

# 語彙の概念化と Wikipedia を用いた英字略語の意味推定手法 Meaning Estimation Method of Alphabetical Abbreviation using Conceptualization of a Word and Wikipedia

後藤 和人<sup>1\*</sup> 土屋 誠司<sup>2</sup> 渡部 広一<sup>2</sup>  
Kazuto Goto<sup>1</sup> Seiji Tsuchiya<sup>2</sup> Hirokazu Watabe<sup>2</sup>

<sup>1</sup> 同志社大学大学院 理工学研究科

<sup>1</sup> Graduate School of Science and Engineering, Doshisha University

<sup>2</sup> 同志社大学 理工学部

<sup>2</sup> Faculty of Science and Engineering, Doshisha University

**Abstract:** In this paper, we propose the method to estimate the meaning of an alphabetical abbreviation. To resolve the polysemy of alphabetical abbreviations, this method uses the concept base, Wikipedia, and, the evaluation approach for semasiological association between sentences. The concept base allows Conceptualization of a word. We apply Calculation of Degree of Association or Earth Mover's Distance (EMD) to evaluate semasiological association between sentences. This paper used 129 articles to evaluate the proposal method. 129 articles included 129 meanings of alphabetical abbreviations and 58 literation of alphabetical abbreviations. The experiments showed that the accuracy of the proposal method was close to 80%.

## 1 はじめに

元来から日本は、外来語を受け入れやすい環境にあるといわれており、数多くの外国の言葉を片仮名として表記し、そのまま使用している。近年になり、よりグローバル化が激しくなると共に、外来語が益々増加する中、外来語の発音を片仮名表記にしないケースが見受けられる。特に、英語の場合、外国語の表記をそのまま利用するシーンも増えてきている。また、英単語などの頭文字をつなげて表記する、いわゆる略語もよく利用されるようになってきている。例えば、「IC」といった英字略語がそれにあたる。

しかし、英字略語は英単語の頭文字から構成される表現であるため、まったく別のことを表現しているにも関わらず、同じ表記になることが多い。先の英字略語「IC」を例にとると、「IC」には「集積回路」という意味や、高速道路などの「インターチェンジ」の意味がある。さらには、ある業界では、これらとはまた別の意味で使用されることもある。

このように、英字略語は便利な反面、いわゆる一般的な単語よりも非常に多くの意味を有する多義性の問題を持っている。そのため、英字略語が利用されている情報は、すべての人が容易に、また、正確に把握で

きているとは言い難い。

特に、知能を有し、情報収集をはじめとする手段により上記問題を解決できる人と比べ、機械にとって、英字略語の理解はより困難な問題といえる。近年、機械は我々の生活・社会と密接に関与し、必要不可欠な存在となりつつある。機械の目指すべき姿は「人と共存する機械（ロボット）」だといえる。二足歩行ができる身体能力に優れたロボットなどが数多く開発されたことにより、その一部が実現されつつある。今後、ロボットが真に「人と共存」するためには、優れた身体能力を持つロボットに「知能」を持たせ、その知能に基づいて会話することが重要になる。

ロボットに知能を持たせるためには、人が日常的に行っている常識的判断を実施できるようにすることが必要になる。そのためには、ロボットに単語を理解させることが必要であり、その手段として、語彙を概念化する研究が行われている。例えば、[1]では、ある概念に対して、当該概念の意味特徴を表す属性と属性の重みを登録したデータベースを構築している。そして、構築したデータベースをもとに、概念同士の意味的な近さを計算できる関連度評価アルゴリズムを提案している。また、[2]では、シソーラスを利用する方法と特異値分解を用いる数学的方法を併用して、単語間の属性空間を構築する方式が提案されている。さらに、[3]は、コーパスから抽出した情報をもとに、格フレームを

\*同志社大学大学院理工学研究科情報工学専攻  
〒610-0394 京都府京田辺市多々羅谷 1-3  
E-mail: eup1102@mail4.doshisha.ac.jp

自動的に構築する手法を提案している。これらの手法は、それぞれアプローチは異なるものの、ロボット（計算機）に単語や文章を理解させることを目的している。語彙を概念化することで、単語の表記に頼らず、単語や単語同士の関係を表現することを実現している。結果、単語の意味を考慮した情報検索や応答が可能となり、常識を有するロボットの実現につながることを期待されている。

さらに、ロボットの知能を発展させるために、ユーザの入力（会話など）に対する応答ではなく、ロボットが起点となり、会話文を提示する研究がなされている[4]。この手法では、新聞記事のヘッドラインを手掛かりに、取得した情報の意味を理解した上で、情報を会話文形式に変換し、会話文を自動生成することを実現している。このとき、新聞記事に英字略語が含まれており、かつ、英字略語の意味を正しく理解できなかった場合、ロボットはユーザに誤った会話文を提示する可能性がある。このように、知能を有するロボットの研究開発において、多義性を有する英字略語の意味を正しく理解できなければ、自然な知的対話の実現は困難であるといえる。

英字略語の意味を明示するために、文章（新聞記事など）の中で最初に英字略語が使用される箇所において、括弧書きでその意味を日本語で併記する処置をとっていることが多い。しかし、よく知られている英字略語にはそのような処置をとらないなど、完全に対処されているわけではない。また、文章中において、最初の箇所のみ上記のような処置がとられており、それ以降はその意味が併記されないことが多い。そのため、途中から文章を読んだ場合、最初にその英字略語が出現した箇所を探すことが必要になる。その結果、解読にはひと手間が必要となり、理解の妨げとなる。さらに、このように英字略語の意味を併記するという処置がとられていないこともある。

これらを踏まえて本論文では、英字略語の意味を推定する手法について提案する。本研究と同様の主旨の研究には、[5]の研究が報告されている。この手法では、先に述べたように新聞記事において英字略語の意味を括弧書きで併記する表現に着目し、英字略語の意味の自動推定を実現している。しかし、前述のように、すべての英字略語に対してこのような表現が適用されているわけではない。そのため、うまく自動推定できない場合がある。また、[6]では、片仮名表記の外来語を英語に復元した後に辞書を用いて日本語訳を獲得する手法が提案されている。しかし、本手法では、対象が片仮名語に限定されており、かつ、多義性を有する語彙には対応できない問題がある。

そこで、本論文では、多義性を有する英字略語に対して、意味の推定を実現する方法を提案する。提案手法では、我々がすでに提案している、あらゆる語彙の

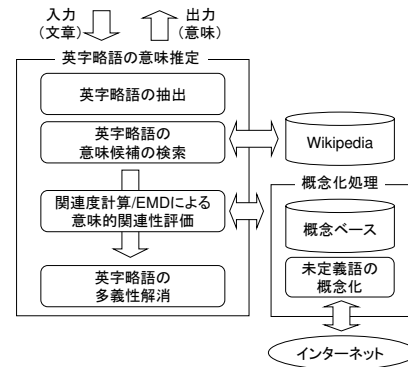


図 1: 英字略語の意味推定手法の概略図。

意味的な近さを判断できるメカニズムを組み合わせることで、英字略語の意味を推定する。具体的には、ある概念から様々な概念を連想する語彙の概念化処理が可能な概念ベース [7, 8, 9], 及び、世界で最も収録語数が多いとされている Wikipedia[10] を使用する。さらに、概念化した語彙の意味的な近さを判断するために、関連度計算 [11], または、Earth Mover's Distance を応用した文章間関連度計算手法 [12] を用いる。これらを用いることで、英字略語の多義性を解消し、英字略語の本来の意味を推定する。この技術により、英字略語の意味を理解しやすくなることができ、自然な知的対話の実現に加え、情報検索や自動要約、情報推薦や自動翻訳など多くのアプリケーションの性能向上に寄与することも期待できる。

## 2 提案手法の概要

図 1 に提案手法である英字略語の意味推定手法の概略図を示す。英字略語が含まれる文章を入力として、その文章から英字略語を抽出する。その英字略語を Wikipedia で検索し、意味が 1 つであれば、その意味を出力する。意味が複数ある場合には、それらの意味と入力文章との意味的な近さを判断し、最も近いと判断した意味を決定する。なお、ここで述べる「意味」とは、英字略語の意味を表現する語、つまり、英字略語の基となっている英単語の日本語での表現を「意味」と定義している。例えば、前述した英字略語「IC」の意味を推定する場合、「集積回路」や「インターチェンジ」という語を「意味」として出力することになる。

3 章では、本論文において使用した要素技術として、語彙を概念化する手法と、概念化した語彙の意味的な近さを判断する手法に関して詳細に説明する。

表 1: 概念ベースの例.

概念	属性
医者	(医師, 0.34), (患者, 0.11), (病院, 0.08), …
病院	(医院, 0.25), (手術, 0.18), (施設, 0.04), …
患者	(病人, 0.52), (看病, 0.21), (治療, 0.12), …
…	…

### 3 使用要素技術

#### 3.1 語彙の概念化処理

##### 3.1.1 概念ベース

概念ベース [7, 8, 9] とは、複数の電子化国語辞書などの見出し語を概念、その語義文に使用されている自立語を概念の意味特徴を表す属性と定義して構築された大規模なデータベースである。本論文で使用した概念ベースは自動的に概念および属性を構築した後、人間の常識に沿った属性の追加や削除を人手で行ったものであり、概念数は約 9 万語である。

概念ベースでは、ある概念  $A$  は  $m$  個の属性  $a_i$  とその属性の重要性を表す重み  $w_i$  の対によって構成されており、以下のように表現することができる。ここで、属性  $a_i$  を概念  $A$  の一次属性と呼ぶ。

$$\text{概念 } A = \{(a_1, w_1), (a_2, w_2), \dots, (a_m, w_m)\}$$

概念ベースの大きな特徴として、属性である単語は概念として必ず定義されている点がある。これにより、概念  $A$  の一次属性である属性  $a_i$  を概念とみなし、更に属性を導くことができる。概念  $a_i$  から導かれた属性  $a_{ij}$  を、元の概念  $A$  の二次属性と呼ぶ。概念ベースの具体例を表 1 に示す。例えば、表 1 のように、概念「医者」の一次属性である「患者」は、概念「患者」としても定義されている。また、この概念「患者」の一次属性である「病人, 看病, 治療, …」は、元の概念「医者」の二次属性ということになる。

このように、概念ベースにより概念の意味特徴を定義し、連鎖できる構造を利用することで語彙の意味の近さを評価することができる。詳細は 3.2 節で説明する。

##### 3.1.2 未定義語の概念化

前節で述べたとおり、概念ベースは複数の電子化国語辞書などを用いて構築されており、大規模かつ品質が高いというメリットがある。しかし、その反面、すべての語彙を網羅できていないという欠点もある。そのため、概念ベースに登録されていない未定義語は概念化されておらず、意味の近さを評価することはできない。

そこで、未定義語については、Web 上の言語情報を利用し、自動的に概念化すること [13, 14] で対処する。具体的には、Web 検索エンジン [15] により未定義語をキーワードとして情報検索し、検索結果の上位 100 件の検索結果ページの内容を取得する。その内容から概念ベースに登録されている自立語のみを抽出し、それらを未定義語の一次属性とする。また、一次属性に対する重みは、情報検索の分野で広く用いられている  $tf/idf$  [16] の考え方を応用することで算出する。

語の網羅性である  $tf$  値は、検索結果ページ  $A$  中出现する自立語  $Word_{A_i}$  の出現頻度  $tfreq(Word_{A_i}, A)$  を、検索結果ページ  $A$  中のすべての自立語の語数  $tnum(A)$  で割ることで算出される。算出式は以下のようになる。

$$tf(Word_{A_i}, A) = \frac{tfreq(Word_{A_i}, A)}{tnum(A)}$$

次に、語の特定性である  $idf$  値については、 $SWeb-idf$  値 [13, 14] を用いる。 $idf$  値の算出には、対象となる全文書空間の情報が必要になる。しかし、Web を利用する場合、Web 上のすべての情報が必要ということになり、正確な  $idf$  値を算出することは現実的には不可能である。

そこで、無作為に選択した固有名詞 1000 個をそれぞれキーワードとして、Web 検索エンジン [15] で検索する。続いて、検索結果の上位 10 件ずつ検索結果ページの内容を取得する。そして、それらの内容に含まれるすべての自立語の集合を疑似的に Web の全情報空間とみなし  $SWeb-idf$  値を算出する。 $SWeb-idf$  値の算出式は以下のように定義される。ここで  $N$  は、固有名詞 1000 個を検索キーワードとした際の各検索結果上位 10 件の合計ページ数 ( $N = 10000$ )、 $df(Word_{A_i})$  は、 $Word_{A_i}$  が出現する検索結果ページ数である。

$$SWeb-idf(Word_{A_i}) = \log \frac{N}{df(Word_{A_i})}$$

この 10000 ページから、複数の国語辞書や新聞などから概念を抽出したデータベースである概念ベースの収録語数である約 9 万語以上の単語数が得られたことから、獲得した 10000 ページを Web の全情報空間とみなしている。なお、固有名詞の選び方を変えても  $SWeb-idf$  値に大きな変化は見られないという報告がなされている [13]。

以上に示した式より、自立語  $Word_{A_i}$  へ付与する重み  $w$  は次の式で定義される。つまり、ある自立語の重みは、網羅性を表す  $tf$  値と特定性を表す  $SWeb-idf$  値を掛け合わせることで与えられる。

$$w = tf(Word_{A_i}, A) \times SWeb-idf(Word_{A_i})$$

Web 上の言語情報を利用するため、品質という面では概念ベースより劣るというデメリットはあるが、す

すべての語彙を概念化できるという大きなメリットがある。また、[13, 14]などの研究成果から、本論文での処理においても十分な性能を確保していると考えられる。

## 3.2 意味的な関連性評価手法

### 3.2.1 関連度計算

関連度計算とは、概念と概念の関連の強さを定量的に評価するものである。概念と概念の間にある関連性を定量的に評価する手法として、ベクトル空間モデルが広く用いられている。しかし、本論文では、概念を定義する属性集合とその重みを含めた一致度に基づいた関連度計算方式を利用している。これは、関連度計算方式が有限ベクトル空間によるベクトル空間モデルよりも良好な結果が得られるという報告がなされているためである [1]。本論文では、重み比率付き関連度計算方式を使用する [11]。

任意の概念  $A, B$  について、それぞれ一次属性を  $a_i, b_j$  とし、対応する重みを  $u_i, v_j$  とする。また、概念  $A, B$  の属性数を  $L$  個,  $M$  個 ( $L \leq M$ ) とする。なお、各概念の一次属性の重みは、その総和が 1.0 となるよう正規化している。

$$\begin{aligned} A &= (a_i, u_i) \mid i = 1 \sim L \\ B &= (b_j, v_j) \mid j = 1 \sim M \end{aligned}$$

このとき、概念  $A, B$  の重み比率付き一致度  $MatchWR(A, B)$  は以下のように定義される。

$$\begin{aligned} MatchWR(A, B) &= \sum_{a_i=b_j} \min(u_i, v_j) \\ \min(\alpha, \beta) &= \begin{cases} \alpha & (\beta > \alpha) \\ \beta & (\alpha \geq \beta) \end{cases} \end{aligned}$$

この定義は、概念  $A, B$  に対し、 $a_i = b_j$  となる属性 (概念  $A, B$  に共通する属性) があった場合、共通する属性の重みの共通部分、つまり、小さい重み分のみ一致するという考えに基づいてる。

次に、属性の少ない方の概念を  $A$  とし ( $L \leq M$ )、概念  $A$  の属性を基準とする。

$$A = \{(a_1, u_1), \dots, (a_i, u_i), \dots, (a_L, u_L)\}$$

そして、概念  $B$  の属性を、概念  $A$  の各属性との重み比率付き一致度  $MatchWR(a_i, b_{x_i})$  の和が最大になるように並び替える。

$$B_x = \{(b_{x_1}, v_{x_1}), \dots, (b_{x_i}, v_{x_i}), \dots, (b_{x_L}, v_{x_L})\}$$

これによって、概念  $A$  の一次属性と概念  $B$  の一次属性の対応する組を決める。対応にあふれた概念  $B$  の属

性は無視する。ただし、一次属性同士が一致する (概念表記が同じ) ものがある場合 ( $a_i = b_j$ ) は、別扱いにする。これは概念ベースには約 9 万語の概念が存在し、属性が一致することは稀であるという考えに基づく。従って、属性の一致の扱いを別にするにより、属性が一致した場合を大きく評価する。具体的には、対応する属性の重み  $u_i, v_j$  の大きさを重みの小さい方にそろえる。このとき、重みの大きい方はその値から小さい方の重みを引き、もう一度、他の属性と対応をとることにする。例えば、 $a_i = b_j$  で  $u_i = v_j + \alpha$  とすれば、対応を決定する箇所は  $(a_i, v_j)$  と  $(b_j, v_j)$  であり、 $(a_i, \alpha)$  はもう一度他の属性と対応させる。このように対応を決め、対応がとれた属性の組み合わせの数を  $T$  個とする。

重み比率付き関連度とは、重み比率付き一致度を比較する概念の各属性間で算出し、その和の最大値を求めることで計算する。概念  $A, B$  の関連度  $DoA(A, B)$  は以下の式で定義される。

$$DoA(A, B) = \sum_{i=1}^T \{ MatchWR(a_i, b_{x_i}) \cdot (u_i + v_{x_i}) \cdot (\min(u_i, v_{x_i}) / \max(u_i, v_{x_i})) / 2 \}$$

以下、重み比率付き一致度を一致度、重み比率付き関連度を関連度と略し、この関連度 [11] を用いる。関連度は概念間の関連の強さを 0~1 の間の連続値で表す。

### 3.2.2 Earth Mover's Distance

前節において、概念間の関連の強さを評価する方法として関連度計算について記載した。関連度計算は関連性が高い順に属性の対応をとることで計算を行う、つまり、1対1で対応をとる手法である。そのため、両概念の中で少ない方の属性の数しか対応がとれない。例えば、概念  $A$  が持つ属性が 3 語、概念  $B$  が持つ属性が 100 語であった場合、概念  $B$  の属性 97 語は計算の対象外となる。本論文では、両概念の属性数に差がある状況にも対応するため、関連度計算に加えて、M対Nで対応をとることができる Earth Mover's Distance (EMD) を用いて意味的な関連性を評価する手法を利用する。

EMD を用いた意味的な関連性評価手法は、ヒッチコック型輸送問題 [17] (需要地の需要を満たすように供給地から輸送を行う際の最小輸送コストを解く問題) で計算される距離尺度である EMD を概念間の関連性評価に適用したものである。EMD は、2つの概念間の関連性を定量的に表現することが可能であり、[12]の研究によりその有用性が報告されている。

EMD とは、2つの離散分布があるとき、一方からもう一方の分布への変換を行う際の最小コストを指す。離



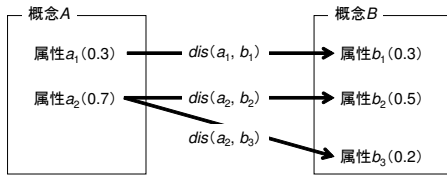


図 2: EMD を用いた意味的な関連性評価手法.

散分布はそれを構成する要素と重みの対の集合で表現される。コスト算出の際には、変換前の離散分布の要素が持つ重みを供給量、変換先の離散分布の要素が持つ重みを需要量と考え、要素間の距離を供給量、需要量にしたがって重みを運送すると考える。できるだけ短い距離で、かつ、需要量に対して効率的に重みを運送する経路が EMD となる。

これを概念間の関連性評価に適用させる際には、概念の一次属性を要素として捉え、一次属性の集合を離散分布と考える。ある概念の離散分布を違う概念の離散分布へ変換すると考えると、その際のコストが最小となる概念が元の概念に最も近い概念となり概念間の関連性評価へ適用することが可能となる。

EMD を用いた意味的な関連性評価手法について、図 2 に示すような簡略図を用いて説明する。ある概念 A と B があったとき、概念 A を概念 B に変換する際のコストを考える。それぞれの概念をそれらの一次属性  $a_i$ ,  $b_j$  の離散分布と考える。EMD では変換コストの算出を行う際に離散分布を構成する要素同士の距離を用いる。EMD を用いた意味的な関連性評価手法では、この距離を一次属性同士の関連性であると考え、一致度によってこれを求める。

属性  $a_i$ ,  $b_j$  の距離  $dis(a_i, b_j)$  は次の式で表される。一致度は関連性が高いと値が大きくなる。また、一致度の最大値は 1 であるため、1 から一致度を引いた値を距離としている。

$$dis(a_i, b_j) = 1 - MatchWR(a_i, b_j)$$

ここで、図 2 の例における  $a_1$  と  $b_1$  の間の変換コスト  $cost(a_1, b_1)$  は次の式で算出される。これは  $a_1$  と  $b_1$  の距離に重みを掛けたものである。 $a_1$  と  $b_1$  が持つ重みは同じく 0.3 であるため供給量と需要量が合致し、 $a_1$  からの重みの運送はこの時点で終了する。

$$cost(a_1, b_1) = dis(a_1, b_1) \cdot 0.3$$

同様にコストの計算を行っていき、最終的にすべての運送経路のコストを足し合わせたものが EMD となる。図 2 の例では概念 A, B の EMD は次のように表される。

$$EMD(A, B) = cost(a_1, b_1) + cost(a_2, b_2) + cost(a_2, b_3)$$

以上のような式で算出された EMD の値の最小値を最適化計算で求め、概念間の関連性（意味的な近さ）を算出している。

## 4 英字略語の意味推定手法

### 4.1 英字略語の抽出

本論文で提案する英語略語の意味推定手法の処理の対象として扱う英字略語とは、英単語の頭文字から構成される表記とする。例えば、商品の型番や「W 杯」のように記号や数字、日本語などアルファベット以外の文字が混じる表記の場合、それらは英字略語ではないものとする。また、1 文字で構成される英字略語の場合、英単語の頭文字ではなく、例えば、S 字カーブの「S」のように、アルファベットの形状などに起因する意味で使用されることがある。本研究は、語彙の意味に着目し、多義性を有する英字略語の意味を推定することを目的としている。そのため、本論文では、2 文字以上のアルファベットのみで構成されている語を英字略語として扱うこととする。

入力として受け付ける情報は、英字略語が含まれている文章とし、その文章から 2 文字以上のアルファベットの羅列を英字略語として抽出する。

### 4.2 Wikipedia による意味候補の検索

4.1 節で抽出した英字略語の意味を Wikipedia[10] で検索する。検索の結果、意味が 1 つであった場合には、その意味を出力することになる。意味を複数有する場合には、次節で述べる意味的な近さに基づく多義性の解消を行うため、それぞれの意味を概念化する。概念化には、3.1.1 節で述べた概念ベースと 3.1.2 節で述べた Web を用いた未定義語の概念化手法を用いる。

一例として、英字略語「IC」を Wikipedia で検索した際の結果を表 2 に示す。英字略語「IC」の場合、11 種類の意味を有していた。

なお、Wikipedia に掲載されている意味の中には、商品の型番やある種のコードなど英字略語ではないもの（表 2 の例では 9 番と 10 番）が多数含まれている。また、英字略語の意味を概念化する際に、不要な情報が付加されていることが多い。そこで、Wikipedia に掲載されている意味候補の中から、表 3 に示した規則を上から順に適用した上で意味候補を取得する。ここで、表 3 におけるストップワードは、表 4 に示した通りである。

表 2: 英字略語「IC」を Wikipedia で検索した際の結果.

<ol style="list-style-type: none"> <li>1. 集積回路 (Integrated Circuit) - 電子機器に用いられる部品。関連：IC カード</li> <li>2. インタークーラー (Inter Cooler)</li> <li>3. インターチェンジ (Inter Change) - 道路交通同士が接続するための合分流構造。</li> <li>4. イメージカラー (Image Color)</li> <li>5. イオンクロマトグラフィー (Ion Chromatography) の略</li> <li>6. インフォームド・コンセント (Informed Consent)</li> <li>7. インターシティ (InterCity) - (特にヨーロッパの) 都市間特別急行列車</li> <li>8. インデックスカタログ - 星団や星雲、銀河を収載した 2 つの星表のこと (Index Catalogue)</li> <li>9. NHK 富山放送局のラジオ第 2 放送・教育テレビのコールサイン (JOIC/JOIC-DTV)</li> <li>10. イリノイ・セントラル鉄道の報告記号 (Illinois Central railroad)</li> <li>11. 間質性膀胱炎 (Interstitial cystitis) の略</li> </ol>
---

表 3: Wikipedia から英字略語の意味候補を取得する規則.

<ul style="list-style-type: none"> <li>● 意味候補にストップワード (表 4 参照) が含まれる場合, 意味候補から除外</li> <li>● 意味候補内の括弧開き ( ( ), 右矢印 ( → ), ハイフン ( - ) より後ろの語を削除</li> <li>● 意味候補内の「など」を削除</li> <li>● 意味候補内の「のこと」, 「の略」, 「の略符」, 「の通称」, 「の愛称」, 「の名称」, 「の英文略称名」より前の語を取得</li> <li>● 意味候補内の「の一つ、」より後ろの語を取得</li> </ul>
---

### 4.3 英字略語の多義性解消

4.2 節で検索した英字略語が複数の意味を有した場合、その多義性を解消する必要がある。具体的には、4.2 節で概念化された意味候補と入力された英字略語を含む文章との意味的な近さを評価することで実現する。この際、概念化された意味候補と英字略語の意味的な近さを評価するため、英字略語も概念化する必要がある。

入力された文章に含まれる自立語をすべて抽出し、それを英字略語の一次属性と見立てる。これにより、英字略語を疑似的に概念化することができる。なお、英字略語の一次属性とした自立語の中には概念ベースに登録されている語と未定義の語が存在する。未定義語については、3.1.2 節で説明した手法によりそれらを概念化する。この処理により、英字略語を概念とし、一次、二次へと属性を展開することができ、意味的な近さを評価することが可能になる。

なお、英字略語を疑似的に概念化する際、概念ベースに登録されていない未定義語である一次属性に対する重みは、3.1.2 節の未定義語の概念化の際の属性への重み付けの考え方と同様にして付与する。

具体的には、語の網羅性である  $tf$  値は、入力された文章  $A$  中に出現する自立語  $Word_A$  の出現頻度  $tfreq(Word_A, A)$  を入力された文章  $A$  中のすべての自立語の語数  $tnum_A$  で割ったもので算出される。算出式は以

下のようになる。

$$tf(Word_A, A) = \frac{tfreq(Word_A, A)}{tnum_A}$$

語の特定性である  $idf$  値については、 $SWeb-idf$  [13, 14] の代わりに  $SA-idf$  値を用いる。これは、疑似的な全文章空間の情報としては、 $tf$  値を算出する際に使用した文章と同じカテゴリ、ジャンルである文章集合を利用する必要があるためである。 $SA-idf$  値の算出式は以下のように定義される。ここで、 $N$  はこの文章集合の全文章数、 $df(Word_A)$  はその文章集合の中で  $Word_A$  が出現する文章数である。

$$SA-idf(Word_A) = \log \frac{N}{df(Word_A)}$$

以上に示した式より、自立語  $Word_A$  へ付与する重み  $w$  は次の式で定義される。

$$w = tf(Word_A, A) \cdot SA-idf(Word_A)$$

このように、概念化された英字略語と意味候補との意味的な近さを 3.2 節で説明した語彙の意味的な近さを評価する手法により評価する。その結果、最も意味的に近いと判断された意味候補を英字略語の意味とする。

表 4: 設定したストップワードのリスト.

型番, 型式, 形式, シリーズ, 略号, 単位, ドメイン, 拡張子, 記号, 符号, 係数, コマンド, 国名コード, 行政区画コード, 県名コード, 郵便コード, 空港コード, 港コード, IATA コード, 航空会社コード, 形式コード, 通貨コード, 言語コード, 作品, 楽曲, 登場, アルバム, コールサイン, 一覧, 上記, その他, 以下
---

## 5 評価実験

### 5.1 実験条件

評価実験のデータとしては、新聞記事から英字略語を無作為に抽出し、当該英字略語が含まれている記事を入力文章とした。使用した記事数は129記事であり、その中で、表記が異なる英字略語の数は58個であった。つまり、129種類の英字略語の意味と58種類の英字略語の表記が含まれる記事の評価実験データとして使用した。また、当該58個の英字略語の表記をWikipediaで検索した結果、859個の意味を取得できた(1つの英字略語の表記につき、平均で14.8個、最少で2個、最多で39個の意味が存在)。提案手法により推定した英字略語の意味が、当該英字略語を含む新聞記事における意味と一致した場合を正答として評価した。

なお、4.3節で述べたとおり、入力した文章を用いて英字略語を疑似的に概念化する際には、3.1.2節での考え方にに基づき属性に重み付けを行う。今回の実験の場合、入力の対象は新聞記事である。そのため、概念ベースに登録されていない未定義語である一次属性に対する重み付けに必要な  $SA-idf$  値の算出には、1か月分の新聞記事集合を疑似的な全文章空間の情報として用いた。

### 5.2 評価結果

英字略語の意味推定結果として、関連度計算を利用した場合に76%、EMDを利用した場合に79%の正答率を得ることができた。意味推定結果の一例を図3に示す。

図3において例1では、英字略語AFCに対して、Wikipediaから取得した13個の意味候補の中から正しい意味を推定できており、英字略語の意味を十分に理解できていることが分かる。

次に例2では、英字略語FSに対して、25個の意味候補の中からEMDでは正しい意味を推定できた一方、関連度計算では意味の推定に失敗している。これは、関連性の高い属性を1対1で対応をとって計算を行うために、他に関連性の高い属性を除外してしまう関連度計算に比べ、全ての属性を計算に利用するEMDが有効に機能したためだと考えられる。

最後に例3では、英字略語FBに対して、15個の意味候補の中から正しい意味を推定することができなかった。これは、入力文章に金銭に関連する語が多く含まれていたため、正しい意味であるフェイスブックより金銭に関連する意味候補である政府短期証券のほうが意味的に近いと判定されたためだと考えられる。

Wikipediaは現存する辞書の中で収録語数が最も多いとされているが、提案手法には、処理の拠り所である辞書に登録されていない単語には対応できないという問題がある。これに関しては、辞書を使用する以上、避けることができない問題である。ただし、評価実験で使用したデータとは異なる100件の新聞記事を無作為に調査した結果、Wikipediaに登録されていない英字略語の出現頻度は約4.5%であった。このことから、新聞記事に登場するような比較的一般的な英字略語を対象とする場合、Wikipediaに未登録の語があることに起因する正答率の低下は5%程度であると考えられる。

## 6 まとめ

本論文では、英単語の頭文字から構成される表現である英字略語に焦点を当て、その意味を推定する手法について提案した。提案手法では、語彙の概念化手法と語彙の意味的な近さを評価できる関連性評価手法を使用する。さらに、Wikipediaを辞書として用いることで、英字略語の多義性を解消し、英字略語の本来の意味を推定することを実現した。

提案手法に対して、129件の新聞記事(英字略語の意味:129種類、英字略語の表記:58種類)を入力文章として評価を行った。評価結果より、関連度計算を使用した場合は76%、EMDを使用した場合は79%と、最高で80%近い正答率を得ることができた。

本論文で提案した技術により、英字略語の意味を理解しやすくすることができ、自然な知的対話の実現や情報検索など多くのアプリケーションの性能向上に寄与することができると思われる。

## 謝辞

本研究の一部は、科学研究費補助金(若手研究(B)24700215)の補助を受けて行った。

例	入力文章	英字略語	英字略語の意味	意味候補の数	評価手法	推定結果	判定
1	… 国際サッカー連盟理事と AFC会長だった2008～11年の間に、 度重なる倫理規定違反が認められたのが理由。 …	AFC	アジア サッカー連盟 (Asian Football Confederation)	13	関連度 計算	アジア サッカー連盟	○
					EMD	アジア サッカー連盟	○
2	… 一方、1Q業績は、 前年同期が四半期FSを作成していないため 前年同期比較はないが、 経常利益は、6億7000万円。 …	FS	財務諸表 (Financial Statements)	25	関連度 計算	Future Systems プロジェクト	×
					EMD	財務諸表	○
3	ソーシャル・ネットワーキング・ サービス世界最大手、米FBの株式上場の際、 情報開示が不公正だったとして、 米マサチューセッツ州の証券監督当局は17日、 上場手続きを取り仕切った幹事社の 米証券大手モルガン・スタンレーに 罰金500万ドル(約4億2000万円)の 支払いを命じた。 …	FB	フェイスブック (FaceBook)	15	関連度 計算	政府短期証券	×
					EMD	政府短期証券	×

図 3: 英字略語の意味推定結果の一例.

## 参考文献

- [1] 渡部広一, 河岡司: 常識的判断のための概念間の関連度評価モデル, 自然言語処理, Vol. 8, No. 2, pp. 39–54 (2001)
- [2] 笠原要, 稲子希望, 加藤恒昭: 単語の属性空間の実現方法, 人工知能学会論文誌, Vol. 17, No. 5, pp. 539–547 (2002)
- [3] 河原大輔, 黒橋禎夫: 格フレーム辞書の漸次的自動構築, 自然言語処理, Vol. 9, No. 5, pp. 93–110 (2005)
- [4] Eriko, Y., Misako, I., Seiji, T., Hirokazu, W.: Computer-generated Conversation Using Newspaper Headline, *Computer Technology and Application*, Vol. 4, No. 8, pp. 387–394 (2013)
- [5] 岡崎直観, 石塚満: 日本語新聞記事からの略語抽出, 第 21 回人工知能学会全国大会論文集, 2G4-4, pp. 1–3 (2007)
- [6] 吉田辰巳, 遠間雄二, 増山繁, 酒井浩之: 可読性の向上を目的とした片仮名表記外来語の換言知識獲得, 電子情報通信学会論文誌, Vol. J88-D-II, No. 7, pp. 1237–1245 (2005)
- [7] 小島一秀, 渡部広一, 河岡司: 連想システムのための概念ベース構成法-属性信頼度の考え方に基づく属性重みの決定, 自然言語処理, Vol. 9, No. 5, pp. 93–110 (2002)
- [8] 広瀬幹規, 渡部広一, 河岡司: 概念間ルールと属性としての出現頻度を考慮した概念ベースの自動精錬手法, 信学技報, NLC2001-93, pp. 109–116 (2002)
- [9] 奥村紀之, 土屋誠司, 渡部広一, 河岡司: 概念間の関連度計算のための大規模概念ベースの構築, 自然言語処理, Vol. 14, No. 5, pp. 41–64 (2007)
- [10] Wikipedia: Wikipedia, <https://jp.wikipedia.org>
- [11] 渡部広一, 奥村紀之, 河岡司: 概念の意味属性と共起情報を用いた関連度計算方式, 自然言語処理, Vol. 13, No. 1, pp. 53–74 (2006)
- [12] 藤江悠五, 渡部広一, 河岡司: 概念ベースと Earth Mover’s Distance を用いた文書検索, 自然言語処理, Vol. 16, No. 3, pp. 25–49 (2009)
- [13] 辻泰希, 渡部広一, 河岡司: www を用いた概念ベースにない新概念およびその属性獲得手法, 第 18 回人工知能学会全国大会論文集, 2D1-02, pp. 1–4 (2004)
- [14] 後藤和人, 土屋誠司, 渡部広一, 河岡司: Web を用いた未知語検索キーワードのシソーラスノードへの割り付け手法, 自然言語処理, Vol. 15, No. 3, pp. 91–113 (2008)
- [15] Google: Google, <http://www.google.co.jp>
- [16] 徳永健伸: 情報検索と言語処理, 東京大学出版会, (1999)
- [17] A.J.Hoffman.: On simple linear programing problems, *In Proc of the Seventh Symposium in Pure Mathematics of the AMS*, pp. 317–327 (1963)

# ソーシャルタグを用いた類似楽曲検索システム

## Content-based music information retrieval using SNS tags

萩原 智彰<sup>1</sup> 山田 誠二<sup>1,2,3</sup>  
Tomoaki HAGIHARA<sup>1</sup> Seiji YAMADA<sup>1,2</sup>

<sup>1</sup> 東京工業大学

<sup>1</sup> Tokyo Institute of Technology

<sup>2</sup> 国立情報学研究所/総合研究大学院大学

<sup>2</sup> National Institute for Informatics/SOUKENDAI

**Abstract:** Recently, increasing digital music collections make us difficult to find new favorite music which we are really interested in. To overcome this problem, content-based music information retrieval has been studied. It enables users to search similar music with users choose, however few method can improve search result interactively. In this paper, we propose a new method that allows users to reflect their search intention. By selecting SNS tags, users can naturally inform their intention to a music retrieval system, and the system utilize user feedback to modify a similarity function. we employ learning to rank technique to utilize tag information to modify a feature vector space.

## 1 序論

### 1.1 研究背景

近年、定額楽曲配信サービスの登場などにより、個人で取り扱える楽曲の量は爆発的に増えているが、ユーザが自分の好む楽曲を見つけることはますます困難になっている。ユーザの楽曲発見を支援するためにいくつかの技術が研究されている。その中でもっとも一般的な手法はユーザの購買履歴を元に推薦を行うものである。このような手法は協調フィルタリングと呼ばれ、amazon.comなどのサービスで実際に利用されている[神島 2007]。しかしながら協調フィルタリングをベースとした推薦システムには、購買履歴が十分な量集まらないと適切な推薦が行えないなどいくつかの問題点が指摘されている[神島 2007]。また、コンテンツの内容を一切考慮せずに推薦楽曲を決定するため、楽曲の特性が推薦結果に反映されないという問題がある。

一方、楽曲の音響信号を利用してユーザの楽曲発見を支援する手法として類似楽曲検索が挙げられる[Casey 2008]。この技術を用いることで、ユーザは楽曲を指定してその楽曲に類似した曲を獲得することができる。類似楽曲検索では、楽曲の音響信号をもとに複数の特徴量を抽出し、特徴ベクトルを作成する。楽曲と楽曲の類似の度合いは、特徴ベクトル同士の距離を事前に定めた類似度関数に基づいて定義する。しかしながらユーザの検索意図は多様であるため、あるひとつの類似度関

数を利用するだけでは、ユーザの求める検索結果が得られない場合がある。そのため、類似度関数を各ユーザごとにパーソナライズする手法が求められる。

本研究では、SNS上のタグから作成したランキングを教師データとする学習手法を利用し、インタラクティブに検索結果の改善を行う類似楽曲検索手法を提案する。提案手法により、ユーザは表示されたタグの中から自分の検索意図を反映したタグを選択することによって検索結果の改善を行えるようになる。

### 1.2 本研究の貢献

本研究では、タグから得られるランキングが、音響信号から得られるランキングよりも優れていると仮定し、ランキング学習手法によって、優れたランキングに検索結果が近づくように学習を行う。

類似楽曲検索システムの検索結果改善をランキング学習問題として定式化した研究はMcFeeらによって既に行われている[McFee 2010]。しかしながらこの研究は事前に学習を行っている。それに対して本研究ではユーザにクエリが与えられてから学習するという、より難しい課題に取り組み、実時間での応答で検索結果改善が可能であることを示した。また、学習の効果を、計算機実験、参加者実験の両方で確認しており、関連分野の研究者にとって有益な知見を提供するものである。

## 2 関連研究

本章では、まず本研究で取り組む課題と似た目的をもつ研究を紹介する。その後、次節で提案手法に取り入れられている手法について解説する。

### 2.1 似た課題に取り組んでいる研究

類似楽曲検索に関する研究の多くは楽曲の音響信号からの特徴抽出手法の改善を行っている。楽曲間の類似の度合いの良し悪しは特徴抽出がどの程度うまくいっているかに強く依存するためである。しかしながらそのようなアプローチに限界があることが Aucouturier らの研究 [Aucouturier 2004] によって示されている。この研究は当時利用されていた特徴抽出手法とそれらに必要なパラメータについて大規模な比較調査を行ったものである。この研究によって、いかにパラメータの調整を注意深く行ったとしてもそれ以上性能の向上しない限界点があること、また、どの楽曲にも似ていないと判断される曲やどの楽曲とも似ていると判断される曲が出てきてしまうことを発見した。今日、この問題は hubs and orphans 問題として知られている。このような限界を超えるために特徴抽出とは別の部分で問題の解決に取り組む研究が行われている。このような方向性の研究にはユーザのフィードバックを利用するものや、ソーシャルタグのような楽曲に付随する情報を利用するものがある。

斎藤らの研究 [斎藤 2011] では、楽曲の集合を可視化した上でユーザに表示し、適合する楽曲を選択してもらい、対話的遺伝的アルゴリズム手法を利用して次に表示する楽曲の選択を行っていくことでユーザの新たな楽曲発見の支援を試みている。また、類似楽曲検索に対して検索結果の個人化を試みた研究として小林らの研究 [?] があげられる。小林らは楽曲と楽曲の類似度をマハラノビス距離の形式で定義し、行列の値をパラメータとし、非適合とされる楽曲が適合とされる楽曲よりも類似度が低くなるよう学習を行うことにより改善を行う。しかしながらこれらの手法はユーザが適合不適合の情報を与える必要がある。実際に曲を聞いて評価しなければならぬためユーザの負担が大きくなる。

類似楽曲検索の分野においては、画像検索と異なりユーザが検索結果を評価するためには実際に楽曲を耳にする必要があるためフィードバックを与えることがより困難である。そこで SNS 上で得られるソーシャルタグを利用することが検討されている。Karydis らの研究 [Karydis 2013] ではタグを元に抽出した特徴量と、音響信号を元にした特徴量とで性能の比較をし、タグを元にした特徴量の方がより性能が高いことを示した。しかし、ソーシャルタグはすべての楽曲に付けられて

いるものではないので、タグのない楽曲に対してもタグによる検索を実現するような技術が必要となる。そのような技術として楽曲から自動的にタグを推定する研究が行われている [Turnbull 2008]。また、タグから得られる特徴量の特徴空間と、音響信号から得られる特徴空間との間を対応付ける関数を学習する手法が提案されている [Ioannis 2011]。しかしこれらの手法は、検索結果として得られるランキングそのものを最適化しているわけではない。ランキングそのものに着目して学習を行う手法は McFee らによって提案されており [McFee 2010]、本研究ともっとも関連性が強い。しかしながら先述の通り、実時間で学習を行っている、参加者実験による性能確認を行っているなどの違いがある。

### 2.2 提案手法に取り入れられている手法

#### 2.2.1 楽曲の特徴抽出

類似楽曲検索では、楽曲を特徴ベクトルで表し、特徴ベクトル間のユークリッド距離の大小で類似の程度を評価する。そのため、どのような特徴量を用いるかは検索結果に大きな影響を与える。楽曲や音声の特徴量としては MFCC (Mel Frequency Cepstrum Coefficient) が広く用いられている。これは人間の聴覚特性に合わせた各周波数成分のスペクトル密度である。MFCC は短時間の信号から抽出するものなので、楽曲から MFCC を抽出する場合、楽曲の時間に比例した複数個の組が出力される。この組をどのようにまとめるかで様々な手法が提案されている。

本研究では Levy らによって提案された手法 [Levy 2006] を採用した。この手法では二つの異なる楽曲間の距離を次式のような形で表す。

$$D(p, q) = (\vec{\mu}_p - \vec{\mu}_q)^T \Sigma_{\mu}^{-1} (\vec{\mu}_p - \vec{\mu}_q) + (\vec{s}_p - \vec{s}_q)^T \Sigma_s^{-1} (\vec{s}_p - \vec{s}_q) \quad (1)$$

ここで、 $\vec{\mu}_p$ ,  $\vec{\mu}_q$  はそれぞれ楽曲  $p$ ,  $q$  の MFCC の各成分の平均ベクトルであり、 $\vec{s}_p$ ,  $\vec{s}_q$  は各成分の標準偏差である。 $\Sigma_{\mu}$ ,  $\Sigma_s$  はデータセットの楽曲を表す特徴ベクトル  $\vec{\mu}$ ,  $\vec{s}$  の共分散行列である。この手法は他の手法に比べ、特徴量の抽出、比較が高速に行えるという性質がある。我々はこのような性質が検索システムに望ましいと考え、この手法の採用を決断した。

#### 2.2.2 タグの特徴抽出

本研究ではタグの情報を利用する。タグは、各タグを一つの単語とみなすことで Bag-of-Words のベクトルに変換できる。Bag-of-Words 表現のベクトルに対す

る重み付けとして TF-IDF がもっとも利用されているが本研究では IDF のみを各タグに対する重み付けとして採用した。本研究では、楽曲はその人気によってタグが付けられた回数の差が非常に大きく、 $tf$  値が大きいタグほどその楽曲の特徴を適切に反映しているとはいえないためである。

female vocalist と female vocalists のように、タグの中には意味的に非常に近いものがある。前述した Bag-of-Words の表現では、これは完全に別々の単語としてみなされてしまう。単語と単語の共起関係を考慮することでこれらの単語の間関係を考慮することができればより望ましい検索結果を導く事ができる。そのような要求に答える技術が Latent semantic Indexing(LSI)[Scott 1988] である。LSI では、高頻度で共起する単語が同一の要素に対応するよう特徴ベクトルの次元削減を行う。機械学習の分野で特徴ベクトルの次元を削減する場合、もっともよく用いられる手法が主成分分析である。しかしながら主成分分析は固有値分解を伴うため、単語-文書行列のような非正方行列には適用できない。そこで LSI では主成分分析の代わりに特異値分解を行うことで次元削減を実現する。

文書数が  $N$  個で、それらの文書に含まれる単語の種類が  $M$  種類であるとき、その単語-文書行列  $C$  は大きさが  $M \times N$  となる。この行列では一つの列が一つの文書を表す。このとき  $C$  は次式 (2) のように分解できる。

$$C = U\Sigma V^T \quad (2)$$

ここで  $U$  は  $CC^T$  の固有ベクトルを、対応する固有値が大きい順に並べた行列である。 $V$  は  $C^TC$  の固有ベクトルを同様に並べた行列である。このとき  $\Sigma$  は  $(i, i)$  成分が、行列  $C$  の  $i$  番目に大きい特異値であり、そのほかの成分がすべて 0 となる対角行列である。行列  $C$  が式 (2) のように分解可能であるとき、 $k$  次元の低階数近似行列  $C_k$  は式 (3) で得られる。

$$C_k = U\Sigma_k V^T \quad (3)$$

ここで  $\Sigma_k$  は  $\Sigma$  の  $j > k$  であるすべての  $(j, j)$  成分を 0 に置き換えた行列である。

本研究では  $k = 200$  として利用した。

### 2.2.3 ランキング学習

特徴ベクトルで表されたデータセットとクエリについて、理想的なランキングが明らかであるときに検索結果をその理想的なランキングに近づけようとする手法としてランキング学習と呼ばれるものがある [Li 2009]。これは何らかの方法で理想的なランキングを機械学習の教師データとして扱えるよう変換し、教師あり学習手法で検索結果の改善を行う。理想的なランキングを

どのように教師データとして利用するかで、いくつかの枠組みが存在する。本論文では、分類手法を用いる枠組みを採用したので、それについて説明する。

あるデータの特徴ベクトルを  $\vec{v}$  とする。そして、あるデータの評価値が  $\vec{w}^T \vec{v}$  のような特徴ベクトルの重みつき和で表されるとする。ここで  $\vec{w}$  は重みベクトルを表す。ここで理想的なランキングにおいて順位が  $i$  番目のデータの特徴ベクトルを  $\vec{v}_i$ 、 $j$  番目の特徴ベクトルを  $\vec{v}_j$  とすると  $i < j$  である  $i, j$  の組に対して次式の制約を可能な限り満たす重みベクトルを学習する。

$$\vec{w}^T \vec{v}_i > \vec{w}^T \vec{v}_j \quad (4)$$

この重みを学習する問題は  $\vec{w}^T (\vec{v}_i - \vec{v}_j) > 0$  と変換すると二値分類問題と考えることができる。重みの学習には分類問題に用いられるパーセプトロンや SVM のような手法が利用可能である。本研究ではパラメータを調整する必要がない Passive Aggressive(PA) 法 [Koby 2006] を中心に使用するアルゴリズムの検討を行った。法については次節で学習手法を紹介する際、詳しく解説する。

## 3 提案手法

### 3.1 提案手法の利用を想定している局面

提案手法がどのような局面での利用を想定しているのかを述べる。提案手法は、ユーザが検索したいクエリ楽曲と、クエリ楽曲に対応するタグの情報が入力として必要になる。タグの情報と音響信号の情報を比較したとき、どちらの方がより類似楽曲検索の役に立つかは先行研究で調査されている [?][Ioannis 2011]。これらの研究はタグの情報の方がより精度の良い検索結果をもたらすことを示している。もし検索対象のデータセットに、楽曲そのものだけでなくタグも含まれるのであれば、そのタグを用いて検索することができる。しかしながら、タグの情報は SNS 等から得られるごく一部の楽曲にしかつけられていない。そのため、データセットを、楽曲の音響信号のみから検索しなければならない状況が発生する。本研究はそのような状況で利用される。つまり、検索したい対象となるデータは楽曲のみしか用意できないが、別の楽曲の集まりについてはタグと楽曲の両方を用意できるという状況である。提案手法では、楽曲の音響信号をから得られる特徴ベクトルを元に類似度を算出する際、ある重みパラメータを用いて類似度の算出の仕方を修正することで検索結果の改善を実現する。次節以降でシステムの詳細を記述する。

### 3.2 教師データ

提案手法ではランキング学習の手法を用いて類似度関数の重みを調整する。ランキング学習を行うためには教師データとなる理想的なランキングが必要となる。本手法では楽曲につけられているタグをもとに得られる検索結果を教師データとして利用した。タグから類似楽曲を得る手続きについて説明する。まず、事前にデータセット中の楽曲について、そのタグをもとに Bag-of-Words で表現されたベクトルを作成し、関連研究の項で解説した IDF による重み付け、LSI による次元圧縮を行って 200 次元の特徴ベクトル  $\vec{f}$  に変換した。そしてこのベクトルを記録した。検索する際は、まず、クエリとして与えられたタグに対し同様の処理を行ってベクトル  $\vec{f}_q$  を作成する。その後コサイン類似度でデータセット中の楽曲との類似度を算出する。

コサイン類似度は大きな値であれば大きな値であるほどより類似していることを表す指標であるので、コサイン類似度の値が大きい順にソートすることによって、類似した楽曲のランキングを作成することができる。

### 3.3 学習手法

教師データとなるランキングが得られたならば次の課題は重みの学習である。本研究ではオンライン学習手法の一つである Passive-Aggressive 法 [Koby 2006] を改良した手法で重みの学習を行った。この改良法の紹介をする前、まず PA 法を本研究に応用するときどのようにすべきかの解説を行う。次に改良法との変更点を示す。

#### 3.3.1 Passive-Aggressive

クエリとなる楽曲を  $q$ 、タグをもとに検索して得られる理想的なランキングにおいてある順位である楽曲をその順位を表す記号  $i, j$  を用いて表現する。今、 $i < j$  であるとき、検索結果を理想的なランキングに近づけるためには式 (5) の制約ができる限り満たされるように重み  $\vec{w}$  の学習を行えばよい。

$$D_n = \vec{w}^T \vec{v}_j > D_p + 1 = \vec{w}^T \vec{v}_i + 1 \quad (5)$$

ここで  $D_n$  は楽曲  $j$  と楽曲  $q$  の距離を表し、 $D_p$  は楽曲  $i$  と楽曲  $q$  の間の距離を表す。クエリ  $q$  との距離が小さい楽曲ほどより似ていると判断される。 $\vec{v}_p, \vec{v}_n$  はそれぞれクエリとの差の大きさを表すベクトルであり、楽曲の音響信号より得られる特徴ベクトル  $\vec{x}_i, \vec{x}_j$  を用いて次式 (6)(7) で定義される。

$$\vec{v}_p = (\vec{x}_i - \vec{x}_q) \circ (\vec{x}_i - \vec{x}_q) \quad (6)$$

$$\vec{v}_n = (\vec{x}_j - \vec{x}_q) \circ (\vec{x}_j - \vec{x}_q) \quad (7)$$

$\circ$  はアダマール積を表し、成分ごとに積を取る処理を表す。制約 (5) は、より上位にランキングされるべき楽曲  $i$  と  $q$  との間の距離が、より下位にランキングされるべき楽曲  $j$  と  $q$  との間の距離よりも 1 以上小さくなることを意味している。

ある楽曲  $i, j$  に対してこの制約を満たす重みは無数にある。しかしながら、制約が満たされる中で、重みの更新量を最小にしようとするとき、その重みの値はただ一つに定まる。このことは次の制約付き最小化問題 (8) を解くことで求められる。

minimize

$$O(\vec{w}) = \frac{1}{2} |\vec{w} - \vec{w}_t|^2 \quad (8a)$$

subject to

$$l_t(\vec{w}, \vec{v}_p, \vec{v}_n) = 0 \quad (8b)$$

ここで  $l_t$  は式 (9) で定義される。

$$l_t(\vec{w}, \vec{v}_p, \vec{v}_n) = \begin{cases} 0 & (\text{if } D_n - D_p \geq 1), \\ 1 - (D_n - D_p) & (\text{otherwise}), \end{cases} \quad (9)$$

このとき最適化問題を解くと次の更新式が得られる。

$$\vec{w}_{t+1} = \vec{w}_t + \tau_t (\vec{v}_n - \vec{v}_p) \quad \text{where } \tau_t = \frac{l_t}{|v_n - v_p|^2} \quad (10)$$

以上のアルゴリズムをまとめると Algorithm 1 となる。ここで  $d$  は特徴ベクトルの次元数を表す。ones( $d$ ) は値がすべて 1 である  $d$  次元の列ベクトルを出力する関数である。

#### 3.3.2 提案手法

前節で説明したアルゴリズムを本研究に適用するといくつかの問題が生じる。それぞれの問題とその解決策を次に示す。

まず、もっとも大きな問題点は計算時間である。楽曲が  $N$  個あるデータセットから順序を考慮せず 2 つ取り出した組み合わせは  $\frac{N(N-1)}{2}$  通りとなる。この組み合わせすべてを学習に利用すると計算時間のオーダーは  $O(N^2)$  となり実時間での応答が難しくなる。これは検索システムとしてふさわしくない。このような問題に対処するために提案手法では学習に利用する楽曲を、ランキング上位  $k$  個と下位  $m$  から得られる組に限定した。これはランキングで順位がそれほど変わらない楽



---

**Algorithm 1** Passive Aggressive
 

---

**Require:**  $\vec{x}_q, R$

```

 $\vec{w} = \text{ones}(d)$ 
for  $i = 0 : N - 1$  do
  for  $j = i : N - 1$  do
     $\vec{v}_p = (\vec{x}_i - \vec{x}_q) \circ (\vec{x}_i - \vec{x}_q)$ 
     $\vec{v}_n = (\vec{x}_j - \vec{x}_q) \circ (\vec{x}_j - \vec{x}_q)$ 
     $D_p = \vec{w}^T \vec{v}_p$ 
     $D_n = \vec{w}^T \vec{v}_n$ 
    if  $D_n - D_p < 1$  then
       $l_t = 1 - (D_n - D_p)$ 
       $\vec{v} = (\vec{v}_n - \vec{v}_p)$ 
       $s = \vec{v}^T \vec{v}$ 
       $\vec{\Delta} = \frac{l_t}{s} \vec{v}$ 
       $\vec{w} = \vec{w} + \vec{\Delta}$ 
    end if
  end for
end for

```

**Ensure:**  $\vec{w}$

---

曲同士を比べても汎化性能の向上に繋がる適切な学習が行えない、という考えに基づいている。このように学習データを限定することで学習にかかる計算時間は  $O(k \cdot m)$  となる。このように学習データを制限する手続きを以後 Algorithm 2 とする。本研究では  $k = 50$ ,  $m = 200$  と設定した。

二つ目の問題点は汎化誤差に関するものである。PA法は与えられたデータに対して必ず正常に分類できるように学習するものなので、はずれ値が与えられると、それを正常に分類できるように学習することによって汎化誤差が悪化する可能性がある。この問題に対処するために、学習によって重みが更新されるたびにその値を保存し、出力の際にそれらの成分ごとの平均を取るという処理を考案した。これは [Carvalho 2006] の手法から着想を得たものである。このような手続きを以後 Algorithm 3 とする。

三つ目の問題点は、特徴量の次元数が低いと調整できる重みが限られるというものである。この問題に対処するため、特徴量をそのまま利用するのではなく、代わりに各特徴量の積を利用するという処理を考案した。これは SVM の多項式カーネルから着想を得たものである。この工夫を導入するとき、解くべき最適化問題は次のようになる。

minimize

$$O(\mathbf{W}) = \frac{1}{2} |\mathbf{W} - \mathbf{W}_t|^2 \quad (11a)$$

subject to

$$l_t(\mathbf{W}, \vec{v}_p, \vec{v}_n) = 0 \quad (11b)$$

ここで  $l_t$  は式 (9) であるが、クエリとの距離  $D_p$ ,  $D_n$  の定義が異なり、式 (12)(13) のように定義される。

$$\vec{v}_p = \text{abs}(\vec{x}_p - \vec{x}_q) \quad (12a)$$

$$D_p = \text{sum}(\mathbf{W} \circ \vec{v}_p \vec{v}_p^T) \quad (12b)$$

$$\vec{v}_n = \text{abs}(\vec{x}_n - \vec{x}_q) \quad (13a)$$

$$D_n = \text{sum}(\mathbf{W} \circ \vec{v}_n \vec{v}_n^T) \quad (13b)$$

この工夫を取り入れた手法を Algorithm 4 とする。また、考案したすべての工夫を取り入れた場合、手続きは Algorithm 5 となる。ここで  $S_p^k$  は LSI より得られるランキングの上位  $k$  個の楽曲を集めた集合、 $S_n^m$  は下位  $m$  個の楽曲を集めた集合を表し、 $S_w.append(\vec{w})$  は配列  $S_w$  に  $\vec{w}$  を追加する処理、 $\text{mean}(S_w)$  は  $S_w$  に含まれる重みベクトル  $\vec{w}$  の各成分の平均を取る処理を表す。 $\text{diag}(d)$  は  $d$  次元で対角成分がすべて 1 である対角行列を出力する関数を表す。 $\mathbf{T}_r$  は非零成分が 1 である大きさ  $(d \times d)$  の上三角行列であり、 $\mathbf{V}$  の非対角成分が二重に足し合わされるのを防ぐ役割を担っている。

---

**Algorithm 5** Proposed Method
 

---

**Require:**  $\vec{x}_q, S_p^k, S_n^m$

```

 $\mathbf{W} = \text{diag}(d)$ 
 $S_w = []$ 
for  $\vec{x}_p$  in  $S_p^k$  do
  for  $\vec{x}_n$  in  $S_n^m$  do
     $\vec{v}_p = \text{abs}(\vec{x}_p - \vec{x}_q)$ 
     $\vec{v}_n = \text{abs}(\vec{x}_n - \vec{x}_q)$ 
     $D_p = \text{sum}(\mathbf{W} \circ \vec{v}_p \vec{v}_p^T)$ 
     $D_n = \text{sum}(\mathbf{W} \circ \vec{v}_n \vec{v}_n^T)$ 
    if  $D_n - D_p < 1$  then
       $l_t = 1 - (D_n - D_p)$ 
       $\mathbf{V} = (\vec{v}_n \vec{v}_n^T - \vec{v}_p \vec{v}_p^T) \circ \mathbf{T}_r$ 
       $s = \text{sum}(\mathbf{V} \circ \mathbf{V})$ 
       $\mathbf{\Delta} = \frac{l_t}{s} \mathbf{V}$ 
       $\mathbf{W} = \mathbf{W} + \mathbf{\Delta}$ 
       $S_w.append(\mathbf{W})$ 
    end if
  end for
end for
 $\mathbf{W}_{output} = \text{mean}(S_w)$ 

```

**Ensure:**  $\mathbf{W}_{output}$

---

我々はこれら Algorithm 2, 3, 4, 5 について、それぞれ計算機実験により汎化性能を確認した。また、もっともすぐれた性能を示したアルゴリズムを用いて参加者実験を行った。実験の詳細について次節に示す。

## 4 実験

提案手法の性能を確認するためにふたつの実験を行った。一つ目の実験は学習による汎化性能を確認する計算機実験である。そしてもう一つは実際にユーザの感性和照らし合わせても望ましい検索結果が得られることを確認する参加者実験である。まず、実験に利用したデータセットの構築について述べ、その後それぞれの実験の詳細を記述する。

### 4.1 実験に用いるデータ

本研究の実験は参加者による評価を含む実験であるため、楽曲のタグ、特徴ベクトルのほかに、楽曲そのものが必要となる。そのようなデータセットは筆者の知る限りでは存在しない。そこで、我々は本実験のために last.fm<sup>1</sup>、7digital<sup>2</sup> の二つの API を利用してデータセットの構築を行う。last.fm は音楽に特化した SNS であり、サービスを通して得られた楽曲のメタデータを公開している。本実験では API を利用し、楽曲 3046 曲のメタデータを獲得した。このメタデータ内にはタイトル、アーティスト名などの一般的な情報のほかに各楽曲につけられたタグの情報を含んでいる。タグそれぞれには、そのタグが対応する楽曲に対して何度付けられたかを示す数が付けられている。本研究では、不正確なタグを除外するために、一つの楽曲に対して前述した数が大きいものから上位 20 個のタグのみを採用した。データに含まれる楽曲の再生可能な音響信号情報を獲得するために 7digital で公開されている API を利用した。7digital API を利用することで各楽曲ごとに約 30 秒の長さの試聴用データを入手できる。本実験ではこの試聴用データをもとに、特徴抽出および参加者実験を行う。特徴抽出は先述のとおり先行研究 [Levy 2006] で用いられている手法を利用した。

### 4.2 計算機実験

一つ目の実験では、テストデータを学習データとテストデータの半分に分け、汎化性能を確認した。クエリとしては、データセット中の楽曲とタグを利用した。学習の際は、学習データについて LSI より得られるランキングの上位 50 曲と下位 200 曲から得られる順序関係の組み合わせ 10000 個を教師データとして採用した。評価指標としては、テストデータの上位 50 曲と下位 200 曲から得られる順序関係の何割が満たされているかで評価した。はじめの 50 個のサンプルで学習時間を計測したところ平均学習時間は 0.502 秒だった。実験

のサンプル数は 3046 である。実験結果を図 1 に示す。Baseline は学習なしで検索した結果を表している。図 1 のエラーバーは標準誤差を表している。図から Alg 5 がもっとも汎化性能において優れていることが明らかになった。

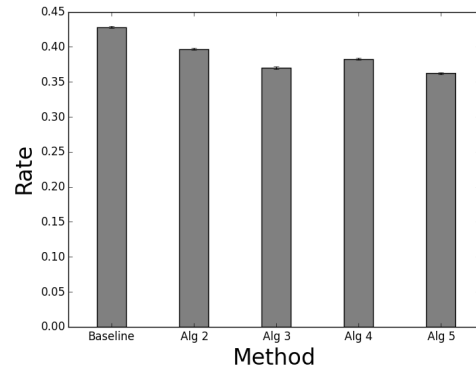


図 1: 計算機実験の結果

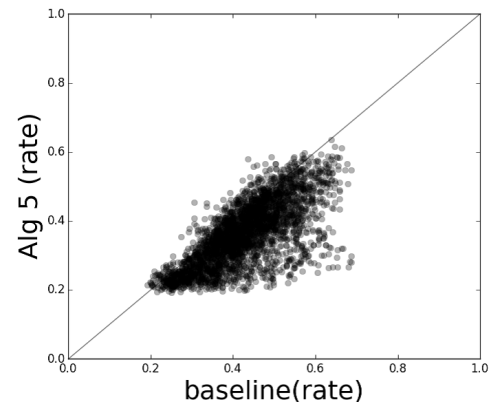


図 2: Baseline との比較

統計量だけでなく、一曲単位で見たとき、Baseline に比べ提案手法が優れているかは図 2 で確認できる。図 2 は学習の前後での、破られる制約の割合をプロットしたものである。左下から右上にの線は境界線を表しており、この境界線より右下にプロットされているならば、学習後の方が破られる制約の割合が低いこと、すなわち学習によってより望ましいランキングに近づいていることを示している。図 2 より多くのクエリで、学習後の結果が改善していることが明らかになった。

<sup>1</sup>The Last.fm Dataset — Million Song Dataset <http://labrosa.ee.columbia.edu/millionsong/lastfm>

<sup>2</sup>7digital Developers <http://developer.7digital.com/>

### 4.3 参加者実験

#### 4.3.1 実験手続き

提案手法に、ユーザに実際に聞いて望ましいと感じる結果が得られていることを確認するため参加者実験を行った。実験参加者は男性 15 名 (平均年齢 25, 最小 23, 最大 32, 標準偏差 2.68) である。また、参加者には実験報酬として 1500 円相当の筆記用具を贈与した。実験の手続きを説明する。参加者には各々が好む楽曲を用意してもらった。実験ではその楽曲をシステムに入力してもらい、事前にこちらが用意したタグの中から、楽曲を表すのにふさわしいと思うものを任意の数 (ただし一つ以上) 指定してもらった。用意したタグを表 1 に示す。これらのタグは実験者が楽曲を選ぶのに役に立つと判断して決定したものである。

表 1: 選択肢として提示したタグ

Tag	
instrumental	relax
rock	classic
pop	male vocalist
jazz	female vocalist
Ballad	beautiful
metal	guitar
alternative	piano
electronic	sad
dance	happy

システムにタグと楽曲が入力されると、学習なしで検索した結果 (Baseline) と、Algorithm 5 で学習をした検索結果からそれぞれ上位 5 曲が無作為な順番で図 3 のように表示される。楽曲名やアーティスト名のようなメタ情報にユーザが惑わされず、楽曲そのものについて評価できるようにするため、検索された楽曲に関する情報は表示されない。ユーザは UI を通して、楽曲の再生、一時停止、停止を行うことができる。検索結果はシャッフルされてから表示されるので、提案手法より得られる楽曲を高く評価するといった恣意的な評価を行うことはできないようになっている。

参加者にはそれぞれの楽曲を聞いてもらい、その楽曲が検索結果としてどの程度ふさわしいかを評価してもらった。評価値は 0 から 1 の間の得点を 0.1 点刻みで付けてもらった。参加者より得られた評価値を元に、学習前後の結果に対して、ランキングの評価指標である nDCG(normalized Discount Cumulative Gain) [Jarvelin 2000] を算出した。

nDCG は情報検索の分野で用いられる指標で検索結果のランキングが理想的なランキングにどれぐらい近

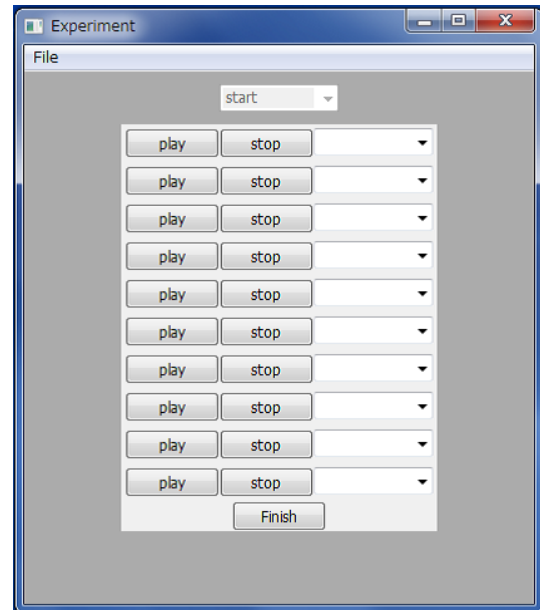


図 3: システムのユーザーインターフェース

いかをあらわす指標である。nDCG には、ユーザのアイテムに対する選好順序の情報を必要とせず、各アイテムについての評価さえ得られれば算出できるというメリットがある。本研究では、楽曲集合に対して選好の順序を記述させるよりも、各楽曲をそれぞれ評価する方がユーザにとって負担が少ないと考え、評価指標として nDCG を採用した。

#### 4.3.2 実験結果

実験結果を図 4 に示す。図 4 のエラーバーは標準誤差を表す。図より、参加者が提案手法から得られる楽曲をより高く評価する傾向にあることが確認された。

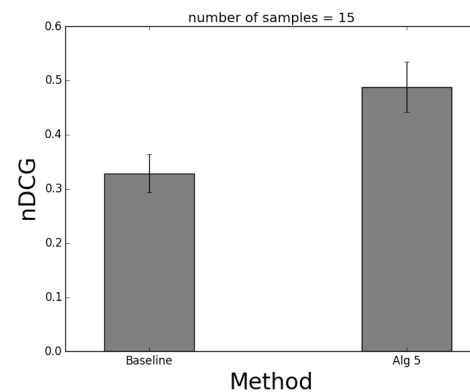


図 4: 参加者実験結果 棒グラフ

## 5 まとめと今後の課題

本論文では、類似楽曲検索において、ユーザがインタラクティブに検索結果の改善を行える方法を考案した。計算機実験の結果、学習によって重みを修正することで、出力をタグより得られるランキングに近づけることができることが確認された。また、参加者実験によって、このように学習して得られるランキングは、学習前のランキングに比べ、ユーザにとってより好ましいものであることがわかった。

一方提案手法は実際にユーザからクエリを受け取ってから学習を行うために、応答に時間がかかるという問題が残っている。また、本研究ではあらかじめ実験者が選択すべきタグの候補を決定していたが、実際にどのタグを用いると検索結果に大きな影響を与えるのかは未調査である。さらに有益なシステムのためには、今後このような課題に取り組むことが求められる。

## 参考文献

- [神寫 2007] 神寫 敏弘 : 推薦システムのアルゴリズム (1), 人工知能学会誌 22 巻 6 号 (2007)
- [Casey 2008] Casey, Michael A., et al. : Content-based music information retrieval: Current directions and future challenges, Proceedings of the IEEE 96.4 668-696 (2008)
- [Li 2009] Li, Hang. : Learning to rank, Tutorial given at ACL-IJCNLP, August. research.microsoft.com/enus/people/hangli/li-acl-ijcnlp-2009-tutorial.pdf (2009).
- [Jarvelin 2000] Jarvelin, K., Kekalainen, J. : IR evaluation methods for retrieving highly relevant documents. In Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval (pp. 41-48). ACM. (2000)
- [斉藤 2011] 斉藤 優理 伊藤 貴之 : 特徴量空間における対話型進化計算を用いた楽曲提示インタフェース, 2011 年度人工知能学会全国大会 (2011)
- [Levy 2006] Mark Levy, Mark Sandler : Lightweight measures for timbral similarity of musical audio, AMCM '06 Proceedings of the 1st ACM workshop on Audio and music computing multimedia Pages 27-36
- [Ioannis 2011] Karydis, Ioannis ; Nanopoulos, Alexandros : Audio-to-Tag mapping: A novel approach for music similarity computation, Multimedia and Expo (ICME), 2011 IEEE International Conference on
- [Carvalho 2006] Vitor R. Carvalho, William W. Cohen : Single-Pass Online Learning: Performance, Voting Schemes and Online Feature Selection, KDD '06 Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining Pages 548-553
- [Koby 2006] Koby Crammer, Ofer Dekel, Joseph Keshet, Shai Shalev-Shwartz, Yoram Singer : Online Passive-Aggressive Algorithms, Journal The Journal of Machine Learning Research archive Volume 7, 12/1/2006 Pages 551-585
- [Scott 1988] Scott Deerwester and Susan T. Dumais and George W. Furnas and Thomas K. Landauer and Richard Harshman : Indexing by latent semantic analysis, JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE, Volume 41, Issue 6, pages 391407, September 1990
- [McFee 2010] Brian McFee, Luke Barrington, Gert Lanckriet : LEARNING SIMILARITY FROM COLLABORATIVE FILTERS, 2010 International Society for Music Information Retrieval.
- [Karydis 2013] Ioannis Karydis, Katia Lida Kermanidis, Spyros Sioutas, Lazaros Iliadis : Comparing content and context based similarity for musical data, Neurocomputing 107 (2013) 6976
- [Aucouturier 2004] Jean-Julien Aucouturier, Francois Pachet : Improving Timbre Similarity: How high's the sky? : Journal of Negative Results in Speech and Audio Sciences (2004)
- [Turnbull 2008] Douglas Turnbull, Luke Barrington, David Torres, and Gert Lanckriet : Semantic Annotation and Retrieval of Music and Sound Effects, IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 16, NO. 2, FEBRUARY 2008

# 類推を用いた数学の学習支援システム

## Learning Support System for Mathematics by using Analogy

草田 圭輔\* 砂山 渡  
Keisuke Kusada Wataru Sunayama

広島市立大学大学院情報科学研究科  
Graduate School of Information Sciences, Hiroshihma City University

**Abstract:** And the knowledge of one field based on, in order to analogize the knowledge of other fields, it is necessary to understand the nature to be a common point of both the knowledge . Even if only the formula is remembered earnestly , and there is a formula of the same meaning of the plural by solved learning method without also understanding kind in learning of mathematics , it'll be the circumstances which can't be applied because those are remembered as the separate formula . In this study , problem of different fields with a common essence , we propose a learning system to support the analogy of both fields on an understanding of the essence .

### 1 はじめに

小学校, 中学校, 高校の児童や生徒が, 計算によって解答を導き出す科目や分野の問題解決過程には, 3 ステップある. 問題の解決方法を説明する「考え方」, 考え方に基づいて問題を解けるようにする「問題の定式化」, そして定式化により与えられる式に基づいて答えを求める「計算」がある. しかし, 「問題の定式化」, つまり, 「公式」を活用して「計算」を行うことによって, 試験で得点が取れるため, 学生のレベルに応じてはこの本質部分が省略されることも少なくない. 生徒達も「公式」のみを覚えれば得点がとれることを知っているため, わざわざ考え方を学ぶ必要がないと考えている. 教師も問題解決過程を全て教えていては, 授業時間が足りなくなる. そのため, 「考え方」に十分時間をかけて教えることができない. しかし, この学習方法では公式やパターンを暗記する必要があることに加え, 類似する他の問題への応用を考えることができなくなる欠点が生じる.

最近までの国内での教育課程実施状況調査や国際的な学力調査の結果分析によると, 計算などの技能の定着について低下傾向は見られないが, 計算の意味を理解することなどに課題が見られる. また, 身に付けた知識や技能を生活や学習に活用することが十分でないといった状況が見られる [1].

算数, 数学の学習目的として, 身に付けたものを生活や他教科等の学習, より進んだ算数・数学の学習へ活用することが重要だといわれている [1]. これより, 学

習した内容を他の問題へと活用することは重要だとわかる. すなわち, 上記で述べた欠点を改善することは重要であるとわかる.

また, 最近では学校の授業において ICT の活用が増えてきている. 平成 26 年に全国の公立学校を対象にアンケートを行ったところ, タブレット端末の台数が前年より 2 倍に増加し, 7 万台以上にも増加した [2]. 今では, ICT を活用して授業を行うことは珍しくなくなってきたため, 数学や算数の学力向上のシステムやアプリが必要不可欠となっている.

その他にも, 学習支援システムを作成する際に注意すべき点がある. 異なる学習分野や, 問題のレベル, 生徒の学力に応用することができない点が挙げられる [3], [4]. システムを作成しても, 汎用性がなければ, 意欲の低下や, 復習になってしまうため, 学習者に良い影響を与えることはできない.

そこで, 本研究では, 共通の考え方(本質)を持つ異なる分野の問題に対して, 本質の理解を促した上で両分野間の類推を支援する学習システムを提案する.

### 2 関連研究

類似関係に注目した事例検索システムによる数学学習支援 [5] では, 人間は新しい問題解決を行う際に, よく過去の問題経験を想起し, その解法を新しい問題に適用すると述べている. そのため, 過去の問題をデータベースに組み込み必要なときに呼び出せるようなシステムを開発している. しかし, まだ呼び出せるだけで, 学習者に提供しているわけではないことに加え, 考え方の理解を促す枠組みを含んでいない.

\*連絡先: 広島市立大学大学院情報科学研究科システム工学専攻  
〒731-3194 広島市安佐南区大塚東三丁目4番1号  
E-mail: kusada@sys.info.hiroshima-cu.ac.jp

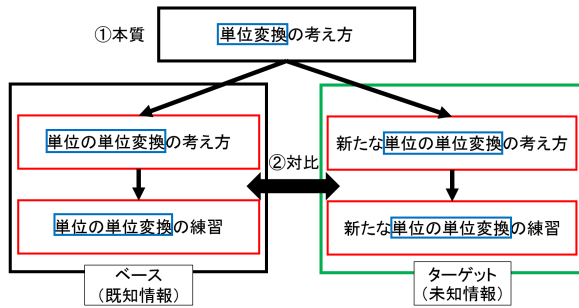


図 1: 類推による学習の枠組み

直感的な意味付けとその繰り返しにより問題の考え方の理解と定着を促す学習システム [6] では、問題を解く過程を重視した。頭の中で考えることを短時間に繰り返す練習により、問題を解く際の「考え方」を身につけられる学習システムを構築した。しかし、基本的な「考え方」を学ぶだけで、他の問題に応用しなかった。数学の学習システムの関連研究としては、「理解度に対応した一次方程式の学習支援システム」がある [7]。文章題の問題を出題し、解答者は数式を入力して答えを導くシステムとなっている。解答を間違えると、データベースから解答例を表示し、学習させる仕組みとなっている。本研究との相違点は、学習時に解答例として、本質を表示させるが、別の問題を表示させて学習を行う点が異なる。

### 3 類推を用いた数学の学習システム

#### 3.1 本研究における類推

本研究で提案する、類推による学習の枠組みを図 1 に示す。学習中に類推を行う本研究の枠組みは、本質と対比から構成される。なお、図中の青枠部分には、類推を用いた学習分野が入る。この図では、例として「単位変換」の場合の枠組みを示している。

本質では学習者が学習分野の「考え方」を理解することを表す。また対比は、新たに学習したい内容（ターゲット：未知情報）を、すでに学習した内容（ベース：既知情報）と照らし合わせながら学習を進めることを表す。

すなわち、ベースとターゲットの共通点となる「考え方」を抑えた上で、両者の類似性に着目させることで、類推を促す。

##### 3.1.1 本質

本質とは、学習過程内の基本的な「考え方」のことを示している。例えば、単位変換においては、「単位変

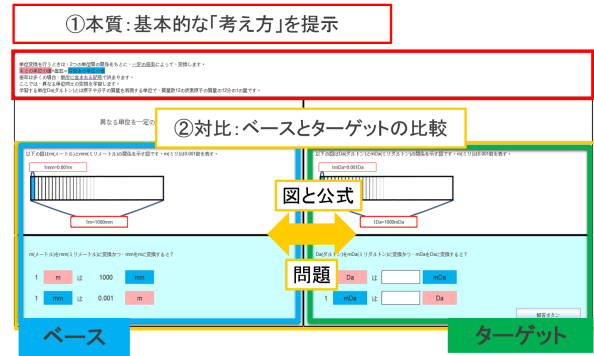


図 2: 類推を用いた学習システムの画面構成

換を行うときは、2つの単位間の関係をもとに、一定の倍率によって、変換する」という説明が相当する。学習分野内では、本質は共通になっている。また、本質の定着を促すために、学習分野で扱う公式やイメージ図、練習問題を用意して定着を図った。

##### 3.1.2 対比

対比とは、ベースを基にしてターゲットの内容を理解することを表す。例えば、1km は 1000m になるという既知情報をもとに、1kSv/h が何 Sv/h になるかを考える場合、既知情報を基に、「k」が「1000 倍」を表すと推測されることから、1000Sv/h だと類推して答える。すなわち、ターゲットの学習を行う際には、ベースでの解法を参考にし、ターゲットの学習に応用する。

この本質と対比を「単位変換」のみではなく、類推を用いることができる様々な学習分野に応用することができる。そこで、応用したものをフローチャートに沿って、学習を進めていく。以下では、数学で類推を用いることができる様々な学習分野を説明する。

上記を基に、類推を用いた学習システムを構築した (図 2)。画面上部が本質、下部が対比の画面となっている。画面下部の各画面についてそれぞれ説明していく。

まず、対比の画面上部について説明する。上部では、図と公式を示しており、学習分野に応じて内容が変わる仕組みとした (図 3)。下部では問題演習の例を示した (図 4)。こちらも学習分野に応じて変わる仕組みとした。ベースとターゲットでは同じ図を示しているが、公式については計算式は同じだが、内容は異なっている。問題演習も計算式は同じだが、内容は異なるようにした。また、ターゲットは空欄に答えを記入するように設計した。



以下の図はm(メートル)とmm(ミリメートル)の関係を示す図です。  
 m(ミリ)は0.001倍を表す。

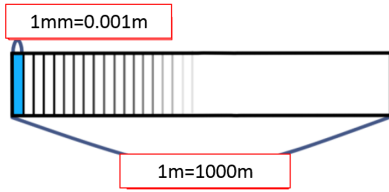


図 3: 単位変換のベースの考え方の例

m(メートル)をmm(ミリメートル)に変換かつ、 mmをmに変換すると?				
1	m	は	1000	mm
1	mm	は	0.001	m

図 4: 単位変換のベースの練習問題の例

### 3.1.3 類推学習の流れと学習ステップ

類推学習のフローチャートに示す(図 5). フローチャートで表した順番で各パネルを見て学習していくことで、本質と対比を意識して学習を進めることが可能になる。また、青い枠で囲まれている部分を1ステップとし、学習対象とする学習分野内では、問題の難易度が異なる複数のステップを用意した上で、ステップごとにこのフローチャートに基づく学習を繰り返す。

## 3.2 類推学習の対象となる学習分野

数学で、類推を行うのにふさわしい学習分野には、以下の例がある。

- 2進数の計算をベースとし、3進数の計算を行う
  - 2進数 1010 と 2進数 0001 を足す計算を基に 3進数 1010 と 3進数 0002 を足す計算を行う
- 単位変換において、既知の単位をベースとし、未知の単位を求める
  - 1km は 1000m を利用し、1kSv/h が 1000Sv/h だと求める
- 二次元ベクトルの考え方を利用し、三次元ベクトルに応用する
  - 1つのベクトルを他の2本のベクトルを利用して求める

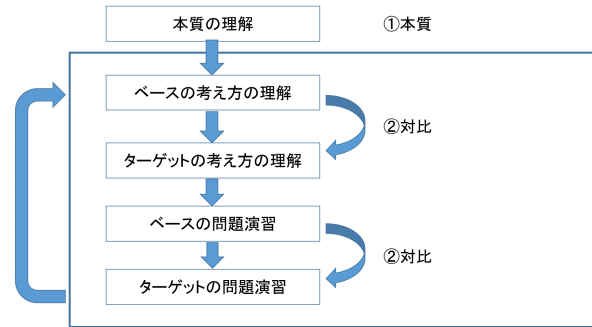


図 5: 類推学習のフローチャート

- 基にする値・割合(量)・比較量において、既知の関係をベースとし、未知の関係を求める
  - 値段、個数、合計金額の関係を利用し、エンジンの回転速度、時間、エンジンの回転数を求める

本研究では、「単位変換」と「基にする値・割合(量)・比較量」の2つの学習分野を対象としてシステムを実装した。以下で、この2つの学習分野の具体的な実装内容について述べる。ただし以下では、「基にする値・割合(量)・比較量」は、学習者の理解のしやすさを考慮して、「基にする値・割合(量)・実際の値」と呼ぶ。

## 3.3 単位変換の学習

### 3.3.1 単位変換の本質

単位変換を行うときは、2つの単位間の関係をもとに、一定の倍率によって、変換する。式で表すと、「もとの単位の値×倍率=変換後の単位の値」となる。倍率は多くの場合、単位に含まれる記号で決まる。また、異なる単位同士の足し算の場合は、どちらかの単位に単位変換を行ってから足し算を行う。また、もとの単位の値と、変換後の単位の値には色をつけた。類推を行う上で重要な対比を行う際に色が合ったほうが、分かりやすいと考えたことによる。変換後の単位の値は倍率に応じて色が変わる仕様に変更した。

### 3.3.2 単位変換の対比

「1m を mm に変換すると 1000mm」、「1mm を m に変換すると 0.001m」といった単位変換をベースとして設定し、「1Sv/h を mmSv/h に変換すると 1000mmSv/h」、「1mmSv/h を Sv/h に変換すると 0.001Sv/h」といった学習者が知らない単位変換をターゲットの学習として行う。本研究では、ベースとして、「m(メートル)と km(キ

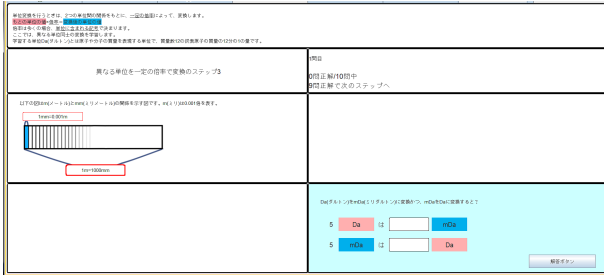


図 6: 単位変換におけるステップ 3 のシステム画面

ロメートル)」、「m と mm(ミリメートル)」、「m と cm(センチ)」、「L(リットル) と dL(デシリットル)」の関係を扱った。この 4 つを選んだ理由は、学習者全員に定着していると考えたことによる。ターゲットは学習者が学びたいと思った単位を自由に選択することができるようにした。

### 3.3.3 単位変換におけるステップ構成

単位変換におけるステップ構成は、6 ステップとした。ステップ 1, ステップ 2 は、簡単な数値を用いて新たな単位の「考え方」を定着させる。ステップ 3 では、定着したかどうか、類推を行うことができているかを調べるために練習問題を行った。このステップでは、図 6 のようなパネルの配置にした。これにより、ターゲットに関する公式やイメージ図を見ることなく、答えを求めることができるかどうかを確かめることができる。そのため、類推を上手く利用できているかを判断することができ、さらに、本質の定着を促すこともできる。

ステップ 4 からステップ 6 は、異なる単位同士の足し算「1m と 1km を足すと何 m になるでしょうか」といった単位変換の応用問題を実装した。ステップ 4, ステップ 5 では、簡単な数値同士の足し算を行い、ステップ 6 はステップ 3 と同様のパネル構成で行った。

また、扱う問題は暗算可能な数値を用いて簡単な計算で済むようにした。同じ問題を続けて行う必要はないため、ランダムで出題した。

## 3.4 基にする値・割合(量)・実際の値の学習

### 3.4.1 基にする値・割合(量)・実際の値の本質

「基にする値・割合(量)・実際の値」とは、「基準量・割合・比較量」と呼ばれる関係のこと。本研究では、直感的に理解できるようにするため、「基にする値・割合(量)・実際の値」と呼ぶこととした。また、本質を説明するパネルに色をつけることで、各値がどの値に該当するかを意識させ、類推を行いやすい環境を作った。

### 3.4.2 基にする値・割合(量)・実際の値の対比

「1 個当たりの値段×個数=合計金額」をベースとして設定し、ターゲットの学習を行った。ターゲットとしては、「1 秒あたりに送られるデータ量×時間=送信された全体のデータ量」といった学習者が知らない関係を学習する。本研究においてベースは、学習者に定着しているであろう上記の関係を使用した。また、システム画面では本質と同様に各値に色をつけた。

### 3.4.3 基にする値・割合(量)・実際の値におけるステップの構成

「基にする値・割合(量)・実際の値」におけるステップの構成は、3 つの値それぞれに、単位変換の学習と同様の 3 ステップを用意した。9 つの全てのステップクリア後に 2 ステップを有する総合問題を設けた。すなわち、合計 11 ステップの学習過程となる。総合問題では、3 つの値それぞれを求める問題がランダムで出題される。

## 4 類推を用いた学習支援の評価実験

### 4.1 実験手順

本システムによって、共通の考え方(本質)を持つ異なる分野の問題に対して、本質の理解を促した上で両分野間の類推の支援がなされているかを実験によって評価する。情報科学を専攻する大学生・大学院生 16 名を被験者とした。事前テストの結果により、結果が均等になるように提案群 8 名と比較群 8 名の 2 群に分けた。また、実験は事前テストを行ってもらった後、学習を 5 日間してもらい、学習が終わってから 3 日後に事後テストとアンケートを実施した。

### 4.2 事前・事後テスト

事前テストは、「システム内でベースとなる内容」「未知の内容かつシステム内で学習する内容」および「未知の内容かつシステム内で学習しない内容」の 3 種類を実施した。

1 種類目の「システム内でベースとなる内容」では、1m は 1000mm といった被験者が知っているだろう簡単な単位変換をテストした。次に、「未知の内容かつシステム内で学習する内容」では、1Da (ダルトン) は 1000mDa (ミリダルトン) 等の被験者が知らない単位の内容をテストした。最後に「未知の内容かつシステム内で学習しない内容」では、実験用に作成した新たな単位を用いてテストを行った。これは、被験者が絶対



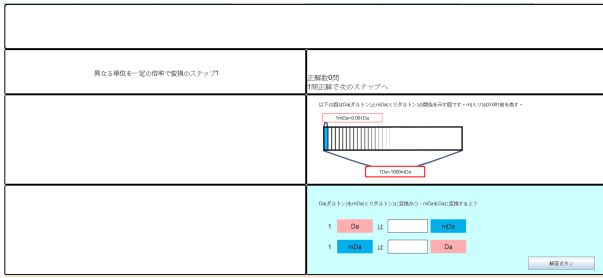


図 7: 比較システムのシステム画面

に知らない単位となる。各種類毎に1分ずつでテストを行った。事後テストは事前テスト同様に、3種類のテストを行った。制限時間も同様に1分で行った。また、事前テスト、事後テスト共に筆記で行った。単位変換だけでなく、基にする値・割合(量)・実際の値に関しても同様に、3種類のテストを事前と事後で行った。

### 4.3 比較システム

比較群には提案システム(図2)の本質とベース部分を除いた図7の比較システムを用いて学習を行ってもらった。学習過程に関しては提案群と少し異なる。その根拠は、比較群は、ターゲットのみの学習を行うため、提案群と同様のステップ構成では不具合が生じると考えたことによる。ただし、提案群と比較群の間で、学習量に差が出ないようにステップ数と問題数の調整を行った。

#### 4.3.1 比較システムのステップ構成

比較システムの「単位変換」は4ステップとした。「基にする値・割合(量)・実際の値」は各値に2ステップ、総合問題に1ステップで、全7ステップで構成されている。提案群における、単位変換のステップ数は全部で6ステップだったが、比較群では、4ステップとなっている。しかし、提案群と比較群の問題数は共に43問に設定した。「基にする値・割合(量)・実際の値」に関しては、提案群は11ステップ、比較群は7ステップとなっている。単位変換と同様に、両群共に問題数は83問に設定した。

### 4.4 評価実験の結果

#### 4.4.1 問題を解く時間の結果

学習開始時と終了時で問題を解く速度に変化があったかを調べた(図8)。提案群に関しては、ベースの演

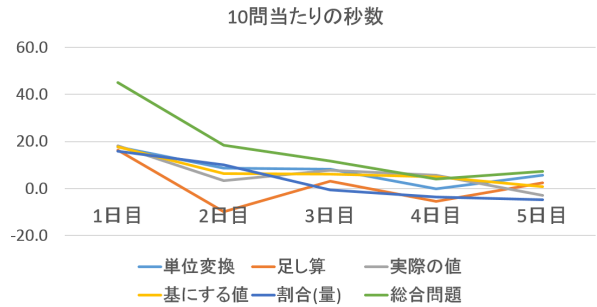


図 8: 学習時の提案群と比較群の時間差(秒: 平均値)

習問題の解き方と、ターゲットの図と公式がないときのタイムを使用している。

1日目の学習では、全ての項目において提案群よりも比較群のほうが、問題を解く速度が速いとわかる。これは、提案群が類推を行うために、本質の理解やベースとターゲットの対比を行うことで、時間がかかってしまったことが原因だと考えられる。

しかし、5日目の学習では提案群と比較群の問題を解く速度が同程度になっていることがわかる。類推を用いて解くことに慣れてきたことや、類推を用いることで学習内容をより理解したことで、問題を解く速度が提案群と比較群で同程度になったと考えられる。特に、基にする値・割合(量)・実際の値を求める問題がランダムで出題されるため、難易度が高いと考えられる「総合問題」において、提案群と比較群の差が5日間で40秒も差が縮んでいることは、類推を用いて学習したことが大きな要因といえる。

このことから、提案群は類推を用いることで、比較群と同程度の速さで問題を解けるようになったといえる。

#### 4.4.2 事前テストと事後テストの結果と考察

学習の前後で、提案システムが有効であるかを調べるために、事前テストと事後テストの結果を調べた。学習前に行った事前テストでは、「単位変換」、「基にする値・割合(量)・実際の値」で、両群共に差はなかった。

学習後に行った事後テストで、t検定を行った結果、「単位変換」、「基にする値・割合(量)・実際の値」両群共に全体の結果に差はなかった。しかし、「単位変換」では、提案群の「学習する単位」と、「未知の単位」の間には、差はなかったが比較群には大きな差があった(表1)。この差に関しては、提案群は類推を用いて学習を行ったため、「未知の単位」にも対応することができたと考えられる。また、「基にする値・割合(量)・実際の値」では、提案群の「学習する関係」と、「未知の関係」の間には、差はあったが比較群ほど大きな差はな

表 1: 「単位変換」の事後テストの結果

	提案群	比較群
学習する単位 (8問)	7.5	7.8
未知の単位 (8問)	7.3	6.3
t検定	p>0.1	p<0.01

表 2: 「基にする値・割合(量)・実際の値」の事後テストの結果

	提案群	比較群
学習する関係 (9問)	8.1	8.0
未知の関係 (9問)	6.3	5.5
t検定	p<0.05	p<0.01

かった(表2)。これも「単位変換」と同様に、提案群が類推を用いて学習を行っていたためだと考えられる。

このことから、提案群は「未知の内容」に関しても類推を用いて、他の問題へと応用することができたといえる。

## 5 おわりに

共通の考え方(本質)を持つ異なる分野の問題に対して、本質の理解を促した上で、両分野間の類推を支援する学習システムを作成した。

作成したシステムを使って実験を行った結果から、類推を用いた学習では未知の問題に応用することができることがわかった。このことから、今回作成したシステムは有効だと考えられる。

今回の実験では、被験者が大学生だったため、中学生や高校生で実験をし、さらに学習期間も長くすることで、現在よりもさらに良い結果が得られると考えられる。

また、今後は他の類推を用いた学習が可能な分野への応用を検討したい。

## 参考文献

- [1] 文部科学省：小学校学習指導要領解説 算数編，(2008)。
- [2] 文部科学省：(URL)[http://www.mext.go.jp/a\\_menu/shotou/zyouhou/1350411.html](http://www.mext.go.jp/a_menu/shotou/zyouhou/1350411.html)

- [3] 荒義明，堀口秀嗣，前田真人：小学生のためのウェブサイト「算数のまとめ：6年生」とe-L倶楽部，日本教育情報学会年会論文集，第24回，210-211，(2008)。
- [4] 江崎光治，佐々木喜一郎，安田孝美：特別支援教育におけるタブレット端末向け学習コンテンツ制作の取り組み，第76回全国大会講演論文集，2014(1)，823-825，(2014)。
- [5] 小島一晃，三輪和久：類似関係に注目した事例検索システムによる数学学習支援，電子情報通信学会技術研究報告．ET，教育工学102(65)，35-40，(2002)。
- [6] 長田佳倫，砂山渡，川本佳代：直感的な意味付けとその繰り返しにより問題の考え方の理解と定着を促す学習システム，日本知能情報ファジィ学会誌，27(5)，723-733，(2015)。
- [7] 宮地功：理解度に対応した一次方程式の学習支援システム，電子情報通信学会技術研究報告．ET，教育工学103(368)，7-12，(2003)。

# ストーリーとライフを用いた学習意欲の向上と維持の枠組み Elevation and Maintenance Framework of Learning Motivation by Story and Life functions

高橋麻祐  
Mayu Takahashi

砂山渡  
Wataru Sunayama

広島市立大学 情報科学部  
Faculty of Information Sciences, Hiroshima City University

**Abstract:** These days, in the education field, in order to foster a solid academic achievement, there is a need to enhance the learning motivation of the children. In this study, we propose the framework that was added to the Functions that can be viewed the Story depending on certain learning outcomes, and Life Function that give a constraint on the number of times that can be addressed to exercise in the typical learning system, as an element to promote the elevation and maintenance of learning motivation.

## 1 はじめに

近年の変化の激しい社会を生き抜いていくために、子どもたちに「生きる力」を身に付けさせる教育が求められている。文部科学省は、生きる力として「確かな学力」「豊かな人間性」「健康と体力」の3つを挙げ、中央教育審議会は、この中で特に「確かな学力」（基礎的な知識技能を修得し、それらを活用して、自ら考え、判断し、表現することにより、様々な問題に積極的に対応し、解決する力）を重視するべきとし、特に重要な視点として学習意欲を高めることとしている [1]。

しかし、これまでのPISA等の調査結果において、日本の子供の学習意欲が必ずしも高くないこと、また学校質問紙調査においても、「生徒は熱意を持って勉強している」と回答した学校の方が、教科の平均正答率が高い傾向が見られる [2] など、学習意欲の向上と維持は、日本の教育にとって重要な課題の1つとして挙げられる。

そこで本研究では、学習システムに、学習者の学習意欲の向上と維持を促す要素を加えた枠組みを提案する。本研究における「学習意欲」は、広辞苑 [3] の意味、“積極的に何かをしようという気持ち”、“種々の動機の中からある一つを選択してこれを目標とする能動的意志活動”、をもとに、学習に対する欲求と意志の組合せとして、「意志の加わった学習に対する欲求」として定義する。

また本研究では、学習者は学習そのものには興味がない、という前提に立ち、学習対象そのものに興味を付与するのではなく、学習環境に、ストーリーとライフという2つの要素を追加することで、学習意欲の向

上と維持を目指す。

以下本論文では、2章で関連研究について述べ、3章で提案する学習意欲の向上と維持の枠組みについて述べる。4章で提案する枠組みの評価実験について述べ、5章で本論文を締めくくる。

## 2 関連研究

これまでのゲーム開発で用いられ蓄積されてきた要素を、ゲーム以外のほかの分野において、行動のための動機付け方法として活用し、モチベーションの向上を促進させる取り組みをゲーミフィケーションという [4][5]。ゲーミフィケーションはユーザの内発的動機づけに関わる要素を有しており、ユーザのモチベーション維持に有効とされる [6]。ゲーミフィケーションの導入によって、学習意欲向上を目指すリーディング Web アプリケーションとして REX (Reading EXercise) がある [7]。この研究では、ゲーム要素として、問題の正誤情報をもとにして敵キャラクターを討伐する対戦形式、学習者群内のランキングの表示、メインの演習とは別に用意されているゲームをプレイするためのチケット制を導入している。メインの演習とは別に、学習者の興味を惹く要素が加えられている点では、本研究と類似している。しかし、学習者の興味を惹く要素が、この研究では意欲の継続を図りにくい英単語ゲームだが、本研究では意欲を継続させるストーリーである点が異なる。また、先行研究では演習に取り組む回数が制限されていないが、本研究では制限を設けることで、学習者に一気に演習を達成できない物足りなさを生み出

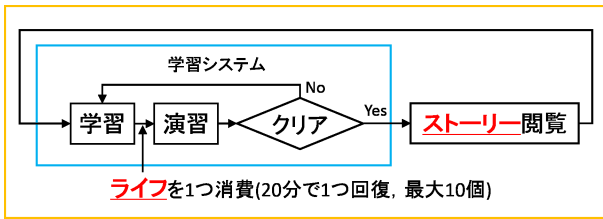


図 1: 学習意欲の向上と維持の枠組み

し、飽きにくくさせる点も異なる。

また、携帯情報端末を利用して、学習者の学習意欲を高める学習支援システムがある [8]。これは、携帯情報端末上でクイズ形式の問題を解くことにより、授業中の講義内容を復習させる学習支援システムとなっている。学習者の学習意欲を高めることを目的として、学習支援システムを開発している点、演習 1 回あたりの時間を短くし、学習を促すようにしている点では、本研究と類似している。また、学習するカテゴリを選択し、演習を行った後に答え合わせを行う流れも共通している。しかし、この研究では、メインの学習とは別に学習者の興味を惹く要素が取り入れられていない点、学習の時間に制限がない点で本研究と異なる。また、この研究は、学習テーマが授業で習った内容の復習であるが、本研究では、学習テーマが授業の復習に限られていない点でも異なる。

### 3 学習意欲の向上と維持の枠組み

本章では、学習意欲の向上と維持の枠組みを実装したシステムとその構成について述べる。図 1 に提案する枠組みを示す。一般的な学習システムにおいては、学習と演習を繰り返す形式で実装されている。学習者は、学習を行った後に演習を行い、学習した内容を定着させる。本研究ではこの一般的な学習システムに、学習意欲を向上と維持を狙う要素として、「ストーリー」と「ライフ」を実装する。以下の各節で、実装した学習システム、ストーリー機能、ライフ機能について述べる。

#### 3.1 学習システム

本枠組みで対象とする学習システムは、学習と演習のセットからなると仮定する。学習は新しい問題に対する考え方を理解するステップ、演習は理解した考え方を実践的に学ぶステップとなる。

#### 3.1.1 学習テーマ

本枠組みが対象とする学習テーマは、1 問 1 答形式で解答が可能な問題集合とする。特に、考え方を学んだ後に、繰り返しの演習によって考え方を理解する形式として、1 問の解答に多くの時間がかからない学習テーマを想定している。1 問の解答に時間がかからないことで、演習に挑戦しやすくなり、演習そのものにかかる時間が短縮されることで、後述のストーリー機能とライフ機能の効果が現われやすくなることを意図している。

考えられる学習テーマには、暗記を必要とする学習分野として、英単語の意味とスペル、元素と元素記号、歴史年号と史実が挙げられる。また、数学における単純な計算問題も対象となる。

#### 3.1.2 学習と演習の形式

学習は、暗記を必要とする学習テーマに対しては、英単語の意味とスペル、のように対となる内容を一覧表の形で提示する。計算問題の場合には、計算の方法を示す説明を表示する。1 つの学習テーマに対して、すべてを 1 度に学習することは困難と考えられるため、学習テーマ内で複数のカテゴリを生成し、各カテゴリごとに順に学習を進める。

演習は、1 問 1 答形式で出題と解答を繰り返し、演習 1 回で 10 問が出題される形式として、1 問 10 点の 100 点満点とする。この演習の繰り返しにより学習で学んだ内容を定着させる。1 つのカテゴリで一定の成績を収めると、次のカテゴリの学習に進むことができるようになる。

#### 3.1.3 実装した学習システム

本論文では、学習テーマとして「世界の国とその首都 (188 カ国)[9]」を実装したシステムを構築した。全ての国を 6 つの地域 (アジア、アフリカ、オセアニア、北アメリカ、南アメリカ、ヨーロッパ) に分け、それぞれの地域を学習テーマ内のカテゴリとして「学習」と「演習」を用意した。各地域の国の数を表 1 に示す。

学習の画面では、国の名前と首都の一覧を表示するようにした。演習では、国の名前が出題され、その首都を予め用意した誤答を含む 4 つの選択肢から解答する形式とした。演習問題 1 問の解答制限時間は 5 秒とした。これは自力での解答を促すための設定とした。またこの条件により、1 回の演習 10 問は 1 分以内に終了し、時間をかけずに気軽に演習に取り組むことができる。

地域	国の数
オセアニア	12
北アメリカ	21
ヨーロッパ	44
アジア	46
南アメリカ	12
アフリカ	53
合計	188

表 1: 各地域の国の数

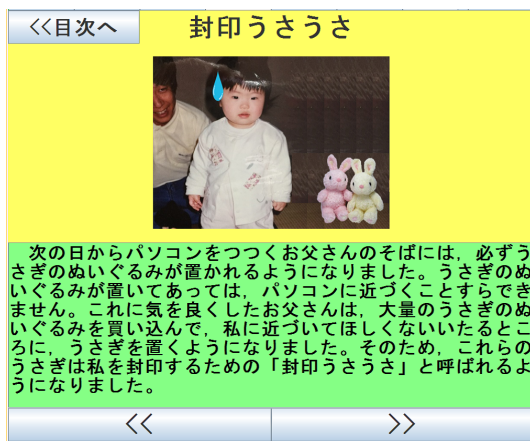


図 2: ストーリー画面の例

### 3.2 ストーリーを用いた学習意欲の向上と維持

本節では、学習意欲を向上させ、それを維持させる要素としての「ストーリー」機能について説明する。

#### 3.2.1 ストーリーの準備

ストーリーは、学習者が各カテゴリの演習をクリアした際の報酬として、クリア時に閲覧可能な要素として実装する。これにより、ストーリーの続きを見ることを目的として演習に取り組ませることで、学習意欲の向上と維持を狙う。

ストーリーの1つの話は「写真」と「本文」から構成され、学習テーマのすべてをクリアした際に一つのストーリーが完結するように、連続性を有する複数の話を用意する。ストーリー画面の例を図2に示す。

ストーリーの準備の仕方にはさまざまな方法が考えられ、既存の漫画や小説をもとに構成する方法も想定される。しかし学習者の好みが変わる場合、複数のストーリーを準備するなどの手間が増える。そこで本研究では、特に学習者の興味を引き、学習意欲の向上



図 3: 目次画面の例

と維持を狙うため、学習者が強く興味をもつ人物の話を用意することを提案する。具体的には、学習者を教える教師の逸話を用意することが望ましいと考えている。

#### 3.2.2 ストーリーの閲覧条件

ストーリーは、学習テーマにおける（カテゴリ数+1）以上の話を用意する。第1話はプロローグとして、学習開始前に閲覧可能として学習者の興味を引き出し、以降は、各カテゴリをクリアしたときに1話ずつ閲覧可能にする。また話の詳細とは別に、図3のようなストーリーの目次を用意して、タイトルをリスト形式で表示することで、継続的に意欲を引き出す。

各カテゴリのクリア条件としては、そのカテゴリの内容を十分に学習したと考えられる条件を設定する。今回実装した学習テーマ「世界の国とその首都」においては、各カテゴリの演習で「90点以上を7回獲得する」ことをクリア条件として設定した。このクリア条件は、たまたま満たされるといことがなく、無駄に長過ぎない回数でクリアとなるように定めた。

#### 3.2.3 実装したストーリー

本論文では、著者の1名の幼少期から現在に至るまでの「生き立ちストーリー」（全26話）、並びに、今後の理想の未来を描いた「将来像ストーリー」（全26話）を実装した。実装した学習テーマの6つの各カテゴリをクリアするたびに、4話ずつ閲覧可能になるように実装した。

### 3.3 ライフを用いた学習意欲の向上と維持

本節では、学習意欲を向上させ、それを維持させる要素としての「ライフ」機能について述べる。



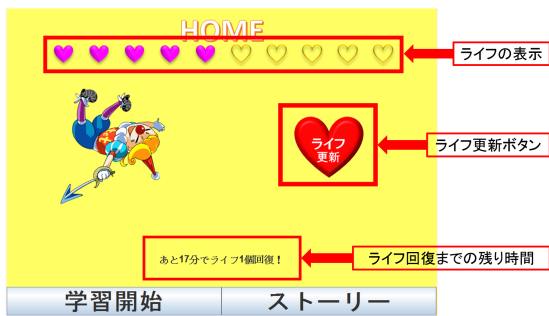


図 4: ライフ表示の画面

### 3.3.1 ライフ機能

ライフは、演習に取り組むことができる回数として実装する。すなわち、無制限に学習に取り組むことができるのではなく、学習回数に制約を設けるための機能として実装する。

学習システムの起動直後には、ライフは 10 個用意されており、1 回の演習に取り組むためにライフを 1 つ消費する。消費されたライフは、消費後 20 分経過すると 1 つ増える。すなわち、演習に連続で取り組むと、約 10 分ですべてのライフを消費し、200 分 (3 時間 20 分) の経過で、ライフが全回復の状態となる。

学習量に制約を設けることで、定期的に学習を継続させること、また学習できない時間を作ることで、「いつでも」ではなく、「今」取り組む必要性を感じさせることで、学習意欲の向上と維持を狙う。

また、いずれかのカテゴリをクリアした際にも、ライフが全回復する。これは、あるカテゴリのクリアによって達成感が大きくなりすぎること、そこで学習をやめてしまわないように、即座に次のカテゴリの学習に移行できるように設定した。

### 3.3.2 実装したライフ表示インタフェース

図 4 に実装したライフ表示の画面を示す。画面上部のピンク色のハートが残りのライフを表す。また画面下部には、ライフが 1 つ回復するまでの時間を明示する。これにより、いつライフが回復するか、あるいは全回復するかをわかりやすくすることで、ライフを効率的に利用した学習計画を立てやすくする。

グループ	ストーリー	ライフ
A	有 (生き立ちストーリー)	有
B	有 (将来像ストーリー)	有
C	無	有
D	無	無

表 2: 被験者のグループ分け (各グループ 7 名)

地域名	国数	問題数
オセアニア	12	5
南アメリカ	12	5
北アメリカ	21	5
ヨーロッパ	44	11
アジア	46	11
アフリカ	53	13
合計	188	50

表 3: テストで出題した地域別問題数

## 4 学習意欲の向上と維持の枠組みの有効性の評価実験

本章では、学習意欲の向上と維持の枠組みの有効性の評価実験について述べる。

### 4.1 実験の設定

実験は、ストーリーを作成した著者一名の友人と知人の理系大学生、大学院生 28 名に対して、平日の 5 日間に、前章で述べた「世界の国とその首都」を学習してもらった。

被験者は表 2 の 4 つのグループに分けた。グループ A とグループ B には、前章で述べたストーリー含むシステムを用いてもらった。「生き立ちストーリー」は、学習者 (実験被験者) の興味を惹くと思われる話として、「将来像ストーリー」は、学習者 (実験被験者) の興味を惹くかどうか分からない話として想定した。

ストーリーが実装されているシステムを利用するグループ A と B には、システムを初めて利用したとき、地域をクリアした (新しいストーリーが閲覧可能になった) ときには必ずストーリーを閲覧するよう指示を与えた。また、全ての被験者に、1 日に最低 10 回は演習に取り組むように指示を与えた<sup>1</sup>。また実験後には、事前テストと同様のテストを行うことを事前に伝えた。

#### 4.1.1 事前テストと事後テスト

世界の国名を問題とし、その首都名を答えるものを 50 問出題した。解答方法は記述式で 1 問 2 点とし、100

<sup>1</sup>ただし、1 日の中で時間を限定して行う実験ではなかったため、都合により挑戦回数が 10 回に満たない場合があることも許容した。

表 4: ストーリーの有無による 5 日間の演習の挑戦回数 (5 日間の合計の被験者中央値)

ストーリー	有	無	t 検定
挑戦回数	93.0	64.4	p<.05

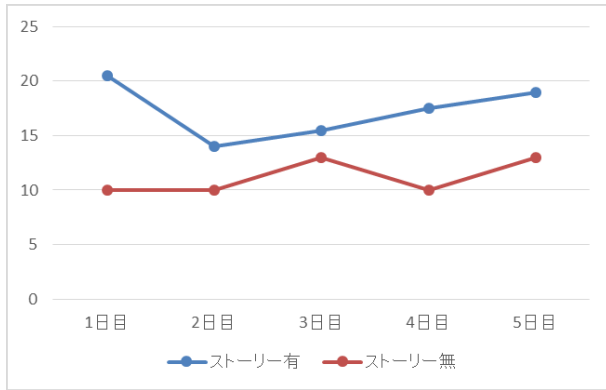


図 5: グループ A・B とグループ C の 1 日の演習の挑戦回数 (被験者中央値) の推移

点満点とした。6つの地域から偏りなく出題するために、1つの地域から最低5問出題した上で、残りは各地域の国数の25%の数の問題を出題した(表3)。解答の制限時間は10分間とした。また、事後テストは2種類行い、事後テスト1は事前テストと全く同じ問題のテストとし、事後テスト2では、事前テストと全く同じ問題で、解答(首都名)の頭文字が記述されているテストとした。

## 4.2 実験結果と考察

### 4.2.1 ストーリーの効果

表4に、ライフとストーリーがあるグループAとBの被験者の中央値と、ライフがあるがストーリーがないグループCの被験者の中央値を示す。ストーリーがあるシステムを用いたグループの被験者は、ストーリーがないグループの被験者に比べて演習の挑戦回数が有意に多かった。このことから、ストーリーが被験者の学習意欲の向上につながられたことがわかる。

図5に、ライフとストーリーがあるグループAとBの被験者の1日の演習の挑戦回数(被験者中央値)の推移と、ライフがあるがストーリーがないグループCの被験者の1日の演習の挑戦回数の推移を示す。ストーリーがあるグループの被験者の方が、ストーリーがない被験者に比べ、すべての日において演習の挑戦回数

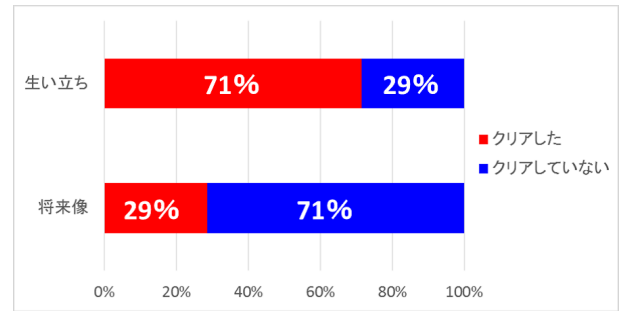


図 6: ストーリーの違いによる全地域をクリアした人の割合

表 5: ライフの有無による演習の挑戦回数 (5 日間の合計の被験者中央値, 有: グループ C, 無: グループ D)

ライフ	有	無	t 検定
挑戦回数	64.4	62.7	n.s.

が多くなった。このことから、ストーリーが被験者の学習意欲の維持につながられたことがわかる。

次に、ストーリーの内容の違いは、学習意欲に差を生むかどうかを検証した。すなわち、「生い立ちストーリー」が実装されているシステムを使用したグループAの被験者と、「将来像ストーリー」が実装されているシステムを使用したグループBの被験者とを比較する。

両グループの演習の挑戦回数(5日間の合計の被験者中央値)は、それぞれ103回と83回と差が生じたが、有意差は確認できなかった。しかし、図6の、全地域をクリアした(ストーリー最終話まで閲覧できた)人の割合は、「生い立ちストーリー」の人が多い結果となり、「生い立ちストーリー」の方が、学習者のより強い興味を引くことができたと考えられる。このことから、用いるストーリーは、学習者が知る人物の実話に基づく話の方が興味を引きやすいことが示唆される。

### 4.2.2 ライフの効果

表5に、ライフがあるシステムを使用したグループCの被験者と、ライフがないシステムを利用したグループDの被験者の挑戦回数(5日間の合計の被験者中央値)を示す。両グループでの挑戦回数の合計には、差が見られなかった。

しかし、図7に示す1日ごとの挑戦回数(被験者中央値)の推移によると、ライフがあるグループの被験

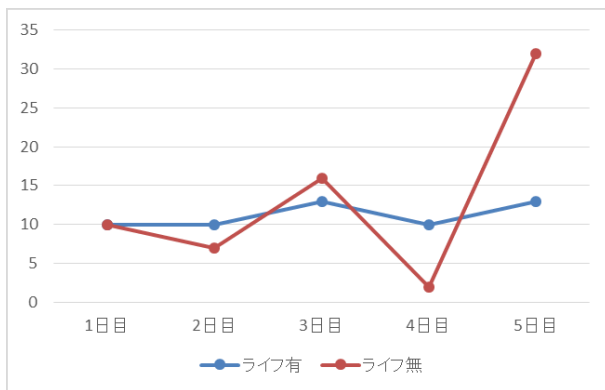


図 7: ライフの有無による演習の挑戦回数 (被験者中央値) の推移

表 6: 事前テスト, 事後テスト 1 および事後テスト 2 の結果 (被験者中央値)

グループ	A	B	C	D	分散分析
事前テスト	8	4	4	4	n.s.
事後テスト 1	20	12	14	8	n.s.
事後テスト 2	28	18	20	16	p<.05

者は、毎日継続的に演習を行っていたのに対して、ライフがなかったグループでは、特に最終日に演習を多く行っていた。これは、5日間の学習の後には、テストを行うことを伝えていたため、ライフがなかったグループの被験者は、いわゆる一夜漬けの学習を行ったことが伺える。

このことから、ライフによって学習機会が制限されることで、今しか学習できない、という限定感が、学習意欲につながったことが伺える。逆にライフがないグループでは、いつでも学習できるという気持ちだが、却って学習意欲の減退につながったと考えられる。

#### 4.2.3 テストの結果

表 6 に事前テスト, 事後テスト 1, および事後テスト 2 の結果を示す。事後テスト 2 において、グループ A の被験者のみ、他のグループの被験者と点数差が生じた。このことから、ストーリーとライフの両機能を用いることで、学習結果にもつながることが期待される。

本研究においては、学習意欲の向上と維持に焦点を当てたシステム構築と実験を行ったため、学習時には 4 つの選択肢からの選択解答方式としていた。そのため、選択肢が限定されたときには正解できたが、自由

記述として解答できるまで知識が定着できず、絶対的な点数が高くならなかったと考えられる。このことから、知識の定着には解答方式、ならびに知識が定着するまで学習の繰り返しを促す仕組みの検討が必要と考えられる。

## 5 おわりに

ストーリーとライフの 2 つの機能を実装することで、学習意欲を向上させ、それを維持する学習システムの枠組みを提案した。評価実験より、提案システムを用いた学習者の学習意欲を向上させ、それを維持できる効果があることを確認した。

今後は、一通りクリアすることですべてのストーリーが見られるのではなく、特定のカテゴリで高い正解率を出すと見られるストーリーを用意するなど、やり込み要素を追加することで、一時的ではない知識の習得の支援を行う枠組みを検討していきたいと考えている。

## 参考文献

- [1] 中央教育審議会, 「初等中等教育における当面の教育課程及び指導の充実・改善方策について (答申)」, 2003.
- [2] 文部科学省, 国立教育政策研究所, 「全国学力・学習状況調査報告書」, 2013.
- [3] 新村出編, 『広辞苑』 第五版, 岩波書店, 1998.
- [4] 矢沢崇史, 高橋稔, 多田伊佐武, 市村哲, 「視力入力を用いて英文読解を支援するゲーミフィケーションの研究」, 情報処理学会研究報告. GN, [グループウェアとネットワークサービス研究報告会], 2015-GN-94 巻, 7 号, 2015.
- [5] 井上朋人, 『ゲーミフィケーション<ゲーム>がビジネスを変える』, NHK 出版, 2012.
- [6] 深田浩嗣, 『ゲーミフィケーションはなぜハマるのかゲーミフィケーションが変える顧客満足』, 2011.
- [7] 大城敬人, 宮岸祐成, 宮崎佳典, 「ゲーミフィケーションによる学習意欲向上を目的としたリーディング Web アプリケーションの構築」, 情報処理学会第 76 回全国大会講演論文集, pp.805-807, 2014.
- [8] 松本潤, 松本裕二, 盧存偉, 「携帯情報端末を利用した学習支援システムの開発と効果検証」, 電子情報通信学会技術研究報告. ET, 教育工学 112(300), 2013.
- [9] 外務省編: 世界の国一覧表, 世界の動き社, 2003.



# ツールの操作履歴の共有によるテキストマイニング時の発想支援

## Creativity Support for Text Mining System by Sharing Tool Operation History

金谷 直哉                      砂山 渡  
Naoya Kanaya                Wataru Sunayama

広島市立大学大学院 情報科学研究科  
Graduate School of Information Sciences, Hiroshima City University

**Abstract:** The total environment for text data mining, includes a number of mining tools, which allows the text analysis in a variety of ways. When performing text mining, it is necessary to analyze the results of the analysis widely collected. However, it is limited in the idea of one person. If it is possible to see how to use the unknown tools, it may be connected to the idea of obtaining new results. Therefore, this study suggests function for TETDM user to share with each other the operation history of the tool.

## 1 はじめに

近年、非常に多くのアプリケーションソフトウェアが登場し競争が行われている。これによりアプリケーションソフトウェアの多くは非常に多機能になり、様々な目的に利用することができるようになった。それに伴い我々はコンピュータを日常的に利用し、様々な活動を効率的に行えるようになったが、便利になる反面、豊富な機能の中から有用な機能を選び出し使いこなすことは非常に難しくなっているという点が課題となっている。

豊富なテキストデータマイニングツールを提供するテキストデータマイニングツールのための統合環境として TETDM[1][2] がある。TETDM では多彩なテキストデータマイニングツールを提供しており、分析に没入して新たな発想を得られる環境の構築を目指している。TETDM で利用できるツールの機能は多様で様々な目的に合わせて使うことができる。しかし、それぞれのツールの使い方は異なるため、目的に応じすべてのツールを使いこなすことは難しい。さらに TETDM では柔軟に新しいツールを開発しシステム上に実装することができ、新たなツールの提供も行われるため、最適なツールの選択を行うには、高い技術が求められる。

そこで本研究では、TETDM 上の機能の 1 つとして、操作履歴を保存する機能、他者の操作履歴を元に操作を再現する機能を実装し、様々な利用者が行ったツール操作を自由に閲覧できるようにすることで、TETDM のツールによるテキストマイニングを行う際の発想を支援し、利用者のテキストマイニングを支援することを目的とする。

## 2 関連研究

本章では TETDM によるマイニングの発想支援の関連研究を、操作履歴共有、操作の再現手法の観点からまとめる。

機能実行履歴の共有によって、多機能なソフトウェアの有用な機能の発見を促す研究がある [3]。この研究は、ソフトウェア使用目的がよく似た人同士で、機能使用履歴を参照可能にすることで、ユーザーがそこから有用な機能発見を可能にすることを目指した研究で、Word、PowerPoint の利用機能を共有する評価実験を行ったところ、平均 16.1 個の未知の機能を発見することができ、被験者はそのうち約 40 % は有用な機能であったと解答した、という結果になった。本研究では、この理論を用いて、TETDM のツール利用の発想支援を目指す。

また視線行動を記録し注視点を可視化することで、着眼スキルを発達させる支援を行う研究がある [4]。この研究では可視化された他ユーザーの注視点と自身の行動を比較することで、知識や思考の違いを取り着眼スキルを発達させることができるミラーエージェントシステムを開発した。視線行動の入出力はマウスカーソルによって行われた。評価実験として自動車運転時の危険予知トレーニングを行い、着眼スキルが有意に向上していたことが分かった。この研究から他者の行動履歴の再現から他者の感覚を感じとることができることを示唆していた。そこで本研究では他者のツール操作の再現から操作の発想を獲得する支援を目指す。

Web アプリケーションの動作検証を目的として、操作のリプレイを再現する手法を開発した研究がある [5]。

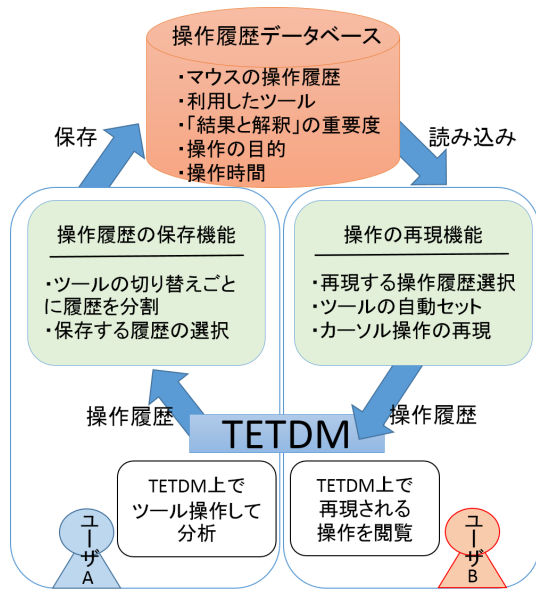


図 1: テキストマイニング時の発想支援の枠組み

このシステムは Capture 部と Replay 部で構成され、Capture 部では Web アプリケーション上で操作が行われると、ページを進む、戻る、ブックマークの操作や、クリック、スクロール、マウス移動などのカーソル操作のイベントを保存する。Replay 部ではこうして保存された操作履歴の内の Web アプリケーションのページ遷移の再現を行う。このシステムではマウスカーソルの操作の記録を行っていることから、より詳細な再現を可能にすることでアプリケーションの操作ガイドに応用できる可能性がある、と述べている。本研究では TETDM のツールを使ったマイニングの発想支援を目的として再現を利用する。

### 3 テキストマイニング時の発想支援

TETDM には数多くの機能が実装されており、TETDM 上でのテキストマイニングの方法も一人ひとり異なる。本研究では、TETDM のツールの操作履歴を再現できる機能を実装し、利用者がお互いに操作履歴を共有できるようにすることで、テキストマイニング時の発想を支援し知識の発散を促すことを目的とする。

TETDM は、ツールを利用して特徴的な結果を得た際に、それを記録するための「結果と解釈」の登録インターフェース (図 2) を備えている。ここで、さまざまな分析結果を得て、より多くの結果と解釈を登録し、知識創発における思考の材料をより多く集めることが、根拠があり質の高いアイデアの生成につながると考え、この結果と解釈の登録数が増えることを目指す。

発想支援の機能は、操作履歴のデータベース、操作

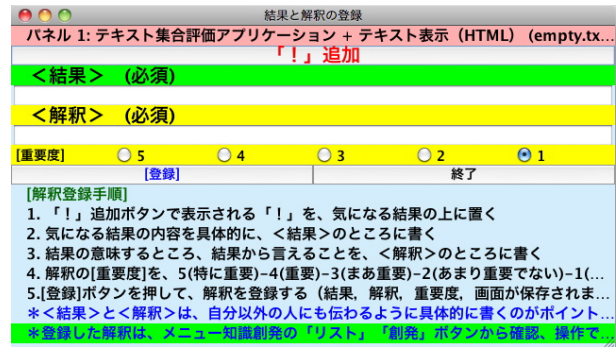


図 2: 結果と解釈の登録インターフェース

表 1: 記録されるツールの操作

TETDMの操作	記録される操作	再現の可否
カーソル移動	○	○
ボタン	○	○
チェックボックス	○	○
スクロールバー	○	○
コンボボックス	○	○
ドラッグ	○	○
ホイール	○	○
テキスト入力	×	×
メニューウィンドウのボタン操作	○	×
ツールの変更操作	○	×
「結果と解釈」の入力	○	×
TETDMのその他の機能	×	×

履歴の保存機能、操作履歴の再現機能で構成されている。テキストマイニング時の発想支援の枠組みを図 1 に示す。以下でこの枠組みの詳細について述べる。

#### 3.1 操作履歴のデータベース

操作履歴保存機能と操作履歴の再現機能の間でデータベースを介して、操作履歴の受け渡しを行う。データベースにはツールを利用した際の以下の情報が保存される。

- 表 1 の操作が行われたディスプレイ上の XY 座標
- 操作履歴でセットされたツールの組み合わせ
- ツールを操作していた時間の長さ
- ツール操作から得られた出力結果を元に登録された「結果と解釈」の重要度の合計点
- PC の OS, ディスプレイサイズ, 処理性能
- 操作した人が保存時に入力した操作の目的

#### 3.2 操作履歴保存機能

##### 3.2.1 操作履歴の分割

TETDM によるテキストマイニングでは、目的とする結果を得るため様々なツールに切り替えながら処理

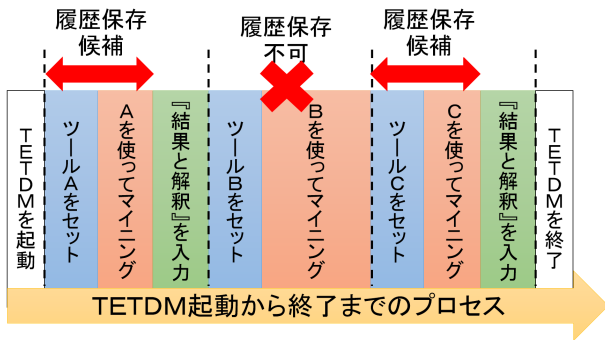


図 3: 操作履歴分割の仕組み

結果の分析を行う。データベースに保存する操作履歴を作成するため、ツールの組み合わせが変更されたタイミングでカーソル操作の操作履歴の分割を行う。そのため、1つの操作履歴には1つの組み合わせのツール操作のみの履歴が保存される。また本機能では、TETDMにあらかじめ用意されたツールの組み合わせを利用したツール変更のタイミングを分割するタイミングとして利用した。図3に操作履歴保存の仕組みを示す。

### 3.2.2 操作履歴の保存

TETDM 操作終了時に、利用者は分割した操作履歴の中から有用な操作を行ったツールを選択し保存する。ツールを操作した後、出力結果を元に「結果と解釈」を登録した場合、そのツール操作は有用な可能性が高いと考えられることから、保存操作を行う際は「結果と解釈」の入力操作を行ったツールのみ選択し保存することができる仕様にした。保存されるデータは図3の仕組みで分割された TETDM のツール操作とそのツール ID、そのツール操作時に登録された「結果と解釈」に与えた重要度の合計点、PC に関する情報があり、保存操作がされたとき1つの再現用操作履歴としてデータベースに保存される。

### 3.2.3 操作履歴保存を行うインターフェース

TETDM の利用を終了する際、有用な結果が得られたツールの操作履歴を選択し保存する機能を TETDM の機能の1つとして実装した。TETDM の終了操作を行うと、TETDM の終了と入れ替わりに操作履歴保存ウィンドウがポップアップし、このウィンドウから操作した履歴の保存を行うことができる。操作履歴保存ウィンドウには、TETDM でセットしたツールの組み合わせ一覧が表示される。一覧にはツールを使って登録された「結果と解釈」の重要度合計点、ツールを操作した時間も合わせて表示される。

操作履歴一覧にはセットしたツールの組み合わせが

全て表示されるが、重要度合計点が0点の場合は選択することができない。保存可能な操作履歴を選択した後、『操作履歴保存』ボタンをクリックするとさらに操作履歴のタイトルと操作の目的を入力するウィンドウがポップアップする。各項目を入力することで、操作履歴の保存が完了する。

## 3.3 操作履歴の再現機能

### 3.3.1 再現する履歴の選択

TETDM 上で再現機能を起動すると、データベースから読み込んだ操作履歴の一覧が表示され見たいものをクリックすることで操作履歴の選択が行える。操作履歴一覧をクリックした段階でカーソル操作の履歴がデータベースから読み込まれ TETDM には自動的に操作履歴で利用されたツールがセットされるが、カーソル操作の再現は行われずさらに再現開始を確定する操作を行う必要がある。利用者は再現を閲覧するかどうか、実際にツールを見ながら選択することができる。

### 3.3.2 再現の実現方法

操作履歴の再現機能では、データベースに保存された操作履歴を参照し TETDM のツールの操作を再現する。操作履歴の再現は、ツール上でのマウスカーソルの動きの再現によって行い、本研究ではカーソルの動作の再現を行うため、Java の Robot クラスを利用した。

本機能ではツール上での操作のうち、クリック、ドラッグ、ホイール操作の再現を行い、TETDM を操作する。操作履歴に記録されたディスプレイ座標にマウスカーソルを移動させ、マウスの左ボタンの入力、ホイールの回転操作の命令を Robot クラスにより入力することで、操作履歴と同じ操作の再現を実現した。

再現中でもカーソルの操作が可能な時間として、一定回数操作が行われるたびに5秒間の再現中断間隔を設けた。再現を開始するとマウスカーソルが自動的に動き始め自由に操作することができなくなり、利用者は自分のしたい操作をするために再現がすべて終わるまで待たなければならなくなる。利用者は再現中断間隔を利用することで、再現が終わるまで待たず操作を中断したり調整をすることが可能である。再現中断間隔までの操作数は再現開始時に選択することができる。

### 3.3.3 再現を行うインターフェース

操作履歴をデータベースから選択し再現の閲覧を行う機能を TETDM の機能の1つとして実装した。TETDM

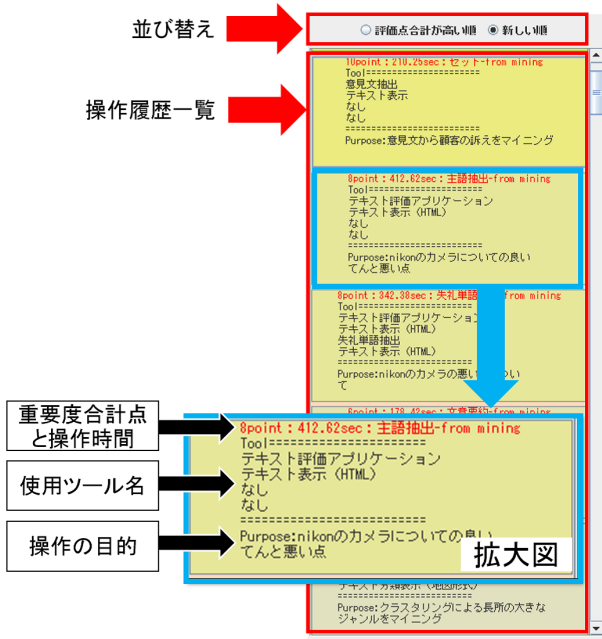


図 4: 再現ウィンドウの操作履歴一覧



図 5: 再現制御ウィンドウ

のメニューウィンドウから『再現』の項目を選択することで、再現ウィンドウがポップアップし再現を閲覧するための操作を行うことができる。このウィンドウでは再現速度、再現中断間隔の設定を変更することができる。

再現中の速度は、再現が始まった後でも再現中断間隔を利用して変更することができる。

さらに再現ウィンドウでは再現したい操作履歴の選択を行うことができる。再現ウィンドウの右側に操作履歴が一覧で表示されており、各操作履歴に関する以下の情報を確認することができる。

- 操作から得られた重要度の合計点
- 操作履歴の操作時間
- 利用されていたツールの種類
- 操作履歴保存時に入力された操作の目的

操作履歴一覧から閲覧する操作履歴を選択しクリックすると、TETDMに操作履歴のツールの組み合わせが自動的にセットされる。最後に『再現開始』ボタンを押すと、再現ウィンドウの表示が消え再現が開始される。

### 3.3.4 再現制御ウィンドウ

再現中断間隔を利用して再現速度の変更や停止操作を行うことができる図5の再現制御ウィンドウを開発した。このウィンドウは再現ウィンドウの『再現開始』ボタンを押すと同時にディスプレイ左上に表示される。

再現制御ウィンドウの上部の三角は現在の再現速度を表し、色のついた三角が多いほど速度が速くなっていることを表している。次に中央の水色のメーターは再現進行度を表し、水色のメーター量で再現がどれくらい進んでいるかいつでも確認することができる。下部の4つのボタンは再現の制御用のボタンで、再現の一時停止、終了、速度変更を行うことができる。

## 4 操作履歴の再現機能の評価実験

本章では、TETDMのツール操作履歴の再現機能を利用し他者のマイニングを閲覧することで、新たなテキストマイニングの発想を獲得し、そこから新たな知識を創発ができるのかを検証した実験について述べる。

### 4.1 実験方法

実験は、TETDMのチュートリアル『初心者』『初級』の課題をクリアしたことがある大学生と大学院生12名を対象として行った。実験では全被験者に同一スペックのPCを使用してもらい、2日間でTETDMによるテキストマイニングを行ってもらった。

マイニング課題においては、通信販売サイトで販売される nikon 製カメラのレビューを収集し使用した。レビューは全40件をまとめたもので、多様な知識を発見できるように評価値5, 4のものと1, 2のものからほぼ同数になるよう収集した。“（被験者名）カメラ”の社員であると仮定して、ライバル会社 nikon に勝つためのカメラのアイデアを考えてもらい、そのために「結果と解釈」機能に、nikon のカメラの特徴とそこから得られた解釈を登録を行ってもらった。

1日目の実験では、被験者にTETDM上でレビュー記事をテキストマイニングしてもらい「結果と解釈」を登録してもらった。「結果と解釈」は最低でも10件以上の登録を目指してもらい、行った操作の履歴を保存してもらった。

2日目の実験では、TETDMでレビュー記事をテキストマイニングし1日目のものに追加する形で「結果と解釈」を登録してもらった。「結果と解釈」は可能な限り登録することを目指してもらった。2日目では再現機能により1日目の他の被験者の操作を閲覧することができ、分析に取り入れることができるようにした。これによって個人では得られなかった分析結果の獲得が可能になるか検証した。



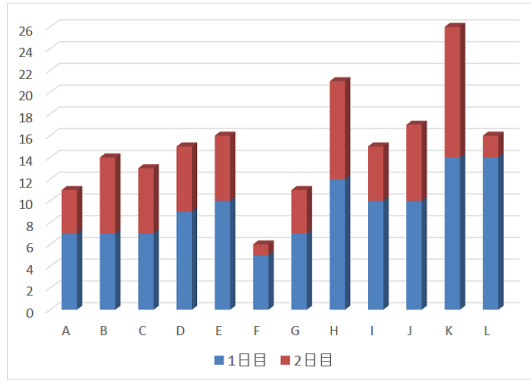


図 6: 登録された「結果と解釈」の数

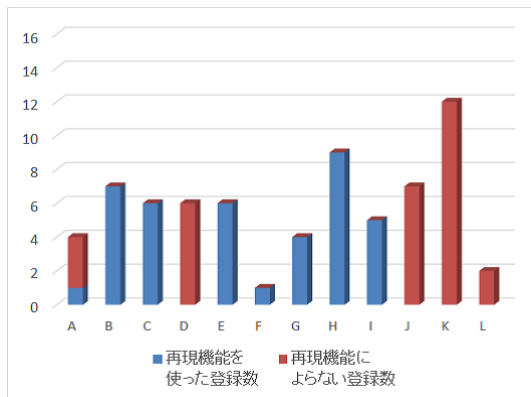


図 7: 操作履歴閲覧後の登録数

## 4.2 結果と考察

### 4.2.1 「結果と解釈」登録数の伸び

TETDMを利用して新しい登録できなくなるまで「結果と解釈」の登録を行った被験者でも、再現機能を利用することで、マイニングの新たな発想を獲得し、新たな「結果と解釈」を入力することができるようになるのかを検証するため1日目の時点の「結果と解釈」数と2日目の累計「結果と解釈」数を比較した。結果を図6に示す。1日目のマイニングでは「結果と解釈」を最低でも10個以上登録することを課題としたが、半数の被験者は10個以上「結果と解釈」として登録する結果を得ることができなかったが、図6を見ると、全被験者において再現機能を使って再度マイニングをすることで、「結果と解釈」数が増えており、マイニングの新発想を得ることができたことがわかる。

### 4.2.2 操作履歴の閲覧から得られた「結果と解釈」

「結果と解釈」の登録数の増加が再現機能によって発生したものあることを検証するため、再現機能を利用して登録された「結果と解釈」の集計を行った。

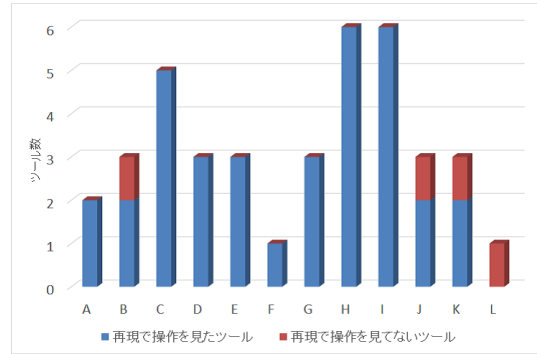


図 8: 2日目の「結果と解釈」登録に繋がったツール数

再現機能を利用し他者の操作を閲覧した直後、再現機能で見たツールを利用して登録された「結果と解釈」は、再現機能の閲覧によって得られたマイニングの発想によって、創発されたものであるといえる。そこで2日目の実験において登録された「結果と解釈」のうち、再現機能を使って操作を見た後そのツールから別のツールに変更せず登録された「結果と解釈」の数の集計を行った。結果を図7に示す。

図7から、再現機能を使い新たに登録された「結果と解釈」が多数あったことが分かる。この被験者の再現機能の使い方を詳細に検証すると、再現を見てその後操作せず「結果と解釈」を登録している場合と、再現を見た後ツールを操作して「結果と解釈」の登録を行った場合が見られた。これは他者の操作が出した出力結果から新たな発見をしたという場合と他者の操作方法から操作方法の発想を得た場合があると考えられる。さらに、再現途中で再現を停止させて、自身で操作を行った例も見られ、これはカーソル制御によって再現を行う利点として考えられる。

### 4.2.3 「結果と解釈」登録に繋がったツール数

図7では操作履歴閲覧直後そのツールを利用した登録のみを操作履歴の閲覧により得られた登録とした。しかし、操作履歴を閲覧してしばらくして、後に「結果と解釈」登録に繋がる可能性も考えられる。

被験者が「結果と解釈」の登録に利用できたツールは『そのツールでマイニングを行う発想がある』といえる。そこで2日目で各被験者が「結果と解釈」の登録に利用したツール数の内、「結果と解釈」を登録するより以前にそのツールの操作履歴を見ていた数がどのくらいあるかについて検証を行った。図8では、2日目の実験で各被験者が「結果と解釈」の登録に利用したツールの種類と、事前に操作履歴を見ていた割合を示す。

図8から、ほとんどの被験者は再現によって見たツールを使って「結果と解釈」の登録を行っていたことが

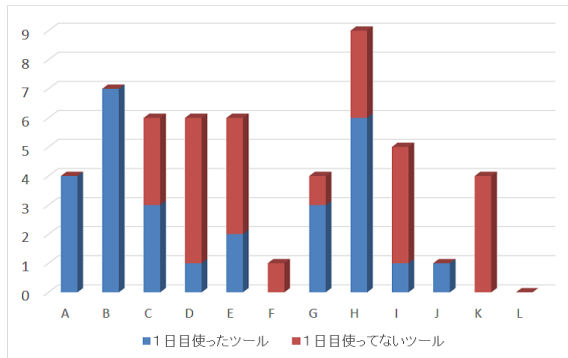


図 9: 2 日目に再現で操作を見たツールで登録した「結果と解釈」数

分かる。このことから操作履歴再現の閲覧が、ツールを使い方の新しい発想に繋がったことが分かる。また、アンケートのコメントに「使わなかったツールを使うきっかけになった」というものが複数あったことから、再現の閲覧がツールの使い方を考えるきっかけになり、それが新しい「結果と解釈」の登録に繋がったことが考えられる。

#### 4.2.4 再現で操作を見たツールで登録した「結果と解釈」数

図 9 では、2 日目に再現機能で操作を見たツールから得られた各被験者の「結果と解釈」数と、その内の 1 日目に使っていないツールから得られた「結果と解釈」数の割合を示す。

この結果、1 日目に使っていないツールを利用して「結果と解釈」を登録する例がほとんどであることが分かった。このことから、再現機能で操作履歴を閲覧する機能は特に、操作方法を知らないツールの使い方を知る、有効に使えなかったツールを活用する発想を得る、使わなかったツールを使うきっかけになるという目的において効果的であると考えられる。または、1 日目に使っていたツールで新たな「結果と解釈」を得ていたという例も複数見られ、これは使い方を知っているツールの新しい使い方の発想を得る目的において効果を発揮したためであると考えられる。

#### 4.2.5 「結果と解釈」で登録された内容の比較

1 日目と 2 日目で同一のツールを利用して登録された「結果と解釈」の比較を行った。失礼単語を抽出するツールを使った登録を見ると 1 日目では「ボタンの操作性はいまいち」と曖昧な文章が登録されていたが、2 日目では「AF ON ボタンが無い」という比較的具体的な文章になっていた。他にも詳細な文章になってい

る例は複数見られ、このことから他者の操作履歴を閲覧することでツールの使い方の発想を獲得することで、以前は行えなかった詳細な分析を行えたと考えられる。

## 5 結論

TETDM を利用したマイニングを行う際の発想を支援するため、TETDM の機能の 1 つとしてツールの操作履歴の保存、保存した操作履歴の再現機能を開発し、操作履歴の共有を可能にした。

操作履歴の再現機能の効果を検証するために評価実験を行った。レビュー記事についてテキストマイニングを行い商品の特徴を発見するという課題を、操作履歴の再現機能を導入して行ってもらったところ、再現機能で他者の操作履歴を閲覧することがより多くの特徴の発見に繋がること、多くのツールを利用したマイニングに繋がることが分かった。

今後の課題として、操作履歴の再現の精度の向上、再現に対応した操作の充実、操作意図の伝達効果の向上によって、マイニング時の発想をより効果的に支援することを目標としていきたい。

## 参考文献

- [1] Total Environment for Text Data Mining (テキストデータマイニングのための統合環境), (URL) <http://tetdm.jp>.
- [2] 砂山渡, 高間康史, 西原陽子, 徳永秀和, 串間宗夫, 阿部秀尚, 梶並知記, 松下光範, Bollegala Danushka, 佐賀亮介, 河原吉伸, 川本佳代: テキストデータマイニングのための統合環境 TETDM, 第 29 回人工知能学会全国大会, 2E3-NFC-01a-1, (2015).
- [3] 森崎修司, 白石裕美, 門田暁人, 大和正武, 松本健一, 鳥居宏次: 機能実行履歴を用いたソフトウェア機能の発見支援システム, 電子情報通信学会論文誌 D - I, Vol.J84 - D - I, No. 6, pp.755-767, (2001).
- [4] 藤本武司, 砂山渡, 山口智浩, 谷内田正彦: 視線行動の可視化による着眼スキル伝達支援, 人工知能学会論文誌, Vol.71, No.1, pp.517-518, (2009).
- [5] 柿元宏晃, 佐野博之, 大園忠親, 新谷虎松: Web 行動リプレイシステムに基づく Web アプリケーション動作検証システムとその応用, 第 71 回情報処理学会全国大会講演論文集, Vol. 19, No. 3, pp.174-183, (2004).

# TETDM 利用者の利用意欲向上のためのゲームモードの開発

## Development of TETDM Game Mode for Elevating Use Motivation

竹岡 駿 砂山 渡  
Shun Takeoka Wataru Sunayama

広島市立大学大学院 情報科学研究科  
Graduate School of Information Sciences, Hiroshima City University

**Abstract:** Total environment for text data mining TETDM, also to people of non-specialists of text mining, is used to analyze the Web page, e-mail, SNS and BBS comments, as an environment that is used to analyze the text information is many opportunities to see it is. In addition, by the user to increase, increased development motivation of tool developers is considered an advantage to be easily added to the total environment occurs. So, with the aim of improve the motivation to use TETDM, we propose a game mode add the elements of the game.

### 1 はじめに

近年、テキストマイニングのツールやシステムが数多く開発されてきている。企業における開発や販売促進などの具体的な分析目標に対して、テキストマイニング専用のツールを使うためには、テキストマイニングやデータの取扱いに関する知識を必要とする。しかし、かな漢字変換や文章の自動校正など既存のソフトウェアに組み込まれ、すでに無意識に利用するテキストマイニングツールが存在し、ネット上での書き込みの履歴から、書き込んだ人の年齢や性別、出身地を推定するなど興味を引かれるサービスが提供されている現状もある。

現在の世の中では、メールなどネットワークを介してお互いの情報を電子的にやりとりすることが多く、情報収集においても、インターネットで検索を行うことが一般的となっている。これらの情報は「テキスト」形式となるため、テキストマイニングツールは、卑近な利用に対してより便利な環境を提供できる可能性を持つとともに、利用者は便利な環境が現われることを潜在的に望んでいると考えられる。

テキストマイニングの専門家以外の利用、ならびに幅広い利用目的を想定した環境として、テキストマイニングのための統合環境 TETDM[1] が開発されている。このソフトウェアは JAVA で記述されているため、Windows と Mac のプラットフォームに依存せず、特別なインストール作業は不要で、必要なファイルをコピーするだけで、すぐに利用を開始できる。

しかし、具体的な利用に向けて意欲を与えられなければ、実際の利用には至らない。そこで本研究では、TETDM の利用に際して、TETDM にゲームの要素を

加えることで、利用者の利用意欲を向上させ、より多くの人がテキストマイニングのツールとして TETDM を利用できるようになることを目指す。特に TETDM は、利用可能なツールを後から任意に追加できる形態となっているため、利用者の増加が開発者の増加につながって環境が改善され、それがまた利用者の増加につながることが期待される。

以下本論文では、2 で研究背景と関連研究、3 で提案する TETDM のゲームモードについて述べる。4 で実装したゲームモードが利用意欲の向上に寄与するかを検証した実験について述べ、5 で本論文を締めくくる。

### 2 関連研究

近年、ソーシャルゲームを代表とする、ゲーム業界が急成長を遂げており、学習意欲向上を目的にゲーム要素を取り入れるシステムは数多く作成されている。

プログラミング学習とゲームを組み合わせた研究として Christine らの「Enhancing Self-Motivation in Learning Programming Using Game-Based Simulation and Metrics」[2] がある。この研究では独自に作成したゲームの戦術を、プログラミングによって実装することでプログラミング学習の意欲向上を狙っている。この研究では、もともと学習意欲があり、プログラミングの知識がある人にシステムを利用してもらい学習意欲の維持をはかっているが、本研究ではテキストマイニングの知識がない人に意欲を湧かせると共に、利用意欲を継続させることを目標としている点が相違点となる。

パソコンでの作業意欲を向上させるためのシステムの開発として倉本らの「作業意欲を持続的に維持向上

させる EELF に基づく主観的比較型エンタテインメントシステム」[3]がある。この研究では、デスクワークの意欲向上を目的とし、パソコンでの作業時間に応じて画面内のキャラクターが成長する要素を取り入れている。本研究との相違点は、キャラクターが成長することではなく、キャラクター多く集めることが目的の点となる。

キャラクターを使用して、意欲向上を目指す研究として、益子らの「キャラクターを用いた学習継続支援ツールの開発」[4]がある。この研究では、パソコン上で文献の蓄積と整理をする際にキャラクターを利用したシステムを開発し、支援を行っている。同じくキャラクターを使用して、意欲向上を目指す研究として渡辺らの「セルラー型授業におけるエージェントキャラクターによる学習支援」[5]がある。この研究ではキャラクター同士の会話形式でプログラミングのヒントを提示している。これらの研究ではキャラクターが作業のヒントを直接しているが、本研究ではキャラクターを利用することでより効率的に作業を進めることができる点が相違点となる。

eラーニング学習とゲームを融合した研究として小島の「ゲームと融合したeラーニング問題集の開発と評価」[6]がある。この研究では、外国語学習のeラーニング教材にゲーム要素を加えることで学習意欲を湧かせることを目的としている。この研究では、多くの問題に正解しなければゲームが進まないが、本研究ではTETDMを利用していれば意識していなくてもゲームを進めることができる点が相違点となる。

また全ての関連研究に言えるが、ゲームの要素とテキストマイニングを組み合わせる研究ではないところも相違点となる。

### 3 TETDM ゲームモード

本章では、TETDM の利用意欲向上に向けて実装したゲームモードについて述べる。図1に、ゲームモードの枠組みを示す。

ゲームモードとして、多くの市販ゲームやソーシャルゲームで実装されている、ゲームへの取り組みの量を表す「経験値」と、ゲーム内通貨となる「コイン」を実装する。これをもとに、「経験値」の増加による「ランク(レベル)」, ならびに「コイン」の利用先として「キャラクター」を実装する。また、利用目的として「ミッション」を実装し、ミッションの達成によっても「経験値」と「コイン」が獲得できるようにする。

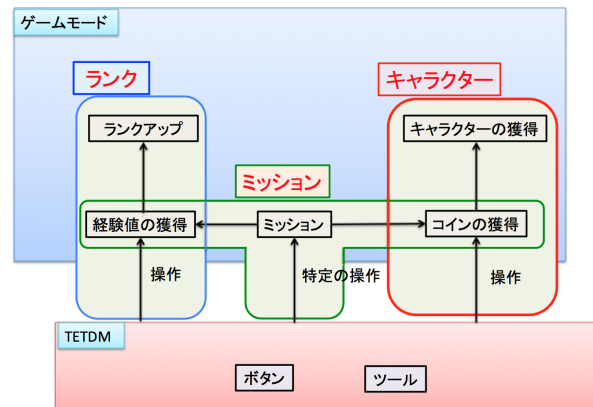


図1: TETDM ゲームモードの枠組み

表1: ボタンやツールの利用回数と獲得経験値

ボタンやツールの利用回数	獲得経験値
1回目	10
2回目	5
3回目	2
4回目以上	1

#### 3.1 経験値

経験値は、統合環境内のボタンとツールの利用、チュートリアルの実施、テキストマイニングプロセスにおける結果と解釈の登録と知識創発、後述のミッションのクリア、によって得ることができる。

##### 3.1.1 統合環境のボタンとツールの利用

経験値は、TETDM 内のボタンを押すたびに、利用回数に応じた表1の数値を獲得できる。すなわち、定常的には各ボタンの利用で1の経験値しか獲得できないが、利用していないボタンをなくし、幅広い操作を行ってもらうことで、より多くの経験値が獲得できるように設定している。また、パネルにセットして利用するツールについても同様に、各ツールの利用回数に応じて、ツールをパネルにセットしたタイミングで表1の経験値を獲得できる。こちらも、より多くのツールを利用してもらうことを意図している。

その他、テキストマイニングプロセスにおける「結果と解釈の登録」を行った場合、また、知識創発ウィンドウを用いて登録された解釈を1つにまとめた際にも経験値が与えられる。



表 2: 連続起動による獲得コイン数

連続起動日数	獲得コイン
1日	100
2日	150
3日	200
4日	250
5日	300
6日	350
7日以上	400

### 3.1.2 チュートリアルのクリア

TETDM 内に現在実装されている、3 種類の利用者向けチュートリアル「初心者」「初級」「文章推敲」の各課題をクリアするごとに、経験値を得ることができ、各課題のクリアによって 10 から 70 の経験値を獲得できるが、経験値が得られるのは 1 度限りとなっている。早い時点で、操作方法を積極的に習得してもらうために設定した。

### 3.1.3 結果と解釈の登録と知識創発

テキストマイニングプロセスにおける結果と解釈の登録と知識創発に関して、結果と解釈を 1 回登録するたびに 5 の経験値、知識創発ウインドウを用いて、解釈を 1 つにまとめた際には、まとめるもの解釈の数 × 10 の経験値が獲得できる。これは、テキストマイニングツールの最も重要な利用目的が知識創発と考えられることから、ツールの出力結果を見て有効な結果を積極的に集めること、ならびに集めた結果を収束させてアイデアの創発に積極的に繋げてもらうことを意図して設定した。

## 3.2 コイン

コインは、統合環境内のボタンとツールの利用、毎日の起動、テキストマイニングプロセスにおける結果と解釈の登録と知識創発、後述のミッションのクリア、によって得ることができる。

統合環境内のボタンとツールの利用と、テキストマイニングプロセスにおける結果と解釈の登録と知識創発、については、獲得できる経験値と同量のコインを得られる。

### 3.2.1 毎日の起動

継続的な利用を促すことを目的として、毎日起動されると獲得できるコインが増加するように設定した。連続起動で獲得できるコインを表 2 に示す。

表 3: ミッションの例

ランクを 500 にしよう
ランクを 1000 にしよう
チュートリアル「利用」の初心者、初級を全てクリアしよう
チュートリアル「利用」の文章推敲を全てクリアしよう
知識創発を 10 回行ってみよう
知識創発を 50 回行ってみよう
TETDM を 50 回起動しよう
TETDM を 100 回起動しよう
TETDM を 30 日間起動しよう
TETDM を 100 日間起動しよう
TETDM を連続 10 日間起動しよう
TETDM を連続 30 日間起動しよう
合計 50 時間利用しよう
合計 100 時間利用しよう
結果と解釈を 50 回登録しよう
結果と解釈を 100 回登録しよう
タイピング(色)で 200 点以上を出そう
処理ツールを 20 種類使用しよう
可視化ツールを 10 種類使用しよう
ツールを 1000 回セットしよう
ツールを 2000 回セットしよう
処理ツール内のボタンを 10 種類押そう
ツール内のボタンを 1000 回押そう
テキストエディタ内のボタンを 100 回押そう
スキルランクを 500 にしよう
スキルランクを 1000 にしよう

## 3.3 ランク

ランク (レベル) は、経験値が 100 増加するごとに 1 増える。多くのゲームで、ランクはユーザのやり込みの度合いを表す数値として用いられ、この値が高いことが一種のステータスとなっている。そのためランクを明示することで、ランクの上昇を目指して積極的に TETDM を利用してもらうことを意図した。

## 3.4 キャラクター

ゲーム内通貨のコインの利用先として、キャラクターおよびキャラクターのコスチュームを購入できる。キャラクターは、「キャラクターなんとか機」[7]を用いて 6 種類用意した。また、各キャラクター向けに 4 種類のコスチュームを用意した。これらの全キャラクターとコスチュームを集めることを目指して、継続的に TETDM を利用してもらうことを意図した。

## 3.5 ミッション

TETDM の利用に関する具体的な目標をミッションとして用意した。各ミッションを達成すると、ミッションごとに定められた経験値やコインが獲得できる。これにより、継続的に TETDM を利用してもらうことを意図した。用意したミッションの例を表 3 に示す。

### 3.6 TETDM ゲームモードのインタフェース

本節では TETDM ゲームモードのインタフェースについて述べる。図 2 に、TETDM の画面とそのゲームモード用のウィンドウを示す。

図上部の左側に、ランク (RANK)、経験値 (EXP)、コインが表示される。右側にはボタンが並べられ、「ショップ」ボタンで、キャラクターやコスチュームの購入、「キャラ」ボタンで、キャラクターやコスチュームの選択、「ミッション」ボタンで、ミッションの表示が行える。また画面中央上部には、選択中のキャラクターが表示される。

## 4 TETDM ゲームモード利用者の利用意欲向上の有効性検証実験

本章では、提案する TETDM ゲームモードが利用者の利用意欲向上につながるかを検証した実験について述べる。

### 4.1 実験手順

被験者は TETDM 利用経験がない大学生、大学院生の男女 16 名とし、これまでの家庭用ゲームやソーシャルゲームの経験が均等になるように、ゲームモードを利用する提案群と、ゲームモードを利用しない比較群の 8 名ずつの 2 グループに分けた。なお、事前アンケートでソーシャルゲームの利用経験を尋ねたところ、16 名中 10 名は現在もよくやっていると回答し、その他の 5 名は現在または昔時々やっていたと回答した。実験期間は 2 週間として、その期間中に行うべき必須課題と任意課題とを表 4 のように設定した上で、それらの課題を行う時間を含め、最低 5 時間は TETDM を利用するように指示を与えた。また、3.5 で述べたゲームモードのミッションについては、実験期間内での達成可能性が高くなるように数値を設定した。

### 4.2 実験結果と考察

#### 4.2.1 利用意欲の向上に関する結果

各被験者の実験期間中の TETDM の利用時間を図 3 に、利用時間の平均値を表 5 に示す。全体的に提案群の被験者の方が多く TETDM を利用しており、提案群と比較群の利用時間には有意差が見られた。このことから、ゲームモードを利用した提案群の方が、より意欲的に TETDM を利用していたことが伺える。

表 4: 被験者に与えた必須課題と任意課題

必須課題
(課題 1) 説明書を用いた TETDM の基本的な操作の習得 インターネット上のニュース記事などを コピー&ペーストして 20 分程度自由に操作する
(課題 2) 利用者向けチュートリアル「初心者」のクリア
(課題 3) 利用者向けチュートリアル「初級」のクリア
任意課題
(課題 1) 以下のようなテキストを入力として与えて さまざまなツール上で、その結果を確認 1) インターネット上のニュース記事 2) インターネット上の情報サイトの内容 3) 電子掲示板や SNS などに書き込まれた コメントやコメント集合 4) 自分や他人が書いたメール 5) 自分が書いたレポートなどの文章
(課題 2) 利用者向けチュートリアル「文章推敲」のクリア
(課題 3) さまざまなテキストを入力として与えて その結果と解釈の登録
(課題 4) 登録された結果と解釈をもとに 登録した内容をひとつの結果としてまとめる知識創発

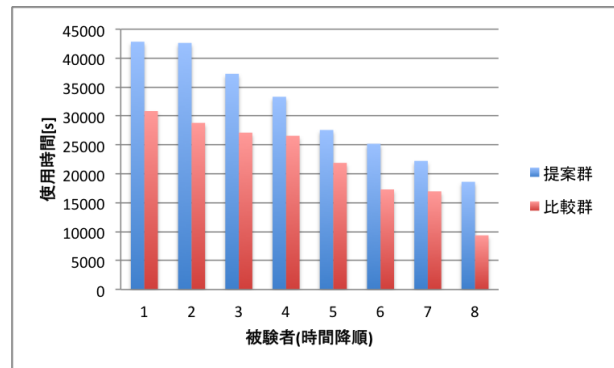


図 3: 各被験者の実験期間中の TETDM 利用時間

#### 4.2.2 ランクに関する結果

図 4 に各被験者の実験終了時のランクを、表 6 に実験終了時のランクの中央値を示す。図 4 から、提案群では 4 人の被験者がランク 100 を超える結果となったが、比較群の被験者は一人もランク 100 を越えなかった。また提案群の一人の被験者は、600 を越えるまでランクを上げていたことなどから、提案群の被験者については、ランクの表示によって、TETDM を意欲的に利用する意志があったと考えられる。

しかし、提案群と比較群全体としては、ランクの数値に有意差が見られなかった。提案群の下位 4 名のランクは、比較群の上位 4 名とあまり差がない。このことから、ゲームモードの存在に対して、すべての利用者が即座に積極的に利用するわけではないことが伺える。

#### 4.2.3 キャラクターに関する結果

表 7 に提案群の被験者によるキャラクターの購入状況を示す。4 名の被験者はすべてのキャラクターを集め

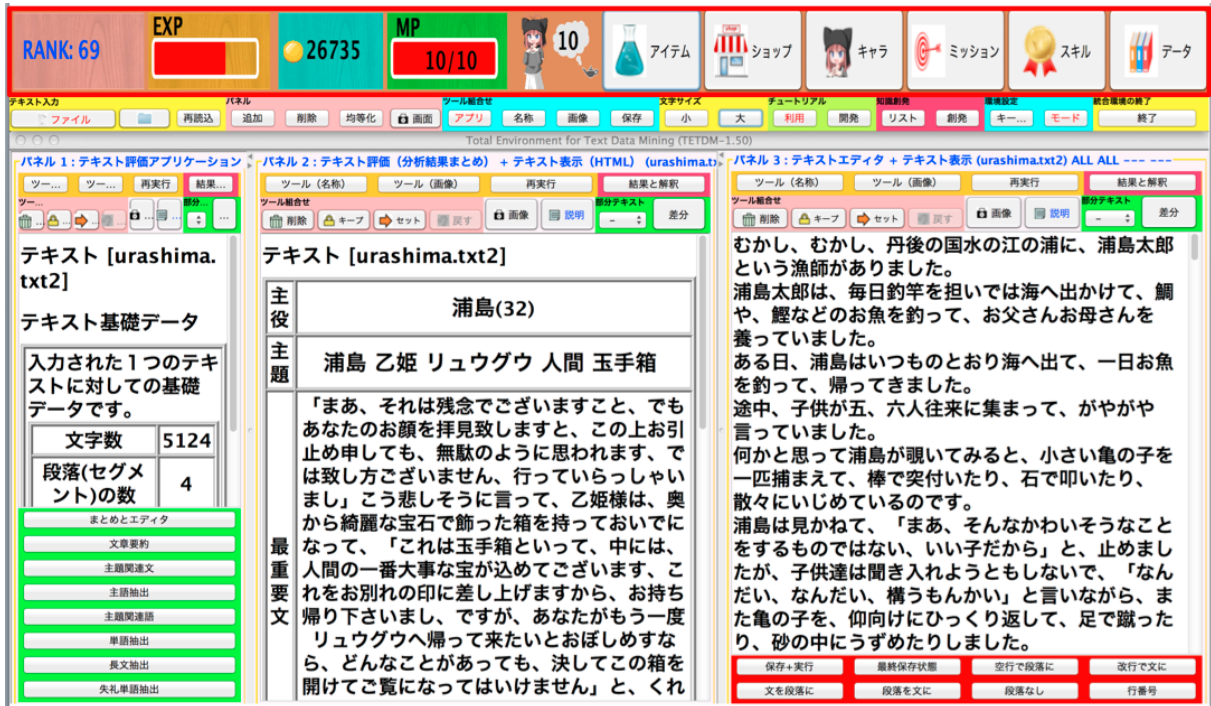


図 2: ゲームモード用のウィンドウ (画面上部)

表 5: TETDM 利用時間の平均値 (時間)

	提案群	比較群	t 検定
平均値	8.7	6.2	$p < .05$

表 6: 被験者の実験終了時のランクの中央値

提案群	比較群	t 検定
88.0	49.5	n.s.

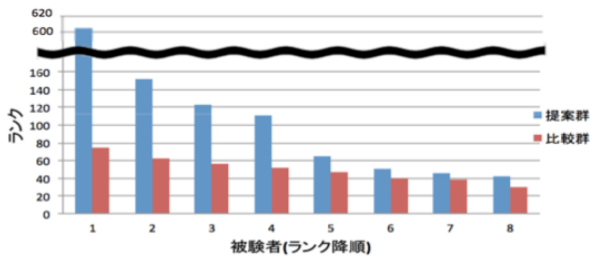


図 4: 被験者の実験終了時のランク

表 7: キャラクター (最大5) とコスチューム (最大18) の購入数

被験者	キャラクター	コスチューム
A	5	0
B	5	15
C	0	0
D	5	1
E	5	7
F	0	1
G	0	0
H	1	0
平均	2.63	3.00

たのに対して、他の4名の被験者はキャラクターには興味を示さなかった。このことから、一部の利用者にはゲームモードは強く効果を発揮する反面、興味がない利用者に対しては効果を上げにくいことがわかる。

#### 4.2.4 ミッションに関する結果

ミッションは提案群にのみ明示されているが、比較群の被験者についても、その操作履歴をもとに達成の

有無を判定した。

提案群、比較群ともに多くの被験者が達成できていたミッションは、必須課題の「初心者、初級チュートリアルをクリア」と、「処理ツールを20種類利用」「可視化ツールを20種類利用」「処理ツール内のボタンを10種類押す」という、普通に使っていれば容易に達成できるものとなった。

「10時間利用しよう」「ツールを1000回セットしよ

う)「ツール内のボタンを 500 回押そう」というミッションは、提案群のみが 3 名から 5 名達成できた課題となっており、目標を明示することで、その達成に向けて利用意欲が生じることが確認された。

#### 4.2.5 実験結果まとめ

今回の被験者の多くはソーシャルゲームの経験があり、提案群の半数の被験者は意欲的 TETDM を利用する結果となった。ゲームモードにより、強い利用意欲が生じた被験者は、ランク上げ、キャラクターの収集、ミッションの達成などに積極的に取り組むことが確認された。しかし、すべての被験者がゲームモードに興味を示すわけではないことも確認された。提案群の被験者のうち、ランクが 100 を越えた被験者の 4 名は、事前アンケートでソーシャルゲームを「現在もよくやっている」「現在も時々やっている」と 2 名ずつ回答した。しかしランクが 100 を越えなかった 4 名の被験者のうち 3 名は、ソーシャルゲームを「現在もよくやっている」、1 名は「全然やっていない」と回答していることから、ソーシャルゲームの利用経験が、必ずしも TETDM ゲームモードの利用意欲につながるわけではないことがわかった。

提案群の被験者に対して、利用意欲が十分に生じなかった理由としては、ゲームモードとしての作り込みが不十分で内容が単純であったことが一因として考えられる。そのため、より利用者の興味を引きやすいゲームモードの内容については、今後検討する必要があると考えている。

一方、ランクが 100 を越えた提案群の 4 名の被験者のうち、3 名は事前アンケートで、テキストマイニングという言葉で「まったく聞いたことがない」と回答していた。このことから、テキストマイニングに興味をなかった利用者に、新規の利用を促すという目的に対して、ゲームモードは一定の効果があったと考えられる。

## 5 結論

本論文では、TETDM の機能の一つとして、TETDM 利用者の利用意欲を向上させるためのゲームモードの開発を行った。評価実験により、提案するゲームモードが利用意欲の向上につながることを検証した。

今後は、より多くの利用者の興味を引けるように改良を加えるとともに、キャラクターを利用したテキストマイニングの支援を実装するなどにより、TETDM によるテキスト分析と、ゲームモードとの関連を強めていきたいと考えている。また、TETDM の未利用者だけでなく、多くの利用者が継続的に利用できるよう、

達成可能な目標を細かく設定することで、利用意欲の維持に繋げていきたい。

最終的に、利用者の拡大により、TETDM 内で動作するツールが幅広く開発され、また利用者の利便性が高まることで、世の中で利用されていないデータが積極的に活用されること、またデータの分析が一般的になることで、多くの利用者にとって、データを元にした根拠のある説明が可能になり、合理的な判断や行動に繋げられることを期待している。

## 参考文献

- [1] 砂山渡, 高間康史, 西原陽子, 梶並知記, 串間宗夫, 徳永秀和: 統合環境 TETDM を用いたマイニングツールの開発と利用の実践, 人工知能学会論文誌, Vol.29, No.1, pp.100-112, (2014)
- [2] Hewijin Christine, Jiau, Jinghong, Cox Chen, Kuo-Feng Ssu : Enhancing Self-Motivation in Learning Programming Using Game-Based Simulation and Metrics, TRANSACTIONS ON EDUCATION, Vol. 52, No. 4, pp.555-562, (2009)
- [3] 倉本 到, 片山 拓馬, 渋谷 雄, 辻野 嘉宏: 懐優館: 作業意欲を持続的に維持向上させる EELF に基づく主観的比較型エンタテインメントシステム, 情報処理学会論文誌, Vol.50, No.12, pp.2807-2818 (2009)
- [4] 益子博貴, 松村敦, 宇陀則彦: キャラクターを用いた学習継続支援ツールの開発, ARG 第 1 回 Web インテリジェンスとインタラクション研究会 (ARG SIG-WI2), (2012)
- [5] 渡辺博芳, 水柿恵: セルフラーニング型授業におけるエージェントキャラクタによる学習支援, 情報科学技術フォーラム一般講演論文集 (N. 教育・人文科学), Vol.3, No.4, pp.335-337, (2004)
- [6] 小島一秀: ゲームと融合した e ラーニング問題集の開発と評価, 第 11 回情報科学技術フォーラム講演論文集 第 3 分冊, RK-011, pp.109-114, (2012)
- [7] キャラクターなんとか機: (URL)  
<http://khemix.sakura.ne.jp/download.shtml>

# 分析結果とその解釈の統合を支援する 知識創発インタフェース

## Knowledge Creation Support by Integrating Analysis Results and Interpretations

西山 知志  
Satoshi Nishiyama

砂山 渡  
Wataru Sunayama

広島市立大学 情報科学部  
Faculty of Information Sciences, Hiroshima City University

**Abstract:** In the process of convergence of knowledge creation in text mining, by repeating the interpretation and integration of the collected analysis result, we are aiming to acquire a straightforward knowledge represented as cause-and-effect relationship. In this study, by the “multi-stage of interpretation” that make giving a repeat interpretation in one of the results, provide assistance leading to the acquisition of general knowledge by extending the width of the interpretation. In addition, by the “cause and result of labeling” to the interpretation, provide assistance to integrate knowledge as a cause-and-effect relationship.

### 1 はじめに

近年、情報技術の発展により、私たちはブログや SNS などを通して、テキスト形式で情報発信を手軽に行えるようになった。そうした中で、インターネット上には膨大なテキストデータが氾濫している状況にある。テキストデータには様々な情報が埋もれており、その中から活用できる有用な知識をいかにして獲得するのが、ビジネスの現場において課題となっている。例えば、企業の視点からテキストデータを利用する際に、顧客の声や風評を抽出した分析結果を、マーケティングや商品開発に活用することが求められている。ただ単に、テキストデータの分析結果を得ることにとどまるだけでは、知識の獲得につなげていくのは難しい。収集した分析結果に対して、それぞれが何を意味しているのかについて、効果的な解釈を行うことが必要になる。

そこで本研究では、テキストデータの分析結果に対する解釈の与え方として 2 つの方法を考えた。1 つ目の方法は解釈の多段化で、解釈を与える際に論理的思考の飛躍を促すことで解釈の幅を広げる支援を行う。2 つ目の方法は解釈への原因、結果のラベル付けで、解釈同士に潜む因果関係を捉える支援を行う。これらの支援をもとに、知識創発を促すことを目的とする。

本研究における知識創発とは、テキストマイニングツールを用いて収集したテキストデータの分析結果に対して、それらを総合的に解釈して 1 つの新たな知識に統合することを指す。多くの分析結果とその解釈が

得られても、まとまりがなければそれらを実際に活用できないので、根本的に信頼性が高い知識を得るために、解釈の統合を行う。本研究の目的が達成された場合、より幅広い解釈の中で解釈同士の因果関係を意識した知識創発を実現できる。

### 2 関連研究

テキストから知識を獲得するための研究として、テキストを対象とした評価情報の分析に関する研究 [1]、出来事の望ましさを判定を目的とした語彙知識獲得 [2]、軽卑表現の情報を活用した知識発見 [3] がある。以上に挙げた 3 つの研究に対して、テキストデータの分析結果から知識を獲得する点で類似しているが、本研究は、得られた複数の結果から最終的に 1 つの新たな知識へ結びつける点で異なる。また、テキストに含まれる文章やキーワード同士の因果関係に関する研究として、接続標識「ため」に基づく文書集合からの因果関係知識の自動獲得 [4]、共起パターンの学習による事象間関係知識の獲得 [5]、構文パターンを用いた因果関係の抽出 [6] がある。以上に挙げた 3 つの研究に対して、分析結果からテキスト内に含まれる因果関係を捉える点で類似しているが、本研究は、分析結果の解釈同士の因果関係を捉える点で異なる。



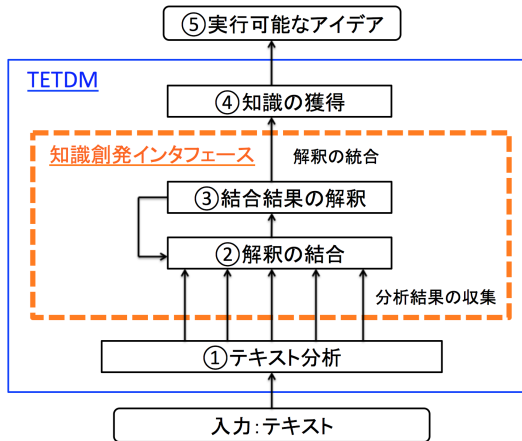


図 1: 従来の知識創発インタフェースの枠組み



図 2: 従来の知識創発インタフェースの画面

### 3 従来の知識創発インタフェース

#### 3.1 TETDM(テキストデータマイニングのための統合環境) への実装

本研究では、知識創発インタフェースを TETDM[7] 中の 1 つの機能として実装する。TETDM とは、テキストデータマイニングのための統合環境のことで、複数のテキストマイニングツールを組み合わせて統一的に扱うことができる。

#### 3.2 知識創発の 5 つのプロセス

従来の知識創発インタフェースの枠組みを図 1 に示す。知識創発を行う際には、TETDM にテキストを入力して最終的に得た知識を実行可能なアイデアにつなげていくまでに、図 1 の丸数字が示す 5 つのプロセスを踏むことになる。また、従来の知識創発インタフェースのシステム画面を図 2 に示す。

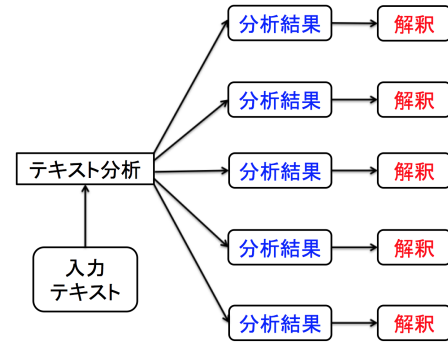


図 3: 知識創発における発散的思考



図 4: 共通点および解釈の記入欄 (赤枠)

#### 3.2.1 テキストの分析結果の収集

TETDM に分析したいテキストを入力する。テキストマイニングツールを用いてテキスト分析を行い、複数の分析結果を得る。このとき、それぞれの分析結果に対して、20 文字程度の文章で解釈を与える。

知識創発において、最終的に 1 つの知識を得るために分析結果を得る行為を繰り返すことを、発散的思考と呼ぶ。図 3 に発散的思考のイメージを示す。

#### 3.2.2 分析結果の解釈の結合

収集した分析結果のそれぞれの解釈の中から、共通点を含む解釈同士を探す。共通点を見つけた解釈同士は後に結合する。このとき、共通点の内容を図 4 赤枠上段のテキストフォームに記入する。

#### 3.2.3 共通点の解釈

解釈同士で見つけた共通点から考えられる解釈を、図 4 赤枠下段のテキストフォームに記入する。図 4 最下の結合ボタンを押すと、記入した共通点の解釈が新たに生成される。

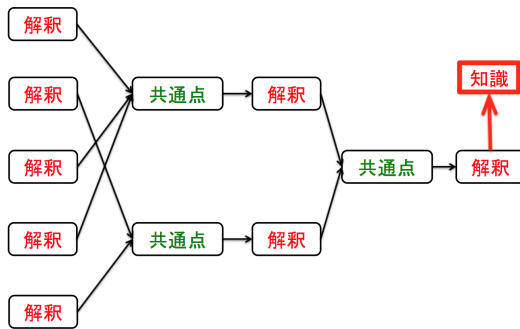


図 5: 知識創発における収束的思考

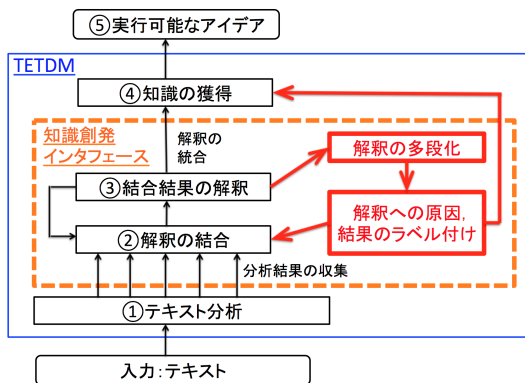


図 6: 提案する知識創発インタフェースの枠組み

### 3.2.4 全ての解釈の統合による知識の獲得

解釈の結合と共通点の解釈を繰り返し、最終的に 1 つに統合された解釈を新たな知識として獲得する。知識創発において、最終的に 1 つの知識を得るために全ての解釈を統合することを、収束的思考と呼ぶ。図 5 に収束的思考のイメージを示す。

### 3.2.5 得られた知識をもとにした実行可能なアイデアの考案

全ての解釈を統合して知識を獲得するだけにとどまってしまうと、その知識のままでは実際に役に立たない。そこで最後に、得られた知識をもとに実際に現場で活用できる、実行可能なアイデアを考案する必要がある。

## 4 提案する知識創発インタフェースのシステム構成

提案する知識創発インタフェースの枠組みを図 6 に示す。図 1 に示した従来の知識創発インタフェースに対して新たに、解釈の多段化、および、解釈への原因、結果のラベル付けの 2 つの機能を実装した。また、提



図 7: 提案する知識創発インタフェースの画面



図 8: 解釈の多段化のイメージ

案する知識創発インタフェースのシステム画面を図 7 に示す。

### 4.1 解釈の多段化

創発とは、部分の性質の単純な総和にとどまらない性質が、全体として現れることを意味する。収集した分析結果の解釈を、ただ単に共通点で整理しながら統合するだけでは、知識の創発にはつながりにくい。知識の創発を促すには、共通点の解釈の幅を広げながら統合していく必要がある。そこで、共通点に対する幅広い解釈を目指すために、解釈の多段化を考えた。図 8 に解釈の多段化のイメージを示す。

共通点に対して、連続して解釈を与えることを可能にした。まず、共通点に対して解釈 1 を与える。解釈 1 に対してさらに発展させた解釈 2 を与える。解釈 2 に対してさらに発展させた解釈 3 を与える。発展させた解釈のパターンは、例えば次の 3 つが考えられる。1 つ目は、「AKB48 が好き」を「アイドル好き」と解釈することで、「AKB48」が「アイドル」に抽象化されるパターンがある。2 つ目は、「スポーツを観戦する」を「サッカーを観戦する」と解釈することで、「スポーツ」が「サッカー」に具体化されるパターンがある。3 つ目は「友達が多い」を「フレンドリーな性格」と解釈することで、「友達が多い」が「フレンドリー」に言い換えが行われるパターンがある。

知識創発において、解釈を行う思考が堅実なままでは一般的な知識の獲得にとどまることになる。一方で、解釈を行う思考に論理的な飛躍が大きいほど、1 度の解釈だけではたどり着きにくかった解釈を得ることが

解釈1 :

解釈2 :

解釈3 :

図 9: 解釈の多段化を行うテキストフォーム

でき、解釈の幅が広がりやすくなる。解釈の多段化により、知識創発を行う上で、解釈の幅を広げることが期待できる。

#### 4.1.1 解釈の多段化のインターフェース

解釈の多段化のインターフェース部分を図9に示す。共通点から考えられる新たな結合結果の解釈を1段目のテキストフォームに記入する。2段目のテキストフォームには、共通点から考えられる2つ目の新たな結合結果の解釈、または、1段目に記入した解釈をさらに発展させた解釈を記入する。3段目のテキストフォームには、共通点から考えられる3つ目の新たな結合結果の解釈、または、2段目に記入した解釈をさらに発展させた解釈を記入する。

## 4.2 解釈への原因，結果のラベル付け

知識創発において価値ある創発を行うためには、収集した分析結果の本質に迫りながら、解釈を統合していくことが必要になる。そのためには、収集した分析結果の解釈から主要な知識を探る力が求められる(図10)。主要な知識とは、根本的な原因と結果の全体像を意味する。結果の全体像が分かると、分析結果から何が起きているのかを理解しやすくなる。また、全ての結果を引き起こした根本的な原因が分かると、知識を活用する明確な方針を立てやすくなる。最終的に、根本的な原因と結果の全体像の関係性を読み取ることで、分析結果の解釈同士に潜む因果関係を捉えることができる。このとき、未知の因果関係の創発結果が得られると、未知知識の発見につながる。一方で、既知の因果関係の創発結果が得られると、既知知識の理解や確認につながる。

そこで、解釈同士に潜む因果関係を意識した知識創発を目指すために、解釈への原因，結果のラベル付けを考えた。知識創発を行う際には、分析結果から何の知識を得たいのかについて、創発テーマを設定する。創発テーマに関して原因を表す解釈には原因ラベルを付ける。一方で、創発テーマに関して結果を表す解釈には結果ラベルを付ける。ラベル付けにより、原因，結果を表す解釈を整理しながら統合することが可能になる。原因ラベルがついた解釈同士を統合することで根本的な原因が分かりやすくなり、結果ラベルの付いた

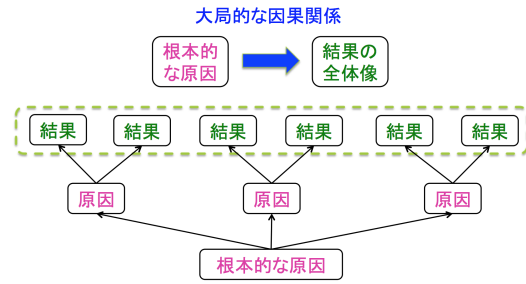


図 10: 主要な知識の探索



図 11: 解釈への原因，結果のラベル付けを行うチェックボックス

解釈同士を統合することで結果の全体像が分かりやすくなる。解釈同士に潜む因果関係を意識した知識創発が期待できる。

#### 4.2.1 解釈への原因，結果のラベル付けのインターフェース

解釈を記入するテキストフォーム右端にある「原因」と「結果」のチェックボックス(図11)をチェックすると、チェックした段の解釈に原因と結果のラベルが結合後に付与される。原因，結果のそれぞれラベル付けがされた解釈は図12の通りにラベルが表示される。3段記入した解釈のうち、創発テーマに関して原因を表す解釈の段には、原因のチェックボックスにチェックを入れる。原因のラベル付けがされた解釈は図13左の通り、システム画面の左側に表示される。一方で、創発テーマに関して結果を表す解釈の段には、結果のチェックボックスにチェックを入れる。結果のラベル付けがされた解釈は図13右の通り、システム画面の右側に表示される。

## 5 知識創発インターフェースの有効性検証実験

### 5.1 実験目的

本実験では、提案した知識創発インターフェースによって、次の2つの有効性を検証することを目的とする。



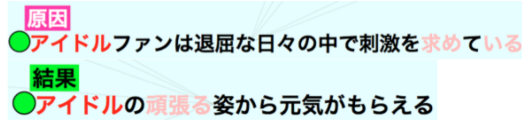


図 12: 原因, 結果のラベル付けがされた解釈



図 13: 解釈結合後の知識創発インタフェースの画面

1. 解釈の多段化により, 解釈の幅を広げることができたのか.
2. 解釈への原因, 結果のラベル付けにより, 解釈同士に潜む因果関係を意識した知識創発が行われたのか.

## 5.2 実験方法

テキストの分析結果の解釈 10 個に対して, 次の 2 つの創発テーマに従い各 2 回ずつ, 合計 4 回の知識創発を被験者に行ってもらった.

1. アイドルのファン心理について述べたテキストの分析結果の解釈 10 個を「ファンがアイドルに求めていること」をテーマに知識創発を行う.
2. ソーシャルゲームの魅力について述べたテキストの分析結果の解釈 10 個を「ユーザがソーシャルゲームに求めていること」をテーマに知識創発を行う.

比較システムは, 解釈の多段化および解釈への原因, 結果のラベル付けの, 2 つの機能を除いたインタフェース (図 2) を用意した. 被験者は大学生 10 人を A と B の 2 グループに分けた. 各被験者が 4 回の知識創発を行った際の創発テーマと使用したシステムについて表 1 に示す.

表 1: 各被験者が知識創発を行った順番

	A グループ 5 人	B グループ 5 人
1 回目	テーマ 1 提案システム	テーマ 1 比較システム
2 回目	テーマ 1 提案システム	テーマ 1 提案システム
3 回目	テーマ 2 比較システム	テーマ 2 提案システム
4 回目	テーマ 2 提案システム	テーマ 2 提案システム

表 2: 解釈の回数 (平均)

	比較システム	提案システム
解釈 1	7.2	6.5
解釈 2	-	5.3
解釈 3	-	3.4
計	7.2	15.2

## 5.3 実験結果と考察

### 5.3.1 解釈の多段化による有効性の実験結果と考察

表 2 に, 各段階ごとの解釈の回数の平均を調べた結果を示す. 表 2 より, 比較システムよりも提案システムの方が解釈を行った回数が増えた. これは, 連続した解釈を可能にしたことで, 1 つの解釈に対して解釈を与える回数が増えたためと考えられる.

次に, 最初に用意された分析結果の文章中に含まれていない名詞が, 解釈を行った文章中にどれだけの種類が新たに登場したかを調べた. 表 3 に, その結果を示す. 表 3 より, 比較システムよりも提案システムの方が, 解釈で新たに登場した分析結果に含まれない名詞数が多くなった. これは, 解釈を重ねることで, 論理的思考の飛躍を促せたためと考えられる.

テーマ 1 において, 解釈の多段化が行われた具体例を挙げる. 1 段階目で「下手でも一生懸命に歌やダンスをする姿がいい」と解釈された後に, 2 段階目で「努力によってパフォーマンスに成長があれば嬉しい」と解釈され, さらに 3 段階目で「ステップアップの過程が見られて更に応援したくなる」と解釈された. このとき, 1 段階目から 2 段階目の解釈で, 「一生懸命」が「努力」に言い換え, 「歌」や「ダンス」が「パフォーマンス」に抽象化, 「いい」が「嬉しい」に具体化が行われた. また, 2 段階目から 3 段階目の解釈で, 「成長」が「ステップアップの過程」に言い換え, 「嬉しい」が「応援したくなる」に具体化が行われた.

以上の実験結果より, 解釈の回数が増えると, 解釈で新たに登場した分析結果に含まれない名詞数が多くなるのが分かった. これは, 解釈の多段化により, 連続した解釈を行うことで, 1 度の解釈だけではたどり着きにくかった新たな解釈が生まれやすくなったためと考えられる. このことから, 提案システムに実装した解釈の多段化により, 解釈の幅を広げられる有効性が示された.

表 3: 解釈で新たに登場した分析結果に含まれない名詞数 (平均)

	比較システム	提案システム
解釈 1	6.4	5.8
解釈 2	-	3.0
解釈 3	-	2.1
計	6.4	10.9

表 4: ラベルが付与された回数と付与率 (平均)

	原因	結果	ラベル付与率
1 回目	2.8	2.6	92.4
2 回目	3.1	3.0	95.5
3 回目	3.4	3.5	97.3

### 5.3.2 解釈への原因, 結果のラベル付けによる有効性の実験結果と考察

表 4 に, 提案システムにおいて, 原因と結果のそれぞれのラベルが付与された回数と, 共通点で結合した回数に対するラベルの付与率を調べた結果を示す. 表 4 より, 提案システムを 3 回使用したいずれの場合においても, 9 割を超える共通点に対する解釈にラベルが付与されていた. これは, 解釈同士の因果関係を探るべく積極的に原因と結果のラベルが使用されたためと考えられる.

次に, 提案システムにおいて, 分析結果を最終的に原因ラベルが付いた解釈 1 個および結果ラベルが付いた解釈 1 個として統合した被験者数を調べた. 表 5 に, その結果を示す. 表 5 より, 提案システムを 3 回使用したいずれの場合においても, 7 割を超える被験者が原因ラベルが付いた解釈 1 個および結果ラベルが付いた解釈 1 個として統合していた. これは, 解釈に原因ラベルを付けたことで, 原因を表す解釈が明示され, 根本的な原因を発見しやすくなったためと考えられる. また, 解釈に結果ラベルを付けたことで, 結果を表す解釈が明示され, 結果の全体像が分かりやすくなったためと考えられる.

テーマ 2 において, 原因ラベルが付いた解釈 1 個と結果ラベルが付いた解釈 1 個に統合された因果関係の具体例を挙げる. 収集した分析結果が最終的に, 原因ラベルが付いた解釈「有名なアニメやゲームとのコラボ企画が楽しい」と, 結果ラベルが付いた解釈「中毒性がありハマると毎日起動する」に統合された. この因果関係から「面白いイベント企画を毎日発生させて飽かさせない工夫をする」実行可能なアイデアが考案された.

以上の実験結果より, 原因と結果のラベルが積極的に使用され, 因果関係を捉えられる形でまとめられていたことが分かった. これは, ラベル付けにより, 解釈を原因と結果に整理することで, 因果関係に基づい

表 5: 最終的に原因 1 個, 結果 1 個として統合した被験者数 (/10 人)

	被験者数
1 回目	7
2 回目	9
3 回目	9

た統合を行いやすくなったためと考えられる. このことから, 提案システムに実装した解釈への原因と結果のラベル付けにより, 解釈同士に潜む因果関係を意識した知識創発が行われる有効性が示された.

## 6 結論

本研究では, 「解釈の多段化」「解釈への原因, 結果のラベル付け」を用いた知識創発インタフェースを提案した. 検証実験により, 解釈の幅を広げ, 因果関係を意識した知識創発を促す有効性を確認した. 今後は, 創発された因果関係の価値を自動評価する方法について検討していきたい.

## 参考文献

- [1] 乾孝司, 奥村学: テキストを対象とした評価情報の分析に関する研究動向, 自然言語処理, Vol.13, No.3, pp.201-241 (2006)
- [2] 乾孝司, 乾健太郎, 松本裕治: 出来事の望ましさ判定を目的とした語彙知識獲得, 言語処理学会第 10 回年次大会発表論文集, pp.91-94 (2004)
- [3] 荻野紫穂, 那須川哲哉, 金山博, 榎美紀: 軽卑表現の情報を活用した知識発見, 言語処理学会第 18 回年次大会発表論文集, pp.58-61 (2012)
- [4] 乾孝司, 乾健太郎, 松本裕治: 接続標識「ため」に基づく文書集合からの因果関係知識の自動獲得, 情報処理学会論文誌, Vol.45, No.3, pp.919-933 (2004)
- [5] 阿部修也, 乾健太郎, 松本裕治: 共起パターンの学習による事態間関係知識の獲得, 自然言語処理, Vol.16, No.5, pp.79-100 (2009)
- [6] 坂地泰紀, 竹内康介, 増山繁, 関根聡: 構文パターンを用いた因果関係の抽出, 言語処理学会第 14 回年次大会論文集, pp.1144-1147 (2008)
- [7] 砂山渡, 高間康史, 西原陽子, 徳永秀和, 串間宗夫, 阿部秀尚, 梶並知記: テキストデータマイニングのための統合環境 TETDM の開発, 人工知能学会論文誌, Vol.28, No.1, pp.1-22 (2013)

# 社会情勢の変化を表す表現の自動収集と可視化の検討

## Automatic Collection and Visualization of the Expressions that Represent Changes of Situations in the Society

尾亦 智弘<sup>1\*</sup> 渋木 英潔<sup>2</sup> 森 辰則<sup>2</sup>  
Tomohiro Omata<sup>1</sup> Hideyuki Shibuki<sup>2</sup> Tatsunori Mori<sup>2</sup>

<sup>1</sup> 横浜国立大学大学院環境情報学府

<sup>1</sup> Graduate School of Environment and Information Sciences, Yokohama National University

<sup>2</sup> 横浜国立大学大学院環境情報研究院

<sup>2</sup> Graduate School of Environment and Information Sciences, Yokohama National University

**Abstract:** In this paper, we discuss automatic collection and visualization of the expressions that represent changes of situations in our society. In creating new products or making a new plan, it is necessary to grasp the change in social situation. There are various ways in which we may find changes of situation in our society. For some of them, such as statistical manners, we have to begin with discussion of how to interpret them. We focus on expression on Web documents that seem to be written as authors' own interpretation of changes of situations in our society. We propose a method to collect them automatically from the Web documents. We also discuss a visualization method to help users to understand the collected expressions easily.

### 1 はじめに

ある技術を活用し新しいものを作る時や、企業などが新企画を立てる際には、それがどれだけ社会のニーズに答えられるかが重要となる。そのため、社会がどのような状況にあるのか、どのような変化が起きているのかといった社会情勢の変化を把握することが必要である。しかし、社会情勢の変化が現れる対象は多岐にわたる上に、統計情報のように解釈の仕方から議論しなければならないような情報もある。そのため、社会情勢の調査を支援するような技術が求められている。

しかし、そのような状況にありながら、現在、社会情勢の変化を収集するような手法は確立されていない。そのため、ニュースや新聞を見るといった手作業での調査が行われている。そこで本稿では、社会情勢の調査時の負担軽減を目的とし、Web上の文書群から、社会情勢の変化を表すものであると書き手が解釈しているであろう表現を自動的に収集する手法を検討する。また、収集した社会情勢の変化を表す表現を利用者に見やすく提示するために可視化についての検討も行う。ここで、社会情勢の変化を表す表現とは、「少子化」や「女性の社会進出」など、我々の社会の状態の変化を指し示す名詞や名詞句である。

社会情勢の変化を表す表現を含む文にはどのようなものがあるか調査を行ったところ、「少子化といった時代の変化…」や「少子化などの世の中の流れ…」といった文が見られた。この文では、「といった」や「などの」といった、例示の表現によって、「時代の変化」や「世の中の流れ」といった、社会情勢の変化の具体事例が属するであろう上位クラスに対して、その実例として「少子化」という表現により指し示されるコトが存在すること示されている。つまり、この文は「<社会情勢の変化の実例を指し示す表現><例示の表現><社会情勢の変化の具体事例が属する上位クラスを指し示す表現>」という構成をしている。このことから、社会情勢の変化の具体事例が属する上位クラスを指し示す表現から、社会情勢の変化の実例を指し示す表現を抽出するという、上位下位関係の抽出によって、社会情勢の変化を表す表現が収集できるのではないかと考えた。また、上位クラスを指し示す表現の種類を増やすことで、より多くの社会情勢の変化を表す表現が獲得できると考えられる。そのため、ブートストラッピング法により、上位クラスを指し示す表現を拡充させることを考える。本稿では、これら考えに基づき、社会情勢の変化を表す表現の収集手法を検討する。

\*連絡先： 横浜国立大学大学院環境情報学府  
〒 240-8501 神奈川県横浜市保土ヶ谷区常盤台 79-1  
E-mail: omata.t@forest.eis.ynu.ac.jp

## 2 関連研究

本節では提案手法に用いた要素技術について述べる。

### 2.1 定型表現を用いた上位下位関係抽出手法

大量のテキストデータから用語間の関係を判別する手法はこれまでに数多く提案されている。X を上位語、Y を下位語としたときに、「Y などの X」といった定型表現を用いる手法が有力な手法として存在している [1, 2]。上記の定型表現を用いることで、X と Y は上位下位関係にあるということが判別できる。このほかにも、安藤ら [3] は、「Y といった X」「Y のような X」などのパターンも上位下位関係を判定するために有力な定型表現であることを分析している。

本研究では、「Y などの X」「Y といった X」「Y のような X」の 3 つ定型表現を抽出に用いる。

### 2.2 ブートストラッピング法

ブートストラッピング法 [4, 5] とは、獲得対象となるクラスのインスタンスをシードとして与え、コーパスからインスタンスと共起するパターンを抽出し、抽出した共起パターンを用いて新たなインスタンスを抽出する。といった手順を反復的に繰り返すことで、少数のインスタンスから大規模なインスタンス集合を再帰的に獲得する手法である。ブートストラッピング法は、語義曖昧性解消、固有表現抽出および関係抽出など自然言語処理の様々なタスクに利用されている。本研究では、ブートストラッピング法を関係抽出に利用する。

ブートストラッピング法では、反復処理を繰り返していくうちにシードインスタンスとは関係のないインスタンスを抽出してしまう「意味ドリフト」という問題がある [6]。これは、ブートストラッピング法の反復過程において再現率が高いが適合率が低いパターン(ジェネリックパターン)を抽出してしまうことに起因する現象である。

意味ドリフトを回避する方法としては、ジェネリックパターンを抽出する前に反復を打ち切ることがあるが、反復を停止する最適な回数はタスクにより異なり、事前に決定することは困難である。Esspresso アルゴリズム [7] は精巧なスコアリング関数を用いて相互再帰的にインスタンスとパターンのスコアを定義し、意味ドリフトによる問題を軽減している。

### 2.3 Espresso アルゴリズム

Espresso アルゴリズムのスコアリング関数は、信頼度の高いパターンと頻繁に共起するインスタンスは信頼度が高く、信頼度の高いインスタンスと頻繁に共起するパターンは信頼度が高いという考えに基づいている。パターン  $p$  とインスタンス  $i$  のスコアはそれぞれ  $r_{\pi}(p)$  と  $r_l(i)$  であり、以下の式を用いて信頼度を計算する。

$$r_{\pi}(p) = \frac{1}{|I|} \sum_{i \in I} \frac{pmi(i,p)}{\max pm_i} r_l(i) \quad (1)$$

$$r_l(i) = \frac{1}{|P|} \sum_{p \in P} \frac{pmi(i,p)}{\max pm_i} r_{\pi}(p) \quad (2)$$

$$pmi(i,p) = \log \frac{|i,p|}{|i,*||*,p|} \quad (3)$$

$P$  はパターン集合、 $I$  はインスタンス集合であり、 $|P|$  と  $|I|$  はパターンとインスタンスの数を表す。 $|i,*|$  はインスタンス  $i$  の頻度、 $|*,p|$  はパターン  $p$  の頻度を表す。 $|i,p|$  はインスタンス  $i$  とパターン  $p$  が共起する回数である。 $pmi$  はインスタンスとパターン間の自己相関情報量を表しており、 $\max pm_i$  は全てのインスタンスとパターンの組み合わせの間における  $pmi$  の最大値である。なお、 $r_{\pi}(p)$  と  $r_l(i)$  の初期値はそれぞれ 1 である。

Espresso アルゴリズムでは、反復過程において (1) 式と (2) 式を適用することで、精度を高く保ちながら再現率を大幅に向上させている。

### 2.4 自己組織化マップ (SOM)

Kohonen によって提案されたニューラルネットワークモデルのひとつであり、トポロジカルマッピングによる教師なし学習を行うことで、多次元の属性値から 2 次元マップを生成し、視覚化することで、効果的なデータ分類手法として注目されている。自己組織化マップ上では類似度の高いデータどうしは近くに、類似度の低いデータどうしは遠くに配置される。[8]

## 3 社会情勢の変化を表す表現

本節では、本研究で抽出対象とする、社会情勢の変化を表す表現について述べ、それが現れる文章にはどのようなものがあるか調査する。そして、どのようにしたらそれを収集できるのかについて考察する。

### 3.1 社会情勢の変化を表す表現

社会情勢の変化は、注目しているある時点での社会の状態と、それ以前の社会の状態の間に違いが見られるものであり、その上で、その状態が変わっていく方向が見られるものが社会情勢の変化を表す表現だと言える。例としては、「少子化」や「女性の社会進出」などのものである。「少子化」では、注目している時点において、子供の数が少なく、それ以前の時点ではそれよりも数が多かったことが分かり、それから、「化」という文字によって、多い状態から少ない状態への推移が分かる。なお、どのような変化を表す表現を抽出すべきかは利用者のニーズによる。そのニーズとして、今の社会情勢の変化が知りたいということが多いため、過去の内容を指し示す表現は対象にしない。

### 3.2 社会情勢の変化を表す表現が現れる文章

次に、社会情勢の変化を表す表現であることが明示されている文章について調査した。その結果を「少子化」を例に、以下に記す。

#### ① 上位クラスの例示として示されているもの

若者の都市部集中、車離れ、晩婚化、少子化といった時代の流れとは別の経済活動や行動様式を持っている。

#### ② 直後の表現によって変化であることが示されるもの

急速な少子化の進行とそれに伴う人口減少は、社会経済全般にわたり、さまざまな影響を及ぼすことが想定されます。

①では、「時代の流れ」という、社会情勢の変化の具体事例が属する上位クラスを指し示すものの表現があり、その中の例示として「少子化」が示されている。②では、「進行」というある物事が進んでいることを表す表現によって、「少子化」が変化であることが示されている。これらのうち①が抽出に有用な表現だと考えられる。その理由としては、「といった」という例示であることを表す表現によって「<社会情勢の変化の実例を指し示す表現><といった><社会情勢の変化の具体事例が属する上位クラスを指し示す表現>」という構造をしていることが挙げられる。②は、直後の表現の膨大さや、文の意味を解釈する必要があるといったことが考えられる。そのため、本研究では①のように、上位クラスを指し示す表現の例示として示されるものを対象に収集を行う。

### 3.3 社会情勢の変化を表す表現の収集に向けたアプローチ

前小節にて、上位クラスの例示として示されるものを対象に収集を行うことを述べた。本小節では、これについてさらに調査し、どのようにして社会情勢の変化を表す表現の収集に取り組むのかというアプローチについて述べる。前小節の①の例に出てきた、例示を表す「といった」と、社会情勢の変化の具体事例が属する上位クラスを指し示す表現の「時代の流れ」の2つを組み合わせた「といった時代の流れ」という表現で検索を行ったところ、以下のような①とは違う文章が得られた。

グローバル化、市場の変化、販売方法の多様化といった時代の流れ、スピードは今後益々想像を絶する勢いで推移することでしょう。

この文章からは、「グローバル化」「市場の変化」「販売方法の多様化」が社会情勢の変化を表す表現として得られる。このため、「<例示の表現><社会情勢の変化の具体事例が属する上位クラスを指し示す表現>」という定型表現を検索質問として検索することで、定型表現の前に現れる、社会情勢の変化を表す表現が得られるのではないかと考えられる。

例示の表現は「といった」のほかにも「などの」「のような」といったものが考えられこれらについても、「時代の流れ」と組み合わせた表現を用いて検索を行い、以下の文章を得た。

少子・高齢化や高度情報化、グローバル化などの時代の流れにより、個人の価値観の多様化や行動の変化が進むほか、雇用システムの変化、社会保障制度の見直し、環境問題への取り組み、男女共同参画の推進など、社会経済システムの改革が迫られている。

文明は、生活環境の大幅な変化、或は産業革命のような時代の流れがあった時に使われているように思います。

これらの例からも、社会情勢の変化を表す表現を得ることができる。

また、例示の表現の場合と同様に、社会情勢の変化の具体事例が属する上位クラスを指し示す表現も「時代の変化」「世の中の流れ」のように、「時代の流れ」以外の表現が考えられる。これについて、「といった」と組み合わせた表現を用いて検索を行い、以下の文章を得た。

社会が急激に変化する中、我が国の教育も、知識基盤社会、グローバル化、人口減少社会といった時代の変化に即した対応が求められており、教育を支える教員についても新たな時代にふさわしい資質能力を備える必要がある。

我々は皆様のよき相談相手となれるよう、日々研鑽するとともに、コンプライアンスや内部統制といった、世の中の流れに合った支援をさせていただきます。

これらの例から、これまでとは異なる社会情勢の変化を表す表現が獲得できていることが分かる。

以上の調査から、「＜例示の表現＞＜社会情勢の変化の具体事例が属する上位クラスを指し示す表現＞」という定型表現を用いた社会情勢の変化を表す表現の抽出は有効であると考えられる。また、例示の表現と社会情勢の変化の具体事例が属する上位クラスを指し示す表現の組み合わせを変えて定型表現の種類を増やすことで、様々な社会情勢の変化を表す表現が抽出できると考えられる。これらの考えに基づき、再帰的処理によって少数のシードから、多くのインスタンスを獲得するブートストラッピング法を用いることを考える。ブートストラッピング法では、まず、抽出したい表現 X が属すクラス C を考えたとき、「X といった C」のような表現に注目し、C を固定しつつ、X に対応する表現を集める。その後、C の種類を増やすために、新たに得られた表現 X' を用いて、「X' といった C'」のような表現を集め、C' の候補を増やす、ということを反復的に繰り返すことによって、少数のシードから徐々に、X ならびに C に対応する表現を拡充させている。

## 4 定型表現を用いたブートストラッピング法による社会情勢の変化を表す表現の収集

本節では、前節にて述べたアプローチに基づく提案手法について述べる。

### 4.1 概要

本研究では、以下のステップにより社会情勢の変化を表す表現の収集を行う。

- ① Espresso アルゴリズムによる社会情勢の変化を表す表現の獲得

- ② Espresso アルゴリズムによる社会情勢の変化を表す上位クラス表現の候補の獲得

- ③ 人手による②で得られた表現のフィルタリング

- ④ ①②③を繰り返す

図 1 に上記ステップの流れを概略図として示す。上記ステップの詳細について次節で述べていく。

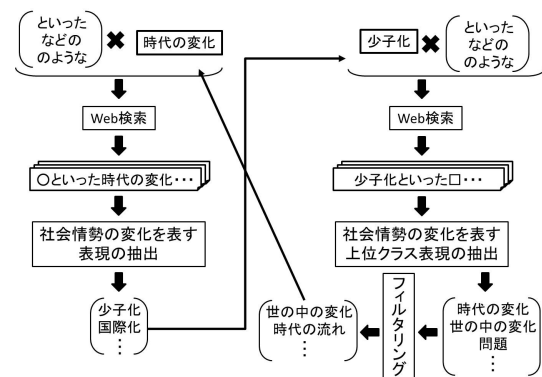


図 1: プロセス概要

### 4.2 Espresso アルゴリズムによる社会情勢の変化を表す表現の獲得

このステップでは、入力として社会情勢の変化の具体事例が属する上位クラスを指し示す表現を与え、例示の表現を組み合わせて定型表現を作成する。作成した定型表現をクエリに Web 検索を行い、定型表現を含む文章を得る。本研究では、例示の表現として「といった」「などの」「のような」の 3 つの表現を用いた。つまり、クエリは「といった○○」「などの○○」「のような○○」(○○は入力として与えた表現)となる。また、Web 検索には Bing<sup>1</sup> を用いた。

得られた文章から検索に用いた定型表現を含む 1 文を抜き出し、KNP により係り受け解析を行う。KNP による解析の結果、定型表現にかかっていると解析されたものを社会情勢の変化を表す表現として抽出する。その際、「少子・高齢化や高度情報化、グローバル化などの時代の流れにより...」の例のように、社会情勢の変化を表す表現が並列していることがある。そのため、句点「、」、助詞「や」、助詞「と」で区切って抽出をする。上記の例では、「少子・高齢化」「高度情報化」「グローバル化」に分けられて抽出される。

その後、Espresso アルゴリズムのスコアリング関数を用いて、得られた表現の信頼度を計算し、その上位 3

<sup>1</sup><http://www.bing.com/?cc=jp>

件を次のステップである、社会情勢の変化の具体事例が属する上位クラスを指し示す表現を抽出する際の入力とする。実験的に抽出を行ったところ、「現在」「私」「今」などのような社会情勢の変化を表す表現ではなく、高頻度で現れるため Espresso アルゴリズムの信頼度が高くなってしまった語があった。そのため、これらの語を排除するため、2文字以下の語は信頼度を計算せず、入力にならないようにした。

### 4.3 Espresso アルゴリズムによる社会情勢の変化の具体事例が属する上位クラス表現の候補の獲得

このステップでは、入力として社会情勢の変化を表す表現を与え、例示の表現と組み合わせて定型表現を作成する。前のステップと同様に、例示の表現としては「といった」「などの」「のような」の3つの表現を用いる。そのため、クエリは「△△といった」「△△などの」「△△のような」(△△は入力として与えた表現)となる。その後、作成した定型表現をクエリに Web 検索を行い、定型表現を含む文章を得る。

得られた文章から検索に用いた定型表現を含む1文を抜き出し、JUMANにより形態素解析を行う。解析の結果、定型表現の後に続く名詞句を社会情勢の変化の具体事例が属する上位クラス表現として抽出する。本研究では、名詞、助詞「の」、接尾辞、連体詞、形容詞のいずれかが続く限り名詞句と見なしている。

その後、Espresso アルゴリズムのスコアリング関数を用いて、得られた表現の信頼度を計算し、表現の信頼度の順位付けを行う。実験的に抽出を行ったところ、「問題」「課題」「私たち」のような意味ドリフトを起こしやすい表現が Espresso アルゴリズムでの信頼度が高くなっていた。そのため、3文字以下の語は入力を計算せず、入力にならないようにした。

### 4.4 人手による表現のフィルタリング

Espresso アルゴリズムによる社会情勢の変化を表す表現獲得の予備実験を行ったところ、反復を繰り返すにつれ社会情勢の変化とは関係のない表現が多く獲得された。これは、1度目の反復の時点で、「取り組み」や「キーワード」のような意味ドリフトを起こすインスタンスが選択されるため、後の反復で意味ドリフトが起き、関係のないインスタンスばかりが獲得されるためである。そこで、各反復において社会情勢の変化を表す上位クラス表現に対して、人手でフィルタリングをかけることを考える。獲得されるインスタンスは300程度あり、その中から社会情勢の変化を表す表現かどうかを確認するのは大きなコストがかかる。これに

対し、人手でのフィルタリングで確認する件数としては、毎回数件を選べばよいとため、各回で10件もない。そのため、社会情勢の変化を表す表現かどうかを確認するコストよりも小さいといえる。そのため、人手でのフィルタリング処理を加えても、十分に本研究の目的を果たせると考えられる。人手でのフィルタリングにより、関係のないインスタンスを取り除き、意味ドリフトの問題を抑えることができるようになることが期待される。

人手によるフィルタリングは、Espresso アルゴリズムによって順位付けされた表現を上位から見ていき、1度用いた表現は選択しないという条件で行い、3件の表現を獲得するまで行うことによって行う。

## 5 評価実験

4章で述べた手法により、社会情勢の変化を表す表現がどの程度獲得できるのかを確認するために実験を行った。

### 5.1 実験方法

加藤らの研究[9]の実験方法を参考にした。社会情勢の変化の具体事例が属する上位クラス表現のシードとして、「時代の変化」「時代の流れ」「世の中の変化」「世の中の流れ」の4つの表現を与え、ブートストラップによる抽出の反復を3回行う。そして、社会情勢の変化を表す表現として得られた文字列の集合に対して人手で評価を行った。なお、本実験の Web 検索には Bing を用い、1回の検索で取得する Web 文書数は50件とした。

### 5.2 実験結果

本研究では Web 文書を対象にしているため、再現率を本質的に求めることはできない。そのため、ブートストラップを3回繰り返して得られた候補に対して人手で確認を行い、そのなかで社会情勢の変化を表す表現であると判断されたものを全正解集合と仮定する。そのときの適合率と再現率のグラフを図2に、反復回数と、得られた表現の異なり数、正解の異なり数の増加の様子を図3に示す。また、抽出された社会情勢の変化を表す表現の正解例とその上位クラス表現の例を、それぞれ表1と表2に示す。なお、図2における、0回目とは、入力のシードで抽出されたものを指す。

<sup>2</sup>企業の社会的責任 (corporate social responsibility) 企業が利益を追求するだけでなく、組織活動が社会へ与える影響に責任をもち、あらゆるステークホルダー (利害関係者: 消費者、投資家等、及び社会全体) からの要求に対して適切な意思決定をすることを指す。



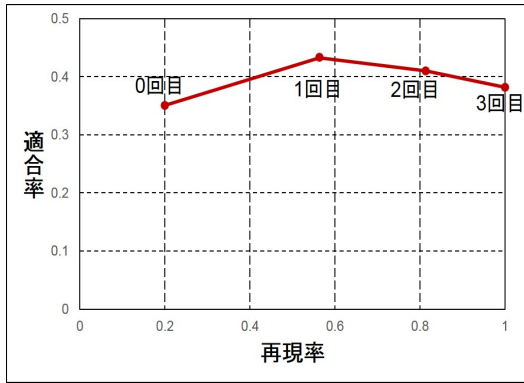


図 2: 収集した社会情勢の変化を表す表現の適合率-再現率グラフ

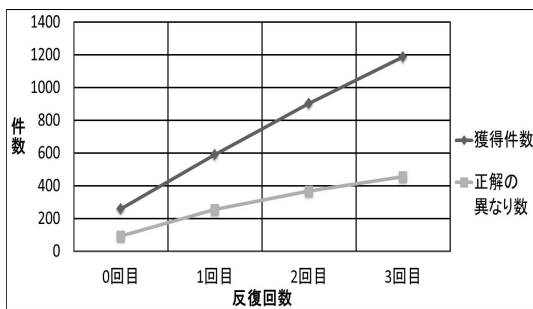


図 3: 獲得件数と正解の異なり数

表 1: 抽出された社会情勢の変化を表す表現の正解例

社会情勢の変化を表す表現	
少子化	高齢化
グローバル化	情報化
女性の社会進出	国際化
人口減少	景気の低迷
経済成長	核家族化
少子・高齢化	技術の進展
CSR <sup>2</sup> の広がり	コモディティ化
産業構造の変化	共働き世帯の増加

表 2: 抽出された社会情勢の変化を表す上位クラス表現の例

社会情勢の変化の具体事例が属する上位クラス表現
環境変化
社会情勢の変化
社会構造の変化
環境の変化
社会の変化
社会環境の変化

## 6 考察

図 2, 3 から, 1 回目の反復で再現率を大幅に上昇させ, その後も多くの正解を得ていることが分かる. また, 適合率が反復を重ねても大きく低下するようなことは起こっていない. これは, 人手でのフィルタリングが効果的に働いたためと考えられる.

しかし, 適合率が 40 % 程度とやや低調な値であった. 誤抽出された表現に対して分析をしたところ, 原因は以下の 4 つに大別された.

### ① 明らかに現在の社会情勢を表していないもの

- 1917 年のロシア革命によるソヴィエト政権の成立
- 大日本帝国の滅亡
- 武士の台頭

### ② 比喩の表現として用いられていたもの

- 今の濁流  
 もとの文章: 今の濁流のような世の中の流れに取り残されようとしてる日本と私(極めて私的な私)がいます。
- この怒涛  
 もとの文章: この怒涛のような時代の変化の渦にのみ込まれて沈まないように、本年は、時代の流れを見据えた運営計画を立案し着実に実行していきたいと思ひます。

### ③ 係り受けの誤りによるもの

- 最短でその日のうちに自宅まで届けるサービスで  
 もとの文章: ネットスーパーは、インターネット上で注文した生鮮品や食料品・日用品を、最短でその日のうちに自宅まで届けるサービスで、働く女性や高齢化世帯の増加、インターネットの普及といった社会環境の変化などで、利用者が急速に拡大している。

### ④ 並列して述べられていたものが区切ったことにより社会情勢の変化であることが分からなくなったもの

- 個人の価値観  
 もとの文章: 当社では、賃貸住宅事業を住宅事業のもうひとつの柱とすべく、「三井の賃貸」ブランドのもと、個人の価値観やライフスタイル、家族構成の多様化といった社会情勢の変化を踏まえた、都心部における良質な賃貸住宅供給を目指した“パーク

「アクシス」シリーズを中心に事業展開しております。

①について、抽出された表現のみで現在の社会情勢の変化を表しているかを判断することは難しいと考えられる。そのため、もとの文章の周りの文脈を見て、話題を判断し、それが過去のものだと判断できる場合は、抽出対象にしないようにするというような処理の改善が必要だと考えられる。

②について、これらの表現は「のような」を用いた定型表現で抽出した際のものであった。例示の表現として使用されていることを期待していたが、比喩の表現として使われていたものがほとんどであった。このことから、「のような」は本研究においてはあまり効果的な表現でなかった可能性があり、今後の十分な検討が必要になる。

③について、まず上記の例の係り受けの結果を以下の図4に示す。この例では、「最短でその日のうちに自

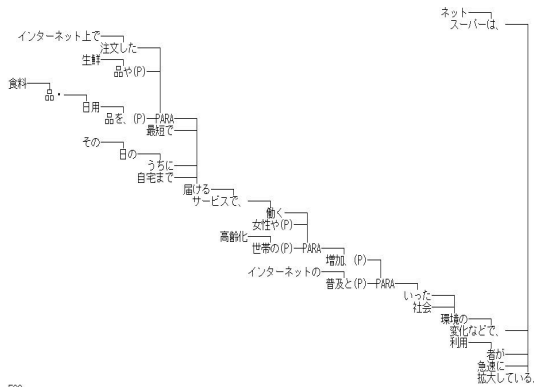


図 4: KNP での解析結果

宅まで届けるサービスで」という表現が「といった社会環境の変化」という表現にかかっていると誤解析されているため、誤って抽出された。これについては、構文解析の以外の所での解決が必要だと思われる。

④について、上記の例は、「個人の価値観の多様化」「ライフスタイルの多様化」「家族構成の多様化」の3つが並列して表記されたものである。これについては、KNP での並列構造の解析結果を上手に利用することで解決が可能と考えられる。

## 7 SOM による社会情勢の変化を表す表現の可視化

前節まででは、社会情勢の変化を表す表現の自動収集手法について述べたが、図3を見ると非常に多くの表現が収集されていることが分かる。また、この表現

の中には、「少子高齢化」、「少子・高齢化」、「少子化」、「高齢化」のように同じような意味の表現も含まれている。この状態で収集した表現の全てを提示するのでは、利用者に負担がかかってしまうと考えられる。そのため、近い意味の表現がまとまるように収集した表現を SOM により可視化し、利用者に見やすく提示する。

SOM で収集した表現を可視化するためには、その表現を対応する高次元データで表す必要がある。本研究では、分布仮説 [10] に基づき、文脈ベクトルを表現に対応する高次元データとする。

### 7.1 文脈ベクトルの作成

文脈ベクトルの作成方法について述べる。対象の表現が収集されたときの文脈だけでは、その表現の文脈としては偏っている可能性がある。そのため、対象の表現を Web で検索し得られ対象の表現を含む一文の集合加えることで、当初とは異なる文脈を集め、その対象の表現が現れうるの文脈を拡張する。また Web 検索時のヒット件数が少なかった表現については文脈ベクトルを作成しない。これはその表現が多様な文脈で使われているかを調べ、長い表現や意味をなさないような表現を対象にしないためである。今回はヒット件数が 10 件に満たないものについては文脈ベクトルを作成しなかった。このようにして拡張した文脈に対して、形態素解析を行い内容語を抽出する。各内容語をベクトルの次元とし、その語の頻度を値としてベクトルを作成する。作成したベクトルはそのベクトルの長さで割ることで正規化する。

### 7.2 SOM による可視化

収集プロセスの1回目で得られた 248 表現について、前小節で述べた方法により文脈ベクトルを作成し、SOM により可視化をする。SOM 作成時の学習回数は 1000 × 1000 回、一辺のノード数は 40 個とした。作成した SOM を図5に示す。

マップの右下に、「少子・高齢化」、「少子高齢化」、「高齢化」がまとまっていることや、マップの右上に「国際化」、「グローバル化」、「グローバル化」、「社会の国際化」、「経済のグローバル化」がまとまっており、近い意味の表現がまとまっていることが確認できる。

### 7.3 考察

マップの右下に注目すると、「少子・高齢化」、「少子高齢化」、「高齢化」に加えて「複合化」も近くに配置されており、近い意味でない表現までまとまっているため、あまり望ましくない結果となった。これらにつ

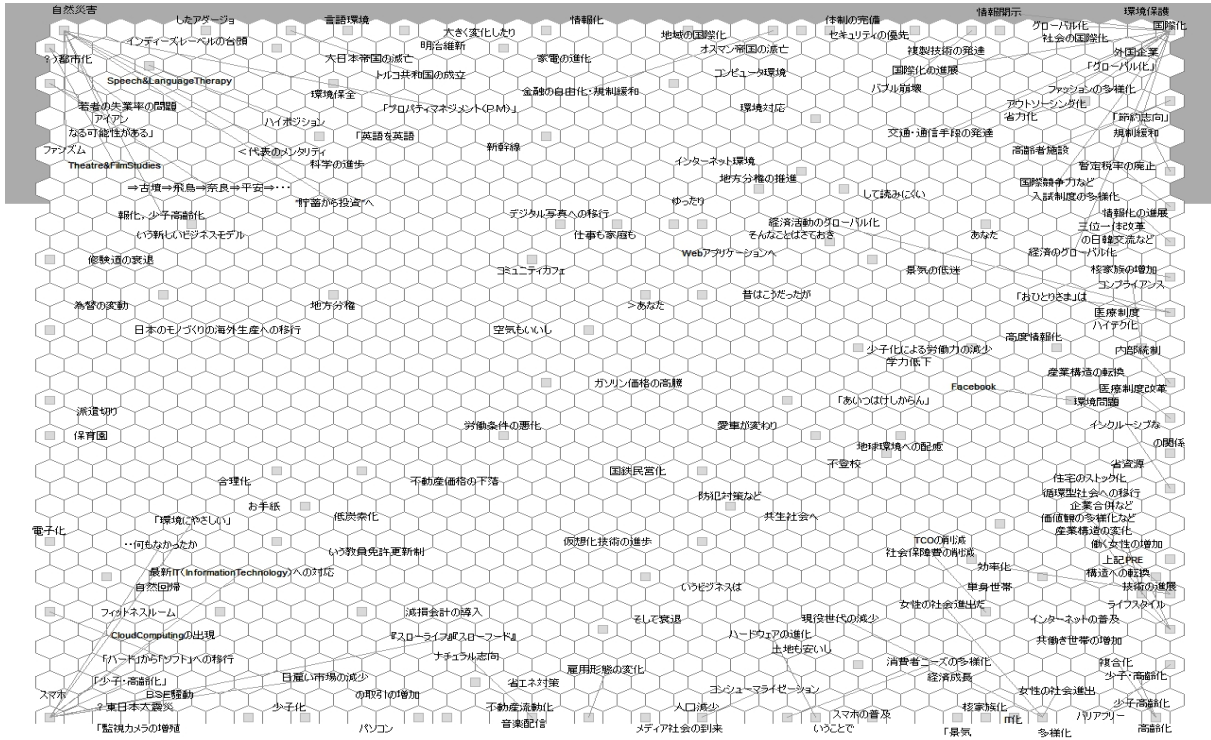


図 5: 作成した自己組織化マップ

いて文脈ベクトルを確認すると、「情報」,「出来る」といった一般的な語が多く共通していることが分かった。これらの一般的な語の影響あると考えられるため、idf などによる語の重み付けや、ストップワードの検討などが必要だと考えられる。

## 8 まとめ

本研究では、例示の表現に着目し、定型表現とブートストラッピング法を組み合わせて、社会情勢の変化を表す表現を自動収集する手法を提案した。適合率が40%程度とやや低調な値ではあったが、多くの社会情勢の変化を表す表現が獲得できた。また、収集した表現をSOMにより可視化した。一般的な語の扱いなどの文脈ベクトルの構築方法については今後の検討が必要である。

## 参考文献

[1] Hearst, M.A.: Automatic Acquisition of Hyponyms from Large Text Corpora, Proceedings of the 14th International Conference on Computational Linguistics, pp.539-545,1992.  
 [2] 安藤まや, 関根聡, 石崎俊: 定型表現を利用した新聞記事からの下位概念単語の自動抽出, 情報処理学会研究報告,2003-NL-157, pp.77-82,2003.

[3] 安藤まや, 関根聡: 上位語・下位語を含む連体修飾表現の言語的分析, 言語処理学会第10回年次大会, 2004.  
 [4] Yarowsky, D.: Unsupervised Word Sense Disambiguation Rivaling Supervised Methods, in Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics, pp.189-196 ,1995.  
 [5] Abney, S.: Understanding the Yarowsky Algorithm, Computational Linguistics, Vol.30, No.3, pp.365-395,2004.  
 [6] Curran, J. R., Murphy, T., and Scholz, B.: Minimising semantic drift with Mutual Exclusion Bootstrapping, in Proceedings of the 10th Conference of the Pacific Association for Computational linguistics, pp.172-180,2007.  
 [7] Pantel, P. and Pennacchiotti, M.: Espresso: Leveraging Generic Patterns for Automatically Harvesting Semantic Relations, in Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics, pp.113-120,2006.  
 [8] T. Kohonen.: The self-organizing map, Proceedings Of The IEEE, Vol.78, No.9, pp1464-1480,1990.  
 [9] 加藤誠, 大島裕明, 小山聡, 田中克己: 共起に基づく Webからの類似関係のブートストラップ抽出, 日本データベース学会論文誌, Vol.8, No.1, pp.11-16,2009.  
 [10] 柴田知秀, 黒橋禎夫: 超大規模ウェブコーパスを用いた分布類似度計算, 言語処理学会第15回年次大会発表論文集, pp.705-708, 2009.

# 発話内容に関する話題の提供による コミュニケーションの継続支援

## Communication Continuation Support by Providing Topics on Conversation Contents

柴田有紀\*      砂山渡  
Yuki Shibata      Wataru Sunayama

広島市立大学 情報科学部  
Faculty of Information Sciences, Hiroshima City University

**Abstract:** A lack of the communication ability I met by the spread of an information terminal in the present society is being a problem. So the contents which are conversation using a head-mounted display of the see-through type when getting communication at the meeting by this research, and, the system to recommend and show the next topic in real time is proposed. Continuation of conversation with the partner who had no dialogue experience so much up to now is suggested by this.

## 1 序論

近年、対人したコミュニケーションが苦手な人が増加している。携帯電話やスマートフォン、パソコン、その他タブレットの普及に伴い、肉声でのコミュニケーションを行う機会の減少に対して、メールやSNS、アプリケーションを用いた文面でのコミュニケーション機会が増加したことによる対人コミュニケーションスキルの低下が原因だと考える。しかし、コミュニケーションスキルの獲得は、社会で生活する上で求められるスキルとなっている。

ARとは、Augmented Realityの略で「拡張現実」と呼ばれる。現実世界でありのままに知覚される情報をもとに、デジタル合成などによって生成される情報を付加し、人の現実認識を強化することである。この技術を用いて、コミュニケーションスキルの獲得を目指すことで、現実世界のコミュニケーションにも適用することのできるコミュニケーションスキルを獲得できるのではないかと考えた。

そこで、本研究においてARとテキストマイニングを用いて、対話の際の発話内容に関する話題の提供によるコミュニケーションの継続支援を提案する。

## 2 関連研究

関連する研究として、ウェブニュースを利用した雑談対話システム[1]がある。対話としてヒト対システム間の発話の継続を図る研究である。

発話内容と類似・関連した情報を元に発話を展開する点において類似している。また、SNSなどを用いてリアルタイムな情報をユーザの「未知情報」として、提供することで「飽きない」会話を展開することで、会話の自然な継続支援を目的としている。

しかし本研究では、発話内容と関連した話題をSNSから取得した上で、一般的な話題に置換して情報を提供することでコミュニケーションの継続を図っている点で相違している。

対人したコミュニケーションの活性化に、SNS上での会話相手と第三者とのコメントを閲覧することで、共通の話題を構築し、実世界でのコミュニケーションにつなげて、会話相手とのコミュニケーションの活性化を図る研究[2]がある。情報端末で、SNS上での情報を閲覧することで、実世界でのSNSでの情報を元に、会話を展開させ会話量の増加を目指す点で、本研究と類似している。

しかし、本研究ではシースルー型のヘッドマウンドディスプレイを装着することで、実世界でのリアルタイムなコミュニケーションの継続支援を試みている点で相違している。

\*連絡先：広島市立大学情報科学部システム工学科  
〒731-3194 広島市安佐南区大塚東三丁目4番1号  
E-mail: shibata@sys.info.hiroshima-cu.ac.jp

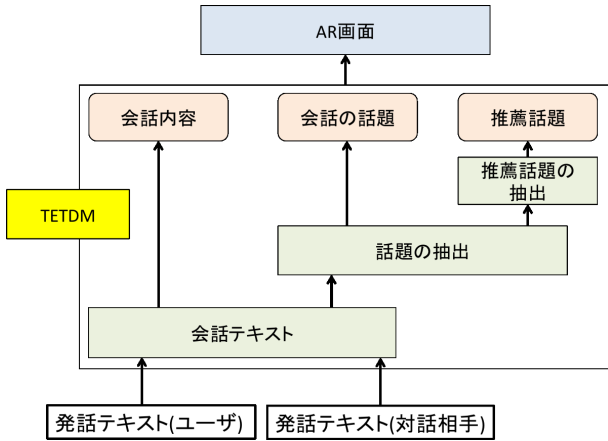


図 1: コミュニケーションの継続支援の枠組み

### 3 コミュニケーションの継続支援の枠組み

図 1 に、コミュニケーション継続支援の枠組みを示す。主な処理は、音声認識による音声のテキスト化、テキストマイニング処理による音声のテキストに沿った会話の話題の抽出と推薦話題の抽出、AR 画面処理による AR 画面のインタフェース設計を行う。

コミュニケーションの継続支援として、音声認識によって取得されたユーザと会話相手のテキストデータを元にテキストデータマイニングを行い、各会話テキストに関連した会話の話題の抽出、会話テキストが逐次保存されるため、会話内容を発話順に並び替え、会話内容の表示を行い、現在の会話に沿った会話の話題の抽出により、インターネットを経由して、推薦話題を取得する、という手順でコミュニケーションの継続支援を行う。会話内容の表示にによって、今何について話しているのかという、会話の内容を把握し、推薦話題を用いた会話の展開が自然に行うことができるように支援する。TETDM についてはテキストマイニング処理の説明の際に説明する。

#### 3.1 音声認識

音声認識には、MacBookPro(OS X Yosemite) に搭載されている「音声入力」機能を利用する。音声認識によって、音声テキストを生成する。

音声テキストは、被験者・対話相手ともにキー自動入力プログラムを使用し、自動的に音声認識の起動、テキストの改行を行う。ユーザ・対話相手ともに同様の種類の MacBookPro を 1 台ずつ計 2 台使用する。音声を入力する際に、ユーザ・対話相手ともに、会話の区切りを意識してゆっくり話してください、大きな声で

はっきりと話してくださいと注意した上で音声認識を行った。

#### 3.2 テキストマイニング環境

本研究では、テキストマイニング処理を行う環境、ならびに結果を提示する画面を、複数のテキストマイニング技術を組み合わせて使用することのできる統合環境 TETDM[4] を用いて作成する。TETDM では、自由にツールを作成して追加することができること、ならびに既存ツールの再利用のために用いた。すなわち、次節で述べる会話テキストを入力として、会話からの話題の抽出の既存ツール、ならびに抽出された話題からの関連話題の選定ツール、ユーザに提示される可視化インタフェース、を TETDM 上に実装した。

#### 3.3 会話テキスト

音声認識で取得した音声テキストをユーザ、対話相手とのテキストを音声の取得時間と伴に文字列として一行ずつ保存する。

ユーザと会話相手は異なるパソコン上で、音声認識を行う。対話相手の音声認識によって得られたテキストデータをパソコン間で構築された、クライアント/サーバーの関係によって、データの送受信を行う。これによって、リアルタイムに近似したデータの保存とテキストマイニングを行うことが可能になってる。図 1 において、対話相手の音声認識を 1 台の MacBookPro で、システム利用者(以降、ユーザと呼ぶ)も MacBookPro を使用する。各会話の順番を並べるために、取得時間を元にユーザ、会話相手の会話を保存したテキストに添えて保存する。音声テキストデータを 2 秒間に再読み込みすることで会話内容をリアムタイムに近似して投影する。

会話内容の表示により、会話のログを残し「今話されている会話」を意識して会話を展開することで、次の話題とのスムーズな話題転換を行うことを期待する。

#### 3.4 会話の話題の抽出

会話からの話題抽出には既存手法 [3] を用いる。会話の話題として、ユーザ、会話相手の音声テキストを同一のファイルへ保存する。1 分間毎で段落を挿入し、段落毎の会話の話題を抽出する。話題を単語として 1 つの単語を抽出する。段落毎の全ての単語に対して評価値を計算する。各単語  $w$  の評価値  $Ev(w)$  を文頻度  $df(w)$  と主語頻度  $sf(w)$  の積で求め、最も評価値の高い単語を会話の話題とする。以下、計算式を示す。

$$Ev(w) = df(w) \times sf(w) \quad (1)$$



表 1: twitter のトレンドから得られた一般的な話題

漫画	ゲーム
誕生日	季節
流行	アニメ
ドラマ	ジャニーズ
お笑い	俳優
女優	趣味

表 2: 推薦話題の候補リスト

カテゴリ	話題	カテゴリ	話題		
趣味系	映画	学生生活	大学		
	音楽		授業		
	読書		経験		
	漫画		高校		
	旅行		中学校		
	スポーツ		小学校		
	ゲーム		部活動		
	楽器		バイト		
	プロフィール系		名前	テレビ番組	先生
			誕生日		アニメ
年齢		ドラマ			
血液型		ジャニーズ			
出身地		AKB			
家族		お笑い			
健康		クイズ			
髪型		俳優			
体型		女優			
衣食住		通勤	興味関心		流行
	通学	ニュース			
	地域	異性			
	食べもの	ファッション			
	休日	有名人			
	季節				
	天気				

### 3.5 推薦話題の抽出

#### 3.5.1 推薦話題の候補の収集

Twitter のトレンドを管理しているサイト [6] を元に、トレンドとなっている単語を 100 個取得した。各トレンドを「つぶやかれている回数」として計算されている「ツイート数」順に並び替え上位 50 単語を推薦話題の候補として列挙した。その後、候補話題を一般的な話題に変換して表 1 を作成した。

#### 3.5.2 推薦話題の候補リストの作成

Twitter のトレンド話題を一般的な話題へ置換した後に、さらに、重複する単語を排除し、表にまとめたものが、表 1 である。

表 1 の単語から、趣味、プロフィール、衣食住、興味関心、テレビ番組というカテゴリを作成し、本実験の被験者に合わせて、「学生生活」といカテゴリを追加した上で、組織化したものを表 2 に示す。各カテゴリ

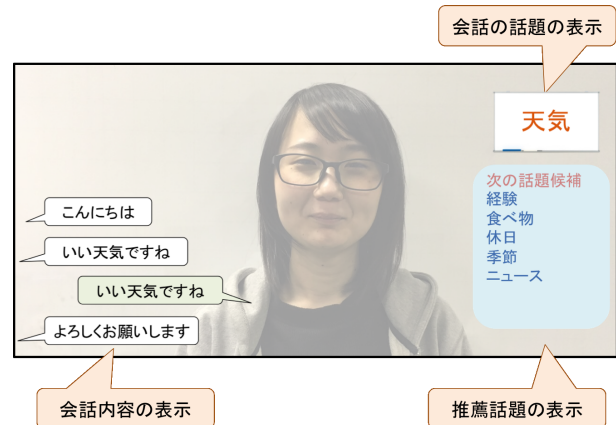


図 2: コミュニケーションの継続支援インターフェース

について、一般的に、どのような事柄について話されているのか、Twitter のトレンドをもとにカテゴリを細分化し、まとめた。

#### 3.5.3 推薦話題の抽出

この表 2 をもとに会話の話題を単語  $W$  とし、推薦話題を  $C_i$  とする。インターネットの検索サイト [7] を利用して、 $W$  AND  $C_i$  として、AND 検索をかける。また、推薦話題として単語を使用した際に、今の話題と推薦話題が一致した時、その単語は再表示されないことになっている。検索ヒット数の順に 5 つの単語を AR 画面に推薦話題として表示する。この動作を 1 分間毎で更新する。

### 3.6 コミュニケーションの継続支援インターフェース

AR 画面に表示する、インターフェースを以下の図 2 に示す。画面の中に表記している、画面の中にある、会話内容の表示が、音声認識によって、取得したユーザと対話相手の音声テキストを時間軸に沿って、表示したものとになっている。白の吹き出しに自分の発話を、緑色の吹き出しに相手の発話を表示し、自然な流れで双方の会話の展開を視覚化している。これによって、今どのような会話が展開されていて、何について話しているのかが一目でわかるようになっている。

また、会話の話題は、前章で説明したように、今話している会話内の重要な単語を 1 つ表示することで、今どのようなことについて会話が展開しているのかを表示するようになっていた。これは、対話相手と展開している会話について、話題転換がされる際に、自然な流れで会話の展開を行えるようになっている。



最後に、推薦話題について、前章でも説明したが、会話の話題は今の話題と関連した話題であるので、次の展開で話す話題として推薦しているものとなっている。ユーザが会話内容に困った際に、推薦話題を見て会話を継続して話せるようになっている。

## 4 コミュニケーション継続支援システムの評価実験

本章では、提案するコミュニケーション継続支援の枠組みを実装したシステムの効果を検証するために行った実験について述べる。

### 4.1 実験手順

提案システムの有効性を検証するために、推薦話題の提供を行わない比較システム（図2の推薦話題の表示がないシステム）を用いて比較実験を行う。

#### 4.1.1 評価実験の事前アンケート

- 問1 普段の会話では、自分から話題をふる、あるいは話題を変える方ですか？
- 問2 普段の会話で、会話中に、今の話題を意識していますか？
- 問3 普段の会話で、会話中に、次の話題を意識していますか？
- 問4 今まであまり話したことが無い人と会話をする時に、話題に困ることはありますか？

問1では、「自分から話題をふる」を5点、「相手に任せる」を1点とした。

問2では、「とても意識している」を5点、「全然意識していない」を1点とした。

問3では、問2と同様とする。

問4では、「よく話題に困る」を5点、「全然話題に困らない」を1点とした。

#### 4.1.2 被験者

初対面の大学生の2人1組のペア10組を作り、提案システムによって被験してもらった提案群と、比較システムによって被験してもらった比較群と2つのグループに分けた。

提案群と比較群で、アンケート結果を基に両群の被験者のコミュニケーション能力を同じにする。また、ア

ンケートでのコミュニケーションに対する自己評価を5点や、1点とした人を被験者にしないこととした。

本実験では、2点から4点までの点数の人を被験してもらった。提案システムと提案システムでの被験者1・被験者2・被験者3は、同一人物であるが、対話相手と異なる人で組み合わせることで、初対面とした。

#### 4.1.3 実験内容

実験内容として、会話を10分間行ってもらう。

ユーザ、会話相手ともに、音声を取得するために骨伝導マイク (Jawbone ICON HD) を装着してもらう。10分間の会話を行うために最初1分間の間で互いに自己紹介をしてもらう。音声認識によって所得した、音声テキストには実験のはじめ、文字列がない状態なので、会話テキストを蓄積するためにも、どのような人でも会話を展開することのできる、自己紹介を行ってもらう。

自己紹介の後に、相手と会話をしてもらう。相手と会話をする気持ちがないと、継続をしようとする気持ちが生まれず、自ら沈黙してしまうことを避けるために、相手と友好的になろうとする意識を維持して会話に臨んでもらうために、「相手を自分の恋人の家族(兄弟・姉妹)だと仮定して会話を行ってください」とあらかじめ会話を行う状況を設定しておく。

ユーザには、システムインタフェースの提示してある内容の説明をし、会話相手には、ユーザからの会話の話題自ら展開しないように心がけることを注意した。

### 4.2 実験の結果と考察

実験の評価には、以下の3つの観点から評価する。

- 推薦話題の利用
- コミュニケーションの量
- アンケート結果

#### 4.2.1 推薦話題の利用

推薦話題の効果を評価するために、話題の転換数を数える。話題の転換数とは、システムを動作した時に、1分間毎で保存される、会話の話題の変化数のことである。

各実験で得られた、話題の転換数を図3にまとめる。

図3より、提案システムの方が比較システムより話題の転換数が多くなっている結果となっている。各システムでの話題の転換数の平均回数を以下の表3にま

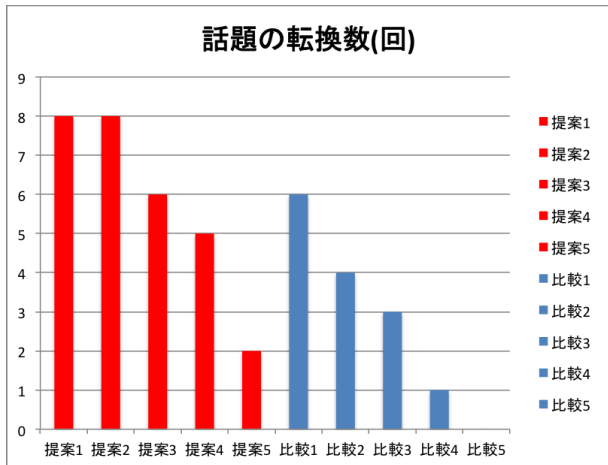


図 3: 話題の転換数

表 3: 話題の転換数 (平均回数)

提案システム	比較システム	t 検定
5.8	2.8	$p < 0.01$

とめた . t 検定の値については、小数第 4 位以下を切り捨てて表記する。

#### 4.2.2 コミュニケーションの量

提案システムと比較システムのシステムの相違点として、推薦話題の表示の有無によって比較実験を行った。比較システムと提案システムでは、提案システムの方が、話題の転換数が平均して、3 回多く話題が切り替わっていることがわかる。このことより、推薦話題の表示があることによって、話題の転換回数が増加したと考えられる。t 検定の値が、0.01 未満であることから、話題の転換数において、有効であったと言える。

以下の図 5 は、提案システムを使用して、被験してもらったある被験者の会話を行ってもらった際の、会話の話題と推薦話題の一覧である。会話の話題と推薦話題の同色の単語に注目すると、まず経過時間が 3 分の時、推薦話題に「ニュース」とあるがその 1 分後の会話の話題が「ニュース」となっていることから、推薦話題を参考にして会話が展開されたと言える。この場合は、「ニュース」と同じ単語であるが、経過時間が 5 分の時には推薦話題が「地域」で、1 分後には会話の話題が「松岡」となっている。これは、会話の内容から関連していたと言える。以降の同色の会話の話題と推薦話題が関連していることが見てとれる。

グラフを参照すると、比較システムの方が沈黙時間が長いに対し、提案システムの方が文字数が多いこと

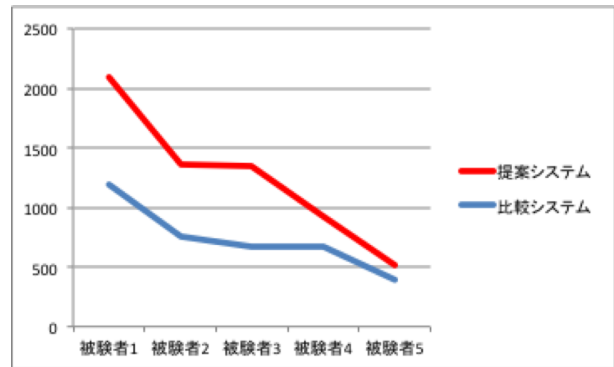


図 4: 各実験における会話内の文字数

表 4: 各実験でのユーザ・対話相手間の発話した文字数の割合

	提案	比較
被験者の文字数 (字)	770.4	576.4
対話相手の文字数 (字)	478.6	159.2
被験者の文字数の割合 (%)	63.5	78.9
対話相手の文字数の割合 (%)	36.4	21.1

がわかった。これにより、提案システムでは、比較システムと比較して、10 分間という制限時間のある会話のなかで、発話量が多いと言えるため、継続してコミュニケーションが行われたと言える。

また、沈黙時間が、比較システムと比較して、提案システムの方が単純沈黙時間では、平均して 72(秒)、沈黙時間では、平均して 89(秒) 減少していることから、提案システムのコミュニケーションの継続の支援に有効だったと言える。

#### 4.2.3 アンケート結果

実験が終了した時点で被験者に事後アンケートを実施した。

- 提案群への事後アンケートの質問

- 問 1 会話で、話題に困ることはありませんでしたか？
- 問 2 表示された「会話の話題」は適切に話題を表していましたか？
- 問 3 表示された「推薦話題」は次の話題の選択に役立ちましたか？
- 問 4 表示された「推薦話題」は「会話の話題」に関係していましたか？

継続時間	会話の話題	推薦話題				
1(分)	東日本	ニュース	地域	大学	旅行	名前
2(分)	東日本	ニュース	地域	大学	旅行	名前
3(分)	変態	アニメ	女優	ニュース	ゲーム	漫画
4(分)	ニュース	大学	高校	映画	スポーツ	名前
5(分)	今日	大学	健康	旅行	地域	家族
6(分)	松岡	名前	アニメ	ゲーム	映画	大学
7(分)	松岡	名前	アニメ	ゲーム	映画	大学
8(分)	予定	大学	高校	地域	スポーツ	ゲーム
9(分)	予定	大学	高校	地域	スポーツ	ゲーム
10(分)	画像					

図 5: 会話中の話題の展開例

表 5: 提案群の事後アンケート結果

被験者	提案 1	提案 2	提案 3	提案 4	提案 5
問 1	4	5	1	4	5
問 2	1	4	3	3	2
問 3	4	4	2	4	5
問 4	2	4	2	3	5

問 1 では、「話題には困らなかった」を 5 点、「話題に困った」を 1 点とした、

問 2 では、「話題の表示は適切だった」を 5 点、「話題の表示は適切でなかった」を 1 点とした、

問 3 では、「とても役にたった」を 5 点、「全然役に立たなかった」を 1 点とした問 4 では、「とても関係していた」を 5 点、「全く関係していなかった」を 1 点とした、

提案群では、5 人中 4 人は話題に困らなかったと回答しており、提案システムがコミュニケーションの継続に寄与したことが伺える。しかし、現在の会話の話題や、現在の会話に関する話題の提供といった点では、被験者により効果が異なる結果となった。これは音声認識システムが、必ずしも発話内容を適切に捉えられていないことにも起因する。しかし、推薦話題として一般的な話題をリストに用意していたことから、何も話題の提案がないよりは、提案があった方がコミュニケーションを続けやすかったということは確認できる。

● 比較群への事後アンケートの質問

問 1 会話で、話題に困ることはありませんでしたか？

問 2 表示された「会話の話題」は適切に話題を表していましたか？

問 1 では、「話題には困らなかった」を 5 点、「話題に困った」を 1 点とした、

問 2 では、「話題の表示は適切だった」を 5 点、「話題の表示は適切でなかった」を 1 点とした、

比較群のアンケート結果から、話題の提供が何もなされない状態だと、次に何の話をするべきか、大きく悩む結果となった。これらのことから、推薦話題の提

表 6: 比較群の事後アンケート結果

被験者	比較 1	比較 2	比較 3	比較 4	比較 5
問 1	1	3	1	2	2
問 2	1	4	2	2	2

示により、コミュニケーションの継続に対して一定の効果は確認できた。しかし会話内容に対する、より適切な話題の提供に対しては課題が残る結果となった。

## 5 結論

本研究として、発話内容に関する話題提供によるコミュニケーションの継続支援を提案した。実験により、発話内容に関する話題提供によるコミュニケーションの継続支援の有用性を検証した。今後は、提供する話題の細分化の是非、話題提供の適切なタイミングなどを検討していきたいと考えている。また、誰もが気軽に HMD などが持ち歩くことのできる社会へ向け、リアルタイムにコミュニケーションを支援できる環境の実現に向け研究を続けていきたい。

## 参考文献

- [1] 水野淳太, 乾健太郎, 松本裕治, ウェブニュースを利用した雑談対話システム, 人工知能学会 第 55 回 SLUD 研究会, pp.1-6, (2009)
- [2] 水口弘紀, 石澤善雄, 村岡優輔, 中尾敏康, 対話のきっかけとなる話題提供によるコミュニケーション活性化技術, NEC 技報 Vol.66 No.1, (2013)
- [3] 後藤賢悟, 砂山渡: AR とテキストマイニングを用いた対話時の好感度推定によるコミュニケーション支援, 第 29 回人工知能学会全国大会, 2E4-NFC-01b-2, (2015)
- [4] TETDM サイト:(URL) <http://tetdm.jp>
- [5] 砂山渡, 高間康史, 西原陽子, 梶並知記, 串間宗夫, 徳永秀和: 統合環境 TETDM を用いたマイニングツールの開発と利用の実践, 人工知能学会論文誌, Vol.29, No.1, pp.100 - 112, (2014).
- [6] ついっぷるトレンド: (URL) <http://tr.twipple.jp>
- [7] 検索サイト: (URL) <https://www.bing.com>

# レイアウト認識に基づく論文構成要素の抽出

## Automatic Role Labeling of OCR Processed Scholarly Papers

岩月憲一<sup>1\*</sup> 加藤恒昭<sup>2</sup> 山口和紀<sup>2</sup>  
Kenichi Iwatsuki<sup>1</sup>, Tsuneaki Kato<sup>2</sup>, Kazunori Yamaguchi<sup>2</sup>

<sup>1</sup> 東京大学教養学部

<sup>1</sup> School of Arts and Sciences, The University of Tokyo

<sup>2</sup> 東京大学大学院総合文化研究科

<sup>2</sup> Graduate School of Arts and Sciences, The University of Tokyo

**Abstract:** Components of scholarly papers bear roles such as title, body, itemization title, or figure. A role label enables advanced searching such as finding papers in which a specified keyword is used in a specified role. In this paper, we propose a fully automatic role labeling method for OCR processed scholarly papers. In the proposed method, we first identify components from the OCR processed images by reconstructing components from incorrectly recognized regions by OCR software. Next, we assign role labels to the components. Our experiment showed that the accuracy of the classification reached 94% in the best case.

## 1 はじめに

学術論文の紙面は、テキストや図表など、複数の構成要素から成り立っている。そして、それぞれの構成要素が、論文タイトル、著者、見出し、本文といった文書内役割を持っている。

論文検索を行う際に、論文全体を対象とした検索を行うよりも、ある特定の役割を持った構成要素に絞って検索を行いたいことがある。この検索意図に応えるためには、予め、論文から抽出した構成要素に対して、正しく文書内役割を特定しておく必要がある。例えば、節タイトルとそれに対応する本文を特定しておけば、ある特定の節に絞った検索が可能となる [3]。しかし、文書内役割は筆者と読者の共通認識に基づき暗黙的に与えられるものであり、論文データ中に文書内役割が直接記述されることはない。従って、構成要素に対し、明示的に文書内役割を特定することが問題となる。本研究は、学術論文の構成要素に対し、文書内役割を特定することを目的とする。

論文構成要素を得るためには、論文 PDF ファイルを直接 XML 等に変換するか、あるいは論文画像をスキャンし、OCR ソフトウェアで処理をするという方法がある。前者の場合、PDF 作成ソフトウェアの仕様により変換後のデータの性質が異なる場合がある [2]。また、デジタルデータから直接 PDF に変換されておらず、紙

媒体をスキャンした PDF については、取り扱うことができない。以上の理由から、本研究では、論文構成要素を得る方法として、論文画像をスキャンし、OCR で処理をするという方法を採用する。

OCR ソフトウェアで論文画像を処理すると、各構成要素と大まかに一致する矩形領域を得ることができる。この時、OCR の誤りによって、得られた矩形領域が実際の論文構成要素と一致しない場合があるので、これを修正する。その後、各構成要素に対して、文書内役割を割り当てる。

## 2 関連研究

先行研究について、(1) 対象とする論文のデータ、(2) 文書内役割特定の手法、(3) 特定する文書内役割の種類、の3つの観点から述べる。

まず、対象とする論文のデータには、2種類あり、PDF を変換し XML 等のデータとして得るもの [2, 5, 6, 7]、論文画像をスキャンし OCR で処理したデータを得るものに分かれる。

次に、文書内役割特定の手法には、大きく分けると2種類あり、1つはルールベースによる手法、もう1つは機械学習による手法である。

ルールベースによる文書内役割特定には、Klink らが提案している手法がある。そこでは、文書には、あらゆる文書に共通する構造と、例えば学術論文のような特定の領域の文書に特化した構造の2種類があると

\*連絡先：東京大学教養学部学際科学科  
〒153-8902 東京都目黒区駒場 3-8-1  
E-mail: 0191699319@mail.ecc.u-tokyo.ac.jp

しており、構成要素に対しそれぞれルールを適用していく。前者では、ページ上部ならばヘッダ、中黒から始まっているテキストならば箇条書き、といったルールが適用される。後者では、表の上にはキャプションが来る、といったルールが適用される。これらのルールが構成要素に複数適用され、最終的に、最も可能性の高い文書内役割が与えられる [4]。

機械学習による文書内役割特定としては、SVMを用いる手法 [2, 6] と CRF [2, 5, 7] を用いる手法の 2 種類がある。学習・推定に際して用いる素性には大きく分けて 2 種類ある。矩形領域の位置座標や大きさといったレイアウト情報から抽出される素性と、テキスト情報から抽出される素性である。いずれの先行研究も、両方の素性を用いている。増田らが用いている素性には他の研究とは異なり、形態素解析を利用して得たテキスト中の名詞や人名の割合が用いられている [6]。また、

Luong らは、処理を 2 段階に分けており、レイアウト情報とテキスト情報を利用して文書内役割を特定し、なおかつ節タイトル部分についてはテキスト情報のみを利用して節の種類を特定している [5]。

最後に、特定する文書内役割の種類について述べる。先行研究には、論文全体を対象としているもの [4, 5, 6] と、一部のみを対象としているもの [2, 7] がある。前者については、タイトル、節タイトル、本文、図、表など 23 種類の役割を特定しているもの [5] から、論文タイトル、著者、本文、ヘッダ、フッタ、ページ番号の 6 種類の役割に留めているもの [6] もある。後者については、論文中の図表とそのキャプションのみを特定しているもの [2]、論文の書誌情報のみ (タイトル、著者、アブストラクト、キーワード) を特定しているもの [7] がある。

本研究では、PDF 作成ソフトに依存せず、かつスキャンされた論文にも対応するため、OCR ソフトウェアで処理したデータセットを用いることにする。

同様に OCR を用いている研究はあるが、いずれも、レイアウト情報の認識誤りを修正する手法については述べていない。しかしながら、Luong らが指摘するように、OCR の誤りによって文書内役割が適切に特定できない場合がある [5] ため、本研究では誤り修正法についても取り組んでいる。

文書内役割特定の手法としては、ルールベースのものは論文の種類によってルールを手動で調節する必要があるため採用せず、機械学習の方法を採用ことにした。今回の問題を系列ラベリング問題として捉えることにより、矩形の列の順序を学習・推定のために用いることができるため、CRF を採用することとした。

文書内役割については、情報検索への応用を前提とするため、論文全体にわたる 19 種類を用意した。

## 3 提案手法

### 3.1 概要

文書内役割の特定にかかる一連の処理は次の手順で行う (図 1)。OCR で論文画像を処理することによって、論文の構成要素を、複数の矩形領域として得ることができる。この時、構成要素と対応して認識されることが望ましい矩形が、分割されて認識されてしまうという誤りが生じる。

このようにして得られた矩形領域を入力とし、まず矩形領域を一次元列に並べ替える。学术论文は、タイトル→著者→見出し→本文といった一定の順序を持っており、この順序の情報を有効に用いたい。そこで、二次元紙面上に配置された論文の構成要素を、論文を読む際の順序に一致するよう一次元列に並べ替える。そして、この順序の情報を用いて以降の処理を行う。

続いて、誤って分割された矩形領域の統合を行う。

最後に、各矩形領域に対応する文書内役割を与える。ここでは、先に与えた順序を利用して、CRF による文書内役割の学習・推定を行う。OCR ソフトウェアで得られる情報には、論文中のテキストそのもの (テキスト情報) に加え、位置座標、大きさなどのレイアウト情報も含まれている。しかし、テキストそのものを利用しようとすると、文字認識の誤りが大きく影響してしまう可能性がある。従って本研究では、レイアウト情報を主に用いることにする。

認識誤りについて、補足しておく。OCR ソフトウェアは、余白などの視覚情報を元に、まとまったテキストや図を 1 つの矩形領域として設定している。そのため、論文の構成要素と、この矩形領域は必ずしも一致しない。このような不一致には、2 つの種類がある。

2 段組の論文にあっては、1 つの節に属するテキストが 2 つの段にわたって記述されている場合、矩形領域としては 2 つに分割されて手に入ることになる。しかしながら、先述の通り読む順序で並べ替えを行うため、2 つのテキストは系列内で連続する。後の文書内役割特定で、2 つとも本文として認識されれば、2 つのテキストを統合し、1 つの節として抽出することが可能になるから、問題にはならない。

他方、図表中に文字を含む場合、これらの文字を認識してしまい、複数のテキストとして認識するため、1 つの図表が複数の矩形領域に分割されてしまう。このまま文書内役割の特定を行おうとしても、実際の図表と比べて、矩形領域の大きさが極端に小さくなってしまったために、図表として認識されにくくなる。このため、分割されている矩形を文書内役割特定の前に統合しておく必要があるのである。

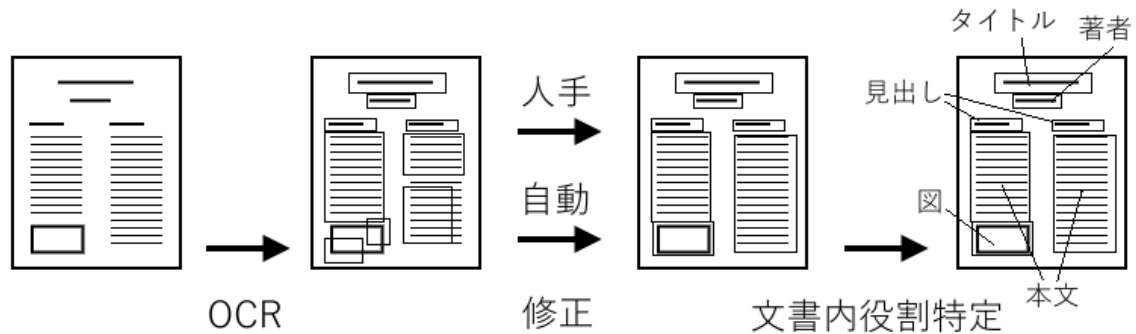


図 1: 処理の流れ

### 3.2 矩形の順序付け

2次元平面上に配置された矩形領域を、1次元の列に並べる手法について述べる。本研究では、論文を読む際の順序に矩形を並べるため、読む順序の特定が必要になる。

読む順序を特定する先行研究としては、Aiello らが次のような手法を提案している [1]。まず、任意の2つの矩形領域の座標の情報をもとに、ルールベースで、ありうる読む順序を全て求める。その後、各矩形のテキストに品詞タグを付与し、それぞれの順序ごとに、矩形の最後2単語の品詞と次の矩形の最初の1単語の品詞が共起する確率を計算し、最も確率の大きい順序に決定するというものである。しかしながら、OCRの文字認識の誤りがあること、および図表が複数の矩形に分割されていることから、各矩形のテキストに対し適切な品詞タグを付与することが困難となるため、本研究で対象としている文書には、この手法を適用することはできない。

本研究では、以下の観察から、ヒューリスティクスに基づいたルールベースで読む順序を特定し、順序付けを行うことにした。

学術論文は、本文部分が2段組となっている場合が多い。従って、左上→左下→右上→右下の順に並べ替えることになるが、大きな図表については、2段にまたがっていることがしばしばある。この場合、図表よりも上の部分と下の部分のそれぞれで、左上→左下→右上→右下の順序に並べ替える必要がある。さらに、図表部分が小さく分割されている可能性があるため、これも考慮する。

### 3.3 矩形領域の誤り修正

構成要素と対応して1つの矩形領域であると認識すべきところ、複数に分割されている矩形を判別し、統合する。

まず、判別の手法であるが、SVMを用いて学習・推定を行った。カーネル関数には、LinearとRadial Basis Function(RBF)をそれぞれ用いた。

矩形のx座標(横方向)、矩形のy座標(縦方向)、矩形の幅、矩形の高さ、矩形内テキストの文字数、矩形の面積に着目し、それぞれ当該矩形の値、1つ前の矩形との差分、1つ後の矩形との差分を学習に用いる素性とした。ここでの1つ前、1つ後というのは、前節で並べ替えた一次元列における順序である。合計18の素性を用いている。

SVMを用いて統合すべきであると判定された矩形を1つの矩形に統合したデータを用意する。統合後の矩形について、位置座標は、最も左上にある矩形のものを採用した。大きさについては、特定された全ての矩形が内部にあるような長方形のうち、最も小さいものを採用した。

### 3.4 文書内役割の推定

各矩形の文書内役割を学習・推定するために、CRFを用いた。今回特定した文書内役割は、謝辞、謝辞タイトル、著者、本文、数式、図、図キャプション、フッタ、脚注、ヘッダ、箇条書き、注釈、ページ番号、参考文献、参考文献タイトル、節タイトル、表、表キャプション、論文タイトルの19種類である。

用いた素性は、矩形の位置座標、矩形の大きさ、フォントサイズ、矩形内文字列の1文字目の文字種、矩形内文字列の行数である。位置座標と大きさについては、



1つ前の矩形との差分も利用した。また、位置座標と大きさは 100px 毎に離散化してある。

## 4 実験と評価

### 4.1 概要

今回の実験に使用したデータセットについて述べる。言語処理学会年次大会発表論文集のうち、2003年、2006年、2009年、2013年の論文からそれぞれ27、23、22、26本ずつ、合計98本の論文を、OCRソフトウェア<sup>1</sup>で処理し、データセットとした。年度によってスキャンの精度に差異があるため、複数の年度から採用した。

次に、矩形認識の誤りを人手で修正統合し、矩形認識誤りのないデータを作った上で、人手で文書内役割を特定することで、正しく文書内役割が割り振られている正解データを作成した。

今回の実験では、矩形の認識誤りの修正と文書内役割の特定という2つの段階があり、合わせて表1のように実験を行った。

表 1: 実験の種類

実験番号	誤り修正	文書内役割特定
(1)	人手	CRF
(2)	SVM	CRF
(3)	なし	CRF

認識誤りの修正は、その効果を検証するため、提案したSVMによる修正に加え、人手で修正した場合と全く修正しなかった場合も行っている。

文書内役割の特定の結果を正解データと比較することによって、CRFによる文書内役割特定の評価を行う。同時に、実験(1)~(3)の結果を比較することで、認識誤りの修正が、文書内役割特定に及ぼす影響について確かめる。

### 4.2 矩形領域の誤り修正

全ての矩形について、統合すべきであるか否かを人手でタグ付けしたデータを用意し、そのデータを用いてSVM<sup>2</sup>で学習・推定を行い、7分割交差検定をしている。

実験の結果は表2、3の通りである。

その後の実験では、RBFカーネルで推定した結果を元に矩形の修正を行ったデータセットを用いている。

### 4.3 文書内役割の推定の評価法

人手で文書内役割を与えた正解データとの比較をすることで、各実験(1)~(3)に対する処理の評価を行うことにする。この場合、それぞれの文書内役割毎に、F値を計算するのが自然である。

しかしながら、実験(1)~(3)の結果は、それぞれ認識誤りの修正の方法が異なるため、矩形列が異なっている。そのため、正解データと、矩形領域が一対一に対応しないために、単純に比較することができない。そこで、以下のような評価法を考える。

まず、正解データでは1つの矩形であるところ、実験データではn個の矩形に分かれてしまっている場合を考える。この時は、複数の矩形で1つの正解データの矩形に対応するので、F値の計算時には、1/n個の矩形として扱う(図2)。

続いて、実験データでは1つの矩形となっているところ、正解データではn個の矩形に対応している場合を考える。この時は、実験データの矩形に付与されたラベルが、対応するn個の矩形全てに付与されたものとして扱う(図3)。

表 2: 矩形領域の誤り修正結果 (Linear カーネル)

正解\推定	分割される	分割されない
分割される	566	784
分割されない	265	5024

表 3: 矩形領域の誤り修正結果 (RBF カーネル)

正解\推定	分割される	分割されない
分割される	558	792
分割されない	250	5039

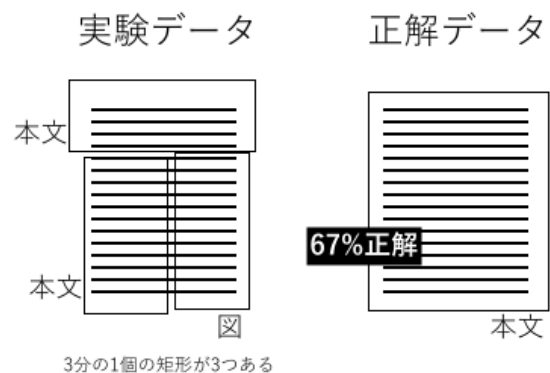


図 2: 矩形が分割された場合の正解率の計算

<sup>1</sup>メディアドライブ社の e.Typist を使用した。

<sup>2</sup>実装には、libsvm (<https://www.csie.ntu.edu.tw/~cjlin/libsvm/>) を用いた。

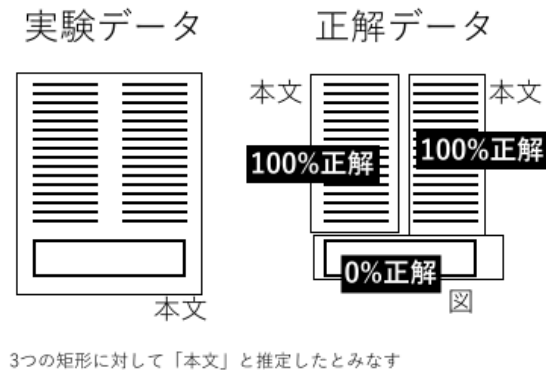


図 3: 矩形が統合された場合の正解率の計算

#### 4.4 文書内役割の推定

CRF<sup>3</sup> を用いて、文書内役割の学習・推定を行った。正解データを用いて学習し、この学習結果を用いて文書内役割の推定を行った。すべて7分割交差検定を行っている。4.1節で説明した3種類の実験の結果を表4にまとめた。

### 5 考察

#### 5.1 矩形認識の誤り修正について

##### 5.1.1 分割された矩形の統合が文書内役割特定にもたらす影響

矩形が分割されてしまう誤りを修正するためには、分割されている矩形を特定する処理と、特定した複数の矩形を統合し分割される前の矩形に戻す処理が必要となる。

表4によれば、人手で分割された矩形を統合している場合、文書内役割特定の正解率は高いことから、分割されている矩形を統合する必要があることが分かる。しかし、SVMを用いて分割された矩形を特定し統合した場合と、一切修正を加えていない場合の、文書内役割特定の正解率に差がないという結果になっている。この原因について考察を加える。

分割された矩形を統合することによる効果をみるために、分割された矩形から一定の割合でランダムに選ばれた矩形を統合した上で、文書内役割の正解率を計算するという実験を行った。

この実験の結果が表5である。全く統合していない場合と、全て統合した場合で、文書内役割特定の正解

率に変化は見られない。これは、CRFによる文書内役割の特定自体が、分割された矩形に対しても、統合された矩形に対しても、効果的でないことを示している。実際に、元々分割されていた矩形のみに注目した文書内役割の正解率を調べると、統合された矩形の割合に関係なく低い正解率となっている(表6)。

分割された矩形を統合した後に付与された文書内役割を見ると、箇条書きとなっている矩形が多い。つまり、もともと図表数式であった矩形が分割され、統合されると、図表数式の特徴を失うということである。例えば、図4に示すように、元々の矩形と統合後の矩形では、位置と大きさが異なる場合が少なくない。従って、分割された矩形を統合する手法を改善する必要があると言える。

なお、表5の結果と、表4のSVM+CRFの結果に差異が見られないのは、SVMによって誤って分割されていないのに分割されていると判断された矩形のうち、修正を加えない場合に文書内役割が正解した矩形は37個(全体の0.6%)であり、これらを統合してしまうことによる文書内役割特定の正解率の低下は非常に小さいためである。

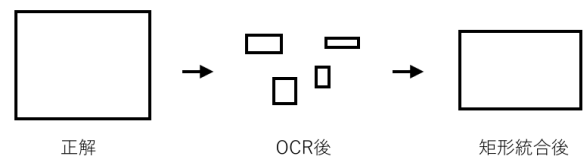


図 4: 統合後の矩形

##### 5.1.2 SVMによる矩形認識の誤り修正

矩形認識誤りの修正において、レイアウト情報を素性とするSVMがどの定後有効か検討する。

SVM (Linear カーネル) によって求められた分離平面の成分を表7にまとめた。これによれば、矩形の幅と、矩形内テキストの文字数が判定に最も大きく寄与していると言える。実際には約6割の分割された矩形は特定に至っていない(表2)。

図5は実際に分割された矩形の幅の分布であるが、分割された矩形の半分程度は比較的大きな矩形が占めている。図6は全矩形の幅の分布であるが、図5と比べると、比較的大きな矩形の大半は分割されていない矩形であり、これらを特定するには、大きさの情報だけでは足りないことが分かる。

<sup>3</sup>実装には、CRF++ (<https://taku910.github.io/crfpp/>) を用いた。

表 4: 実験結果

文書内役割	(1) 人手+CRF			(2)SVM+CRF			(3) なし+CRF		
	P	R	F	P	R	F	P	R	F
謝辞 (ack)	1.00	0.50	0.67	1.00	0.25	0.40	0.75	0.25	0.38
謝辞タイトル (ackt)	0.17	0.67	0.27	0.75	0.25	0.38	1.00	0.33	0.50
著者 (auth)	0.95	0.93	0.94	0.62	0.93	0.75	0.54	0.93	0.68
本文 (body)	0.88	0.91	0.89	0.59	0.73	0.65	0.57	0.72	0.64
数式 (equ)	0.75	0.77	0.76	0.06	0.01	0.02	0.05	0.01	0.02
図 (fig)	0.74	0.72	0.73	0.44	0.22	0.29	0.40	0.22	0.29
図キャプション (figc)	0.76	0.82	0.79	0.59	0.33	0.43	0.55	0.33	0.41
フッタ (ftr)	0.99	1.00	1.00	0.93	1.00	0.96	0.91	1.00	0.95
脚注 (ftnt)	0.82	0.71	0.76	0.55	0.59	0.57	0.54	0.60	0.57
ヘッダ (head)	0.97	1.00	0.99	0.91	0.86	0.89	0.90	0.87	0.88
簡条書き (list)	0.76	0.72	0.74	0.28	0.69	0.40	0.27	0.68	0.38
注釈 (note)	0.33	0.15	0.21	0.33	0.25	0.29	0.42	0.25	0.31
ページ番号 (page)	0.99	1.00	0.99	0.98	1.00	0.99	0.96	1.00	0.98
参考文献 (ref)	0.88	0.77	0.82	0.84	0.81	0.82	0.80	0.81	0.80
参考文献タイトル (reft)	0.76	0.84	0.80	0.68	0.49	0.57	0.64	0.50	0.56
節タイトル (stitle)	0.95	0.98	0.96	0.76	0.51	0.61	0.73	0.53	0.61
表 (tab)	0.80	0.83	0.82	0.68	0.41	0.51	0.59	0.42	0.49
表キャプション (tabc)	0.80	0.80	0.80	0.66	0.45	0.54	0.59	0.46	0.52
論文タイトル (title)	0.97	0.94	0.96	0.81	0.80	0.81	0.78	0.80	0.79
正解率	0.86			0.61			0.61		

表 5: 分割された矩形の正解率と文書内役割の正解率

分割されていると判定する確率	文書内役割の正解率
0	0.61
0.2	0.61
0.4	0.61
0.6	0.61
0.8	0.61
1	0.61

表 6: 分割された矩形に注目した文書内役割の正解率

分割されていると判定する確率	文書内役割の正解率
0	0.16
0.2	0.15
0.4	0.16
0.6	0.17
0.8	0.17
1	0.13

### 5.1.3 複数の矩形が統合されてしまう認識誤り

ここで、OCRによる矩形認識の誤りについて、改めて述べる。認識誤りには、2種類ある。1つは、矩形が複数に分割されてしまうという誤りであり、もう1つは、複数の矩形が1つの矩形として認識されてしまうという誤りである(図3)。これまで、前者に着目して修正を行ってきたが、後者の認識誤りの影響について考察する。

今回の実験では、正解データにおける約3割の矩形が、複数個で1つの矩形として認識されてしまっている。本来異なる文書内役割を持つ複数の矩形が1つにまとめられてしまうと、推定の段階で1つしか文書内役割を与えられないため、正解率が大きく低下することになる。

表 7: 分離平面の成分

成分	値
x座標の差分(1つ前)	0.08
x座標	0.37
x座標の差分(1つ後)	-0.18
y座標の差分(1つ前)	0.35
y座標	-0.15
y座標の差分(1つ後)	-0.85
幅の差分(1つ前)	-0.24
幅	-2.74
幅の差分(1つ後)	2.36
高さの差分(1つ前)	0.10
高さ	-0.84
高さの差分(1つ後)	-0.03
文字数の差分(1つ前)	1.05
文字数	-4.95
文字数の差分(1つ後)	2.05
面積	1.05

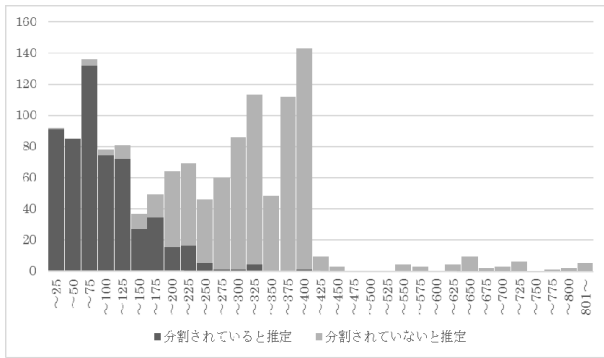
## 5.2 文書内役割の推定について

### 5.2.1 CRFを利用した文書内役割の推定

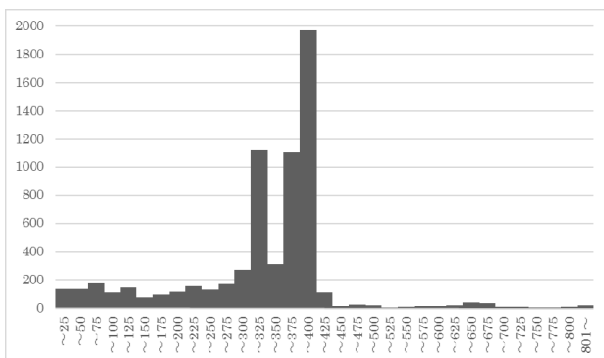
実験(1)における、CRFによる文書内役割特定における混同行列を表9に示した。

謝辞、謝辞タイトルのF値(表4)が低いのは、そもそも謝辞を掲載している論文数がデータセット中に少なく、十分な学習ができなかったためである。

最も目立つのは、本文と簡条書きの混同である。特に本文は最も矩形数が多いため、F値0.89と他の矩形と比較しても目立って低いわけではないにも拘わらず、全体の正解率を下げている。この原因を考察する。



横軸は幅 [px], 縦軸は矩形数  
 図 5: 分割された矩形の幅の分布



横軸は幅 [px], 縦軸は矩形数  
 図 6: 全矩形の幅の分布

当該矩形の文書内役割が「本文」「箇条書き」となる素性関数の中で、重みが大きいもの上位 10 位のうち共通するものが 4 含まれている。これは、本文と箇条書きには、類似点が多いということの意味する。そのため、ある程度の混同が生じると思われる。

そもそも、箇条書きは本文部分に記述されるものであるから、位置座標の点では、本文と大差ない。また、大きさについても、複数の項目を列挙する箇条書きでは、ある程度の行数と幅を持ち、本文と大差ない。大きく異なるであろう部分は、箇条書きの場合、最初の文字が「・」や「(」などの記号になり得るという点である。実際に、重み 5 位に「最初の文字が記号」が来ている。逆に言えば、箇条書きと本文を区別するためには、レイアウト情報では足りず、行頭の記号だけでなく、それに続く文の長さ、体言止めかどうか、などのテキスト情報が必要になると思われる。

続いて、図、表、数式の混同を減らす方法について検討する。矩形認識の誤りがない場合であっても、F 値は高くない。また、混同行列を見ると、この 3 つはお互いに混同されていることが分かる。いずれも、レイアウト情報を考えると、ある程度の大きさを持った矩形になっていると考えられる。テキスト情報につい

ては、図の場合はグラフなどの文字列が認識される場合があり、表の場合は各セルの文字列が、数式はアルファベット・数字が認識される場合がある。これらの文字列の区別が付きにくい場合があると想像される。3 つの文書内役割を特定する際に用いられる素性関数の重みを調べたところ、大きいもの上位 1 位に来るのは、いずれも「フォントサイズが不明」である。文字列を認識しなかった際にもこの値が設定されるためである。他に上位に来ているものとしては、図・表の場合にはキャプションの情報がある。表については、表キャプションは通常表の上部に来るため、読む順序に並べ替えた後、表キャプション→表の順に矩形が並ぶので、表キャプションさえ特定できれば、表の特定は難しくない。実際に、表を特定する素性関数の 3 位には、「1 つ前の矩形が表キャプション」が来ている。ところが、図表は縦に連続して複数紙面に現れることがしばしばある。この場合、表キャプションと表が、あるいは図と図キャプションが交互に出現するため、図表とキャプションの上下関係が逆転されて学習される場合がある。実際に、図の特定にも、重み上位 3 位の素性関数に「1 つ前の矩形が図キャプション」を素性に持つものが来ている。図表数式の正解率を上げるためには、レイアウト情報、テキスト情報以外の情報、例えば一定の長さの曲線や直線の有無、数学記号の有無といった画像処理によって得られる情報などを用いる必要があると考えられる。

以上で本文と箇条書き、図表数式の混同について述べてきたが、検索意図と文書内役割という観点から、これらの混同の重要度について述べる。箇条書きは通常本文の一部と考えることができるため、これらを区別しない状況も十分にあり得る。また、図表に含まれる文字列や数値の情報のうち、重要な部分は本文中でも言及されることがあり、図表を区別して検索する必要がない場合も考えられる。これらを考慮し、箇条書きを本文として、図表数式を図として（表キャプションも図キャプションとする）文書内役割の特定を行った場合、正解率は表 8 の通りとなり、一定の上昇がみられる。

表 8: 文書内役割をまとめた場合の正解率

実験種別	正解率
人手+CRF	0.94
SVM+CRF	0.69
なし+CRF	0.70

表 9: 混同行列 (実験 (1))

		推定された文書内役割																			
		ack	ackt	auth	body	equ	fig	figc	fttr	ftnt	head	list	note	page	ref	reft	stitle	tab	tabc	title	
正解	ack	6			6																
	ackt		8														3	1			
	auth			192	5							1				1	5				2
	body	1		1	1823				18		2	136		9	2	7	3	7			
	equ					151	36	1										7			
	fig					34	210	2										41			4
	figc				8		3	221					10	1	3	1	2	1	1		19
	fttr								103												
	ftnt				7				2	80		17		2				4			1
	head										75										
	list		2			185			10	13		676	2	1	3	3	32				16
	note					3			2			6	3	1			1	1			3
	page													299							
	ref					7								20		96	1	1			
	reft					3			3			3		3	85		4				
	stitle			1	3				3			12		1				1083	1		1
	tab					11	36			3									224		
tabc			1	5		1	35													220	
title				4						2											96

## 6 おわりに

学術論文の構成要素抽出においては、PDF データを直接用いるよりも、論文画像を OCR ソフトウェアで認識した方が、スキャンデータしかない論文も利用でき、PDF 作成ソフトウェアの使用にも左右されないという利点がある。しかしながら、OCR による認識誤りの修正をしなければ、十分に実用的であるとは言えない。

認識誤りには、論文構成要素が複数に分割されてしまうものと、複数の論文構成要素が 1 つにまとめられてしまうものが存在する。本研究では、SVM を用いて分割された矩形の統合を行った。矩形のレイアウト情報から、分割されているか否かを学習・推定したが、レイアウト情報のみでは判別が難しいことが分かった。また、複数の構成要素が 1 つにまとめられる誤りの修正を行わなければ、文書内役割の推定の正解率は、一切の認識誤りがない場合と比べて、大きく低下することが分かった。

矩形認識の誤りがなければ、レイアウト情報を主要な素性として、CRF によって文書内役割を学習・推定する手法は、正解率 0.86 とおおむね良好な結果を出すことができた。本研究では、図、表、数式の混同ならびに本文、箇条書きの混同がみられたが、この混同を減らすためには、レイアウト情報に加え、テキスト情報や図形的情報を用いる必要があると言える。

## 参考文献

- [1] Aiello, M., Monz, C., Todoran, L., Worring, M.: Document understanding for a broad class of documents, *International Journal on Document Analysis and Recognition*, Vol. 5, No. 1, pp. 1–16 (2003)
- [2] 梶本達矢, 太田学, 高須淳宏: 学術論文からの構成要素抽出手法の改良, In *Proc. of The 7th Forum on Data Engineering and Information Management* (2015)
- [3] 加藤恒昭, 岩月憲一, 山口和紀: 文書構造に基づく対話的情報アクセスにむけて, 人工知能学会 第 10 回インタラクティブ情報アクセスと可視化マイニング研究会, pp. 1–8 (2015)
- [4] Klink, S., Dengel, A., Kieninger, T.: Document structure analysis based on layout and textual features, In *Proc. of International Workshop on Document Analysis Systems*, pp. 99–111 (2000)
- [5] Luong, M., Nguyen, T.D., Kan, M.: Logical Structure Recovery in Scholarly Articles with Rich Document Features, *International Journal of Digital Library Systems*, Vol. 1. No. 4, pp. 1–23 (2012)
- [6] 増田勝也, 丹治信, 植松すみれ, 美馬秀樹: 研究動向分析のための論文のデジタルテキスト化とマイニングシステム, 言語処理学会第 20 回年次大会発表論文集, pp. 792–795 (2014)
- [7] Ohta, M., Inoue, R., Takasu, A.: Empirical evaluation of CRF-based bibliography extraction from research papers, In *Proc. of the IADIS International Conference on Information Systems*, Vol. 7, No. 2, pp. 18–26 (2012)