our project, a broad range purposeful textual inference system, was built on two pillars: 1) an innovative learning and inference approach emphasizing joint inference over a component-based architecture, and 2) a textual inference approach that supports relational analysis in multiple NLP tasks. The project focused on studying and developing four algorithmic components. The first was the aforementioned generic purposeful textual inference capability. The other three components were: (2) a Sentence Level Extended Semantic Role Labeling component that provides a complete and coherent predicate-argument representation of sentences covering multiple predicate types; (3) a Discourse Analysis component that addresses discourse phenomena including relations between events, temporal grounding of events and relations, and time lining of events; and (4) a Profiling component that provides a new way of representing, aggregating, and supporting the use of knowledge about concepts and entities in NLP. This document describes the significant progress made towards the goals stated above in all aspects of the projects. We developed new problem formulations, new algorithmic tools, and new software packages that significantly improves the community's natural language understanding capabilities in all areas mentioned above.

15. SUBJECT TERMS
Natural Language Understanding; Joint Inference; Machine Learning; Entity Linking and Wikification; Semantic Role Labeling; Events; Multilingual Named Entity Recognition and Entity Linking; Discourse Processing.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					JAMES M. NAGY
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (*Include area code*)
U	U	U	UU	26	

TABLE OF CONTENTS

1.0 SUMMARY

The following report summarizes the contribution by the UIUC-Cognitive Computation Group team in the course of the DARPA DEFT program (Award No. FA8750-13-2-0008; Dan Roth, PI).

The objective of DARPA's DEFT program is to create capabilities for deep natural language understanding and use them to aid analysts in identifying information sources that contain new developments of interest.

The goal of the Cognitive Computation Group team has been to combine Natural Language Processing (NLP), Machine Learning (ML), and Knowledge Representation and Reasoning (KRR) techniques into new technologies that support the DEFT mission.

Our project, **a broad range purposeful textual inference system**, was built on two pillars:

1) An innovative learning and inference approach emphasizing joint inference over a component-based architecture, and
2) A textual inference approach that supports relational analysis in multiple NLP tasks.

The project focused on studying and developing four algorithmic components.

1) The first was the aforementioned generic purposeful textual inference capability.
2) The other three components were:
    a. A Sentence Level Extended Semantic Role Labeling component that provides a complete and coherent predicate-argument representation of sentences covering multiple predicate types;
    b. A Discourse Analysis component that addresses discourse phenomena including relations between events, temporal grounding of events and relations, and time lining of events; and
    c. A Profiling component that provides a new way of representing, aggregating, and supporting the use of knowledge about concepts and entities in NLP.

The rest of this document describes the significant progress made towards the goals stated above in all aspects of the projects.

We developed new problem formulations, new algorithmic tools, and new software packages that significantly improves the community's natural language understanding capabilities in all areas mentioned above.

2.0 INTRODUCTION

Our project, **a broad range purposeful textual inference system**, has made significant progress and advanced the state of the art, sometimes contributing to a phase shift in automating natural language understanding. Our advances also facilitated the development of tools that could help analysts access information in ways they could not before.

Key accomplishments include:

a. **Joint Inference:** We developed the Constraint Conditional Model (aka Integer-Linear-Programming (ILP) formulation) for structured prediction problems in natural language processing. Our inference formulation has had an impact on all areas in NLP, from syntax to summarization, to information extraction, to multiple tasks in semantics. This work has changed the way the research community thinks about global inference in natural language processing and has had an impact on all areas of natural language understanding. In addition, we developed a state-of-the-art structured learning package (Illinois-SL) and studied an amortized inference approach that cuts down on the computational cost of the inference by an order of magnitude.

b. **Profiling Concepts and Entities:** Our work on Wikification and entity linking has been the most impactful line of work in this research area over the last few years. From the initial work on the Illinois Wikifier that contributed to the general architecture of all follow-up approaches to this problem, through the work on global inference and coherence, to the cross-lingual Wikification approach we developed, and up to the final stages of the projects (and just after it) – where we developed a general approach to multilingual embeddings and a way to use it in the context of Entity Linking, and the development of a neural approach to learning representations of entities and to the "entity-driven" approach to entity linking.
The work on Wikification, entity representation and linking, and its cross-lingual aspects and the neural approaches for it have all served to establish the key approaches taken by the research community towards these problems.

c. **Discourse Structure and Representation:** We have studied and developed multiple discourse components building on our joint inference formulation and on two key innovations: a new semantic language model, and a new dataless (0-shot) approach that we developed in the context of topic classification and extended to event extraction, co-reference resolution, and discourse parsing. Along with this, we developed models for event extraction and started to develop an approach for consolidation of events. In addition, we studied and developed state-of-the-art models for temporal and quantitative extraction and for temporal and quantitative reasoning, supporting time lining of events as well as mapping quantitative arguments in text to equations. Our work on temporal reasoning and, even more, the work on quantitative reasoning, has helped to establish the initial dataset and algorithmic approaches for this newly studied set of problems.

d. We developed **shallow semantic representations** that extend the standard semantic role labeling (prop-bank) annotation with multiple additional relations, including nominal, prepositions, comma-driven relations, light verbs, and others. As a result we also augmented and improved event extraction and contributed to the discussion around the definition of events and of relations among events.

3.0 METHODS, ASSUMPTIONS, AND PROCEDURES

In the course of the program we developed multiple frameworks, learning and inference formulations, and natural language understanding algorithmic tools; we used those to develop and deliver state-of-the-art Information Extraction and Natural Language Understanding software tools. We describe below our key contributions.

3.1 Extended Sematic Role Labeling

Semantic Role Labeling (SRL) is a shallow semantic analysis at the sentence level that was found to be very useful in downstream natural language understanding tasks. Earlier approaches to SRL focused on verb predicates and the semantic relations they express. In the course of DEFT we have augmented SRL with a large number of relations, including Prepositions [Srikumar & Roth TACL'13] (where we also developed a new taxonomy of relations expressed by prepositions), commas [Arivazhagan et al. AAAI'16] (where we also developed a new taxonomy of relations commas participate in, such as lists, appositions, locative, etc.), Light Verbs, and a few others. In several cases we also developed data resources for the new relations. A few of the other relations we studied include quantities [Roy et al. TACL'15, Roy & Roth EMNLP'15, Roy et al. EMNLP'16, Roy & Roth TACL'18] (more details on that below) and temporal relations [Ning et al. EMNLP'17, Zhao et al NAACL'12] (likewise, see below), as well as Verb Phrase Ellipsis, Metonymy, Light verb constructions, Phrasal verb constructions, Multiword expressions, Preposition Phrase Attachment, Sentence Specificity, Compound nouns, Negation, and Attribution. Our Question Answering project makes use of the Extended SRL and exhibits the benefits of the additional semantic relations [Khashabi et al. AAAI'18].

One key property of our approach is that we do not rely on exhaustive global annotation to all the phenomena, but rather rely on individual annotation the community has developed over the years for independent phenomena. We then exploit interactions among these phenomena as a way to improve performance on individual components.

3.2 Discourse Structure and Representation

In the course of DEFT we have made very significant progress on this component in several directions.

a. We have developed a state-of-the-art entity co-reference module. This was developed in a few stages over a few years, with several innovations including an integrated approach to mention detection and co-reference that resulted in a significant improvement over existing mentions detection modules, a probabilistic model that exploits the "left-to-right" leaning of the co-reference clusters, co-reference of "Winograd Schemas", and others. [Peng et al. CoNLL'15; Peng et al. NAACL'15, Samdani et al. ICML'14; Chang et al. EMNLP'13].

b. We have developed a state-of-the-art discourse parser that identifies and classifies implicit and explicit discourse markers, within and across sentences. [Peng et al. CoNLL'17, Peng & Roth ACL'16, Song et al. CoNLL'15]

The major step we have done in the last part of the project, building on earlier advances (and, in particular, the Extended SRL mentioned above) has been the development of a **semantic language model (SemLM).** [Peng & Roth ACL'16; Peng et al. CoNLL'17]. The basic assumption underlying this research direction is that natural language understanding often requires deep semantic knowledge. We suggested that some important aspects of semantic knowledge can be modeled as a language model if done at an appropriate level of abstraction and exhibited that by developing two distinct models that capture semantic frame chains and discourse information while abstracting over the specific mentions of predicates and entities. We then extended this model to capture additional aspects of semantics including actions, entities and emotions. Our semantic language models thus jointly models important aspects of semantic knowledge – frames, entities, and sentiments. One of the key applications of SemLM, in addition to supporting improvements in tasks such as co-reference and discourse parsing, is that of "Story Completion" [Chaturvedi et al. EMNLP'17]. Automatic story comprehension is a fundamental challenge in Natural Language Understanding and can enable computers to learn about social norms, human behavior and common sense. Based on our Semantic Language Model approach, we developed a story comprehension model that explores three distinct semantic aspects: (i) the sequence of events described in the story, (ii) its emotional trajectory, and (iii) its plot consistency. We judge the model's understanding of real-world stories by inquiring whether, like humans, it can develop an expectation of what will happen next in a given story. Specifically, we use it to predict the correct ending of a given short story from possible alternatives. The model uses a hidden variable to weigh the semantic aspects in the context of the story. Our experiments demonstrate the potential of our approach to characterize these semantic aspects, and the strength of the hidden variable based approach. The model outperforms the state-of-the-art approaches and achieves best results on a publicly available dataset.

We have made multiple significant contributions to Events identification and co-reference as well as to the study of Relations between events, with a focus on Time Lining of Events, and Consolidation of Events. [Wities et al. EACL'17; Peng et al. EMNLP'16; Upadhyay et al. COLING'16; Upadhyay et al. EVENTS-NAACL'16; Lu & Roth ACL'12]
Identifying temporal relations between events is an essential step towards natural language understanding. However, the temporal relation between two events in a story depends on, and is

often dictated by, relations among other events. Consequently, effectively identifying temporal relations between events is a challenging problem even for human annotators. We suggest that it is important to consider these dependencies while learning to identify temporal relations and, consequently, propose a *structured learning* approach to address this challenge. As a byproduct, we show that this new approach also provides a new perspective on handling missing relations, a known issue that hurts existing methods. As we show, the proposed approach results in significant improvements on the two commonly used data sets for this problem. [Ning et al. EMNLP'17].

Considering multiple related events that are described in a single or multiple documents brings up the question of consolidating the information and integrating the event representation. In a recent work in an EACL workshop [Wities et al. EACL'17] we proposed a knowledge representation that aims to represent information conveyed jointly in a set of texts. Our initial approach does so by consolidating event triples produced by an Open Information Extraction system, using entity and predicate co-reference, while modeling information containment between co-referring elements via lexical entailment. We suggest that generating this representation can be a useful step in the NLP pipeline, to give semantic applications an easy handle on consolidated information across multiple texts.
These works built on our main Event modules – event extraction, typing, and co-reference resolution. The key approach we have developed here is that of dataless event identification and co-reference. [Song et al. IJCAI'16; Song & Roth NAACL'15; Song & Roth AAAI'14] While it is clear that an important aspect of natural language understanding involves recognizing and categorizing events and the relations among them, these tasks are quite subtle, and annotating training data for machine learning based approaches is an expensive task, resulting in supervised systems that attempt to learn complex models from small amounts of data, which they over-fit. The direction we have developed addresses this challenge by developing an unsupervised event detection and co-reference system. We view these tasks as semantic similarity problems, between event text and an ontology of types, or between event mentions in respective events. This allows us to use large amounts of out-of-domain text data. Notably, our semantic relatedness function exploits the structure of the text by making use of a semantic-role-labeling based representation of an event. We show that the unsupervised approach to event identification is competitive with the top supervised methods, and, more significantly, our unsupervised event co-reference approach outperforms the state-of-the-art supervised methods on the benchmark data sets, and supports significantly better transfer across domains. The key reason our results are so good, we think, is that our focus on learning a representation that is, essentially, independent of the task, allows us to use a lot more data than is available to supervise the specific task; this results in the generation of robust representations that require little tuning to fit a specific task.
In addition to this main task we have also investigated a few subtler issues that are necessary to understanding events and using it in downstream natural language understanding tasks. Specifically, we have looked at the identification of *noteworthy events* in news articles, and the idea of "first story detection," via the notion of a news-peg, a concept borrowed from the political science literature, that serves as a more fine-grained measure of noteworthiness than a summary.

3.3 **Profiling** Concepts **and Entities**

In the course of DEFT, we have developed several Entity Linking (Wikification) approaches, each with different advantages, catering to different needs, and with different levels of supervision available.

Our work in this area set the standards in this domain and has been very influential.

Some of the key contributions we have made in this area include the following:
[Tsai & Roth AAAI'18;
Gupta et al. EMNLP'17;
Mayhew et al. EMNLP'17;
Tsai & Roth TACL'16;
Tsai & Roth NAACL'16;
Tsai et al. CoNLL'16;
Tsai & Roth COLING'16;
Cheng & Roth EMNLP'13]:

(i) Cross-lingual Wikifier: we have developed an Entity Linking approach that can be used to Wikify all languages in Wikipedia into the English Wikipedia. In particular, the approach we developed for this problem was shown to have an important byproduct – it helped to improve cross-lingual Named Entity Recognition (NER); for example, in Spanish, where our system was ranked first in the Text Analysis Conference (TAC) evaluation, we have shown 10% improvements over existing Spanish NER systems.

(ii) As part of our work on cross-lingual Wikification and NER, we have developed a general approach to multilingual embeddings. [Upadhyay et al. ACL'16; Ling et al. ACL'16; Wieting et al. TACL'15] This has since been extended to support multiple lexical semantics tasks, such as lexical entailment.

(iii) We have developed an Entity Linking approach that maps entities in text to knowledge base entries other than Wikipedia. The challenge in this problem is that, unlike Wikipedia, standard knowledge bases will not have the rich textual context for each entity. We therefore developed an indirect supervision approach that exploits the little redundancy there is in the target knowledge bases as a way to train the entity linking model.

(iv) We have developed a second generation Wikifier that can handle additional encyclopedic resources. It is different from the previous generation in that it is an "entity" driven approach rather than a "distance" driven approach. Since its development in the late stage of DEFT this model has been influential on all Neural based approaches to Wikification.

(v) We have developed a Fine-Named-Entity Recognition module that combines Wikification and fine NER. Our work here is based on the observation that, unlike coarse NER, to type entities in a fine way, we often cannot use information in the context, but instead they need to be "wikified" to get their types.

3.4 Purposeful Textual Inference System

In the course of DEFT we have developed a generic textual inference system, and then focused on three fundamental questions.

a. Lexical similarity and Entailment: we developed new phrasal embeddings and similarity metrics based on these, with the goal of improving lexical entailment. We have shown that the use of a paraphrase data set as a way to better train a similarity metric yields very significant improvements. Following this work we have made several steps in the study of embeddings for similarity and entailment. In one study we considered the size of the embeddings, where we studied the effect of limited precision data representation and computation on word embeddings, and presented a systematic evaluation of word embeddings with limited memory – including methods that directly train the limited precision representation with limited memory. Our results show that it is possible to use and train an 8-bit fixed-point value for word embedding (instead of the standard 64 bits) without loss of performance in word/phrase similarity and dependency parsing tasks. This is important because memory issues are becoming significant in practical systems that use huge numbers of embeddings. In a second two-part study, we developed new multilingual embeddings both for similarity and for lexical entailment. This work is still in progress and it focuses on studying the type of NLP resources that are needed in order to develop embeddings that can support entailment. [Upadhyay et al. ACL'16; Ling et al. ACL'16; Wieting et al. TACL'15]

b. We have focused our inference work on Quantity Entailment and Quantitative Reasoning tasks. This has been a multi-step process where, for the most part, we make use of the abstraction of math word problems as a way to pursue quantitative reasoning. Math word problems form a natural abstraction to a range of quantitative reasoning problems, such as understanding financial news, sports results, and casualties of war, and more. Solving such problems requires several pieces of domain knowledge associated with mathematics such as understanding transactions, dimensional analysis, subset relationships, etc., and often, as we show, it is better to add this information declaratively. We have developed three generations of system in this area, as well as data resource, and set the standards for work on quantitative reasoning from natural language text. [Roy et al. TACL'15, Roy & Roth EMNLP'15, Roy et al. EMNLP'16, Roy & Roth TACL'18]

c. We have developed state-of-the-art tools for temporal information extraction and for Temporal Reasoning. Our approach, that is an outcome of our Events work discussed above, allows us to map a time interval to events appearing in text and then order the events temporally. [Ning et al. EMNLP'17, Zhao et al. NAACL'12]

3.5 Joint Inference Component-supporting Architecture (JOINCA)

We have made significant progress in our joint learning and inference framework, developing a better understanding for how to support global decisions using tractable, decomposable methods in the context of all the NLP work mentioned above. Virtually all the work mentioned makes progress within this joint inference framework. In particular, we developed several generic structured learning package and a generic inference modulethat takes as input an objective function (learned via multiple models) and a set of declarative constraints. With this general formulation, we allow the use of multiple inference algorithm, both approximate and exact. [Lee & Roth ICML'15; Lee et al. NIPS WRKSP'15; Chang et al. ECML'13; Samdani et al. NAACL'12; Chang et al. Machine Learning Journal'12]

A key algorithmic contribution over the last few years has been that of the development of an Amortized Inference Paradigm, which allows us to cut the cost of inference using integer linear programs by an order of magnitude. [Chang et al. AAAI'15; Kundu et al. ACL'13; Srikumar et al. EMNLP'12]

In addition, we have developed in [Roth AAAI'17] a general approach to Incidental Supervision, a framework that gets around the need to train supervised machine learning models even in cases when there is not enough directly supervised data, and it is impossible to scale up to the levels of supervised annotation needed.

4.0 RESULTS AND DISCUSSION

In the course of the DEFT program we have developed and delivered a large number of algorithmic components, addressing all the research topics above, all with state-of-the-art performance. Many of the software packages we developed have been integrated into the ADEPT framework. All are available on the cogcomp web page: http://cogcomp.org/page/software/ and https://github.com/CogComp. They are being used by the research community, and sometimes also commercially.

Deliverables:

a. Tokenizer/sentence segmenter, Part of Speech tagger, and Shallow Parser. Integrated into ADEPT framework.
b. State-of-the-art named entity recognition that supports all Wikipedia languages. Integrated into ADEPT framework. Used in Tinkerbell end-to-end system for TAC Cold Start++ evaluation.
c. State-of-the-art cross-lingual Wikifier and nil clustering. Integrated into ADEPT framework, used in Tinkerbell end-to-end system for TAC Cold Start++ evaluation; nil clustering component used in BBN end-to-end system for TAC Cold Start++ evaluation.
d. State-of-the-art entity co-reference resolution package. Integrated into ADEPT framework, and used in BBN end-to-end system for TAC Cold Start++ evaluation.
e. Extended Semantic Role Labeling package.
f. An RTE (Recognizing Textual Entailment) module (for lexical entailment). Integrated into ADEPT framework.
g. Event extraction, temporal extraction and time lining of events. Used in Tinkerbell end-to-end system for TAC Cold Start++ evaluation.
h. Quantity extraction and an algebra word problem solver.
i. Semantic language model.
k. Cross-lingual embeddings (multiple versions appropriate for different settings).

NOTE: The deliverables integrated into the DEFT ADEPT framework in coordination with BBN were done so as individual modules, rather than as a monolithic combined component with a single point of access. This was done with the intention to allow straightforward integration with components from other teams, which a monolithic approach does not allow.

5.0 CONCLUSIONS

The reports describe the contributions made by the UIUC project, **a broad range purposeful textual inference system**, in the course of the DEFT program. We have made significant progress, sometimes resulting in a phase shift, and we have developed new problem formulations, new algorithmic tools, and new software packages that significantly improve the community's natural language understanding capabilities in all areas mentioned above.

6.0 REFERENCES

1. Srikumar, V., Roth, D., "Modeling Semantic Relations Expressed by Prepositions," *TACL - 2013*
2. Arivazhagan, N., Christodoulopoulos, C., Roth, D., "Labeling the Semantic Roles of Commas," *AAAI - 2016*
3. Roy, S., Vieira, T., Roth, D., "Reasoning about Quantities in Natural Language," *TACL - 2015*
4. Roy, S., Roth, D., "Solving General Arithmetic Word Problems," *EMNLP - 2015*
5. Roy, S., Upadhyay, S., Roth, D., "EQUATION PARSING : Mapping Sentences to Grounded Equations," *EMNLP - 2016*
6. Roy, S., Roth, D., Mapping to Declarative Knowledge for Word Problem Solving *TACL - 2018*
7. Ning, Q., Feng, Z., Roth, D., "A Structured Learning Approach to Temporal Relation Extraction," *EMNLP - 2017*
8. Zhao, R., Do, Q., Roth, D., A Robust Shallow Temporal Reasoning System *NAACL - 2012*
9. Khashabi, D., Khot, T., Sabharwal, A., Roth, D., "Question Answering as Global Reasoning over Semantic Abstractions," *AAAI - 2018*
10. Peng, H., Chang, K.-W., Roth, D., "A Joint Framework for Coreference Resolution and Mention Head Detection," *CoNLL - 2015*
11. Peng, H., Khashabi, D., Roth, D., "Solving Hard Coreference Problems," *NAACL - 2015*
12. Samdani, R., Chang, K.-W., Roth, D., "A Discriminative Latent Variable Model for Online Clustering," *ICML - 2014*
13. Chang, K.-W., Samdani, R., Roth, D., A Constrained Latent Variable Model for Coreference Resolution *EMNLP - 2013*
14. Peng, H., Chaturvedi, S., Roth, D., "A Joint Model for Semantic Sequences: Frames, Entities, Sentiments," *CoNLL - 2017*
15. Peng, H., Roth, D., "Two Discourse Driven Language Models for Semantics," *ACL - 2016*
16. Song, Y., Peng, H., Kordjamshidi, P., Sammons, M., Roth, D., "Improving a Pipeline Architecture for Shallow Discourse Parsing," *CoNLL - 2015*
17. Chaturvedi, S., Peng, H., Roth, D., "Story Comprehension for Predicting What Happens Next," *EMNLP - 2017*
18. Wities, R., Shwartz, V., Stanovsky, G., Adler, M., Shapira, O., Upadhyay, S., Roth, D., Martinez Camara, E., Gurevych, I., Dagan, I., A Consolidated Open Knowledge Representation for Multiple Texts *Proceedings of the 2nd LSDSem Workshop, in EACL - 2017*
19. Peng, H., Song, Y., Roth, D., "Event Detection and Co-reference with Minimal Supervision," *EMNLP - 2016*
20. Upadhyay, S., Gupta, N., Christodoulopoulos, C., Roth, D., "Revisiting the Evaluation for Cross Document Event Coreference," *COLING - 2016*
21. Upadhyay, S., Christodoulopoulos, C., Roth, D., "Making the News - Identifying Noteworthy Events in News Articles," *The 4th Workshop on EVENTS, NAACL - 2016*
22. Lu, W., Roth, D., Automatic Event Extraction with Structured Preference Modeling *ACL - 2012*
23. Song, Y., Upadhyay, S., Peng, H., Roth, D., "Cross-lingual Dataless Classification for Many Languages," *IJCAI - 2016*

24. Song, Y., Roth, D., "Unsupervised Sparse Vector Densification for Short Text Similarity," *NAACL - 2015*
25. Song, Y., Roth, D., "On Dataless Hierarchical Text Classification," *AAAI - 2014*
26. Tsai, C.-T., Roth, D., "Learning Better Name Translation for Cross-Lingual Wikification," *AAAI - 2018*
27. Gupta, N., Singh, S., Roth, D., "Entity Linking via Joint Encoding of Types, Descriptions, and Context," *EMNLP - 2017*
28. Mayhew, S., Tsai, C.-T., Roth, D., "Cheap Translation for Cross-Lingual Named Entity Recognition," *EMNLP – 2017*
29. Tsai, C.-T., Roth, D., Concept Grounding to Multiple Knowledge Bases via Indirect Supervision *TACL - 2016*
30. Tsai, C.-T., Roth, D., "Cross-lingual Wikification Using Multilingual Embeddings," *NAACL - 2016*
31. Tsai, C.-T., Mayhew, S., Roth, D., "Cross-Lingual Named Entity Recognition via Wikification," *CoNLL - 2016*
32. Tsai, C.-T., Roth, D., Illinois Cross-Lingual Wikifier: Grounding Entities in Many Languages to the English Wikipedia *COLING Demonstrations - 2016*
33. Cheng, X., Roth, D., "Relational Inference for Wikification," *EMNLP - 2013*
34. Upadhyay, S., Faruqui, M., Dyer, C., Roth, D., "Cross-lingual Models of Word Embeddings: An Empirical Comparison," *ACL - 2016*
35. Ling, S., Song, Y., Roth, D., "Word Embeddings with Limited Memory," *ACL - 2016*
36. Wieting, J., Bansal, M., Gimpel, K., Livescu, K., Roth, D., "From Paraphrase Database to Compositional Paraphrase Model and Back," *TACL - 2015*
37. Lee, C.-p., Roth, D., "Distributed Box-Constrained Quadratic Optimization for Dual Linear SVM," *ICML - 2015*
38. Lee, C.-p., Chang, K.-W., Upadhyay, S., Roth, D., Distributed Training of Structured SVM *OPT Workshop, NIPS - 2015*
39. Samdani, R., Chang, M.-W., Roth, D., Unified Expectation Maximization *NAACL - 2012*
40. Chang, M.-W., Ratinov, L., Roth, D., Structured Learning with Constrained Conditional Models*Machine Learning* (2012) pp. 399-431
41. Chang, K.-W., Upadhyay, S., Kundu, G., Roth, D., "Structural Learning with Amortized Inference," *AAAI - 2015*
42. Kundu, G., Srikumar, V., Roth, D., "Margin-based Decomposed Amortized Inference," *ACL - 2013*
43. Srikumar, V., Kundu, G., Roth, D., On Amortizing Inference Cost for Structured Prediction *EMNLP - 2012*
44. Roth, D., "Incidental Supervision: Moving beyond Supervised Learning," *AAAI - 2017*

7.0 APPENDIX: PUBLICATIONS AND PRESENTATIONS

7.1 Publications:

1. Ning, Q., Wu, H., Peng, H., Roth, D., “Improving Temporal Relation Extraction with a Globally Acquired Statistical Resource”. Submitted to NAACL’18.
2. Ning, Q., Feng, Z., Wu, H., Roth, D., “Joint Reasoning for Temporal and Causal Relations”. Submitted to NAACL’18.
3. Upadhyay, S., Gupta, N., Roth, D., “Exploiting Multilingual Supervision for Cross-lingual Entity Linking”. In submission.
4. Upadhyay, S., Vyas, Y., Carpuat, M., Roth, D., “Robust Cross-lingual Hypernymy Detection using Dependency Context.” In submission.
5. Wu, H., Muddireddy, P., Roth, D., “Fine-Grained Entity Typing - A Detailed Study”. In submission.
6. Roy, S., Roth, D., Mapping to Declarative Knowledge for Word Problem Solving *TACL - 2018*
7. Khashabi, D., Khot, T., Sabharwal, A., Roth, D., “Question Answering as Global Reasoning over Semantic Abstractions,” *AAAI - 2018*
8. Tsai, C.-T., Roth, D., "Learning Better Name Translation for Cross-Lingual Wikification," *AAAI - 2018*
9. Wities, R., Shwartz, V., Stanovsky, G., Adler, M., Shapira, O., Upadhyay, S., Roth, D., Martinez Camara, E., Gurevych, I., Dagan, I., A Consolidated Open Knowledge Representation for Multiple Texts *Proceedings of the 2nd LSDSem Workshop, in EACL - 2017*
10. Duncan, C., Chen, L.-W., Peng, H., Wu, H., Upadhyay, S., Gupta, N., Tsai, C.-T., Sammons, M., Roth, D., "UI CCG TAC-KBP2017 Submissions: Entity Discovery and Linking, and Event Nugget Detection and Co-reference," *Text Analysis Conference (TAC) - 2017*
11. Ning, Q., Feng, Z., Roth, D., "A Structured Learning Approach to Temporal Relation Extraction," *EMNLP - 2017*
12. Chaturvedi, S., Peng, H., Roth, D., "Story Comprehension for Predicting What Happens Next," *EMNLP - 2017*
13. Mayhew, S., Tsai, C.-T., Roth, D., "Cheap Translation for Cross-Lingual Named Entity Recognition," *EMNLP - 2017*
14. Gupta, N., Singh, S., Roth, D., "Entity Linking via Joint Encoding of Types, Descriptions, and Context," *EMNLP - 2017*
15. Peng, H., Chaturvedi, S., Roth, D., "A Joint Model for Semantic Sequences: Frames, Entities, Sentiments," *CoNLL - 2017*
16. Khashabi, D., Khot, T., Sabharwal, A., Roth, D., "Learning What is Essential in Questions," *The Conference on Computational Natural Language Learning (CoNLL) - 2017*
17. Tsai, C.-T., Roth, D., Concept Grounding to Multiple Knowledge Bases via Indirect Supervision *TACL - 2016*
18. Wang, C., Song, Y., Roth, D., Zhang, M., Han, J., "World Knowledge as Indirect Supervision for Document Clustering," *ACM TKDD - 2016*
19. Roy, S., Upadhyay, S., Roth, D., "EQUATION PARSING : Mapping Sentences to Grounded Equations," *EMNLP - 2016*

20. Peng, H., Song, Y., Roth, D., "Event Detection and Co-reference with Minimal Supervision," *EMNLP - 2016*
21. Rozovskaya, A., Roth, D., "Grammatical Error Correction: Machine Translation and Classifiers," *ACL - 2016*
22. Sammons, M., Christodoulopoulos, C., Kordjamshidi, P., Khashabi, D., Srikumar, V., Vijayakumar, P., Bokhari, M., Wu, X., Roth, D., "EDISON: Feature Extraction for NLP, Simplified," *LREC - 2016*
23. Ling, S., Song, Y., Roth, D., "Word Embeddings with Limited Memory," *ACL - 2016*
24. Song, Y., Upadhyay, S., Peng, H., Roth, D., "Cross-lingual Dataless Classification for Many Languages," *IJCAI - 2016*
25. Arivazhagan, N., Christodoulopoulos, C., Roth, D., "Labeling the Semantic Roles of Commas," *AAAI - 2016*
26. Tsai, C.-T., Mayhew, S., Peng, H., Sammons, M., Mangipundi, B., Reddy, P., Roth, D., "Illinois CCG Entity Discovery and Linking, Event Nugget Detection and Co-reference, and Slot Filler Validation Systems for TAC 2016," *Text Analysis Conference (TAC 2016) - 2016*
27. Roth, D., "Incidental Supervision: Moving beyond Supervised Learning," *AAAI - 2017*
28. Upadhyay, S., Gupta, N., Christodoulopoulos, C., Roth, D., "Revisiting the Evaluation for Cross Document Event Coreference," *COLING - 2016*
29. Tsai, C.-T., Roth, D., Illinois Cross-Lingual Wikifier: Grounding Entities in Many Languages to the English Wikipedia *COLING Demonstrations - 2016*
30. Upadhyay, S., Christodoulopoulos, C., Roth, D., "Making the News - Identifying Noteworthy Events in News Articles," *The 4th Workshop on EVENTS, NAACL - 2016*
31. Tsai, C.-T., Mayhew, S., Roth, D., "Cross-Lingual Named Entity Recognition via Wikification," *CoNLL - 2016*
32. Peng, H., Roth, D., "Two Discourse Driven Language Models for Semantics," *ACL - 2016*
33. Upadhyay, S., Faruqui, M., Dyer, C., Roth, D., "Cross-lingual Models of Word Embeddings: An Empirical Comparison," *ACL - 2016*
34. Tsai, C.-T., Roth, D., "Cross-lingual Wikification Using Multilingual Embeddings," *NAACL - 2016*
35. Sammons, M., Peng, H., Song, Y., Upadhyay, S., Tsai, C.-T., Reddy, P., Roy, S., Roth, D., "Illinois CCG TAC 2015 Event Nugget, Entity Discovery and Linking, and Slot Filler Validation Systems," *TAC - 2015*
36. Wieting, J., Bansal, M., Gimpel, K., Livescu, K., Roth, D., "From Paraphrase Database to Compositional Paraphrase Model and Back," *TACL - 2015*
37. Song, Y., Peng, H., Kordjamshidi, P., Sammons, M., Roth, D., "Improving a Pipeline Architecture for Shallow Discourse Parsing," *CoNLL - 2015*
38. Roy, S., Roth, D., "Solving General Arithmetic Word Problems," *EMNLP - 2015*
39. Peng, H., Chang, K.-W., Roth, D., "A Joint Framework for Coreference Resolution and Mention Head Detection," *CoNLL - 2015*
40. Fei, Z., Khashabi, D., Peng, H., Wu, H., Roth, D., "ILLINOIS-PROFILER: Knowledge Schemas at Scale," *Workshop on Cognitive Knowledge Acquisition, Applications, IJCAI - 2015*
41. Lee, C.-p., Roth, D., "Distributed Box-Constrained Quadratic Optimization for Dual Linear SVM," *ICML - 2015*

42. Kordjamshidi, P., Wu, H., Roth, D., "Saul: Towards Declarative Learning Based Programming," *IJCAI - 2015*
43. Wang, C., Song, Y., El-Kishky, A., Roth, D., Zhang, M., Han, J., "Incorporating World Knowledge to Document Clustering via Heterogeneous Information Networks," *KDD - 2015*
44. Wang, C., Song, Y., Roth, D., Wang, C., Han, J., Ji, H., Zhang, M., "Constrained Information-Theoretic Tripartite Graph Clustering to Identify Semantically Similar Relations," *IJCAI - 2015*
45. Song, Y., Roth, D., "Unsupervised Sparse Vector Densification for Short Text Similarity," *NAACL - 2015*
46. Peng, H., Khashabi, D., Roth, D., "Solving Hard Coreference Problems," *NAACL - 2015*
47. Roy, S., Vieira, T., Roth, D., "Reasoning about Quantities in Natural Language," *TACL - 2015*
48. Lee, C.-p., Chang, K.-W., Upadhyay, S., Roth, D., Distributed Training of Structured SVM *OPT Workshop, NIPS - 2015*
49. Sammons, M., Song, Y., Wang, R., Kundu, G., Tsai, C.-T., Upadhyay, S., Ancha, S., Mayhew, S., "Overview of UI-CCG Systems for Event Argument Extraction, Entity Discovery and Linking, and Slot Filler Validation," *TAC - 2014*
50. Chang, K.-W., Upadhyay, S., Kundu, G., Roth, D., "Structural Learning with Amortized Inference," *AAAI - 2015*
51. Rozovskaya, A., Roth, D., "Building a State-of-the-Art Grammatical Error Correction System," *TACL - 2014*
52. Song, Y., Roth, D., "On Dataless Hierarchical Text Classification," *AAAI - 2014*
53. Rozovskaya, A., Chang, K.-W., Sammons, M., Roth, D., Habash, N., "The Illinois-Columbia System in the CoNLL-2014 Shared Task," *CoNLL - 2014*
54. Wu, H., Fei, Z., Dai, A., Mayhew, S., Sammons, M., Roth, D., "IllinoisCloudNLP: Text Analytics Services in the Cloud," *LREC - 2014*
55. Rozovskaya, A., Roth, D., Srikumar, V., "Correcting Grammatical Verb Errors," *EACL - 2014*
56. Goldwasser, D., Roth, D., "Learning from Natural Instructions," *Machine Learning - 2014*
57. Samdani, R., Chang, K.-W., Roth, D., "A Discriminative Latent Variable Model for Online Clustering," *ICML - 2014*
58. Chang, K.-W., Samdani, R., Roth, D., A Constrained Latent Variable Model for Coreference Resolution *EMNLP - 2013*
59. Cheng, X., Roth, D., "Relational Inference for Wikification," *EMNLP - 2013*
60. Rozovskaya, A., Roth, D., "Joint Learning and Inference for Grammatical Error Correction," *EMNLP - 2013*
61. Rozovskaya, A., Chang, K.-W., Sammons, M., Roth, D., "The University of Illinois System in the CoNLL-2013 Shared Task," *CoNLL - 2013*
62. Kundu, G., Srikumar, V., Roth, D., "Margin-based Decomposed Amortized Inference," *ACL - 2013*
63. Chang, K.-W., Srikumar, V., Roth, D., "Multi-core Structural SVM Training," *ECML-PKDD - 2013*
64. Srikumar, V., Roth, D., "Modeling Semantic Relations Expressed by Prepositions," *TACL - 2013*

65. Goldwasser, D., Roth, D., "Leveraging Domain-Independent Information in Semantic Parsing," *ACL - 2013*
66. Lu, W., Roth, D., Automatic Event Extraction with Structured Preference Modeling *ACL - 2012*
67. Srikumar, V., Kundu, G., Roth, D., On Amortizing Inference Cost for Structured Prediction *EMNLP - 2012*
68. Chang, M.-W., Ratinov, L., Roth, D., Structured Learning with Constrained Conditional Models*Machine Learning* (2012) pp. 399-431
69. Samdani, R., Chang, M.-W., Roth, D., Unified Expectation Maximization *NAACL - 2012*
70. Zhao, R., Do, Q., Roth, D., A Robust Shallow Temporal Reasoning System *NAACL – 2012*

7.2 Presentations:

7.2.1 Distinguished Lectures and Keynote Talks

1. Computer Science Distinguished Lecture Series, Northwestern University, Chicago IL, October 2017. *Natural Language Understanding with Incidental Supervision.*
2. IJCAI John McCarthy Award Lecture., IJCAI 2017, Melbourne, Australia, August 2017. *The Necessity of Learning and Reasoning: A Natural Language Understanding Perspective.*
3. Illinois Health Data Analytics Summit, University of Illinois, May 2017. *Natural Language Processing in Support of Healthcare.*
4. NSF Smart & Connected Health Visioning Meeting, Boston University, March 2017. *Natural Language Processing in Support of Healthcare.*
5. American Bar Association (ABA) Antitrust Law Spring Meeting, Washington DC, March 2017. *Expert Testimony of the Future of Machine Learning in the Legal Domain.*
6. Workshop on India's Tryst with Artificial Intelligence, Bangalore, India, January 2017. *Making Sense of Unstructured Data: The Emergence of AI.*
7. NIPS 2016 Workshop on Cognitive Computation, Barcelona, Spain, December 2016. A keynote speech on *Natural Language Understanding with Common Sense Reasoning*.
8. 15th Conference of the Italian Association for Artificial Intelligence, Genoa, Italy, November 2016. A keynote speech on *Inducing Semantics with Minimal (or No) Supervision.*
9. Computer Science Distinguished Lecture Series, Northeastern University, Boston MA, November 2016. *Making Sense of (And Trusting) Unstructured Data*.
10. The Conference of the Association of Computational Linguistics. The 7th Workshop on Cognitive Aspects of Computational Language Learning, August 2016. Berlin, Germany. An invited talk on *Starting from Scratch in Semantic Role Labeling*.
11. The North American Conference of the Association of Computational Linguistics, The 4th Workshop on EVENTS, June 2016. San Diego, CA. An invited talk on *Events in Natural Language Text*.
12. IBM Research, White Planes, NY., September 2016. An invited talk on *Inducing Semantics with Minimal (or No) Supervision.*
13. The University of Amsterdam, a EU workshop on Semantic Processing. June 2016. Amsterdam, The Netherlands. An invited talk on *Inducing Semantics with Minimal (or No) Supervision*.
14. University of Wisconsin, Madison, WI, May 2016. Department of Computer Science Distinguished Lecture Series on Data Management and Analysis. An invited talk on *Making Sense of (And Trusting) Unstructured Data.*
15. Rutgers University, New Brunswick, NJ, April 2016. Department of Homeland Security Retreat. An Invited Lecture on *Information Trustworthiness*.
16. University of Pennsylvania, Philadelphia, PA, February 2016. Department of Computer Science Distinguished Lecture Series. An invited talk on *Constraints Driven Learning and Inference for Natural Language Understanding*.
17. AAAI'16, Phoenix, AZ. A workshop on Declarative Learning Based Programming, January 2016. A keynote speech on *Declarative Learning Based Programming*.
18. Boston University, Boston, MA., January 2016. An invited talk on *Constraints Driven Learning and Inference for Natural Language Understanding.*

19. The University of Utrecht, The Netherlands, October 2015. A workshop on Common Sense and Logic for Reasoning in Natural Language. Keynote Speaker. *Common Sense Reasoning for Natural Language Understanding.*
20. NLPCC, Nanchang, China, October 2015. The 4th China Computer Federation Conference on Natural Language Processing & Chinese Computing. Keynote Speaker. *Learning and Inference for Natural Language Understanding.*
21. TSD, Plzen, Czech Republic, September 2015. The 18th International Conference on Text, Speech and Dialogue. Keynote Speaker. *Learning and Inference for Natural Language Understanding.*
22. IJCAI'15, Buenos Aires, Argentina, July 2015. The 10th International Workshop on Neural-Symbolic Learning and Reasoning (NeSy'15). Distinguished Workshop Speaker. *Natural Language Understanding with Common Sense Reasoning.*
23. Microsoft Research, Redmond, WA, July 2015. An Invited Lecture in the MSR Faculty Summit. *Common Sense Reasoning for Natural Language Understanding.*
24. Data Science Initiative, Distinguished Lecture Series, Boston University, Boston MA, April 2015. *Learning and Inference for Natural Language Understanding*.
25. Wolfram Research, Champaign, IL, March 2015. *Progress in Natural Language Understanding*.
26. Google, Mountain View, CA, February 2015. *Top Ten Challenges in Natural Language Understanding*.
27. Singapore University of Technology and Design, Singapore, December 2014. *Making Sense of Unstructured Data*.
28. Singapore National University, Singapore, December 2014. *Learning and Inference for Natural Language Understanding*.
29. Advanced Digital Sciences Center, Singapore, December 2014. Workshop on Natural Language Processing. A keynote talk on *Learning and Inference for Natural Language Understanding*
30. Ben-Gurion University, Beer-Sheva, Israel, December 2014. *Learning and Inference for Natural Language Understanding*.
31. Cornell University, Ithaca, NY, November 2014. *Learning and Inference for Natural Language Understanding*.
32. Rochester Institute of Technology, Rochester, NY, October 2014. Distinguished Computational Linguistics Lecture on *Learning and Inference for Natural Language Understanding.*
33. Rutgers University, New Brunswick, NJ, October 2014. A Fusion Fest Workshop in honor of Paul Kantor. An Invited Lecture on *Making Sense of Unstructured Data*.
34. Andreessen Horowitz, Academic Roundtable. Palo Alto, CA. September 2014. *Data Science: Making Sense of Unstructured Data.*
35. AutoML, an ICML workshop, June 2014. Beijing, China. A keynote speech on *Learning Based Programming.*
36. EACL'14, The European Conference on Computational Linguistics. Gutenberg, Sweden, April 2014. A keynote speech on *Learning and Inference for Natural Language Understanding.*
37. Allen Institute for AI (AI2), Seattle, WA., March 2014. A Distinguished Lecture Series talk on *Learning and Inference for Natural Language Understanding*.

38. ITA 2014, The Information Theory and Applications Workshop, San Diego, CA, February 2014. An invited talk on *Amortized Integer Linear Programming Inference*.
39. NIPS 2013 Workshop on Output Representation Learning, Lake Tahoe, CA, December 2013. A keynote speech on *Amortized Integer Linear Programming Inference*.
40. Fondazione Bruno Kessler, The Center for Information and Communication Technology, Trento, Italy, November 2013. Distinguished Lecture Colloquium on *Amortized Integer Linear Programming Inference.*
41. JSSP 2013 - Joint Symposium on Semantic Processing, Trento, Italy, November 2013. A keynote speech on *Computational Frameworks for Supporting Textual Inference*.
42. Institute of Computational Linguistics, Distinguished Lecture Colloquium, The University of Heidelberg, Heidelberg, Germany, October 2013. *Better Natural Language Analysis and Amortized Integer Linear Programming.*
43. The CIKM Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR'13), San Francisco, CA, October 2013. A keynote speech on *Computational Frameworks for Semantic Analysis and Wikification.*
44. The University of Washington & Microsoft Research Summer Institute on Understanding Situated Language in Everyday Life, July 2013. A keynote speech on *Starting from Scratch in Semantic Role Labeling.*
45. The Second AAAI workshop on Combining Constraint Solving with Mining and Learning, July 2013. A keynote speech on *Amortized Integer Linear Programming Inference.*
46. Inferning: Interactions between Learning and Inference, an ICML workshop, June 2013. A keynote speech on *Amortized Integer Linear Programming Inference*.
47. Structured Learning: Inferring Graphs from Structured and Unstructured Inputs, an ICML Workshop, June 2013. A keynote speech on *Decomposing Structured Prediction via Constrained Conditional Models.*
48. 22nd Annual Belgium-Netherlands Conference on Machine Learning (BENELEARN-2013), Nijmegen, the Netherland, June 2013. A keynote speech on *Constrained Conditional Models: Towards Better Semantic Analysis of Text.*
49. KU Leuven Distinguished Lecture Series, Leuven, Belgium, May 2013. *Constrained Conditional Models: ILP Formulations for Natural Language Understanding.*
50. Computational Science and Engineering Center, University of Illinois. Keynote Speech at the Annual Meeting, April 2013. *Making Sense of and Trusting, Unstructured Data*.
51. A COLING Workshop on Information Extraction & Entity Analytics on Social Media Data, December 2012. A keynote speech on *Constraints Driven Information Extraction and Trustworthiness.*
52. The Annual Italian Operation Research Meeting (AIRO 2012), Salerno, Italy, September 2012. A keynote speech on *Constrained Conditional Models Integer Linear Programming Formulations for Natural Language Understanding*.
53. The 2012 Workshop on Statistical Relational AI (STAR-AI 2012), Catalina Island, CA, August 2012. A keynote speech on *Constrained Conditional Models Integer Linear Programming Formulations for Natural Language Understanding*.
54. An NAACL'12 Workshop on "From Words to Actions", Montreal, Canada, June 2012. A keynote speech on *Learning from Natural Instructions*.
55. Semantic Representation and Inference, A Workshop sponsored by the NSF and the Stanford Center for Language and Information (CLSI), Stanford, CA, March 2012. *Constrained Conditional Models for Natural Language Understanding.*

7.3 Other Invited Talks (Colloquia Talks)

1. IBM Research, White Plains, NY, September 2016. *Inducing Semantics with Minimal (or No) Supervision..*
2. Boston University, Boston, MA, January 2016. *Constraints Driven Learning and Inference for Natural Language Understanding.*
3. Peking University, Beijing, China, October 2015. *Learning and Inference for Natural Language Understanding.*
4. Charles University, Prague, Czech Republic, September 2015. *Learning and Inference for Natural Language Understanding.*
5. INRIA Lille, France, May 2015. *Learning, Inference and Supervision for Structured Prediction Tasks.*
6. INRIA, Paris, France, May 2015. *Learning and Inference for Natural Language Understanding.*
7. Wolfram Research, Champaign, IL, March 2014. *Progress in Natural Language Understanding.*
8. Google, Mountain View, CA, February 2015. *Top Ten Challenges in Natural Language Understanding.*
9. Ben-Gurion University, Beer-Sheva, Israel, December 2014. *Learning and Inference for Natural Language Understanding.*
10. Singapore National University, Singapore, December 2014. *Learning and Inference for Natural Language Understanding.*
11. Singapore University of Technology and Design, Singapore, December 2014. *Making Sense of Unstructured Data.*
12. Cornell University, Ithaca, NY, November 2014. *Learning and Inference for Natural Language Understanding.*
13. University of Rochester, Rochester, NY, October 2014. "Big Picture Series" Lecture on *Learning and Inference for Natural Language Understanding.*
14. Microsoft Research, Beijing, China, June 2014. *Learning and Inference for Natural Language Understanding.*
15. Rensselaer Polytechnic Institute (RPI), Troy, NY, April 2014. *Learning and Inference for Natural Language Understanding.*
16. Columbia University, NYC, NY, March 2014. *Learning and Inference for Natural Language Understanding.*
17. Google, Mountain View, CA, March 2014. *Learning and Inference for Natural Language Understanding.*
18. University of California, Santa Cruz. February, 2014. *Learning and Inference for Natural Language Understanding.*
19. *SEM, The Second Joint Conference on Lexical and Computational Semantics. Atlanta, GA, June 2013. A Panel Presentation on *Extended Semantic Role Labeling*.
20. Google, New York, NY, August 2013. *Better Natural Language Analysis and Amortized Integer Linear Programming.*
21. IBM Research, White Plains, NY, February 2013. *Making Sense of and Trusting, Unstructured Data.*
22. West Point Military Academy, Network Science Center, West Point, NY, February 2013. *Making Sense of, and Trusting, Unstructured Data.*

23. New York City Natural Language Processing Seminar, City University of NY, NYC, NY, January 2013. *Constrained Conditional Models: Integer Linear Programming Formulations for Natural Language Understanding.*
24. Health Informatics Technology Center, Workshop at the University of Illinois, Champaign, IL, November 2012. *Constraints Driven Information Extraction in the Medical Domain.*
25. Johns Hopkins University, The Center for Language and Speech Processing, Baltimore, MD, September 2012. *Constrained Conditional Models: Integer Linear Programming Formulations for Natural Language Understanding*.
26. University of Illinois Technology Showcase, Champaign, IL, April 2012. *Making Sense of Unstructured Data.*
27. Illinois Informatics Institute Lecture Series, Champaign, IL, March 2012. *Making Sense of Unstructured Data.*
28. University of Colorado, Boulder, CO, March 2012. *Learning from Natural Instructions*.
29. Princeton Plasma Physics Laboratory, Princeton, NJ, February 2012. *Learning and Reasoning for Natural Language Understanding.*

8.0 LIST OF ACRONYMS AND ABBREVIATIONS

AAAI	Association for the Advancement of Artificial Intelligence (Conference)
ACL	(Annual Meeting of the) Association for Computational Linguistics
ADEPT	All-Domain Execution and Planning Technology
COLING	International Conference on Computational Linguistics
CoNLL	SIGNLL Conference on Computational Natural Language Learning
EACL	European chapter of the Association for Computational Linguistics (Conference)
EMNLP	Conference on Empirical Methods in Natural Language Processing (Proceedings)
ICML	International Conference on Machine Learning
IJCAI	International Joint Conference on Artificial Intelligence
ILP	Integer-Linear-Programming
JOINCA	Joint Inference Component-supporting Architecture
KRR	Knowledge Representation and Reasoning
ML	Machine Learning
NAACL	North American Chapter of the ACL (Conference)
NER	Named Entity Recognition
NIPS	Neural Information Processing Systems Conference
NLP	Natural Language Processing
RTE	Recognizing Textual Entailment
SemLM	Semantic Language Model
SIGNLL	ACL's Special Interest Group on Natural Language Learning
SRL	Semantic Role Labeling
TAC	Text Analysis Conference
TACL	Transactions of the Association for Computational Linguistics