

採点者の意図を加味した自由記述採点補助システム

朝日 信成, 藤森 誠, 松田 基弘, 伊藤 一成, Martin J. DÜRST

青山学院大学理工学部

1 はじめに

試験では主に選択形式, 穴埋め形式, 記述形式などの問題形式が用いられている. 特に記述形式の設問は, 公平な評点を下すことが他の形式に比べ, 困難である. なぜなら, 解答の自由度が大きく, また大量に採点を行うことで採点者の採点基準が揺らぐ恐れが生じるからである. 実際, 採点に慣れてくると点数の基準を上げ, 既に採点した解答と差をつけて採点する傾向があるという報告がある [1]. これをハロー効果と呼ぶ. そこで, 計算機に採点作業を完全に任せる自動採点システムが多く提案されている [2].

英語を対象とした研究が先行しているが, 近年日本語もシステムの構築や自動採点方法などに関する成果が出始めている [3][4].

先行研究では, 数百字から数千字程度の長い文章記述や小論文を主な採点対象としている. 一方我々は, 事象の説明や短い意見を述べさせる小規模自由記述形式を主な研究対象としている. 論述形式では, 起承転結といった文章の構成や文章のつながり方が採点のポイントとなるが, 小規模自由記述形式では物事を的確な言葉で説明されているかや出題者の問いに対して肯定/否定, 賛成/反対なのかといった点がポイントとなる. 今回は, 文章を構文解析器にかけて, その文書構造に基づいて基準点を定めることとしている. しかしながら, 表層的な解析技術だけで, 点数を全自動で決定することは現在の技術レベルからいって非常に危険である. そこで本稿では, あくまで採点者側に主導権があり, その採点行為を補助するインタラクティブ型の採点システムを試作したので, 報告する.

2 構文解析を利用した基準点作成

まず, 重要語句の出現の有無で基準点を算定する単純な方法が考えられる. しかし, 文中で正しく重要語句が使用されているかはまた別問題であり, 意味もな

A Grading Support System for Free Form Problems Considering the Grader's Intention
Nobushige ASAHI, Makoto FUJIMORI, Motohiro MATSUDA, Kazunari ITO and Martin J. DÜRST
Department of Integrated Information Technology, College of Science and Engineering, Aoyama Gakuin University
5-10-1 Fuchinobe, Sagamihara, Kanagawa 229-8558, Japan
{nobushige,fujimori,matsuda}@sw.it.aoyama.ac.jp,
{kaz, duerst}@it.aoyama.ac.jp

```

彼女は——|
20 歳ではないと——|
                いうのは——|
人に——|
                よって——|
                とらえ方が——|
                異なるので, ——|
                命題ではない.

```

図 1: 構文解析結果の例

く重要語句を羅列した解答にも高得点を付与する恐れが生じる. この問題を回避するためには単語が文の構成要素として正確に使用されているか否かも判断すべきである. 文節の存在を意識した単語の検索を行うことで文章の構成要素として使用されているか判断する. 特定の依存関係や係り受け関係が存在する場合, その関係に対して採点者が定めた点数を加算することで個々の解答に対して採点基準を作成する. 単なる単語の羅列には分節は存在しないのでこの方法を用いることで単語の羅列と文章を分けて評価できる. 本稿で説明する依存関係とは文章を構文解析した際に得られる分節の係り受け関係からわかる単語の関係性のことである. 説明のために“彼女は 20 歳ではないというのは人によってとらえ方が異なるので, 命題ではない.”という文章を構文解析器 KNP[5]を利用して解析した例を図 1 に示す. 結果から“異なるので”は“命題ではない”に対して係り受け関係を持つと判定できる. 単語が含まれる文節に係り受け関係があるので, それぞれの単語同士も係り受け関係に近いものがあると判断し, これを依存関係とする. また同一の分節に含まれる場合にも依存関係があると判断する.

3 プロトタイプシステム

今回実装したプロトタイプシステムについて概説する. 関係の登録フェーズと採点フェーズから構成される.

3.1 関係の登録

関係の登録フェーズについてスクリーンショットを参照しながら説明する. スクリーンショットを図 2 に示す. 画面左側には解答者の解答一覧が表示される (図 2

の1)．氏名，学籍番号等の個人情報表示されないで，先入観なしに採点できる．その情報を元に解答群の中に存在する単語をクリックするか，テキストボックス(図2の2)に直接キーボードから文字列を入力することで解析の対象とする単語対を登録できる．ちなみに解答群に含まれる動詞，形容詞，副詞，名詞などの出現頻度をグラフ形式で表示させることも可能である．登録済みの単語対の全情報は単語入力用テキストボックスの下に提示される(図2の3)．解答群中に多く含まれる単語は解答群の傾向を表しているため，単語選択に有益な情報である．

解答群はあらかじめKNPで構文解析されてデータベースに格納されている．そのデータを基に，指定された単語間の依存関係の有無を調べ，それぞれの解答がその依存関係を含むか否かを表示する(図2の4)．さらに新しい関係の登録を決める情報として指定したそれぞれの単語に依存関係を持つ頻度が高い単語を提示する(図2の5)．これを参考に再び単語を入力することで新しい単語間の依存関係を順次登録していく．

“命題ではない”，“命題とは言えない”，“命題ではありません”など同じ意味表現でも様々な記述が存在する．それぞれの表現ごとに繰り返し点数を決めることは負担であり，また揺れが生じる原因にもなる．代表的な例だけ対応すればよい仕組みが望ましい．よって複数の単語群を類似した意味をなす言葉のクラスとして登録できるようにした．これにより，一度に多くの表現に対応することができる．さらに一度登録したクラス情報はデータベースに格納され，別問題の関係の登録作業や採点作業に再利用できる．既に登録されている依存関係に使われている単語と同じクラスに新しく単語を登録した場合，各解答の依存関係の有無表示を更新する．

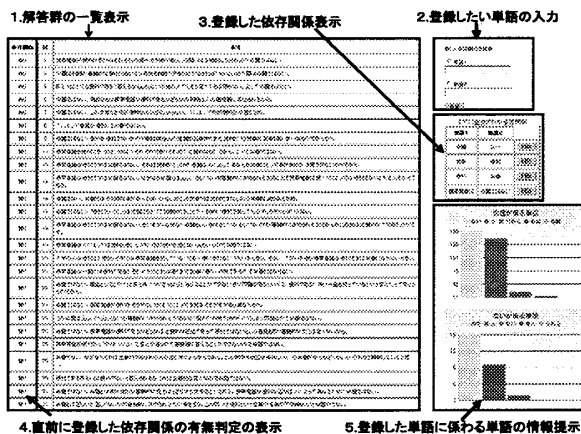


図2: 関係の登録時のスクリーンショット

3.2 採点

次に採点フェーズについて説明する．スクリーンショットを図3に示す．先に登録した依存情報に対応するスライダバーを操作することで点数を設定できる(図3の1)．各点数を更新すると各解答に基準点が表示され(図3の2)．円グラフに基準点の分布が示される(図3の3)．基準点は，条件が満たされる依存関係の点数の和としている．またそれぞれの解答の横についてテキストボックスに直接点数を入れることで点数の補正を行うこともできる(図3の4)．デフォルトでは基準点が入力されている．最終的に登録ボタンが押された時点でテキストボックスに入力されている値が採点となる．

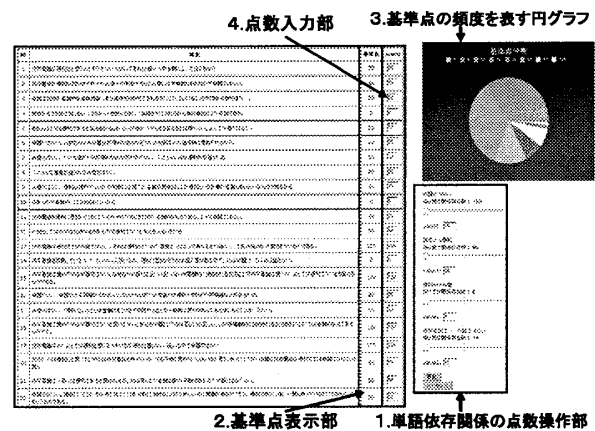


図3: 採点時のスクリーンショット

4 まとめ

採点者主導の採点補助システムのプロトタイプを構築した．今後追加機能の実装や評価実験を行っていきたいと考えている．

参考文献

- 池田央：テストの科学 試験にかかわるすべての人に，日本文化科学社(1992)．
- 石岡恒憲：小論文およびエッセイの自動評価採点における研究動向，人工知能学会誌，Vol. 23, No. 1, pp. 17-24 (2008)．
- 長坂悦敬：Web 環境における記述問題自動採点エンジンの開発，情報処理振興事業協会平成13年度成果報告集(2001)．
- 笠原要，天野成昭，佐藤哲司：小論文の自動採点，人工知能学会 第18回全国大会(2004)．
- 黒橋禎夫，河原大輔：日本語構文解析システムKNP version 2.0 使用説明書．