

Solving Hard Coreference Problems *Haoruo Peng, Daniel Khashabi and Dan Roth***Summary:**

Although much research is going on in Coreference resolution it is still an unsolved problem due to the ambiguous nature of the Natural Language understanding. Existing works performs poorly on pronouns, specifically when gender or plurality information doesn't help. In this paper, authors aim to improve coreference resolution by addressing these problems. In many sentences we cannot resolve the pronouns based on gender or plurality information, many works like (Rahman and Ng 2012), (Levesque et al. 2011), (Chang et al., 2013) was able to achieve this with 73.05% precision. Authors believe that addressing these hard problems requires significant amounts of background knowledge along with an inference paradigm that can make use of this background knowledge in supporting the coreference decision making. The paper is concerned about following questions:

- How can we represent background knowledge and how can we infer it for coreference decision making?

The knowledge required is represented based on key predicates in the sentence called Predicate Schemas. In this paper, authors develop the notion of Predicate Schemas, instantiate them with automatically acquired knowledge, and show how to compile it into constraints that are used to resolve coreference within a general Integer Linear Programming (ILP) driven approach to coreference resolution. Specifically, authors are interested in two types of Predicate Schemas first specifies one predicate with its subject and object, it provides information of subject and object preferences of a given predicate. The second specifies two predicates with a semantically shared either subject or predicate, it gives the information about the role preferences of one predicate among other. Authors instantiate these schemas by acquiring statistics in an unsupervised way from multiple resources including the Gigaword corpus, Wikipedia, Web Queries and polarity information. The main obstacle for this type of instantiation is that however, these resources can provide relatively strong evidence that loses impact in standard training due to sparsity. To overcome this author, compile Predicate Schemas knowledge automatically at inference time and made use of the ILP driven framework to make decisions. The main contributions of this paper can be summarized are 1. Authors proposed the Predicate Schemas representation and study two specific schemas. 2. Authors showed that in a given context, Predicate Schemas can be automatically compiled into constraints and we could infer from it. 3. Authors address hard pronoun resolution problems as a standard coreference problem.