

Resolving Complex Cases of Definite Pronouns: The Winograd Schema Challenge *Altaf Rahman and Vincent Ng***Summary:**

A definite pronoun is a pronoun that refers to something specific. So, a personal pronoun is a definite pronoun. For a given sentence human can easily resolve the pronoun with their word knowledge. Whereas state of art coreference resolvers lack world knowledge or inferring that knowledge and they cannot resolve these complex pronouns. In this paper, authors examine the task of resolving complex cases of definite pronouns that appear in the sentence. Specifically, each sentence (1) has two clauses separated by a discourse connective, where the first clause contains two or more candidate antecedents and the second clause contains the target pronoun and (2) the target pronoun agrees in gender, number, and semantic class with each candidate antecedent, but does not have any overlap in content words with any of them. Since these kinds of pronouns don't appear frequently many resolvers doesn't handle these problems. But, Levesque (2011) has argued that the problem of resolving the difficult pronouns in a carefully chosen set of twin sentences, which he refers to as the Winograd Schema Challenge, could serve as a conceptually and practically appealing alternative to the well-known Turing Test. Levesque believes that if any resolver can resolve correctly with high probability then we are a step closer to intelligent machines. Authors created a dataset that contains 941 sentence pairs where each sentence contains exactly two candidate antecedents and the whole dataset is divided into training and test set by 70/30 ratio. The question paper is concerned about are:

- whether we can assign a rank value to the pronoun and the two candidate antecedents such that we can determine which of the two candidate antecedents is the correct antecedent for the target pronoun in each sentence?

The main challenge the way to the solution is getting the world knowledge humans have and inferring from that knowledge, which needs to have a deeper understanding of the text. To overcome this challenge author categorized the linguistic features for resolving difficult pronouns into eight components and inferred text from a variety of sophisticated knowledge sources. The eight different features are derived from (1) Narrative Chains which are partially ordered sets of events centered around a common protagonist, aiming to provide special kind of knowledge for scripts. (2) By using Google authors generate six queries for each sentence and from the counts returned they are creating four features based on few rules. Authors are intending to capture world knowledge like lions rather than zebras are predators. Such knowledge may not be expressed as a relation and hence may not be easily captured using specific patterns. (3) Using Frame Net authors replace the two antecedents with their extracted semantic roles and generate the search queries from resulting sentence (4) Some sentences involves in comparing two antecedents by using Heuristic Polarity authors resolves these kinds of sentences (5) Authors employed Opinion Finder which has a pre-trained classifier for annotating the phrases in a sentence with their contextual polarity values (6) By Connective based relations authors automate human method for resolving pronouns. Semantic compatibility using Google web search has potential problems like precision and recall. (7) To address these potential problems, authors compute knowledge of selectional preferences from a large, unannotated corpus. (8) Authors exploit the coreference-annotated training documents by creating lexical features from them. To tackle this question authors employed a ranking based machine learning approach to combine the features derived from these eight knowledge sources. Results are expressed in terms of accuracy, which is the percentage of correctly resolved target pronouns. our system achieves an accuracy of 73.1%, significantly outperforming the Combined resolver by 17.9 points in accuracy. Error analysis indicates that further gains could be achieved via more accurate sentiment analysis and induction of world knowledge from corpora or the Web.