A Text Summarizer for Newspaper Articles on *Sassho-jiken*

Teiji Furugori and Rihua Lin

Department of Computer Science
The University of Electro-Communications
Chofu, Tokyo, Japan
Furugori@cs.uec.ac.jp

Abstract: We describe a text summarization system for Japanese newspaper articles on *sassho-jiken* (murders and bodily harms). The summary production is done through five processes: preprocessing, information extraction, conceptualization, sentence generation, and postprocessing. The qualities of summaries are sometimes rough, but coherent, and the system proves its effectiveness of understanding and generating sentences in Japanese.

1 Introduction

Documents we produce are said to double every ten years and, in the Internet society, the amount of materials we need to read have exceeded our capacity of reading them. Here is a reason for renewed interest in text summarization by computers [MM01, MM02].

Automatic text summarization can be extractive or generative. We can produce a summary by extracting certain sentences from text or generating sentences that reflect the content in the text. The the former is basically very mechanical and easy to implement. But it has intrinsic deficiencies in cohesiveness and coherence. The latter may be cognitive and difficult to implement. However, along this line lies a real hope that the summary produced may reach that of human beings in quality.

Human summarizers interpret source text, internalize its content, and produce summaries using the internal structure. This is not usually the approach the text summarizers by computers have taken, however. Here, the researchers use statistical quantities (e.g., word frequencies) or linguistic cues (e.g., cue phrases) and select important sentences to form summaries. There are a few studies that take the interpretation approach in text summarization. DeJong [DJ01], for an example, produced a summary from news about events (e.g., earthquake) using event schemata. The work on progress by Hovy and Lin [HL01] seems to be a robust and well-rounded system to produce summaries both in extraction and generation.

In this paper, we describe a system that produces summaries from Japanese newspaper articles on *sassho-jiken* (murders and bodily harms). It "understands" the content of an article, extracts necessary information, builds conceptual structure, and finally generates summaries from it.

2 Analysis of *Sassho-jiken* Articles

A text is not a mere sequence of sentences. It has certain structure [MT01]. It is obvious that we fail to take its meaning when randomizing the sentences in it.

An observation reveals that a *sassho-jiken* article consists of three scenes: Main-event, Background, and Consequence. Each of them consists of a sequence of Actions and States. Main-event has main-events or actions, where $0<i\leq 2$. Background or Consequence has j backgrounds or k consequences, or actions and/or states, where $j,k\geq 0$. Main-events describe the nature of a *sassho-jiken*, e.g., who did what's to whom and why. Backgrounds explain the surrounding situations on Main-event, e.g., "the suspect was poor." Consequences typically consist of two kinds of actions: those of the police and the court, e.g., "the police arrested the suspect."

We find a number of participants in the scenes. There are Person, Time, and Place involved in a *sassho-jiken*. Each of them has its members. A person in a *sassho-jiken* belongs to one of three types: victim, offender, and third-party person. The type of a person may be identified in Japanese by the suffix attached to his or her name. If it is one of さん, ちゃん, and くん, it refers to either victim or third-party person. An offender appears without suffix, or with such nouns as 容疑者 (suspect) and 犯人 (culprit) attached right after his or her name.

In Japanese, a participant, including person, is identified through its position, the case marker attached to it, the selectional restriction of the verb (or its nominal form) associated with it in a sentence, and/or the types of nouns. For instance, when we see a sentence, Xが Yを刺す, (X stabs Y.), we know that the verb 刺す (*to stab*) takes at least two cases: offender and victim. The offender assumes subject position with case marker が or は and the victim assumes the object position with case marker を. In another example, 警察がXを逮捕. (The police arrests X.), we know that the verb 逮捕 (nominal form of *to arrest*) must have a public-office (in this case the police) as its subject position and the suspect or offender as its object position.

We see that many of the verbs and nouns associated with each scene are rather distinctive, except for backgrounds. For example, 刺す indicates a main-event and 逮捕 a consequence.

Each participant appears with certain attributes accompanied to it in a *sassho-jiken* article. A person appears with some or all of its name, age, sex, occupation, and address. A time appears with some or all of its year, month, day, and clock. A place appears with some or all of its prefecture, city, section, and number.[1]

3 Processes of Summarization

We extract information on the scenes that constitute a *sassho-jiken*, represent it in frames that are interconnected, and generate summaries from the frame representation. We do preprocessing and postprocessing before and after these processes.

Preprocessing and postprocessing may be necessary in any type of document processing. For our summarization, we eliminate という (*it is reported*) expression and replace 同署 (*the same police station*) and 同所 (*the same place*) with their proper department and place names in the preprocessing process. We change the sentences in a summary to the past tense obligatory and passivize the sentences, when necessary.

The Article 1 below shows a typical example appeared on a newspaper. In this article,

---

[1] Attributes attached to the participants are culture-dependent. For instance, the age, sex, occupation, etc. are considered very important to describe the persons involved in incidents and accidents in Japan. It may not be so in other countries.

only one type of elimination is to occur for the portions boldfaced.

Article 1 二十九日午前十一時ごろ、横須賀市小矢部二丁目の市道で、近くに住む男性(78)が男二人に棒で頭を殴られ、持っていたバッグを奪われた。男性によると、バッグには現金五十万円が入っていたという。男は二十歳一三十歳くらい。男は逃走した。横須賀署が強盗事件として調べている。調べでは、男性は買い物をして帰る途中だったという。

We put the content of the article in frames (See Fig. 1). To do so, we perform morphological analysis by JUMAN, a morphological analyzer [KN01], and sentence analysis by checking, chunk by chunk, the types of words with a dictionary we provided, paying particular attention to the words for participants and verbs for actions and states [SI.01]. Here, we note a special provision: we put any sentence as it is in the conceptual structure when we fail to analyze it, or find certain clue word that signifies an importantness of the sentence in the *sassho-jiken*, e.g., 調べでは (*the Police reports that...*).

Scenes
Main-event: **Main-event1**
Background: **Background1**
Consequence: **Consequence1**
Actions and States
Main-event1 [e-type]: 殴る; [offender]:Person2; [victim]:Person1; [arms]: 棒; [place]:Place1; [time]:Time1; [reason]: ; [deprivation]: ;
Main-event2 [e-type]: 奪う; [offender]:Person2; [victim]:Person1; [object-robbed]: バッグ; [place]:Place1; [time]:Time1; [reason]: ; [deprivation]: ;
Background1 [b-type]: ; [content]: 男性によると、バッグには現金五十万円が入っていた。
Background2 [b-type]: 調べでは; [content]: 男性は買い物をして帰る途中だった。
Consequence1 [e-type]: 逃走する; [offender]:Person2; [origin]: ; [destination]: ;
Consequence2 [e-type]: 調べている; [public-office]: 横須賀署; [crime-committed]: 強盗;
Participants
Person1 [p-type]:victim; [name]: ; [age]: 78; [sex]: 男性; [occupation]: ; [address]: ;
Person2 [p-type]:offender; [name]: ; [age]: ; [sex]: 男; [occupation]: ; [address]: ;
Time1 [year]: ; [month]: ; [day]: 29; [clock]: 午前11;
Place1 [prefecture]: ; [city]: 横須賀; [section]: 小矢部; [number]: 2 ;

Fig. 1: Conceptual Structure of Article 1

The frames contain Main-event, Background, and Consequence. A main-event consists of 8 attributes: [e-type], [offender], [victim], [arms], [place], [time], [reason], and [deprivation]. The [e-type] takes a verb used for the main-event. The rest is all related to [e-type] and each of them assumes a value suggested by its attribute name. However, we restrict [deprivation] to be what has happened to the [victim]. The value of [offender] or [victim] is a pointer to a person frame, [place] is to a place frame, and [time] is to a time frame. [Reason] is a word that indicates the reason why the offender committed the crime and [deprivation] is a word that indicates what has happened to the victim in Main-event.

A consequence is an action or a state that is related to the police and the court. Its frame depends on the verb used for it. For instance, the verb 逮捕する (*to arrest*) takes five attributes: [e-type], [public-office], [offender], [crime-committed], [place], and

[time]. Here, [e-type] is a verb used for consequences. [Public-office] assumes the police department involved, or 'the police' as the default value when the name is not supplied. The frame for a background depends on the verb used, too.

We know by dictionary definitions how many attributes the verb for a consequence or a background takes, beside its [e-type] or [b-type] that is common to all the consequences or backgrounds.

A person frame takes five attributes: [p-type], [name], [age], [sex], [occupation], and [address]. Here, [p-type] indicates a type of person and [address] takes a pointer to a place frame. A place frame takes four attributes: [prefecture], [city], [section], and [number]. A time frame takes four attributes: [year], [month], [day], and [clock].

4 Summary Generation

We use an augmented transition network type of grammar to produce summaries. A summary is generated by combining what are in Main-event, Consequence, and Background.

The networks for verbalizing scenes and participants contain many optional rules. That is, scenes are optional, except a Main-event. Within a scene, its participants or constituents are optional. Within a constituent, many of its members are optional. Thus, we can generate a variety of phrases and sentences using them, together with the transformation rules in the postprocessing process. For instance, *X stabs Y* can be changed to:

12歳の少年が52歳の男性を刺す。(A 12 years old boy stabs a man of 52 years old.)
小学六年生の少年が会社員の鈴木太郎さんを刺す。
(A boy in the 6th grade stabs Suzuki, a salaried man.)

and many others using the network for Person. When we use a passive transformation, the last sentence is changed to:

会社員の鈴木太郎さんが小学六年生の少年に刺される。
(Taro Suzuki, a salaried man, is stabbed by a boy in the 6th grade.)

The followings are actual summaries produced from Article 1.

Summary 1 男が男性を殴った。男は男性からバッグを奪った。
Summary 2 二〇一九年午前十一時、横須賀市小矢部二丁目で、78歳の男性が二十歳ー二十歳の男に棒で殴られた。男性は男にバッグを奪われた。
Summary 3 二〇一九年午前十一時、男が横須賀市小矢部二丁目で男性を棒で殴った。男性は買い物をして帰る途中だった。男は逃走した。

Summary 1 is produced using Main-event only with obligatory cases for the verbs 殴る and 奪う. Summary 2 used optional cases for Place and Time, and a passive transformation in the postprocessing process. Summary 3 chose Main-event1, Background1, and Consequence1. We say that the shortest summary we can produce in our system is roughly equivalent to a headline in a *sassho-jiken*.

A weakness of the system lies in its analytical power for analyzing sentences and expressing them in the frames used. Interestingly, however, the weakness looks like strength producing practical summaries. When we get a summary by sentence extraction, the sentences picked up tend to lose structural integrity as a text, 'wholeness' or compactness in the content and smoothness as a discourse.

The summaries produced by our system, on the other hand, look smooth as they are generated from the scenes and include some important sentences through the use of clue words or phrases, though this measure was a compromise to make up a fault in sentential analysis. The summaries are also acceptable in content as they are compact and cover the main event occurred and its surrounding actions and states.

5 Conclusion

Text has its own structure. It may be possible to get a summary by sentence extraction from it. But we are unable to get a short summary by the sentence extraction. The Article 1 is typical news for a *sassho-jiken*, but even when we choose the first sentence, the summary exceeds 30% of the original text. On top of it, we are unable to produce cohesive and coherent summaries by extracting sentences. We never reach the level of human summarizers, however the methods we employ in this direction.

It is impossible to get a summary that resembles to the one by human beings without understanding and reorganizing the original text. Our system intends to implement some processes in this direction. But of course, this is a problem easier stated than implemented and in fact we had to make the big compromise in implementation: when we failed to analyze a sentence, we chose the whole sentence as either a consequence or a background and used it to produce a summary. We encountered a number of other difficulties, too: for instance, it is obvious that the frames are unable to grasp complex human relations (e.g., social as well as kinship relations among the persons involved). Nevertheless, the summaries are coherent, within the limitations, and the system shows a good way of understanding and generating sentences for *sassho-jiken's*.

Bibliography

[DJ01] DeJong, G.: An over view of the FRUMP system, In Lehnert, W.G. and Ringle, H. M.eds.: Strategies for natural language processing, Erlbaum, pp. 149-175, 1982.

[HL01] Hovy, E. and Lin, C.: Automated text summarization in SUMMARIST, " In [MM02], pp.18-24.,1982, 1997.

[KN01] Kurohashi, S. and Nagao, M.: Nihongo Keitaiso Kaiseki Sisutemu JUMAN (Japanese Morphological Analyzer), version 3.61, Kyoto University, 1999.

[MM01] Mani, I. and Maybury, M. eds.: Intelligent scalable text summarization. Universidad Nacional de Educacion a Distancia, Madrid, 1997.

[MM02] Mani, I. and Maybury, M. eds.: Advances in Automatic Text summarization, MIT Press, London, 1999.

[MT01] Mann, W. C. and Thompson, S. A.: Rhetrical structure theory: Toward a functional theory of text organization, *Text*, 8(3):243-282, 1988

[SL01] Schank, R., Lebowitz, M. and Birnbaum, L.: An Integrated Understander, Am. J. of Computational Linguistics, Vol.6, No.1, pp.13-30, .1980.