

Detecting contradiction in Text

Akashdeep Patel,Sachin Yadav
{akashdp,sachinky}@iitk.ac.in

Advisor : Dr. Amitabh Mukerjee
Department of Computer Science and Engineering,
IIT Kanpur, India

October 22, 2013

1 Motivation

Detecting conflicting statements is a foundational text understanding task with applications in information analysis. Contradictions arise from relatively obvious features such as antonymy, negation, or numeric mismatches. They also arise from complex differences in the structure of assertions, discrepancies based on world-knowledge, and lexical contrasts. Topical relation between the documents is also important. In many situations a mismatch in named entities or numerical expressions is a strong clue for a contradiction between texts T1 and T2, given their dependency parse trees are similar. System to detect contradictions needs to be more fine-grained than common systems of entailment.

2 Our Implementation

We will initially try the following approach on for english and later try to do the same for hindi also. We detect contradiction at sentence level. First a filter is set to remove cases of non-coreferent events

- by checking whether the root of the hypothesis graph is aligned in the text graph.
- on the basis of continuum of topicality

Following steps are involved to detect contradiction at sentence level.

- Split complex sentences into simple sentences that contain only one predicate each.
- Calculate the minimum cost alignment for each pair of simple sentences from T1 and T2.
- Calculate the total minimum cost alignment of all simple sentences in T2 with the simple sentences in T1.
- Calculate the degree of lexical contradiction of the aligned chunks (bunsetsus) in the aligned simple sentences.
- Calculate the degree of contradiction between T1 and T2

2.0.1 Dataset

We use the data used by Marneff et al.(2008) for english.For Hindi, we still need to search/create data.

2.0.2 Related Work

Little work has been done in the field on contradiction detection.Till now only english and japnese language have been studied for detecting contradiction.In a paper by Marneff et al.(2008),typology of contradiction was proposed although their approach was not able to give good result in cases where contradiction was not identifiable simply using lexical terms and syntax.While Daniel et al. (2013) follow some of the steps used by Marneff et al.(2003) with a novel approach of cost assignment to the sentences and lexical matching.. The PASCAL Recognizing Textual Entailment (RTE) Challenges (Dagan et al., 2006; Bar-Haim et al., 2006; Giampiccolo et al., 2007) focused on textual inference in any domain. Condoravdi et al. (2003) first recognized the importance of handling entailment and contradiction for text understanding. Harabagiu et al. (2006) provide the first empirical results for contradiction detection. But their focus was on specific kinds of contradiction. The recent work of Deniel Andrade, Massaki Tsuchida and others focus on contradiction detection to level. Marrie Catherine, Anna N. Rafferty and Christopher D. Manning define different types of contradictions and provide results for one.

2.1 References

- M.-C. De Marneffe, A. N. Rafferty, and C. D. Manning. Finding contradictions in text. Proceedings of ACL-08: HLT, pages 1039-1047, 2008
- Detecting Contradiction in Text by Using Lexical Mismatch and Structural Similarity,Daniel Andrade,Masaaki Tsuchida
- K. Zhang and D. Shasha. Simple fast algorithms for the editing distance between trees and related problems. SIAM journal on computing, 18(6):1245-1262, 1989.