

自然言語表現に基づく学生アンケート分析システム

A Student Questionnaire Analysis System based on Natural Language Expressions

酒井 哲也[†]
Tetsuya Sakai

石田 崇[‡]
Takashi Ishida

後藤 正幸[§]
Masayuki Goto

平澤 茂一[¶]
Shigeichi Hirasawa

1. はじめに

近年、通信基盤の整備およびワープロ・メールソフトの普及により、教育現場において教員が配布する教材や学生が提出する答案が電子化テキストとして蓄積されることは常識となりつつある。蓄積されたテキストデータから有用な知識を発掘する技術としては情報検索やテキストマイニングがあり、製品化も行われている(例えば [12])。本研究では、自由記述形式の学生アンケートに対して情報検索技術を適用した新しい分析手法を示し、その精度評価を行い、活用例を示す。学生の考えていることを効率的に把握し、分析結果を授業改善やクラス分けなどの支援に役立てることが我々の最終目標である [1, 2, 3, 4, 5, 9]。なお、アンケートデータは選択式質問・回答と記述式質問・回答から成っており、[1, 3, 4]では両方を用いた統計処理を試みているが、本研究では記述式部分のみに着目している。

2. 従来研究との比較

教育分野におけるテキストデータに対して情報検索技術を適用したものとしては、例えば長坂ら [7]、森田ら [6]の研究がある。長坂ら [7]は、市販の情報検索エンジンを用いて学生の解答と模範解答の双方をベクトル化し、類似度計算に基づく記述問題の採点支援を行っている。この手法は、タームの一致だけでなく共起も考慮しているため「概念検索」と呼ばれるが、このようにボトムアップにテキストデータをベクトル化する分析法では、分析者独自の視点を導入することが難しい。また、通常の情報検索技術ではテキストデータから抽出した自立語をもとにベクトルを生成するため、例えば付属語を含む「～したい」「～したいわけではない」などの微妙な自然言語表現に着目した分析を行うことが難しい。これに対し我々の手法では、意味役割解析技術 [8] の適用により、例えば「学生の望んでいること」や「学生が不安に思っていること」といった独自の視点をトップダウンに設定し、自然言語表現に着目した分析を行うことができる。

一方、森田ら [6]は、記述問題の解答を頻度の高いタームにより分類するシステムを提案している。しかし、テキスト中に頻出する語は必ずしもそのテキストを特徴づける語とはいえない。情報検索の分野では、タームが多くのテキストに出現するものか、あるいは特定のテキストに偏って出現するものかに基づく重み付け (*idf* と呼ばれる) が有効であることが知られている。我々のシステムでは、この考え方を発展させた確率型検索モデルに基づき代表語および代表文の抽出を行っている [9]。この方法では、ひとつ以上のテキストにおいてタームが出現す

る確率と、上記以外のテキストにおいてタームが出現する確率の比較により代表語を決定するため、森田らのようにストップワードを設定せずともテキスト中にたまたま頻出する語を排除できる。さらに本研究では、クラスタリングおよび意味役割解析の適用により、例えば「多くの学生が不安に思っていることは何か」「授業に対する要望でユニークなものはあるか」などの視点から代表語および代表文を選出することが可能である。

3. 意味役割解析を用いた分析システムの提案

本アンケート分析システムは、情報検索システム BRIDGE [8, 10] を用いて実装されており、意味役割解析により自然言語の言い回しに着目した分析を行う。

意味役割解析は、以下の手順で正規表現マッチングによりテキストデータに任意のラベルを付与する [8]。

1. テキストデータを形態素解析する。
2. 「テキスト分割ルール」により、形態素解析結果を文や節などの単位(単位文と呼ぶ)に分割する。
3. 「意味役割解析ルール」により、各単位文に0個以上の意味役割を付与する。(意味役割解析ルール間には優先度を設定することができる。)

これにより、例えば「～したい」という文に「意思」という意味役割を付与したり、人名・地名の品詞タグが付与された文に「人名」「地名」という意味役割を付与することができる。なお現状では、意味役割解析ルールの正規表現の収集を人手で行う必要がある。

意味役割解析は、検索の高精度化 [8]、検索結果の提示 [10]、質問応答 [11] などにおいて活用されている。本システムも、ある学生の自由記述と類似した他の学生の自由記述を意味役割の一致度に基づき検索する「類似学生検索機能」や、各自由記述から意味役割やタームの統計量に基づいて文を抜粋する「意味役割要約」機能を有している。本稿では、5章にてアンケートを対象とした場合の意味役割解析の精度評価を行い、次に6章にて上記の解析結果を活用して実際に分析を行った例を示す。

4. アンケートの収集方法

本研究では、早稲田大学理工学部経営システム工学科の学生を対象とした講義「コンピュータ工学」の初回に実施したアンケートの自由記述部分を分析対象とした。2002年度と2003年度の二度にわたって実施したアンケートのうち2002年度分(132名分)を意味役割解析ルール作成のための「学習データ」として用い、2003年度分(135名文)を未知の「評価用データ」として用いた。

図1に2003年度分の自由記述の設問を示す。なお、2002年分の自由記述の設問は「コンピュータについて考えていることを自由に書いてください」という単一の質問であった [9]。2003年度に回答フィールドを設けた理

[†](株) 東芝 研究開発センター 知識メディアラボラトリー

[‡]早稲田大学大学院 理工学研究科

[§]武蔵工業大学 環境情報学部

[¶]早稲田大学 理工学部

(コンピュータ知識(自由記述))「コンピュータ」に関するあなたの「知識」と「経験」について自由に述べてください。(250 - 300文字)

(将来設計(自由記述))あなたは将来どのような業界(例えば、銀行、電機メーカー、商社等)、職種(例えば、営業、人事、研究等)に就職したいと考えていますか?自由に述べてください(250 - 300文字)

(講義のイメージ(自由記述))「コンピュータ工学」と聞いてどのような講義を想像していますか?講義で扱われると思うテーマや、自分にとってのこの講義の位置づけ等について自由に述べてください(200 - 250文字)

(分野への意欲(自由記述))コンピュータに関連する分野についてどのようなテーマに興味を持っていますか?自由に述べてください(250 - 300文字)

(講義への要望(自由記述))どのような講義を希望していますか?講義の内容や運営方法(例えば、評価方法、レポート課題、試験、出席管理、教材等)に対する要望を述べてください(250 - 300文字)

図 1: 2003 年度「コンピュータ工学」初回アンケートの自由記述設問部分

由は、学生に情報を確実に記入してもらうため、また、システムによる分析を容易にするためである。

5. 意味役割解析の評価実験

「コンピュータ工学」のアンケート自由記述を分析するため、今回は図 1 の設問を参考にして以下の 8 つの意味役割を定義し、2002 年分の自由記述をもとに意味役割解析ルールを作成した。

1. WILLING:YES 意欲, 意思がある!「~したいと思う」「頑張る」などの表現を含む。
2. WILLING:NO 意欲, 意思がない!「~したくない」「~する気がおきない」などの表現を含む。
3. INTERESTED:YES 興味がある!「興味がある」「関心がある」「楽しみだ」などの表現を含む。
4. INTERESTED:NO 興味がない!「興味がない」「関心がない」「拒否反応がある」などの表現を含む。
5. KNOWLEDGE:YES 知識, 理解がある!「知識が増えた」「理解した」などの表現を含む。
6. KNOWLEDGE:NO 知識, 理解がない!「知識不足」「理解できなかった」「~さえ知らない」などの表現を含む。
7. REQUEST 授業への要望!「講義に期待する」「~をお願したい」などの表現を含む。
8. ANXIETY 不安なこと, 心配事!「心配だ」「気にかかる」「怖い」などの表現を含む。

なお、今回は例えば WILLING:YES と WILLING:NO は同時に付与されることがないようにルールを構成した。このため「~はやりたいが~はやりたくない」のような表現を含む文があってもいずれか一方のみが付与される。一方、例えば「~には興味があるので是非やりたい」という文に対して WILLING:YES と INTERESTED:YES の両方を付与することは可能である。

上記の意味役割解析ルールを用いて、未知のデータすなわち 2003 年分の自由記述 135 名分 (3,662 文) に対する意味役割解析の精度評価を行った。正解データは人手で作成し、評価指標としては再現率 (recall)、適合率 (precision)、F-measure を用いた。システムが適切な意味役割を付与した文数を *hit*、システムが適切な意味役割を付与し損ねた文数を *miss*、システムが不適切な意味役割を付与した文

数を *false_alarm* とすると、 $recall = hit / (hit + miss)$ 、 $precision = hit / (hit + false_alarm)$ 、また

$$F\text{-measure} = \frac{(1 + \beta^2) * precision * recall}{\beta^2 * precision + recall} \quad (1)$$

である。ここでは $\beta = 1$ (再現率と適合率の重要度を同等と見なす) として評価する。

表 1 に 2003 年分の各文に対してシステムが付与した意味役割および正解意味役割の回答フィールド別分布を示す。(括弧内が正解文数を表す。) 今回の実験では回答フィールドを参考にして意味役割を設定したが、例えば回答フィールド「コンピュータ知識」には KNOWLEDGE:NO (知識, 理解がない) に分類されるべき文が多く、次に WILLING:YES (意欲がある) に分類されるべき文が多いことがわかる。また、回答フィールド「講義への要望」には当然 REQUEST (要望) に分類されるべき文が多いが、その他にも意欲や興味、知識に関する記述が混在していることがわかる。

表 2 に 2003 年分の自由記述に対する精度評価結果を示す。F-measure の平均は 0.61 であり、これは例えば人手の要約との比較によるテキスト要約の評価値の水準と同程度であり、概ね良好と言える。KNOWLEDGE:YES や WILLING:NO に関する評価値が低いのは、評価用データ (2003 年分) と同様、学習データ (2002 年分) においてもこれらに該当する例文が非常に少なかったためである。今後、年度毎に学習データを蓄積し、意味役割解析ルールを改良すれば未知のデータに対する分類精度は向上すると考えられる。

6. 意味役割を利用したアンケート分析例

6.1 代表文選択における利用

我々は [9] において、情報検索における擬似適合フィールドバック技術の応用により、指定された 1 つ以上の自由記述から代表語および代表文が簡単に抽出できることを示した。即ち、自由記述の形態素解析結果より得られた各ターム t に対して選出基準 ow を算出し、 ow の高いものから T 個選んで代表語とする。

$$rw_t = \log \frac{(r_t + 0.5)(N - n_t - R + r_t + 0.5)}{(n_t - r_t + 0.5)(R - r_t + 0.5)} \quad (2)$$

$$ow_t = r_t * rw_t \quad (3)$$

ここで、 N は自由記述の総数、 n_t はこのうち t を含むものの数、 R は指定された自由記述の数、 r_t はこのうち t を含むものの数である。また、代表文のスコアは上記 T 個の代表語に関する rw の和により算出できる [9]。

ここでは上記の方法で得られた代表文をさらに意味役割により絞りこんで分析を行う例を示す。表 3 は、2003 年分の自由記述に対して回答フィールド毎に階層型クラスタリングを行い [9]、クラスタ数 50 における最大クラスタ (すなわち多数派の学生の集まり) から代表語および代表文を選出し、さらに代表文の中で 1 つ以上の意味役割が付与されたものを表示したものである。代表語、代表文の数はそれぞれ 10 および (紙面の制約上) 3 としており、代表文中の代表語には下線を付与している。なお、

表 1: システムが付与した意味役割 (括弧内は正解意味役割) の回答フィールド別分布

	コンピュータ知識	将来設計	講義のイメージ	分野への意欲	講義への要望	合計
WILLING: YES	86 (54)	390 (320)	145 (82)	204 (176)	154 (20)	979 (652)
WILLING: NO	0 (0)	17 (12)	4 (4)	5 (0)	1 (0)	27 (21)
INTERESTED: YES	32 (25)	98 (70)	34 (23)	239 (214)	34 (3)	437 (335)
INTERESTED: NO	6 (4)	4 (4)	12 (8)	12 (16)	5 (4)	39 (36)
KNOWLEDGE: YES	15 (25)	0 (0)	11 (1)	6 (0)	2 (0)	34 (29)
KNOWLEDGE: NO	119 (112)	13 (3)	13 (28)	41 (13)	34 (16)	252 (172)
REQUEST	2 (1)	0 (0)	0 (5)	5 (8)	271 (379)	283 (393)
ANXIETY	8 (6)	5 (0)	5 (8)	2 (1)	6 (2)	29 (17)

表 2: 意味役割解析の評価結果

	miss	hit	false alarm	recall	precision	F-measure ($\beta = 1$)
WILLING: YES	34	618	361	0.95	0.63	0.76
WILLING: NO	8	13	14	0.62	0.48	0.54
INTERESTED: YES	37	298	139	0.89	0.68	0.77
INTERESTED: NO	13	23	16	0.64	0.59	0.61
KNOWLEDGE: YES	22	7	27	0.24	0.21	0.22
KNOWLEDGE: NO	44	128	124	0.74	0.51	0.60
REQUEST	131	262	21	0.67	0.93	0.78
ANXIETY	3	14	15	0.82	0.48	0.61
平均				0.70	0.56	0.61

形態素解析の時点で異表記吸収処理を行っているため、例えば「コンピュータ」と「コンピューター」、「レポート」と「リポート」は同じタームとして扱われている。

例えば「コンピュータ知識」の最大クラスタを見ると、学生の大半は「コンピュータ工学」といっても「パソコン」「ワード」「エクセル」といったイメージしか持っておらず比較的知識が少ないという事実が迅速に把握できる。また、「将来設計」の最大クラスタを見ると「経営システム工学科で学んだことが役に立たなくてもよい」という学生や「ここで学んだ知識や技術を生かした職業に就きたい」という学生を対比させて概観することができる。同様に「講義への要望」の最大クラスタでは「学生の評価の際には出席やレポートを重視してほしい」という旨の要望やこれと反対の要望を概観することができる。

「講義のイメージ」および「分野への意欲」の最大クラスタでは、代表文に2つの意味役割が付与されているものがある。このうち、*がついているものは false alarm、すなわち「興味のあるものかもわからない」「興味を持つということはない」という表現が INTERESTED: YES の「興味のある」「興味をもつ」という表現を含むルールに過剰マッチしてしまったものである。前述の通り、意味役割解析の精度は今後向上させていく予定であるが、意味役割をあくまで参考として分析を行う限り、現状の精度でもある程度有用であると思われる。

また、以上では多数派意見に焦点を当てた分析例を示したが、クラスタリング結果から少数派意見・ユニークな意見に着目して同様の分析を行ったり、成績と関連付けて分析を行ったりすることも可能である [9]。

6.2 テキスト視覚化のための利用

個々の学生の自由記述を精読する場合にも意味役割解析結果は利用できる。図2は、2名の学生の自由記述を、付与された意味役割別に集計して対比しながら視覚化した例である。学生Aについては WILLING: YES に分類さ

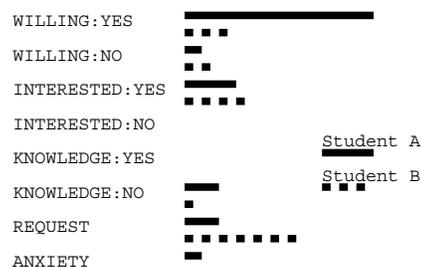


図 2: 学生 A, B の自由記述分類結果表示例

れた文が多く、学生Bについては REQUEST に分類された文が多いことがわかる。意味役割解析結果が完璧でなくとも、このような視覚化により「この学生は積極性がありそうだ」「この学生は心配事が多そうだ」といったおおまかな印象を踏まえた上で文章を読み進めることは有用であると思われる。特に、学生の心配事や要望は見落とさないことが望ましいので、これらに対する意味役割解析の再現率は重要となってくる。なお従来は、アンケート結果を教員が精読し (a) 文系指向か理系指向か (b) 基礎重視か応用重視か (c) 学習意欲の有無などを主観的に判断し手作業により分類を行っていた。

7. まとめ

本研究では、自然言語の言い回しに着目した学生アンケート分析システムを提案した。意味役割解析の精度は、アンケートデータの収集を継続し学習データを増やすことによりさらに向上が可能である。今後、本システムを実際に授業改善やクラス分けに役立て、またテストの答案などアンケート以外のテキストデータへ適用範囲を拡張することを検討していく。

謝辞 本研究プロジェクト開始当初からアンケート収集・

表 3: クラスタ数 50 の場合の各最大クラスタの代表語上位 10 件, および意味役割が付与された代表文上位 3 件

スコア	意味役割	代表文
「コンピュータ知識」の最大クラスタ (代表語: エクセル ワード 大学 授業 パソコン 高校 機会 パワー 一年 コンピュータ)		
11.3	INTERESTED: YES	高校の時の授業でパソコンを少しがじり, ... 大学に入学して今までやったことのなかったワード, エクセル, プログラミングなどの授業を通して, ほとんどできなかったパソコンが少しづつできるようになり興味ができました.
10.6	KNOWLEDGE: NO	大学に入り, 授業で, ワードやエクセルを扱うようになり, ... ちょっとは使えるようにはなったものの, パソコンの原理などについては, まだ無知な状態です.
10.1	WILLING: YES	一年のときにワードやエクセルをやった程度なのでこの機会にコンピューターについて知識をつけて, IT 化の現在において社会に出ても十分通用するように自由に使いこなせるようになりたいと思います.
「将来設計」の最大クラスタ (代表語: システム 経営 工学 就職 将来 希望 職種 職業 業界 金融)		
9.7	WILLING: YES	... 将来どんな職種につきたいかという点, 直接的に経営システム工学科で学んだことが役に立たなくてもよいと思っている.
9.4	WILLING: YES	経営システム工学科の目的である企業などの組織における目的達成または問題発見と問題解決を日々考えるような職業に就きたいということです.
9.4	WILLING: YES	せっかく浪人してまであこがれの経営システム工学科に入学できたわけですから, ここで学んだ知識や技術を生かした職業に就きたいと思っています.
「講義のイメージ」の最大クラスタ (代表語: 言葉 意味 勉強 自分 興味 仕組み 内容 必修 講義 専門)		
8.2	WILLING: YES	... 興味のあるものもわからないですけど, 学科が指定した必修科目ということは, 絶対自分のためになる科目だと思うので, ちゃんと勉強をしていきたいと思っています.
8.0	WILLING: YES	... コンピュータの歴史や仕組みなどの専門的な内容を勉強するよりは, “パソコンを使ってみんなでホームページを作ってみよう” というほうが楽しいと思う...
7.7	KNOWLEDGE: NO	... コンピュータに接してきたなかで, 何度か見たこと聞いたことのある言葉や仕組みなどがあるのですが, その意味をあまり正確に把握していなかったり, まったく知らなかったりといったことがよくありました.
「分野への意欲」の最大クラスタ (代表語: 自分 分野 興味 コンピュータ 質問 知識 授業 プログラム 言語 勉強)		
7.0	INTERESTED: YES*	コンピューターに詳しくないので, 特に分野に興味を持つということはないが, 自分のホームページを作るために HTML を勉強したりはしたことがある.
6.2	KNOWLEDGE: NO	INTERESTED: YES
6.2	WILLING: YES	INTERESTED: YES
6.2	INTERESTED: YES	INTERESTED: YES
6.2	WILLING: YES	INTERESTED: YES
6.2	INTERESTED: YES	INTERESTED: YES
「講義への要望」の最大クラスタ (代表語: 授業 出席 評価 期末 確認 リポート 重視 試験 理解 テスト)		
9.8	REQUEST	... 評価方法も出席やレポート重視ではなく, 最終的な期末試験における知識の定着, 活用具合によって決めて欲しいと思います.
8.8	REQUEST	評価方法としては試験のみの評価ではなく毎回の出席やレポートなども加味した評価で 1 時間 1 時間をさらに集中できるよりよい授業を期待しています.
8.4	REQUEST	評価方法に関しては, 期末試験はもとより, 出席点, レポートの出来具合なども加味してほしいと思う.

分析に協力してくれた伊藤潤君 (現 任天堂) に感謝する.

参考文献

- 後藤, 酒井, 伊藤, 石田, 平澤: 選択式・記述式アンケートからの知識発見, PC カンファレンス, pp. 43-46 (2003).
- 後藤, 伊藤, 石田, 酒井, 平澤: ベイズ統計を用いた文書ファイルの自動分析手法, 経営情報学会 2003 年度秋季全国研究発表大会, pp. 28-31 (2003).
- 平澤, 石田, 伊藤, 後藤, 酒井: 授業に関する選択式・記述式アンケートの分析, 私立大学情報教育協会平成 15 年度大学情報化全国大会, pp. 144-145 (2003).
- 石田, 伊藤, 後藤, 酒井, 平澤: 授業モデルとその検証, 経営情報学会 2003 年度秋季全国研究発表大会, pp. 226-229 (2003).
- 伊藤, 石田, 後藤, 酒井, 平澤: PLSI を利用した文書からの知識発見, FIT 2003 論文集 Vol. 2, pp. 83-84 (2003).
- 森田, 北, 高瀬, 林: 記述式の解答を即時に講師が把握するためのシステム, FIT 2002 情報技術レターズ LN-1, pp. 233-234 (2002).
- 長坂悦敬, 阿手雅博: 記述問題の自動評価を目指した教育支援情報システムによる Interactive Education, 情報教育方法研究第 3 巻, 第 1 号, pp. 37-42 (2000).
- 酒井, 小山, 鈴木, 真鍋: 意味役割解析に基づく高適合英語文書の検索, FIT 2002 情報技術レターズ LD-8, pp. 67-78 (2002).
- 酒井, 伊藤, 後藤, 石田, 平澤: 情報検索技術を用いた効率的な授業アンケートの分析, 経営情報学会 2003 年度春季全国研究発表大会, pp. 182-185 (2003).
- Sakai, T. et al.: BRIDGE over a Language Barrier: Cross-Language Information Access by Integrating Translation and Retrieval, *IRAL 2003 Proceedings*, pp. 65-76 (2003).
<http://acl.lidc.upenn.edu/W/W03/W03-1109.pdf>
- Sakai, T. et al.: ASKMi: A Japanese Question Answering System based on Semantic Role Analysis, *RIA O 2004 Proceedings*, pp. 215-231 (2004).
- 東芝 KnowledgeMeister:
http://cn.toshiba.co.jp/prod/km2/index_j.htm