

Sentiment and Belief:

How to Think about, Represent, and Annotate Private States

A Tutorial

Owen Rambow
Columbia University
rambow@ccls.columbia.edu

Janyce Wiebe
University of Pittsburgh
wiebe@cs.pitt.edu

1 Tutotial Description

1.1 Introduction

Over the last ten years, there has been an explosion in interest in sentiment analysis, with many interesting and impressive results. For example, the first twenty publications on Google Scholar returned for the Query “sentiment analysis” all date from 2003 or later, and have a total citation count of 12,140. The total number of publications is in the thousands. Partly, this interest is driven by the immediate commercial applications of sentiment analysis.

Sentiment is a “private state” (Wiebe, 1990). However, it is not the only private state that has received attention in the computational literature; others include belief and intention. In this tutorial, we propose to provide a deeper understanding of what a private state is. We will concentrate on sentiment and belief. We will provide background that will allow the tutorial participants to understand the notion of a private state as a cognitive phenomenon, which can be manifested linguistically in communication in various ways. We will explain the formalization in terms of a triple of state, source, and target. We will discuss how to model the source and the target. We will then explain in some detail the annotations that have been made. The issue of annotation is crucial for private states: while the MPQA corpus (Wiebe et al., 2005; Wilson, 2007) has been around for some time, most research using it does not make use of many of its features. We believe this is because the MPQA annotation is quite complex and requires a deeper understanding of the phenomenon of “private state”, which is what the annotation is

getting at. Furthermore, there are currently several efforts underway of creating new versions of annotations, which we will also present.

The larger goal of this tutorial is to allow the tutorial participants to gain a deeper understanding of the role of private states in human communication, and to encourage them to use this deeper understanding in their computational work. The immediate goal of this tutorial is to allow the participants to make more complete use of available annotated resources. We propose to achieve these goals by concentrating on annotated corpora, since this will allow participants to both understand the underlying content (achieving the larger goal) and the technical details of the annotations (achieving the immediate goal).

1.2 Current Work on Annotating Sentiment

To date, the computational analyses of sentiment are often fairly superficial. Much work in sentiment analysis and opinion mining is at the document level (Pang et al., 2002). There is increasing interest in more fine-grained levels: sentence-level (McDonald et al., 2007), phrase-level (Choi and Cardie, 2008; Agarwal et al., 2009), aspect-level (Hu and Liu, 2004; Titov and McDonald, 2008), etc. Sentiments toward entities and events (“eTargets”) expressed in blogs, newswire, editorials, etc. are particularly important. A system that could recognize sentiments toward entities and events would be valuable in an application such as Automatic Question Answering, to support answering questions such as “Toward whom/what is X negative/positive?” “Who is negative/positive toward X ?” (Stoyanov et al., 2005).

Or, to augment an automatic wikification system (Ratinov et al., 2011), which could include information about whom or what the subject supports or opposes. A recent NIST evaluation – The Knowledge Base Population (KBP) Sentiment track¹ — aims at using corpora to collect information regarding sentiments expressed toward or by named entities. Annotated corpora of reviews (Hu and Liu, 2004; Titov and McDonald, 2008), widely used in NLP, often include annotations of targets that are aspects of products or services. As such, they are somewhat limited, excluding, e.g., events or agents of events.

A widely used corpus is Version 2 of the MPQA opinion annotated corpus (Wiebe et al., 2005; Wilson, 2007). It is entirely span-based, and contains no eTarget annotations. However, it provides an infrastructure for sentiment annotation that is not provided by other sentiment NLP corpora, and is much more varied in topic, genre, and publication source. MPQA 3.0 (Deng and Wiebe, 2015), which was recently created, adds entity- and event-target (*eTarget*) annotations to the MPQA 2.0 annotations (Wilson, 2007).² The MPQA annotations consist of *private states*, states of a *source* holding an *attitude*, optionally toward a *target*. An important property of sources is that they are *nested*, reflecting the fact that private states and speech events are often embedded in one another. There are several types of attitudes included in MPQA 2.0, including sentiment, arguing, and intention. The tutorial will focus on sentiments (while also discussing the others), which are defined in (Wilson, 2007) as positive and negative evaluations, emotions, and judgements. In the future, eTargets may be added to private states with other types of attitudes.

This tutorial will present the original MPQA annotation scheme (V2) and its recent extension to include eTarget annotations (V3), which we believe is a valuable new resource for the community.

1.3 Belief Annotations

Compared to sentiment, belief has received far less attention in the computational community. There have been several efforts at annotating belief recently. The most complete is FactBank (Saurí and

Pustejovsky, 2009), which represents the source of the belief, the target, the strength, and the polarity (using a system of 10 tags which cover strength and polarity). Following (Wiebe et al., 2005), the sources are nested, reflecting the same nesting of private states we also observe for sentiment. FactBank is a rich and complex annotation; the so-called LU corpus of Diab et al. (2009) was created independently, and represents a subset of the annotations of FactBank. The LU corpus annotates only the writer’s belief in the propositions in the text, only distinguishes 3 types of belief, but does clearly represent the target. Unlike FactBank, which is annotated on top of the Penn Treebank, the LU corpus represents a diverse set of texts. The recent annotations at the LDC for the DARPA DEFT project follow the simplicity of the LU corpus annotations, but extend the tagset of the LU corpus to four tags. An annotation effort in the spring of 2015 will include the source of the belief. The LDC effort is important since it covers a new domain – web discussion forums. Its size is an order of magnitude larger than that of FactBank or the LU corpus (about 800,000 words). This tutorial will discuss these resources and compare the annotations.

1.4 Integration Issues

Sentiment and belief are very similar: most importantly, they are both private states. They both involve a holder and a target, and within the broad categories of sentiment and belief there are subdivisions, which can affect the strength of the private state. There is an important difference though: while the target of a sentiment can be an entity or an event (state of affairs), belief can only target a state of affairs. In addition to being similar types of phenomena, the same linguistic means can convey sentiment and belief at the same time: the utterance *I regret that I am leaving tomorrow* reveals both the utterer’s sentiment and belief towards the leaving event. Despite these interactions between sentiment and belief, there has been no attempt to jointly annotate or predict sentiment and belief. The tutorial will use examples to show the interaction between sentiment and belief, and discuss some issues that arise in joint annotation and tagging.

¹<http://www.nist.gov/tac/2014/KBP/Sentiment/index.html>

²Available at <http://mpqa.cs.pitt.edu>

2 Tutorial Contents

1. Introduction: an overview over the issue of private states, and how they relate to other well-known concepts such as the BDI (belief-desire-intention) model (Bratman, 1999 1987), related work in NLP (such as RST (Mann and Thompson, 1987) and dialog act tagging), linguistic semantics (for example, the notion of veridicity (Karttunen, 1971) and modality), and cognitive science. (45 minutes)
2. Representing sentiment: a presentation of early work, of MPQA V2 (with nested sources, and attitude, expressive-subjective element, and target span annotations), and of MPQA Version 3 (extension of MPQA V2 to eTargets). (45 minutes)
3. Break (15 minutes)
4. Representing belief: a presentation of FactBank, the LU corpus, and the ongoing LDC annotation under the DARPA DEFT program. (30 minutes)
5. Integration and looking forward: a discussion of how sentiment and belief interact, and how we can integrate their annotations, including a discussion of a General Modality Annotation Scheme. (45 minutes)

3 Tutorial Instructors

3.1 Owen Rambow

Owen Rambow is a Senior Research Scientist at the Center for Computational Learning Systems at Columbia University. He is also the co-chair of the Center for New Media at the Data Science Institute at Columbia University. He has been interested in modeling cognitive states in relation to language for a long time, initially in the context of natural language generation (Rambow, 1993; Walker and Rambow, 1994). More recently, he has studied belief in the context of recognizing beliefs in language (Diab et al., 2009; Prabhakaran et al., 2010; Danlos and Rambow, 2011; Prabhakaran et al., 2012). He is currently involved in the DARPA DEFT Belief group, working with other researchers and with the LDC to define annotation standards and evaluations. He has

recently led the pilot evaluation for belief recognition (in English) in the DARPA DEFT program.

He has been the PI or co-PI on many other Government grants from the NSF, DARPA, and IARPA. He has been the Chair of the North American Chapter of the Association for Computational Linguistics. He has been on the editorial board of *Computational Linguistics*, and has served as chair or area chair for several major conferences. <http://www.cs.columbia.edu/~rambow>

3.2 Janyce Wiebe

Janyce Wiebe is Professor of Computer Science and Professor and Co-Director of the Intelligent Systems at the University of Pittsburgh. She has worked on issues related to private states for some time, originally in the context of tracking point of view in narrative (Wiebe, 1994), and later in the context of recognizing sentiment in other genres such as news articles (Wilson et al., 2005). She has approached the area from the perspective of corpus annotation (Wiebe et al., 2005; Deng et al., 2013), lexical semantics (Wiebe and Mihalcea, 2006), and discourse (Somasundaran et al., 2009). In addition to continuing these lines of research, she has recently begun investigating implicatures in opinion analysis (Deng and Wiebe, 2014).

She has received funding for her research from NSF, NIH, DARPA, ONR, NSA, ARDA, and Homeland Security. She was Program Chair of NAACL 2000 and Program Co-Chair of ACL-IJCNLP 2009. She has been on the editorial board of *Computational Linguistics* and is currently an action editor for *Transactions of the ACL*. <http://people.cs.pitt.edu/~wiebe/>

References

- Apoorv Agarwal, Fadi Biadisy, and Kathleen Mckeown. 2009. Contextual phrase-level polarity analysis using lexical affect scoring and syntactic N-grams. In *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, pages 24–32, Athens, Greece, March. Association for Computational Linguistics.
- Michael E. Bratman. 1999 [1987]. *Intention, Plans, and Practical Reason*. CSLI Publications.
- Yejin Choi and Claire Cardie. 2008. Learning with compositional semantics as structural inference for sub-sentential sentiment analysis. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 793–801. Association for Computational Linguistics.
- Laurence Danlos and Owen Rambow. 2011. Discourse relations and propositional attitudes. In *Proceedings of CID 2011 - Fourth International Workshop on Constraints in Discourse*.
- Lingjia Deng and Janyce Wiebe. 2014. Sentiment propagation via implicature constraints. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 377–385, Gothenburg, Sweden, April. Association for Computational Linguistics.
- Lingjia Deng and Janyce Wiebe. 2015. Entity/event-level sentiment annotation. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (short paper)*. Association for Computational Linguistics, May.
- Lingjia Deng, Yoonjung Choi, and Janyce Wiebe. 2013. Benefactive/malefactive event and writer attitude annotation. In *ACL 2013 (short paper)*. Association for Computational Linguistics.
- Mona Diab, Lori Levin, Teruko Mitamura, Owen Rambow, Vinodkumar Prabhakaran, and Weiwei Guo. 2009. Committed belief annotation and tagging. In *Proceedings of the Third Linguistic Annotation Workshop*, pages 68–73, Suntec, Singapore, August. Association for Computational Linguistics.
- Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. ACM.
- Lauri Karttunen. 1971. Some observations on factivity. *Research on Language & Social Interaction*, 4(1):55–69.
- William C. Mann and Sandra A. Thompson. 1987. Rhetorical Structure Theory: A theory of text organization. Technical Report ISI/RS-87-190, ISI.
- Ryan McDonald, Kerry Hannan, Tyler Neylon, Mike Wells, and Jeff Reynar. 2007. Structured models for fine-to-coarse sentiment analysis. In *Annual Meeting-Association For Computational Linguistics*, page 432. Citeseer.
- Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pages 79–86. Association for Computational Linguistics.
- Vinodkumar Prabhakaran, Owen Rambow, and Mona Diab. 2010. Automatic committed belief tagging. In *Coling 2010: Posters*, pages 1014–1022, Beijing, China, August. Coling 2010 Organizing Committee.
- Vinodkumar Prabhakaran, Michael Bloodgood, Mona Diab, Bonnie Dorr, Lori Levin, Christine D. Piatko, Owen Rambow, and Benjamin Van Durme. 2012. Statistical modality tagging from rule-based annotations and crowdsourcing. In *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics*, pages 57–64, Jeju, Republic of Korea, July. Association for Computational Linguistics.
- Owen Rambow. 1993. Rhetoric as knowledge. In *Proceedings of the ACL Workshop on Intentionality and Structure in Discourse Relations*, Columbus, OH.
- Lev Ratinov, Dan Roth, Doug Downey, and Mike Anderson. 2011. Local and global algorithms for disambiguation to wikipedia. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 1375–1384. Association for Computational Linguistics.
- Roser Saurí and James Pustejovsky. 2009. Factbank: a corpus annotated with event factuality. *Language Resources and Evaluation*, 43:227–268. 10.1007/s10579-009-9089-9.
- Swapna Somasundaran, Galileo Namata, Janyce Wiebe, and Lise Getoor. 2009. Supervised and unsupervised methods in employing discourse relations for improving opinion polarity classification. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 170–179, Singapore, August. Association for Computational Linguistics.
- Veselin Stoyanov, Claire Cardie, and Janyce Wiebe. 2005. Multi-Perspective Question Answering using the OpQA corpus. In *Proceedings of the Human Language Technologies Conference/Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP-2005)*, pages 923–930, Vancouver, Canada.
- Ivan Titov and Ryan T McDonald. 2008. A joint model

- of text and aspect ratings for sentiment summarization. In *ACL*, volume 8, pages 308–316. Citeseer.
- Marilyn Walker and Owen Rambow. 1994. The role of cognitive modeling in achieving communicative intentions. In *Proceedings of the Seventh International Workshop on Natural Language Generation*, Kennebunkport, ME.
- Janyce Wiebe and Rada Mihalcea. 2006. Word sense and subjectivity. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, pages 1065–1072, Sydney, Australia, July. Association for Computational Linguistics.
- Janyce Wiebe, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language ann. *Language Resources and Evaluation*, 39(2/3):164–210.
- Janyce M. Wiebe. 1990. Identifying subjective characters in narrative. In *Proceedings of the 13th International Conference on Computational Linguistics (COLING'90)*.
- Janyce Wiebe. 1994. Tracking point of view in narrative. *Computational Linguistics*, 20(2):233–287.
- Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the Human Language Technologies Conference/Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP-2005)*, pages 347–354, Vancouver, Canada.
- Theresa Wilson. 2007. *Fine-grained Subjectivity and Sentiment Analysis: Recognizing the Intensity, Polarity, and Attitudes of private states*. Ph.D. thesis, Intelligent Systems Program, University of Pittsburgh.