

Semantic Lexicon Expansion for Concept-Based Aspect-Aware Sentiment Analysis

Anni Coden¹(✉), Dan Gruhl², Neal Lewis²(✉), Pablo N. Mendes²(✉),
Meena Nagarajan², Cartic Ramakrishnan², and Steve Welch²

¹ Thomas J. Watson Research Center, Hawthorne, NY, USA
anni@us.ibm.com

² IBM Research Almaden, 650 Harry Rd, San Jose, CA, USA
{dgruhl,nrlewis,pnmendes,MeenaNagarajan,carticramakrishnan}@us.ibm.com

Abstract. We have developed a prototype for sentiment analysis that is able to identify aspects of an entity being reviewed, along with the sentiment polarity associated to those aspects. Our approach relies on a core ontology of the task, augmented by a workbench for bootstrapping, expanding and maintaining semantic assets that are useful for a number of text analytics tasks. The workbench has the ability to start from classes and instances defined in an ontology and expand their corresponding lexical realizations according to target corpora. In this paper we present results from applying the resulting semantic asset to enhance information extraction techniques for concept-level sentiment analysis. Our prototype (Demo at <http://bit.ly/1svngDi>) is able to perform SemSA's Elementary Task (Polarity Detection), Advanced Task #1 (Aspect-Based Sentiment Analysis), and Advanced Task #3 (Topic Spotting).

1 Introduction

Detecting the sentiment expressed in text is a challenging task riddled by the inherent ambiguity and contextual nature of human languages. Consider, for a moment, what is the sentiment expressed by the sentence “I had a cold beer in a cold dining room.” Based on common knowledge (which can be location specific), beer is best enjoyed cold, which implies a positive sentiment. But is a cold dining room good or bad? This determination depends on the context of the sentence - e.g. on a very hot and humid summer day one may enjoy a cold room, however when coming into the house from shoveling snow, a warm room would be more desirable.

The above example illustrates that background knowledge and contextual information are important pieces in trying to solve the sentiment analysis puzzle. We propose a core ontology enriched by semantic lexicon expansion to tackle the most trivial sentiment analysis tasks, while alleviating more complex problems such as the aforementioned sentence. The domain model allows the association of concepts and *a priori* polarity information - such as ‘beer’ (a food concept) and ‘cold temperature’ (a temperature concept). Is a ‘cold’ glass of white wine good

or should it be served at room temperature? In order to help discover concept mentions in text for extending the ontology, we used a Semantic Asset Management Workbench to create and expand semantic lexicons. The workbench allows users to expand the ontology’s coverage of concept and opinion mentions in text, easing and speeding up the creation of resources to aid in the interpretation of the same text through the eyes of different cultures and contexts.

This paper is organized as follows. Section 2 provides an overview of related work. Section 3 describes the core ontology developed and knowledge bases used. Section 4 describes the semantic lexicon expansion. Section 5 presents the sentiment analysis module. Section 6 presents evaluation results. Finally, Sect. 7 presents conclusions and future work.

2 Related Work

Sentiment Analysis and Aspect Detection have gained much attention over the last several years - see [6] for a survey. With the growth of social media and various review sites, rich data sources are becoming more accessible, and industrial use cases increasingly apparent. We find that most approaches rely on semantic lexicons, stressing the need for methods to create and maintain high quality lexical information per category.

Related work in aspect extraction and sentiment analysis has generally had a narrower focus as compared to ours. Blinov and Kotelnikov [1] perform sentiment analysis on verbs and adjectives only, restricting sentiment to narrow descriptive semantics, excluding contextual queues (e.g. “A burger and fries for \$25”). Furthermore, it relies on linguistic features (e.g. POS tagging) that are known to be harder to accurately extract in informal text (e.g. Twitter). Schouten, Frasinicar and de Jong [7] present a method that relies heavily on training corpus and co-occurrence based algorithms, restricting aspect terms to training corpus exposing a risk of over-fitting. Wagner et al. [8] presents a method that performs well with a combination of rules sets that account for domain specific sentiment terms and multiple distance metrics, combined with machine learning to boost their rule sets. However, the approach does not account for conflicting sentiment cases, as well as non-obvious expressions of negation (e.g. “The management was less than accommodating”). They do note that rule based systems will suffer in accuracy when encountering unforeseen terms.

Our work is related to recent advances in concept-level sentiment analysis [2] and relies on techniques ranging from keyword spotting, through endogenous NLP, to noetic NLP [9]. Our model captures entities and aspects, as well as opinions about these aspects or entities. Our focus is on rapidly expanding the model’s lexical coverage to new domains and languages.

3 Ontology and Knowledge Bases

We designed an ontology to model online reviews – i.e. textual comments provided by a customer with opinions about some entity or aspect of that entity.

Each **Review** contains potentially multiple sentences, and each sentence contains 0 to N item reviews (**ItemReview**) and associated opinions (**Opinion**). For example, one review could state that a customer likes the food but dislikes the service. Another might state that the customer likes one food item and dislikes another item in the same category. It is also possible that reviewers provide item reviews with both positive and negative opinions about the same item, in which case we consider that review item as having a polarity **conflict**. Moreover, when contextual knowledge is needed but not present, the system may classify the sentiment as **vague**.

Each **ReviewItem** refers to a mention of an RDF resource in a sentence – i.e. it represents a surface form or the `rdfs:label` of a resource appearing in a certain position in the textual content of a review. The model is able to include review items that are aspects of other items. Aspects include parts-of, containment, or other characteristics of items. For example, a review may target a shop’s floorplan, and offer opinions about the outside seating space (a part of the shop’s floorplan). An opinion may also be directed at the review target resource itself, in which case the aspect is the resource itself – e.g. ‘the restaurant was great’.

The RDF resources included as instances of our model may come from any number of knowledge bases (KBs). In the current prototype, we have imported instances from DBpedia 3.9 [4], and lexicalizations from the DBpedia Lexicalizations Dataset [5]. We focused on instances relating to Books, DVDs, Electronics, Restaurants, and Kitchen&Housewares. We expanded the lexicalizations through our Semantic Asset Management Workbench (see Sect. 4). Besides identifying new lexicalizations for existing concepts, this expansion enables the system to detect items or aspects that are in a known category, but that do not have a URI in the imported knowledge bases. Consequently, the system may produce blank (skolemized) nodes when it cannot find a suitable URI in the current KB. This allows for an incremental approach to maintaining and evolving the core ontology used by the system, as new terms can be later added to the KB or new lexicalizations can be associated to their corresponding URIs¹ (Fig. 1).

4 Semantic Asset Management Workbench

We have developed a Semantic Asset Management Workbench (SAMW) that allows an analyst to draw on a number of techniques for developing, expanding and refining lexical entries in an ontology. Starting with a seed set of terms (usually anywhere between 3 and 30), the system finds all occurrences of these terms in a corpus and collects a set of patterns composed of the 0–6 tokens on the left and right of each occurrence of terms in the corpus. It then examines the corpus to find other words that match these patterns. The results are scored for confidence, support and prevalence. The users are then prompted to examine the top (up to 100) candidates and select which results to add to the lexicon. The system then iterates, taking these new terms, creating an even larger set of patterns and reprocessing the corpus to find more potential matches. Having

¹ This process is currently handled manually, but algorithmic support is possible.

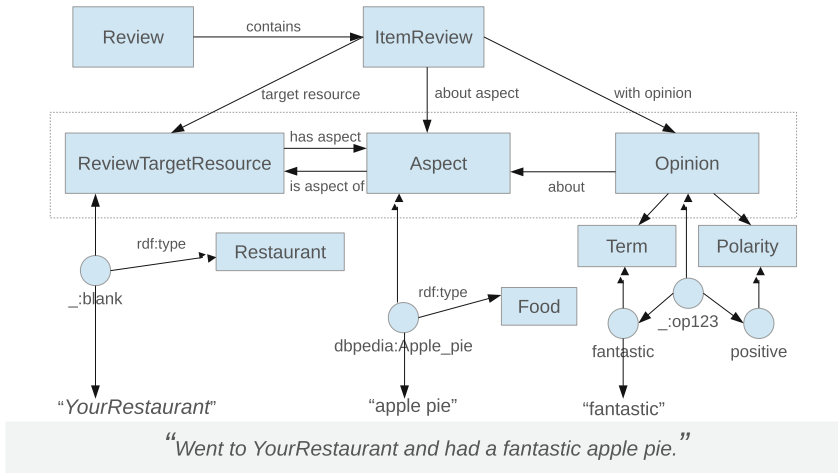


Fig. 1. Concept-based aspect-aware sentiment analysis ontology with examples.

the human in the loop helps to contain conceptual drift (e.g. is water a food?) and focus the lexicon on the concepts as necessary for the task at hand. Thus one key characteristic of SAMW is mutual discovery: it draws from user input to discover more terms, and provides output back to the users that prompts them to make new discoveries.

We started by defining a set of semantic classes of interest and adding them to the core ontology, namely Books, DVDs, Electronics, Restaurants, and Kitchen & Housewares. For the types that existed in DBpedia, we bootstrapped SAMW with entity names from DBpedia. Since the main objective in this particular task is to understand user opinions, we also included classes for positive, negative and neutral valence opinion terms. For those classes, we seeded SAMW with 3–5 manually created examples. For each semantic class, we can also define a set of aspect categories. For example, restaurants have aspects categories in ambience, food, price and service². Additionally, valence lexicons were created, negative, positive and true neutral opinions in a food context (which is somewhat suggestive - e.g. is so-so really neutral?).

We ran 5 to 50 iterations per lexicon on a variety of ‘open’ and ‘closed’ corpora and acquired between 29 and 1126 terms per category. This let us find rarer terms such as ‘sopaipillas’ or ‘mole sauce’ in food, more esoteric opinion terms such as ‘exquisite’ or ‘viable’ for positives in food. We note that SAMW identified opinion terms that have the potential to differ by domain. For example, you wouldn’t say a food is very compact or blazing fast, nor would you say a laptop is ‘flavorful’ or ‘intimate’. Valence varies by domain too, a ‘small’ camera is usually a positive opinion while a small car might not be, and SAMW is able to make such distinctions.

² We used the same categories as the SemEval’14 Task 4.

5 Sentiment Analysis Component

We have developed a sentiment analysis component that extracts sentiment at the item level (e.g. ‘MyRestaurant’), at the aspect level (e.g. ‘MyRestaurant’s rice’), at the item category level (e.g. **Food** and **Restaurant**), or at the review level – i.e. aggregating opinions of multiple items into a final assessment of the overall sentiment in the review. It computes the sentiment of a sentence based on the sentiments of the concepts expressed within a that sentence. Inference across multiple sentences is planned for our future work.

In our prototype, each sentence is processed to produce constituency and dependency parses using OpenNLP³ and ClearNLP⁴ [3]. In addition, we use the aforementioned semantic lexicons from our core ontology, therefore considering concepts under the following categories: **1.** Aspects and ReviewTarget Resources (**AR**) – e.g. beer, wine, dining room; **2.** Positive Opinion Terms (**Pos**) express in general a positive sentiment – e.g. like, good, happy; **3.** Negative Opinion Terms (**Neg**) express in general a negative sentiment – e.g. death, bad, unhappy; **4.** Polarity Inversion Terms (**Inv**) used to invert the polarity of a sentiment – e.g. not, cannot, will not, but, however; **5.** Association concepts **AC**(**concept**, **opinion**, **sentiment**) describing the prior polarity for an opinion term given a concept, where **concept** and **opinion** are instances in one of the above lexicons – e.g. (**beer**, **cold**, **positive**). Clearly “negative concepts” can be used in a positive sense; for instance, the phrase “death by chocolate” is used to refer to very rich chocolate desserts delighting many people. Our model is able to capture these cases through the association concepts.

Our algorithm performs the following steps: **1.** Extract the concepts and opinion terms discussed in each sentence based on our semantic lexicons **AR**, **Pos**, and **Neg**; **2.** Identify the syntactical association between concepts based on the parse of the sentence. **3.** Query our knowledge base for semantic/sentiment (**AC**) associations. **4.** Special processing is done to identify lists, parenthesized expressions and hyphenated expressions. **5.** Polarity inversion: a. Identify the concepts specified in **Inv**; b. Identify the part of the sentence the polarity inversion applies using syntactic parsing constructs and rules;

6 Results

We evaluated our system’s performance on SemSA’s Elementary Task (Polarity Detection), Advanced Task #1 (Aspect-Based Sentiment Analysis), and Advanced Task #3 (Topic Spotting). Precision (P), Recall (R) and F1 results are shown in Tables 1, 2 and 3.

We also performed a preliminary evaluation on the SemEval’14 Task 4 (Restaurants) dataset⁵ and obtained good results for aspect term extraction of non-composed terms of length 1 (72% of the dataset, with F1=0.829) and length 2

³ <http://opennlp.apache.org>

⁴ <http://clearnlp.com>

⁵ <http://alt.qcri.org/semEval2014/task4/>

Table 1. Task 0.

| Team | P | R | F1 |
|-------------|------|------|-------------|
| NCU | 0.78 | 0.57 | 0.66 |
| ours | 0.66 | 0.59 | 0.62 |
| FBK | 0.42 | 0.47 | 0.44 |

Table 2. Task 1.

| Team | P | R | F1 |
|-------------|------|------|-------------|
| FBK | 0.25 | 0.26 | 0.25 |
| ours | 0.24 | 0.14 | 0.18 |
| UNI-NEGEV | 0.12 | 0.05 | 0.07 |

Table 3. Task 3.

| Team | P |
|-------------|-------------|
| ours | 0.33 |
| FBK | 0.17 |

(19% of the data, F1=0.655). In future work we plan to address term compositionality and reevaluate terms that are longer than 3 tokens (9% of the dataset, current F1=0.389).

7 Conclusion

We have presented a prototype for concept-based aspect-aware sentiment analysis. Our system relies on a core ontology of the task that allows us to model reviews based on the resources that they target, aspects of these resources as well as opinion terms related to these aspects or target resources. The ontology allows the definition of *a priori* concept-based opinion polarity to account for differences in expected polarity when one says ‘cold beer’ (positive) versus ‘cold room’ (negative). In order to expand the lexical forms in our ontology, we employed a Semantic Asset Management Workbench that empowers users to discover new terms and learns from the discoveries to improve its discovery process. This workbench allowed us to acquire new terms, name variations, as well as specialized opinion terms to particular categories that may not make sense for other categories (e.g. ‘flavorful’ for food and ‘blazing fast’ for a laptop).

Our system was the best performing system in SemSA’s Task 3, and second best in Tasks 0 and 1, ranking highest of all systems if considering the 3 tasks.

References

1. Blinov, P., Kotelnikov, E.: Distributed representations of words for aspect-based sentiment analysis at SemEval. In: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pp. 140–144 (2014)
2. Cambria, E., Schuller, B., Xia, Y., Havasi, C.: New avenues in opinion mining and sentiment analysis. *IEEE Intell. Syst.* **28**(2), 15–21 (2013)
3. Choi, J.D.: Optimization of Natural Language Processing Components for Robustness and Scalability. Ph.D thesis, Boulder, CO, USA (2012). AAI3549172
4. Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P., Hellmann, S., Morsey, M., van Kleef, P., Auer, S.: DBpedia - a large-scale. multilingual knowledge base extracted from wikipedia. *Seman. Web J.* (2014)
5. Mendes, P., Jakob, M., Bizer, C.: DBpedia for NLP: A Multilingual Cross-domain Knowledge Base. In: Calzolari, N., Choukri, K., Declerck, T., Uan, M., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S. (eds) Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC’12), Istanbul, Turkey, May 2012. European Language Resources Association (ELRA) (2012)

6. Pang, B., Lee, L.: Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.* **2**(1–2), 1–135 (2008)
7. Schouten, K., Frasincar, F., de Jong, F.: Commit-p1wp3: A co-occurrence based approach to aspect-level sentiment analysis. In: *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pp. 203–207 (2014)
8. Wagner, J., Arora, P., Cortes, S., Barman, U., Bogdanova, D., Foster, J., Tounsi, L.: Dcu: Aspect-based polarity classification for semeval task 4. In: *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pp. 223–229 (2014)
9. White, B., Cambria, E.: Jumping NLP curves: a review of natural language processing research. *IEEE Comput. Intell. Mag.* **9**(2), 20–30 (2014)