Machine Reading for Abstractive Summarization of Customer Reviews in the Touristic Domain

Ehab Hassan[1], Davide Buscaldi[1], Aldo Gangemi[1,2]

[1] LIPN CNRS UMR 7030, Laboratoire d'Informatique de Paris Nord, Villetaneuse, France
{ehab.hassan,davide.buscaldi}@lipn.univ-paris13.fr

[2] STLab, ISTC-CNR, Rome, Italy
aldo.gangemi@istc.cnr.it

Abstract : Abstractive summarization is the task of producing a concise representation from a more complex text or a set of texts. This is a useful task especially in the summarization of customer reviews. In this paper we present an abstractive summarization method based on a machine reader and sentiment analysis dictionaries. We carried out a preliminary evaluation of the method on 15 hotel reviews from the Opinosis collection.

Mots-clés : Machine Reading, Abstractive Summarization, Opinion Analysis

1 Introduction

Text summarization is a task consisting in the production of a concise description of a longer, more complex text. Usually, summarization approaches can classified into two types: *extractive* and *abstractive*. In the first case, the original text is reduced to a smaller one, keeping the most important fragments. In the latter, a new text is produced on the basis of the context of the original one. Therefore, abstractive summarization needs a deeper comprehension of the underlying semantics, where extractive summarization can be considered as a shallower task, where the semantics does not play an important role.

One of the most recent applications of the abstractive approaches is the summarization of product reviews and opinions Ganesan *et al.* (2010). This is particularly useful in cases where there are many reviews and most of them are redundant: a user may have to read a great quantity of text before being able to obtain a precise idea of the qualities and the disadvantages of a product.

Machine readers have been introduced by Etzioni *et al.* (2006) as tools for text understanding. They combine different text analysis layers (Part-Of-Speech tagging, syntactic analysis, disambiguation, named entity recognition) to produce a rich semantic representation of the text, which is the reason why we chose to apply them to user reviews in the touristic domain for the abstractive summarization of opinions.

The rest of the paper is structured as follows: in Section 2 we describe the process to extract the opinions and the features from the user reviews. In Section 3 we describe the summarization steps, while in Section 4 we show the experiments carried out and the obtained results. Finally, in Section 5 we draw some conclusions about this preliminary work.

2 Features and Opinion Extraction

The first step consists in the identification of features (or aspects) that are the object of evaluation by users. When users write a review of an hotel, for instance, they usually evaluate not the hotel in its entirety, but specific features of the hotel. Then, we need to find the attributes used to express the opinions. We assumed that such attributes are usually expressed as adjectives. In order to extract the features with their associated attributes from the user reviews, we perform a deep semantic parsing of text, obtaining a RDF Linked-Data-ready graph representation of the text. We employ a large variety of machine reading systems, as implemented in the FRED tool [1] Presutti *et al.* (2012), which extracts knowledge (named entities, senses, taxonomies, relations, events) from text, resolves it onto the Web of Data, adds data from background knowledge, and represents all that in RDF and OWL.

FRED is a tool to automatically transform knowledge extracted from text into RDF and OWL, i.e. it is a *machine reader* for the Semantic Web. It is available as a RESTful API and as a web application. In its current form, it relies upon several NLP components: Boxer[2] for the extraction of the basic logical form of text, BabelNet Navigli & Ponzetto (2010) for word sense disambiguation, and Apache Stanbol[3] for named entity resolution.

Since review features are usually nouns or noun phrases in user reviews, we only interested in features that appear explicitly as nouns or nouns phrases in the reviews. Applying a SPARQL query to the semantic graph produced by FRED, we can extract these features with their opinion words:

```
PREFIX dul: <http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#>
PREFIX vnrole: <http://www.ontologydesignpatterns.org/ont/vn/abox/role/>
PREFIX boxing: <http://www.ontologydesignpatterns.org/ont/boxer/boxing.owl#>
PREFIX boxer: <http://www.ontologydesignpatterns.org/ont/boxer/boxer.owl#>
PREFIX : <http://www.ontologydesignpatterns.org/ont/boxer/test.owl#>
PREFIX d0: <http://www.ontologydesignpatterns.org/ont/d0.owl#>
PREFIX schemaorg: <http://schema.org/>
PREFIX fred: <http://www.ontologydesignpatterns.org/ont/fred/domain.owl#>
PREFIX pos: <http://www.ontologydesignpatterns.org/ont/fred/pos.owl#>
PREFIX BE: <http://www.essepuntato.it/2008/12/earmark#>
SELECT  distinct ?Feature ?neg ?qlt
WHERE {
{
        {?Feature rdf:type ?FeatureType}.
        {{?FeatureType pos:boxerpos pos:n}
          UNION{?FeatureType rdfs:subClassOf* ?FeatureType_1.?FeatureType_1 pos:boxerpos pos:n}}.
          {{?FeatureType dul:hasQuality ?qlt}
            UNION{?Feature dul:hasQuality ?qlt}}
}
OPTIONAL {?sit boxing:involves ?Feature . ?sit boxing:involves ?qlt . ?sit boxing:hasTruthValue ?neg}
}
```

This SPARQL query allows to extract noun features with their modifiers e.g. logical negations, and adverbial qualities (opinion words). Logical negations are very important to determine the polarity of features.

3 Summarization

The system performs the summarization in three main steps: (1) - identify features that have been commented on by customers; (2) - identify opinion words and their polarity, and deciding whether each opinion word is positive, negative, or neutral; and (3) - summarize the results

[1]http://wit.istc.cnr.it/stlab-tools/fred

[2]http://svn.ask.it.usyd.edu.au/trac/candc/wiki/boxer

[3]http://stanbol.apache.org

using the redundant opinions. Given an input composed by a set of user reviews, the system first extracts all the features that appear explicitly as nouns or noun phrases in the reviews and have at least one opinion word associated with them, together with their attributes and eventually the associated logical negation. Then, we used three sentiment lexicons (SentiWordNet (Baccianella *et al.* (2010)), AFINN[4] and Liu (2012)) to detect the polarity of the opinion words. We had three types of polarity which can be assigned to opinion word (e.g. positive, negative, or neutral). Afterward, we kept the features that have positive and negative polarity and deleted the neutral ones. To generate the final summary, we regrouped the remaining features by measuring the similarity between them. We used the WordNet:Similarity package by Pedersen *et al.* (2004) to measure the similarity between the features and considered that two features can be grouped together (i.e. consider as synonyms) if their Lin similarity score is greater than 0.5. We didn't take into account this method to group together the attributes since the WordNet:Similarity package does not offer good semantic similarity measures for adjectives (the only one is the Lesk measure which is not as reliable as the Lin one).

4 Experiments and Results

We started with 15 reviews of a randomly picked hotel from the Opinosis collection by Ganesan *et al.* (2010). We analyzed the reviews with FRED, extracting 140 attributed features. In Figure 1 we show a subset of the features retrieved from the 15 reviews and the associated attribute/opinion word.

Area	Clean	Hotel	Warm	Room	Great
Area	Nice	Location	Not_Better	Room	Huge
Area	Pleasant	Location	Excellent	Room	Large
Bathroom	Nice	Location	Excellent	Room	Spacious
Bathroom	Spacious	Location	Great	Room	Spotless
Bathroom	Spacious	Location	Perfect	Roomy	Clean
Bed	Comfortable	Location	Perfect	Service	Excellent
Bed	Perfect	Location	Right	Service	Great
Bed	Quiet	Person	Happy	Staff	Friendly
Door	Fantastic	Person	Nervous	Staff	Helpful
Hallway	Dark	Person	Picky	Staff	Helpful
Hallway	Loud	Person	Pleased	Stay	Enjoyable
Hotel	Not_Excellent	Rate	Awesome		
Hotel	Big	Rate	Excellent		
Hotel	Friendly	Rate	Great		
Hotel	Great	Rate	Great		
Hotel	Great	Restaurant	Incomplete		
Hotel	Great	Restaurant	Unavailable		
Hotel	Nice	Room	Comfortable		
Hotel	Upscale	Room	Excellent		

Figure 1: An excerpt of the 140 attributed features extracted from the 15 reviews. The highlighted features have been grouped together on the basis of their WordNet::Similarity distance.

The next step was to find the polarity of each attribute. For instance, "comfortable" has a positive polarity in all three dictionaries, and "nervous" has negative polarity in all dictionar-

[4]https://github.com/abromberg/sentiment_analysis/tree/master/AFINN

ies. We reduced the three polarities to a single value and then found the redundant attributed features. Therefore, the 15 reviews were summarized to the attributed features in Table 1.

Feature	Attribute	Freq	Polarity	Feature	Attribute	Freq	Polarity
Staff	Helpful	2	+	Hotel	Great	3	+
Location	Perfect	2	+	Location	Excellent	2	+
Rate	Great	2	+	Rate	Bad	2	-
Room	Spacious	3	+	Room	Nice	2	+
Bed	Comfortable	2	+				

Table 1: The result of the summarization of the 15 reviews.

5 Conclusions

Although this is a very preliminary work, we were able to reduce effectively the complete set of opinion to a synthetic table of features and attributes. Further directions may be to combine the attributes that are very similar (“perfect", “excellent"), using semantic similarity measures developed for Semeval STS[5], and find a way to deal with conflicting ratings. We need also to carry out a more comprehensive evaluation and compare to other summarization methods, such as the one proposed by Popescu & Etzioni (2007).

References

BACCIANELLA S., ESULI A. & SEBASTIANI F. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *LREC*, volume 10, p. 2200–2204.

ETZIONI O., BANKO M. & CAFARELLA M. (2006). Machine reading. In *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI)*.

GANESAN K., ZHAI C. & HAN J. (2010). Opinosis: a graph-based approach to abstractive summarization of highly redundant opinions. In *Proceedings of the 23rd International Conference on Computational Linguistics*, p. 340–348: Association for Computational Linguistics.

LIU B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, **5**(1), 1–167.

NAVIGLI R. & PONZETTO S. P. (2010). Babelnet: Building a very large multilingual semantic network. In *Proceedings of the 48th annual meeting of the association for computational linguistics*, p. 216–225: Association for Computational Linguistics.

PEDERSEN T., PATWARDHAN S. & MICHELIZZI J. (2004). Wordnet:: Similarity: measuring the relatedness of concepts. In *Demonstration papers at HLT-NAACL 2004*, p. 38–41: Association for Computational Linguistics.

POPESCU A.-M. & ETZIONI O. (2007). Extracting product features and opinions from reviews. In A. KAO & S. POTEET, Eds., *Natural Language Processing and Text Mining*, p. 9–28. Springer London.

PRESUTTI V., DRAICCHIO F. & GANGEMI A. (2012). Knowledge extraction based on discourse representation theory and linguistic frames. In *EKAW: Knowledge Engineering and Knowledge Management that matters*.

[5] http://alt.qcri.org/semeval2015/task2/