

T5 を用いた一括質問・答弁形式の議事録からの答弁の要約

櫻 惇志[†] 太刀岡勇氣[†]

[†] 株式会社デンソーアイティラボラトリ

E-mail: †{keyaki.atsushi,tachioka.yuki}@core.d-itlab.co.jp

あらまし 本研究では、一括質問・答弁形式の会議の議事録において、質問の要約が与えられた場合に、その質問と対応する答弁の要約を生成するタスクに取り組む。答弁の要約生成の手順として、1. 与えられた質問要約に対応する質問の特定、2. 質問と対応する答弁のマッチング、答弁の要約の3段階にて行う。手順1, 2の手順については単語の共起・重み付けによって行う。手順3は、転移学習モデルであるT5による要約を行う。その際、大規模日本語コーパスによって事前学習されたモデルに対して、事前に抽出された「答弁-答弁要約ペア」を訓練データとしてファインチューニングを行うことで、一括質問・答弁形式の会議の議事録に特化した要約モデル構築を目指す。評価実験の結果、T5を用いた答弁の要約を行うことで、ナイーブな手法と比較して高精度な結果を示した。

キーワード 一括質問・答弁, 地方議会議事録, 答弁要約, T5

1 はじめに

近年、公共データの活用に期待が高まり、自治体のオープンデータ化が促進されている[1],[2]。国会や地方議会での議論は一般市民にとっても重要な情報が多く含まれており、特に地方議会では地域コミュニティに密接した施策や支援について言及されることも多い。これらの情報を有効に活用することでより豊かな生活を送ることが可能となることを期待される。その一方で、国会や地方議会での議論は一括質問・答弁形式となっており、また、一般市民には馴染みの低い特徴的な表現が頻出する[3]。一括質問・答弁形式とは、まず質問者が複数の質問を行い、その後で回答者が個々の質問に対して答弁を行う形式のことであり、答弁に対応する質問は、数個の質問から多い場合には数十個の質問に及ぶ。つまり、一問一答形式のように質問と答弁が対応付けられておらず、中には質問・答弁全体が非常に長文となる場合もある。更には、その言い回しの特殊性により、非専門家である一般市民が会議録の生データから有用な情報を取得するのは難易度が高い。

そこで、自然言語処理技術によって、国会や地方議会の会議録データを一般市民が利活用しやすくなるように支援することは非常に重要である。本研究では、特に、一括質問・答弁形式の会議録から、質問・答弁ともに要約された一問一答形式に変換して提示することに取り組む。その際、本タスクに取り組むうえで、NTCIR-16 QA Lab-PoliInfo-3 タスク[4]¹のQuestion Answering サブタスクのデータセットを利用する。Question Answering サブタスクの目的は、会議録の質問の要約が与えられたときに、会議録から質問に対応する答弁を見つけだし、要約した結果を返すことであり、本研究の目的と合致する。答弁の要約生成の手順として、

- (1) 与えられた質問要約に対応する質問の特定
- (2) 質問と対応する答弁のマッチング

(3) 答弁の要約

の3段階にて行う。手順1, 2の手順については単語の共起・重み付けによって行う。手順3は、転移学習モデルであるT5による要約を行う。その際、大規模日本語コーパスによって事前学習されたモデルに対して、事前に抽出された「答弁-答弁要約ペア」を訓練データとしてファインチューニングを行うことで、一括質問・答弁形式の会議録に特化した要約モデル構築を目指す。

以降、2節にて NTCIR-16 QA Lab-PoliInfo-3 タスク及び Question Answering サブタスクの概要について述べる。その後、3節にて提案手法である質問の要約を入力とした答弁の要約について議論を行い、4節にて評価実験結果について報告を行う。

2 地方議会議事録データ

2.1 NTCIR QA Lab-PoliInfo タスク

本節では NTCIR-16 QA Lab-PoliInfo-3 について述べる。NTCIR-16 QA Lab-PoliInfo-3 は政治課題を解決する上で適切な情報を提示することを目的としたタスクであり、NTCIR-14 QA Lab-PoliInfo タスク[5]、NTCIR-15 QA Lab-PoliInfo-2 タスク[6]に流れを汲む。いずれのタスクも地方議会議事録を対象として有用な情報の抽出・判定・提示を目指す。NTCIR-14 QA Lab-PoliInfo タスクには下記の3種類のサブタスクが含まれる。

Segmentation Task 引用(要約)された文に対応する会議録中の範囲(最初と最後の文)を推定する。

Summarization Task 会議録中の文章を、話者の意図を歪めないように要約する。

Classification Task 政策(トピック)に対して、文を「根拠つきで賛成」、「根拠つきで反対」、「その他」の3つのクラスに割り当てる。

に対して、NTCIR-15 QA Lab-PoliInfo-2 タスクは下記の4種

¹ : <https://poliinfo3.net/>

類のサブタスクが含まれる。

Stance classification task 政治家の発言から、政治家の立場を推定する。

Dialog summarization task 地方議会における対話の構造を考慮しながら、自動要約する。

Entity linking task 政治家の発言に含まれる政治用語を対象として、表記揺れや曖昧性の問題を解決しながら、発言の根拠となる記事や一次情報と結びつける。

NTCIR QA Lab-PoliInfo は上記の 2 ラウンドの取り組みを経て、NTCIR-16 QA Lab-PoliInfo-3 ではより高度かつ有用性の高い下記の 4 種類のサブタスクに取り組んでいる。

Question Answering 会議録の質問の要約が与えられたときに、会議録から質問に対応する答弁を見つけだし、要約した結果を返す。

QA Alignment 一括質問一括答弁形式の質問と答弁が与えられたとき、それぞれの質問と答弁を対応付ける。

Fact Verification 要約と会議録が与えられたとき、要約の内容が会議録中に存在するかを判定し、存在する場合はその範囲を特定する。

Budget Argument Mining 予算情報が与えられたときに、会議録から議論している部分を見つけ、結びつけるとともに、Claim(主張), Premise(根拠) を推定する

前述の通り、本研究グループでは、一般市民が議会議事録の概要を把握するうえでより利便性の高い情報を提示するために、質問・答弁ともに要約された一問一答形式に変換する Question Answering サブタスクに取り組む。

2.2 Question Answering サブタスク

Question Answering サブタスクにて、オーガナイザーより提供されるデータについて説明する。それぞれ、オープンデータからクロールされ、整形したものが JSON フォーマットにて提供される。

質問要約ファイル 都議会だより²に含まれる、議員の質問の要約

会議録ファイル 2001 年以降の東京都議会の議事録³に含まれる議員・答弁者(説明員)の発言

質問要約ファイルには、下記のフィールドが含まれる: ID, 会議名 (Meeting), 日付 (Date), 質問者の発言全体の要旨 (Headlines), サブトピック (SubTopic), 質問者 (QuestionSpeaker), 質問要約 (QuestionSummary), 答弁者 (AnswerSpeaker), 答弁の要約 (AnswerSummary)⁴。図 1 に具体例を示す。また、会議録ファイルは下記のフィールドを含む: 年・回 (Volume), 号 (Number), 日付 (Date), 会議録の表題 (Title), 発言者の議席番号または役職 (SpeakerPosition), 発言者の氏名 (SpeakerName), 発言に対応する質問者の氏名と役職 (QuestionSpeaker), 発言者の氏名と役職 (Speaker), 発言 (Utterance)。図 2 に具体例

2: 「都議会だより」は、議会の活動をお知らせする広報紙。質問と答弁それぞれの要約を含む。 <https://www.gikai.metro.tokyo.jp/newsletter/>

3: <https://www.gikai.metro.tokyo.jp/record/proceedings/>

4: ただし出題ファイルでは空欄である。

```
[
  {
    "ID": "PoliInfo3-QA-v20211120-244-03-1-001",
    "Meeting": "平成13年第3回定例会",
    "Date": "2001-09-26",
    "Headlines": ["首都機能を一層高め", "東京新時代を築くべき"],
    "SubTopic": "都市再生",
    "QuestionSpeaker": "松本文明(自民党)",
    "QuestionSummary": "首都圏再生へ大集会の開催を。",
    "AnswerSpeaker": "知事",
    "AnswerSummary": "七都県市の連携を基本として検討する。"
  },
  {
    "ID": "PoliInfo3-QA-v20211120-244-03-1-002",
    "Meeting": "平成13年第3回定例会",
    "Date": "2001-09-26",
    "Headlines": ["首都機能を一層高め", "東京新時代を築くべき"],
    "SubTopic": "羽田空港国際化",
    "QuestionSpeaker": "松本文明(自民党)",
    "QuestionSummary": "実現すべき。見解を伺う。",
    "AnswerSpeaker": "知事",
    "AnswerSummary": "深夜などの未使用枠の活用で国際化は可能。"
  },
  ...
]
```

図 1 質問要約ファイルの具体例

```
[
  {
    "Volume": "2011-2",
    "Number": "2",
    "Date": "2011-06-23",
    "Title": "平成二十三年東京都議会議事録第八号",
    "SpeakerPosition": "知事",
    "SpeakerName": "石原慎太郎",
    "QuestionSpeaker": "山下太郎(民主党)",
    "Speaker": "知事",
    "Utterance": "山下太郎議員の代表質問にお答えいたします。"
  },
  {
    "Volume": "2011-2",
    "Number": "2",
    "Date": "2011-06-23",
    "Title": "平成二十三年東京都議会議事録第八号",
    "SpeakerPosition": "知事",
    "SpeakerName": "石原慎太郎",
    "QuestionSpeaker": "山下太郎(民主党)",
    "Speaker": "知事",
    "Utterance": "まず、被災地支援についてですが、震災発生から三カ月たちましたが、依然として被災地には大地震がもたらした惨状が広がっており、被災者の生活再建も遅々として進んでおりません。"
  },
  ...
]
```

図 2 議事録ファイルの具体例

を示す。

3 提案手法

提案手法では、会議録の質問の要約が与えられたときに、会議録から質問に対応する答弁を見つけだし、要約した結果を返す。そのために必要な事前準備である、議事録の発言から個別の質問・答弁への分割と、自動要約器の構築を行う。

3.1 個別の質問・答弁への分割

議事録の発言は、図 2 の通り、議事録の発言は一文が含まれており、同一話者の連続した発言を結合することで質問もしくは答弁を取り出すことが可能である。このままでは質問・答弁は一括形式となっているため、一件の質問もしくは答弁ごとに分割する。その際、質問もしくは答弁の区切りとなる表現(例えば、「まず」「初めに」「次に」「次いで」「最後に」など)を用いて分割を行う。詳細な処理手順については文献 [7] に示す。

質問要約：被災地が真に必要とする支援に継続して取り組むべき。知事の見解は。

会議録（一部抜粋）

1. 対応する質問の特定

山下太郎（民主党）：都は、児童の心のケアや、災害時要援護者の救護など、医療...
山下太郎（民主党）：また、各県、道の復興、もしくは再生を目指し、独自復興計画を...
山下太郎（民主党）：このように被災地が取り組むべき課題は山積し、日々刻々地域ごと...
山下太郎（民主党）：被災地、被災者が真に必要とする支援に今後とも継続して取り組むべきと考えますが、知事の見解を伺います。
山下太郎（民主党）：現在、都内には福島県などから自主避難してきた約五千名の避難者...
山下太郎（民主党）：故郷から遠く、いつ帰れるのかという思いを持って...
：
：
知事：山下太郎議員の代表質問に答えます。

2. 質問と答弁のマッチング

知事：まず、被災地支援についてですが、震災発生から三カ月たちましたが、依然...
知事：電力、農林水産物等の供給の多くを被災地に依存してきた首都東京が、全国の先頭に立ち、刻々被災地を支援していくことは当然であります。
知事：眼前に立ちほだかる瓦れきの撤去や仮設住宅の建設、被災地域の経済再生など...
知事：後手後手に回っておりますが、東京は、今後とも復興に向けみずから踏み出す
知事：被災地、被災者を強力に後押ししていきたいと思っております。
：

3. 答弁の要約

答弁要約：全国の先頭に立ち刻々被災地を支援するのは当然。今後も強力に後押しする。

図 3 質問要約からの答弁要約の生成手順

3.2 自動要約器の構築

文書分類、機械翻訳、自動要約といった多様な自然言語処理タスクにおいて state-of-the-art な性能を示す転移学習モデルである Text-to-Text Transfer Transformer (T5) [8] を用いる。なお、大規模日本語コーパスを用いた事前学習済みのモデル `sonoisa/t5-base-japanese`⁵ に対して、本タスクに特化させるため、会議録の答弁要約モデルにファインチューニングを行う。そのため、タスクオーガナイザーから提供される訓練データから、「答弁-要約答弁」ペアの抽出を行い、上記事前学習済みのモデルに対して追加の学習を行う⁶。

3.3 要約された質問に対応する答弁の要約

答弁の要約は下記の 3 段階にて行う：

- (1) 与えられた質問要約に対応する質問の特定
- (2) 質問と対応する答弁のマッチング
- (3) 答弁の要約

概要図を図 3 に示す。

質問された入力要約は「被災地が真に必要とする支援に継続して取り組むべき。知事の見解は。」とする。まず、手順 1 では、会議録を参照して、対応する質問を特定する。次に、手順 2 では、質問と対応する答弁とのマッチングを行う。これらの手順においては、質問要約と質問、質問と答弁、それぞれの類似度を算出して、最も高い類似度を持つ質問及び答弁とのマッチングを行う。なお、類似度算出には、単語の共起語数、Jaccard 係数、BM25、BERT [9] を採用する。また、二段階検索の効果を検証するため、BM25 によって得られた上位の結果に対して

表 1 各手法の性能比較

手法	ROUGE-1 F measure
共起語数	.299
Jaccard 係数	.273
BM25	.301
BERT	.142
BM25 + BERT	.172
ナイーブ手法 (タスクオーガナイザー)	.0767

BERT でリランキングを行う BM25 + BERT も評価する。続いて、手順 3 では、選択された答弁から要約答弁を生成する。その際、3.2 節にて構築した自動要約器を用いる。

4 評価実験

提案手法の有用性を評価するため行った評価実験を報告する。図 3 の手順 1, 2 における類似度算出手法である共起語数、Jaccard 係数、BM25 それぞれの性能の評価を行った。なお、評価尺度は ROUGE-1 [10] の F measure である。また、比較手法としてタスクオーガナイザーによって submit されたナイーブな手法 (答弁の末尾 40 文字を要約とする手法) との性能比較結果を表 1 に示す。表より、BM25 において最も精度が高く、次に共起語数、Jaccard 係数と続いた。BERT や二段階検索モデルである BM25 + BERT の性能は低くなったが、いずれにおいてもナイーブ手法と比較して大きく精度が改善した。

5 おわりに

本研究では、一括質問・答弁形式の会議の議事録において、要約された質問を入力されれば、質問と対応する答弁を要約し

5 : <https://github.com/sonoisa/t5-japanese>

6 : 学習は 8 epoch 行った。

て提示するシステムを提案した。その際、提案手法では答弁の要約には転移学習モデルである T5 を用いた。評価実験の結果、提案手法を用いることで、ナイーブな手法と比較して性能が改善した。

文 献

- [1] 大向 一輝. 日本におけるオープンデータの進展と展望. In *情報管理*, number 56, pages 440–447, 2013.
- [2] 木村 泰知, 関根 聡, and 乾 健太郎. 地方議会会議録の要約に向けて. In *言語処理学会 第 24 回年次大会 発表論文集*, 2018.
- [3] 安達 康昭 and 山本 和英. 特徴的冗長表現に着目した国会会議録要約. In *情報処理学会研究報告*, volume NL157-15/FI72-15, pages 107–114, 2003.
- [4] Yasutomo Kimura, Hideyuki Shibuki, Hokuto Ototake, Yuzu Uchida, Keiichi Takamaru, Madoka Ishioroshi, Masaharu Yoshioka, Tomoyoshi Akiba, Yasuhiro Ogawa, Minoru Sasaki, Kenichi Yokote, Kazuma Kadowaki, Tatsunori Mori, Kenji Araki, Teruko Mitamura, and Satoshi Sekine. Overview of the NTCIR-16 QA Lab-PoliInfo-3 Task. In *Proceedings of the 16th NTCIR Conference on Evaluation of Information Access Technologies (NTCIR-16)*, pages 101–112, 2020.
- [5] Yasutomo Kimura, Hideyuki Shibuki, Hokuto Ototake, Yuzu Uchida, Keiichi Takamaru, Kotaro Sakamoto, Madoka Ishioroshi, Teruko Mitamura, Noriko Kando, Tatsunori Mori, Harumichi Yuasa, Satoshi Sekine, and Kentaro Inui. Overview of the NTCIR-14 QA Lab-PoliInfo Task. In *Proceedings of the 14th NTCIR Conference on Evaluation of Information Access Technologies (NTCIR-14)*, pages 289–315, 2019.
- [6] Yasutomo Kimura, Hideyuki Shibuki, Hokuto Ototake, Yuzu Uchida, Keiichi Takamaru, Madoka Ishioroshi, Teruko Mitamura, Masaharu Yoshioka, Tomoyoshi Akiba, Yasuhiro Ogawa, Minoru Sasaki, Kenichi Yokote, Tatsunori Mori, Kenji Araki, Satoshi Sekine, and Noriko Kando. Overview of the NTCIR-15 QA Lab-PoliInfo-2 Task. In *Proceedings of the 15th NTCIR Conference on Evaluation of Information Access Technologies (NTCIR-15)*, pages 101–112, 2020.
- [7] 太刀岡 勇気 and 櫻 惇志. 一括質問・答弁形式の議事録からの質問・答弁のリンキング. In *14 回データ工学と情報マネジメントに関するフォーラム (第 20 回日本データベース学会年次大会) (DEIM 2022)*, G31-4, 2022.
- [8] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. In *arXiv*, 2019.
- [9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019)*, pages 4171–4186, 2019.
- [10] Chin-Yew Lin and Eduard Hovy. Automatic Evaluation of Summaries Using N-gram Co-Occurrence Statistics. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology (NAACL 2003)*, pages 71–78, 2003.