Progress Report
Review Summarization

Prepared by Group 23

Group Leader: Ajay Kumar Jaiswal (15CS60R10)
Anoop A (15CS60R21)
Debanjan Paul (15CS60R32)
Dimpi Saikia (15CS60R08)
Gowtham Nayak (15CS60R22)
Kalyani Roy (15CS60R20)
Prasen Kumar Sharma (15IT60R23)
Prasenjit Dey (15CS60R04)
Priya Shree (15IT60R19)

Group Mentor: Abhishek Sikchi

Table of Contents

1. Objective

Summary of product reviews is very useful for both customers and manufacturers. Looking at one review the product cannot be evaluated that whether it's right to go for that particular product and if we go the other way that viewing more than reviews then of course that is a nice way of getting the right concept about the product, but it is time taking work and hence less feasible.

An effective summary of opinions can give users an idea of how good or bad is the the product based on different parameters or aspects. The summary is more abstract in the sense that it captures the parameter, prioritizing them according to number of times they are used. On the other hand manufacturers can use the summarized product reviews to do improvements on the product. So, there is a need of a system built not only to extract sentiments but also justification opinions.

We propose an aspect based abstractive summarization system for product reviews using discourse structure of aspects present in the reviews. Here aspects represents the important parameters or qualities or anything that is evaluated in a the review of the product and the product itself.

The idea is to apply a discourse parser to each customer review and obtain a discourse tree representation for all such reviews. The discourse tree is modified such that all aspects are the leaves of the trees. After creation of aspect trees, we aggregate all the discourse trees to generate a graph, which are called Aspect Rhetorical Relation Graph (ARRG). From ARRG, we select a sub graph representing the most important aspects and the rhetorical relations between them using a PageRank algorithm. Now, this sub graph is transformed to generate an aspect tree. Finally, a natural language summary is generated by applying a template-based NLG framework.

The approach used is an aspect-based abstract from multiple reviews of a product. The proposed system is responsible for content selection and structuring strategy for review summarization, which assumes no prior domain knowledge, by taking advantage of the discourse structure of reviews.

The product-independent template-based NLG framework is used to generate an abstract based on the selected content, without relying on deep syntactic knowledge or sophisticated NLG methods. This framework can effectively convey the distribution of opinions or the reviews.

2. Dataset and Processing details

We conduct our experiment using customer reviews of a digital camera. We use manually annotated aspects and their associated sentiment from the same dataset. We sort the aspects of product based on dir-moi. Then, for each aspect, we generate a sentence based on a simple template "quantifier + polarity-verb".
http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#datasets

3. Proposed Approach

- Summarization Framework:

At a high-level, our summarization framework involves generating a summary from multiple in-put reviews based on an Aspect Hierarchy Tree (AHT) that reflects the importance of aspects as well as the relationships between them. In our framework, an AHT is generated automatically from the set of input reviews, where each sentence of every review is marked by the aspects presented in that sentence and the polarity of opinions over them. polarity/strength (P/S) information given as input to the system. P/S scores are integer values in the range [-3, +3], where +3 is the most positive and -3 is the most negative polarity value. The first component of our system applies a dis-course parser to each review and obtains a dis-course tree representation for every review. The discourse trees are then modified such that every leaf node only contains the aspect words. The output of the first component is an aspect-based discourse tree (ADT) for every review. In the second component, we aggregate the ADTs and generate a graph called Aggregated Rhetorical Relation Graph (ARRG). The third component of our framework is responsible for content selection and structuring. It takes ARRG as input, runs Weighted PageRank, and selects a sub graph representing the most important aspects. Finally it transforms the selected sub graph into a tree and provides an AHT as output. The generated AHT is the input of the last component which generates a natural language summary by applying micro planning and sentence realization. We now describe each component of our framework in more detail.

- Discourse Parsing

Any coherent text is structured so that we can derive and interpret the information. This structure shows how discourse units (text spans such as sentences or clauses) are connected and relate to each other. Discourse analysis aims to reveal this structure. We divide a text into minimal atomic units, called Elementary Discourse Units (EDUs). It then forms a tree representation of a discourse called a Discourse Tree (DT) using rhetorical relations (e.g., Elaboration, Explanation, etc) as edges, and EDUs as leaves. EDUs linked by a rhetorical relation are also distinguished based on their relative importance in conveying the author's message: nucleus is the central part, whereas satellite is the peripheral part. After obtaining the DTs, we remove all words from the text spans of each EDU, except the aspect words. Thus, for each review, we have a DT where a leaf node represents the aspects occurring in the corresponding EDU. Note that there may be EDUs containing no aspects in a review. In such cases, we keep the corresponding node and mark it with no aspect. We call the resulting tree an Aspect-based Discourse Tree (ADT) which will be used in the next components.

- Aspect Rhetorical Relation Graph (ARRG)

In the second component, we aim at generating an ARRG for a product, based on the ADTs which are the output from the previous component. ARRG is a directed graph in which we allow multiple edges between two vertices. In ARRG, vertices represent aspects. We associate to each aspect/node an importance measure that aggregates all the P/S values that the aspect receives in all the reviews.
The direct measure of importance of the aspect is defined as:

$$dir\text{-}moi(a) = \sum_{ps \in PS(a)} ps^2$$

In ARRG, an edge with label r, w from node u to node v, $u \xrightarrow{\text{r, w}} v$ indicates the existence of a relation r with confidence w between two aspects u and v. Also, the direction of the arrow indicates that u and v occurred in the satellite and nucleus spans respectively.

To build ARRG, we use all the ADTs that are output of the previous component (one for each review). From each ADT(j) , we extract all tuples of the form (u, r, v, w) in which u is an aspect occurring in a satellite span, v is an aspect occurring in a nucleus span, r is a relation type and w is the weight of the tuple computed as follows:

$$w = 1 - 0.5\frac{|\text{EDUs between u and v}|}{|\text{total EDUs in } ADT_j|} - 0.5\frac{d_r}{d} \quad (2)$$

Where, |.| indicates cardinality of a set. d indicates the depth of the ADT(j) and dr indicates the depth of the sub-tree of ADT(j) rooted at relation r. Equation 2 weighs a tuple based on two factors:(i) the relative distance of the EDUs in which the two aspects u and v participating in relation r occur. The intuition is that aspects occurring in close proximity to each other are more related; and (ii) the depth of the sub-tree at the point of the relation relative to the depth of the whole ADT(j) . This is because as we move from leaves to the root of a DT, the accuracy of the rhetorical structure has been shown to decrease. Notice that every two aspects u and v may be related by the same relation more than once in an ADT for a review. Thus, we might have i tuples with the same u, r, and v but confidence weights which are not necessarily the same. From every ADT (j) , we extract all (u, r, v, w, i, j) and select the one with maximum confidence. We then aggregate the selected tuples extracted from different reviews. Putting these two steps together, for every two aspects u and v related by relation r, we obtain a single tuple $(u, r, v, \hat{w})$ where

$$\hat{w} = \sum_j \max_i w_{ij}$$

- Content Selection and Structuring

The content of the summary is selected by extracting from ARRG a sub graph containing the most important aspects. Such content is then structured by transforming the sub graph into an aspect hierarchy.

- Sub graph Extraction

In ARRG aspects/nodes are weighted by how frequently and strongly they are evaluated in the re-views and edges are weighted by how frequently and strongly the corresponding aspects are rhetorically related in the discourse trees. In content selection, we want to extract aspects that not only have high weight, but that are also linked with heavy edges to other heavy aspects. This problem can be effectively addressed by Weighted Page Rank (WPR). WPR takes the importance of both the in-links and out-links of the aspects into account and distributes rank scores based on the weights of relations between aspects. In this way, the heavier aspect nodes, which are either in the nuclei of many relations or in the satellites of relations with other heavy aspects, are promoted. We then update the weight of nodes (aspects) with the new score from WPR. Finally, we rank nodes based on their updated score moi and select the top N aspects:

$$moi(a) = \alpha dir\text{-}moi(a) + (1 - \alpha)WPR(a) \quad (4)$$

Here α is a coefficient that can be tuned on a development set or can be set to 0.5 without tuning.

- Aspect Sub graph to Aspects Hierarchy Transformation

In this step, we generate a hierarchical tree structure for aspects. Such a tree structure helps to navigate over aspects and can be easily traversed to find certain aspects and their relation to their parent or children. The hierarchy of aspects also matches the intuition that the root node is the most frequent and general aspect (often the product) and as the depth increases, nodes represent more specific aspects of the product with less frequency and weight. To obtain a hierarchical tree structure from the extracted sub graph, we first build an undirected graph as follows: we merge the edges connecting two nodes and consider the sum of their weights as the weight of the merged graph. We also ignore the relation direction for the purpose of generating the tree. We then find the Maximum Spanning Tree of the undirected sub graph and set the highest weighted aspect as the root of the tree. This process results in a useful knowledge structure of aspects with their associated weight and sentiment polarity connected with the rhetorical relations called Aspect Hierarchical Tree (AHT).

- ## Abstract Generation

The automatic generation of a natural language summary in our system involves the following tasks.

- ### Micro planning

Once the content is selected and structured, it is passed to the micro planning module which per-forms lexical choice. Lexical choice is an important component of micro planning. Lexical choice is formulated in our system based on a “formal” style, language “variability” and “fluent” connectivity among other lexical units.

Quantifiers: for each aspect, a quantifier is selected based on both the absolute and relative number of users whose opinions contributed to the evaluation of the aspect.

Polarity verbs: for each aspect, a polarity verb is selected based on the average sentiment polarity strength for that aspect. Although the average, in most cases, can be a good metric to evaluate the polarity of an aspect, it fails when the distribution of evaluations is centered on zero, for instance, if there are equal numbers of positive and negative evaluations (i.e., controversial). To partially solve this problem, we first check whether the aspect evaluation is controversial by applying the formula proposed by (Carenini and Cheung, 2008). In the case of controversialist, our micro planner selects a lexical item to express the controversialist of the aspect. In other cases, we use the average and select the polarity verb based on that.

Connectives: In order to form more fluent and readable sentences and to increase the language variability, we randomly select our connectives from the list shown in Table 1. Moreover, when a parent aspect (excluding the root in AHT) has two children, they are connected by one of the coordinating conjunction “[and, similarly]” if they agree on polarity, and they will be connected by a choice of “[on the contrary, in contrast]” otherwise. As an alternative we could have selected connectives based on the discourse relations specified in the aspects tree. However, this is left as future work.

Quantifiers:
if (relative-number == 1) : [“All users (x people) who commented about the aspect”, “All costumers (x people) that reviewed the aspect”, ...]
if (relative-number >= 0.8) : [“Almost all users commented about the aspect and they”, “Almost all costumers mentioned the aspect and they”, ...]
if (relative-number >= 0.6) : [“Most users commented about the aspect and they mainly”, “Most shoppers mentioned aspect and they”, ...]
if (relative-number >= 0.45) : [“Almost half of the users commented about the aspect and they”, 'Almost 50% of the shoppers mentioned the aspect and they”, ...]
if (relative-number >= 0.2) : [“About y% of the reviewers commented about the aspect and they”, “Around y% of the shoppers mentioned the aspect and they”, ...]
if (relative-number >= 0.0) : [“z reviewers commented about the aspect and in overall they”, “z shoppers mentioned about the aspect and they”, ...]

Polarity verbs:
if (controversial(aspect)) : [“had controversial opinions about it”, “expressed controversial opinions about this feature”, ...]
else: if (average <= −2) : [“hated it”, “felt that it was very poor', 'thought that it was very poor”, ...]
if (average <= −1) : [“disliked it”, “felt that it was poor”, “thought that it was poor”, ...]
if (average < 0) : [“did not like it”, “felt that it was weak”, “thought that it was weak”, ...]
if (average == 0) : [“did not express any strong positive or negative opinion about it”, ...]
if (average <= +1) : [“liked it”, “felt that it was fine”, “thought that it was satisfactory”, ...]
if (average <= +2) : [“absolutely liked it”, “really liked this feature”, “felt that it was a really good feature”, “thought that it was really good”, ...]
if (average <= +3) : [“loved it”, “felt that it was great”, “thought that it was great”, ...]

Connectives
[“Also, related to the aspect”, “Accordingly, ”, “Moreover, regarding the aspect, ” ,“In relation to the aspect, ”, “Talking about the aspect, ”, ...]

Micro planning strategy for lexical choice. The selected lexical items will fill the template in the realization step

- Sentence Realization

The realization of our abstract generation is per-formed by applying a rather simple and comprehensive template-based strategy. Depending on the specific lexical choice in micro planning step, an appropriate template and corresponding fillers are selected as shown in given template below:

Sentence realization templates:
First sentence templates: *if (polarity-agreement(root,highest-weighted-child) & connecting-relation == {elaboration, explain, cause, summary, same-unit, background, evidence, justify}):* "*quantifier* + *polarity-verb* + 'mainly because of the' + *highest-weighted-child*" *else: "quantifier + polarity-verb"*
First level children (aspects) sentences templates: *"connective + ',' + quantifier + ' ' + polarity-verb"*
Supporting sentences templates: *if (#children(aspect)==1): "connective + quantifier + verb "* *elseif (#children(aspect)>1 & polarity-agreement(children)): "connective + quantifier + verb + {and, similarly, while, ...} + quantifier + verb"* *elseif (#children(aspect)>1 & !polarity-agreement(children)): "connective + quantifier + verb + {but, in contrast, on contrary, ...} + quantifier + verb"*

4. Results Obtained:

1. We have obtained the discourse parsed tree of the reviews and have identified the aspects with their polarity strength. Rhetorical relation among the EDUs are also identified. A small snapshot of the result is provided below:

```
( Nucleus (span 1 3) (rel2par Joint)

        ( Satellite (leaf 1) (rel2par Attribution) (text _!_!I
want to start off!__!) )

        ( Nucleus (span 2 3) (rel2par span)

          ( Satellite (leaf 2) (rel2par Attribution) (text
_!_!saying!__!) )

          ( Nucleus (leaf 3) (rel2par span) (text _!_!that this
camera is small for a reason . <s>!__!) )

  )
```

2. From the Discourse tree, we have generated the Aspect based discourse tree which defined the underlying aspect of EDU and the Rhetorical relation among them.

```
room,Evaluation,small,0.261905
room,Evaluation,camera,0.333333
room,Evaluation,size,0.357143
memory,Elaboration,size,0.547619
size,Manner-Means,camera,0.761905
camera,Evaluation,size,0.166667
camera,Evaluation,memory,0.261905
camera,Contrast,small,0.5
size,Contrast,small,0.47619
memory,Contrast,small,0.380952
```

```
memory,Elaboration,camera,0.52381
```

3. We have generated the ARRG of the product based on the ADTs based on the output of previous component. The output snippet is provided below:

```
camera,Elaboration,auto mode,0.23
photo quality,Background,auto mode,0.75
camera,Elaboration,photo quality,0.5
camera,Elaboration,auto mode,0.33
camera,Elaboration,photo quality,0.2
####,####,####,####
camera,Elaboration,control,0.5
control,Contrast,auto mode,0.66
camera,Elaboration,auto mode,0.075
camera,Elaboration,control,0.2
camera,Elaboration,auto mode,0.375
####,####,####,####
```

4. We have generated the tuples with highest strength from the output of previous state and used it for creating the summary.

```
photo quality,Background,auto mode,0.75
camera,Elaboration,photo quality,0.5
camera,Elaboration,auto mode,0.705
camera,Elaboration,control,0.5
control,Contrast,auto mode,0.66
```

5. We have used the above tuple set of graphs to generate our review summary and we have received the following results :

Many reviewers commented about camera.They were happy with photo quality with auto mode. They found it satisfactory camera because of its photo quality.They found it good camera because of its auto mode.They found it satisfactory camera because of its control.They said it have satisfactory control but still auto mode.

5. Work plan for rest of Semester:

- We will be working on the ranking algorithms to find the tuples with highest underlying strength to become part of the review.
- We will be improving the results of the ARRG graph for the better results and also working on the quantization of reviewers to find the percentage metric of the count who have described about a particular quality of product.
- We will be trying to write a more organized template for sentence realization that provides a more decent review output.
- We will be trying to build an aspect related dataset for a product with it aspect strength understanding its strength in our review dataset.

6. Appendix

Team Member	Workload
Prasenjit Dey (15CS60R04)	Discourse Parse Tree, ADT and ARRG Generation
Ajay Kumar Jaiswal (15CS60R10)	Discourse Parse Tree, Sentence Realization and Aspect Filtering
Debanjan Paul (15CS60R32)	Discourse Parse Tree, ADT and ARRG Generation
Anoop A (15CS60R21)	Discourse Parse Tree, Sentence Realization and Aspect Filtering
Dimpi Saikia (15CS60R08)	Discourse Parse Tree, Sentence Realization and Aspect Filtering
Priya Shree (15IT60R19)	ARRG and ADT Generation and Report Preparation
Kalyani Roy (15CS60R20)	Report Preparation and ADT Generation
Prasen Kumar Sharma (15IT60R23)	Report Preparation and Aspect Identification
Gowtham Nayak (15CS60R22)	Report Preparation and Aspect Identification

7. References:

- Abstractive Summarization of Product Reviews Using Discourse Structure http://www.aclweb.org/anthology/D/D14/D14-1168.pdf
- Review Synthesis for Micro-Review Summarization http://www.cs.uoi.gr/~tsap/publications/wsdm348-nguyen.pdf
- Opinosis: A Graph-Based Approach to Abstractive Summarization of Highly Redundant Opinions http://lexitron.nectec.or.th/public/COLING-2010_Beijing_China/PAPERS/pdf/PAPERS039.pdf
- Micro opinion Generation http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.232.5398&rep=rep1&type=pdf