Connotation Frames: Typed Relations of Implied Sentiment in Predicate-Argument Structure

Hannah Rashkin    Sameer Singh    Yejin Choi
Computer Science & Engineering
University of Washington
{hrashkin, sameer, yejin}@cs.washington.edu

Abstract

Through a choice of a predicate (e.g., “violate”), a writer can convey subtle sentiments and value judgements toward the arguments of a verb (e.g., projecting the agent as an “antagonist” and the theme as a “victim”). We introduce connotation frames to encode the rich dimensions of implied sentiment, value judgements, and effect evaluation as typed relations that these choices influence, and propose a factor graph formulation that captures the interplay among different types of connotative relations at the lexicon-level. Experimental results confirm that our model is effective in predicting connotative sentiments compared to strong baselines and existing sentiment lexicons.

1 Introduction

People communicate their opinions in many different ways. In domains such as product reviews and social media, it is often through explicit and direct communication; for example, “a thoroughly enjoyable movie with some interesting thoughts on AI.”

The past decade of sentiment analysis research has achieved most success in these domains, where explicit sentiments are considerably prevalent (Pang and Lee, 2008; Liu and Zhang, 2012).

However, in many domains, including journalism and political discourse (Thomas et al., 2006; Somasundaran and Wiebe, 2010), people commonly express their opinions through more subtle and nuanced language, maintaining objectivity on the surface while influencing the readers’ judgments with their ideological standpoints. Anti-climate policy advocates, for instance, might cherry-pick their preferred climate science research and use careful rhetorics to promote their stance through seemingly factual content and objective statements (Boydstun et al., 2013; Greene and Resnik, 2009). Interpreting the true intent and the perspective of the author frequently requires reading between the lines, i.e., inferring the implied messages and subtext. Central to the challenge is beyond what is being directly said, it is often how the writer conveys information that communicates the author’s intention (Greene and Resnik, 2009).

Consider, for example, “The trial court refused defense requests to provide ...”. Although the author does not state it directly, we can infer that they are likely to be negative towards “the trial court” while being more sympathetic towards the “defense requests.” What makes it possible to read between the lines is the author’s choice of predicate “refuse”; from the author’s viewpoint, the trial court “refused” rather than “turned down” or “did not grant” the requests, subtly implying that, in their eyes, the court was being somewhat unreasonable.

In other words, the choice of the predicate frames the likely configuration of the world surrounding it, echoing the very insight of Fillmore’s Frame Semantics (Fillmore, 1982). Unlike most work in frame semantics that defines frames with arguments that appear directly in text (Baker et al., 1998; Palmer et al., 2005), in this work, we also consider the writer and the reader in describing the frames, even if not explicitly mentioned in text. After all, the author is always there, communicating
with implicit readers (Wiebe et al., 2005).

We propose Connotation Frame as a representation framework to describe the implied and predictive sentiment invoked by the choice of the predicate. An example is shown in Figure 1. Here, the predicate violate is a negative sentiment word, as listed in existing sentiment lexicons (Baccianella et al., 2010; Wiebe et al., 2005). The polarity of sentiment from the agent \(a_{agent}\) toward the theme \(a_{theme}\) is negative, i.e., \(P(a_{agent} \rightarrow a_{theme}) = -\). Since the sentiment is generally reciprocal, it is also likely that \(P(a_{theme} \rightarrow a_{agent}) = -\), or at least \(P(a_{theme} \rightarrow a_{agent}) \neq +\).

Interestingly though, not all paired sentiments are negative. The verb “violate” frames its agent as a villain, and its theme as a victim, which means \(P(a_{writer} \rightarrow a_{agent}) = -\), while \(P(a_{writer} \rightarrow a_{theme}) = +\). In other words, the writer is in support of or sympathetic toward \(a_{theme}\), and wants the readers to share similar sentiments, by way of framing the event through the word “violate”. Indeed, anything that can be violated must be something that is intrinsically valuable. Therefore, when placing an entity (e.g., a person, or an issue) as \(a_{theme}\) of “violate”, the implication is that \(a_{theme}\) is something valuable. Under the connotation frame, we can also encode the intrinsic value of each argument implied by the predicate as a unary relation, for example, \(V(a_{theme}) = +\). Additionally, we can also encode the effect of the predicate applied to its argument, as studied in the recent work of Wiebe and Deng (2014) and Choi and Wiebe (2014). For example, being violated means something bad happened to the theme, hence, \(E(a_{theme}) = -\), while \(E(a_{agent}) = 0\), where 0 denotes neutral. Finally, entities involved in an event have internal states, which denote their general outlook. The connotation frame can condense the possible emotional responses of an entity (e.g., happiness, aggression, fear) as another unary relation, \(S(a_i)\). For example, if a patient is saddened by the event, then, under the connotation frame, \(S(a_{patient}) = -\).

As illustrated above, the polarity of different aspects of sentiment, value judgements, and the evaluation of effects can differ from that of the more conventional notion of sentiment, which has been studied extensively (Wilson et al., 2005; Baccianella et al., 2010; Wiebe et al., 2005; Velikovich et al., 2010; Kaji and Kitsuregawa, 2007; Kamps et al., 2004; Takamura et al., 2005; Adreiksvaia and Beggler, 2006; Feng et al., 2013). We propose to organize these various aspects of sentiment, values, effects as connotation frames with typed relations. Of the several different components in connotation frames, our main focus in this paper will be on the directed sentiment relations \(P(a_1 \rightarrow a_2)\). To learn these relations, we construct a factor graph incorporating multi-relational constraints across different types of decisions and find the polarity assignment using loopy belief propagation.

Experiments demonstrate that it is possible to infer these implied sentiments using corpus-driven statistics and global inference, and the proposed approach is able to label inferred sentiments with around 62% accuracy, outperforming two strong baselines by up to 15%. We also show that these results, when used to annotate effect relations, is competitive with the state-of-the-art +/-EffectNet Lexicon (Choi and Wiebe, 2014), outperforming it on some metrics. We include several case studies to show how the resulting connotation frames can help gain insights into biased language in journalism.

### 2 Connotation Frame

Given a predicate \(v\), we define a connotation frame \(F(v)\) as a collection of typed relations and their polarity assignments. Of the different types of re-
values are important elements of frames that help the writer feel positively towards the agent but negatively towards the theme. We also differentiate between interconnected polarities that are neutral and neutral between the agent and the theme. We also differentiate these sentiments as positive, negative, or neutral. However, there are further distinctions that can be made, as well. For example, if someone “scares” another person, the person who has been scared is “fearful.” Unlike becoming “angry” and “aggressive,” which are also negative states, “fearfulness” is sympathetic, and the perspective of the writer towards the scared person is probably positive due to their sympathy for them. On the other hand, the subject has an “aggressive” state, which belongs to the COMPLIANCE frame in FrameNet (Baker et al., 1998), is associated with a frame element NORMS that are generally regarded as something valuable. Similarly, “save,” which belongs to the RESCUING frame, concerns an action that is applied to an ASSET frame element. The values interact with sentiment inference as follows. If an action has a positive effect on a theme \( a_t \) (\( \mathcal{E}(a_t) = + \)), and \( a_t \) is valuable (\( \mathcal{V}(a_t) = + \)), then it is likely that the agent \( a_o \) is positive toward \( a_t \) (\( \mathcal{P}(a_o \rightarrow a_t) = + \)), and so are the writer and the readers. Of course there may be exceptions to these inference rules, as the interpretation of the sentiment is ultimately contextual (Akkaya et al., 2009).

In this paper we focus on learning the perspective elements of the connotation frames, leaving value induction as future work. To enhance the learning however, we explore models that encode values as prior knowledge (approximated using sentiment lexicons) and capture the interplay between values and perspectives through potential functions (§3).

### 2.1 Inference Rules

Given a predicate, the polarity assignments of typed relations are interdependent. Most notably, if the writer feels positively towards the agent but negatively towards the theme then it is likely that the agent and theme do not feel positively towards each other. This insight is related to that of Wiebe and Deng (2014), but differs in that the polarities are predicate-specific and do not rely on knowledge of prior sentiment towards the arguments. These interdependencies are summarized in Table 2. There is a notable pattern: if three non-neutral polarities \( \mathcal{P}(e_1 \rightarrow e_2), \mathcal{P}(e_1 \rightarrow e_3), \mathcal{P}(e_1 \rightarrow e_2) \) exist, and \( \mathcal{P}(e_1 \rightarrow e_2) \oplus \mathcal{P}(e_1 \rightarrow e_3) = \mathcal{P}(e_2 \rightarrow e_3) \), then the polarity values are invalid. We will use these rules in our model (§3).

There are other dependencies that help simplify the tasks. First, the directedsentiments between the agent and the theme are likely to be reciprocal, or at least do not conflict with + and − simultaneously. Therefore, we collapse them into one bidirectional relationship, e.g., \( \mathcal{P}(a_1 \leftrightarrow a_2) \). In addition, we assume the predicted \(^1\) perspective from the reader \( r \) to an argument \( \mathcal{P}(r \rightarrow a) \) are likely to be the same as the implied perspective from the writer \( w \) to the same argument \( \mathcal{P}(w \rightarrow a) \). Therefore, we collapse the readers into the writer. Lifting these assumptions will be future work. For simplicity, our model only explores the polarities involving the agent and the theme. We also differentiate between polarized inferences (when none of the interconnected polarities are neutral) and neutral inferences (when at least one is neutral).

### 2.2 Values and Frames

Values are important elements of frames that help to interpret (implied) sentiment. For example, “viol-
Table 2: Table of inference rules between polarities: Using the known polarities on the left, we can infer the predicted polarities on the right. These inference rules hold in most general cases; however, there are exceptions specific to certain predicates.

3 Modeling Connotation Frames

We present a factor graph representation (Figure 2) to model the connotation frames. There are two types of random variables in the model: nodes \( v \) that represent the polarities associated with verbs, and nodes \( n \) associated with the polarity of valuable nouns. Specifically, for each verb \( w_i \), the factor graph contains three nodes, \( v_{i,w} \), \( v_{i,w,o} \), \( v_{i,s} \), that represent \( \text{writer} \rightarrow \text{subject} \), \( \text{writer} \rightarrow \text{object} \), \( \text{subject} \rightarrow \text{object} \) polarities, respectively. For each noun, there is a vertex \( n_j \) representing the polarity of its value. All these variables, both for verbs and for nouns, take polarity values from the set \( \{-, \text{Neutral}, +\} \).

There are a number of properties and dependencies between the verb and noun polarities that we would like to encode in the graph as factors. Each factor neighbors a subset of the variables and assigns a positive scalar value to its neighbors, the product of which defines the joint distribution over the values of all the variables (Kschischang et al., 2001). Given the set of factors that we describe next (also illustrated in Figure 2), we define probability for any assignment to the verb and noun polarities as follows:

\[
P(v, n) \propto \prod_{i,k} \psi^{\text{seed}}(v_{i,k}) \prod_i \psi^{\text{inf}}(v_{i,w}, v_{i,w,o}, v_{i,s}) \prod_{j,j'} \psi^{\text{sim}}(n_j, n_{j'}) \prod_{i,i',k} \psi^{\text{pref}}(v_{i,k}, v_{i',k}) \prod_{i,j,k} \psi^{\text{pref}}(v_{i,k}, n_j)
\]

Seed Factors We include background knowledge on a subset of variables as unary factors. Incorporating this knowledge as factors, as opposed to fixing the variables as observed, affords us the flexibility of representing noise in the labels as soft evidence. Labeled seed facts are weighted more strongly (0.7) while the label of the nearest neighbor in the training data provide lighter factors (0.5).

Inference Factors We include inference factors to enforce the constraints defined by the inference rules (§2.1). For each verb \( w_i \), we include a factor between nodes \( v_{i,w}, v_{i,w,o}, v_{i,s} \), and \( v_{i,s} \), with a potential that encourages the adherence to inference constraints with the differentiation between polarized inferences and neutral inferences, as follows:

\[
\psi^{\text{inf}}(v_{i,w}, v_{i,w,o}, v_{i,s}) = \begin{cases} 
.9 & \text{polarized inferences} \\
.5 & \text{neutral inferences} \\
.1 & \text{otherwise}
\end{cases}
\]

Similarity Factors To propagate connotation frames across the corpus, we also include factors that connect pairs of similar verbs and pairs of similar nouns. Here, we estimate the similarity between two words, \( \text{sim}(w_i, w_j) \) as the cosine similarity of word2vec pre-trained word vectors (Mikolov et al., 2013). To reduce loops and promote faster convergence, the factors are only introduced when the similarity is greater than a threshold (we use 0.35 and 0.5 respectively), defined as:

\[
\psi^{\text{sim}}(v_{i,k}, v_{j,k}) \propto \begin{cases} 
\text{sim}(w_i, w_j) & v_{i,k} = v_{j,k} \\
1 - \text{sim}(w_i, w_j) & \text{otherwise}
\end{cases}
\]

\[
\psi^{\text{sim}}(n_i, n_j) \propto \begin{cases} 
\text{sim}(w_i, w_j) & n_i = n_j \\
1 - \text{sim}(w_i, w_j) & \text{otherwise}
\end{cases}
\]
Selectional Preference Factors These factors encourage the nouns that tend to fill the argument slots of a verb frame to have the correct corresponding frame. Therefore, we include pairwise factors between verb and noun vertices that represent such selectional preferences. If a noun has a high co-occurrence of being used as the subject or object of a given verb, then there is a factor connecting the \( w \rightarrow s \) or \( w \rightarrow o \) aspect of the verb to that noun. These factors are only drawn for noun-verb pairs with a co-occurrence above a chosen threshold (0.025). The potential, \( \psi_{\text{sel}} \), is determined using the selectional preference, \( p(n|v) \) based on Google syntactic \( n \)-gram occurrences of the verb (Goldberg and Orwant, 2013), that have been re-scaled, defined as:

\[
\psi_{\text{sel}}(v_i^k, n_j) \propto \begin{cases} 
  p(n_j|w_i) & v_i^k = n_j \\
  1 - p(n_j|w_i) & \text{otherwise}
\end{cases}
\]

Loopy Belief Propagation We use loopy belief propagation to annotate the polarity values for the unknown verbs. In the belief propagation algorithm, messages are iteratively passed between the nodes to their neighboring factors and vice versa. Each message \( \mu \), containing a scalar for each value \( x \in \{-, 0, +\} \), is defined from each node \( v \) to a neighboring factor \( a \) as follows:

\[
\mu_{a \rightarrow v}(x) \propto \prod_{a^* \in N(v)} \mu_{a^* \rightarrow v}(x)
\]

and from each factor \( a \) to a neighboring node \( v \) as:

\[
\mu_{a \rightarrow v}(x) \propto \sum_{x', x'_{v^a} = x} \psi(x') \prod_{v^a \in N(a)} \mu_{v^a \rightarrow v}(x'_{v^a})
\]

At conclusion of message passing, the probability of a specific polarity associated with node \( v \) being equal to \( x \) is proportional to \( \prod_{a \in N(v)} \mu_{a \rightarrow v}(x) \). Although belief propagation is not guaranteed to converge on our loopy model, in practice, it often converges and obtains accurate predictions.

4 Experiments

We first describe crowd-sourced annotations (§4.1), then present the empirical results on perspective (§4.2) and effect prediction (§4.3), and conclude with qualitative analysis of a large corpus (§4.4).

4.1 Data and Crowdsourcing

To seed the factor graph with prior knowledge, and also to evaluate the resulting lexicon, we designed an Amazon Mechanical Turk annotation study. We gathered 1207 transitive verbs from a set of news articles on issues related to abortion clinics. Given that abortion cases tend to be divisive, we hypothesize that there should be more opinionated language found in these articles.\(^3\) For each verb, we ask

\(^3\)The articles come from 30 news sources indicated by others as possibly exhibiting liberal or conservative leanings (Mitchell et al., 2014; Center for Media and Democracy, 2013; Center for Media and Democracy, 2012; HWC Library, 2011). We collected 1024 news articles: 827 from possibly-left-leaning and 797 from possibly-right-leaning sources.
We use part of our data as a development set (237 while the remaining 825 which are mapped to the perspective labels. A few verbs are used for three
lectional preference factors and scaling potentials, our factor graph with the baselines. The initial factor graph contains only the seed set factors and the verb similarity factors to propagate this seed information. To this end, we include the inference factors to have per-verb consistency, and finally, incorporate the noun subgraph (noun similarity and selectional preference factors).

### Baselines
To contrast the difference between the polarities marked by conventional sentiment lexicons and the directed perspectives we aim to learn, we present SENTIMENT LEXICON baseline using the well-established lexicon of Wilson et al. (2005). Additionally, to contrast the contribution of the global inference over a local decision, we include NEAREST NEIGHBOR as the second baseline, which labels each verb in the test set with the polarities of the most similar verb from the training set according to the cosine similarity of the words’ corresponding word2vec embeddings.

### Results
We summarize the results in Table 4. All versions of our factor graph outperform the baselines in terms of strict accuracy, a much harder measure than non-conflicting accuracy. The $F_{1}^{-}$ for Writer→Obj is consistently small or often zero due to the fact that only a small percentage (∼ 2%) of the annotations of this aspect are negative. The most significant increase in accuracy by our model can be attributed to the inference factors. This suggests that the inference rules enforced in a soft manner by these factors correct previous errors.

Table 5 shows additional insights on subsets of data with varying characteristics: top 50% most frequent verbs and most confidently-labeled verbs

---

**Figure 3:** Example of Yes/No survey questions for annotators on Amazon Mechanical Turk

five workers to answer questions that ask about the subject and object’s relationship with each other, which are mapped to the perspective labels. A few example questions are included in Figure 3.

The percent agreement on these annotations, and other quality measures, are included in Table 3. The $w → o$ aspect is quite unbalanced and rarely annotated as negative. This is probably due to its passive position in the sentence; annotators felt that the object could not be considered harmful or negative without more context. Annotators were also asked to provide information about the effect of the verb on its subject and object. Notably, the subject did not have an effect 90% of the time, and when there is an effect on either the subject or object, the good and bad effects are roughly equal.\(^4\)

For the noun subgraph, we focus on words that appear as objects of verbs in the JUDGEMENTS frame, as they are likely to be charged words. Using Google syntactic n-grams (Goldberg and Orwant, 2013), we obtain 1167 such nouns and their selectional preference statistics with respect to the verbs. We will share the annotated data at anonymous.url.

### 4.2 Perspective Prediction

We use part of our data as a development set (237 verbs) for selecting cutoffs for similarity and selectional preference factors and scaling potentials, while the remaining 825 verbs are used for three-fold cross validation. All the associated aspects of each word appear together in the same fold. For evaluation, we measure the average accuracy and F1 scores on the three perspectives: $w → s$, $w → o$, and $s → o$. We compare three different versions of our factor graph with the baselines. The initial factor graph contains only the seed set factors and the verb similarity factors to propagate this seed information. To this end, we include the inference factors to have per-verb consistency, and finally, incorporate the noun subgraph (noun similarity and selectional preference factors).

| Aspect | Human % Agreement Distribution |
|--------|-------------------------------|
|        | Acc. | Strict | NC | % + | % - | % 0 |
| $w → s$ | 79.2 | 65.3 | 100.0 | 40.3 | 30.9 | 28.8 |
| $w → o$ | 72.7 | 55.2 | 100.0 | 57.5 | 1.9 | 40.6 |
| $s → o$ | 78.7 | 64.6 | 100.0 | 40.0 | 30.9 | 29.1 |

\(^4\)Cases without a clear majority vote or where two votes directly conflict (positive vs negative) were discarded, as they may indicate word-sense disambiguation issues or words that can be used in varying context (Akkaya et al., 2009). This leads us to 1062 labeled verbs in total with 874, 744, and 923 labels for $w → s$, $w → o$, and $s → o$ respectively.

\(^5\)This implies that when the writer wants to create a frame projecting negative sentiment toward an entity $a$, they are more likely to place $a$ in a more volitional position (agent) of a predicate, rather than in a more passive role (theme). These annotated labels are in part supported by the large-scale corpus analysis in Table 8, where Obama (Romney) is being criticized more often according to left- (right-) leaning sources.
Table 4: Accuracy of Perspective Prediction on cross-validated test set (over all verbs).

| Aspect   | Algorithm                        | Accuracy | F₁   |
|----------|----------------------------------|----------|------|
|          |                                  | Strict   | Non-conflict | + | - | 0 |
|          |                                  | 52.5     | 99.3  | 35.0 | 64.4 | 55.1 |
| Writer → Subj | Sentiment Lexicon                | 59.4     | 90.0  | 55.6 | 71.0 | 51.1 |
|          | 1-Nearest Neighbor               | 60.6     | 89.3  | 54.9 | 72.6 | 49.5 |
|          | FG (v_sim + seed factors)        | 63.1     | 91.0  | 60.8 | 74.3 | 52.3 |
|          | + inference factors              | 63.2     | 91.5  | 61.2 | 75.0 | 51.2 |
|          | + noun subgraph                  |          |       |      |      |     |
| Writer → Obj | Sentiment Lexicon                | 48.2     | 97.2  | 29.2 | 5.6 | 60.8 |
|          | 1-Nearest Neighbor               | 58.9     | 98.8  | 63.7 | 0.0 | 54.3 |
|          | FG (v_sim + seed factors)        | 61.7     | 98.5  | 71.9 | 0.0 | 41.2 |
|          | + inference factors              | 62.9     | 98.5  | 71.5 | 0.0 | 47.9 |
|          | + noun subgraph                  | 62.3     | 98.6  | 71.5 | 0.0 | 45.8 |
| Subj → Obj | Sentiment Lexicon                | 51.4     | 98.9  | 36.1 | 64.7 | 54.7 |
|          | 1-Nearest Neighbor               | 59.6     | 90.3  | 60.5 | 70.6 | 51.5 |
|          | FG (v_sim + seed factors)        | 60.2     | 88.1  | 57.4 | 72.1 | 51.0 |
|          | + inference factors              | 62.4     | 89.8  | 62.9 | 74.1 | 52.5 |
|          | + noun subgraph                  | 61.9     | 90.3  | 63.3 | 73.5 | 50.5 |

Table 5: Accuracy of the full factor graph model over different subsets of verbs. “Most frequent verbs” are the 50% most used verbs in our dataset. “Most confident verbs” are all verbs that have 2 or more aspects voted unanimously by AMT annotators.

| Data                          | Writer → Subj | Writer → Obj | Subj → Obj | All   | Most Frequent Verbs | Most Confident Verbs |
|-------------------------------|---------------|--------------|------------|-------|--------------------|---------------------|
|                               | Human S NC    | Human S NC   | Human S NC |       |                    |                     |
| All                           | 79.2 63.2 91.5| 72.7 62.3 98.6| 78.7 61.9 90.3|       |                    |                     |
| Most Frequent Verbs           | 79.4 63.6 92.0| 73.1 65.9 98.4| 79.5 64.6 91.2|       |                    |                     |
| Most Confident Verbs          | 89.6 80.9 91.4| 99.8 77.0 97.0| 99.8 80.0 89.9|       |                    |                     |

(verbs where annotators voted unanimously on at least two aspects). The substantial performance boost (> 10%) in the latter set indicates that the task at hand is challenging in part because reading the implied sentiment can be difficult for human judges as well, and that the system shows very high performance (nearly 80% on strict measure) when human annotators have the highest agreements. The list of most confident labeled verbs includes words like “maim,” “damage,” “defend,” and “survive.”

4.3 Effect Prediction

We create a factor graph similar to one in §3 for effect predictions, but without the connections for selectional preference. More concretely, we include verb similarity factors and unary factors, the latter consisting of annotated factors over the seed set of annotated verbs and weaker prior factors defined by the post-processed output from the perspective results. As a comparison point, we also deterministically convert the perspective predictions from §4.2 to the effect predictions based on rules from Table 2 that infer effect from perspective polarities. Finally, we also report state-of-the-art EFFECTWORDNET of Choi and Wiebe (2014).6

Results As shown in Table 6, the deterministic conversion from the perspective predictions achieves competitive results, and the global inference using the factor graph raises the performance further. The performance of the perspective conversion is especially surprising as it does not use any effect annotations, further demonstrating that the notions of implied sentiments in connotation frames are closely related to verb effects.

4.4 Analysis of a Large News Corpus

Journalistic news texts, although written with the intent of objective reporting, often express the biases of the writer in an implicit manner. Since in the last section we show that our proposed approach is effective at identifying the implied sentiment of the verbs, we apply the resulting lexicon to explore occurrences of these implied sentiments.

Data From the Stream Corpus 2014 dataset (KBA, 2014), we select 70 million news articles that originate from a variety of online

6Since this lexicon provides a single polarity for the general effect of each verb, instead of using a different value for both the subject and object, we use this value for both the subject and object.
Table 6: Evaluation of the effect a verb has on its subject and object.

| Algorithm | Subject Accuracy | Object Accuracy |
|-----------|------------------|-----------------|
| +/- Effect Lexicon (Choi and Wiebe, 2014) | $F^+_1$ | $F^-_1$ | $F^+_0$ | $F^-_0$ |
| Conversion from Perspectives | 70.9 | 16.1 | 17.1 | 82.5 | 58.9 | 42.8 | 46.2 | 67.6 |
| Effect Factor Graph | 63.6 | 17.0 | 0.0 | 76.6 | 57.8 | 52.9 | 65.7 | 55.1 |

Table 7: Online news sources selected for analysis.

| Left | Right |
|------|-------|
| huffingtonpost.com | foxnews.com |
| slate.com | nypost.com |
| thinkprogress.org | washingtonpost.com |
| newswEEK.com | theamericanconservative.com |
| politico.com | weeklystandard.com |
| washingtonpost.com | nationalreview.com |
| nytimes.com | townhall.com |
| latimes.com | lifepostnews.com |
| newyorker.com | nationalrightslifenews.org |
| salon.com | breitbart.com |
| theguardian.com | wnd.com |
| thedailybeast.com | city-journal.org |
| opednews.com | thehill.com |
| rhrealitycheck.org | humanevents.com |
| npr.org | theblaze.com |

Table 8: Frequent subjects and objects for a few polarizing verbs in left- and right-biased news sources. We show correct labels in the $P(w \rightarrow \cdot)$ column. Our system made a mistake on suffer.

Table 9: Implied Sentiments via Verbs

sources such as newspapers, TV, and web-only. We extract subject-verb-object relations for this subset using the direct dependencies between nouns and verbs as identified by the BBN Serif system, with an additional heuristic of treating nouns that appear before the verb as subjects, obtaining 280 million unique tuples of the form $(url, subject, verb, object, count)$.

**Implied Sentiments via Verbs** There are a number of verbs that our approach found to be extremely indicative of writer’s sentiment towards the verb arguments. Exploring the usage of such verbs can provide an insight into the biases of different writers, for example between *left*- and *right*-winged news sources. To this end, we aggregate over the subset of 30 URLs for which we know their leaning (as described in the footnote in Section 4.1 and shown in Table 7), compute the most common subjects and objects for such verbs, and list some of the frequent ones in Table 8. There are a number of interesting observations here, for example for verbs like “accuse”, “attack”, and “criticize” that denotes that the writer is negative towards the subject and positive towards the objects, we observe left and the right are against (“Limbaugh”, “McCain”, etc.) and (“Obama”, “Biden”, etc.), respectively. It is also worth noting that for verbs such as “accuse”, a noun that appears as a subject in left-leaning sources often appears in as an object in the right-winged sources, and vice versa.

**Estimating Entity Polarities** A crucial advantage of the connotation frames is that not only do they capture the writer’s sentiment towards the subject or the object, but also define the implicit sentiment between the arguments of the verbs. Over a large corpus, this can enable analysis of sentiments of specific entities towards other nouns, and thus provide insight into their opinions even when they are not being expressed explicitly. Figure 4, for example, presents the polarity of entities “Obama” and “Romney” towards a selected set of nouns, by computing the average estimated polarity (using our lexicon) over triples where one of these entities...
Figure 4: Average sentiment of Obama and Romney (as subjects) to selected nouns (as their objects), aggregated over a large corpus using the learned lexicon (§4.2). The line indicates identical sentiments, i.e. Romney is more positive towards the nouns that are above the line.

appear as a subject. Apart from showing nouns that both entities are positive (“audience”, “fund”) or negative (“heat”, “jab”) towards, we can also see interesting examples in which Obama is more positive (below the line: “administration”, “Kerry”, “Democrats”) and ones where Romney is more positive (“opposition”, “GOP”, “America”).

5 Related Work

Most prior work on sentiment lexicons focused on the overall polarity of words without taking into account their semantic arguments (Wilson et al., 2005; Baccianella et al., 2010; Wiebe et al., 2005; Velikovich et al., 2010; Kaji and Kitsuregawa, 2007; Kamps et al., 2004; Takamura et al., 2005; Adreevskaia and Bergler, 2006). Several recent studies began exploring more specific and nuanced aspects of sentiment such as connotation (Feng et al., 2013), good and bad effects (Choi and Wiebe, 2014), and evoked sentiment (Mohammad and Turney, 2010). Drawing inspirations from them, we present connotation frames as a unifying representation framework to encode the rich dimensions of implied sentiment, value judgements, and effect evaluation, and propose a factor graph formulation that captures the interplay among different types of connotation relations.

Goyal et al. (2010a; 2010b) investigated how characters (protagonists, villains, victims) in children’s stories are affected by certain predicates, which is related to the effect relations studied in this work. While Klenner et al. (2014) similarly investigated the relation between the polarity of the verbs and arguments, our work introduces new perspective types and proposes a unified representation and inference model. There have been growing interests for modeling framing (Greene and Resnik, 2009; Hasan and Ng, 2013), biased language (Recasens et al., 2013) and ideology detection (Yano et al., 2010). All these tasks are relatively less studied, and we hope our connotation frame lexicon will be useful for them.

Sentiment inference rules have been explored by the recent work of Wiebe and Deng (2014) and Deng and Wiebe (2014). In contrast, we make a novel conceptual connection between inferred sentiments and frame semantics, organized as connotation frames, and present a unified model that integrates different aspects of the connotation frames. Finally, in a broader sense, what we study as connotation frames draws a connection to schema and script theory (Schank and Abelson, 1975). Unlike most prior work that focused on directly observable actions (Chambers and Jurafsky, 2009; Frermann et al., 2014; Bethard et al., 2008), we focus on implied sentiments that are framed by predicate verbs.

6 Conclusion

In this paper, we presented a novel system of connotative frames that define a set of sentiment-based relations for a predicate. Our factor graph formulation effectively predicts connotative sentiment between predicate arguments and can also be applied to related problems such as identifying effect. Experimental results — both quantitative and qualitative — demonstrate that statistical patterns exist in the way we imply and project sentiment. Our work suggests new research avenues on learning connotation frames, and their applications to deeper understanding of social and political discourse. All the learned connotation frames and annotations will be shared at anonymous.url.
References

Alina Adreevskaia and Sabine Bergler. 2006. Mining wordnet for fuzzy sentiment: Sentiment tag extraction from wordnet glosses. In 11th Conference of the European Chapter of the Association for Computational Linguistics, pages 209–216.

Cem Akkaya, Janyce Wiebe, and Rada Mihalcea. 2009. Subjectivity word sense disambiguation. In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1 - Volume 1, EMNLP ’09, pages 190–199, Stroudsburg, PA, USA. Association for Computational Linguistics.

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner, and Daniel Tapias, editors, Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC’10), Valletta, Malta, May. European Language Resources Association (ELRA).

Collin F Baker, Charles J Fillmore, and John B Lowe. 1998. The Berkeley frameNet project. In Proceedings of the 17th International Conference on Computational linguistics-Volume 1, pages 86–90. Association for Computational Linguistics.

Steven Bethard, William J Corvey, Sara Klingenstein, and James H Martin. 2008. Building a corpus of temporal-causal structure. In LREC.

Amber E Boydstun, Justin H Gross, Philip Resnik, and Sara Klingenstein. 2009. Subjectivity word sense disambiguation. In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1 - Volume 1, EMNLP ’09, pages 190–199, Stroudsburg, PA, USA. Association for Computational Linguistics.

Stephan Greene and Philip Resnik. 2009. More than 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.

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.

Lingjia Deng and Janyce Wiebe. 2014. Sentiment propagation via implicature constraints. In Proceedings of EACL.

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.

Charles Fillmore. 1982. Frame semantics. Linguistics in the morning calm, pages 111–137.

Lea Frermann, Ivan Titov, and Manfred Pinkal. 2014. A hierarchical bayesian model for unsupervised induction of script knowledge. In Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics, Gothenburg, Sweden.

Yoav Goldberg and Jon Orwant. 2013. A dataset of syntactic-ngrams over time from a very large corpus of English books. In Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 1: Proceedings of the Main Conference and the Shared Task: Semantic Textual Similarity, pages 241–247, Atlanta, Georgia, USA, June. Association for Computational Linguistics.

Amit Goyal, Ellen Riloff, and Hal Daumé III. 2010a. Automatically producing plot unit representations for narrative text. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, EMNLP ’10, pages 77–86, Stroudsburg, PA, USA. Association for Computational Linguistics.

Amit Goyal, Ellen Riloff, and Hal Daumé III, and Nathan Gilbert. 2010b. Toward plot units: Automatic affect state analysis. In Proceedings of HLT/NAACL Workshop on Computational Approaches to Analysis and Generation of Emotion in Text (CAET), Los Angeles, CA.

Stephan Greene and Philip Resnik. 2009. More than words: Syntactic packaging and implicit sentiment. In Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, NAACL ’09, pages 503–511, Stroudsburg, PA, USA. Association for Computational Linguistics.

Kazi Saidul Hasan and Vincent Ng. 2013. Frame semantics for stance classification. CoNLL-2013, 124.

HWC Library. 2011. Consider the Source: A Resource Guide to Liberal, Conservative, and Nonpartisan Periodicals. www.ccc.edu/colleges/washington/departments/Documents/PeriodicalsPov.pdf. Compiled by HWC Librarians in January 2011.
Nobuhiro Kaji and Masaru Kitsuregawa. 2007. Building lexicon for sentiment analysis from massive collection of html documents. In *EMNLP-CoNLL*, pages 1075–1083.

Jaap Kamps, MJ Marx, Robert J Mokken, and Maarten De Rijke. 2004. Using wordnet to measure semantic orientations of adjectives.

TREC KBA. 2014. Knowledge Base Acceleration Stream Corpus. http://trec-kba.org/kba-stream-corpus-2014.shtml.

Manfred Klenner, Michael Amsler, Nora Hollenstein, and Gertrud Faas. 2014. *Verb Polarity Frames: a New Resource and Its Application in Target-Specific Polarity Classification*. Universitätsbibliothek Hildesheim.

F.R. Kschischang, B.J. Frey, and H.-A. Loeliger. 2001. Factor graphs and the sum-product algorithm. *Information Theory, IEEE Transactions on*, 47(2):498–519, Feb.

Bing Liu and Lei Zhang. 2012. A survey of opinion mining and sentiment analysis. In *Mining text data*, pages 415–463. Springer.

Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig. 2013. Linguistic regularities in continuous space word representations. In *HLT-NAACL*, pages 746–751.

Amy Mitchell, Jeffrey Gottfried, Jocelyn Kiley, and Katerina Eva Matsa. 2014. Political Polarization & Media Habits. www.journalism.org/2014/10/21/political-polarization-media-habits/. Produced by Pew Research Center in October, 2014.

Saif M Mohammad and Peter D Turney. 2010. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 26–34. Association for Computational Linguistics.

Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The proposition bank: An annotated corpus of semantic roles. *Computational linguistics*, 31(1):71–106.

Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1-2):1–135.

Marta Recasens, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. 2013. Linguistic models for analyzing and detecting biased language. In *ACL (1)*, pages 1650–1659.

Roger C Schank and Robert P Abelson. 1975. *Scripts, plans, and knowledge*. Yale University.

Swapna Somasundaran and Janyce Wiebe. 2010. Recognizing stances in ideological on-line debates. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 116–124. Association for Computational Linguistics.

Hiroya Takamura, Takashi Inui, and Manabu Okumura. 2005. Extracting semantic orientations of words using spin model. In *Proceedings of ACL-05, 43rd Annual Meeting of the Association for Computational Linguistics*, Ann Arbor, US. Association for Computational Linguistics.

Matt Thomas, Bo Pang, and Lillian Lee. 2006. Get out the vote: Determining support or opposition from Congressional floor-debate transcripts. In *Proceedings of EMNLP*, pages 327–335.

Leonid Velikovich, Sasha Blair-Goldensohn, Kerry Hannan, and Ryan McDonald. 2010. The viability of web-derived polarity lexicons. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, HLT ’10, pages 777–785, Stroudsburg, PA, USA. Association for Computational Linguistics.

Janyce Wiebe and Lingjia Deng. 2014. An account of opinion implicatures. *CoRR*, abs/1404.6491.

Janyce Wiebe, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, 39(2-3):165–210.

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pages 347–354. Association for Computational Linguistics.

Tae Yano, Philip Resnik, and Noah A. Smith. 2010. Shedding (a thousand points of) light on biased language. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*, CSLDAMT ’10, pages 152–158, Stroudsburg, PA, USA. Association for Computational Linguistics.