

デジタルアニーラを用いたパーソナライズド抽出型要約の高速求解

高津 弘明¹ 柏川 貴弘² 木村 浩一² 安藤 涼太³ 松山 洋一¹
¹ 早稲田大学 ² 富士通研究所 ³ 内外切抜通信社
 takatsu@pcl.cs.waseda.ac.jp {kashikawa, k.kimura}@jp.fujitsu.com
 ando@naigaipc.co.jp matsuyama@pcl.cs.waseda.ac.jp

概要

ニュース記事に代表されるまとまった量の情報を効率的に伝える音声対話システムを開発している。伝達対象の文書の要点を説明するための対話のシナリオを主計画と呼ぶ。主計画生成問題を、文書の談話構造と合計発話時間を制約として興味度が最大となる文集合を文書集合の中から抽出する組合せ最適化問題として定式化した。この組合せ最適化問題の求解速度を高速化するため、イジングマシン的一种である富士通デジタルアニーラ¹⁾を使用した。ニュース記事と芸術作品の解説テキストに対して談話構造と興味度を付与したデータセットを用いた実験で、デジタルアニーラにより制約違反のない準最適解が実用的な時間で得られることを確認した。

1 はじめに

人々のメディアへの接触時間は年々増加し、「世の中の情報量は多すぎる」「世の中の情報のスピードは速すぎる」といった意識が人々の間で高まっている [1]。このような社会背景を踏まえ、我々は、人々が興味のある情報を、限られた時間で、音声対話により効率的に消費することを支援する技術の開発を進めている。

我々の音声対話システムは、要点説明のための主計画と補足説明のための副計画から構成されるシナリオに基づいて対話を進行させる [2]。ユーザが受け身で聴いているとき、システムは主計画の内容を伝える。主計画を生成する問題を、一貫性と効率性の観点から、話題が異なる N 個の文書から各文書の談話構造を制約としてユーザが興味のある文書を抽出し、その内容を音声で T 秒以内に伝える問題として定式化する。この問題は整数線形計画問題として定式化できるが [3]、NP 困難であり、問題の規

1) <https://www.fujitsu.com/jp/digitalannealer/>

模が大きくなると、最適解を見つけるのに膨大な時間が必要になる。そこで、本研究では、組合せ最適化問題を解くことに特化した非ノイマン型のコンピュータであるイジングマシンを用いることで高速化を試みる。イジングマシンの中でも、量子現象に着想を得たデジタル回路により、組合せ最適化問題を高速に解くアーキテクチャである富士通デジタルアニーラを使用し、ユーザごとにパーソナライズした主計画を高速に生成できることを示す。

2 デジタルアニーラ

量子コンピューティング技術は大きく量子ゲート方式とイジングマシン方式に分類される。量子ゲート方式は、0 と 1 の重ね合わせ状態が取れる量子ビットを操作する量子ゲートを用いて汎用的な計算を行う方式である [4, 5]。イジングマシン方式は、+1 と -1 の 2 値を取るスピンとスピン間の相互作用でエネルギーが規定されるイジングモデルに問題をマッピングして解を探索する組合せ最適化問題に特化した方式である。イジングマシン方式には、超電導回路で量子ビットを操作する量子アニーリング方式 [6, 7, 8] と、デジタル回路で古典ビットを操作するシミュレーテッドアニーリング方式 [9, 10, 11, 12, 13] がある。デジタルアニーラは、シミュレーテッドアニーリング方式のイジングマシンであり、量子現象に着想を得たデジタル回路 (Digital Annealing Unit: DAU) により組合せ最適化問題を高速に解くアーキテクチャである [11, 12, 13]。

2.1 第二世代デジタルアニーラ

第二世代デジタルアニーラ (DA2) は、シミュレーテッドアニーリングに基づく手法であるアニーリングモード (DA2-AM) [11] またはマルコフ連鎖モンテカルロ法 (Markov-Chain Monte Carlo: MCMC) に基づく手法であるパラレルテンパリングモー

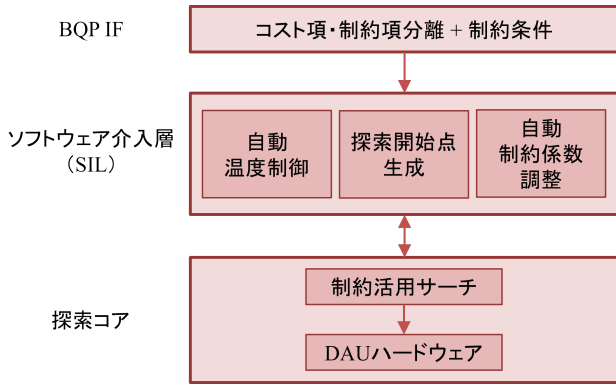


図1 第三世代デジタルアニーラの構成 [13]

ド (DA2-PTM) [12] を使用して, QUBO (Quadratic Unconstrained Binary Optimization) 形式で定式化された組合せ最適化問題の解を高速に探索する. QUBO モデルは, イジングモデルにおけるスピンの代わりに 0 と 1 のバイナリ変数を用いたモデルである. DA2 は問題のビット数に応じて, ビット数パラメータを 1024, 2048, 4096, 8192 の中から指定する必要がある. DA2 は 4096 ビット以下の問題を 64 ビット精度で, 8192 ビット以下の問題を 16 ビット精度で扱うことができる.

2.2 第三世代デジタルアニーラ

第三世代デジタルアニーラ (DA3) は, 最大 10 万ビットのバイナリ二次計画 (Binary Quadratic Programming: BQP) 問題を入力して, ソフトウェア介入層 (Software Intervention Layer: SIL) と探索コアが協調して最適解または近似解を求めるソフトウェアとハードウェアのハイブリッド求解システムである (図 1) [13]. 制約項をエネルギー関数に含めた QUBO 形式による求解は, 制約条件の数や変数の規模が増大すると最適解や近似解への到達が難しくなる [14]. その問題を解決するため, DA3 は, エネルギー関数をコスト項と制約項に分離し, 制約項の重みを自動的に調整する仕組みを持つ. さらに, DA3 は, バイナリ変数列の総和が 1 となる等式制約 (1hot 制約) および線形不等式制約を直接設定するインタフェースを持つ. 探索コアでは, 制約活用サーチが, ソフトウェア介入層で生成された探索開始点を起点に, 制約項と自動調整された制約係数による制約違反の影響度および, 1hot 制約や不等式制約などの条件に従って大域的に解を探索し, 良解を抽出する. そして, DAU ハードウェアが良解周辺の最適解を MCMC で高速に探索する.

表 1 変数の定義

x_{ki}	文 s_{ki} が選択されたか
y_{kij}	文 s_{ki} と文 s_{kj} が選択されたか
b_{ki}^u	文 s_{ki} に対するユーザ u の興味度
r_{kij}	文 s_{ki} と文 s_{kj} の類似度
t_{ki}	文 s_{ki} の発話時間 (秒)
T	合計発話時間の上限値 (秒)
L_{kl}	文書 d_k と文書 d_l の抽出文数の差の絶対値の上限
$f_k(i)$	文 s_{ki} の係り先の文 ID を返す関数
D_N^u	ユーザ u に対して選択された N 個の文書 ID の集合
S_k	文書 d_k に含まれる文 ID の集合
C_{km}	文書 d_k の m 番目のチャンクに含まれる文 ID の集合

3 手法

3.1 ILP モデル

N 個の文書からユーザ u が興味のあるような文を抽出し, その内容を音声で T 秒以内に伝える要約問題を考える. 主計画の要件は, ユーザにとって興味がある内容であること, ストーリーが一貫していること, 冗長でないことである. そこで, 要約問題を, 談話構造と合計発話時間を制約とし, 文に対する興味度の高さや文間の類似度の低さのバランスで目的関数を定義した整数線形計画 (Integer Linear Programming: ILP) 問題として定式化する.

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

s.t.

$$\forall k, i, j: x_{ki} \in \{0, 1\}, y_{kij} \in \{0, 1\} \\ \sum_{k \in D_N^u} \sum_{i \in S_k} t_{ki} x_{ki} \leq T \quad (2)$$

$$\forall k \neq l: \sum_{i \in S_k} x_{ki} - \sum_{i \in S_l} x_{li} \leq L_{kl} \quad (3)$$

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

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

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

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

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

変数の定義を表 1 に示す. k 番目の文書 d_k の i 番目の文を s_{ki} と表記する. r_{kij} は, 文 s_{ki} と文 s_{kj} の内容語の Bag-of-Words のコサイン類似度である.

式 2 は要約の合計発話時間を T 秒以下にする制約である. 式 3 は抽出する文数の文書間での偏りを L_{kl} 文以下に抑える制約である. 式 4 は文 s_{ki} を抽

出する場合、談話依存構造木における親ノードの文 s_{kj} も抽出しなければならないとする制約である。式 5 はあるチャンクの文 s_{ki} を抽出する場合、同チャンクに含まれる他の文も抽出しなければならないとする制約である。式 6 から式 8 は文 s_{ki} と文 s_{kj} が選ばれた際に $y_{kij} = 1$ とするための制約である。

L_{kl} は、合計発話時間の上限値 T と文書数 N 、文の平均発話時間 \bar{t} に基づいて以下の式で計算する。

$$L_{kl} = L + \lceil \max(\max(\bar{n} - |S_k|, 0), \max(\bar{n} - |S_l|, 0)) \rceil \quad (9)$$

$$L = \left\lceil \frac{\bar{n}}{\sqrt{N}} + 0.5 \right\rceil \quad (10)$$

$$\bar{n} = \frac{T}{\bar{t} \times N} \quad (11)$$

\bar{n} は一つの文書から抽出が期待される文の数を表し、 L はその値を文書数の平方根で割り、四捨五入した値である。

3.2 QUBO モデル

DA2 で組合せ最適化問題を解くために、上記 ILP モデルに基づき、QUBO モデルのエネルギー関数 (ハミルトニアン) H を以下のように定義する。

$$H = H_0 + \lambda_1 H_1 + \lambda_2 H_2 + \lambda_3 H_3 + \lambda_4 H_4 \quad (12)$$

$$H_0 = - \sum_{k \in D_N^u} \sum_{i < j \in S_k} b_{ki}^u b_{kj}^u (1 - r_{kij}) x_{ki} x_{kj} \quad (13)$$

$$H_1 = \left(T - \sum_{k \in D_N^u} \sum_{i \in S_k} t_{ki} x_{ki} - \sum_{n=0}^{\lfloor \log_2(T-1) \rfloor} 2^n y_n \right)^2 \quad (14)$$

$$H_2 = \sum_{k \neq l \in D_N^u} \left(L_{kl} - \left(\sum_{i \in S_k} x_{ki} - \sum_{i \in S_l} x_{li} \right) - \sum_{n=0}^{\lfloor \log_2(L_{kl}-1) \rfloor} 2^n z_n \right)^2 \quad (15)$$

$$H_3 = (x_{ki} - x_{ki} x_{kj=f_k(i)})^2 \quad (16)$$

$$H_4 = \sum_{k \in D_N^u} \sum_{m \in C_k} \sum_{i \in C_{km}} \left(\sum_{j \in C_{km}} x_{kj} - |C_{km}| \times x_{ki} \right)^2 \quad (17)$$

ここで、 y_n と z_n は不等式制約を等式制約に変換するために導入したスラック変数である。 $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ は各制約に対する重みである。

QUBO モデルはシミュレーテッドアニーリングに基づく手法 [11] やパラレルテンパリング [12] によって解くことができる。これらの手法では異なる初期値で並列にアニーリングすることで複数の解候補を得ることができる。しかしながら、これらの手法では制約違反が発生しないことを保証していない。そ

のため、後処理で、得られた解候補の中から制約違反スコアが最小なもののうち、目的関数のスコアが最大な解を採用する。制約違反スコア E_{total} は以下の式で計算する。

$$E_{total} = E_{dep} + E_{chunk} + E_{bias} + E_{time} \quad (18)$$

$$E_{bias} = \sum_{k < l \in D_N^u} \max(\hat{L}_{kl} - L_{kl}, 0) \quad (19)$$

$$E_{time} = \max(\hat{T} - T, 0) \quad (20)$$

ここで、 E_{dep} は係り受け関係の違反件数、 E_{chunk} はチャンクの違反件数、 \hat{T} は抽出された文集合の合計発話時間、 \hat{L}_{kl} は文書 d_k と文書 d_l で抽出された文数の差の絶対値を表す。

3.3 DA3 モデル

DA3 は線形不等式制約を直接設定するインターフェースを備えているため、不等式制約を QUBO 化する必要がない。そのため、エネルギー関数はコスト項のみ (式 13) で定義する。制約に関しては式 5 を 0 以上 0 以下の不等式に変換し、他の不等式制約 (式 2, 式 3, 式 4) とともに、DA3 のインターフェースに設定する。ソフトウェア介入層が不等式間の重み付けを自動的に調整し、制約活用サーチがこれらの不等式制約を評価しながら解の探索を行う。

4 実験

4.1 実験設定

データセットには談話構造と興味度が付与されたニュース記事データセット [15] とミュージアムデータセット [16] を使用した。ニュース記事データセットは 1200 個のニュース記事の話題と文 (一記事あたり 15 文~25 文) に 6 段階の興味度が付与された 2017 人分 (一人あたり 6 記事) のデータセットであり、一人あたりの問題の規模は 4096 ビット以下である。ミュージアムデータセットは e 国宝で公開されている東京国立博物館の 400 個の国宝・重要文化財とその解説文 (一作品あたり 4 文~14 文) に 6 段階の興味度が付与された 2485 人分 (一人あたり 8 作品) のデータセットであり、一人あたりの問題の規模は 2048 ビット以下である。ニュース記事データセットを用いた実験では $N = 6$ 個の記事を $T = 450$ 秒以内に説明する設定で、ミュージアムデータセットを用いた実験では $N = 8$ 個の作品を $T = 330$ 秒以内に説明する設定で、要約性能を評価した。

文の発話継続長を計算するために音声合成器

表2 情報伝達効率

	イテレーション回数	ミュージアムデータセット (2048 ビット問題)					ニュース記事データセット (4096 ビット問題)				
		処理時間 (秒)	被覆率	除外率	EoIT ₁	EoIT ₂	処理時間 (秒)	被覆率	除外率	EoIT ₁	EoIT ₂
CPU-CBC	-	8.57	0.601	0.550	0.574	0.559	102	0.616	0.637	0.626	0.633
DA2-AM	10 ³	0.0612	0.558	0.516	0.536	0.524	0.199	0.524	0.570	0.546	0.560
DA2-PTM	10 ³	0.249	0.558	0.513	0.535	0.521	0.749	0.537	0.561	0.549	0.556
	10 ⁴	1.73	0.569	0.521	0.544	0.530	6.44	0.550	0.574	0.562	0.569
DA3	-	2.07	0.601	0.545	0.572	0.555	2.10	0.613	0.630	0.621	0.627

として、ニュース記事データセットでは AITalk²⁾ を、ミュージアムデータセットでは Google Cloud Text-to-Speech³⁾ を使用した。文間の間 (ま) を 1 秒に設定し、合成音声ファイルの再生時間にこの間 (ま) の時間を加算した値を t_{ki} とした。

ILP モデルは CPU 上で分枝切除法⁴⁾ [17, 18] (CPU-CBC) により解いた。具体的には PuLP⁵⁾ の PULP_CBC_CMD ソルバーを使用し、スレッド数は 30 に設定した。処理時間は以下の構成の Google Compute Engine⁶⁾ 上で計測した。OS:Ubuntu 18.04, CPU:Xeon (2.20 GHz, 32 コア), メモリ:64 GB。

DA2-PTM のレプリカ数および DA2-AM のアニーリングの繰り返し回数は 128 に設定した。DA2 の性能は主に λ の値と一回のアニーリングにおける探索回数 (イテレーション回数) によって決まるため、制約違反が発生しないようにこれらの値を調整した。最終的に、DA2-AM と DA2-PTM の $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ は、ニュース記事データセットではそれぞれ $\{10^2, 10^6, 10^{10}, 10\}$ と $\{10^2, 10^5, 10^9, 10\}$ に、ミュージアムデータセットではそれぞれ $\{10^2, 10^7, 10^{10}, 10\}$ と $\{10^2, 10^6, 10^9, 10\}$ に設定した。

CPU-CBC をベンチマークとして、DA2-AM と DA2-PTM および DA3 の性能を評価した。評価指標には EoIT _{β} (Efficiency of Information Transmission) [19] を使用した。興味度が 4 以上 (興味あり) の文の被覆率を C 、興味度が 3 以下 (興味なし) の文の除外率を E としたとき、EoIT _{β} は、重み付き F 値 [20] に基づき、以下の式で定義される。

$$\text{EoIT}_{\beta} = \frac{(1 + \beta^2) \times C \times E}{\beta^2 \times C + E} \quad (21)$$

4.2 実験結果

実験結果を表 2 に示す。各評価指標と処理時間の値は平均値である。まず、CPU-CBC と DA2-AM

- 2) <https://www.ai-j.jp/product/voiceplus/manual/>
- 3) <https://cloud.google.com/text-to-speech/?hl=jp>
- 4) <https://projects.coin-or.org/Cbc>
- 5) <https://coin-or.github.io/pulp/>
- 6) <https://cloud.google.com/compute/?hl=jp>

を比較すると、EoIT に関しては、DA2-AM は CPU-CBC よりも低かったものの、処理時間に関しては、DA2-AM の方が圧倒的に早いことが確認された。次に、DA2-AM と DA2-PTM を比較すると、イテレーション回数が千回のときは、EoIT に差はほとんど見られなかったが、処理時間は DA2-AM の方が早いことが分かった。一方で、DA2-AM はイテレーション回数を増やしても性能に変化が見られなかったが、DA2-PTM はイテレーション回数を増やすにしたがって EoIT が向上した。しかしながら、それに伴い処理時間も増加した。次に、DA3 の結果に着目すると、EoIT は CPU-CBC に近い値を示した。DA2 では、問題の規模が 2048 ビットから 4096 ビットに上がると、処理時間は増加し、EoIT は CPU-CBC との差が大きくなる傾向にあったが、DA3 では、問題の規模が大きくなって同じ処理時間で、高い EoIT を示した。一方で、処理時間に関しては、DA2-AM が最も早い値を示した。

15 文から 25 文のニュース記事 6 個を要約する問題では、4096 ビット必要であり、DA2-AM で処理すると 1 人あたり約 0.2 秒かかるため、10 万人の主計画を用意するのに、一つの DAU でおよそ 5 時間半かかる計算になる。以上のことから、この程度の規模の問題設定であれば、デジタルアニーラを用いることで、前の晩のニュースから各ユーザにパーソナライズした対話のシナリオを、翌日の通勤・通学時までに用意することが可能だと言える。

5 おわりに

音声対話システムにおける要点説明のための対話のシナリオである主計画を生成する問題を、文書の談話構造と合計発話時間を制約として興味度が最大となる文集合を文書集合の中から抽出する組合せ最適化問題として定式化した。シミュレーテッドアニーリングベースのイジングマシンである富士通デジタルアニーラを用いることで、実用的な時間で制約違反のない準最適解が得られることを確認した。

参考文献

- [1] 博報堂DYメディアパートナーズ(編). 広告ビジネスに関わる人のメディアガイド2020: メディア環境のこれからと今. 宣伝会議, pp. 20–40, 2020.
- [2] 高津弘明, 福岡維新, 藤江真也, 林良彦, 小林哲則. 意図性の異なる多様な情報行動を可能とする音声対話システム. 人工知能学会論文誌, Vol. 33, No. 1, pp. 1–24, 2018.
- [3] Hiroaki Takatsu, Ryota Ando, Hiroshi Honda, Yoichi Matsuyama, and Tetsunori Kobayashi. Personalized extractive summarization with discourse structure constraints towards efficient and coherent dialog-based news delivery. In **Proceedings of the 12th International Workshop on Spoken Dialog System Technology**, 2021.
- [4] 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 Viallonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. **Nature**, Vol. 574, No. 7779, pp. 505–510, 2019.
- [5] Laszlo Gyongyosi. Unsupervised quantum gate control for gate-model quantum computers. **Scientific Reports**, Vol. 10, No. 1, pp. 1–16, 2020.
- [6] 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. Quantum annealing with manufactured spins. **Nature**, Vol. 473, No. 7346, pp. 194–198, 2011.
- [7] 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. Architectural considerations in the design of a superconducting quantum annealing processor. **IEEE Transactions on Applied Superconductivity**, Vol. 24, No. 4, pp. 1–10, 2014.
- [8] Masaaki Maezawa, Go Fujii, Mutsuo Hidaka, Kentaro Imafuku, Katsuya Kikuchi, Hanpei Koike, Kazumasa Makise, Shuichi Nagasawa, Hiroshi Nakagawa, Masahiro Ukibe, and Shiro Kawabata. Toward practical-scale quantum annealing machine for prime factoring. **Journal of the Physical Society of Japan**, Vol. 88, No. 6, 2019.
- [9] Masanao Yamaoka, Chihiro Yoshimura, Masato Hayashi, Takuya Okuyama, Hidetaka Aoki, and Hiroyuki Mizuno. A 20k-spin Ising chip to solve combinatorial optimization problems with CMOS annealing. **IEEE Journal of Solid-State Circuits**, Vol. 51, No. 1, pp. 303–309, 2016.
- [10] Takuya Okuyama, Masato Hayashi, and Masanao Yamaoka. 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**, pp. 1–6, 2017.
- [11] Maliheh Aramon, Gili Rosenberg, Elisabetta Valiante, Toshiyuki Miyazawa, Hirotaka Tamura, and Helmut G. Katzgraber. Physics-inspired optimization for quadratic unconstrained problems using a Digital Annealer. **Frontiers in Physics**, Vol. 7, No. 48, pp. 1–14, 2019.
- [12] Satoshi Matsubara, Motomu Takatsu, Toshiyuki Miyazawa, Takayuki Shibasaki, Yasuhiro Watanabe, Kazuya Takemoto, and Hirotaka Tamura. 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**, pp. 667–672, 2020.
- [13] 中山寛, 小山純平, 米岡昇, 宮澤俊之. 解説: 第三世代デジタルアニーラの技術, pp. 1–9, 2021.
- [14] Fred Glover, Gary Kochenberger, and Yu Du. Quantum bridge analytics I: A tutorial on formulating and using QUBO models. **4OR: A Quarterly Journal of Operations Research**, Vol. 17, No. 4, pp. 335–371, 2019.
- [15] Hiroaki Takatsu, Takahiro Kashikawa, Koichi Kimura, Ryota Ando, and Yoichi Matsuyama. Personalized extractive summarization using an ising machine towards real-time generation of efficient and coherent dialogue scenarios. In **Proceedings of the 3rd Workshop on Natural Language Processing for Conversational AI**, pp. 16–29, 2021.
- [16] 高津弘明, 安藤涼太, 中野鐵兵, 柏川貴弘, 木村浩一, 松山洋一. イジングマシンを用いたミュージアムガイドの発話計画最適化. 第12回対話システムシンポジウム, pp. 26–31, 2021.
- [17] John E. Mitchell. Branch-and-cut algorithms for combinatorial optimization problems. **Handbook of Applied Optimization**, pp. 65–77, 2002.
- [18] Manfred Padberg and Giovanni Rinaldi. A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems. **SIAM Review**, Vol. 33, No. 1, pp. 60–100, 1991.
- [19] Hiroaki Takatsu, Mayu Okuda, Yoichi Matsuyama, Hiroshi Honda, Shinya Fujie, and Tetsunori Kobayashi. Personalized extractive summarization for a news dialogue system. In **Proceedings of the 8th IEEE Spoken Language Technology Workshop**, pp. 1044–1051, 2021.
- [20] Nancy Chinchor. MUC-4 evaluation metrics. In **Proceedings of the 4th conference on Message understanding**, pp. 22–29, 1992.

付録

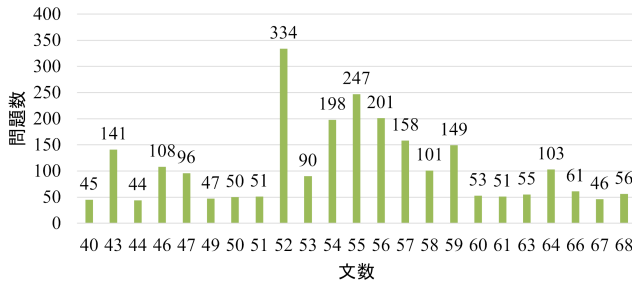


図2 文数ごとの問題数 (ミュージアムデータセット)

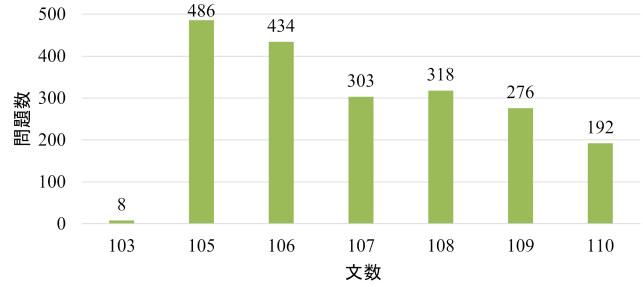
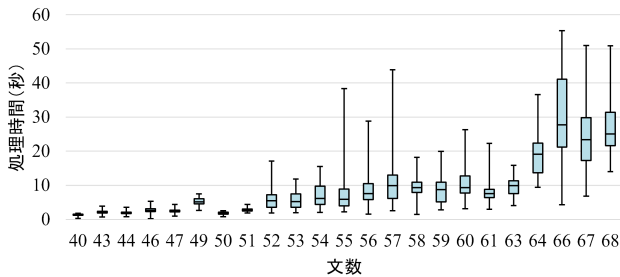
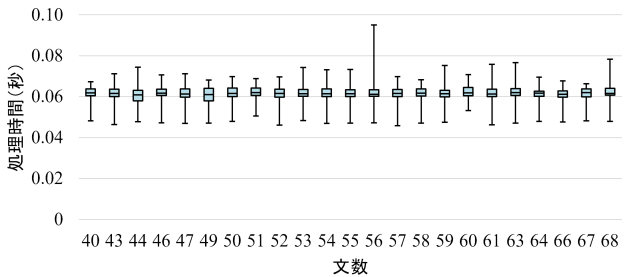


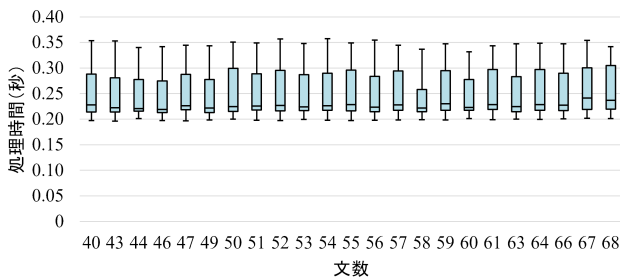
図3 文数ごとの問題数 (ニュース記事データセット)



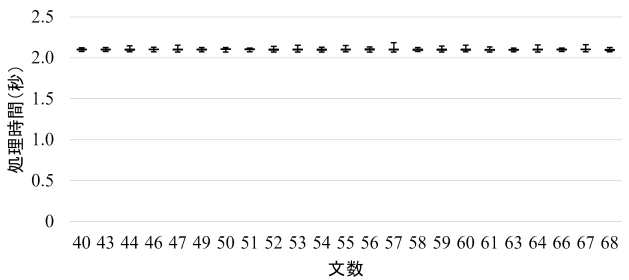
CPU-CBC (スレッド数 = 30)



DA2-AM (イテレーション回数 = 1000)

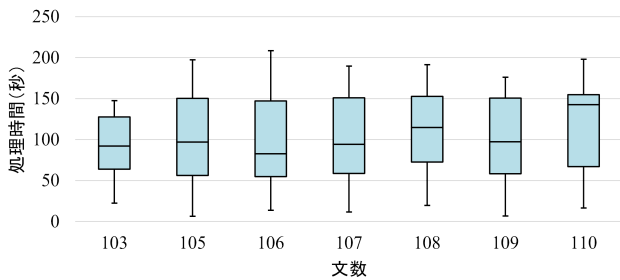


DA2-PTM (イテレーション回数 = 1000)

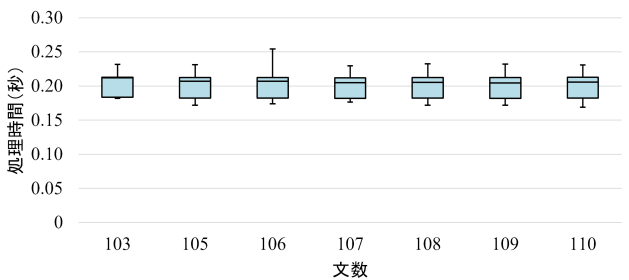


DA3

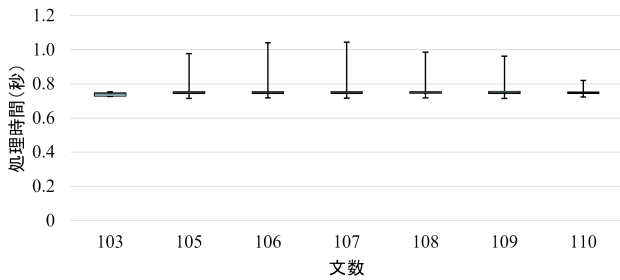
図4 文数ごとの処理時間 (ミュージアムデータセット)



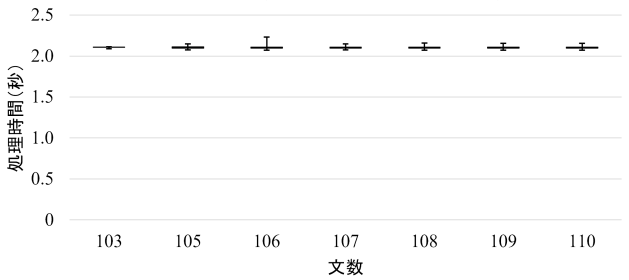
CPU-CBC (スレッド数 = 30)



DA2-AM (イテレーション回数 = 1000)



DA2-PTM (イテレーション回数 = 1000)



DA3

図5 文数ごとの処理時間 (ニュース記事データセット)