A Conceptual Framework for Inferring Implicatures

Janyce Wiebe
Department of Computer Science
University of Pittsburgh
wiebe@cs.pitt.edu

Lingjia Deng
Intelligent Systems Program
University of Pittsburgh
lid29@pitt.edu

Abstract

While previous sentiment analysis research has concentrated on the interpretation of explicitly stated opinions and attitudes, this work addresses a type of opinion implicature (i.e., opinion-oriented default inference) in real-world text. This work describes a rule-based conceptual framework for representing and analyzing opinion implicatures. In the course of understanding implicatures, the system recognizes implicit sentiments (and beliefs) toward various events and entities in the sentence, often of mixed polarities; thus, it produces a richer interpretation than is typical in opinion analysis.

1 Introduction

This paper is a brief introduction to a framework we have developed for sentiment inference (Wiebe and Deng, 2014). Overall, the goal of this work is to make progress toward a deeper automatic interpretation of opinionated language by developing computational models for the representation and interpretation of opinion implicature (i.e., opinion-oriented default inference) in language. In this paper, we feature a rule-based implementation of a conceptual framework for opinion implicatures, specifically implicatures that arise in the presence of explicit sentiments, and events that positively or negatively affect entities (goodFor/badFor events). To eliminate interference introduced by the noisy output of automatic NLP components, the system takes as input manually annotated explicit-sentiment and event information, and makes inferences based on that input information. Thus, the purpose of this work is to provide a conceptual understanding of (a type of) opinion implicature, to provide a blueprint for realizing fully automatic systems in the future.

Below, we give terminology, overview the rule-based system, and then present the rule schemas. Finally, via discussion of an example from the MPQA opinion-annotated corpus (Wiebe et al., 2005), we illustrate the potential of the framework for recognizing implicit sentiments and writer-level sentiments that are not anchored on clear sentiment words, and for capturing interdependencies among explicit and implicit sentiments.

We have developed a graph-based computational model implementing some rules introduced below (Deng and Wiebe, 2014). Moreover, in ongoing work, we have proposed an optimization framework to jointly extract and resolve the input ambiguities.

2 Terminology

The building blocks of our opinion implicature framework are subjectivity, inferred private states, and benefactive/malefactive events and states.

Subjectivity. Following (Wiebe et al., 2005; Wiebe, 1994), subjectivity is defined as the expression of private states in language, where private states are mental and emotional states such as speculations, sentiments, and beliefs (Quirk et al., 1985). Subjective expressions (i.e., opinions) have sources (or holders): the entity or entities whose private states are being expressed. Again following (Wiebe et al., 2005; Wiebe, 1994), a private state is an attitude held by a source toward (optionally) a target. Sentiment and belief are types of attitudes. Subjectivity is the linguistic expression of private states. Subjectivity is a pragmatic notion: as the sentence is interpreted in context, a private state is attributed to a source in that context (Banfield, 1982). By sentiment expression or explicit sentiment, we mean a subjective expression where the attitude type of the expressed
private state is sentiment.

There are many types of linguistic clues that contribute to recognizing subjective expressions (Wiebe, 1994). In the clearest case, some word senses give rise to subjectivity whenever they are used in discourse (Wiebe and Mihalcea, 2006). Other clues are not as definitive. For example, researchers in NLP have begun to develop lexicons of connotations (Feng et al., 2011), i.e., words associated with polarities out of context (e.g., war has negative connotation and sunshine has positive connotation (Feng et al., 2013)). However, words may be used in context with polarities opposite to their connotations, as in Ghenghis Kan likes war.

Inferred Private States and Opinion Implicatures. We address private states inferred from other private states, where the attitude type of both is sentiment. Inference is initiated by explicit sentiment subjectivity. We borrow the term implicature from linguistics, specifically generalized conversational implicature. Grice (1967; 1989) introduced the notion to account for how more can be pragmatically communicated than what is strictly said - what is implicated vs. what is said (Doran et al., 2012). Generalized conversational implicatures are cancellable, or defeasible.

Analogously, we can treat subjectivity as part of what is said, and the private-state inferences we address to be part of what is implicated. Opinion implicatures are default inferences that may not go through in context.

Benefactive/Malefactive Events and States. This work addresses sentiments toward, in general, states and events which positively or negatively affect entities. Various lexical items and semantic roles evoke such situations. We adopt one clear case in this work (Deng et al., 2013): ⟨agent, event, object⟩ triples, where event negatively (badFor) or positively (goodFor) affects the object. An event that is goodFor or badFor is a gfbf event. Note that we have annotated a corpus with gfbf information and the speaker’s sentiment toward the agents and objects of gfbf events (Deng et al., 2013).

3 Overview

In this section, we give an overview of the rule-based system to provide an intuitive big picture of what it can infer, instead of elaborating specific rules, which will be introduced in Section 4.

The system includes default inference rules which apply if there is no evidence to the contrary. It requires as input explicit sentiment and gfbf information (plus any evidence that is contrary to the inferences). The data structure of the input and the output are described in Section 3.1. The rules are applied repeatedly until no new conclusions can be drawn. If a rule matches a sentiment or event that is the target of a private state, the nesting structure is preserved when generating the conclusions. We say that inference is carried out in private state spaces, introduced in Section 3.2. Finally in Section 3.3, an example is provided to illustrate what the system is able to infer.

3.1 Data Structure

The system builds a graphical representation of what it knows and infers about the meaning of a sentence. A detailed knowledge representation scheme is presented in (Wiebe and Deng, 2014). Below is an example from the MPQA corpus.

Ex(1) [He] is therefore planning to trigger [wars] here and there to revive [the flagging arms industry].

There are two gfbf events in this sentence: ⟨He, trigger, wars⟩ and ⟨He, revive, arms industry⟩. The system builds these nodes as input (as printed by the system):

8 writer positive believesTrue
4 He revive flagging arms industry
6 writer positive believesTrue
1 He trigger wars

The system’s printout does not show all the structure of a node. Consider node 8. It has a source edge to the node representing the writer, and a target edge to node 4, which in turn has an agent edge to the node representing He and a object edge to the node representing flagging arms industry. The nodes also have attributes which record, e.g., what type of node it is (node 8 is a privateState and node 4 is a gfbf), polarity (if relevant), etc.

The graph is directed. For example, node 4 is a child of 8. A specification for the input is that each root node must be a sentiment or believesTrue
node whose source is the writer. Inference proceeds by matching rules to the graph built so far and, when a rule successfully fires, adding nodes to the graph.

### 3.2 Private State Spaces

The approach adopted here follows work on reasoning in belief spaces and belief ascription in natural language (Martins and Shapiro, 1983; Rapaport, 1986; Slator and Wilks, 1987). Other than private states of the writer, all propositions and events must be the target of some private state. In the simplest case, the writer believes the proposition or event he/she describes in the document, so the proposition or event is nested under a writer positive believesTrue node.

We want to carry out inferences within private state spaces so that, for example, from S positive believesTrue P, & P ⇒ Q, the system may infer S positive believesTrue Q. However, we are working with sentiment, not only belief as in earlier work, and we want to allow, as appropriate, these types of inferences: from S sentiment toward P, & P ⇒ Q, infer S sentiment toward Q. For example, if I’m upset my computer is infected with a virus, then I’m also upset with the consequences (e.g., that my files may be corrupted).

A private state space is defined by a path where the root is a believesTrue or sentiment node whose source is the writer, and each node on the path is a believesTrue or sentiment node. Two paths define the same private state if, at each corresponding position, they have the same attitude type, polarity, and source. P is in a private state space if P is the target of the rightmost node on a path defining that space.

### 3.3 An Example

Now we have introduced the data structure and the private state spaces, let’s see the potential conclusions which the system can infer before we go into the detailed rules in the next section.

Ex(2) However, it appears as if [the international community (IC)] is tolerating [the Israeli] campaign of suppression against [the Palestinians].

The input nodes are the following.

- writer negative sentiment
- IC positive sentiment
- Israel suppression Palestinians

The gfbf event (Israeli, suppression, Palestinians) is a badFor event. According to the writer, the IC is positive toward the event in the sense that they tolerate (i.e., protect) it. However and appears as if are clues that the writer is negative toward IC’s positive sentiment.

Given these input annotations, the following are the sentiments inferred by the system just toward the entities in the sentence; note that many of the sentiments are nested in private state spaces.

- writer positive sentiment
- Palestinians
- writer negative sentiment
- Israel
- writer negative sentiment
- IC
- writer positive believesTrue
- Israel negative sentiment
- Palestinians
- writer positive believesTrue
- IC negative sentiment
- Palestinians
- writer positive believesTrue
- IC positive sentiment
- Israel
- writer positive believesTrue
- IC positive believesTrue
- Israel negative sentiment
- Palestinians

Note that for the sentiments between two entities other than the writer (e.g., Israel negative toward Palestinians), they are nested under a writer positive believesTrue node. This shows why we need private state spaces. The writer expresses his/her opinion that the sentiment from Israel toward Palestinians is negative, which may not be true outside the scope of this single document.

### 4 Rules

Rules include preconditions and conclusions. They may also include assumptions (Hobbs et al., 1993). For example, suppose a rule would successfully fire if an entity S believes P. If the entity S is not the writer but we know that the writer believes P, and there is no evidence to the contrary (i.e. there is no evidence showing that the entity S doesn’t believe P), then we’ll assume that S believes it as well, if a rule “asks us to”.

Thus, our rules are conceptually of the form:

\[ P_1, ..., P_j : A_1, ..., A_k/Q_1, ..., Q_m \]

where the Ps are preconditions, the As are assumptions, and the Qs are conclusions. For the Qs to be concluded, the Ps must already hold; there
must be a basis for assuming each $A$; and there
must be no evidence against any of the $A$s or $Q$s.

Assumptions are indicated using the term “Ass-
sume”, as in rule 10, which infers sentiment from
connotation:

**rule10:**
(Assume Writer positive ... believesTrue) $A$ gfbf $T$ &
$T$’s anchor is in connotation lexicon $\Rightarrow$
Writer sentiment toward $T$

The first line contains an assumption, the sec-
ound line contains a precondition, and the third con-
tains a conclusion.

**rule8:**
$S$ positive believesTrue $A$ gfbf $T$ &
$S$ sentiment toward $T$ $\Rightarrow$
$S$ sentiment toward ideaOf $A$ gfbf $T$

For example, applying rule 8 to “The bill would
curb skyrocketing health care costs;” from the
writer’s ($S$’s) negative sentiment toward the costs
($T$) expressed by skyrocketing, we can infer the
writer is positive toward the event {bill, curb, costs} ($A$ gfbf $T$) because it would decrease the
costs.

Note that, in rule 8, the inference is (sentiment toward object) $\Rightarrow$ (sentiment toward event).
Rules 1 and 2 infer in the opposite direction.

**rule1:**
$S$ sentiment toward $A$ gfbf $T$ $\Rightarrow$
$S$ sentiment toward ideaOf $A$ gfbf $T$

For rule 1, why “ideaOf $A$ gfbf $T$”? Because the
purview of this work is making inferences about
attitudes, not about events themselves. Conceptu-
ally, ideaOf coerces an event into an idea, raising
it into the realm of private-state spaces. Reasoning
about the ideas of events avoids the classification
of whether the events are realis (i.e., whether they
did/will happen).

Rule 9 infers sentiment toward the agent in a
gfbf event.

**rule9:**
$S$ sentiment toward $A$ gfbf $T$ &
$A$ is a thing &
(Assume $S$ positive believesTrue ... substantial) $A$ gfbf $T$ $\Rightarrow$
$S$ sentiment toward $A$

By default, the system infers the event is in-
tentional and that the agent is positive toward the
event; if there is evidence against either, the infer-
ence should be blocked.

**rule6:**
$A$ gfbf $T$, where $A$ is animate $\Rightarrow$
$A$ intended $A$ gfbf $T$

**rule7:**
$S$ intended $S$ gfbf $T$ $\Rightarrow$
$S$ positive sentiment toward
ideaOf $S$ gfbf $T$

So far, the preconditions have included only one
sentiment. Rule 3 applies when there are nested
sentiments, i.e., sentiments toward sentiments.

**rule3.1:**
$S_1$ sentiment toward
$S_2$ sentiment toward $Z$ $\Rightarrow$
$S_1$ agrees/disagrees with $S_2$ that
isGood/isBad $Z$ &
$S_1$ sentiment toward $Z$

**rule3.2:**
$S_1$ sentiment toward
$S_2$ pos/neg believesTrue substantial $Z$ $\Rightarrow$
$S_1$ agrees/disagrees with $S_2$ that
isTrue/isFalse $Z$ &
$S_1$ pos/neg believesTrue substantial $Z$

**rule3.3:**
$S_1$ sentiment toward
$S_2$ pos/neg believesShould $Z$ $\Rightarrow$
$S_1$ agrees/disagrees with $S_2$ that
should/shouldNot $Z$ &
$S_1$ pos/neg believesShould $Z$

Among the subcases of rule 3, one shared con-
clusion is $S_1$ agrees/disagrees with $S_2$ *, which de-
pends on the sentiment from $S_1$ toward $S_2$. The
reason there are subcases is because the attitude
types of $S_2$ are various, which determine the in-
ferred attitude type of $S_1$.

By rule 3, given the sentiment between $S_1$ and
$S_2$, we can infer whether $S_1$ and $S_2$ agree. Simi-
larly, we can infer in the opposite direction, as rule
4 shows.

**rule4:**
$S_1$ agrees/disagrees with $S_2$ that * $\Rightarrow$
$S_1$ sentiment toward $S_2$

Two other rules are given in (Wiebe and Deng,
2014).

5 Inferences for An Example from
MPQA Corpus

This section returns to the example from the
MPQA corpus in Section 3.1, illustrating some in-
teresting inference chains and conclusions.
Recall that the input for Ex(1) in Section 3.1 is:

8 writer positive believesTrue
4 He revive flagging arms industry
6 writer positive believesTrue
1 He trigger wars
The first inference is from connotation to sentiment since the word war is in the connotation lexicon.

**rule10:**
(Imagine Writer positive ... believesTrue) A gfbf T &
T’s anchor is in connotation lexicon ⇒ Writer sentiment toward T

**Assumptions:**
6 writer positive believesTrue
1 He trigger wars

**Infer:**
17 writer negative sentiment
2 wars

From the writer’s negative sentiment toward wars, the system infers a negative sentiment toward trigger wars, since triggering wars is good.

For them:

**rule8:**
S positive believesTrue A gfbf T &
S sentiment toward T ⇒ S sentiment toward A gfbf T

**Preconditions:**
6 writer positive believesTrue
1 He trigger wars
17 writer negative sentiment
2 wars

**Infer:**
28 writer negative sentiment
1 He trigger wars

On the other hand, since the agent, He, is animate and there is no evidence to the contrary, the system infers that the triggering event is intentional, and that He is positive toward the idea of his performing the event:

**rule6:**
38 writer negative sentiment
20 He positive intends
1 He trigger wars

**Infer:**
41 writer negative sentiment
25 He positive sentiment
26 ideaOf
1 He trigger wars

Continuing with inference, since the writer has a negative sentiment toward the agent’s positive sentiment, the system infers that the writer disagrees with him (rule 3) and thus that the writer is negative toward him (rule 4):

**rule3.1:**
S1 sentiment toward
S2 sentiment toward Z ⇒ S1 agrees/disagrees with S2 that
isGood/isBad Z &
S1 sentiment toward Z

**Preconditions:**
41 writer negative sentiment
25 He positive sentiment
26 ideaOf
1 He trigger wars

**Infer:**
50 writer disagrees with He that
49 isGood
26 ideaOf
1 He trigger wars

30 writer negative sentiment
26 ideaOf
1 He trigger wars

Then rule 4 works on node 50 and infers:

**rule4 ⇒ Infer:**
55 writer negative sentiment
3 He

In addition to the sentiment related to the wars, we have also drawn several conclusions of sentiment toward the arms industry. For example, one of the output nodes related to the arms industry is:

32 writer positive believesTrue
31 He positive sentiment
5 flagging arms industry

The MPQA annotators marked the writer’s negative sentiment, choosing the long spans therefore . . . industry and therefore planning . . . here and there as attitude and expressive subjective element spans, respectively. They were not able to pinpoint any clear sentiment phrases. A machine learning system trained on such examples would have difficulty recognizing the sentiments. The system, relying on the negative connotation of war and the gfbf information in the sentence, is ultimately able to infer several sentiments, including the writer’s negative sentiment toward the trigger event.

6 Conclusions

While previous sentiment analysis research has concentrated on the interpretation of explicitly stated opinions and attitudes, this work addresses opinion implicature (i.e., opinion-oriented default inference) in real-world text. This paper described a rule-based framework for representing and analyzing opinion implicatures which we hope will contribute to deeper automatic interpretation of subjective language. In the course of understanding implicatures, the system recognizes implicit sentiments (and beliefs) toward various events and entities in the sentence, often of mixed polarities; thus, it produces a richer interpretation than is typical in opinion analysis.

Acknowledgements. This work was supported in part by DARPA-BAA-12-47 DEFT grant #12475008 and National Science Foundation grant #IIS-0916046. We would like to thank the anonymous reviewers for their feedback.
References

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

G. Bruder and J. Wiebe. 1995. Recognizing subjectivity and identifying subjective characters in third-person fictional narrative. In Judy Duchan, Gail Bruder, and Lynne Hewitt, editors, *Deixis in Narrative: A Cognitive Science Perspective*. Lawrence Erlbaum Associates.

Lingjia Deng and Janyce Wiebe. 2014. Sentiment propagation via implicature constraints. In *Meeting of the European Chapter of the Association for Computational Linguistics* (EACL-2014).

Lingjia Deng, Yoonjung Choi, and Janyce Wiebe. 2013. Benefactive/malefactive event and writer attitude annotation. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers), pages 120–125, Sofia, Bulgaria, August. Association for Computational Linguistics.

Ryan Doran, Gregory Ward, Meredith Larson, Yaron McNabb, and Rachel E. Baker. 2012. A novel experimental paradigm for distinguishing between ‘what is said’ and ‘what is implicated’. *Language*, 88(1):124–154.

Song Feng, Ritwik Bose, and Yejin Choi. 2011. Learning general connotation of words using graph-based algorithms. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 1092–1103, Edinburgh, Scotland, UK., July. Association for Computational Linguistics.

Song Feng, Jun Seok Kang, Polina Kuznetsova, and Yejin Choi. 2013. Connotation lexicon: A dash of sentiment beneath the surface meaning. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers), pages 1774–1784, Sofia, Bulgaria, August. Association for Computational Linguistics.

Herbert Paul Grice. 1967. Logic and conversation. The William James lectures.

H Paul Grice. 1989. *Studies in the Way of Words*. Harvard University Press.

Jerry R. Hobbs, Mark E. Stickel, Douglas E. Appelt, and Paul Martin. 1993. Interpretation as abduction. *Artificial Intelligence*, 63(1-2):69–142, October.

João Martins and Stuart C. Shapiro. 1983. Reasoning in multiple belief spaces. In *IJCAI*.

Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman, New York.

William J. Rapaport. 1986. Logical foundations for belief representation. *Cognitive Science*, 10(4):371–422.