

Entity/Event-Level Sentiment Detection and Inference

Lingjia Deng

Intelligent Systems Program

University of Pittsburgh

lid29@pitt.edu

Abstract

Most of the work in sentiment analysis and opinion mining focuses on extracting explicit sentiments. Opinions may be expressed implicitly via inference rules over explicit sentiments. In this thesis, we incorporate the inference rules as constraints in joint prediction models, to develop an entity/event-level sentiment analysis system which aims at detecting both explicit and implicit sentiments expressed among entities and events in the text, especially focusing on but not limited to sentiments toward events that positively or negatively affect entities (*+/-effect events*).

1 Introduction

Nowadays there is an increasing number of opinions expressed online in various genres, including reviews, newswire, editorial, blogs, etc. To fully understand and utilize the opinions, much work in sentiment analysis and opinion mining focuses on more-fined grained levels rather than document-level (Pang et al., 2002; Turney, 2002), including sentence-level (Yu and Hatzivassiloglou, 2003; McDonald et al., 2007), phrase-level (Choi and Cardie, 2008), aspect-level (Hu and Liu, 2004; Titov and McDonald, 2008), etc. Different from them, this work contributes to the sentiment analysis at the **entity/event-level**. 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 “Who is negative/positive toward X ?” (Stoyanov et

al., 2005). It could also be used to facilitate the entity and event resolution (e.g. wikification system (Ratinov et al., 2011)). 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. We will compare the entity/event-level sentiment analysis task to other fine-grained level sentiment analysis tasks in Section 2, and propose to annotate a new entity/event-level sentiment corpus in Section 3.

The ultimate goal of this proposal is to develop an entity/event-level sentiment analysis system which aims at detecting both explicit and implicit sentiments expressed among entities and events in the text. Previous work in sentiment analysis mainly focuses on detecting explicit opinions (Wiebe et al., 2005; Johansson and Moschitti, 2013; Yang and Cardie, 2013). But not all the opinions are expressed in a straight forward way (i.e. explicitly). Consider the example below.

EX(1) It is great that the bill was defeated.

There is a positive sentiment, *great*, explicitly expressed. It is toward the clause *the bill was defeated*. In other words, the writer is explicitly positive toward the event *defeating bill*. Previous work may stop here. However, it is indicated in the sentence that the writer is negative toward *the bill* because (s)he is happy to see that the bill was defeated. The negative sentiment is implicit. Compared to detecting the explicit sentiment, it requires inference to recognize the implicit sentiment.

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

Now consider example Ex(2).

Ex(2) It is great that the bill was passed.

In Ex(2), the writer’s sentiment toward *the bill* is positive, because (s)he is happy to see that the bill was passed. The writer is positive toward the events in both Ex(1) and Ex(2). But different events lead to different sentiments toward the bill. The *defeat* event is harmful to the bill, while the *pass* event is beneficial to the bill. We call such events are named **+/-effect events** (Deng et al., 2013)². Many implicit sentiments are expressed via the +/-effect events, as we have seen in Ex(1) and Ex(2). Previously we have developed rules to infer the sentiments toward +/-effect events (Deng and Wiebe, 2014). An introduction of the rules will be given in Section 4.

This proposal aims at embedding the inference rules and incorporating +/-effect event information into a computational framework, in order to detect and infer both explicit and implicit entity/event-level sentiments. An overview of this proposed work will be presented in Section 5. Later, we will discuss the methods we propose to extract explicit entity/event-level sentiment in Section 6, and talk about how to incorporate the rules to jointly infer implicit sentiments and disambiguate the ambiguities in each step in Section 7. The contributions of this thesis proposal are summarized in Section 8.

2 Related Work

Sentiment Corpus. Annotated corpora of reviews (e.g., (Hu and Liu, 2004; Titov and McDonald, 2008)), widely used in NLP, often include target annotations. Such targets are often aspects or features of products or services, which can be seen as entities or events that are related to the product. However, the set of aspect terms is usually a pre-defined and closed set. (As stated in SemEval-2014: “we annotate only aspect terms naming particular aspects”.) For an event in newsire (e.g. a terrorist attack), it is difficult to define a closed set of aspects. Recently, to create the Sentiment Treebank (Socher et al., 2013), researchers crowdsourced annotations of movie review data and then overlaid the annotations

²It was initially named as **goodFor/badFor event** (Deng et al., 2013; Deng and Wiebe, 2014). Later we renamed it as +/-effect event (Deng et al., 2014; Choi and Wiebe, 2014).

onto syntax trees. Thus, the targets are not limited to aspects of products/services. However, turkers were asked to annotate small and then increasingly larger segments of the sentence. Thus, all the information of the sentence is not shown to turkers when they annotate the span. Moreover, in both corpora of reviews and Sentiment Treebank, the sources are limited to the writer.

+/-Effect Event. Some work have mined various syntactic patterns (Choi and Cardie, 2008), proposed linguistic templates (Zhang and Liu, 2011; Anand and Reschke, 2010; Reschke and Anand, 2011) to find events similar to +/-effect events. There has been work generating a lexicon of patient polarity verbs (Goyal et al., 2012). We define that a *+effect* event has positive effect on the theme (e.g. pass, save, help), while a *-effect* event has negative effect on the theme (e.g. defeat, kill, prevent) (Deng et al., 2013). A +/-effect event has four components: the agent, the +/-effect event, the polarity, and the theme. Later, Choi and Wiebe (2014) have developed sense-level +/-effect event lexicons.

Sentiment Analysis. Most work in sentiment analysis focuses on classifying explicit sentiments and extracting explicit opinion expressions, sources and targets (Wiebe et al., 2005; Wiegand and Klakow, 2012; Johansson and Moschitti, 2013; Yang and Cardie, 2013). There is some work investigating features that directly indicate implicit sentiments (Zhang and Liu, 2011; Feng et al., 2013). In contrast, to bridge between explicit and implicit sentiments via inference, we have defined a generalized set of inference rules and proposed a graph-based model to achieve sentiment propagation between the sentiments toward the agents and themes of +/-effect events (Deng and Wiebe, 2014). But it requires each component of an +/-effect event from manual annotations as input. Later we use an Integer Linear Programming framework to reduce the need of manual annotations in the same task (Deng et al., 2014).

3 Corpus of Entity/Event-Level Sentiment: MPQA 3.0

The MPQA 2.0 (Wiebe et al., 2005; Wilson, 2007) is a widely-used, rich opinion resource. It includes editorials, reviews, news reports, and scripts of interviews from different news agencies, and covers

a wide range of topics³. The MPQA annotations consist of *private states*, states of a *source* holding an *attitude*, optionally toward a *target*. Since we focus on sentiments, we only consider the attitudes which types are sentiments⁴. MPQA 2.0 also contains *expressive subjective element (ESE)* annotations, which pinpoint specific expressions used to express subjectivity (Wiebe et al., 2005). We only consider ESEs whose polarity is positive or negative (excluding those marked neutral).

To create MPQA 3.0, we propose to add entity-target and event-target (*eTarget*) annotations to the MPQA 2.0 annotations. An **eTarget** is an entity or event that is the target of an opinion (identified in MPQA 2.0 by a sentiment attitude or positive/negative ESE span). The eTarget annotation is anchored to the head word of the NP or VP that refers to the entity or event.

Let’s consider some examples. The annotations in MPQA 2.0 are in the brackets, with the subscript indicating the annotation type. The eTargets we add in MPQA 3.0 are boldfaced.

Ex(3) When the Imam [issued the fatwa against]_{sentiment} [Salman **Rushdie** for **insulting** the Prophet]_{target} ...

In Ex(3), *Imam* has a negative sentiment (*issued the fatwa against*) toward the target span, *Salman Rushdie for insulting the Prophet*, as annotated in MPQA 2.0. We find two eTargets in the target span: *Rushdie* himself and his act of *insulting*. Though *the Prophet* is another entity in the target span, we don’t mark it because it is not negative. This shows that within a target span, the sentiments toward different entities may be different. Thus it is necessary to manually annotate the eTargets of a particular sentiment or ESE.

In the following example, the target span is short.

Ex(4) [**He**]_{target} is therefore [**planning to trigger wars**]_{sentiment} ...

He is George W. Bush; this article appeared in the early 2000s. The writer is negative toward Bush because (the writer claims) he is planning to trigger

wars. As shown in the example, the MPQA 2.0 target span is only *He*, for which we do create an eTarget. But there are three additional eTargets, which are not included in the target span. The writer is negative toward Bush planning to trigger wars; we infer that the writer is negative toward the idea of triggering wars and thus toward war itself.

We carried out an agreement study to show the feasibility of this annotation task (Deng and Wiebe, 2015). Two annotators together annotated four documents, including 292 eTargets in total. To evaluate the results, the same agreement measure is used for both attitude and ESE eTargets. Given an attitude or ESE, let set A be the set of eTargets annotated by annotator X , and set B be the set of eTargets annotated by annotator Y . Following (Wilson and Wiebe, 2003; Johansson and Moschitti, 2013), which treat each set A and B in turn as the gold-standard, we calculate the average F-measure $agr(A, B) = (|A \cap B|/|B| + |A \cap B|/|A|)/2$. The $agr(A, B)$ is 0.82 on average over the four documents, showing that this annotation task is feasible. In the future we will continue annotating the MPQA corpus.

We believe that the corpus will be a valuable new resource for developing entity/event-level sentiment analysis systems and facilitating other NLP applications in the future.

4 Inference Rules

Previously we have proposed rules to infer sentiments toward +/-effect events and the components (Deng and Wiebe, 2014). The rule used to infer sentiments in Ex(1) in Section 1 is listed below.

$$\begin{aligned} & \text{writer positive } (E_2 \text{ -effect } E_3) \Rightarrow \\ & \text{writer positive } E_2 \text{ \& writer negative } E_3 \end{aligned}$$

The rule above can be explained as: the writer is positive toward the defeating event (-effect) with the agent (E_2) being implicit and the bill (E_3) being the theme, so that the writer is negative toward the bill. However, these rules are limited to sentiments toward the particular type of event, +/-effect events. Later we develop more rules to infer sentiments toward all types of entities and events (Wiebe and Deng, 2014). One of the rules and an example sentence is:

³Available at <http://mpqa.cs.pitt.edu>

⁴The other types of attitudes include belief, arguing, etc.

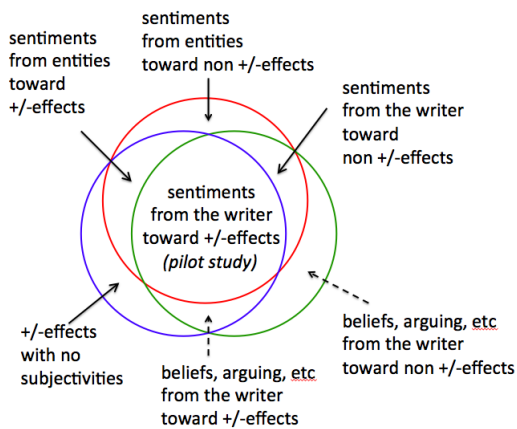


Figure 1: Overview of Subtasks.

$$E_1 \text{ positive } (E_2 \text{ positive } E_3) \Rightarrow \\ E_1 \text{ positive } E_2 \ \& \ E_1 \text{ positive } E_3$$

Ex(5) Great! Mike praised my project!

The rule above can be explained as: if *Mike* (E_2) is positive toward *project* (E_3), and the speaker (E_1) is positive about that positive sentiment, then we could infer: (1) the speaker is positive toward Mike, because the speaker is glad that Mike holds the sentiment, implying that the two entities agree with each other. (2) Because the speaker agrees with Mike, the speaker is positive toward project.

5 Overview

The ultimate goal of this proposed work is to utilize the +/-effect events information and inference rules to improve detecting entity/event-level sentiments in the documents. There are ambiguities in each step of the whole task. We decompose this task into several subtasks, as shown in Figure 1. In this section, we illustrate what are the ambiguities in each subtask.

(1) The region in the blue circle in Figure 1 represents the +/-effect events and the components to be identified. The ambiguities come from: (1.1) Which spans are +/-effect events? (1.2) Which NPs are the agents, which are the themes? (1.3) What is the polarity of the +/-effect event? (1.4) Is the polarity reversed (e.g. negated)?

(2) The region in the red circle represents sentiments we need to extract from the document. The ambiguities are: (2.1) Is there any explicit sentiment? (2.2) What are the sources, targets and polari-

ties of the explicit sentiments? (2.3) Is there any implicit sentiment inferred? (2.4) What are the sources, targets and polarities of the implicit sentiments?

(3) The region in the green circle represents all types of subjectivities of the *writer*, including sentiments, beliefs and arguing. The ambiguities are similar to those in the red circle: (3.1) Is there any subjectivity of the *writer*? (3.2) What are the targets and polarities of the subjectivity?

Though there are many ambiguities, they are interdependent. Inference rules in Section 4 define dependencies among these ambiguities. Our pilot study identifies and infers the writer’s sentiments toward +/-effect events and the components (Deng et al., 2014). We first develop local classifiers using traditional methods to generate the candidates of each ambiguity. Each candidate is defined as a variable in an Integer Linear Programming (ILP) framework and four inference rules are incorporated as constraints in the framework. The pilot study corresponds to the intersection of the three regions in Figure 1. The success of it encourages us to extend from the intersection to all the regions with *solid lines* pointed to: the sources of sentiments are not limited to only the writer but all entities, and the targets of sentiments are not only the +/-effect events and the components, but all the entities and events. The pilot study used a simplified version of the set of rules in (Wiebe and Deng, 2014). In this proposal, we will use the full set.

In summary, this proposal focuses on (a) extracting +/-effect events and the components, and (b) extracting explicit and implicit sentiments. For subtask (a), we propose to utilize the +/-effect event lexicon (Choi and Wiebe, 2014) and semantic role labeling tools to generate candidates of each ambiguity. For subtask (b), we will discuss how to extract explicit sentiments in the next section. Finally, we will discuss how to simultaneously infer implicit sentiments and disambiguate the ambiguities listed above in a joint model in Section 7.

Gold Standard. The MPQA 3.0 proposed in Section 3 and the KBP sentiment dataset will be used as gold standard in this thesis.

Note that, although the two regions with *dashed lines* pointed to are out of scope in this proposal, we can adopt the framework in this proposal to jointly analyze sentiments and beliefs in the future.

6 Explicit Entity/Event-Level Sentiment

To fully utilize the off-the-shelf resources and tools in the span-level and phrase-level sentiment analysis (Wiegand and Klakow, 2012; Johansson and Moschitti, 2013; Yang and Cardie, 2013; Socher et al., 2013; Yang and Cardie, 2014), we will use the opinion spans and source spans extracted by previous work. To extract **eTargets**, which are newly annotated in the MPQA 3.0 corpus, we propose to model this subtask as a classification problem: Given an extracted opinion span returned by the resources, a discriminative classifier judges whether a head of NP/VP in the same sentence is the correct eTarget of the extracted opinion. Two sets of features will be considered.

Opinion Span Features. Several common features used to extract targets will be used, including Part-Of-Speech, path in the dependency parse graph, distance of the constituents on the parse tree, etc (Yang and Cardie, 2013; Yang and Cardie, 2014).

Target Span Features. Among the off-the-shelf systems and resources, some work extracts the target spans in addition to the opinions. We will investigate features depicting the relations between a NP/VP head and the extracted target spans, such as whether the head overlaps with the target span. However, some off-the-shelf systems only extract the opinion spans, but do not extract any target span. For a NP/VP head, if the target span feature is false, there may be two reasons: (1) There is a target span extracted, but the target span feature is false (e.g. the head doesn't overlap with the target span). (2) There is no target span extracted by any tool at all.

Due to this fact, we propose three ways to define target span features. The simplest method (M1) is to assign zero to a false target span feature, regardless of the reason. A similar method (M2) is to assign different values (e.g. 0 or -1) to a false target span feature, according to the reason that causes the feature being false. For the third method (M3), we propose the Max-margin SVM (Chechik et al., 2008). Unlike the case where a feature exists but its value is not observed or false, here this model focus on the case where a feature may not even exist (structurally absent) for some of the samples (Chechik et al., 2008). In other words, the Max-margin SVM deals with features that are known to

be non-existing, rather than have an unknown value. This allows us to fully utilize the different structures of outputs from different state-of-the-art resources.

7 Implicit Entity/Event-Level Sentiment

The explicit sentiments extracted from Section 6 above are treated as input for inferring the implicit sentiment. We are pursuing such a **joint prediction model** that combines the probabilistic calculation of many ambiguities under the constraints of the dependencies of the data, defined by inference rules in the first order logic. Every candidate of every ambiguity is represented as a variable in the joint model. The goal is to find an optimal configuration of all the variables, thus the ambiguities are solved. Models differ in the way constraints are expressed. We plan to mainly investigate undirected lifted graphical models, including Markov Logic Network, and Probabilistic Soft Logics.

Though our pilot study (Deng et al., 2014) and many previous work in various applications of NLP (Roth and Yih, 2004; Punyakanok et al., 2008; Choi et al., 2006; Martins and Smith, 2009; Somasundaran and Wiebe, 2009) have used Integer Linear Programming (ILP) as a joint model, by setting the dependencies as constraints in the ILP framework, there is one limitation of ILP: we have to manually translate the first order logic rules into the linear equations and inequations as constraints. Now we have more complicated rules. In order to choose a framework that computes the first order logic directly, we propose the Markov Logic Network (MLN) (Richardson and Domingos, 2006).

The MLN is a framework for probabilistic logic that employ weighted formulas in first order logic to compactly encode complex undirected probabilistic graphical models (i.e., Markov networks) (Beltagy et al., 2014). It has been applied to various NLP tasks to achieves good results (Poon and Domingos, 2008; Fahrni and Strube, 2012; Dai et al., 2011; Kennington and Schlangen, 2012; Yoshikawa et al., 2009; Song et al., 2012; Meza-Ruiz and Riedel, 2009). It consists of a set of first order logic formula, each associated with a weight. The goal of the MLN is to find an optimal grounding which maximizes the values of all the satisfied first order logic formula in the knowledge base (Richardson and Domingos,

2006). We use the inference rules in Section 4 as the set of first order logic formula in MLN, and define atoms in the logic corresponding to our various kinds of ambiguities. Thus, solving the MLN is to assign true or false value to each atom, that is solving the ambiguities at the same time. For example, $\text{THEME}(x,y)$ represents that the +/-effect event x has a theme y , $\text{TARGET}(x,y)$ represents that the sentiment x has a target y , $\text{POS}(s,x)$ represents that s is positive toward x . The inferences used in Ex(1) and Ex(5) are shown in Table 1.

It is great that the bill was defeated.
$(\text{THEME}(x, y) \wedge \text{POLARITY}(x, -effect)) \Rightarrow$ $(\text{POS}(s, x) \Leftrightarrow \text{NEG}(s, y))$
$(\text{THEME}(\text{defeat}, \text{bill}) \wedge \text{POLARITY}(\text{defeat}, -effect)) \Rightarrow$ $(\text{POS}(\text{writer}, \text{defeat}) \Leftrightarrow \text{NEG}(\text{writer}, \text{bill}))$
Great! Mike praised my project!
$(\text{TARGET}(x, y) \wedge \text{POLARITY}(x, positive)) \Rightarrow$ $(\text{POS}(s, x) \Leftrightarrow \text{POS}(s, y))$
$(\text{TARGET}(\text{praised}, \text{project}) \wedge$ $\text{POLARITY}(\text{praised}, positive)) \Rightarrow$ $(\text{POS}(\text{speaker}, \text{praised}) \Leftrightarrow \text{POS}(\text{speaker}, \text{project}))$

Table 1: Examples and Inference Rules. In each box, line 1: sentence. Line 2: inference rule. Line 3: presenting the sentence in the rule.

Though MLN is a good choice of our task, it has a limitation. Each atom in the first order formula in MLN is boolean value. However, as we stated above, each atom represents an candidate of ambiguity returned by local classifiers, which may be numerical value. We can manually set thresholds for the numerical values to be boolean values, or train a regression over different atoms to select thresholds, but both methods need more parameters and may lead to over-fitting. Therefore, we propose another method, Probabilistic Soft Logic (PSL) (Broecheler et al., 2010). PSL is a new model of statistical relation learning and has been quickly applied to solve many NLP and other machine learning tasks in recent years (Beltagy et al., 2014; London et al., 2013; Pujara et al., 2013; Bach et al., 2013; Huang et al., 2013; Memory et al., 2012; Beltagy et al., 2013). Instead of only being boolean value, the atom in PSL could have numerical values. Given the atoms being numerical, PSL uses the *Lukasiewicz t-norm* and

its corresponding co-norm to quantify the degree to which a grounding of the logic formula is satisfied (Kimmig et al., 2014).

Not limited to the lifted graphical models proposed above, other graphical models are attractive to explore. The Latent Dirichlet Allocation (LDA) (Blei et al., 2003), is widely used in sentiment analysis (Titov and McDonald, 2008; Si et al., 2013; Lin and He, 2009; Li et al., 2010). Li et al. (2010) proposed a LDA model assuming that sentiments depend on each other, which is similar to our assumption that the implicit sentiments depend on explicit sentiment by the inference rules. There is work combining LDA and PSL together (Ramesh et al., 2014), which may be another exploration for us.

8 Contributions

The proposed thesis mainly contributes to sentiment analysis and opinion mining in various genres such as newswire, blogs, editorials, etc.

- Develop MPQA 3.0, an entity/event-level sentiment corpus. It will be a valuable new resource for developing entity/event-level sentiment analysis systems, which are useful for various NLP applications including opinion-oriented Question Answering systems, wikification systems, etc.
- Propose a classification model to extract explicit entity/event-level sentiments. Different from previous classifications in sentiment analysis, we propose to distinguish opinion span features, which are applicable to all the data samples, and target span features, which may be structure absent for some samples (i.e. features do not exist at all).
- Propose a joint prediction framework aims at utilizing the +/-effect events information and inference rules to improve detecting entity/event-level sentiments in the documents and disambiguate the followed ambiguities in each step simultaneously.

Acknowledgement. Thank my advisor Dr. Janyce Wiebe for her very helpful suggestions in this thesis proposal. Thank the anonymous reviewers for their useful comments.

References

- Pranav Anand and Kevin Reschke. 2010. Verb classes as evaluativity functor classes. In *Interdisciplinary Workshop on Verbs. The Identification and Representation of Verb Features*.
- Stephen H Bach, Bert Huang, and Lise Getoor. 2013. Learning latent groups with hinge-loss markov random fields. In *Infering: ICML Workshop on Interactions between Inference and Learning*.
- Islam Beltagy, Cuong Chau, Gemma Boleda, Dan Garrette, Katrin Erk, and Raymond Mooney. 2013. Montague meets markov: Deep semantics with probabilistic logical form. In *2nd Joint Conference on Lexical and Computational Semantics: Proceeding of the Main Conference and the Shared Task, Atlanta*, pages 11–21. Citeseer.
- Islam Beltagy, Katrin Erk, and Raymond Mooney. 2014. Probabilistic soft logic for semantic textual similarity. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1210–1219, Baltimore, Maryland, June. Association for Computational Linguistics.
- David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022.
- Matthias Broecheler, Lilyana Mihalkova, and Lise Getoor. 2010. Probabilistic similarity logic. In *Uncertainty in Artificial Intelligence (UAI)*.
- Gal Chechik, Jeremy Heitz, Gal Elidan, Pieter Abbeel, and Daphne Koller. 2008. Max-margin classification of data with absent features. *The Journal of Machine Learning Research*, 9:1–21.
- Yejin Choi and Claire Cardie. 2008. Learning with compositional semantics as structural inference for subsentential sentiment analysis. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 793–801. Association for Computational Linguistics.
- Yoonjung Choi and Janyce Wiebe. 2014. +/-effectwordnet: Sense-level lexicon acquisition for opinion inference. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1181–1191, Doha, Qatar, October. Association for Computational Linguistics.
- Yejin Choi, Eric Breck, and Claire Cardie. 2006. Joint extraction of entities and relations for opinion recognition. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, EMNLP '06*, pages 431–439, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Hong-Jie Dai, Richard Tzong-Han Tsai, Wen-Lian Hsu, et al. 2011. Entity disambiguation using a markov logic network. In *IJCNLP*, pages 846–855. Citeseer.
- 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. Mpqa 3.0: An entity/event-level sentiment corpus. In *Conference of the North American Chapter of the Association of Computational Linguistics: Human Language Technologies*.
- Lingjia Deng, Yoonjung Choi, and Janyce Wiebe. 2013. Benefactive/malefactive event and writer attitude annotation. In *ACL 2013 (short paper)*. Association for Computational Linguistics.
- Lingjia Deng, Janyce Wiebe, and Yoonjung Choi. 2014. Joint inference and disambiguation of implicit sentiments via implicature constraints. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 79–88, Dublin, Ireland, August. Dublin City University and Association for Computational Linguistics.
- Angela Fahrni and Michael Strube. 2012. Jointly disambiguating and clustering concepts and entities with Markov logic. In *Proceedings of COLING 2012*, pages 815–832, Mumbai, India, December. The COLING 2012 Organizing Committee.
- Song Feng, Jun Sak Kang, Polina Kuznetsova, and Yejin Choi. 2013. Connotation lexicon: A dash of sentiment beneath the surface meaning. In *Proceedings of the 51th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Sofia, Bulgaria, August. Association for Computational Linguistics.
- Amit Goyal, Ellen Riloff, and Hal Daum III. 2012. A computational model for plot units. *Computational Intelligence*, pages 466–488.
- 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.
- Bert Huang, Angelika Kimmig, Lise Getoor, and Jennifer Golbeck. 2013. A flexible framework for probabilistic models of social trust. In *Social Computing, Behavioral-Cultural Modeling and Prediction*, pages 265–273. Springer.
- Richard Johansson and Alessandro Moschitti. 2013. Relational features in fine-grained opinion analysis. *Computational Linguistics*, 39(3):473–509.
- Casey Kennington and David Schlangen. 2012. Markov logic networks for situated incremental natural language understanding. In *Proceedings of the 13th An-*

- nual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 314–323. Association for Computational Linguistics.
- Angelika Kimmig, Lilyana Mihalkova, and Lise Getoor. 2014. Lifted graphical models: a survey. *Machine Learning*, pages 1–45.
- Fangtao Li, Minlie Huang, and Xiaoyan Zhu. 2010. Sentiment analysis with global topics and local dependency. In *AAAI*.
- Chenghua Lin and Yulan He. 2009. Joint sentiment/topic model for sentiment analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 375–384. ACM.
- Ben London, Sameh Khamis, Stephen H. Bach, Bert Huang, Lise Getoor, and Larry Davis. 2013. Collective activity detection using hinge-loss Markov random fields. In *CVPR Workshop on Structured Prediction: Tractability, Learning and Inference*.
- André F. T. Martins and Noah a. Smith. 2009. Summarization with a joint model for sentence extraction and compression. In *Proceedings of the Workshop on Integer Linear Programming for Natural Language Processing - ILP '09*, pages 1–9, Morristown, NJ, USA. Association for Computational Linguistics.
- 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*, volume 45, page 432. Citeseer.
- Alex Memory, Angelika Kimmig, Stephen Bach, Louiqa Raschid, and Lise Getoor. 2012. Graph summarization in annotated data using probabilistic soft logic. In *Proceedings of the 8th International Workshop on Uncertainty Reasoning for the Semantic Web (URSW 2012)*, volume 900, pages 75–86.
- Ivan Meza-Ruiz and Sebastian Riedel. 2009. Multilingual semantic role labelling with markov logic. In *Proceedings of the Thirteenth Conference on Computational Natural Language Learning: Shared Task*, pages 85–90. Association for Computational Linguistics.
- 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.
- Hoifung Poon and Pedro Domingos. 2008. Joint unsupervised coreference resolution with Markov Logic. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pages 650–659, Honolulu, Hawaii, October. Association for Computational Linguistics.
- Jay Pujara, Hui Miao, Lise Getoor, and William Cohen. 2013. Knowledge graph identification. In *The Semantic Web-ISWC 2013*, pages 542–557. Springer.
- Vasin Punyakanok, Dan Roth, and Wen-tau Yih. 2008. The importance of syntactic parsing and inference in semantic role labeling. *Computational Linguistics*, 34(2):257–287.
- Arti Ramesh, Dan Goldwasser, Bert Huang, Hal Daume III, and Lise Getoor. 2014. Understanding mooc discussion forums using seeded lda. In *9th ACL Workshop on Innovative Use of NLP for Building Educational Applications*. ACL.
- 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.
- Kevin Reschke and Pranav Anand. 2011. Extracting contextual evaluativity. In *Proceedings of the Ninth International Conference on Computational Semantics, IWCS '11*, pages 370–374, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Matthew Richardson and Pedro Domingos. 2006. Markov logic networks. *Machine learning*, 62(1-2):107–136.
- Dan Roth and Wen-tau Yih. 2004. A linear programming formulation for global inference in natural language tasks. In *CONLL*.
- Jianfeng Si, Arjun Mukherjee, Bing Liu, Qing Li, Huayi Li, and Xiaotie Deng. 2013. Exploiting topic based twitter sentiment for stock prediction. In *ACL (2)*, pages 24–29.
- Richard Socher, Alex Perelygin, Jean Y Wu, Jason Chuang, Christopher D Manning, Andrew Y Ng, and Christopher Potts. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1631–1642. Citeseer.
- Swapna Somasundaran and Janyce Wiebe. 2009. Recognizing stances in online debates. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 226–234, Suntec, Singapore, August. Association for Computational Linguistics.
- Yang Song, Jing Jiang, Wayne Xin Zhao, Sujian Li, and Houfeng Wang. 2012. Joint learning for coreference resolution with markov logic. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natu-*

- ral Language Learning*, pages 1245–1254. 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.
- Peter D Turney. 2002. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 417–424. Association for Computational Linguistics.
- Janyce Wiebe and Lingjia Deng. 2014. An account of opinion implicatures. *arXiv*, 1404.6491[cs.CL].
- 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.
- Michael Wiegand and Dietrich Klakow. 2012. Generalization methods for in-domain and cross-domain opinion holder extraction. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 325–335. Association for Computational Linguistics.
- Theresa Wilson and Janyce Wiebe. 2003. Annotating opinions in the world press. In *Proceedings of the 4th ACL SIGdial Workshop on Discourse and Dialogue (SIGdial-03)*, pages 13–22.
- 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.
- Bishan Yang and Claire Cardie. 2013. Joint inference for fine-grained opinion extraction. In *ACL (1)*, pages 1640–1649.
- Bishan Yang and Claire Cardie. 2014. Context-aware learning for sentence-level sentiment analysis with posterior regularization. In *Proceedings of ACL*.
- Katsumasa Yoshikawa, Sebastian Riedel, Masayuki Asahara, and Yuji Matsumoto. 2009. Jointly identifying temporal relations with markov logic. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-Volume 1*, pages 405–413. Association for Computational Linguistics.
- Hong Yu and Vasileios Hatzivassiloglou. 2003. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the 2003 conference on Empirical methods in natural language processing*, pages 129–136. Association for Computational Linguistics.
- Lei Zhang and Bing Liu. 2011. Identifying noun product features that imply opinions. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 575–580, Portland, Oregon, USA, June. Association for Computational Linguistics.