Fuzzy Set Theory-Based Belief Processing for Natural Language Texts

Ralf Krestel and **René Witte**
Fakultät für Informatik
Universität Karlsruhe (TH), Germany
krestel|witte@ipd.uka.de

Sabine Bergler
Department of Computer Science and Software Engineering
Concordia University, Montréal, Canada
bergler@cs.concordia.ca

Introduction

The growing number of publicly available information sources makes it impossible for individuals to keep track of all the various opinions on one topic. The goal of our *artificial believer* system[1] we present in this paper is to extract and analyze opinionated statements from newspaper articles.

Beliefs are modeled with a fuzzy-theoretic approach applied after NLP-based information extraction. A fuzzy believer models a human agent, deciding what statements to believe or reject based on different, configurable strategies.

Related Work

Existing systems mainly deal with either the extraction of *opinions* or the identification of *entailment* between two sentences (see the PASCAL challenges (Dagan, Glickman, & Magnini 2005; Bar-Haim *et al.* 2006)). Our work combines these previously separated tasks to solve the complex problem of modeling a human newspaper reader.

The extraction of *reported speech*, together with creating structures based on source and reporter has been the main focus in (Bergler *et al.* 2004). It also incorporates an algorithm to assign credibility to statements according to the reporting verb used. Our system extends this work to process statements in order to (i) find common topics, (ii) determine their compatibility, and (iii) create beliefs based on heuristics, thereby creating a user agent.

System Overview

Our system extracts reported speech out of newspaper articles and generates beliefs for subsequent processing by the system. The limitation to reported speech allows a clear attribution of statements to a *source* and enables the system to infer the credibility of a statement using *evidential analysis* (Bergler 1992). The extracted reported speech structures can then be represented as beliefs attributed to a source with different levels of nesting (Ballim & Wilks 1991).

To identify statements belonging to the same topic, our system extracts predicate-argument structures using the output of parsers. To find the kind of relation between two statements about one topic, we employ fuzzy set theory.

Our fuzzy believer system is implemented in GATE (Cunningham *et al.* 2002). For preprocessing it uses standard components shipped with GATE. The components we developed comprise: *(1)* Reported Speech Extraction, *(2)* Profile Generation, *(3)* Predicate-Argument Extraction, and *(4)* Fuzzy Believing. We describe these components below.

Reported Speech Extraction. To extract opinionated statements, we developed a grammar written in JAPE[2] that recognizes 6 different reported speech patterns grouped around 50 verbs often used in reported speech.

Profile Generation. In the next step, we group the statements extracted in the first step by source entities using coreference chains computed by a fuzzy system (Witte & Bergler 2003). E.g., "The president," "Mr. Reagan," and "he" may refer to the same entity, so we can assign all statements from these sources to one profile.

Predicate-Argument Extraction. The extracted reported speech statements are parsed by both RASP (Briscoe, Carroll, & Watson 2006) and MiniPar (Lin 1998). According to a set of rules, we extract subject/verb/object triples (predicate-argument structures) from the parsers' output.[3]

Fuzzy Believer. This component has to fulfill three tasks: *(1)* Group the statements according to their topics into *domains*, e.g., all statements about "The President knowing about the diversion," see Figure 1. *(2)* Identify the *polarity* of the statements in each domain, i.e., which statements contradict each other. And *(3)* implement *strategies* for what to believe, e.g., trust all newly made statements, hold onto old beliefs, or rely on what the majority believes.

The first task is handled by two heuristics (semantic, based on WordNet[4], and syntactic, based on string similarity) that compare the extracted predicate-argument structures (PAS) of two statements. If the heuristics recognize a similarity degree higher than a given threshold between two statements, they are grouped into one domain (topic).

The second task is solved by using fuzzy set theory and representing all statements as degrees of similarity between the verbs of the statements in one domain. This similarity is again computed using WordNet together with the detection of negations and antonyms.

[1] http://www.ipd.uka.de/~durm/tm/believer/

[2] Java Annotations Patterns Engine, see http://gate.ac.uk/

[3] The triples for the example in Figure 1, extracted from Wall Street Journal articles on the Iran-Contra Affair, are: [president-know-diversion], [president-authorize-diversion], [Reagan-know-diversion], [president-know-diversion], and [Reagan-involved-policy].

[4] WordNet, http://wordnet.princeton.edu

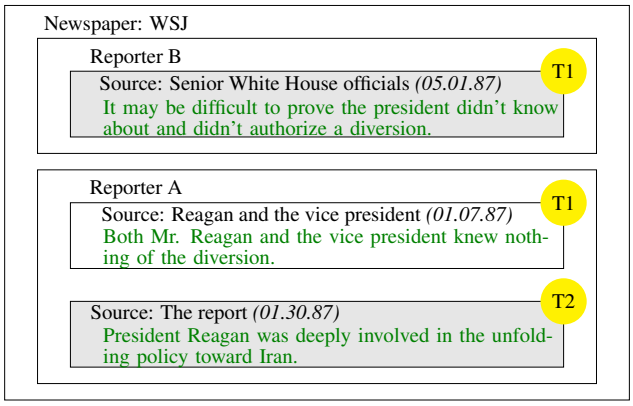

Figure 1: System's beliefs based on majority strategy

For the third task we use three fuzzy set operations (*Union*, *Expansion*, and *Revision*, see (Witte 2002)) to model various belief strategies: believe (1) all, (2) all new, (3) old statements, (4) certain reporter/source/newspaper, (5) majority, or (6) weighted majority—combining 4 & 5.

System Output. In Figure 1, we see the LaTeX-output generated by our system. The sentences in the inner boxes show the extracted reported speech statements that are grouped according to their source, reporter, and newspaper. The circles in the top right of each box show the *id* of the corresponding domain. The statements with gray background are believed by the system employing the *majority* strategy.

Evaluation

So far, we performed a detailed evaluation of the individual components of our system.

Reported Speech Extraction. In order to evaluate the reported speech extraction component, we randomly picked 7 newspaper articles from the Wall Street Journal corpus. The articles contain about 400 sentences (∼6100 words), among them 133 reported speech constructs. For the detection of reporting verb and source, our system achieved a recall value of 0.83 and a precision value of 0.98. This results in an F-measure of 0.90.

Domain Finding. The evaluation of the domain finding component includes the comparison of the results obtained with RASP, MiniPar, and manually annotated predicate-argument structures. The test data we use is taken from the MSR[5] corpus and comprises 116 paraphrase pairs. We treat all sentences as content of a reported speech construct. The best result for recall is 81% and best precision value obtained 52% with a different configuration. The rather shallow semantic approach sets a practical limit to the achievable results. This can be inferred by comparing the numbers obtained using manually parsed predicate-argument structures with the numbers obtained from the parsers. It shows that there is space for improvement on the side of the parsers as well as on the side of the PAS extractor. But a precision of 55% and a recall of 85%, as obtained for the best configuration of the system using manually parsed PASs, shows that it needs more and/or better heuristics to obtain significant improvements.

Polarity Finding. The data that comes closest to the conditions we need are the entailment pairs of the PASCAL challenge corpus (Bar-Haim *et al.* 2006). We tested different configurations and computed accuracy for two settings. For one experiment, we included all results in the evaluation counting the entailment pairs that were not grouped into the same domain by the domain classification as non-entailing. Here, the best results were around 55% accuracy. The other test setting only considered the sentence pairs that were actually grouped into the same domain by the domain classification component yielding an accuracy of 58% using MiniPar extracted PASs. More elaborate heuristics could probably increase these numbers, comparable to the PASCAL challenge, where participating systems also started with around 50% accuracy, but improved over the years.

Conclusions and Future Work

We developed an artificial believer system that can be applied in different scenarios: *(1)* companies evaluating product reviews on web sites or blogs, *(2)* governmental organizations interested in dispositions of people, or *(3)*, as we demonstrated here, assist individuals in news analysis.

Apart from the evaluation described above, tests of the system on actual newspaper articles showed accepted and rejected beliefs that reflect the desired results. Embedding the system within an Internet agent and measuring its effectiveness for a real user will be the next major step.

References

Ballim, A., and Wilks, Y. 1991. *Artificial Believers: The Ascription of Belief*. Lawrence Erlbaum Associates, Inc.

Bar-Haim, R.; Dagan, I.; Dolan, B.; Ferro, L.; Giampiccolo, D.; Magnini, B.; and Szpektor, I. 2006. The Second PASCAL Recognising Textual Entailment Challenge. In *Proc. of the Second PASCAL Challenges Workshop on Recognising Textual Entailment*.

Bergler, S.; Doandes, M.; Gerard, C.; and Witte, R. 2004. Attributions. Technical Report SS-04-07, 16–19. Stanford, CA, USA: AAAI Press.

Bergler, S. 1992. *The Evidential Analysis of Reported Speech*. Ph.D. Dissertation, Brandeis University, Massachusetts, USA.

Briscoe, E.; Carroll, J.; and Watson, R. 2006. The Second Release of the RASP System. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*.

Cunningham, H.; Maynard, D.; Bontcheva, K.; and Tablan, V. 2002. GATE: A framework and graphical development environment for robust NLP tools and applications. In *Proc. of the 40th Anniversary Meeting of the ACL*.

Dagan, I.; Glickman, O.; and Magnini, B. 2005. The PASCAL Recognising Textual Entailment Challenge. In *Proc. of the PASCAL Challenges Workshop on Recognising Textual Entailment*.

Lin, D. 1998. Dependency Based Evaluation of MINIPAR. In *Proc. of the Workshop on the Evaluation of Parsing Systems, First Intl. Conf. on Language Resources and Evaluation*.

Witte, R., and Bergler, S. 2003. Fuzzy Coreference Resolution for Summarization. In *Proc. of 2003 Intl. Symposium on Reference Resolution and Its Applications to Question Answering and Summarization (ARQAS)*, 43–50.

Witte, R. 2002. Fuzzy Belief Revision. In *9th Intl. Workshop on Non-Monotonic Reasoning (NMR'02)*, 311–320.

[5] http://research.microsoft.com/research/nlp/msr_paraphrase.htm