D-005

問題文と数式を考慮した数学の類題検索
Search for Similar Problems in Mathematics Considering Texts and Formulas

居樹 彩乃[1)]　湯本 高行[1)]
Ayano Sueki　Takayuki Yumoto

1　はじめに

数学の問題の類似度算出技術は教育現場において重要である。そのため、数学の学習方法の中で解法を習得したい問題の類題を演習することは基本的な勉強法であるといえる。しかしながら、特に高校数学においては解法の種類が多く、類題を見つけるために時間がかかってしまう現状がある。その上、類題か否かの判断は指導者や学生の経験値に依存する。以上の点から、高校数学において類題を探すことは非常に困難であるといえる。
数学問題の類似性評価を扱う先行研究として、上江洲ら [1] 研究がある。その中では、数式を木構造と認識して類似度を算出し、日本語の類似度は単語の出現頻度で表現しており、非常に単純な計算問題のみしか扱っていなかった。また、BERT モデルを数式の構造を考慮可能とした事前学習モデルである MathBERT[2] も存在する。しかし、このモデルのみでは日本語を扱うことは不可能である。
そこで本研究では大問単位での類似度を算出し、高校数学のどの単元においても適用可能な手法を提案する。問題の類似度算出にあたって、文章部分数式部分でそれぞれ別のアプローチを行う。文章部分では数学用語の集合間類似度を算出し、数式部分では公式の学習済みモデルを用いたベクトルの類似度を求める。これら 2 つの類似度を用いて問題間類似度を算出する。

2　問題間の類似度算出方法

本研究では文章部分と数式部分で別の処理を行い、どちらの要素も考慮した類似度を算出する方法を提案する。数学の問題における「問題文」と「解説文」を結合したものを本研究では 1 つの問題とする。ここで、問題 p と問題 q の類似度を式 (1) と定義する。ただし、$sim_{text}(p,q)$ は文章の類似度、sim_{math} は数式の類似度、α はパラメータ $(0 \leq \alpha \leq 1)$ とする。

$$sim(p,q) = (1-\alpha)sim_{text}(p,q) + \alpha \cdot sim_{math}(p,q) \quad (1)$$

2.1　文章部分の類似度の算出方法

2.1.1　数学用語集合

数学の問題の文章部分においての類似性を判断する際に、共通の数学用語の有無が重要な役割を果たしていると仮定した。そこで数学用語集合は数学の公式集から収集する。本研究では『新課程 チャート式 基礎からの数学』(以降、青チャートと記載) 数学 IA、IIB[3][4] に記載されている公式一覧のページに含まれる日本語の抽出を行った。この抽出した日本語に対して ginza [1)]で形態素解析を行い、名詞と固有名詞の全 178 単語を数学用語集合とした。

2.1.2　文章部分の類似度

文章部分の類似度を判定するために、集合間の類似度を表す Jaccard 係数、Dice 係数、Simpson 係数で算出を行う。この中で、最も精度が良かったものを採用し、文章部分の類似度とする。それぞれの類似度算出方法に関しては、まず問題 i, j に含まれる用語集合をそれぞれ X_i, X_j、数学用語集合を K、問題 i, j に含まれる数学用語集合をそれぞれ Y_i, Y_j とすると、Y_i, Y_j は式 (2) で求めることができる。

$$Y_i = X_i \cap K, Y_j = X_j \cap K \quad (2)$$

これを用いて、それぞれの類似度は式 (3), 式 (4), 式 (5) で求める。

$$Jaccard = \frac{|Y_i \cap Y_j|}{|Y_i \cup Y_j|} \quad (3)$$

$$Dice = \frac{2|Y_i \cap Y_j|}{|Y_i| + |Y_j|} \quad (4)$$

$$Simpson = \frac{|Y_i \cap Y_j|}{\min(|Y_i|, |Y_j|)} \quad (5)$$

2.2　数式部分の類似度

数式の処理の全体像を図 1 で示す。まず、問題に含まれる全ての数式を文埋め込みモデルでベクトル化を行う。次にそれぞれの問題に含まれるベクトル化された数式同士のコサイン類似度行列を計算し、各行の最大値を抽出する。この最大値行列の平均値を問題 1,2 における数式部分の類似度とする。

2.2.1　数式のトークナイズ

まず、ここで述べる数式部分とは数学の問題を TeX 形式で表現した際に、￥[￥] または￥(￥) で囲まれている部分を抽出したものである。このとき抽出した数式部分の長さが 2 以下の場合は除外をしている。この理由として、抽出した部分に「点 A」の「A」などといったアルファベット 1 文字なども含まれてしまうことを防ぐためである。その上、数式部分は図 2 のように『=』または改行で区切り、それを 1 つの数式の要素としている。ただし、その数式に『=』が 1 つしか含まれていない場合は、区切らないまま処理を行っている。この処理に関しては、問題のレイアウト上の都合から改行のタイミングが異なることによって、複数の意味をもつ数式が同一のものと処理されることを防ぐためである。

これに対して MathBERT のトークナイザーでトークン化を行って以降の処理を行う。実際にトークナイズを行った例を以下の図 3 で示す。

MathBERT のトークナイザーを使用する理由としては、図 3 のトークン化例のように数式の構造を考慮した上でトークナイズすることが可能であるからである。

2.2.2　数式のベクトル表現

数式のベクトル化には教師なし SimCSE[5] を用いる。教師なし SimCSE では、同じ文から作成したベクトルのペアを正例ペアとし、違う文から作成したベクトルのペ

1）　兵庫県立大学大学院情報科学研究科
1）　https://github.com/megagonlabs/ginza

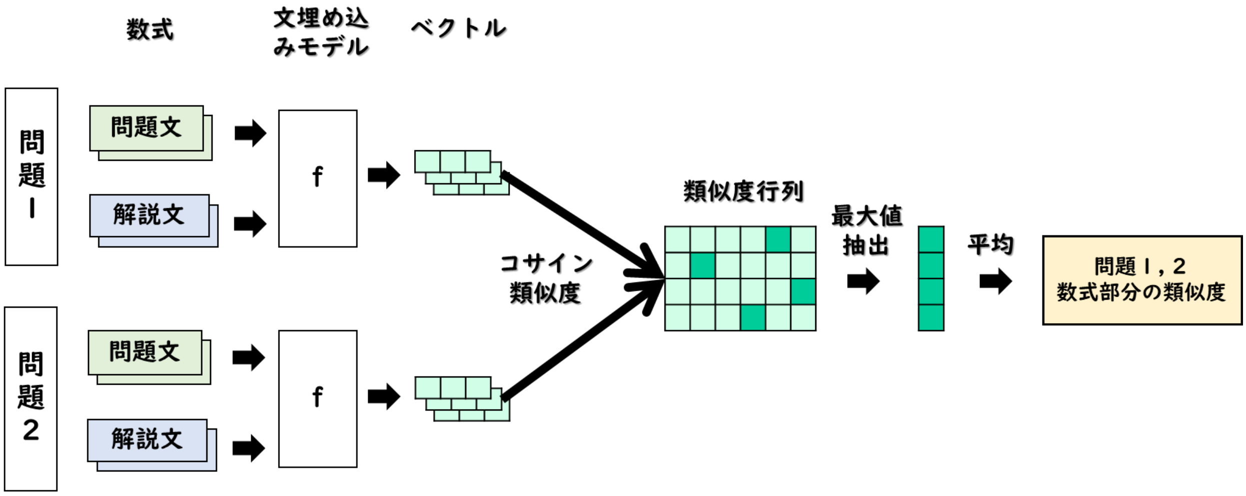

図 1　数式部分の類似度

分割する例

$x^2 + 4x + 4 = (x+2)^2 - 2^2 + 4 = (x+2)^2$

$\Rightarrow x^2{+}4x+4\,,\ (x+2)^2 - 2^2 + 4\,,\ (x+2)^2$

分割しない例

$x = 4$

図 2　数式の区切り例

元の数式　$\sin^2\theta = \frac{1}{2}$

TeX形式　¥sin^2 ¥theta = ¥frac{1}{2}

トークン化　'[CLS]', '¥¥', 'sin', '^', '2', '¥¥', 'theta', '=', '¥¥', 'frac', '{', '1', '}', '{', '2', '}', '[SEP]'

図 3　数式のトークン化例

アを負例ペアとして、正例ペアのベクトルの類似度が高く、負例ペアのベクトルの類似度が低くなるように文の埋め込みモデルを学習する。本研究ではデータセットは青チャート数学Ⅰ A、Ⅱ B に記載されている公式 (全 385 種) を TeX 形式で表現したものを用意し、ベースモデルとトークナイザーはともに MathBERT のものを使用した。これらを用いて正例ペアは 385 件、負例ペアは 147840 件で学習エポック数は 1、バッチサイズは 64 で学習を行った。文埋め込みモデルでは MathBERT の CLS トークンを使用し、内部で Dropout を行うことで毎回異なるベクトルを出力するようになっている。

次に問題文と解答文にそれぞれ含まれる数式を抽出し、MathBERT のトークナイザーを用いてトークン化を行う。これを入力として先述した事前学習モデルでベクトル化 (768 次元) を行った。

2.2.3　数式部分の類似度の算出方法

2.2.2 の方法でベクトル化を行った数式同士の類似度をコサイン類似度で算出する。これらの類似度を用い、問題間で最も類似度が高い組み合わせの平均をその問題同士の類似度とする。ここで、n 個の数式 f_i からなる問題 p と m 個の数式 g_j からなる問題 q の類似度 $sim_{math}(p,q)$ の算出方法は式 (6) と表すことができる。

$$sim_{math}(p,q) = \frac{1}{n}\sum_{i=1}^{n}\max_j \cos(f_i, g_j) \qquad (6)$$

3　実験

ここでは 3 つの実験を行う。まず 1 つめは 2.1.2 で示した 3 つの集合間類似度の中でどれが文章部分の類似度として適切か検証を行う。この中で最も精度がよかったものを文書部分の類似度算出方法として採用する。2 つめは式 (1) における適当なパラメータの調査である。全データと単元別においてどのパラメータで精度がよくなるのかを調査する。加えて、文章部分の類似度と数式部分の類似度がどのように問題の類似度を影響を与えているのかを調べる。3 つめはベースラインとの比較である。ベースラインと提案手法の精度などの比較を行って、提案手法の有用性を検証する。

3.1　テストデータについて

本研究では青チャートの数学 IA、IIB に記載されている内容を使用している。ただし、この青チャートに記載されている EXERCISES の問題と、それに対応する青チャートの例題のうち、文章部分以外を TeX 形式に変換する処理を行っている。これらの類題ペアを EXERCISES の 1 ページから 1 ペアずつ選択してテストデータとした。なお抽出する際には、ある EXERCISES の問題と対応する例題が 1 題となるような問題を原則として選んでいる。この類題ペアを 85 ペア用意し、評価に使用した。

3.2　集合間類似度の比較

3.2.1　実験方法

ここでは、Jaccard 係数、Dice 係数、Simpson 係数のうちのどれを用いるべきかを検証する。式 (3), 式 (4), 式 (5) を用いて文章部分の類似度を計算し、それぞれにおいて順位付けを行う。その後、上位 1 件、上位 5 件に正

解ペアが検出できているかを表す HitRate で比較を行う。ただし、*HitRate@k* は以下の式 (7) で定義する。

$$\frac{\text{上位}k\text{件の中に含まれている正解ペアの総数}}{\text{テストデータの総数}} \tag{7}$$

また同時に、「問題文と解説文の集合間類似度」、「問題文のみの集合間類似度」、「解説文のみの集合間類似度」、「問題文のみの集合間類似度と解説文のみの類似度の平均値」においてそれぞれ HitRate にどのような違いが見られるのか検証を行った。

3.2.2　実験結果と考察

上記の条件で Jaccard 係数、Dice 係数、Simpson 係数それぞれにおける HitRate の結果についてまとめたものを表 1〜3 に示す。

どの集合間類似度においても問題文と解説文で分割して考えるのではなく、まとめて考えた方が精度が良くなるという結果になった。これは、問題文のみは検出される数学用語が少ないため集合間類似度が 0 になったり、計算できないケースが多く存在したが要因だと考えられる。集合間類似度の算出方法の比較に関しては、極端に大きな差は存在しなかった。その原因としては数式用語集合に含まれる要素を事前に抽出しておくことにより、それぞれの集合の要素数が極端に異なる差が存在しなかったことであると考えられる。以上より、ここでは 4 つの算出方法の中でも安定した HitRate があり、類似性の包含関係も考慮できる Simpson 係数を日本語の類似度算出方法として採用する。

3.3　パラメータ α の設定

3.3.1　実験方法

ここではパラメータ設定に関する 4 つの実験を行った。まず 1 つめは、文章部分と数式部分の割合における最適なパラメータの調査である。式 (1) における α の値を 0 から 1 の範囲において 0.05 ずつ増加させて、最も HitRate が高くなるパラメータを調査した。2 つめは、単元別の最適なパラメータの調査である。テストデータに単元情報を追加し、1 つめと同様の手順でそれぞれの単元において最適なパラメータを調査した。なお、追加した単元とは青チャートにおいてその問題が含まれている章のタイトルのことである。3 つめの実験は、文章部分の類似度と数式部分の類似度の分布と HitRate の可視化を行った。ある問題とその類題の類似度が「文章部分の類似度のときのみ 1 位」、「数式部分の類似度のときのみ 1 位」、「どちらの場合でも 1 位」、「どちらの場合でも 1 位でない」の 4 つのラベル付をした上で可視化を行っている。4 つめは、問題の類似度の分布の特徴を調査するために、それぞれの問題を「代数」、「幾何」、「関数」、「データ」の 4 分野に分けてグループ化を行い、3 つめの実験の同様の可視化を行った。この分類は文部科学省が公開している高等学校学習指導要領 [6] を参考に独自に設定したものである。各単元との対応は以下の表 4 に示す。

3.3.2　実験結果と考察

問題全体で検証　まず、高校数学の問題全体の類似度算出における文章部分と数式部分のの最適なパラメータを調査するために、α の値 ($0 \leq \alpha \leq 1$) と、それを用いて算出した類似度に基づく HitRate の推移を図 4 に示す。

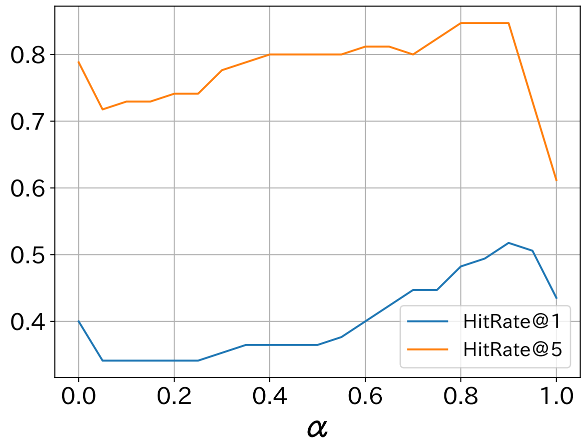

図 4　α の変化による HitRate の推移

図 4 より、式 (1) における最適なパラメータは 0.90 であるといえる。上位 5 位に関しては、パラメータが 1.0 の時を除いて全て HitRate が 0.7 を超えていることから、数学用語のみで大まかな単元の推定を行うことが可能であると考えられる。しかしながら上位 1 位の推移より、その単元の中で解法を判別するには数式の類似度を考慮する必要性があるといえる。これは同じ単元内では共通の数学用語が多く存在するため、用語のみでは判別しきれないことが要因だと考えられる。

単元別に検証　次に数学の単元の違いよる最適なパラメータの変化をまとめたものが表である。ここでは上位 1 件、上位 5 件における HitRate が最も高いときのパラメータを最適なパラメータとしている。それに加えて、上位 1 件と上位 5 件の HitRate の平均値が最も高くなるときのパラメータを一番右の列に記載している。

表 6 より「集合と命題」、「図形と計量」、「統計的な推測」は問題全体の最適なパラメータである 0.9 より値が低くなっている。これらの単元は、上位 1 位が同単元内であるものの類題ペアではない問題が多く存在した。ここから、この 3 つの単元は式変形よりも登場する数学用語の方が解法に影響しているといえる。

類題ペアの類似度分布　ここでは文章部分と数式部分の類似度が正解ペア発見にどのような影響を与えているのかを示している。なお、横軸が数式部分のみの類似度、縦軸が文章部分のみの類似度の分布を示したものが図 5 である。ただし、文章部分のときのみ 1 位を Top1(text)、数式部分のときのみ 1 位を Top1(math)、どちらの場合でも 1 位を Top1(both)、どちらの場合でも 1 位でないを not Top1 というラベルで表現している。

図 5 より、同じ問題でもかなり類似度の値に差があることがうかがえる。さらに、文章部分のみで類題を発見できているものは、数式のみの類似度においても発見できている。ここから使う用語が同じならば登場する式もおおよそ同じであると推測される。また文章部分の類似度が横並びの直線が見受けられることから、1 つの要素が含まれないとなった場合の影響が大きくなってしまっていると考えられる。

分野ごとの類似度分布　ここでは、代数 (algebra)、関数 (function)、幾何 (geometry)、データ (data) の 4 分野に分類した際の類似度分布を図 6 で示す。このとき、横軸を数式のみの類似度、縦軸を文章部分のみの類似度と

表 1　Jaccard 係数の HitRate

	問題文と解説文	問題文のみ	解説文のみ	類似度の平均
HitRate @ 1	0.482	0.212	0.953	0.212
HitRate @ 5	0.824	0.233	0.965	0.224

表 2　Dice 係数の HitRate

	問題文と解説文	問題文のみ	解説文のみ	類似度の平均
HitRate @ 1	0.482	0.412	0.329	0.447
HitRate @ 5	0.824	0.718	0.741	0.765

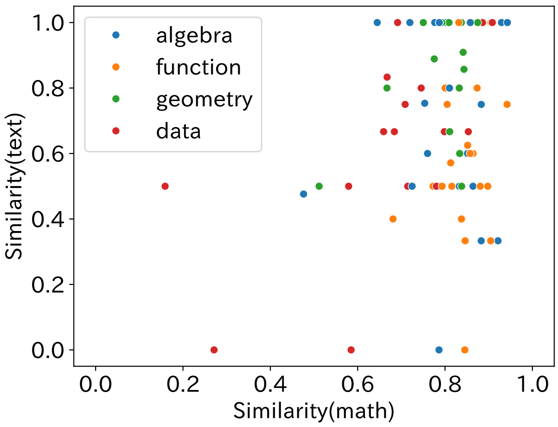

図 5　正解ペアの類似度分布

している。

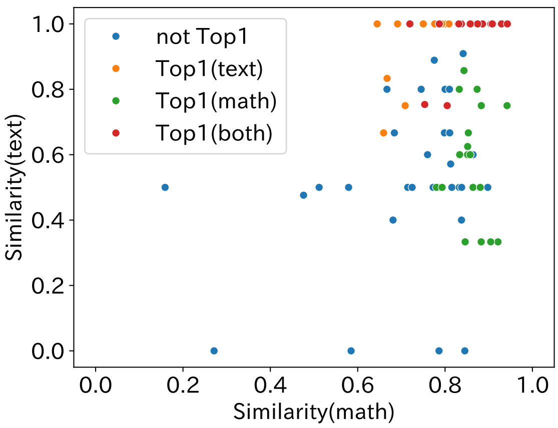

図 6　分野ラベルつき正解ペアの類似度分布

図 6 より、関数分野において正解ペアの数式部分の類似度は大きいが文章部分の類似度にばらつきがあることがうかがえる。ここから、関数分野の類似性において数式部分が重要な要素となっていると言える。一方、幾何の分野は他の分野と比べて文章部分の類似度が高くなっていることから、文章部分の類似度が重要となっていると言える。ここから、関数などは式変形、幾何はキーワードが類題と判定する際のポイントとなっていると考えられる。また、ここからデータの分野は他と比べて明らかに類似度が低いものがあるといえる。この要因として、「場合の数」などの単元は数式同士や数学用語が共通しているしているから似ているのではなく、解法の手順が似ていることが類題として判断する基準となっていることが考えられる。

3.4　ベースラインとの比較

3.4.1　実験方法

ベースライン手法として MathBERT と tf-idf を用い、提案手法との３つの比較を行う。まず１つめは、3.1 のテストデータを用いて 3.2.1 と同様の方法で HitRate を算出し、その比較を行った。このとき、提案手法のパラメータは 3.3 で算出した最適パラメータの 0.9 を使用している。一方、比較対象の手法の類似度は以下のように算出している。

MathBERT を用いた類似度算出方法　問題に含まれる数式を MathBERT のトークナイザーとモデルを使用してベクトル化行い、提案手法と同様に式 (6) を用いた値を類似度とした。

tf-idf を用いた類似度算出方法　テストデータとして使用した青チャートコーパスにおいて tf-idf を算出し、その問題同士のコサイン類似度を tf-idf での類似度とした。ここで、ある問題 q に含まれる単語 w の出現回数を $n_{q,w}$、問題 q における総単語数を N_q とする。このとき、ある問題 q に含まれる単語 w の出現頻度 $TF(q,w)$ を式 (8) で定義する。

$$TF(w,q) = \frac{n_{q,w}}{\sum_{i=1}^{N_q} n_{q,w_i}} \tag{8}$$

また、ある文書における単語 w の逆文書頻度 $idf(w)$ は、全文書数を $|C|$、単語 w が含まれる文書数を $df(w,C)$ として式 (9) と定義する。

$$idf(w,C) = \log \frac{|C|}{df(w,C)} + 1 \tag{9}$$

したがって、式 (8) と式 (9) を用いて tf-idf を以下の式 (10) で定義する。

$$TFIDF(w,q,C) = TF(w,q) \cdot idf(w,C) \tag{10}$$

２つめの実験は、正解ペアとならなかった問題の議論をするために各手法における上位 1 件の「正解ペア未発見数」「正解ペア単元一致率」「正解ペア単元 HitRate」の比較を行った。ここでの「正解ペア未発見数」はある問題と類似度が最も高い問題が正解ペアではなかった件数を表している。それに加えて、提案手法については不正解単元ペアの例を調査した。これはある問題と類似度が最も高い問題の単元が異なっているケースから数件を抽出したものである。

最後の３つめの実験では、各手法ごとに全ての問題ペアの類似度の分布を調査した。ここで使用した問題ペア

表 3　Simpson 係数の HitRate

	問題文と解説文	問題文のみ	解説文のみ	類似度の平均
HitRate @ 1	0.471	0.506	0.435	0.376
HitRate @ 5	0.835	0.753	0.77647	0.776

表 4　各単元の分類

分類	単元	
代数	数と式 数学と人間の活動 複素数の方程式	集合と命題 式と証明
幾何	図形と計量 図形と方程式	図形の性質
関数	2 次関数 指数関数と対数関数 微分法	三角関数 積分法 数列
データ	データ分析 確率	場合の数 統計的な推測

は、精度比較に使用した青チャートの EXERCISES の問題 85 題とそれに対応する類題 85 題全ての組み合わせであり、合計は $(85 \times 85 = 7225)$ ペアである。

3.4.2　実験結果と考察

各手法の HitRate の比較　提案手法、MathBERT、tf-idf それぞれについて、上位 1 件、上位 5 件の HitRate を算出し、まとめたものが表 6 である。

表 6 より、上位 1 件と上位 5 件どちらにおいても提案手法の HitRate が最も良いといえる。それに加え上位 5 件の結果をみると、どの手法も HitRate が 0.7 を越えていることがわかる。これは各単元におけるテストデータ数が 5 件前後であることから、どの手法も HitRate が高くなっている可能性が考えられる。したがって、単元を推定することにはベースラインの手法でも十分有効であるが、同じ解法の問題かどうかを判定するには文章と数式どちらのアプローチも必要なことがうかがえる。

未発見問題に関する比較　ここでは提案手法、MathBERT、tf-idf それぞれにおいて正解ペアが発見できなかった問題について、その総数と単元が一致しているのかをまとめたものが表 7 である。

表 7 より、単元推定のタスクにおいても提案手法が最もよい結果となった。ベースラインと比較しても、やはり日本語と数式の両方の要素を考慮して類似性を判定することで高い精度を実現できていると考えられる。次に、提案手法において正解ペアの単元も一致しなかった 23 件のうち、5 件について実際どの単元が類似度が最も高いとされたか示したのが表 8 である。

ここから、ほとんどの問題が単元は異なるものの同じ概念を用いる問題で、さらに解法において包含関係があることが分かった。例えば問題番号 77 は数列の計算の中で指数を計算する問題であった。このような例のように、最も類似度が高かった問題同士は何かしらの共通要素が存在していた。

各手法における類似度比較　提案手法、MathBERT、tf-idf それぞれについて精度比較に使用した不正解のものも含む問題間の類似度 $(85 \times 85 = 7225$ペア$)$ の分布を表したものが図 7 である。

表 5　単元別の最適なパラメータ α

	HR@1 が最大	HR@5 が最大	HR@1 と HR@5 の平均が最大
数と式	0.725	0.450	0.775
集合と命題	0.675	0.625	0.625
2 次関数	0.950	0.850	0.900
図形と計量	0.000	0.475	0.000
データ分析	0.825	0.450	0.775
場合の数	0.900	0.000	0.900
確率	0.850	0.775	0.850
図形と性質	0.925	0.900	0.900
数学と人間の活動	0.950	0.825	0.900
式と証明	0.800	0.450	0.800
複素数の方程式	0.900	0.325	0.325
図形の方程式	0.900	0.400	0.800
三角関数	0.775	1.000	1.000
指数関数と対数関数	0.500	0.900	0.900
微分法	0.925	0.650	0.925
積分法	0.775	0.500	0.775
数列	0.975	0.900	0.950
統計的な推測	0.000	0.475	0.000

※ HR は HitRate の略

表 6　各手法の HitRate

	提案手法	MathBERT	tf-idf
HitRate @ 1	0.518	0.471	0.388
HitRate @ 5	0.847	0.741	0.729

MathBERT で算出した問題同士の類似度は 0.9 前後に極端に多い上に最小値が 0.75 となっており、数式同士の「似ていない」という情報を表現しきれていなかった。一方、提案手法は幅広く類似度を算出できていることがうかがえる。この要因として、提案手法は数式をベクトル化する際に数学の公式集合を用いた教師なし SimCSE のモデルを使用している。そのため、数式に含まれる要素が似ていても公式としての違いをとらえ、数式同士の「似ていない」という要素を表現できたのではないかと考えられる。

4　おわりに

本研究では、文章と数式で別の処理によって高校数学問題同士の類似度を算出する手法を提案した。文章部分は数学用語における Simpson 係数を用いて類似度を算出を行った。数式部分においては、数式の公式集で教師なし SimCSE[5] を行ったモデルでベクトル化に変換した問題に含まれる数式同士でコサイン類似度を計算し、その最大値平均を数式の類似度とした。この 2 つの類似度を用いて問題の類似度を求める際の最適なパラメータは 0.9 であった。この算出方法は上位 1 件と上位 5 件どちらにおいてもベースラインよりも HitRate が良いとい

表 7 各手法の不正解のうち上位 1 件における単元 HitRate

	提案手法	MathBERT	tf-idf
正解ペア未発見数	41	44	53
正解ペア単元一致数	18	14	14
正解ペア単元 HitRate	0.439	0.341	0.264

表 8 提案手法における不一致単元ペアの例

問題番号	問題の単元	最高類似度の問題の単元
9	2 次関数	式と証明
30	確率	場合の数
35	図形の性質	図形と計量
40	数学と人間の活動	数と式
77	数列	指数関数と対数関数

う結果になった。さらに、単元推定においてもベースラインと比較して提案手法が最も良いという結果であった。ゆえに、数学の問題の類似度を算出する際に日本語と数式どちらものアプローチを行うことは有効であるといえる。

さらに、MathBERT のモデルをベースに数学の公式集のデータを用いてで教師なし SimCSE で学習を行ったモデルを作成した。これを用いて数式のベクトル化を行うことによって MathBERT のみでは表現できなかった、数式同士の「似ていない」という要素を表現することができていた。

今後の課題としては、複数の解法や単元を含む問題の処理である。本研究で使用したテストデータは原則、その問題に対して青チャート内の類題が 1 つであるようなものを選んでいる。それが 1 対多や多対多だった場合、類似度算出にどのような影響があるか議論する必要がある。

謝辞

本研究は JSPS 科学研究費助成事業 24K15195 による助成を受けたものです。ここに記して謝意を表します。

参考文献

[1] 上江洲 弘明, 谷口 哲也, 高井 勇輝, 西岡 圭太, 中川 勇人, 数学問題の類似性評価, 第 38 回ファジィシステムシンポジウム 講演論文集 (FSS2022 オンライン), 2022

[2] Shuai Peng, Ke Yuan, Liangcai Gao, Zhi Tang, MathBERT: A Pre-Trained Model for Mathematical Formula Understanding, arXiv:2105.00377v1 [cs.CL], 2021

[3] チャート研究所, 新課程 チャート式 基礎からの数学 I+A, 2022

[4] チャート研究所, 新課程 チャート式 基礎からの数学 II+B, 2022

[5] Tianyu Gao, Xingcheng Yao, Danqi Chen, SimCSE: Simple Contrastive Learning of Sentence Embeddings, Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pp.6894–6910, 2021

[6] 文部科学省, 高等学校学習指導要領（平成 30 年告示）解説 数学編 理数編, https://www.mext.go.jp/content/20230217-mxt_kyoiku02-100002620_05.pdf, 2018

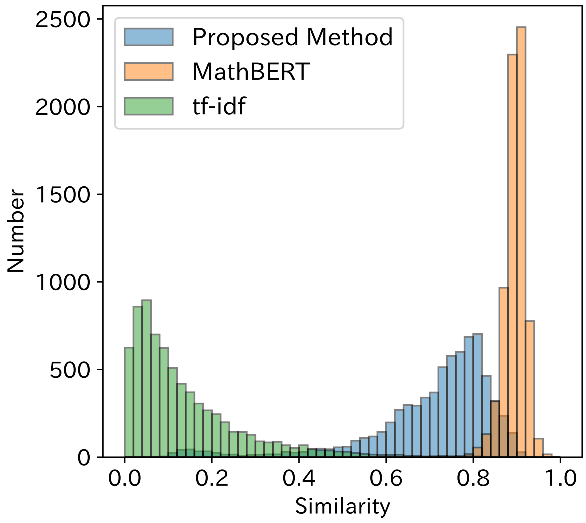

図 7 問題ペアの類似度の分布