Capture of Evidence for Summarization: An Application of Enhanced Subjective Logic

Sukanya Manna, B. Sumudu U. Mendis, and Tom Gedeon

School of Computer Science, The Australian National University, Canberra, ACT 0200, Australia
{sukanya.manna,sumudu.mendis,tom.gedeon}@anu.edu.au

Abstract. In this paper, we present a method to generate an extractive summary from a single document using subjective logic. The idea behind our approach is to consider words and their co-occurrences between sentences in a document as evidence of their relatedness to the contextual meaning of the document. Our aim is to formulate a measure to find out 'opinion' about a proposition (which is a sentence in this case) using subjective logic in a closed environment (as in a document). Stronger opinion about a sentence represents its importance and are hence considered to summarize a document. Summaries generated by our method when evaluated with human generated summaries, show that they are more similar than baseline summaries.

Keywords: subjective logic, opinions, evidence, events, summarization, information extraction.

1 Introduction

It is sometime necessary to analyse a single document for intelligent decision making purpose in the absence of prior domain knowledge. In such a scenario, significant sentences from a document or rather gist of that document can only let an user know about what it is all about. Based on this filtered information, the user can decide what kind of measures to be taken to perform the analysis; thus, single-document summarization is one of the best ways to do this.

One way to do text summarization is by text extraction, which means to extract pieces of original text on statistical basis or heuristic methods and put them together to a new shorter text with as much information as possible preserved [9]. The concept of extracting significant sentences from a document for generating extractive summaries has drawn attention in the literature.

In this paper, our approach is not mere assigning scores to a sentence. When a document is looked from the perspective of human, they analyse it by finding what the main idea of the source text is and filtering what is essential in the information conveyed by the text. In [10], the authors have pointed out that a given piece of text is interpreted by different person in a different fashion especially in the way how they understand and interpret the context. Thus we see

M.J. Zaki et al. (Eds.): PAKDD 2010, Part II, LNAI 6119, pp. 291–298, 2010.

that human understanding and reasoning is *subjective* in nature unlike propositional logic which deals with either truth or falsity of a statement. So, to deal with this kind of situation we used *subjective logic* to find out sentences which are significant in the context and can be used to summarize a document.

2 Modeling 'Opinions' about a Sentence in a Document Using Subjective Logic

In this section, we present how we formulate 'opinion' about a sentence using subjective logic. Subjective logic [3] is a logic which operates on subjective beliefs about the world, and use the term opinion to denote the representation of a subjective belief. An opinion can be interpreted as a probability measure containing secondary uncertainty, and as such subjective logic can be seen as an extension of both probability calculus and binary logic. It is a type of probabilistic logic that explicitly takes uncertainty and belief into account. It is suitable for modeling and analysing situations involving uncertainty and incomplete knowledge [3], [4].

2.1 Interpretation of Evidence in a Document

How can we define evidence in a document? This is what we are building here automatically. We consider words, phrases or co-occurrence of words, or a sentence itself to be evidence present in a document. Now, based on this, our basic motivation is to formulate 'opinion' about a proposition, which is a sentence in this case. Stronger the opinions about a sentence, more is its significance in the document. These opinions are measured by probability expectation of a sentence. Greater the probability expectation, more significant is the sentence. If probability expectation of two sentences are similar, then we need to look at the sentence with lower uncertainty to fetch the important one [4] between two.

Assumptions: We propose the following framework for the practical application of subjective logic in a document computing context.
1. All the words or terms (removing the stop words) in a document are atomic.
2. The sentences are unique, i.e., each of them occur only once in a given document.

Representation of a document: A document consists of sentences. In this paper, a sentence is considered to be a set of words. In a document, sentences are separated by stop marks (".", "!", "?"). Terms (stop words excluded) are extracted and the frequencies (i.e. number of occurrences) of the words in each sentence are calculated.

Let us now define the notations which we will be using in the rest of the equations and explanations. Θ is the frame of discernment. We represent a document as a collection of words, which is

$$\Theta = D_w = \{w_1, w_2, ..., w_n\} \tag{1}$$

where, D_w is a document consisting of words like w_1, $w_2 ... w_n$ and $|D_w| = n$. Now,

$$\rho(\Theta) = \{\{w_1\}, \{w_2\}, ..., \{w_1, w_2, w_3, ..., w_n\}\} \equiv 2^{\Theta} \tag{2}$$

$$|\rho(\Theta)| = 2^n \tag{3}$$

Since a document is a collection of sentences, it can be represented as

$$D_s = \{s_1, s_2, ..., s_m\} \tag{4}$$

where m is a finite integer and each s_i is an element of $\rho(\Theta)$. Each sentence is comprised of words, which belong to the whole word collection of the document D_w. We thus represent each sentence by,

$$S_l = \{w_i w_k ... w_r\} \in \Theta \tag{5}$$

where, $1 \leq i,\ k,\ r \leq n$ and $S_l \in \rho(\Theta)$.

2.2 Definitions of 'Subjective Logic' and Our Conceptualization

In fig.1, we present a model of a document with four sentences (s_1, s_2, s_3, and s_4) and five words (w_1, w_2, w_3, w_4, and w_5) respectively. Let frequency of occurrence of each word in each sentence be one for simplicity. The words and sentences (atomic and non atomic states) represent evidence. Now, we use the original definitions from [4], and explain our formulation.

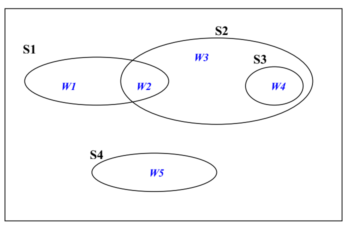


Fig. 1. Example of a document

The first step in applying evidential reasoning is to define a set of possible situations which is called the frame of discernment, Θ. A frame of discernment delimits a set of possible states of the world, exactly one of which is assumed to be true at any one time. In the given example, total number of all possible states are 2^5 for 5 words given.

Definition 1 (Belief Mass Assignment). *Let Θ be a frame of discernment. If with each substate $x \in 2^{\Theta}$ a number $m_{\Theta}(x)$ is associated such that:*

1. $m_{\Theta}(x) \geq 0$
2. $m_{\Theta}(\emptyset) = 0$
3. $\sum_{x \in 2^{\Theta}} m_{\Theta}(x) = 1$

then m_{Θ} is called a belief mass assignment in Θ, or BMA for short. For each substate $x \in 2^{\Theta}$, the number $m_{\Theta}(x)$ is called the belief mass of x.

We calculate BMA for each event by,

$$m(x) = \frac{F(x)}{Z}, \tag{6}$$

where $F(x) = \sum_{k=1}^{N} f_{x_k}$, where N is the total number of sentences in the document, $x \in 2^{\Theta}$, and f_{x_k} is the frequency of occurrence of event x in sentence k. In words, it is the total frequency of that event in all the sentences (or the whole document).

$$Z = \sum_{\substack{\forall x \neq \emptyset \\ f_x \neq 0}} F(x), \quad x \in 2^{\Theta} \tag{7}$$

Z is the total frequency of the all the events which has valid evidence of truth (whose frequency is non zero). In the given example, we have 7 valid states and their corresponding frequencies in the document are: $\{F(w_1) = 1,\ F(w_2) = 2,\ F(w_3) = 1,\ F(w_4) = 2,\ F(w_5) = 1, F(w_1, w_2) = 1,\ F(w_2, w_3, w_4) = 1\}$. Therefore, $Z = 9$ in this case. Using (6), we calculate BMA for each of the states (or events) in the given example.

Definition 2 (Belief Function). *Let Θ be a frame of discernment, and let m_Θ be a BMA on Θ. Then the belief function corresponding with m_Θ is the function $b:\ 2^{\Theta} \rightarrow [0,1]$ defined by:*

$$b(x) = \sum_{y \subseteq x} m_\Theta(y), \quad x, y \in 2^{\Theta} \tag{8}$$

We calculate the belief of a sentence of the example as, $b(s_1) = m(w_1) + m(w_2) + m(w_1, w_2)$.

Definition 3 (Disbelief Function). *Let Θ be a frame of discernment, and let m_Θ be a BMA on Θ. Then the disbelief function corresponding with m_Θ is the function $d:\ 2^{\Theta} \rightarrow [0,1]$ defined by:*

$$d(x) = \sum_{y \cap x = \emptyset} m_\Theta(y), \quad x, y \in 2^{\Theta}. \tag{9}$$

We calculate disbelief of s_1 by $d(s_1) = m(w_3) + m(w_4) + m(w_5)$.

Definition 4 (Uncertainty Function). *Let Θ be a frame of discernment, and let m_Θ be a BMA on Θ. Then the uncertainty function corresponding with m_Θ is the function $u:\ 2^{\Theta}\ [0,1]$ defined by:*

$$u(x) = \sum_{\substack{y \cap x \neq \emptyset \\ y \not\subseteq x}} m_\Theta(y), \quad x, y \in 2^{\Theta}. \tag{10}$$

From Josang's idea, we can get the **Belief Function Additivity** which is expressed as:

$$b(x) + d(x) + u(x) = 1, \quad x \in 2^{\Theta},\ x \neq \emptyset. \tag{11}$$

Now, one can simply calculate the uncertainty of a sentence by using (11), i.e., $u(s_1) = 1 - (b(s_1) + d(s_1))$.

Definition 5 (Relative Atomicity). *Let Θ be a frame of discernment and let $x, y \in 2^{\Theta}$. Then for any given $y \neq \emptyset$ the relative atomicity of x to y is the function $a : 2^{\Theta} \rightarrow [0,1]$ defined by:*

$$a(x/y) = \frac{|x \cap y|}{|y|}, \quad x, y \in 2^{\Theta},\ y \neq \emptyset. \tag{12}$$

It can be observed that $x \cap y = \emptyset \Rightarrow a(x/y) = 0$ and that $y \subseteq x \Rightarrow a(x/y) = 1$. In all other cases relative atomicity will be a value between 0 and 1. The relative atomicity of an atomic state to its frame of discernment, denoted by $a(x/\Theta)$, can simply be written as $a(x)$. If nothing else is specified, the relative atomicity of a state then refers to the frame of discernment. In this case, we get the following relative atomicity for sentence s_1 as:

$a(s_1/w_1) = \frac{|s_1 \cap w_1|}{|w_1|} = \frac{1}{1} = 1$
$a(s_1/w_2) = \frac{|s_1 \cap w_2|}{|w_2|} = \frac{1}{1} = 1$
$a(s_1/\{w_1, w_2\}) = a(s_1, s_1) = \frac{|s_1 \cap \{w_1, w_2\}|}{|\{w_1, w_2\}|} = \frac{2}{2} = 1...$
$a(s_1/w_5) = a(s_1/s_4) = \frac{|s_1 \cap w_5|}{|w_5|} = \frac{0}{1} = 0$
Likewise, we calculate the atomicity for other sentences.

Definition 6 (Probability Expectation). *Let Θ be a frame of discernment with BMA m_{Θ} then the probability expectation function corresponding with m_{Θ} is the function $E : 2^{\Theta} \rightarrow [0,1]$ defined by:*

$$E(x) = \sum_{y} m_{\Theta}(y) a(x/y), \quad y \in 2^{\Theta}. \tag{13}$$

So, for the given example, we calculate $ProbExp$ for sentence s_1 as follows: $E(s_1) = m(w_1)a(s_1/w_1) + m(w_2)a(s_1/w_2) + m(\{w_1, w_2\})a(s_1/\{w_1, w_2\}) + ... + m(w_5)a(s_1/w_5)$ For compactness and simplicity of notation we will in the following denote belief, disbelief, uncertainty, relative atomicity and opinion functions as b_x, d_x, u_x, a_x and ω_x respectively. Thus opinion (ω_{s_1} or $\omega(s_1)$)about a sentence s_1 can be expressed using these four parameters as, $\omega(s_1) = (b(s_1),\ d(s_1),\ u(s_1),\ a(x))$.

In this context, we order sentences based on descending order of their probability expectation and ascending order of their uncertainty; sentence with stronger 'opinion' has greater significance in a document.

3 Method

3.1 Data Processing

In this experiment we used DUC2001 data set [1] for evaluation. The documents are grouped based on a specific topic. Our main aim is to see how our model works on single documents for content analysis purposes, so we focussed on this kind of data set unlike other information retrieval areas. These documents were parsed, tokenized, cleaned, and stemmed. The cleaning is done by removing the stop words. DUC2001 comes with human generated summaries and baseline summaries, providing a good platform for evaluation.

3.2 Generation of Summaries

Summaries are broadly classified into text extraction and text abstraction [7], [5]. For text extraction, sentences from the documents are used as summaries and for text abstraction important pieces of information are extracted and then stitched together to form summaries following some linguistic rules. This evidence based model can be used as a text extraction as we use the original sentences. We compared our method and DUC baseline summaries with the human generated summaries provided by them.

Evidence based model (PEU): In sec.2, we described our method of sentence ranking; subjective logic based where we ranked the sentences based on the descending order probability expectation and ascending order of uncertainty (PEU) of that sentence in the document. We took 30% [2] of the top ranked sentences and used them as summary.

3.3 Evaluation by ROUGE

ROUGE [6] stands for Recall-Oriented Understudy for Gisting Evaluation. It includes measures to automatically determine the quality of a summary by comparing it to other (ideal) summaries created by humans. ROUGE is a recall based metric for fixed length summaries. The measures count the number of over lapping units such as n-gram, word sequences, and word pairs between the computer-generated summary to be evaluated and the ideal summaries created by humans.

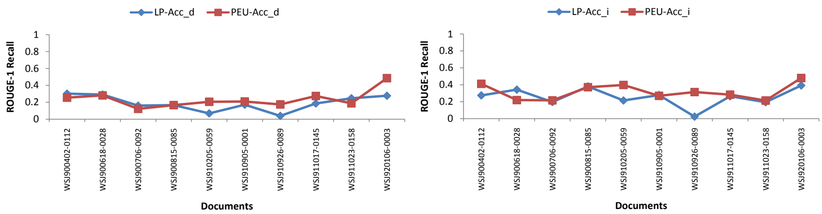
In this experiment, we present the result with ROUGE-1 (n-gram, where n=1) at 95% confidence level. ROUGE is sensitive to the length of the summaries [8]; hence we fixed the length to 100 words for the evaluation.

4 Results

We used DUC2001 dataset for this experiment. Among different document sets, we presented here the evaluation with 'daycare', 'healthcare' and 'pres92'. We compared our method (PEU) and baseline summaries (denoted by LP) with two different human assessors. For each set the assessors are different. The average (table.1) results show that our method out performs the baseline summaries.

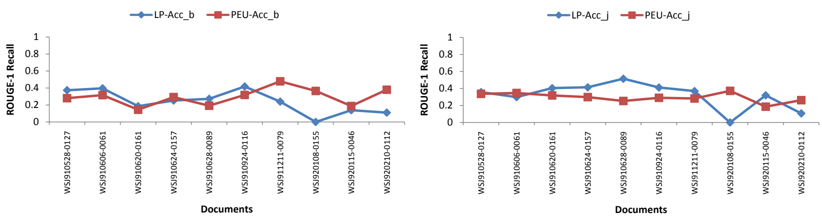
Figures 2(a) and 2(b) show the evaluation comparison using daycare data, where Acc_d and Acc_i are two human assessors. In both the figures 2(a) and 2(b), our method outperforms the baseline summaries (90% of documents with Acc_d and 80% of documents with Acc_i).

Now, figures 3(a) and 3(b) present the results with healthcare data set. Here Acc_b and Acc_j are the two human assessors. There was no baseline summary for the 7th document in this series. So, in both the figures we have value as 0 in the comparison results. In fig.3(a), ROUGE score for PEU with Acc_b is higher



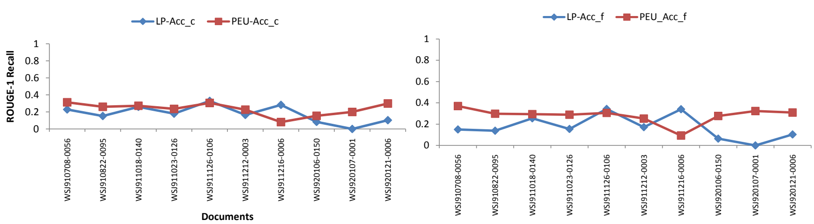
(a) Comparison with Assessor d (b) Comparison with Assessor i

Fig. 2. Daycare dataset



(a) Comparison with Assessor b (b) Comparison with Assessor j

Fig. 3. Healthcare dataset



(a) Comparison with Assessor c (b) Comparison with Assessor f

Fig. 4. Pres92 dataset

than baseline on average except for 30% of the documents. In fig.3(b), baseline shows higher similarity with Acc_j than PEU in 60% documents.

In figures 4(a) and 4(b), PEU outperforms the baseline summaries (90% of documents with each assessors). In table.1, Acc1 and Acc2 are alias of human assessors used in each case. Except for Acc2 in healthcare data, PEU has outperformed all the baseline summaries.

From these results, we can see that summaries produced by humans are abstract. So overlap with human generated of summaries with automated ones can vary a lot unless they are compared with extractive summaries created by humans selecting the original sentences from documents.

Table 1. Summary of all three sets of results (ROUGE-1 Recall)

	LP-Acc1	LP-Acc2	PEU-Acc1	PEU-Acc2
daycare	0.19	0.26	0.24	0.32
healthcare	0.27	0.35	0.29	0.29
*pres*92	0.20	0.19	0.24	0.28
Avg	0.22	0.27	0.25	0.29

5 Conclusion

In this paper we presented an evidence based sentence extraction method for single document summarization. We used enhanced subjective logic to formulate the whole process. Here standard methods for evaluating data are used; in the whole process we figured out that summarization is subjective to the user. In our system we basically used word frequency and co-occurrence concept for formulating subjective logic; rather superficial knowledge. But the results are good in the sense that they have outperformed baseline summaries as illustrated in the results. For our future work we will extend this method to perform deeper semantic analysis of the text and redefine some features of subjective logic in document computing context.

References

1. DUC (2001), The document understanding conference, `http://duc.nist.gov/`
2. Dalianis, H.: SweSum-A Text Summarizer for Swedish (2000), `http://www.dsv.su.se/%7Ehercules/papers.Textsumsummary.html`
3. Jøsang, A.: Artificial reasoning with subjective logic. In: Proceedings of the Second Australian Workshop on Commonsense Reasoning, vol. 48 (1997) Perth:[sn]
4. Jøsang, A.: A logic for uncertain probabilities. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 9(3), 279–311 (2001)
5. Liddy, E.D.: The discourse-level structure of empirical abstracts: an exploratory study. Inf. Process. Manage. 27(1), 55–81 (1991)
6. Lin, C.-Y.: Rouge: A package for automatic evaluation of summaries, Barcelona, Spain, July 2004, pp. 74–81. Association for Computational Linguistics (2004)
7. Lin, C.Y., Hovy, E.: Identifying topics by position. In: Proceedings of the fifth conference on Applied natural language processing, pp. 283–290. Morgan Kaufmann Publishers Inc., San Francisco (1997)
8. Lin, C.Y., Hovy, E.: Automatic evaluation of summaries using n-gram co-occurrence statistics. In: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, vol. 1, p. 78. Association for Computational Linguistics (2003)
9. Mani, I., Maybury, M.T.: Advances in automatic text summarization. MIT Press, Cambridge (1999)
10. Pardo, T.A.S., Rino, L.H.M., Nunes, M.G.V.: Extractive summarization: how to identify the gist of a text. In: The Proceedings of the 1st International Information Technology Symposium–I2TS, Citeseer, pp. 1–6 (2002)