Personalized Extractive Summarization Using an Ising Machine Towards Real-time Generation of Efficient and Coherent Dialogue Scenarios

Hiroaki Takatsu[1] **Takahiro Kashikawa**[2] **Koichi Kimura**[2] **Ryota Ando**[3] **Yoichi Matsuyama**[1]

[1]Waseda University, Tokyo, Japan
[2]FUJITSU LIMITED, Kanagawa, Japan
[3]Naigai Pressclipping Bureau,Ltd., Tokyo, Japan

takatsu@pcl.cs.waseda.ac.jp {kashikawa, k.kimura}@jp.fujitsu.com
ando@naigaipc.co.jp matsuyama@pcl.cs.waseda.ac.jp

Abstract

We propose a personalized dialogue scenario generation system which transmits efficient and coherent information with a real-time extractive summarization method optimized by an Ising machine. The summarization problem is formulated as a quadratic unconstraint binary optimization (QUBO) problem, which extracts sentences that maximize the sum of the degree of user's interest in the sentences of documents with the discourse structure of each document and the total utterance time as constraints. To evaluate the proposed method, we constructed a news article corpus with annotations of the discourse structure, users' profiles, and interests in sentences and topics. The experimental results confirmed that a Digital Annealer, which is a simulated annealing-based Ising machine, can solve our QUBO model in a practical time without violating the constraints using this dataset.

1 Introduction

As mobile personal assistants and smart speakers become ubiquitous, the demand for dialogue-based media technologies has increased since they allow users to consume a fair amount of information via a dialogue form in daily life situations. Dialogue-based media is more restrictive than textual media. For example, when listening to an ordinary smart speaker, users can not skip unnecessary information or skim only for necessary information. Thus, it is crucial for future dialogue-based media to extract and efficiently transmit information that the users are particularly interested in without excess or deficiencies. In addition, the dialogue scenarios generated based on the extracted information should be coherent to aid in the proper understanding. Generating such efficient and coherent scenarios personalized for each user generally takes more time as the information source size and the number of target users increase. Moreover, the nature of conversational experiences requires personalization

in real time. In this paper, we propose a personalized extractive summarization method formulated as a combinatorial optimization problem to generate efficient and coherent dialogue scenarios and demonstrate that an Ising machine can solve the problem at high speeds.

As a realistic application of the proposed personalized summarization method for a spoken dialogue system, we consider a news delivery task (Takatsu et al., 2018). This news dialogue system proceeds the dialogue according to a primary plan to explain the summary of the news article and subsidiary plans to transmit supplementary information though question answering. As long as the user is listening passively, the system transmits the content of the primary plan. The personalized primary plan generation problem can be formulated as follows:

> From N documents with different topics, sentences that may be of interest to the user are extracted based on the discourse structure of each document. Then the contents are transmitted by voice within T seconds.

Specifically, this problem can be formulated as an integer linear programming (ILP) problem, which extracts sentences that maximize the sum of the degree of user's interest in the sentences of documents with the discourse structure of each document and the total utterance time T as constraints. Because this ILP problem is NP-hard, it takes an enormous amount of time to find an optimal solution using the branch-and-cut method (Mitchell, 2002; Padberg and Rinaldi, 1991) as the problem scale becomes large.

In recent years, non-von Neumann computers called Ising machines have been attracting attention as they can solve combinatorial optimization problems and obtain quasi-optimal solutions instantly (Sao et al., 2019). Ising machines can solve combi-

natorial optimization problems represented by an Ising model or a quadratic unconstrained binary optimization (QUBO) model (Lucas, 2014; Glover et al., 2019). In this paper, we propose a QUBO model that generates an efficient and coherent personalized summary for each user. Additionally, we verify that our QUBO model can be solved by a Digital Annealer (Aramon et al., 2019; Matsubara et al., 2020), which is a simulated annealing-based Ising machine, in a practical time without violating the constraints using the constructed dataset.

The contributions of this paper are three-fold:

- The ILP and QUBO models for the personalized summary generation are formulated in terms of efficient and coherent information transmission.
- To evaluate the effectiveness of the proposed method, we construct a Japanese news article corpus with annotations of the discourse structure, users' profiles, and interests in sentences and topics.
- Experiments demonstrate that a Digital Annealer, which is a simulated annealing-based Ising machine, can solve our QUBO model in a practical time without violating the constraints.

The rest of this paper is organized as follows. Section 2 discusses the related work. Section 3 overviews the annotations of the discourse structure and interest data collection. Section 4 details the proposed method. Section 5 describes the Digital Annealer. Section 6 evaluates the performance of the proposed method. Section 7 provides the conclusions and prospects.

2 Related work

2.1 Discourse structure corpus

Typical datasets for discourse structure analysis are RST Discourse Treebank (Carlson et al., 2001), Discourse Graphbank (Wolf and Gibson, 2005), and Penn Discourse Treebank (Prasad et al., 2008). RST Discourse Treebank is a dataset constructed based on rhetorical structure theory (Mann and Thompson, 1988). Some studies have annotated discourse relations to Japanese documents. Kaneko and Bekki (2014) annotated the temporal and causal relations for segments obtained by decomposing the sentences of the balanced corpus of contemporary written Japanese (Maekawa et al., 2014) based on segmented discourse representation theory (Asher and Lascarides, 2003). Kawahara et al. (2014) proposed a method to annotate discourse

relations for the first three sentences of web documents in various domains using crowdsourcing. They showed that discourse relations can be annotated in many documents over a short amount of time. Kishimoto et al. (2018) confirmed that making improvements such as adding language tests to the annotation criteria of Kawahara et al. (2014) can improve the annotation quality.

When applying discourse structure analysis results to tasks such as document summarization (Hirao et al., 2013; Yoshida et al., 2014; Kikuchi et al., 2014; Hirao et al., 2015) or dialogue (Feng et al., 2019), a dependency structure, which directly expresses the parent-child relationship between discourse units, is preferable to a phrase structure such as a rhetorical structure tree. Although methods have been proposed to convert a rhetorical structure tree into a discourse dependency tree (Li et al., 2014; Hirao et al., 2013), the generated trees depends on the conversion algorithm (Hayashi et al., 2016). Yang and Li (2018) proposed a method to manually annotate the dependency structure and discourse relations between elementary discourse units for abstracts of scientific papers, and then constructed SciDTB.

In this study, we construct a dataset suitable to build summarization or dialogue systems that transmit personalized information while considering the coherence based on the discourse structure. Experts annotated the inter-sentence dependencies, discourse relations, and chunks, which are highly cohesive sets of sentences, for Japanese news articles. Users' profiles and interests in the sentences and topics of news articles were collected via crowdsourcing.

2.2 Personalized summarization

As people's interests and preferences diversify, the demand for personalized summarization technology has increased (Sappelli et al., 2018). Summaries are classified as generic or user-focused, based on whether they are specific to a particular user (Mani and Bloedorn, 1998). Unlike generic summaries generated by extracting important information from the text, user-focused summaries are generated based not only on important information but also on the user's interests and preferences. Most user-focused summarization methods rank sentences using a score calculated considering user's characteristics and subsequently generate a summary by extracting higher-ranked sentences

(Díaz and Gervás, 2007; Yan et al., 2011; Hu et al., 2012). However, such conventional user-focused methods tend to generate incoherent summaries. Generic summarization methods, which consider the discourse structure of documents, have been proposed to maintain coherence (Kikuchi et al., 2014; Hirao et al., 2015; Xu et al., 2020).

To achieve both personalization and coherence simultaneously, we propose ILP and QUBO models to extract sentences based on the degree of user's interest and generate a personalized summary for each user while maintaining coherence based on the discourse structure.

3 Datasets

We constructed a news article corpus with annotations of the discourse structure, users' profiles, and interests in sentences and topics. Experts annotated the inter-sentence dependencies, discourse relations, and chunks for the Japanese news articles. Users' profiles and interests in the sentences and topics of news articles were collected via crowdsourcing.

3.1 Discourse structure dataset

Two web news clipping experts annotated the dependencies, discourse relations, and chunks for 1,200 Japanese news articles. Each article contained between 15–25 sentences. The articles were divided into six genres: sports, technology, economy, international, society, and local. In each genre, we manually selected 200 articles to minimize topic overlap. The annotation work was performed in the order of dependencies, discourse relations, and chunks. The discourse unit was a sentence, which represents a character string separated by an ideographic full stop.

3.1.1 Dependency annotation

The conditions in which sentence j can be specified as the parent of sentence i are as follows:

- In the original text, sentence j appears before sentence i.
- The flow of the story is natural when reading from the root node in order according to the tree structure and reading sentence i after sentence j.
- The information from the root node to sentence j is the minimum information necessary to understand sentence i.
- If it is possible to start reading from sentence i, the parent of sentence i is the root node.

3.1.2 Discourse relation annotation

A discourse relation classifies the type of semantic relationship between the child sentence and the parent sentence. We defined the following as discourse relations: *Start*, *Result*, *Cause*, *Background*, *Correspondence*, *Contrast*, *Topic Change*, *Example*, *Conclusion*, and *Supplement*.

An annotation judgment was made while confirming whether both the definition of the discourse relation and the dialogue criterion were met. The dialogue criterion is a judgment based on whether the response is natural according to the discourse relation. For example, the annotators checked whether it was appropriate to present a child sentence as an answer to a question asking the cause, such as "Why?" after the parent sentence.

3.1.3 Chunk annotation

A chunk is a highly cohesive set of sentences. If a parent sentence should be presented with a child sentence, it is regarded as a chunk.

A *hard chunk* occurs when the child sentence provides information essential to understand the content of the parent sentence. Examples include when the parent sentence contains a comment and the child sentence contains the speaker's information or when a procedure is explained over multiple sentences.

A *soft chunk* occurs when the child sentence is useful to prevent a biased understanding of the content of the parent sentence, although it does not necessarily contain essential information to understand the parent sentence itself. An example is explaining the situation in two countries related to a subject, where the parent sentence contains one explanation and the child sentence contains another.

3.2 Interest dataset

Participants were recruited via crowdsourcing. They were asked to answer a profile questionnaire and an interest questionnaire. We used 1,200 news articles, which were the same as those used in the discourse structure dataset. We collected the questionnaire results of 2,507 participants. Each participant received six articles, one from each genre. The six articles were distributed so that the total number of sentences was as even as possible across participants. Each article was reviewed by at least 11 participants.

3.2.1 Profile questionnaire

The profile questionnaire collected the following information: gender, age, residential prefecture, occupation type, industry type, hobbies, frequency of checking news (daily, 4–6 days a week, 1–3 days a week, or 0 days a week), typical time of day news is checked (morning, afternoon, early evening, or night), methods to access the news (video, audio, or text), tools used to check the news (TV, newspaper, smartphone, etc.), newspapers, websites, and applications used to check the news (Nihon Keizai Shimbun, LINE NEWS, SNS, etc.), whether a fee was paid to check the news, news genre actively checked (economy, sports, etc.), and the degree of interest in each news genre (not interested at all, not interested, not interested if anything, interested if anything, interested, or very interested).

3.2.2 Interest questionnaire

Participants read the text of the news article and indicated their degree of interest in the content of each sentence. Finally, they indicated their degree of interest in the topic of the article. The degree of interest was indicated on six levels: 1, not interested at all; 2, not interested; 3, not interested if anything; 4, interested if anything; 5, interested; or 6, very interested.

4 Methods

We propose an integer linear programming (ILP) model and a quadratic unconstraint binary optimization (QUBO) model for the personalized summary generation in terms of efficient and coherent information transmission.

4.1 Integer linear programming model

We considered a summarization problem, which extracts sentences that user u may be interested in from the selected N documents and then transmits them by voice within T seconds. The summary must be of interest to the user, coherent, and not redundant. Therefore, we formulated the summarization problem as an integer linear programming problem in which the objective function is defined by the balance between a high degree of interest in the sentences and a low degree of similarity between the sentences with the discourse structure as constraints. This is expressed as

$$\text{max.} \quad \sum_{k \in D_N^u} \sum_{i < j \in S_k} b_{ki}^u b_{kj}^u \left(1 - r_{kij}\right) y_{kij} \quad (1)$$

Table 1: Variable definitions in the interesting sentence extraction method

x_{ki}	Whether sentence s_{ki} is selected
y_{kij}	Whether both s_{ki} and s_{kj} are selected
b_{ki}^u	Degree of user u's interest in s_{ki}
r_{kij}	Similarity between s_{ki} and s_{kj}
t_{ki}	Utterance time of s_{ki} (seconds)
T	Maximum summary length (seconds)
L	Maximum bias in the number of extracting sentences between documents
$f_k(i)$	Function that returns the parent ID of s_{ki}
D_N^u	IDs of the selected N documents for user u
S_k	Sentence IDs contained in document d_k
C_{km}	Sentence IDs contained in chunk m of d_k

s.t.

$$\forall k, i, j: \quad x_{ki} \in \{0, 1\}, \ y_{kij} \in \{0, 1\}$$

$$\sum_{k \in D_N^u} \sum_{i \in S_k} t_{ki} x_{ki} \leq T \quad (2)$$

$$\forall k < l: \quad \left| \sum_{i \in S_k} x_{ki} - \sum_{i \in S_l} x_{li} \right| \leq L \quad (3)$$

$$\forall k, i: \quad j = f_k(i), \quad x_{ki} \leq x_{kj} \quad (4)$$

$$\forall k, m, i \in C_{km}: \quad \sum_{j \in C_{km}} x_{kj} = |C_{km}| \times x_{ki} \quad (5)$$

$$\forall k, i, j: \quad y_{kij} - x_{ki} \leq 0 \quad (6)$$

$$\forall k, i, j: \quad y_{kij} - x_{kj} \leq 0 \quad (7)$$

$$\forall k, i, j: \quad x_{ki} + x_{kj} - y_{kij} \leq 1 \quad (8)$$

Table 1 explains each variable. Here, the i-th sentence of the k-th document is expressed as s_{ki}. r_{kij} represents the cosine similarity between the bag-of-words constituting s_{ki} and s_{kj}.

Equation 2 is a constraint restricting the utterance time of the summary to T seconds or less. Equation 3 is a constraint restricting the bias of the number of extracting sentences between documents to L sentences or less. Equation 4 is a constraint in which the parent s_{kj} of s_{ki} in the discourse dependency tree must be extracted when s_{ki} is extracted. Equation 5 is a constraint requiring other sentences in the chunk to be extracted when extracting s_{ki} in a chunk. Equations 6–8 are constraints that set $y_{kij} = 1$ when s_{ki} and s_{kj} are selected.

The maximum bias in the number of extracting sentences between documents L is calculated by the following formulas, which are based on the maximum summary length T, the number of documents N, and the average utterance time of the

sentences $\bar{t}$.

$$L = \left\lfloor \frac{\bar{n}}{\sqrt{N}} + 0.5 \right\rfloor \quad (9)$$

$$\bar{n} = \frac{T}{\bar{t} \times N} \quad (10)$$

$\bar{n}$ represents the expected number of sentences to be extracted from one document. L is the value obtained by dividing $\bar{n}$ by the square root of the number of documents and rounding the result.

4.2 Quadratic unconstrained binary optimization model

To solve the combinatorial optimization problem with an Ising machine, the problem must be formulated with the Ising model or the quadratic unconstraint binary optimization (QUBO) model (Lucas, 2014; Glover et al., 2019). The energy functions of the Ising model and the QUBO model are described by the quadratic form of spin variables $\{-1, +1\}$ and binary variables $\{0, 1\}$, respectively. Ising machines search for the lowest energy state of the Ising model or the QUBO model. Here, we formulated the QUBO model based on the above ILP model. The energy function (or Hamiltonian) $\mathcal{H}$ of the QUBO model is defined as

$$\mathcal{H} = \mathcal{H}_0 + \lambda_1 \mathcal{H}_1 + \lambda_2 \mathcal{H}_2 + \lambda_3 \mathcal{H}_3 + \lambda_4 \mathcal{H}_4 \quad (11)$$

$$\mathcal{H}_0 = - \sum_{k \in D_N^u} \sum_{i < j \in S_k} b_{ki}^u b_{kj}^u \left(1 - r_{kij}\right) x_{ki} x_{kj} \quad (12)$$

$$\mathcal{H}_1 = \left(T - \sum_{k \in D_N^u} \sum_{i \in S_k} t_{ki} x_{ki} \right. $$
$$\left. - \sum_{n=0}^{\lfloor \log_2(T-1) \rfloor} 2^n y_n \right)^2 \quad (13)$$

$$\mathcal{H}_2 = \sum_{k < l \in D_N^u} \left(L - \left(\sum_{i \in S_k} x_{ki} - \sum_{i \in S_l} x_{li} \right) \right.$$
$$\left. - \sum_{n=0}^{\lfloor \log_2(L-1) \rfloor} 2^n z_n \right)^2 \quad (14)$$

$$\mathcal{H}_3 = \left(x_{ki} - x_{ki} x_{kj=f_k(i)} \right)^2 \quad (15)$$

$$\mathcal{H}_4 = \sum_{k \in D_N^u} \sum_{m \in C_k} \sum_{i \in C_{km}} \left(\sum_{j \in C_{km}} x_{kj} \right.$$
$$\left. - |C_{km}| \times x_{ki} \right)^2 \quad (16)$$

where y_n and z_n are slack variables introduced to convert inequality constraints into equality constraints, λ_1, λ_2, λ_3, and λ_4 are the weight coefficients for each constraint.

QUBO models are solved by a simulated annealing-based method (Aramon et al., 2019) or parallel tempering (known as replica-exchange Monte Carlo) (Matsubara et al., 2020). In these methods, multiple solution candidates can be obtained by annealing in parallel with different initial values. However, these methods do not guarantee that constraint violations will not occur. Therefore, the solution candidate with the lowest constraint violation score and the highest score of the objective function is adopted. The constraint violation score E_{total} is calculated as

$$E_{total} = E_{dep} + E_{chunk} + E_{bias} + E_{time} \quad (17)$$

$$E_{bias} = \max\left(\hat{L} - L, 0\right) \quad (18)$$

$$E_{time} = \max\left(\hat{T} - T, 0\right) \quad (19)$$

where E_{dep} is the number of dependency errors, E_{chunk} is the number of chunk errors, $\hat{L}$ is the maximum bias in the number of extracted sentences between documents, $\hat{T}$ is the total utterance time of the extracted sentences.

5 Digital Annealer

Quantum computing technologies are categorized into two types: quantum gate computers (Arute et al., 2019; Gyongyosi, 2020) and Ising machines. Quantum gate computers are for universal computing, whereas Ising machines specialize in searching for solutions of combinatorial optimization problems. Ising machines can be subdivided into two categories: quantum annealing machines (Johnson et al., 2011; Bunyk et al., 2014; Maezawa et al., 2019) and simulated annealing machines (Yamaoka et al., 2016; Okuyama et al., 2017; Aramon et al., 2019; Matsubara et al., 2020). Quantum annealing machines search solutions using quantum bits, which are made of quantum devices such as a superconducting circuit. By contrast, simulated annealing machines use a digital circuit. Digital Annealer (Aramon et al., 2019; Matsubara et al., 2020) is a type of simulated annealing machines with a new digital circuit architecture, which is designed to solve combinatorial optimization problems efficiently.

This study uses a Digital Annealing Unit (DAU) of the second generation Digital Annealer. The

DAU has an annealing mode (Aramon et al., 2019) or parallel tempering mode (Matsubara et al., 2020) to solve QUBO models. It can handle up to 4,096 binary variables with 64-bit precision or as many as 8,192 binary variables with 16-bit precision. Hereinafter, the DAU in the annealing mode is referred to as DAU-AM, and the DAU in the parallel tempering mode is referred to as DAU-PTM.

6 Experiments

Using the constructed dataset, we evaluated the performance of the personalized summarization method for dialogue scenario planning. The ILP model was solved by the branch-and-cut method[1] (Mitchell, 2002; Padberg and Rinaldi, 1991) on the CPU. Hereinafter, this method is referred to as CPU-CBC. We used CPU-CBC as a benchmark and compared the performance of CPU-CBC, DAU-AM and DAU-PTM.

6.1 Experimental setup

We used 2,117 participants data which the answer time of the six articles was between 5 and 30 minutes. The performance was evaluated for two cases. The first transmitted $N = 6$ articles in $T' = 450$ seconds and the second transmitted the top $N = 3$ articles with high interest in the topic in $T' = 270$ seconds. Sentences in the news articles was synthesized by AITalk 4.1[2] to calculate the duration of speech. The maximum summary length T was calculated as $T = T' - (N - 1) \times (q_d - q_s)$, where T' denotes the total utterance time of the primary plan, q_s denotes the pause between sentences, and q_d denotes the pause between documents. Here, $q_s = 1$ second and $q_d = 3$ seconds. The value obtained by adding q_s to the playback time of the synthesized audio file was set as t_{ki}.

The PULP_CBC_CMD solver of the PuLP[3], which is a Python library for linear programming optimization, was used to solve the ILP model. Python and PuLP versions were 3.7.6 and 2.4, respectively. The parameter for the number of threads of PULP_CBC_CMD was set to 30. The execution time of the solving function was measured on the Google Compute Engine[4] with the following settings: OS, Ubuntu 18.04; CPU, Xeon (2.20 GHz,

[1] https://projects.coin-or.org/Cbc
[2] https://www.ai-j.jp/product/voiceplus/manual/
[3] https://coin-or.github.io/pulp/
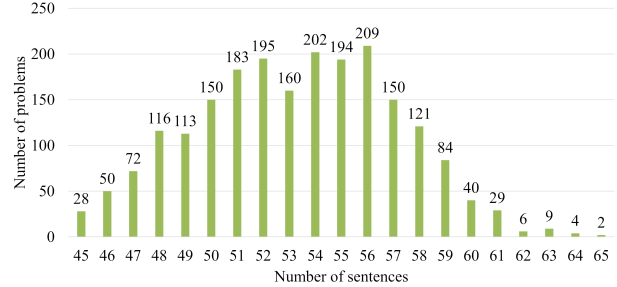[4] https://cloud.google.com/compute/?hl=en

Figure 1: Number of problems for a given number of sentences when $N = 3$
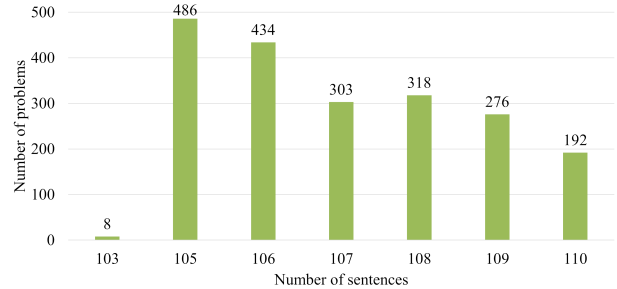

Figure 2: Number of problems for a given number of sentences when $N = 6$

32 cores); Memory, 64 GB.

Figure 1 shows the number of problems for a given number of sentences when $N = 3$. These are the top three articles with the highest degree of interest in the topic among the six articles distributed to each participant. Figure 2 shows the number of problems for a given number of sentences when $N = 6$. Since the six articles were distributed so that the total number of sentences was as even as possible across participants, the variation in the number of sentences was small.

The DAU must set the number of bits parameter from 1,024, 2,048, 4,096, or 8,192, depending on the problem size. In the experimental setting of $N = 3, T' = 270$, the 2,090 QUBO problems were less than 2,048 bits and the 27 QUBO problems were less than 4,096. On the other hand, in the experimental setting of $N = 6, T' = 450$, the 2,017 QUBO problems were less than 4,096 bits and the 100 QUBO problems were less than 8,192. In the latter experiment, these 100 participants data were excluded because the calculation precision of the DAU decreased when the problem size exceeds 4,096 bits.

The number of replicas in DAU-PTM and the number of runs of annealing in DAU-AM were 128. Since the performance of the DAU mainly depended on λ and the number of searches in one annealing (#iteration), these parameters were adjusted to prevent constraint violations.

Table 2: Information transmission efficiency of the summaries ($N = 3, T' = 270$)

	Parameters of the Digital Annealer					Efficiency of information transmission				Processing
	λ_1	λ_2	λ_3	λ_4	#iteration	Coverage	Exclusion rate	$EoIT_1$	$EoIT_2$	time (sec)
CPU-CBC (30 threads)	-	-	-	-	-	0.687	0.726	0.672	0.696	18.9
DAU-AM	10^2	10^6	10^{10}	10	10^3	0.638	0.612	0.584	0.593	0.0570
DAU-PTM	10^2	10^5	10^9	10	10^3	0.656	0.637	0.608	0.618	0.245
					10^4	0.669	0.661	0.627	0.639	1.76

Table 3: Information transmission efficiency of the summaries ($N = 6, T' = 450$)

	Parameters of the Digital Annealer					Efficiency of information transmission				Processing
	λ_1	λ_2	λ_3	λ_4	#iteration	Coverage	Exclusion rate	$EoIT_1$	$EoIT_2$	time (sec)
CPU-CBC (30 threads)	-	-	-	-	-	0.638	0.667	0.639	0.651	102
DAU-AM	10^2	10^6	10^{10}	10	10^3	0.538	0.585	0.552	0.568	0.199
DAU-PTM	10^2	10^5	10^9	10	10^3	0.553	0.577	0.556	0.565	0.749
					10^4	0.570	0.591	0.572	0.580	6.44

6.2 Evaluation metrics

We used $EoIT_\beta$ (efficiency of information transmission) (Takatsu et al., 2021) as the evaluation metric. When C is the coverage of sentences annotated as "very interested," "interested," or "interested if anything," and E is the exclusion rate of the sentences annotated as "not interested at all," "not interested," or "not interested if anything," $EoIT_\beta$ is defined based on the weighted F-measure (Chinchor, 1992) as

$$EoIT_\beta = \frac{(1 + \beta^2) \times C \times E}{\beta^2 \times C + E} \quad (20)$$

When $\beta = 2$, the exclusion rate is twice as important as the coverage. Compared to textual media, which allows readers to read at their own pace, dialogue-based media does not allow users to skip unnecessary information or skim necessary information while listening. Consequently, we assumed that the exclusion rate is more important than the coverage in information transmission by spoken dialogue and set $\beta = 2$.

6.3 Experimental results

Tables 2 and 3 show the results of $N = 3, T' = 270$ and $N = 6, T' = 450$, respectively. The values of the evaluation metrics and processing times represent the average. DAU-AM had a shorter processing time than DAU-PTM for the same number of iterations. However, in terms of the EoIT, DAU-PTM was higher than DAU-AM. The EoIT of DAU-PTM improved as the number of iterations increased, but the processing time also increased. On the other hand, increasing the number of iterations did not improve the EoIT of DAU-AM.

Figures 3 and 4 show the distributions of processing time for each number of sentences in the experimental settings of $N = 3, T' = 270$ and $N = 6, T' = 450$, respectively. The number of iterations for DAU-AM and DAU-PTM was 1,000. The processing time of CPU-CBC became longer as the number of sentences increased. Even if the number of sentences was the same, the processing time varied widely. On the other hand, the processing time of DAU-AM and DAU-PTM changed according to the size of the QUBO problems and the number of iterations, regardless of the number of sentences. At $N = 6$, the size of all QUBO problems was less than 4,096, and the processing time of DAU was almost constant regardless of the number of sentences. At $N = 3$, the size of the QUBO problems was less than 2,048 when the number of sentences was 59 or less, and less than 4,096 when the number of sentences was 63 or more. On the other hand, when the number of sentences was between 60 and 62, problems with less than 2,048 and problems with less than 4,096 were mixed due to the number of chunks.

Although the DAU is inferior to CPU-CBC in the EoIT, its processing time is considerably shorter than CPU-CBC. In addition, the processing time of the DAU can be estimated in advance based on the size of the QUBO problems and the number of iterations.

6.4 Discussion

The DAU does not guarantee that constraint violations will not occur. Although the DAU did not induce constraint violations in the parameter settings shown in the Tables 2 and 3, constraint
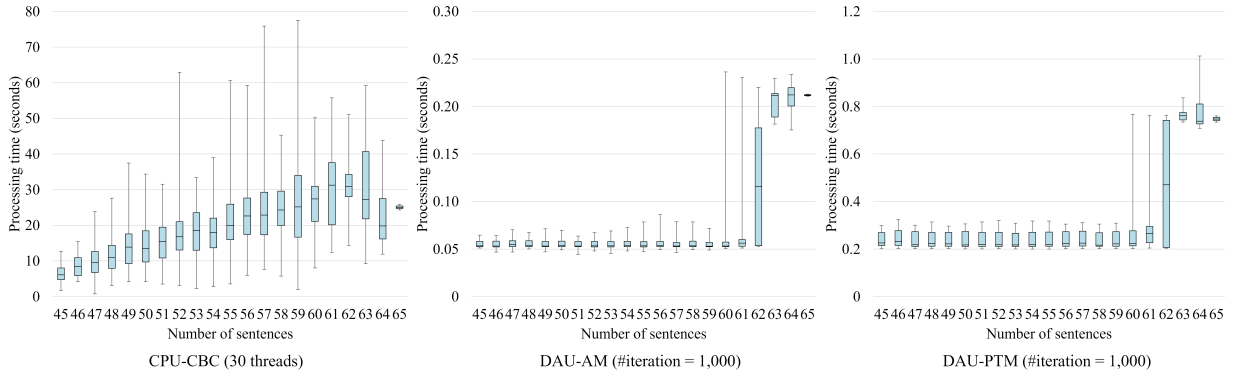
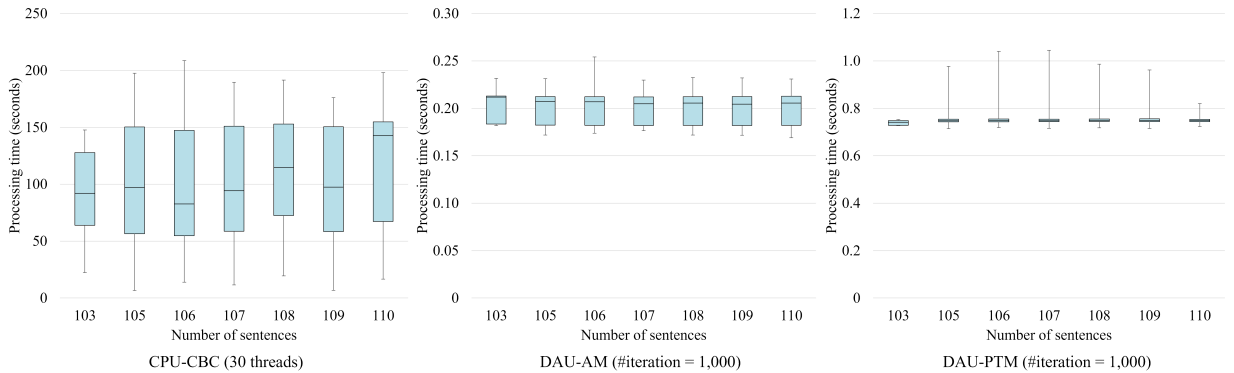Figure 3: Processing time for each number of sentences ($N = 3, T' = 270$)

Figure 4: Processing time for each number of sentences ($N = 6, T' = 450$)

violations occurred when the number of iterations was too small or the value of λ was inappropriate.

In the case that delivers six articles consisting of 15 to 25 sentences, one DAU can generate personalized summaries on the scale of 100,000 users within 6 hours since about 0.2 seconds per person was necessary to generate a summary. In other words, an application with 100,000 users can prepare personalized summaries of the previous night's news by the next morning.

The second generation Digital Annealer can only handle problems up to 8,192 bits. However, a third generation Digital Annealer, which can solve problems on the scale of 100,000 bits, is currently under development (Nakayama et al., 2021). Spoken dialogue systems capable of extracting and presenting personalized information for each user from a huge amount of data in real time will be developed in the near future.

7 Conclusion

This paper proposed a quadratic unconstraint binary optimization (QUBO) model for the real-time personalized summary generation in terms of effi-cient and coherent information transmission. The model extracts sentences that maximize the sum of the degree of user's interest in the sentences of documents with the discourse structure of each document and total utterance time as constraints. To evaluate the effectiveness of the proposed method, we constructed a Japanese news article corpus with annotations of the discourse structure, users' profiles, and interests in sentences and topics. Experiments demonstrated that our QUBO model could be solved by a Digital Annealer, which is a simulated annealing-based Ising machine, in a practical time without violating the constraints using the dataset. In the future, we will verify the effect of personalized scenarios on the spoken dialogue system.

Acknowledgments

References

Maliheh Aramon, Gili Rosenberg, Elisabetta Valiante, Toshiyuki Miyazawa, Hirotaka Tamura, and Helmut G. Katzgraber. 2019. Physics-inspired optimization for quadratic unconstrained problems using a Digital Annealer. *Frontiers in Physics*, 7(48):1–14.

Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. 2019. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505—-510.

Nicholas Asher and Alex Lascarides. 2003. *Logics of conversation: Studies in natural language processing*. Cambridge University Press.

Paul I. Bunyk, Emile M. Hoskinson, Mark W. Johnson, Elena Tolkacheva, Fabio Altomare, Andrew J. Berkley, Richard Harris, Jeremy P. Hilton, Trevor Lanting, Anthony J. Przybysz, and Jed Whittaker. 2014. Architectural considerations in the design of a superconducting quantum annealing processor. *IEEE Transactions on Applied Superconductivity*, 24(4):1–10.

Lynn Carlson, Daniel Marcu, and Mary Ellen Okurovsky. 2001. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of the 2nd SIGdial Workshop on Discourse and Dialogue*, pages 1–10.

Nancy Chinchor. 1992. MUC-4 evaluation metrics. In *Proceedings of the 4th conference on Message understanding*, pages 22–29.

Alberto Díaz and Pablo Gervás. 2007. User-model based personalized summarization. *Information Processing and Management*, 43(6):1715–1734.

Song Feng, Kshitij Fadnis, Q. Vera Liao, and Luis A. Lastras. 2019. DOC2DIAL: A framework for dialogue composition grounded in business documents. In *Proceedings of the 33rd Conference on Neural Information Processing Systems*, pages 1–4.

Fred Glover, Gary Kochenberger, and Yu Du. 2019. Quantum bridge analytics I: A tutorial on formulating and using QUBO models. *4OR: A Quarterly Journal of Operations Research*, 17(4):335—-371.

Laszlo Gyongyosi. 2020. Unsupervised quantum gate control for gate-model quantum computers. *Scientific Reports*, 10(1):1–16.

Katsuhiko Hayashi, Tsutomu Hirao, and Masaaki Nagata. 2016. Empirical comparison of dependency conversions for RST discourse trees. In *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 128–136.

Tsutomu Hirao, Masaaki Nishino, Yasuhisa Yoshida, Jun Suzuki, Norihito Yasuda, and Masaaki Nagata. 2015. Summarizing a document by trimming the discourse tree. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 23(11):2081–2092.

Tsutomu Hirao, Yasuhisa Yoshida, Masaaki Nishino, Norihito Yasuda, and Masaaki Nagata. 2013. Single-document summarization as a tree knapsack problem. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1515–1520.

Po Hu, Donghong Ji, Chong Teng, and Yujing Guo. 2012. Context-enhanced personalized social summarization. In *Proceedings of the 24th International Conference on Computational Linguistics*, pages 1223–1238.

M. W. Johnson, M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk, E. M. Chapple, C. Enderud, J. P. Hilton, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, C. J. S. Truncik, S. Uchaikin, J. Wang, B. Wilson, and G. Rose. 2011. Quantum annealing with manufactured spins. *Nature*, 473(7346):194–198.

Kimi Kaneko and Daisuke Bekki. 2014. Building a Japanese corpus of temporal-causal-discourse structures based on SDRT for extracting causal relations. In *Proceedings of the EACL 2014 Workshop on Computational Approaches to Causality in Language*, pages 33–39.

Daisuke Kawahara, Yuichiro Machida, Tomohide Shibata, Sadao Kurohashi, Hayato Kobayashi, and Manabu Sassano. 2014. Rapid development of a corpus with discourse annotations using two-stage crowdsourcing. In *Proceedings of the 25th International Conference on Computational Linguistics*, pages 269–278.
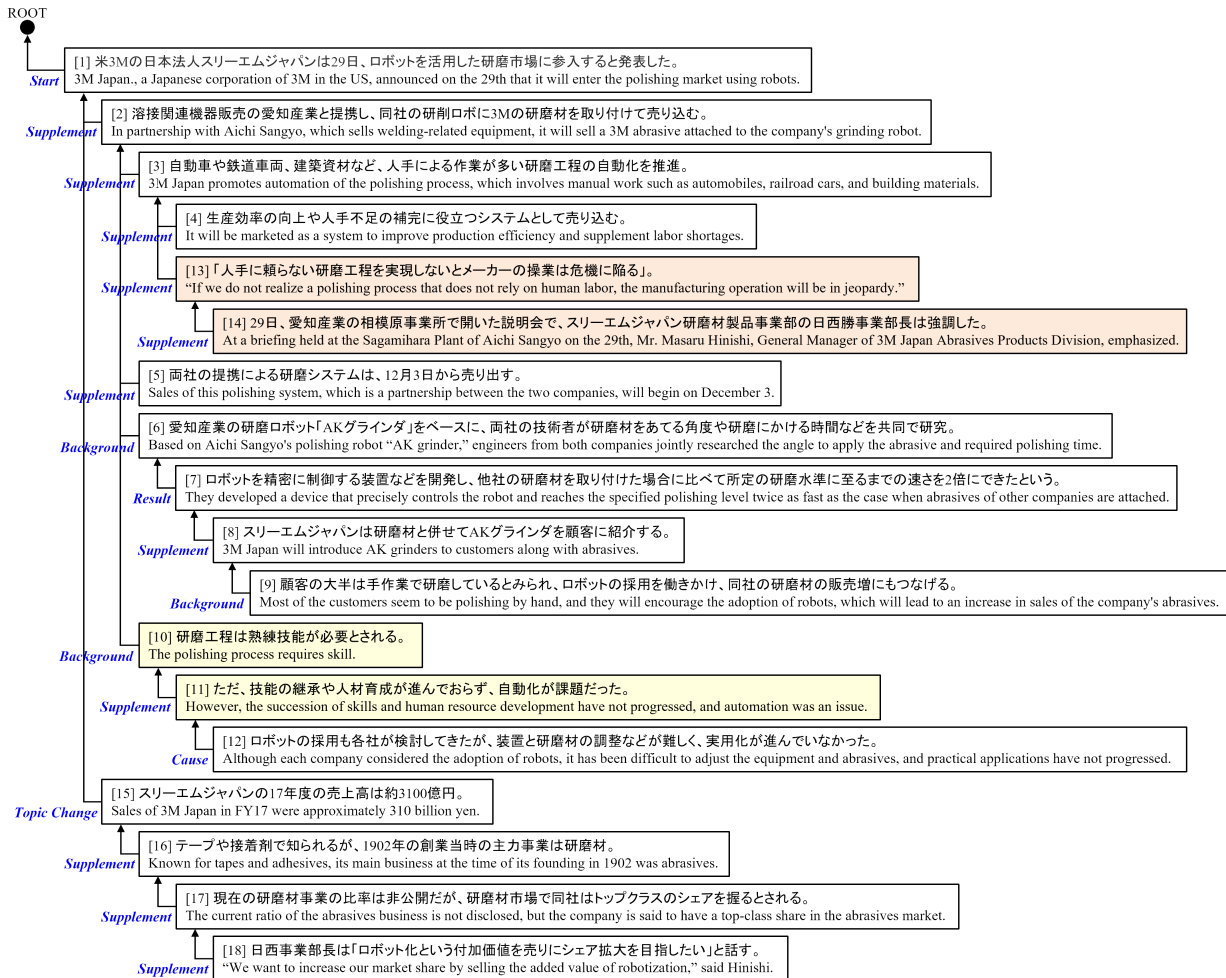
Yuta Kikuchi, Tsutomu Hirao, Hiroya Takamura, Manabu Okumura, and Masaaki Nagata. 2014. Single document summarization based on nested tree structure. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 315–320.

Yudai Kishimoto, Shinnosuke Sawada, Yugo Murawaki, Daisuke Kawahara, and Sadao Kurohashi. 2018. Improving crowdsourcing-based annotation of Japanese discourse relations. In *Proceedings of the 11th International Conference on Language Resources and Evaluation*, pages 4044–4048.

Sujian Li, Liang Wang, Ziqiang Cao, and Wenjie Li. 2014. Text-level discourse dependency parsing. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 25–35.

Andrew Lucas. 2014. Ising formulations of many NP problems. *Frontiers in Physics*, 2(5):1–15.

Kikuo Maekawa, Makoto Yamazaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka, and Yasuharu Den. 2014. Balanced corpus of contemporary written Japanese. *Language Resources and Evaluation*, 48(2):345–371.

Masaaki Maezawa, Go Fujii, Mutsuo Hidaka, Kentaro Imafuku, Katsuya Kikuchi, Hanpei Koike, Kazumasa Makise, Shuichi Nagasawa, Hiroshi Nakagawa, Masahiro Ukibe, and Shiro Kawabata. 2019. Toward practical-scale quantum annealing machine for prime factoring. *Journal of the Physical Society of Japan*, 88(6).

Inderjeet Mani and Eric Bloedorn. 1998. Machine learning of generic and user-focused summarization. In *Proceedings of the 15th National / 10th Conference on Artificial Intelligence / Innovative Applications of Artificial Intelligence*, pages 820–826.

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

Satoshi Matsubara, Motomu Takatsu, Toshiyuki Miyazawa, Takayuki Shibasaki, Yasuhiro Watanabe, Kazuya Takemoto, and Hirotaka Tamura. 2020. Digital Annealer for high-speed solving of combinatorial optimization problems and its applications. In *Proceedings of the 25th Asia and South Pacific Design Automation Conference*, pages 667—-672.

John E. Mitchell. 2002. Branch-and-cut algorithms for combinatorial optimization problems. *Handbook of Applied Optimization*, pages 65–77.

Hiroshi Nakayama, Junpei Koyama, Noboru Yoneoka, and Toshiyuki Miyazawa. 2021. Description: Third generation Digital Annealer technology. https://www.fujitsu.com/jp/group/labs/en/documents/about/resources/tech/techintro/3rd-g-da_en.pdf.

Takuya Okuyama, Masato Hayashi, and Masanao Yamaoka. 2017. An Ising computer based on simulated quantum annealing by path integral Monte Carlo method. In *Proceedings of the 2017 IEEE International Conference on Rebooting Computing*, pages 1–6.

Manfred Padberg and Giovanni Rinaldi. 1991. A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems. *SIAM Review*, 33(1):60–100.

Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. 2008. The penn discourse treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation*, pages 2961–2968.

Masataka Sao, Hiroyuki Watanabe, Yuuichi Musha, and Akihiro Utsunomiya. 2019. Application of Digital Annealer for faster combinatorial optimization. *FUJITSU SCIENTIFIC & TECHNICAL JOURNAL*, 55(2):45–51.

Maya Sappelli, Dung Manh Chu, Bahadir Cambel, David Graus, and Philippe Bressers. 2018. SMART journalism: Personalizing, summarizing, and recommending financial economic news. *The Algorithmic Personalization and News (APEN18) Workshop at ICWSM*, 18(5):1–3.

Hiroaki Takatsu, Ishin Fukuoka, Shinya Fujie, Yoshihiko Hayashi, and Tetsunori Kobayashi. 2018. A spoken dialogue system for enabling information behavior of various intention levels. *Journal of the Japanese Society for Artificial Intelligence*, 33(1):1–24.

Hiroaki Takatsu, Mayu Okuda, Yoichi Matsuyama, Hiroshi Honda, Shinya Fujie, and Tetsunori Kobayashi. 2021. Personalized extractive summarization for a news dialogue system. In *Proceedings of the 8th IEEE Spoken Language Technology Workshop*, pages 1044–1051.

Florian Wolf and Edward Gibson. 2005. Representing discourse coherence: A corpus-based study. *Computational Linguistics*, 31(2):249–287.

Jiacheng Xu, Zhe Gan, Yu Cheng, and Jingjing Liu. 2020. Discourse-aware neural extractive text summarization. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5021–5031.

Masanao Yamaoka, Chihiro Yoshimura, Masato Hayashi, Takuya Okuyama, Hidetaka Aoki, and Hiroyuki Mizuno. 2016. A 20k-spin Ising chip to solve combinatorial optimization problems with CMOS annealing. *IEEE Journal of Solid-State Circuits*, 51(1):303–309.

Rui Yan, Jian-Yun Nie, and Xiaoming Li. 2011. Summarize what you are interested in: An optimization

framework for interactive personalized summarization. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 1342–1351.

An Yang and Sujian Li. 2018. SciDTB: Discourse dependency treebank for scientific abstracts. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 444–449.

Yasuhisa Yoshida, Jun Suzuki, Tsutomu Hirao, and Masaaki Nagata. 2014. Dependency-based discourse parser for single-document summarization. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pages 1834–1839.

Appendix

ROOT

Start — [1] 米3Mの日本法人スリーエムジャパンは29日、ロボットを活用した研磨市場に参入すると発表した。
3M Japan., a Japanese corporation of 3M in the US, announced on the 29th that it will enter the polishing market using robots.

Supplement — [2] 溶接関連機器販売の愛知産業と提携し、同社の研削ロボに3Mの研磨材を取り付けて売り込む。
In partnership with Aichi Sangyo, which sells welding-related equipment, it will sell a 3M abrasive attached to the company's grinding robot.

Supplement — [3] 自動車や鉄道車両、建築資材など、人手による作業が多い研磨工程の自動化を推進。
3M Japan promotes automation of the polishing process, which involves manual work such as automobiles, railroad cars, and building materials.

Supplement — [4] 生産効率の向上や人手不足の補完に役立つシステムとして売り込む。
It will be marketed as a system to improve production efficiency and supplement labor shortages.

Supplement — [13]「人手に頼らない研磨工程を実現しないとメーカーの操業は危機に陥る」。
"If we do not realize a polishing process that does not rely on human labor, the manufacturing operation will be in jeopardy."

Supplement — [14] 29日、愛知産業の相模原事業所で開いた説明会で、スリーエムジャパン研磨材製品事業部の日西勝事業部長は強調した。
At a briefing held at the Sagamihara Plant of Aichi Sangyo on the 29th, Mr. Masaru Hinishi, General Manager of 3M Japan Abrasives Products Division, emphasized.

Supplement — [5] 両社の提携による研磨システムは、12月3日から売り出す。
Sales of this polishing system, which is a partnership between the two companies, will begin on December 3.

Background — [6] 愛知産業の研磨ロボット「AKグラインダ」をベースに、両社の技術者が研磨材をあてる角度や研磨にかける時間などを共同で研究。
Based on Aichi Sangyo's polishing robot "AK grinder," engineers from both companies jointly researched the angle to apply the abrasive and required polishing time.

Result — [7] ロボットを精密に制御する装置などを開発し、他社の研磨材を取り付けた場合に比べて所定の研磨水準に至るまでの速さを2倍にできたという。
They developed a device that precisely controls the robot and reaches the specified polishing level twice as fast as the case when abrasives of other companies are attached.

Supplement — [8] スリーエムジャパンは研磨材と併せてAKグラインダを顧客に紹介する。
3M Japan will introduce AK grinders to customers along with abrasives.

Background — [9] 顧客の大半は手作業で研磨しているとみられ、ロボットの採用を働きかけ、同社の研磨材の販売増にもつなげる。
Most of the customers seem to be polishing by hand, and they will encourage the adoption of robots, which will lead to an increase in sales of the company's abrasives.

Background — [10] 研磨工程は熟練技能が必要とされる。
The polishing process requires skill.

Supplement — [11] ただ、技能の継承や人材育成が進んでおらず、自動化が課題だった。
However, the succession of skills and human resource development have not progressed, and automation was an issue.

Cause — [12] ロボットの採用も各社が検討してきたが、装置と研磨材の調整などが難しく、実用化が進んでいなかった。
Although each company considered the adoption of robots, it has been difficult to adjust the equipment and abrasives, and practical applications have not progressed.

Topic Change — [15] スリーエムジャパンの17年度の売上高は約3100億円。
Sales of 3M Japan in FY17 were approximately 310 billion yen.

Supplement — [16] テープや接着剤で知られるが、1902年の創業当時の主力事業は研磨材。
Known for tapes and adhesives, its main business at the time of its founding in 1902 was abrasives.

Supplement — [17] 現在の研磨材事業の比率は非公開だが、研磨材市場で同社はトップクラスのシェアを握るとされる。
The current ratio of the abrasives business is not disclosed, but the company is said to have a top-class share in the abrasives market.

Supplement — [18] 日西事業部長は「ロボット化という付加価値を売りにシェア拡大を目指したい」と話す。
"We want to increase our market share by selling the added value of robotization," said Hinishi.

Dependency, Discourse relation, and Chunk（hard chunk or soft chunk）

Figure 5: Annotation example of the discourse structure. [∗] indicates the sentence position in the original text. Dependency annotation, discourse relation annotation, and chunk annotation criteria are described in 3.1.1, 3.1.2, and 3.1.3, respectively.

Definition of the discourse relations

Start means that the parent of the sentence is the root node. *Result* means that the child sentence is the result of the parent sentence. *Cause* means that the child sentence is the cause of the parent sentence and includes event origin, the basis of the author's claim, and the reason. *Background* means that the parent sentence states facts or events, and the child sentence provides the background or premise. *Correspondence* means that the child sentence answers the question in the parent sentence. It also includes countermeasures or examinations of problems or issues. *Contrast* means that the parent sentence and the child sentence have a contrasting relationship. *Topic Change* means that the topic of the parent sentence changes in the child sentence, which includes a subtopic level change. *Example* means that the child sentence provides an instance to support the statement in the parent sentence. *Conclusion* means that the child sentence is a summary or conclusion of the story up to the parent sentence. *Supplement* means that the child sentence provides details or supplementary information about what is stated in the parent sentence. In a broad sense, although the above discourse relations are also included in supplement, herein, *Supplement* covers the inter-sentence relations that are not included in the aforementioned discourse relations.

Quality of the discourse annotations

A one-month training period was established, and discussions were held until the annotation criteria of the two annotators matched. To validate the inter-rater reliability, the two annotators annotated the same 34 articles after the training period. The Cohen's kappa of dependencies, discourse relations, and chunks were 0.960, 0.943, and 0.895, respectively. To calculate kappa of the discourse relations, the comparison was limited to the inter-sentence dependencies in which the parent sentence matched. To calculate kappa of the chunks, we set the label of the sentence selected as the hard chunk, soft chunk, and other to "1, 2, and 0," respectively. Then we compared the labels between sentences. Given the high inter-rater reliability, we concluded that the two annotators could cover different assignments separately.

Table 4: Example answers of the interest questionnaire. Interest level in the topic is assigned to the title [0]. (1, not interested at all; 2, not interested; 3, not interested if anything; 4, interested if anything; 5, interested; 6, very interested)

	U_1	U_2	⋯	U_{12}	U_{13}
[0] 研磨工程ロボ化、3Mが愛知産業と提携 Polishing process robotization, 3M partners with Aichi Sangyo	5	4		3	3
[1] 米3Mの日本法人スリーエムジャパンは29日、ロボットを活用した研磨市場に参入すると発表した。 3M Japan., a Japanese corporation of 3M in the US, announced on the 29th that it will enter the polishing market using robots.	5	4		3	4
[2] 溶接関連機器販売の愛知産業と提携し、同社の研削ロボに3Mの研磨材を取り付けて売り込む。 In partnership with Aichi Sangyo, which sells welding-related equipment, it will sell a 3M abrasive attached to the company's grinding robot.	3	3		3	1
[3] 自動車や鉄道車両、建築資材など、人手による作業が多い研磨工程の自動化を推進。 3M Japan promotes automation of the polishing process, which involves manual work such as automobiles, railroad cars, and building materials.	6	4		3	4
[4] 生産効率の向上や人手不足の補完に役立つシステムとして売り込む。 It will be marketed as a system to improve production efficiency and supplement labor shortages.	6	4		5	4
[5] 両社の提携による研磨システムは、12月3日から売り出す。 Sales of this polishing system, which is a partnership between the two companies, will begin on December 3.	3	4		3	1
[6] 愛知産業の研磨ロボット「AKグラインダ」をベースに、両社の技術者が研磨材をあてる角度や研磨にかける時間などを共同で研究。 Based on Aichi Sangyo's polishing robot "AK grinder", engineers from both companies jointly researched the angle to apply the abrasive and required polishing time.	6	4		3	1
[7] ロボットを精密に制御する装置などを開発し、他社の研磨材を取り付けた場合に比べて所定の研磨水準に至るまでの速さを2倍にできたという。 They developed a device that precisely controls the robot and reaches the specified polishing level twice as fast as the case when abrasives of other companies are attached.	5	4		3	1
[8] スリーエムジャパンは研磨材と併せてAKグラインダを顧客に紹介する。 3M Japan will introduce AK grinders to customers along with abrasives.	3	4		2	1
[9] 顧客の大半は手作業で研磨しているとみられ、ロボットの採用を働きかけ、同社の研磨材の販売増にもつなげる。 Most of the customers seem to be polishing by hand, and they will encourage the adoption of robots, which will lead to an increase in sales of the company's abrasives.	4	4		4	1
[10] 研磨工程は熟練技能が必要とされる。 The polishing process requires skill.	5	4		4	1
[11] ただ、技能の継承や人材育成が進んでおらず、自動化が課題だった。 However, the succession of skills and human resource development have not progressed, and automation was an issue.	5	4		3	1
[12] ロボットの採用も各社が検討してきたが、装置と研磨材の調整などが難しく、実用化が進んでいなかった。 Although each company considered the adoption of robots, it has been difficult to adjust the equipment and abrasives, and practical applications have not progressed.	5	3		3	1
[13] 「人手に頼らない研磨工程を実現しないとメーカーの操業は危機に陥る」。 "If we do not realize a polishing process that does not rely on human labor, the manufacturing operation will be in jeopardy."	5	4		4	1
[14] 29日、愛知産業の相模原事業所で開いた説明会で、スリーエムジャパン研磨材製品事業部の日西勝事業部長は強調した。 At a briefing held at the Sagamihara Plant of Aichi Sangyo on the 29th, Mr. Masaru Hinishi, General Manager of 3M Japan Abrasives Products Division, emphasized.	3	3		3	1
[15] スリーエムジャパンの17年度の売上高は約3100億円。 Sales of 3M Japan in FY17 were approximately 310 billion yen.	3	4		3	1
[16] テープや接着剤で知られるが、1902年の創業当時の主力事業は研磨材。 Known for tapes and adhesives, its main business at the time of its founding in 1902 was abrasives.	4	4		3	1
[17] 現在の研磨材事業の比率は非公開だが、研磨材市場で同社はトップクラスのシェアを握るとされる。 The current ratio of the abrasives business is not disclosed, but the company is said to have a top-class share in the abrasives market.	4	4		3	1
[18] 日西事業部長は「ロボット化という付加価値を売りにシェア拡大を目指したい」と話す。 "We want to increase our market share by selling the added value of robotization," said Hinishi.	4	3		3	1

Table 5: Examples of user profiles

|		U_1	U_{13}
Gender		Male	Female
Age		Forties	Twenties
Address		Kanagawa Prefecture	Mie Prefecture
Occupation type		Administrative worker	Professional / technical worker
Industry type		Manufacturing industry	Medical or welfare
Frequency of checking news		Daily	Daily
Typical time of day news is checked		Morning, night	Morning, evening
Methods to access the news		Video, text	Video
Tools used to check the news		TV, newspaper, smartphone / tablet	TV, smartphone / tablet
Newspapers, websites, and applications used to check the news		Yomiuri Shimbun, NHK, LINE NEWS, SmartNews, Yahoo! News	NHK, LINE NEWS
Whether a fee was paid to check the news		Free	Free
News genre actively checked		Politics, economy, international, society, sports, entertainment	Incident / accident, science / medical, sports, entertainment
Degree of interest in each news genre	Politics	Very interested	Not interested if anything
	Economy	Very interested	Not interested if anything
	International	Very interested	Not interested if anything
	Society	Interested if anything	Not interested if anything
	Local	Interested if anything	Not interested if anything
	Incident / accident	Interested if anything	Interested
	Science / medical	Not interested if anything	Interested
	Sports	Very interested	Interested
	Entertainment	Interested if anything	Interested
Hobbies		Watching sports, watching movies, watching dramas, listening to music, board games, video games / computer games / smartphone games, gambling, driving / touring	Sports / exercise, watching sports, watching movies, watching dramas, watching anime, eating out / visiting cafes, fashion, shopping, taking photos or videos, SNS, traveling

Statistics of the dataset
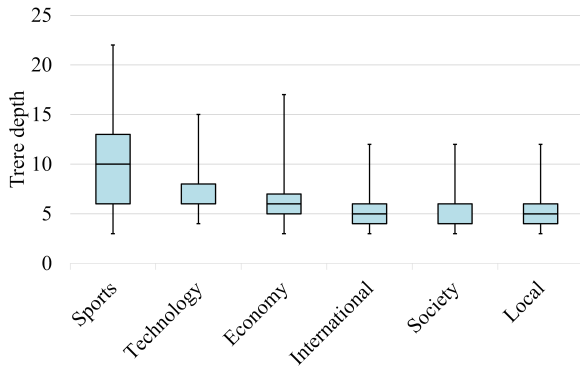
Figure 6: Tree depth distribution, which is the maximum number of sentences from the root node to the leaf nodes for each article. Average tree depth per article is 6.5 sentences.
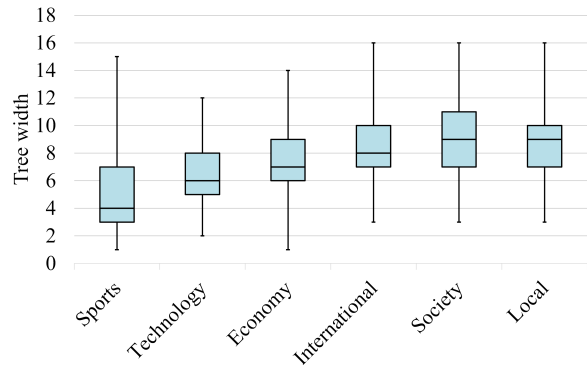

Figure 7: Tree width distribution, which is the number of leaf nodes for each article. Average tree width per article is 7.5 sentences.
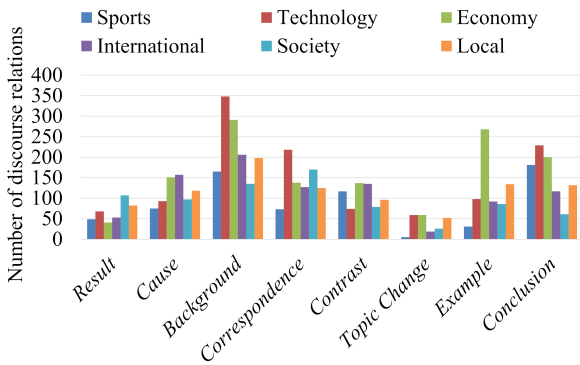

Figure 8: Frequencies of discourse relations in the dataset: *Start* = 1,221; *Result* = 400; *Cause* = 691; *Background* = 1,343; *Correspondence* = 851; *Contrast* = 638; *Topic Change* = 220; *Example* = 709; *Conclusion* = 920; and *Supplement* = 14,609.
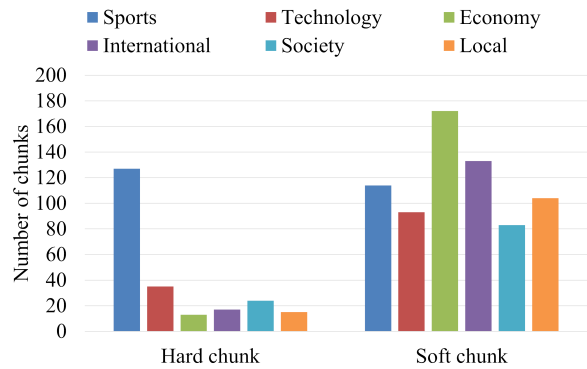

Figure 9: Frequencies of chunks in the dataset. There are 231 hard chunks and 699 soft chunks. Average number of sentences per chunk is 2.15.

(a) Percentage of interest level in the topic

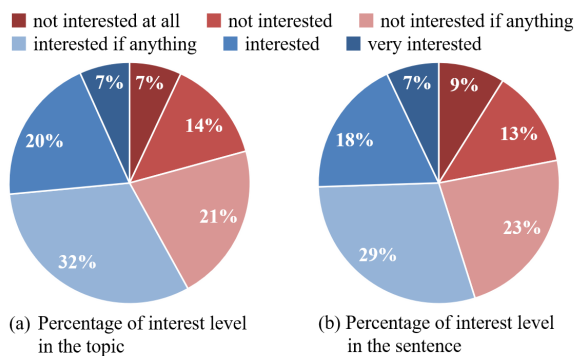(b) Percentage of interest level in the sentence

Figure 10: Percentage of each interest level calculated from the questionnaire results, which includes 15,042 articles and 268,509 sentences.
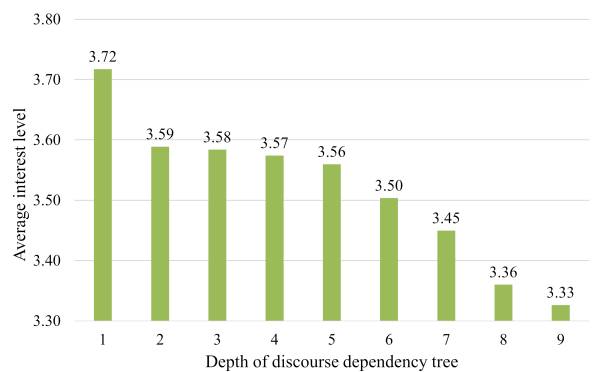

Figure 11: Average interest level of the sentences for each depth of the discourse dependency tree.