

# 歴史的類推を促すための現代の因果関係検索学習支援システム

松丸 健太† 池尻 良平‡ 澄川 靖信\*

拓殖大学†\* 東京大学‡

## 1 はじめに

歴史教育では、暗記だけでなく歴史的思考力の育成が重視されている。特に近年では歴史資料を用いて過去を構築する力に加え、歴史と現代を関連させる力（以降、歴史的類推と呼ぶ）も重視されている[1]。この重要性は歴史学習支援の研究領域でも認められており、現代のニュースに類似する過去の因果関係を検索するシステムを用いて歴史的類推を促進しうる授業設計が提案されている[2]。しかし、この授業設計は学習者が既に一定の歴史を学習していることを想定しているため、学んだばかりの知識を使いながら歴史的類推を促進させることが難しい。

本研究では、学んだばかりの過去の因果関係を用いて現代で生じている諸問題の解決方略を考える足場かけを提供する歴史学習支援システムを実現するために、過去の因果関係を入力すると類似する現代の因果関係を検索するアルゴリズムを提案する。

## 2 提案手法

本手法は、特定の時期の記述でしか使われない時間的固有表現を非時間的固有表現に変換した後、現代と過去の因果関係どうしの類似度を算出し、ランキング形式で類似度の高い順に現代の因果関係を表示する。以下に、時間的固有表現の検出方法をまず述べる。その後、非時間的固有表現への変換と類似度の算出方法について述べる。

時間的固有表現を抽出するために、まず、TagME[3]を文章に適用し、固有表現とそれらのWikipedia記事を取得する。もしこの固有表現が人、組織、出来事のいずれかであれば直ちに時間的固有表現とみなす。もし固有表現が国であれば存続年数を解析する。固有表現の種類を明らかにするために、TagMEの結果から得られるWikipedia記事のカテゴリを取得する。これらの字面にpeople, organization, event, countryが含まれるかを解析することで上記4種類の固有表現の型を判定する。各国が時間的固有表現かどうかを解析するために、図1に示す国の存続年数

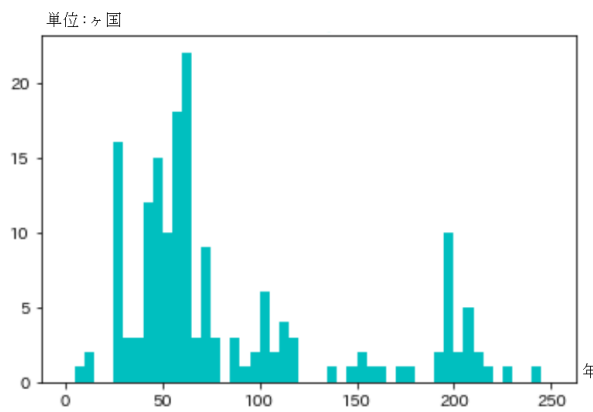


図1 国の存続年数分布

分布をWikipediaから事前に解析した。この結果、120年を境界とする二峰性を確認できるので、存続年数が120年以下ならその国を時間的固有表現とみなす。時間的固有表現を検出した場合、時間的固有表現の字面をその固有表現の型に置換する。

因果関係どうしの類似度を算出するために、まず、各因果関係に対して特徴ベクトルを作成する。本研究では、関連する出来事を時系列順に並べたものを因果関係と定義し、出来事の記事ごとに特徴ベクトルを作成する。文章の特徴ベクトルはlatent semantic analysisによって求められるトピック分布を用いる。この後、入力される過去の因果関係とDBにある現代の因果関係の組み合わせに対して二部グラフを構築して各現代の因果関係との類似度を求める。この二部グラフは、因果関係の各出来事を節、コサイン類似度で計算した節どうしの類似度をラベルに持つ辺によって定義する。この二部グラフ上で、因果関係の類似度を計算するように、解となる辺集合に交点を持たないように最大重みマッチングを求めるevent causal relationship measurement (ECM)[4]を適用する。

ECMを適用すると類似度を表す数値が得られるので、これらの値の上位k件を検索結果として出力する。

## 3 実験

評価用データには歴史の教科書を参考に専門家が手作業で作成した過去の因果関係データセット[5]と、2016年に起きた出来事としてWikipediaに記録されたものを手作業で時系列順にトピックとしてまとめた現代の因果関係デー

Causal Relationship Search based Learning System for Historical Analogy Promotion

† Kenta Matsumaru, Takushoku University

‡ Ryohei Ikejiri, The University of Tokyo

\* Yasunobu Sumikawa, Takushoku University

タセット[6]を用いる。これらのデータセットには因果関係ごとにカテゴリが付与されているので、本稿では入力と出力のカテゴリが一致しているかどうかを評価した。なお、両者のデータセットが定義するカテゴリのうち共通する **Diplomacy**, **Reign**, **Literature**, **Environment**, **Technology** の 5 カテゴリだけを使用した。実際に使用した評価基準は、上位 1 件に対しては適合率 (P), 再現率 (R), F1 値 (F1), 上位 10 位に対して mean average precision (MAP) をそれぞれ使用した。比較対象として、1) Jaccard 係数を用いた文章の類似度を求める手法 (Jaccard), 2) 潜在意味解析とコサイン類似度を用いた手法 (cos), 3) 時間的固有表現変換を行わずに二部グラフを構築する ECM (ECM) の 3 つを用いる。

結果を表 1 に示す。全ての結果において提案手法が最も高い精度が得られたことがわかる。特に上位 1 位に注目する適合率・再現率・F1 値は 15~5.5%ポイントも 3 つの比較手法より提案手法が良い結果だった。この結果から、因果関係を検索するときは、特定の期間にしか出現しない固有表現の置換と時系列順に文章の類似度を評価することが重要であることがわかる。

提案手法の結果を深く分析するために上位 1 位として検索できた因果関係と入力とのカテゴリの一致率を分析した結果を図 2 に示す。縦軸は過去の因果関係のカテゴリであり、横軸は現代の因果関係のカテゴリである。この結果から **Diplomacy** と **Reign** では高い精度が得られたことがわかる。精度が低い結果を分析するために、

表 1 適合率, 再現率, F1 値, MAP 値の比較

	P	R	F1	MAP
Jaccard	39.1%	37.5%	37.7%	64.7%
cos	41.7%	39.4%	40.7%	58.8%
ECM	48.9%	45.0%	46.2%	67.2%
提案手法	54.4%	51.4%	53.1%	71.3%

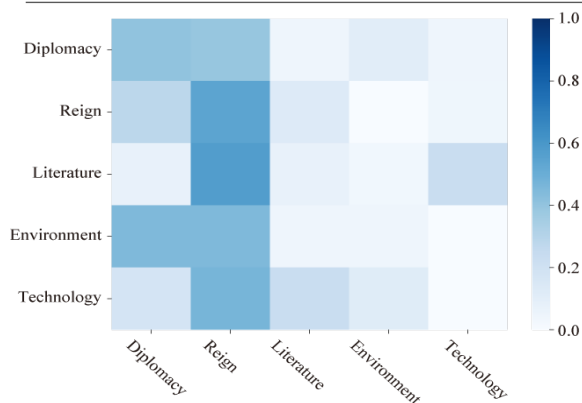


図 2 文章どうしのカテゴリ一致度

カテゴリどうしの類似度と依存度を Jaccard 係数と相互情報量を用いて測定したところ、2 つの値共に、同一カテゴリ内の文章どうしのスコアが **Diplomacy** と **Reign** よりも低かった。また誤検索したカテゴリとの相互情報量が比較的高い傾向が確認できた。誤検索の結果を手作業で確認したところ、本実験では 2 つのデータセットが共通して定義するカテゴリを評価に用いたが、例え同一名称のカテゴリでも対象とする出来事の範囲が厳密な整合性がないことがわかった。例えば、現代のデータセットの **Diplomacy** カテゴリには国際関係や経済関連のニュースを含んでいるが、これらのニュースは国内統治や海外侵略にも関係するものだった。

#### 4 おわりに

本稿では過去の因果関係を入力すると類似する現代の因果関係を出力する検索アルゴリズムを提案した。本手法は時間的固有表現の抽出と置換を行い、因果関係どうしの類似度を求める。本手法の有効性をデータセットを用いて評価したところ比較対象よりも高い精度が得られた。

今後の課題として、本アルゴリズムを用いた学習方法の設計がある。本手法を用いると歴史的類推がどの程度促進されるのかを確認する。

謝辞：本研究の一部は科研 (#22K02884) の助成を受けたものである。

#### 参考文献

- [1] 文部科学省 (2018) . 高等学校学習指導要領解説地理歴史編 (平成 30 年 8 月改訂版) . [https://www.mext.go.jp/content/20220802-mxt\\_kyoiku02-100002620\\_03.pdf](https://www.mext.go.jp/content/20220802-mxt_kyoiku02-100002620_03.pdf) (参照日 2022 年 11 月 18 日)
- [2] 池尻良平, 澄川靖信. 真正な社会参画を促す世界史の授業開発—その日のニュースと関連した歴史を検索できるシステムを用いて—. 社会科学研究, 84:37–48, 2016.
- [3] Ferragina, P. and Scaiella, U.: Fast and Accurate Annotation of Short Texts with Wikipedia Pages, IEEE Software, Vol. 29, No. 1, pp. 70–75 (2011).
- [4] Sumikawa, Y.: Event Causal Relationship Retrieval. WI-IAT '21. pp. 318–325, 2021.
- [5] Ikejiri, R. and Sumikawa, Y. Past Causalities and Event Categories for Connecting Similar Past and Present Causalities. <https://doi.org/10.5281/zenodo.3601707>, 2019.
- [6] Hoang, T.A., Vo, K.D., Nejdil, W.: W2e: A Worldwide-Event Benchmark Dataset for Topic Detection and Tracking. CIKM '18. pp. 1847–1850, 2018.