

# 記述式試験対応採点処理システムの開発

田中 忠芳<sup>A</sup>, 益野 一哉<sup>B</sup>, 王 宝禮<sup>B</sup>

## 1 はじめに

大学入学後、学びのパラダイムは転換すると考えられ、学生の学びを転換するために、大学教育は重要な役割を担っている。とりわけ、歯科医学教育分野においては、専門基礎科目と専門科目との接続、専門知識の臨床への応用など、重要な課題が残っている。昨今の歯科医学を取り巻く環境の変化に呼応して、歯科医師養成の目標や過程が見直されてきている<sup>1)</sup>。

一般に、教育の現場においては、学修者の深い理解と思考力・判断力・表現力を育成するために、記述式試験問題を出題することが推奨される。ところが、記述式試験の採点作業において、量的かつ質的な負荷が教員に加わることから、記述試験問題の出題が敬遠され、学修者の深い理解と思考力・判断力・表現力の育成が後回しにされることが少なくない。

文部科学省は、平成29年7月13日に「大学入学テスト実施方針」を取りまとめ、その中で、大学入試センター試験に代わるテストとして平成32年度から「大学入学共通テスト」を実施すること、マークシート式問題における知識の深い理解と思考力・判断力・表現力を重視した作問への見直し、記述式問題の導入、英語4技能評価にかかる民間の資格・検定試験の活用などを示した<sup>2)</sup>。

現在、大学入試テスト委員会等が進めている試行調査(プレテスト)でも判明しているとおり、記述式問題の解答を採点するコスト、受験者の自己採点と実際の結果との開き等が問題視されている。

このように、記述式問題採点作業の負担軽減、記述式問題の採点コストの削減、採点基準の定式化などを目指した、記述式試験に対応した採点処理システム開発は、緊急の課題である。

本研究では、記述式問題の採点処理を e-Learning システムへと組み込むため、想定される解答を4パターンに分類し、これらを自然言語処理的なアプローチ等を用いることによって採点の半自動化を図った。解答群の類似度を比較したところ、記述式試験に対応し

た採点処理システムに資する有意義な知見が得られた。これらについて報告する。

## 2 自由記述形式への e-Learning システム対応

「大学入学共通テスト」は、新たな試験体制として、既存のマークシート形式、英語4技能評価にかかる民間の資格・検定試験の活用、記述式問題の導入を謳っており、e-Learning のシステム対応としては、既存の仕組みが利用可能な前者2つと異なり、自由記述を評価する必要のある記述式問題への対応が新たに必要となる。

記述式問題の解答パターンとして、次の4パターンを想定した。

- ・単語解 … 1単語にて構成される解答。
- ・短文解 … 80~120文字程度、2センテンス以内で構成される解答。規準としては第2回「大学入学共通テスト」プレテストで設定された最大文字数までとし、長文解については考慮しないものとする。
- ・数値解 … 1数値にて構成される解答。単位設定まで設問に指示されているものとする。
- ・数式解 … 数式で構成される解答。

想定した4パターンに対して、それぞれに対しての対応方法を検討した。

- ・単語解 → 解答として用意されている複数の単語についてシソーラスを参照し、設定された階層内に含まれる単語を取得し、重み付けを設定した上で、これとマッチングする。
- ・短文解<sup>3-5)</sup> → 形態素解析、文章マッチング解析。
- ・数値解 → 単純な完全一致。
- ・数式解 → 青木らによって開発・運用が進められている e-Learning システム WebStudy などを用いられている<sup>6)</sup> 数式処理ソフトウェア Mathematica が提供する標準型への還元処理を利用する。

## 3 解答の類似度比較の結果

今回、「炎症」を短文で説明する問いを設定し、その

A: 金沢工業大学

B: 大阪歯科大学

問いに対する正答例として、次の3通りの正答を準備した。

正答A：炎症とは、生体が微生物の侵入や物理的・化学的刺激などを受けて、5大徴候として発熱（熱感）・発赤（赤くなる）・腫脹（はれる）・疼痛（いたみ）などからの症状を呈し、局所に機能障害を起こすことである。

正答B：炎症とは、生体の恒常性を構成する解剖生理学的反応の一つであり、恒常性を正常に維持する非特異的防御機構の一因である。生体に、これらの異常が生じると5大徴候として、発赤、発熱、腫脹、疼痛、機能障害を特徴とする。

正答C：炎症は、発赤、はれ、熱、痛みを4特徴とする病変であり、機能障害がつけ加えられて5つの徴候を表わすものが炎症の5大徴候である。炎症は物理的・化学的外力や微生物など、なんらかの刺激に対して生体組織の示す一種の防衛反応とされている。

また、誤答例として、次の誤答を作成した。

誤答：炎症についてはよく分からないが、私はオレジとリンゴではリンゴが好きだ。特に青森産のリンゴが好きだ。私は以前は青森に住んでいたが、今は東京に住んでいる。

これらの文章間の類似度比較をした。その結果を表1に示す。文章間の類似度としては cosine 距離を利用している。また、文章のベクターデータへの変換は形態素解析を用いて品詞へと分解した後、名詞情報に対して TF-IDF 値を算出した<sup>7)</sup>。

表1 文章間の類似度比較

文章比較パターン	類似度 [%]
誤答 vs. 正答 A	0.948%
誤答 vs. 正答 B	0.667%
誤答 vs. 正答 C	2.43%
正答 A vs. 正答 B	26.2%
正答 A vs. 正答 C	40.8%
正答 B vs. 正答 C	23.1%

誤答と正答3つとの比較結果と、正答間での比較結果との間で、1~2桁程度の差が得られた。なお、値が高いほど類似したものと判断される。

#### 4 考察と展望

表1の結果から、今回開発したシステムを用いることで、記述式試験問題に対応した採点処理が十分に可能であることが示唆された。今回 TF-IDF 値の算出の際、IDF は正答も回答も合わせて算出したが、これを分割するなどによって本手法の拡張性を活かした判定精度の向上が見込まれる。

短文解について、現行では「解答例+解答規準」という形式になっており、「解答規準」についても併せてチェックする必要がある。「解答規準」は、今後、大きく変わっていく可能性もある。また、現在の「解答規準」仕様のみをトレースするのではなく、「e-Learning のシステムとして自由記述対応するならば」といった姿勢でシステム開発の方が妥当であると考えられる。

今後、規準として、たとえば非正解例の追加などが考えられ、それらを想定してシステム構築を進めていく予定である。

#### 引用・参考文献

- 1) 三浦公嗣：歯学教育に求めるもの，日本歯科医学教育学会雑誌，vol.24，2008，p.127-131.
- 2) 文部科学省：高大接続改革「大学入学者選抜改革について」大学入学共通テスト  
[http://www.mext.go.jp/a\\_menu/koutou/koudai/detail/1397731.htm](http://www.mext.go.jp/a_menu/koutou/koudai/detail/1397731.htm) (2019年12月26日参照)
- 3) 高橋勇，宮川勝年，小高知宏 他：Web サイトからの剽窃レポート発見支援システム，電子情報通信学会論文誌. D，情報・システム，vol.90，no.11(通号 431)，2007，p.2989~2999.
- 4) 村田剛志：参照の共起性に基づく Web コミュニティの発見，人工知能学会論文誌，vol.16，no.3 B，2001，p.316-323.
- 5) 松尾豊，石塚満：語の共起の統計情報に基づく文書からのキーワード抽出アルゴリズム，人工知能学会論文誌，vol.17，no.3 D，2002，p.217-223.
- 6) 青木謙二，久保田真一郎，鍵山茂徳：e ラーニングによる理系基礎学力評価システムとコンテンツ開発，情報処理学会研究報告コンピュータと教育 (CE)，vol.2005，no.123(2005-CE-082)，2005，p.25-32.
- 7) 徳永健伸：情報検索と言語処理，東京大学出版会，1999.