

# The Decomposition of Human-Written Book Summaries

Hakan Ceylan and Rada Mihalcea

University of North Texas  
Computer Science Department  
{hakan,rada}@unt.edu

**Abstract.** In this paper, we evaluate the extent to which human-written book summaries can be obtained through cut-and-paste operations from the original book. We analyze the effect of the parameters involved in the decomposition algorithm, and highlight the distinctions in coverage obtained for different summary types.

## 1 Introduction

Books represent one of the oldest forms of written communication. Despite this fact, given that a large fraction of the electronic documents available online and elsewhere consist of short texts such as Web pages or news articles, the focus of natural language processing techniques to date has been on the automation of methods targeting short documents. We are witnessing however a change: an increasingly larger number of books become available in electronic format, in projects such as Gutenberg, Google Books, or the Million Books project. Similarly, a large number of the books published in recent years are often available in electronic format. Thus, the need for language processing techniques able to handle very large documents such as books is becoming increasingly important.

In this paper, we focus on the problem of book summarization. In particular, we address the first step in automatic summarization, and analyze the extent to which human-written summaries of books can be obtained through extractive methods. We use a decomposition algorithm to automatically identify matches between sentences in the summary and sentences in the book, and thus determine the potential coverage of extractive summarization.

Our work is inspired by the decomposition algorithm previously proposed by Jing & McKeown for single-document summarization [4]. In this paper, we study the applicability of the algorithm to book summaries, and analyze the effect of its various parameters on the coverage of the decomposition. We show that even for long documents such as books, a significant number of summary sentences can be obtained through cut-and-paste operations from the original book. In turn, this coverage depends on the type of summaries being analyzed, with significant differences observed between objective (plot) summaries and interpretative summaries.

## 2 Related Work

To our knowledge, with the exception of [6], no research work to date was specifically concerned with the automatic summarization of books. There is, however, a large and growing body of work concerned with the summarization of short documents, with evaluations typically focused on news articles. In particular, a significant number of summarization systems have been proposed during the recent Document Understanding (DUC) / Text Analysis (TAC) conferences – evaluations that usually draw the participation of 20–30 teams every year.

In terms of analysis of the composition of human-written summaries, the work most closely related to ours is the decomposition algorithm proposed in [4], which analyzed the extent to which single-document summaries of news articles are created by using cut-and-paste operations from the text. More recently, their technique has been applied to the analysis of human summaries of Japanese broadcast news [7], and has also been adapted for the analysis of multi-document human summaries [1].

A related study [5], explored the usefulness and coverage of extractive algorithms for single-document summarization. The study found that the best extractive systems are still 15%–24% below the upper bound obtained through agreement calculated over manual summaries, suggesting that there is still room for improving the quality of methods for extractive summarization.

## 3 Decomposition Algorithm

In order to analyze the sentences in human-written summaries, we used the summary sentence decomposition methodology proposed by Jing & McKeown in [4]. This methodology is based on the assumption that the human summarizers often extract phrases from the original source, and then make further editions to compose the summary sentence. The technique, also referred to as *cut-and-paste*, is supported by the studies in [2] and [3].

The goal of the decomposition analysis of a summary sentence is to discover the original sentences from which the cut-and-paste is done. The task is very difficult, as the extracted phrases of the source sentence can go through several transformations as a result of the editions that human summarizers perform. The transformations may make the summary sentence become quite different as compared to the phrases extracted from the source text.

### 3.1 Cut-and-Paste Operations

Jing & McKeown analyze 120 sentences from 15 different human-written summaries, and identify six major operations performed during cut-and-paste summarization. These operations are sentence reduction, sentence combination, syntactic transformation, lexical paraphrasing, generalization/specification, and reordering. Human summarizers often use one or a combination of these operations.

Determining whether a phrase in a summary sentence is a result of lexical paraphrasing or a generalization/specification of phrase(s) of the original text is a difficult problem, and it is omitted in both our and Jing & McKeown's studies. Hence, in the decomposition algorithm, we only consider the sentence reduction, sentence combination, syntactic transformation, and reordering operations.

The sentence reduction operation refers to extracting a sentence from the original source and then removing certain words or phrases from it. Sentence combination is the process of combining two or more sentences from the original source and merging them into one sentence. Note that it is possible to combine only parts of the sentences, hence this operation is often used together with the sentence reduction operation. Syntactic transformation refers to modification of the syntactic structure of a sentence, such as word reordering or passive transformations. Finally, the reordering operation is concerned with the position of the sentence in the summary with respect to the sentences in the original text that are used to construct it.

### 3.2 Problem Formulation

Based on the assumptions discussed in the previous section, the sentence decomposition problem translates into finding the words of a summary sentence inside the original text. If the words come from a single sentence, then we can conclude that either the original sentence is included as-is in the summary, or that the sentence reduction operation is used to eliminate some of the words. If the position of the words is changed with respect to the original sentence, then syntactic transformations are also involved. Further, if some of the words come from different sentences, then we can conclude that the sentence combination operation is used.

Thus the problem is formulated as follows. Each summary sentence is represented as a sequence of words  $(w_1, w_2, \dots, w_n)$ , and each word  $w_i$  can be represented as a set  $S_i$  of tuples  $(P, L)$ , where  $P$  is the position of the sentence in the source document which contains  $w_i$ , and  $L$  is the position of  $w_i$  within the sentence. The tuple  $(P, L)$  is also referred to as the document position of the word. For example, the tuple  $(4, 12)$  for a word  $w$  means that  $w$  appears in the 12th position of the 4th sentence of the source document. Hence, for each summary sentence, there are  $M = \prod_{i=1}^n |S_i|$  possible ways to compose it, where  $|S_i|$  denotes the cardinality of the set  $S_i$ . We are interested in finding the set that will select the most likely document position for each word. The next section describes the algorithm that attempts to do this task in an efficient way.

### 3.3 Algorithm

The relation between the document position of a word and the position of the preceding words can be used to assign a likelihood probability to the current document position. This likelihood can be estimated using an N-gram model. In our study, we follow [4] and use a bigram model. Hence, we specify the probability  $P(w_i = (P, L)_j | w_{i-1} = (P, L)_k)$  where  $(P, L)_j \in S_i$  and  $(P, L)_k \in S_{i-1}$ , as the

probability of the word  $w_i$  coming from document position  $(P, L)_j$  given that the word  $w_{i-1}$  comes from position  $(P, L)_k$ .

Intuitively, given two adjacent words in a summary sentence, we want to find these two words in the original text. We identify the following cases. First, we can find the words in the same sentence, next to each other, and following the same order. This would be the ideal case, and we refer to it as case 1. We can still find these words in the same sentence, in the same order, but not necessarily consecutively as there might be other words in between. This is case 2. The third option, case 3, is where we find the words in the same sentence but in different order. In the fourth option, case 4, the order is retained, however we fail to find them in the same sentence but we find them in neighboring sentences, where the neighboring is determined via a window parameter. Case 5 is the same as case 4 but this time the order is also reversed. The final option, case 6, is when we find the words in non-neighboring sentences in any order.

Each document position can be seen as a state, and each of the cases defined above can be seen as a transition from one of the states of the word  $w_i$  to one of the states of the adjacent word  $w_{i+1}$  (if the adjacent word has no states, then we consider the states of the next word,  $w_{i+2}$ ). These transitions can be represented using the bigram model, and by assigning a probability to each of the cases defined above, we define a probability for each possible transition. By building these states and transition probabilities for the entire summary sentence, we are in fact constructing a first-order Hidden Markov Model (HMM). We can use the Viterbi algorithm to find the most likely sequence of states in this HMM. The algorithm associates a probability for the current state based on the probability of the previous states and the transition probability from the previous state to the current state. Hence we define a probability for state  $w_i = (P, L)_j$  as  $P(w_i = (P, L)_j) = \max_{(P, L)_k \in S_{i-1}} P(w_{i-1} = (P, L)_k) \times P(w_i = (P, L)_j | w_{i-1} = (P, L)_k)$ . We also define  $P(w_1 = (P, L)_j) = 1$  for all  $j$ , which makes sure that every first word of a sentence has an equal chance of being selected. Now we can approximate the probability  $\max P(w_1, w_2, \dots, w_n)$  as  $\max_{(P, L)_j \in S_i} P(w_n = (P, L)_j)$ . In order to find the most likely sequence, we keep a back pointer to the previous state that is selected in each step.

For cases 1 through 5, we use the same probabilities as given in [4], namely for case 1 the probability is 1.0, for case 2 is 0.9, for case 3 is 0.8, for case 4 is 0.7 and for case 5 is 0.6. For case 6, we use the probability as a parameter, and change it during our analysis. Note that these probabilities are not tuned for a particular task, but rather just assigned intuitively.

## 4 Data Set

Unlike the summarization of short documents, which benefits from the data sets made available through the annual DUC/TAC evaluations, there are no publicly available data sets that can be used for the evaluation of methods for book summarization. The lack of such data sets is not surprising since even for humans the summarization of books is more difficult and time consuming than the summarization of short news documents.

We use a strategy similar to [6], and construct a data set starting from the observation that several English and literature courses make use of books that are sometimes also available in the form of abstracts, meant to ease the access of students to the content of the books. Many of these books are classics that are already in the public domain, and thus for most of them we are able to find the online electronic version of the books on sites such as Gutenberg or Online Literature. Unlike [6], who only used two publishers of summaries, we have identified and used nine different publishers that make summaries available online for books studied in the U.S. high-school and college systems.

For our analysis, we target two kinds of summaries. First, we are interested in summaries that describe the plot of the book, in a manner as objective as possible, without any interpretation from the summary writer. We refer to these summaries as *objective summaries*. Second, we are also interested in analysis summaries that describe, in a subjective manner, the interpretation of the facts and of the main story in the book. The publishers often provide these summaries under the *Notes/Analysis/Interpretation* sections. We refer to these summaries as *interpretative summaries*. Objective and interpretative summaries are collected at chapter level for each book.

Figure 1 shows two samples drawn from an objective and an interpretative summary for “The Red Badge of Courage,” as made available by Cliff’s Notes.

---

OBJECTIVE SUMMARY: *As the novel opens, the soldiers of a regiment are waiting for battle. After one of the men, a tall soldier, suggests that a battle is imminent, other soldiers argue against the notion. One of the young soldiers, Henry, a private, returns to the hut where the regiment is camped and thinks about war. He recalls his desire to enlist in the army, his mother’s refusal to support the idea, and his eventual decision to enlist over her objections.*

---

INTERPRETATIVE SUMMARY: *The overriding impression of this first chapter is one of conflict. The Union soldiers await a physical battle with the Confederate troops in the area. The eminent external conflict is paralleled by the fight raging in Henry’s mind. As the book opens, the reader sees the main character, a soldier waiting for his first battle, ironically engaged in an internal conflict with his own thoughts.*

---

**Fig. 1.** Sample objective and interpretative summaries

From the entire set of 64 books that we found with summaries available from at least two publishers, we selected only those that had both objective and interpretative summaries available online. Moreover, we also placed a constraint on the length of these summaries, and require that the interpretative summaries be at least as long as the corresponding objective summaries, when considering the same source.

This left us with a final data set of 31 books, which is used in the analyses described below. The books in this collection have an average length of 87,000

words. The average length for the objective and interpretative summaries is 6,800 words per summary.

## 5 Parameter Analysis for the Decomposition Algorithm

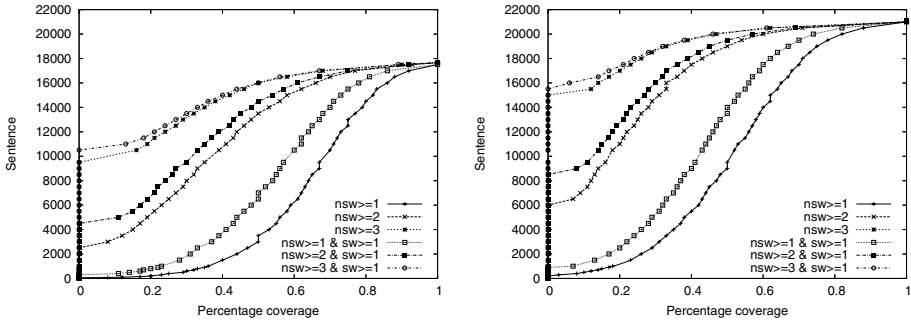
The data set used in [4] is composed of news articles and their summaries, and the parameters were selected intuitively based on this data genre. In this study, rather than taking the same parameters for granted, we decided to see the effects of the parameters when the decomposition algorithm is applied on book summaries. Therefore we start our evaluation by analyzing these parameters. Note that our goal is not to tune these parameters for a specific application, but rather to see how different parameter values affect the decomposition algorithm.

There are three parameters whose value can affect the algorithm. First, the window size parameter  $w$ , which is the number of neighboring sentences that can be considered during a sentence combination operation alongside the current document sentence. Second, the probability  $p$  of finding a word outside the current document sentence or outside the neighboring sentences that are within the window size. Finally, the last parameter is a rule that specifies the number of stop words,  $sw$ , and/or the number of non-stop words,  $nsw$ , that need to be found from a document sentence in order to consider that sentence as being involved in the cut-and-paste process. The values used in [4] for these parameters are  $w = 3$ ,  $p = 0.5$ , and for the final rule, they enforced that a document sentence has to contribute to the summary sentence with either two or more non-stop words or a non-stop word plus one or more stop words, which in logical form can be written as  $nsw \geq 2 \vee (nsw \geq 1 \wedge sw \geq 1)$ .

Moreover, in their analysis Jing & McKeown [4] assumed that a summary sentence is formed as a result of cut-and-paste operations on document sentences if and only if at least half of the words in the summary sentence can be found in document sentences. Rather than using the same assumption, we instead compute a *coverage* for each summary sentence, which is simply the percentage of the words in the summary sentence that are found in the source document. This allows us to create plots and visualize the effect of varying coverage. In addition, we can always specify a cutoff value for the coverage and discard the sentences that fall below the cutoff. Note that when we have a cutoff value of 0.5, we would be making the same assumption as [4].

### 5.1 Number of Words in Contributing Document Sentences

We start by analyzing the third parameter, namely the number of stop words and non-stop words in a contributing document sentence. We note that even at chapter level, the books have significantly greater length than news articles. Since there is a substantially larger number of sentences, the probability of finding a summary word or phrase in multiple sentences is also greatly increased. Therefore we start our experiments by gradually strengthening the conditions, making them more restrictive. Meanwhile, we keep the parameters  $w$ , and  $p$  fixed to 3



**Fig. 2.** (a) Objective summaries (b) Interpretative summaries

and 0.5 respectively, which are the values used in [4]. As our least restrictive rule, we start with the simple condition,  $(nsw \geq 1)$ . The next rules in order of strength are:  $(nsw \geq 2) \vee (nsw \geq 1 \wedge sw \geq 1)$ ,  $(nsw \geq 2)$ ,  $(nsw \geq 3) \vee (nsw \geq 2 \wedge sw \geq 1)$ ,  $(nsw \geq 3)$ ,  $(nsw \geq 4) \vee (nsw \geq 3 \wedge sw \geq 1)$ , and  $(nsw \geq 4)$ .

We show the results in Figure 2. The figure plots the coverage for both the objective and interpretative summaries in our collection. For each summary sentence, we measure its coverage as the percentage of words in the sentence that are obtained from the original text through cut-and-paste operations. The Y axis shows all the sentences in all the summaries in our collection, in increasing order of their coverage; we analyze 17,665 objective and 21,077 interpretative summary sentences from 55 sources.

**Table 1.** Percentage of sentences in the objective and interpretative summaries that are constructed by cut-and-paste operations, as a function of cutoff value and word restriction rules

Rules	Cutoff					
	0.2	0.3	0.4	0.5	0.6	0.7
Objective summaries						
$(nsw \geq 1)$	98.9	96.7	92.4	83.9	66.0	41.9
$(nsw \geq 2) \vee (nsw \geq 1 \wedge sw \geq 1)$	95.9	90.0	80.0	65.5	42.1	20.8
$(nsw \geq 2)$	71.7	55.6	39.7	26.3	14.2	0.69
$(nsw \geq 3) \vee (nsw \geq 2 \wedge sw \geq 1)$	62.2	46.3	31.7	20.1	10.6	0.54
$(nsw \geq 3)$	37.2	25.9	17.0	10.9	0.61	0.35
$(nsw \geq 4) \vee (nsw \geq 3 \wedge sw \geq 1)$	33.6	23.6	15.6	10.0	0.57	0.33
$(nsw \geq 4)$	19.9	15.0	10.4	0.71	0.42	0.26
Interpretative summaries						
$(nsw \geq 1)$	94.7	87.5	74.9	57.7	35.0	17.0
$(nsw \geq 2) \vee (nsw \geq 1 \wedge sw \geq 1)$	88.5	75.5	57.5	37.8	18.2	0.78
$(nsw \geq 2)$	48.2	30.7	17.9	10.2	0.53	0.30
$(nsw \geq 3) \vee (nsw \geq 2 \wedge sw \geq 1)$	40.5	24.7	14.1	0.81	0.44	0.27
$(nsw \geq 3)$	19.5	12.1	0.73	0.48	0.31	0.20
$(nsw \geq 4) \vee (nsw \geq 3 \wedge sw \geq 1)$	17.8	11.2	0.68	0.46	0.30	0.20
$(nsw \geq 4)$	10.5	0.75	0.51	0.37	0.26	0.18

For both the objective and interpretative summaries, increasing the restrictiveness of the conditions results in less coverage. Further, although the results are similar for both objective and interpretative summaries, as the coverage rate increases, the drop in the number of sentences is much quicker for the interpretative summaries. The curves for the objective summaries are less steep because the cutoff affects these summaries to a lesser extent than the interpretative ones.

In order to see this difference, we apply different cutoffs on these results, and report in Table 1 the percentage of the sentences that exceed the cutoff value. This value represents the percentage of the sentences in the summary that are constructed by cut-and-paste operations based on the assumptions.

The difference between the objective and interpretative summaries is clear from the results in Table 1. For example, by applying the second rule and a cutoff of 0.5 (the same assumptions as in [4]), we can identify 65.5% of the objective summary sentences as constructed by cut-and-paste operations. Using the same assumptions, the number dramatically reduces to 37.8% for interpretative summaries. For the remainder of the analyses in this section, for consistency with [4], we fix this parameter to the second rule,  $(nsw \geq 2) \vee (nsw \geq 1 \wedge sw \geq 1)$ .

### 5.2 Probability Assigned to Distant Words

Next, we analyze the parameter  $p$ , which specifies the bias of the algorithm to consider combinations of phrases from sentences that are distant from each other, i.e., outside the window size  $w$ . In order to see this bias, we vary  $p$  in 0.1 increments from 0.0 to 0.5 while keeping the other parameters fixed.

We plot the results in Figure 3. Once again, we show the coverage of the sentences in the summaries, for both the objective and the interpretative summaries in our data set. A summary of the results are also displayed in Table 2, which shows the percentage of sentences in the objective and interpretative summaries that are constructed by cut-and-paste operations.

As seen in the figure and in the table, the coverage is greatly effected when  $p$  is reduced to 0, which only allows for the combination of sentences found within the

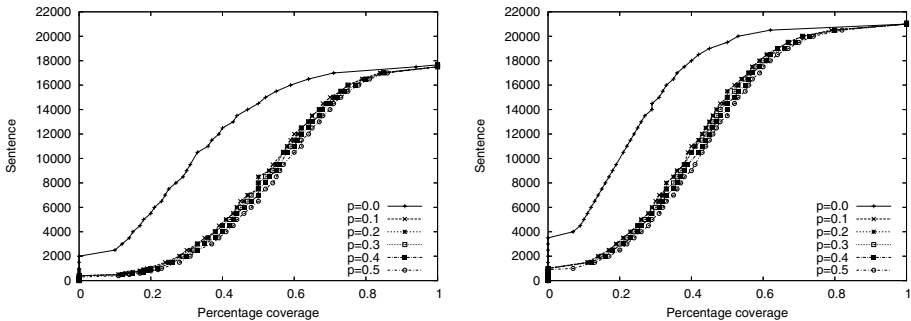


Fig. 3. (a) Objective summaries (b) Interpretative summaries



**Table 2.** Percentage of sentences in the objective and interpretative summaries that are constructed by cut-and-paste operations varying with cutoff and  $p$

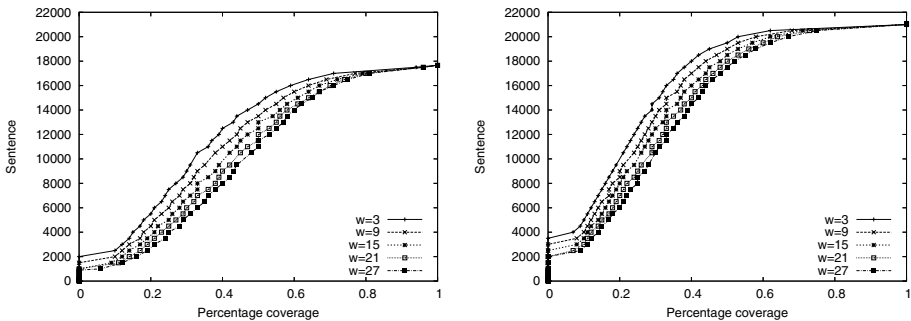
Probability	Cutoff					
	0.2	0.3	0.4	0.5	0.6	0.7
Objective summaries						
$p = 0.0$	70.1	49.4	31.1	18.8	0.91	0.44
$p = 0.1$	94.3	86.5	74.2	57.8	34.6	16.3
$p = 0.2$	94.4	86.8	74.6	58.4	35.1	16.6
$p = 0.3$	94.8	87.5	75.9	56.9	36.5	17.2
$p = 0.4$	95.4	88.7	77.7	62.5	38.8	18.8
$p = 0.5$	95.9	90.0	80.0	65.5	42.1	20.8
Interpretative summaries						
$p = 0.0$	54.0	30.8	15.6	0.75	0.34	0.20
$p = 0.1$	85.5	70.3	50.0	30.7	14.1	0.60
$p = 0.2$	85.7	70.7	50.6	31.2	14.4	0.61
$p = 0.3$	86.3	71.8	52.2	32.5	15.1	0.64
$p = 0.4$	87.4	73.4	54.8	34.9	16.5	0.69
$p = 0.5$	88.5	75.5	57.5	37.8	18.2	0.78

specified window size. This shows that there is a substantial amount of sentences in the summary that are formed by the combination of *distant* sentences.

### 5.3 Window Size

The final parameter that we are investigating is the window size  $w$ . This parameter is related to  $p$  so with increasing  $w$  we expect to gain some of the sentences that are lost when  $p$  is reduced. Thus, for this experiment, we only consider the cases where  $p = 0$ , the third parameter set to the rule  $(nsw \geq 2) \vee (nsw \geq 1 \wedge sw \geq 1)$ , and we set  $w$  to increasingly larger values for a randomly selected delta of 6, namely 3, 9, 15, 21, 27. The results are presented in Figure 4.

As expected, increasing the window size allows the algorithm to combine the sentences that are distant from each other. For example, for objective summaries,



**Fig. 4.** (a) Objective summaries (b) Interpretative summaries

with a window size of 27, and a cutoff value of 0.5, the percentage of the sentences that are found by the algorithm is increased to 39.8% compared to 18.8% when  $p = 0$  with the same parameters and cutoff value.

#### 5.4 The Decomposition of Book Summaries

In their analysis of 1,642 summary sentences, Jing & McKeown found that 42% of these sentences match to a single sentence in the document, 36% of them match to two sentences, and only 3% match to three or more sentences. They conclude that 78% of the summary sentences are constructed by cut-and-paste operations by only counting the ones that use one or two document sentences. Their study was however limited to news articles. In comparison, for book summaries, even at chapter level, we typically have much larger compression rates, so we believe that it is reasonable for a summary sentence to use three or more document sentences. See for instance the example in Figure 5, which is a sentence from a human-written summary for *Moby Dick* by Herman Melville, along with three sentences in the original book that contribute to this summary sentence.

---

SUMMARY SENTENCE (51 WORDS): *Flask is the last person down at the table and the first one to leave; since Flask had become an officer he had never known what it was to be otherwise than hungry, more or less, for what he eats does not relieve his hunger as keep it immortal in him.*

---

BOOK SENTENCES: *Sentence 33 (8 words): Flask was the last person down at the dinner, and Flask is the first man up.*

*Sentence 38 (20 words): Therefore it was that Flask once admitted in private, that ever since he had arisen to the dignity of an officer, from that moment he had never known what it was to be otherwise than hungry, more or less.*

*Sentence 39 (13 words): For what he ate did not so much relieve his hunger, as keep it immortal in him.*

---

**Fig. 5.** Sample summary sentence and three original sentences used to compose it

Not only this example demonstrates the fact that a summary sentence can be constructed from three document sentences, but it also shows the necessity for either a non-zero probability  $p$  or a larger window-size  $w$ .

We analyze 17,665 objective, and 21,077 interpretative summary sentences from a total of 55 sources. In order to be able to compare our results with [4], we use the parameters  $w = 3$ ,  $p = 0.5$ , and the rule  $(nsw \geq 2) \vee (nsw \geq 1 \wedge sw \geq 1)$ . We show the results in Table 3. For objective summaries, 34.5% of the sentences cannot make the cutoff and hence are not identified as constructed by cut-and-paste operations. On the other hand, 48.4% of the sentences use one to four document sentences, and 17.1% of them use five or more. For interpretative summaries, 62.2% of them cannot make the cutoff, and 25.3% use one to four document sentences, and 12.5% use five or more.

**Table 3.** Percentage of the number of document sentences used to compose the summary sentences

Summary type	Number of document sentences used				
	1	2	3	4	5 or more
Objective	5.3%	13.8%	15.6%	13.7%	17.1%
Interpretative	3.1%	6.3%	8.6%	7.3%	12.5%

## 6 Conclusions and Future Work

In this paper, we applied the sentence decomposition algorithm to book summaries, and analyzed the extent to which human-written book summaries can be obtained through cut-and-paste operations from the original book. Specifically, we concentrated on two separate chapter-level summary sources for books: one that attempts to give a summary by redescribing the events in the book in a compact form, which we refer to as objective summaries, and one that attempts to give a summary by capturing the deep meaning of the story by describing the author’s ideas and thoughts, which we refer to as interpretative summaries.

As a result of our analysis, based on certain assumptions concerning the value of the parameters used in the algorithm, we found that about 48.4% of the objective summary sentences can be reconstructed from the original document by using cut-and-paste from one to four document sentences. The same analysis leads to only 25.3% cut-and-paste sentences in the interpretative summaries. We can therefore conclude that humans use very little extraction from the original document when writing interpretative summaries, and thus extractive summarization is not-suitable for this summary type. On the other hand, about half of the human-written objective summary sentences are constructed from the original document, which indicates that extractive summarization can be used to create objective book summaries.

Our results differ from those reported by Jing & McKeown, who performed their analysis on short articles from the news domain. First, the decomposition algorithm fails to find a match for 34.5% of the objective summary sentences compared to only 19% in the study of Jing & McKeown. Although some of those sentences are still constructed by cut-and-paste, the algorithm fails to find them due to other transformations applied to the sentences. This demonstrates that heavy editing operations such as paraphrasing or generalization/specification are more frequently encountered in the construction of sentences in book summaries. Further, while Jing & McKeown found that only 3% of the summary sentences are constructed by using three or more document sentences, we find that sentences in book summaries tend to use a larger number of source sentences from the document. This is due to the lengthy nature of the books and the larger compression ratio of their summaries. In the data set we analyzed, the average compression ratio per chapter was 92.5% for objective summaries, and 89% for interpretative summaries, compared to a ratio of only 50-80% typical for shorter documents.

A decomposition analysis for book summaries not only helps us make the distinction between two summary styles (e.g., objective vs. interpretative), but it also helps us see how humans transform the document sentences into summary sentences, and which document sentences are selected for inclusion into a summary through cut-and-paste operations. In future work, we plan to use these results to train a machine learning model for book summarization. Furthermore, we also plan to extend our work to analyze the subjectivity of the objective and interpretative summary sources, and measure the chapter agreements across different summary sources.

## Acknowledgments

This work was partially supported by an award from Google Inc. We are grateful to Stan Szpakowicz for useful comments on an earlier draft of this paper.

## References

1. Banko, M., Vanderwende, L.: Using n-grams to understand the nature of summaries. In: Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics, Boston, Massachusetts, USA (2004)
2. Endres-niggemeyer, B., Neugebauer, E.: Professional summarising: No cognitive simulation without observation. In: Proceedings of the International Conference in Cognitive Science (1995)
3. Fries, U.: Summaries in newspapers: A textlinguistic investigation. In: Fries, U. (ed.) *The Structure of Texts*, pp. 47–63. Gunter Narr Verlag Tübingen (1997)
4. Jing, H., McKeown, K.R.: The decomposition of human-written summary sentences. In: *SIGIR 1999: Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 129–136. ACM, New York (1999)
5. Lin, C., Hovy, E.: The potential and limitations of sentence extraction for summarization. In: Proceedings of the HLT/NAACL Workshop on Automatic Summarization, Edmonton, Canada (May 2003)
6. Mihalcea, R., Ceylan, H.: Explorations in automatic book summarization. In: Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2007), Prague, Czech Republic (2007)
7. Tanaka, H., Kumano, T., Nishiwaki, M., Itoh, T.: Analysis and modeling of manual summarization of japanese broadcast news. In: Proceedings of the International Joint Conference on Natural Language Processing, Jeju Island, Korea (2005)