Aspect Oriented Opinion Summarization

Shima Gerani, Yashar Mehdad, Giuseppe Carenini, Raymond NG
Department of Computer Science
University of British Columbia
gerani, mehdad, carenini, rng@cs.ubc.ca

1 Introduction

The large amount of available user opinions and reviews and the fast rate of their growth makes it very challenging to read through, analyze and understand them in a short amount of time. An effective opinion summarization system is thus needed to help users in grasping the opinion of people about the key aspects of a topic or an entity. In this work, we tackle the problem of aspect-based opinion summarization, which takes a set of user reviews for an entity and generates a summary considering important aspects, their relations, their sentiments and textual evidences, as expressed in the reviews.

Existing opinion summarization systems (e.g., (Hu and Liu, 2004; Carenini et al., 2013a; Di Fabbrizio et al., 2011; Ganesan et al., 2010; Carenini et al., 2013b) often assume that there are no inter- or intra-review relations between the aspects or they use prior knowledge, such as ontologies, to link the aspects. Moreover, the generated summaries are typically extractive. Extractive methods are usually less preferred by users (Carenini et al., 2013b) because they tend to generate summaries that are too verbose and rather incoherent (Ganesan et al., 2010).

In this paper we rely on the rhetorical structure of reviews to infer the importance of aspects as well as the association between them. We argue that the rhetorical role of the text spans in which an aspect occur can help determining the importance of the aspect for opinion summarization. For example, if an aspect occurs more frequently in the nucleus spans of several reviews, its importance should be considered higher compared to other aspects occurring less in the nucleus spans. Also, by looking into the type of rhetorical relation between two spans covering two aspects, we may be able to infer the existence and type of association between the two aspects. For example, two aspects "camera" and "photo" occur in the sentence "I like this camera because it takes great photos". According to the rhetorical structure of this sentence, there exists a *cause* relation between the two segments "I like this camera" and "because it takes great photos". The first segment is the nucleus of this relation and the second segment is the satellite. Therefore, we can say that aspect "camera" is more important than aspect "photo". Also, the aspect "photo" is related to camera if we want to explain the reason behind the satisfaction from camera. In this paper, we use such information to select a subset of important aspects that are highly connected and their coexistence in the subset of selected aspects is necessary to better convey an opinion about a topic. We aim to investigate whether the use of discourse structure in finding important aspects and their relation is effective in content selection for opinion summarization. Our paper highlights the following contributions:

1- Using the rhetorical structure of reviews to estimate the importance of aspects and to find the association types and their weights between aspects.

2- Proposing a novel content selection strategy for opinion summarization based on the importance and associated relations between aspects.

2 Opinion Summarization Framework

2.1 Discourse Parsing

Rhetorical Structure Theory (RST) (Mann and Thompson, 1988) is one of the most popular theories of discourse. RST divides a text into some minimal atomic units, called elementary Discourse Units (EDUs). It then deposits tree representation of a discourse called Discourse Tree (DT). The leaves of a DT are EDUs. The adjacent EDUs are linked by a *rhetorical relation* (e.g., ELABORATION, EXPLANATION). The aggregation of related parts forms the *spans* which represent the internal nodes of DT. Spans can be linked to other units or spans so that the text is connected together into a hierarchical structure. The highest level span in DT represents the whole text. A span linked by a rhetorical relation can be either a NUCLEUS or SATELLITE. The nucleus span is considered to be more central compared to its related span in conveying the message of the author.

In order to parse the reviews to their DT, we use a state-of-the-art discourse parser (Joty et al., 2013). This discourse parser constructs a DT for every sentence using the intra-sentential parser, and then runs a multi-sentential parser on the resulting sentence-level DTs. After parsing the sentences, we keep the aspects in each sentence and remove other words in the sentence.

2.2 Content Selection

Content selection is the process of selecting aspects that are most important to the users. Input to the content selection system is a set of aspects discussed in a set of reviews about an entity, along with the polarity and strength of opinion expressed about each aspect in every sentence in which the aspect is mentioned. The polarity expresses whether the opinion is positive or negative and the strength shows the degree of the sentiment which is expressed as integer from 1 to 3. Our content selection strategy for opinion summarization of reviews of an entity involves the following steps:

1- Building Aspect Rhetorical Relation Graph (ARRG): we generate the ARRG based on the document level rhetorical structure of reviews of the entity. ARRG=(V,E) is a directed graph where we allow multiple edges between two vertices. In ARRG, vertices represent aspects of the entity. Directed edges describe the rhetorical relationship between spans of a document where the aspects occurred, of which the head represents the aspect occurring in the satellite part of the relation. An edge with label r from node u to node v in ARRG indicates that in one or more documents, features u and v have occurred in two different spans S_u and S_v, and the rhetorical relation r holds between the two spans where S_u is the satellite and S_v is the nucleus of the relation. To build ARRG, we first extract all rhetorical dependencies in every review. Knowing which aspect is mentioned in each span participating in a relation, we can extract tuples of the form (u, r, v, w_i) in which u is the aspect occurring in the satellite span, v is the feature occurring in the nucleus span, r is the relation type and w_i is the weight of the relation computed according the following equation:

$$w_i = 1 - 0.5\frac{|edu_u - edu_v|}{\text{number of edus in } D} - 0.5\frac{d_r}{d} \quad (1)$$

where, edu_u and edu_v indicate the EDUs in which aspects u and v, participating in relation r, occur respectively. d indicates the depth of the DT and d_r indicates the level of tree in which the relation occurred, in the review D.

From each review, for every (u, r, v), we extract all (u, r, v, w_i) and we only keep the one with maximum w_i. Finally, we aggregate the tuples extracted from different reviews. To aggregate multiple tuples of the form (u, r, v, w_i), we sum the w_i values and form a single tuple (u, r, v, w) where $w = \sum_i(w_i)$.

2- Selecting the most important aspects for summarization: there are two key factors in evaluating the importance of as aspect: *i)* how frequently it is mentioned in the reviews; and *ii)* its relation between other aspects and their weights based on the discourse structure. In other words, nodes with heavier links and higher frequency, in ARRG, are highly important. In order to extract such subgraph, we run Weighted Page Rank (Xing and Ghorbani, 2004) on ARRG which promotes higher frequency aspect nodes that are either in the nuclei of many relations or in the satellites of relations with highly important aspects. The extracted subgraph is the input for the surface realization step.

2.3 Summary Generation

After content selection, the automatic generation of a natural language summary involves the following additional tasks (Reiter and Dale, 2000): *(i)* structuring the content by ordering and grouping the selected aspects as well as by specifying discourse relations between them; *(ii)* lexical selection; and *(iii)* produce text from the output of the lexical selection. For the first task, we select, from the subgraph extracted in the previous step, a tree rooted at the most frequent aspect using breadth-first search. The relations between the aspects are kept. For the rest, we have adapted the Summarizer of Evaluative Arguments (SEA) system (Carenini et al., 2013b). For lack of space we do not discuss the details here.

References

Giuseppe Carenini, Jackie Chi Kit Cheung, and Adam Pauls. 2013a. Multi-document summarization of evaluative text. *Computational Intelligence*, 29(4):545–576.

Giuseppe Carenini, Jackie Chi Kit Cheung, and Adam Pauls. 2013b. Multi-document summarization of evaluative text. *Computational Intelligence*, 29(4):545–576.

Giuseppe Di Fabbrizio, Ahmet Aker, and Robert Gaizauskas. 2011. Starlet: Multi-document summarization of service and product reviews with balanced rating distributions. In *Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops*, ICDMW '11, pages 67–74, Washington, DC, USA. IEEE Computer Society.

Kavita Ganesan, ChengXiang Zhai, and Jiawei Han. 2010. Opinosis: A graph based approach to abstractive summarization of highly redundant opinions. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING '10).*

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*, KDD '04, pages 168–177, New York, NY, USA. ACM.

Shafiq Joty, Giuseppe Carenini, Raymond Ng, and Yashar Mehdad. 2013. Combining intra- and multi-sentential rhetorical parsing for document-level discourse analysis. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 486–496, Sofia, Bulgaria, August. Association for Computational Linguistics.

William C. Mann and Sandra A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.

Ehud Reiter and Robert Dale. 2000. *Building Natural Language Generation Systems*. Cambridge University Press, New York, NY, USA.

Wenpu Xing and Ali Ghorbani. 2004. Weighted pagerank algorithm. In *Proceedings of the Second Annual Conference on Communication Networks and Services Research*, CNSR '04, pages 305–314. IEEE Computer Society.