

# 世界史大論述問題を解くための 質問応答システムにおける各国史の利用

福原 優太<sup>†1</sup> 阪本 浩太郎<sup>†1†2</sup> 渋谷 英潔<sup>†1</sup> 森 辰則<sup>†1</sup>

<sup>†1</sup>横浜国立大学 <sup>†2</sup>国立情報学研究所

E-mail: {yuta\_f,sakamoto,shib,mori}@forest.eis.ynu.ac.jp

## 1 はじめに

近年、現実世界を想定した質問応答の取り組みが盛んであり、その一つとして、大学入試問題をコンピュータに解かせる「ロボットは東大に入れるか」プロジェクト<sup>1</sup>やNTCIR<sup>2</sup>のQA Lab[1]がある。QA Labでは大学入試の世界史問題を対象としたタスクが設定されており、二次試験における論述問題も出題されている。論述問題は解答の制限字数から、500字前後で解答する大論述問題と、150字以内で解答する小論述問題に分類できる。我々は文献[2]などにおいて、大論述問題を解答するシステムの開発を行っている。このシステムでは、大学入試の世界史問題を教科書や用語集などを知識源とした情報要求課題と捉えるとともに、論述問題を制限字数以内に収める必要があることから、情報要求の存在する抽出型の複数文書要約と位置付けて解答している。つまり、システムは、要求された情報を知識源から探してくる前半の検索部と、探した情報を適切にまとめる後半の要約部の2段階で構成されている。このシステムは一定の成果を収めているが、解答に含めるべき記述が検索できないなど改善の余地がいくつか残されている。

我々は文献[3]において、模範解答<sup>3</sup>と知識源との対応関係を調査した結果、知識源と対応付けられる模範解答中の記述の割合が文字単位で85.1%であることが分かった。したがって、システムの検索部における問題は、知識源の不足よりも、知識源中の適切な記述を見つけられないことが大きいといえる。適切な記述を見つけられない原因としては、適切な記述中の表現が問題文の表現と一致しないことと、適切な記述中の語句が問題文に書かれていないことの2通りが考えられる。後者に焦点を当てると、問題文中の語句から適切な記述中の語句へと拡張する枠組みが必要である。例えば、世界史を勉強した人間の場合、「ナポレオン」という語句から「エジプト遠征」や「ワーテルローの戦い」といった語句を連想することが可能であり、これに類するモジュールを実現したい。

輝かしい古代文明を建設したエジプトは、その後、連続として5000年の歴史を営んできた。その歴史は、豊かな国土を舞台とするものであるが、とりわけ近隣や遠方から到来して深い刻印を残した政治勢力と、これに対するエジプト側の主体的な対応との関わりを抜きにしては、語るができない。  
(1)エジプトに到来した側の関心や、進出にいたった背景  
(2)進出をうけたエジプト側がとった政策や行動  
の両方の側面を考えながら、エジプトが文明の発祥以来、いかなる歴史的展開をとってきたか概観せよ。解答は、解答欄の(イ)を使用して18行以内とし、下記の8つの語句を必ず1回は用いたうえで、その語句の部分に下線を付せ。  
<指定語句>  
アクティウムの海戦、イスラム教、オスマン帝国、サラディン、ナイル川、ナセル、ナポレオン、ムハンマド・アリー

図 1: 大論述問題の例 (東京大学, 2001 年)

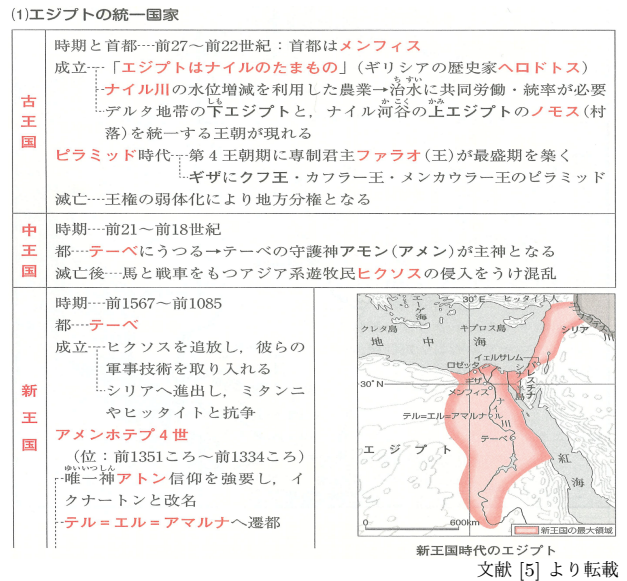


図 2: 各国史の例 (エジプト史)

文献[4]によると、東京大学入試の大論述問題には大きく4つのタイプがあり、その1つに「1つの国、あるいは地域における歴史的流れを叙述させるタイプ」<sup>4</sup>がある。このタイプの大論述問題の例を、図1に示す。この例では、「エジプト」についてのイベントがまとめられている知識源を利用できれば、問題文中に書かれていない語句を推測できるのではないかと考えられる。そのような知識源の一つとして、図2に示すよう

<sup>4</sup>他の3つは、「いくつかの地域における歴史的展開やそれらの関係について叙述させるタイプ」、「いくつかの要素を指定し、それと各地の歴史的展開との関係を叙述させるタイプ」、「2つ以上の歴史的対象を比較・対象して叙述させるタイプ」である。

<sup>1</sup><https://21robot.org/>

<sup>2</sup><http://research.nii.ac.jp/ntcir/index-ja.html>

<sup>3</sup>赤本の解答を模範解答とした。

表 1: 各国史に付与したテーマ

evt:弱いエージェントの統率 evt:家系・王朝の消滅 evt:遷都 evt:混乱 evt:進軍 evt:発見 evt:技術の使用 evt:国家の独立 evt:戦争 evt:領土の喪失 evt:戦いにおける勝利 evt:支援 evt:併合 evt:対立・敵対 evt:領土の拡大 evt:国際会議 evt:条約・協定の締結 evt:道の開通 evt:軍事介入 evt:始まる evt:政治上の革命 evt:戦争における敗北 evt:契約書への署名・調印 evt:民主化運動 evt:侵略 evt:王朝・国家の成立*	evt:最盛期の到来 evt:組織の分立 evt:侵攻 evt:場所からの追放 evt:国家間の戦い evt:弱いエージェントによる成功 evt:支配 evt:遠征 evt:敗北 evt:貿易 evt:組織の設置 evt:人の徴集 evt:組織への編入 evt:出兵 evt:開戦 evt:承認 evt:組織の管理 evt:法・制度・体制への反対 evt:国家の独立運動 evt:協力 evt:法・制度・政策の施行 evt:組織を代表する evt:人物への攻撃 evt:民族移動 evt:宗教運動
---	--

\* このテーマのみ独自に設定した。

な「各国史」が挙げられる。各国史では、複数のテーマ（戦争、政治など）をもとに、ある特定の国における歴史的に重要なイベントが端的に記載されている。以上の背景から、本稿では、大学入試世界史大論述問題を解くための質問応答システムにおける、各国史を知識源として利用した検索拡張手法を提案する。なお、文献 [2] で用いる論述を解答するための知識源を解答知識源、本稿の検索拡張のための知識源を検索拡張知識源と定義する。

## 2 各国史

本稿では、各国史として文献 [5] を電子化して用いた。解答知識源となる教科書や参考書は、記述されるイベントの数が多く、歴史的に重要なイベントも相対的にそうでないものも、詳細な流れで記述されている。それに対し、各国史は、記述されるイベントの数が少なく、歴史的に重要なイベントのみが概要として記述されている。また、各国史の記述は、図 2 に示すように、体言止めや記号による記述など、論述問題の解答としてそのまま用いるのに適切ではない。論述問題には制限字数があり、歴史的な重要性の低いイベント全てを解答に含めることはできない一方で、重要なイベントを幅広く網羅した上で、各イベントを詳細に記述することが要求される。従って、各国史で解答の骨子となる語句を網羅した上で、それらの語句に対応する記述を解答知識源から探すという方法をとる。

## 3 検索拡張知識源からの検索

各国史に記述されるイベントは相対的に少ないとはいえ、全てを解答に含めるには多すぎるため、何らか

ナポレオンのエジプト遠征 (1798-99) ...これを契機にナショナリズムが高揚  
 イタリア遠征 ...ナポレオン=ボナパルトの名声が高まる  
 ナポレオンのエジプト遠征 ←第2回対仏大同盟が結成される  
 第一帝政 ...1804年5月:皇帝ナポレオン1世として開始→これに対して、第3回対仏大同盟が結成される

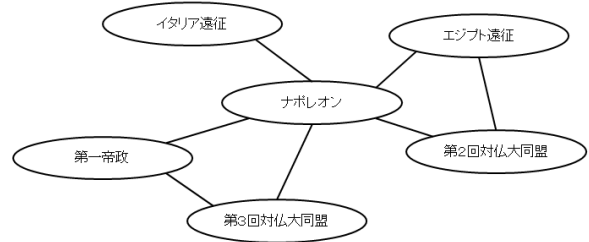


図 3: 固有表現ネットワーク構築の例

の方法で絞り込む必要がある。図 1 の問題文を読むと、「エジプト」が対象であるとともに、「(1) エジプトに到来した側の関心や、進出にいたった背景」という各国の進出、すなわち「戦争」に関連するテーマ、「(2) 進出をうけたエジプト側がとった政策や行動」という「政策」に関連するテーマで解答するように書かれている。こういったテーマを手がかりとすることで、各国史から適切な語句を絞り込むことを考える。文献 [6] では、世界史分野におけるオントロジーを提案しており、歴史的イベントを非常に網羅的かつ詳細に定義している。我々は、世界史オントロジーのイベント集合を参考として各国史の各文にテーマとなる 51 種類のタグ<sup>5</sup>を付与した。付与したタグを表 1 に示す。これにより、テーマに関連する各国史中の文集を求めることができる。以上をテーマによる検索拡張と定義する。

テーマとは別に語句を絞り込む方法として、例えば、「ナポレオン」から「エジプト遠征」、「エジプト遠征」から「ロゼッタストーン」といったように、関連する語句を次々と挙げていく方法が考えられる。基本的に、問題文の指定語句（図 1 参照）に相当する語句は固有表現と考えられる。また、「ナポレオン」と「エジプト遠征」の間の関連性は、例えば、「ナポレオンのエジプト遠征 (1798-99) ...これを契機にナショナリズムが高揚」といった各国史の記述で共起しているかどうかで判断できる。従って、本稿では、各国史の記述に含まれる固有表現をノードとして抽出し、同一記述に含まれる固有表現の間にリンクを張ることで、図 3 に示すようなネットワーク構造（以降、固有表現ネットワーク）を構築する。固有表現ネットワークのリンクを辿ることで、ある固有表現から別の固有表現へと拡張することができる。ただし、各国史の性質上、国名は他の固有表現と比較して非常に多く出現するため、国名をノードとすると殆どのノードが国名にリンクされる歪な固有表現ネットワークが構築される。そのような固有表現ネットワークを用いると、不必要な記述が過剰に検索されるため、国名は固有表現ネットワークのノードとして用いないこととした。以上を固有表

<sup>5</sup>世界史オントロジーのイベント集合を参考としたタグが 50 種類、独自に設定したタグが 1 種類である。

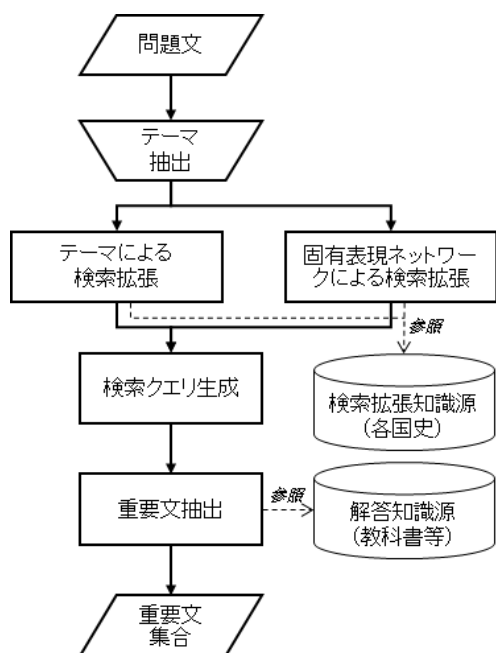


図 4: 提案手法の流れ

現ネットワークによる検索拡張と定義する。

問題文で示される指定語句は解答の重要な手がかりであり、記述すべきイベントの時間的な制約を暗示している。例えば、「ナポレオン」であれば 18 世紀後半から 19 世紀前半の人物であることから、「ナポレオン」を含む各国史の記述も同時期のイベントについてのもと考えられる。図 2 に示すように、各国史は時系列順に書かれていることから、指定語句を含む最初と最後の記述が、時系列的にも最初と最後のイベントとなっている。従って、その範囲外の記述を検索対象から除外することで、さらに絞り込むことができると考えられる。この考え方は、テーマによる検索拡張にも固有表現ネットワークによる検索拡張にも適用することができる。

## 4 提案手法

提案手法の流れを図 4 に示す。提案手法は 2 段階で構成され、第 1 段階（検索クエリ生成まで）では検索拡張知識源から問題文に関連する語句を抽出し、第 2 段階（重要文検索以降）では、第 1 段階で抽出した語句を検索クエリとして解答知識源から解答に含めるべき記述を抽出する。

テーマ抽出では、問題文から「戦争」や「政策」といったテーマを抽出する。この処理は将来的に自動化することが望ましいが、現時点では精度的に困難であるため、第一著者が人手で行った。テーマによる検索拡張では、3 節で付与したタグを用いて、検索拡張知識源からテーマに関連する文集合を求め、その中に含まれる固有表現の集合を求める。固有表現ネットワークによる検索拡張では、問題文の指定語句を起点ノ

ドとして、固有表現ネットワークのリンクを辿り、 $n$  個先までのノードに対応する固有表現の集合を求める。ただし、固有表現ネットワークを構築する際、形態素解析器 MeCab<sup>6</sup> の解析結果において、「品詞細分類 3」が「国」と判定された固有表現については除外する。検索クエリ生成では、両方の検索拡張で求めた固有表現集合の和集合を検索クエリとして求める。重要文抽出では、文献 [2] と同様に、検索クエリの固有表現を含む全てのパッセージ<sup>7</sup> を解答知識源から検索し、当該の固有表現を含むパッセージ中の文を抽出する。

## 5 実験

提案手法が従来手法 [2] の検索部と比較して、どのように改善できたのか実験する。実験には、東京大学の 2001 年度と 1999 年度の入試問題を用いた。固有表現ネットワークにおける  $n$  の値（何個先までリンクを辿るか）は、予備実験の結果から 1 とした。評価指標として、文単位の再現率  $R_s$  と適合率  $P_s$  を用い、それぞれ以下の式で計算する。

$$R_s = \frac{\text{検索された重要文数}}{\text{重要文数}} \quad (1)$$

$$P_s = \frac{\text{検索された重要文数}}{\text{検索された文数}} \quad (2)$$

重要文とは、模範解答と対応付けられる解答知識源中の文である。同様に、パッセージ単位の再現率  $R_p$  と適合率  $P_p$  を以下の式で計算する。

$$R_p = \frac{\text{検索された重要パッセージ数}}{\text{重要パッセージ数}} \quad (3)$$

$$P_p = \frac{\text{検索された重要パッセージ数}}{\text{検索されたパッセージ数}} \quad (4)$$

重要パッセージとは重要文を含むパッセージである。

## 6 考察

実験結果を表 2 に示す。2001 年度と 1999 年度の結果を比較すると、従来手法の適合率に若干の差がみられるものの、全体としては同程度の結果となり、ひとまず年度による影響はないものとして考察を行う。従来手法と提案手法を比較すると、適合率が下がる一方で再現率が大きく向上している。提案手法が大論述問題解答システムの検索部にあたることを考慮すると、解答に不要な記述は後続の要約部により取り除かれる機会が残されている。その意味で、検索部では、解答に必要な記述を網羅的に検索できることが重要であり、適合率よりも再現率を重視した方が良いと考えられる。したがって、提案手法のパッセージの再現率が、2001 年

<sup>6</sup><http://taku910.github.io/mecab/>

<sup>7</sup>教科書であれば 1 つの段落を、用語集であれば 1 つの見出し語の解説を 1 パッセージとした。

表 2: 重要文 (パッセージ) 検索の再現率と適合率

	文				パッセージ			
	再現率		適合率		再現率		適合率	
	2001 年	1999 年	2001 年	1999 年	2001 年	1999 年	2001 年	1999 年
従来手法 [2]	0.133	0.123	0.024	0.118	0.656	0.695	0.080	0.394
提案手法	0.315	0.301	0.020	0.017	0.918	0.966	0.046	0.036

表 3: 検索拡張 (各国史からの抽出結果) の再現率と適合率

	再現率		適合率	
	2001 年	1999 年	2001 年	1999 年
テーマによる検索拡張	0.846	0.647	0.647	0.647
固有表現ネットワークによる検索拡張	0.615	1.000	0.485	0.654
両方の検索拡張結果の積集合	0.462	0.647	0.632	0.846
両方の検索拡張結果の和集合	1.000	1.000	0.510	0.567

度で 0.656 から 0.918 に、1999 年度で 0.695 から 0.966 に大きく向上したことは改善と考えている。

次に、テーマによる検索拡張と固有表現ネットワークによる検索拡張がそれぞれどのように影響しているかを考察する。検索拡張の結果を評価するために、どちらの出力も固有表現ではなく検索拡張知識源の記述 (以降、拡張記述) とし、拡張記述が模範解答と対応付けられる場合を重要拡張記述として、以下の式で再現率  $R_e$  と適合率  $P_e$  を求めた。

$$R_e = \frac{\text{出力された重要拡張記述数}}{\text{重要拡張記述数}} \quad (5)$$

$$P_e = \frac{\text{出力された重要拡張記述数}}{\text{出力された拡張記述数}} \quad (6)$$

拡張記述は、それぞれ以下のように求めた。テーマによる検索拡張では、固有表現を求める直前のテーマに関連する文を拡張記述とした。固有表現ネットワークによる検索拡張では、リンク元とリンク先の両方の固有表現を含む文を拡張記述とした。また、両方の検索拡張が出力した拡張記述の積集合と和集合に対する再現率と適合率も計算した。結果を表 3 に示す。適合率の観点からはテーマによる検索拡張の方が固有表現ネットワークによる検索拡張よりも高い値となった。再現率では年度によって異なるためどちらが高いとはいえないが、和集合とすることで再現率が 1.000 となり、相補的な関係となっているといえる。積集合の結果を見ると、それぞれ単体での結果よりも適合率が向上しており、これを上手く利用することで解答知識源の検索の適合率を向上できないか今後検討していきたい。

## 7 まとめ

本稿では、大学入試世界史大論述問題を解くための質問応答システムにおける、各国史を知識源として利用した検索拡張手法を提案した。提案手法は、各国史で解答の骨子となる語句を網羅した上で、それらの語句に対応する記述を解答知識源から探すという 2 段階

の処理を行う。東京大学の 2 年分の入試問題を用いて実験した結果、従来手法より適合率が低下したものの再現率が大きく向上し、要約部の前処理としての観点からは改善されたと考える。今後は、適合率の向上も考慮しつつ、100%の再現率となるよう改善していきたい。

## 謝辞

本研究の一部は、JSPS 科研費 16K00296 の助成を受けたものである。

## 参考文献

- [1] Hideyuki Shibuki, Kotaro Sakamoto, Madoka Ishioroshi, Yoshinobu Kano, Teruko Mitamura, Tatsunori Mori, Noriko Kando. Overview of the NTCIR-13 QA Lab-3 Task. Proceedings of the 13th NTCIR Conference, 2017.
- [2] 阪本浩太郎, 中山周, 渋谷英潔, 石下円香, 森辰則, 神門典子. 東大入試世界史第 1 問 (大論述問題) を解く質問応答システムの検討. 言語処理学会第 22 回年次大会発表論文集, pp. 553-556, 2016.
- [3] 福原優太, 阪本浩太郎, 渋谷英潔, 森辰則. 世界史論述問題における模範解答-知識源の対応を表すアノテーションの検討. 言語処理学会第 23 回年次大会発表論文集, pp. 593-596, 2017.
- [4] 佐藤貢, 東大の世界史 25 年 [第 5 版], 教学社, 2016.
- [5] 今泉博, 飯屋園巖, 近藤泉子編. 流れがわかる各国別・地域別世界史 B の整理. 株式会社山川出版社, 第 3 版, 2017.
- [6] 川添愛, 宮尾祐介, 松崎拓也, 横野光, 新井紀子. 出来事の成立・不成立の判断をサポートするイベントオントロジー. 人工知能学会第 28 回全国大会, 2014.