# Chapter #

## Low-Level Annotations and Summary Representations of Opinions for Multi-Perspective Question Answering

Claire Cardie
Department of Computer Science
Cornell University
Ithaca, NY 14850, USA
cardie@cs.cornell.edu

Janyce Wiebe, Theresa Wilson, Diane Litman
Department of Computer Science
Intelligent Systems Program
University of Pittsburgh
Pittsburgh, PA  15260, USA
wiebe@cs.pitt.edu

### Abstract[1]

While much recent progress has been made in research on fact-based question answering, our work aims to extend question-answering research in a different direction ─ to handle multi-perspective question-answering tasks, i.e. question-answering tasks that require an ability to find and organize opinions in text. This paper proposes an approach to multi-perspective question answering that views the task as one of opinion-oriented information extraction. We describe an annotation scheme developed for the low-level representation of opinions, and present results of interannotator agreement studies. Next, we propose the use of opinion-oriented scenario templates to act as a summary representation of the opinions expressed in a document, a set of documents, or an arbitrary text segment.  Finally, we outline an approach for the automatic construction of opinion-based summary representations.

## 1.  Introduction

Current research in question answering focuses primarily on the development of methods for finding short answers to factual questions such as, *Where is the Susquehanna River?* and *When was Derek Jeter born?*  While much progress has been made in recent years in fact-based question answering (e.g. Voorhees (2001)), our work aims to extend question-answering (QA) methods to handle *multi-perspective question-answering* tasks. We are interested in developing

techniques to support the answering of opinion-based questions of the following sort:

- What is the general opinion from the African press towards the recent presidential election in Zimbabwe?

- What was the world-wide reaction to the 2001 annual U.S. report on human rights?

- Has there been any change in the official opinion from China towards the 2001 annual U.S. report on human rights since its release?

It should be clear from these questions that the ability to find and organize opinions in on-line text is of critical importance in building successful multi-perspective question-answering (MPQA) systems. In response, this paper proposes an approach to multi-perspective question answering that views the task as one of *opinion-oriented information extraction*.

Very generally, an information extraction (IE) system takes as input an unrestricted text and "summarizes" the text with respect to a topic or domain of interest: it finds useful information about the domain and encodes that information in a structured form that is suitable for populating databases (e.g. Lehnert and Sundheim (1991), Chinchor *et al.* (1993)). In the domain of natural disasters, for example, an IE system extracts, for each disaster event,the type of disaster, the date and time it occurred, any objects damaged or destroyed, and so on. Event-based IE systems typically operate in two stages. First, the system identifies low-level *template relations* that encode domain-specific facts as expressed in the text (e.g., that the disaster, "a tornado", occurred on a particular day, "today", and so on). Next, the system merges the extracted template relations into a *scenario template* that summarizes the entire event.

In contrast, we hypothesize that an information extraction system for MPQA tasks should rely on a set of opinion-oriented template relations that identify each expression of opinion in a text along with its source (i.e., the agent expressing the opinion), its type (e.g., positive, negative, uncertain), and its strength (e.g., strong, weak). Once identified, these low-level relations can then be combined to create an opinion-based scenario template — a summary representation of the opinions expressed in a document, a group of documents,

or an arbitrary text span. This summary representation, in turn, acts as the primary knowledge source that supports a variety of MPQA tasks.

In the sections that follow, we describe the template relation annotation scheme developed for the low-level representation of opinions and present the results of interannotator agreement studies We then motivate the need for summary representations of opinion and propose a method for their automatic construction using techniques from information extraction with a discourse-based approach to "perspective segmentation." We conclude with a discussion of the potential uses of summary representations of opinions for a variety of MPQA tasks.

## 2. Low-level Opinion Annotations

In this section, we present a detailed scheme for annotating expressions of opinions, beliefs, emotions, speculations and other private states in newspaper articles. *Private state* is a general term covering mental and emotional states that cannot be directly observed or verified (Quirk et al., 1985). Private states are expressed using *subjective* language (Banfield, 1982; Wiebe, 1994).

The annotation scheme presented here was developed as part of a U.S. government-sponsored project (ARDA AQUAINT NRRC) to investigate multiple perspectives in question answering (Wiebe et al., 2003). We implemented the scheme in GATE, a General Architecture for Text Engineering (Cunningham et al., 2002). General Instructions for annotating opinions and specific instructions for downloading and using GATE to perform the annotations are available at http://www.cs.pitt.edu/~wiebe.

### 2.1 Expressing Private States in Text

**Private States, Speech Events, and Expressive Subjective Elements.** There are two main ways that private states are expressed in language. Private states may be explicitly mentioned, or they may be expressed indirectly through the way things are described or the particular wording a speaker or writer uses. An example of an explicitly mentioned private state is "frustrated" in sentence (1).

(1) Western countries were left frustrated and impotent after Robert Mugabe formally declared that he had overwhelmingly won Zimbabwe's presidential election.

The second way that private states are expressed is indirectly using *expressive subjective elements* (Banfield, 1982). For example, the private states in sentences (2) and (3) are expressed entirely by the wording and the style of language that is used.

    (2) The time has come, gentlemen, for Sharon, the assassin, to realize that injustice cannot last long.

    (3) "We foresaw electoral fraud but not daylight robbery," Tsvangirai said.

In (2), the phrases "The time has come," "gentlemen," "the assassin," and "injustice cannot last long" express the negative attitude of the writer. Similarly, in (3), "fraud" and "daylight robbery" reflect the anger of the speaker, Tsvangirai. These phrases are examples of expressive subjective elements.

When looking for opinions and other private states in text, an annotator must consider speech events as well as explicitly mentioned private states. In this work, we use *speech event* to refer to any event of speaking or writing. However, the mere presence of a speech event does not indicate a private state. Both sentences (3) above and (4) below contain speech events indicated by "said". Sentence (3) is opinionated, while in (4) the information is presented as factual. For speech terms such as "said," "added," "told," "announce," and "report," an annotator determines if a private state is present mainly by looking inside the scope of the speech term for expressive subjective elements.

    (4) Medical Department head Dr. Hamid Saeed said the patient's blood had been sent to the Institute for Virology in Johannesburg for analysis.

**Nested Sources**  The *source* of a speech event is the speaker or writer. The *source* of a private state is the experiencer of the private state, i.e. the person whose opinion or emotion is being expressed. The writer of an article is a source, because he or she wrote the sentences composing the article, but the writer may also write about other people's private states and speech events, resulting in multiple sources in a single sentence. For example, each of the following sentences has two sources: the writer of the sentences, and Sue, who is the source of a speech event in (5) and of private state in (6).

    (5) Sue said, "The election was fair."

    (6) Sue thinks that the election was fair.

Note, however, that we don't really know what Sue says or thinks. All we know is what the writer tells us. Sentence (6), for example, does not directly present Sue's speech event but rather Sue's speech event according to the writer. Thus, we have a natural nesting of sources in a sentence. In (5) and (6), the nested source of "said," and "thinks," is <writer, Sue>. Expressive subjective elements may also have nested sources. In sentence (3), "fraud" and "daylight robbery" are expressive subjective elements, both with nested source <writer, Tsvangirai>.

## 2.2 Annotating Private States in Text

For this work, annotators are asked to mark the word spans in each sentence that indicate speech events or private states. Specifically, there are two main types of annotations: 1) an annotation type for explicit private states and speech events, and 2) an annotation type for expressive subjective elements. An annotator marks a span of a given type, and then sets a number of attributes to characterize the annotation. The attributes for each annotation type, listed in Table 1, are described Wilson and Wiebe (2003). The remainder of this section focuses on one particular attribute that is marked on private state and speech event annotations, the *onlyfactive* attribute.

| Explicit private states/speech events | Expressive subjective elements |
|---|---|
| nested-source<br>onlyfactive: *yes, no*<br>overall-strength: *low, medium, high, extreme*<br>on-strength: *neutral, low, medium, high, extreme*<br>attitude-type*: *positive, negative, both*<br>attitude-toward*<br>is-implicit<br>Minor | nested-source<br>strength: *low, medium, high, extreme*<br>attitude-type*: *positive, negative, other* |

Table 1 Attributes for the two main annotation types. For attributes that take on one of a fixed set of values, the set of possible values are given. *Exploratory attributes.

The *onlyfactive* attribute is the core of the annotation scheme. This attribute is used to indicate whether the source of the private state or speech event is indeed expressing an emotion, opinion, or other private state. Any expression that is an explicit private state (e.g. "think", "hope," "want") or a private state mixed with speech (e.g. "berate," "praise") is *onlyfactive=no*; however, *neutral* speech

events (e.g. "said," "added") may be either *onlyfactive=yes* or *onlyfactive=no*, depending on their contents. For example, the annotation for "said" in sentence (3) would be marked *onlyfactive=no*, but the annotation for "said" in sentence (4) would be marked *onlyfactive=yes*. Note that, even if *onlyfactive=no*, the sentence may express something the nested source believes is factual. Consider the sentence "John criticized Mary for smoking." John expresses a private state (his negative evaluation of Mary's smoking). However, this does not mean that John does not believe that Mary smokes.

We performed an annotation study to measure agreement for the *onlyfactive* attribute. The data in the study consists of English-language versions of foreign news documents from a variety of publications and countries. Three annotators (A, M, and S) independently annotated 13 documents with a total of 210 sentences. The articles are from a variety of topics and were selected so that 1/3 of the sentences are from news articles reporting on objective topics (objective articles), 1/3 of the sentences are from news articles reporting on opinionated topics ("hot topic" articles), and 1/3 of the sentences are from editorials.

We measure agreement for the *onlyfactive* attribute by looking at the set of private state/speech event span annotations on which the annotators agree (82%). The Kappa agreement scores for all annotator pairs are given in Table 2. For the *onlyfactive* judgments, i.e. whether or not an opinion or other private state is being expressed, the annotators have an average pairwise Kappa of 0.81. Under Krippendorf's (1980) scale, this allows for definite conclusions.

| | All Expressions | | All Sentences | |
|---|---|---|---|---|
| | Kappa | agree | Kappa | agree |
| A&M | 0.84 | 0.91 | 0.75 | 0.89 |
| A & S | 0.84 | 0.92 | 0.84 | 0.94 |
| M& S | 0.74 | 0.87 | 0.72 | 0.88 |

Table 2  Pairwise Kappa scores and overall percent agreement for *onlyfactive* judgments and derived sentence level judgments.

For many applications, it is useful to be able to determine which sentences, not just which expressions, are subjective. We define sentence-level classifications in terms of our lower-level annotations as follows. If an annotator marked one or more *onlyfactive=no* expressions in the sentence, we consider the annotator to

have judged the sentence to be subjective. Otherwise, we consider the annotator to have judged the sentence to be objective. The pairwise agreement results for these derived sentence-level annotations are given on the right-hand side of Table 2. The average pairwise Kappa for sentence-level agreement is 0.77. Note that the agreement is lower than that for onlyfactive judgments because explicit private-state and speech event expressions upon which the annotators did not agree are now included.

## 3. Summary Representation of Opinions

We expect the opinion-based annotation scheme described above to support a wide variety of end-to-end applications in the area of multi-perspective question answering. In addition, we anticipate the need to go create **summary representations of opinions** that would provide concise, and ultimately user-tailored summaries of the opinions expressed in an article, in a set of articles, or in any arbitrary segment of text.

We propose to view these summary representations as information extraction (IE) scenario templates. Rather than traditional event-oriented IE scenario templates, ours will instead summarize opinion-oriented scenarios to accommodate the MPQA task. We postulate that methods from information extraction (see Cardie (1997)) will be largely adequate for the automatic creation of opinion-based summary representations. Machine learning methods can be applied to acquire lexico-syntactic patterns for the automatic extraction of low-level opinion relations. The low-level expressions of opinion can then be combined to create the opinion-based scenario template summary representation.

In the subsections below, we first provide a concrete example of an MPQA summary representation for a portion of one opinion-oriented article. We then describe a number of issues in the automatic creation of summary representations. This is work in progress; we do not present empirical results.

### 3.1 An Example

An MPQA scenario template is meant to encode a summary of the opinions expressed throughout one or more texts or text spans. They are summaries in that they merge and make inferences from the low-level opinion annotations that

have been identified in the text. As an example, consider the text in Figure 1, which is the first ten sentences of one document (#20.20.10-3414) from the Human Rights portion of our MPQA annotated corpus. Given the first sentence of the document,

> The Annual Human Rights Report of the US State Department has been strongly criticized and condemned by many countries,

our MPQA system should produce the following low-level opinion annotations:

> PRIVATE STATE/ SPEECH EVENT: nested-source= <**writer**>, *onlyfactive=yes*

> EXPRESSIVE-SUBJ: nested-source =<**writer**>, *strength=medium*.

In particular, from the writer's perspective, the sentence can be classified as *onlyfactive=yes*. In addition, the lexical cue "strongly" indicates some (*medium*) amount of EXPRESSIVE SUBJectivity.

A similar analysis of the remainder of the text fragment would produce the low-level annotations of Figure 1. It should be clear that the representation of opinions at this level is difficult for humans to absorb. One possible MPQA



Figure 1 MPQA sample text. First ten sentences from document #20.20.10-3414 from the Human Rights portion of the MPQA collection along with its associated low-level opinion annotations. Annotations are shown in roughly the same position as the original text spans from which they were derived.

summary representation for the sample text is shown in Figure 2. The summary indicates that there are three primary opinion-expressing sources in the text — <**writer**>; <**writer**, **many countries**> and <**writer**, Chinese **info**rmation **office**>. Furthermore, these sources expressed the following opinions:

- the writer expressed a negative attitude (of medium strength) towards the human rights report;

- the writer expressed a mildly negative attitude towards the United States;

- according to the writer, many countries expressed a strongly negative attitude towards the human rights report;

- according to the writer, the Chinese information office expressed a negative attitude (of medium strength) towards the United States.

Although expressed in graphical form in Figure 2, the summary representation is equivalent to the relational scenario template representation used in traditional IE systems.

Acquiring portions of the summary representation requires making inferences across related groups of the low-level annotations. One example is the association of a *strength* with scenario template *attitude* relations that cover multiple low-level *attitude-type* annotations. In addition, the *subjectivity index* associated with each nested source illustrates the kind of summary statistic that might be inferred from a collection of low-level opinion annotations. Here, the subjectivity index indicates that four out of ten sentences of the **writer** include subjective language, and that all propositions attributed to **many countries** and the Chinese **info**rmation **office** include subjective content. Summary representations can be encoded as in-line annotations (e.g. XML or SGML),
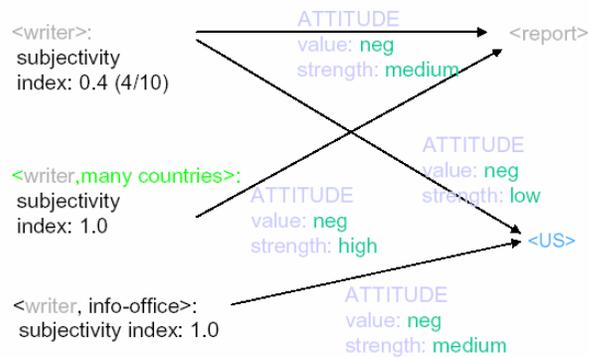


Figure 3: MPQA summary representation for the text sample from document #20.20.10.3414.

allowing for direct querying by the end-user.

Finally, there are many user-specified options for generation of summary representations. The user might want summaries that focus only on particular agents, particular classes of agents, particular attitude types or attitude strengths. The user might also want to specify a particular level of nested source, e.g. creating the summary from the point of view of only the most nested sources.

### 3.2 Automatic Creation of Summary Representations of Opinion

In the paragraphs below, we discuss some of the issues involved in the automatic creation of MPQA summary representations.

**Perfect low-level annotations**   Even given a complete and accurate set of low-level opinion annotations, building a summary representation will still be non-trivial. In particular, the MPQA system would need an accurate noun phrase coreference subsystem to identify the various terms and phrases used to refer to each opinion source and each object towards which some opinion was expressed. In particular, the object of an *attitude-towards* relation is often very difficult for even human readers to determine and is often not explicitly expressed in the text. Another issue in creating summary representations from perfect low-level annotations is how to handle the presence of conflicting opinions from the same source.

**Imperfect low-level annotations**   The situation becomes much harder, of course, when the MPQA summary representation is to be built on top of automatically generated low-level opinion annotations, which are likely to be incomplete and inaccurate.   The situation will be akin to the information extraction task of "merging" automatically extracted and errorful template relations into a scenario template. The noun phrase coreference system will continue to be important in this situation — it will need to provide accurate links between coreferent sources as well as coreferent objects. On the other hand, since our goal is to derive a summary representation, we may be able to use redundancy in the expression of opinions to discard incorrect low-level annotations.  For example, if the low-level opinion annotations associated with one source indicate that eight out of ten expressions from the source are negative

evaluations of object X, then the system might be able to discard the remaining two positive evaluation of X as errors.

**Cross-document coreference** Effective MPQA will require collation of information across documents, since the summary representation may span multiple documents. For example, if a user wants to know the range of perspectives on topic X, then the system will need to perform cross-document coreference w.r.t. topic X as well as w.r.t. the various agents that express views on the topic.

**Segmentation and perspective coherence** Accurate grouping of the low-level annotations — e.g., according to source, topic, negative or positive attitude — will be critical in deriving usable summary representations of opinion. For this task, we believe that it will be important to develop text segmentation methods based on a new notion of perspective coherence. In the natural language processing literature, the term *segmentation* refers to breaking up a document into smaller chunks — or segments — that are locally coherent. Depending on factors such as corpus type and application need, different notions of coherence have been proposed as the basis of segmentation. In the area of information retrieval, for example, text segmentation has usually been based on semantic coherence. Segmentation is performed by placing segment boundaries at points of semantic discontinuity (Hearst, 1997). In the area of discourse analysis, segmentation has instead been based on notions of informational (Hobbs, 1979), and/or intentional coherence (Grosz and Sidner, 1986; Passonneau and Litman, 1993).

To incorporate text segmentation for the derivation of MPQA summary representations, we propose extending our opinion annotation scheme to denote "perspective segments" that will identify sentence spans expressing coherent perspectives. As with other notions of segmentation, perspective segmentation will likely involve merging and performing shallow inferences across sentences.

Consider an example segment produced during an informal manual clustering study we performed as part of the summer workshop. Workshop participants were asked to label opinions, where each opinion could be described by a single sentence, or by a segment consisting of a sentence span. The excerpt in Figure 3 illustrates a sample segmentation from our coding exercise. In particular, four

out of seven coders placed sentences 3-8 through 3-10 in the same segment; a fifth coder placed the beginning of this segment one sentence earlier.

---

**3-6** Mugabe described the opposition as "donkey being controlled by the British," the former colonial power.  (*nested-source*=<**writer**, **Mugabe**>, *onlyfactive=no*)

**SEGMENT BEGIN (1 CODER)**
**3-7** The fledgling MDC won 57 of 120 elected seats in June 2000 parliamentary elections as Mugabe's popularity plunged amid economic devastation and chaos.
(*nested-source*=<**writer**>, *onlyfactive=no*)

**SEGMENT BEGIN (4 CODERS)**
**3-8** The U.S. State Department released a human rights report on Zimbabwe Monday that accused the government of extra-judicial killings, undermining the independence of the judiciary and waging a "systematic campaign of violence targeting supporters and potential supporters of the opposition."  (*nested-source*=<**writer**, **US report**>, *onlyfactive=no*)

**3-9** Security forces tortured opponents, ruling party militants abducted people, and police arrested opposition supporters who were themselves the victims of crimes. **3-10** Freedom of the press and freedom of assembly were also severely restricted, the report said."
(*nested-source*=<**writer**, **US report**>, *onlyfactive=no*)

**SEGMENT END (5 CODERS)**
**3-11** In his speech on Monday, Mugabe thanked African leaders for refusing to buckle to pressure to suspend Zimbabwe from the Commonwealth of Britain and its former territories at a summit of the 54-nation grouping in Australia. (*nested-source*=<**writer**, **Mugabe**>, *onlyfactive=no*) (*nested-source*=<**writer**, **Mugabe**, **African leaders**>, *onlyfactive=no*)

---

Figure 3: An Example Document Excerpt with Human **Segmentations** and ANNOTATIONS.

The segment consisting of sentences 3-8 through 3-10 seems to illustrate one way in which perspective coherence can be defined in terms of the sentence-level annotations: merge sentences into a segment when a single source. Linguistic phenomena such as the use of "his" in 3-11 to refer to Mugabe (who was most recently mentioned outside the segment) lends support this  analysis.

Sentence 3-7 provides an interesting borderline case, as one coder also included this sentence in the segment.  Sentences 3-8 through 3-10 can be seen as providing evidence for the expressive-subjective element "chaos."  A more sophisticated notion of perspective coherence might be to cluster evidence

together then include it with the sentence(s) expressing the opinion that the evidence supports (sentence 3-7).

In terms of aiding the creation of higher-level summary representations of opinions, an MPQA system might use segmentation information to ignore the presence of factive sentences that are providing evidence for an opinion when trying to merge a sequence of opinionated sentences into a larger segment.

## 4. Uses for Summary Representations of Opinion in MPQA

In addition to providing a short, optionally tailored summary of the opinions expressed in texts, we anticipate other uses for MPQA summary representations.

**Direct querying**    When the summary representation is stored as a set of document annotations, it can be directly queried directly by the end-user using XML "grep" utilities.

**Collective perspectives**   The summary representations can be used to describe the collective perspective w.r.t. some issue or object presented in an individual article, or across a set of articles.

**User-specified views**    The summary representations can be tailored to match (some types of) user-specified constraints, e.g., to summarize documents from the perspective of a particular writer, individual, government, ideology, or news service, or to summarize documents w.r.t. a particular issue or object.

**Perspective profiles**   The MPQA summary  representation  would be the basis for creating a perspective "profile" for specific sources/agents, groups, news sources, etc.  The profiles, in turn, would serve as the basis for detecting changes in the opinion of agents, groups, countries, etc. over time.

**Credibility assessment**   Credible news reports might be distinguished from less credible reports that are laden with extremely subjective language.

**Debugging**   Because the summary representation is more readable than the low-level annotations, summary representations can be used to aid debugging of the low-level opinion annotations on which they were based.

**Gold Standard "answer keys"** Once the gold standard for the low-level opinion annotations has been created for a collection, however, it might be possible to largely automate the creation of gold standards for various MPQA summary representations.

# 5. Conclusions

This paper proposed an information-extraction approach to finding and organizing opinions in naturally occurring text as a means for supporting multi-perspective question answering. We first presented a low-level annotation scheme for representing opinion-based "template relations", i.e., localized and individual expressions of opinions, and described the results of inter-annotator agreement studies.We next proposed one possible summary representation for concisely encoding the collection of opinions expressed in a document, a set of documents, or an arbitrary text segment. These are akin to the scenario templates produced by event-based information extraction systems. We hypothesize that information extraction techniques can be used to identify the low-level opinion annotations automatically, and that existing methods from information extraction for merging template relations can be extended via perspective segmentation for use in multi-perspective question answering. We concluded with a brief discussion of how opinion-based summary representations might be used to support a variety of multi-perspective question answering tasks.

## 6. References

Banfield, A. 1982. *Unspeakable Sentences*. Routledge and Kegan Paul, Boston.

Cardie, C. 1997. Empirical Methods in Information Extraction. *AI Magazine*, 18(4):65–79.

Chinchor, N., Hirschman, L., and Lewis, D. 1993. Evaluating Message Understanding Systems: An Analysis of the Third Message Understanding Conference (MUC-3). *Computational Linguistics*, 19(3):409–449.

Cunningham, H., Maynard, D., Bontcheva, K., and Tablan, V. 2002. GATE: A framework and graphical development environment for robust nlp tools and applications. *Proc. of the 40th Annual Meeting of the Association for Computational Linguistics*.

Grosz, B. and Sidner, C. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

Hearst, M. 1997. Texttiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics*, 23(1):33–64.

Hobbs, J. 1979. Coherence and coreference. *Cognitive Science*, 3:67–90.

Krippendorf, K. 1980. *Content Analysis: An Introduction to its Methodology*. Sage Publications, Beverly Hills.

Lehnert, W. and Sundheim, B. 1991. A performance evaluation of text analysis technologies. *AI Magazine*, 12(3):81–94.

Passonneau, R. and Litman, D. 1993. Intention-based segmentation: Human reliability and correlation with linguistic clues. *Proceedings of 31st Annual Meeting of the Association for Computational Linguistics (ACL-93)*. Association for Computational Linguistics. 148–155.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. 1985. *A Comprehensive Grammar of the English Language*. Longman, New York.

Voorhees, E. 2001. Overview of the TREC-9 Question Answering Track. E. Voorhees, editor, *Proceedings of the Ninth Text Retrieval Conference TREC 9*. NIST Special Publication. 500-249.

Wiebe, J., Breck, E., Buckley, C., Cardie, C., Davis, P., Fraser, B., Litman, D., Pierce, D., Riloff, E., Wilson, T., Day, D., and Maybury, M. 2003. Recognizing and organizing opinions expressed in the world press. *Working Notes - New Directions in Question Answering (AAAI Spring Symposium Series)*.

Wiebe, J.. 1994. Tracking point of view in narrative. *Computational Linguistics*, 20(2):233–287.

Wilson, T. and Wiebe, J. 2003. Annotating opinions in the world press. *Proceedings of 4th Workshop on Discourse and Dialogue (SIGDIAL093)*. ACL SIGdial.