

[英語]

# 1 英語問題への挑戦から分かった 技術的課題



東中竜一郎 (NTT コミュニケーション科学基礎研究所)

東ロボくんの2016年の英語の模試(ベネッセ模試)の成績は95点(200点満点)であった。英語の模試の最初の受験が2014年でそのときの成績がまさに同じ95点であったので、点数だけを見るとこの2年間何もしていなかったように見えるが決してそうではない。この2年間で多くの取り組みを重ねてきており、実際、今年の東ロボくんはこれまでで一番賢かった。図-1は歴代の東ロボくんに、ここ3回の模試を解かせたときの点数を示したものである。しかし、東ロボくんは今伸び悩んでいる。図-2は今年の東ロボ英語ソルバーの問題別(各問題の回答手法は後続のコラム参照)の得点率を示したものである。これを見ると、会話文完成問題以降について、まったく点数を伸ばせていない。これらは、複数文問題・長文問題であり、一文を超えた処理・文脈の理解が必要となってくるものである。深層学習を含むさまざまな手法を試みたものの点数を伸ばせなかった。試験問題のような高度な処理を行うための学習データ量が十分でないことや、文章の意味をどう扱うべきかという基本的な問題への取り組みがまだまだ足りないことが課題である。

(2017年3月31日受付)

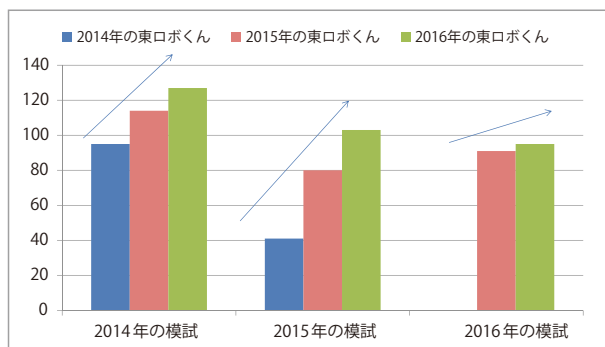


図-1 歴代の東ロボくんの得点比較。出題形式の変化のため2014年の東ロボくんについては2016年の模試の得点を算出できなかった

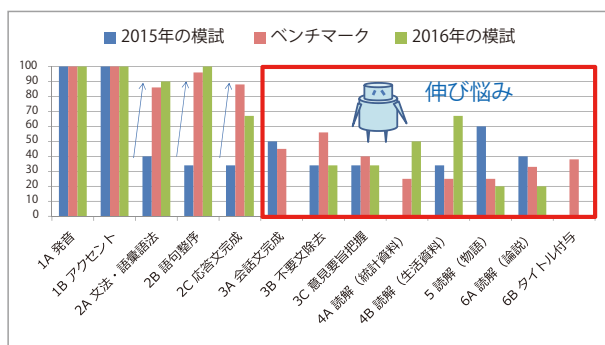


図-2 問題別の得点率。「2015年の模試」は2015年の東ロボくんの結果。「ベンチマーク」と「2016年の模試」は2016年の東ロボくんのベンチマークデータ(独自に準備した評価データ)と2016年の模試に対する結果

東中竜一郎(正会員) higashinaka.yuichiro@lab.ntt.co.jp

1999年慶應義塾大学環境情報学部卒業。2001年同大学院政策・メディア研究科修士課程、2008年博士課程修了。博士(学術)。2001年日本電信電話(株)入社。現在、NTTコミュニケーション科学基礎研究所所属。対話システムや質問応答システムの研究に携わる。

## センター試験英語問題における単文・複数文問題の自動解答手法

杉山弘晃 (NTT コミュニケーション科学基礎研究所) 堂坂浩二 (秋田県立大学システム科学技術学部)  
南 泰浩 (電気通信大学情報理工学研究所) 成松宏美 (NTT コミュニケーション科学基礎研究所)

文法・語彙語法問題は、単文中の空所に最もふさわしい単語・句を、4つの候補の中から選ぶ穴埋め問題である。東ロボ英語ソルバーは、空所に各選択

肢を埋め込んだ文の対数尤度を言語モデルで計算し、もっともらしい選択肢を選択する。19億文の大規模コーパスで言語モデルを学習し、大学入試センタ

## 1 英語問題への挑戦から分かった技術的課題

一試験の過去問等を使って評価したところ、解法の正解率は 85.7% であった。

また、単語並べ替え問題である語句整序問題や、語句を組み合わせて会話の応答文を作る応答文完成問題に対して、同様の手法で回答したところ、過去問の正解率はそれぞれ 96.2%, 87.5% であった。こうした単語の並びに関する問題では、言語モデルによる手法でおおむね回答可能であることが分かる。

会話文完成問題は会話中の空所に相応しい文を 4 つの選択肢から選び、会話文を完成させるという問題である。この問題を解くため、東口ボ英語ソルバーは、会話の流れの自然さを推定し、最も自然な流れとなる選択肢を選ぶ<sup>1)</sup>。会話の流れの自然さは、隣接発話らしさと感情極性の流れの自然さの 2 つのスコアの重み付き和とする。隣接発話らしさは 2 つの発言が会話中で隣接して現れる確率を表し、感情極性の流れの自然さは連続する発言間で感情極性の変化が小さいほど高い値をとる。各スコアの算出はデータから Support Vector Machine (SVM) により学習した認識器により行う。大学入試センター試験の過去問等を使って評価したところ、正解率は 45% であった。なお、Deep Neural Network (DNN) を用いた手法も開発し比較してみたが、その正解率はほとんどランダムベースラインと変わらず、SVM と組み合わせても 34% 程度にとどまった。会話の流れを DNN で自動的に推定するには膨大なデータが必要となるため、現時点では人間が適切な特徴量を選定する必要があると言える。

不要文除去問題は、文章中に埋め込まれた不要な文を 4 つの文候補から選択する問題である。この問題に対し、word2vec (単語間距離を共起関係に基づいて推定するモデル) に基づく手法、Entity Grid (談話の結束性) に基づく手法、Convolutional NN (CNN) に基づく手法を試したが、word2vec に基づく手法の正解率が 56.7% と最も高く、ほかはランダムなベースラインと大差なかった。

意見要旨把握問題は、4 つの要旨候補の中から対話中の話者の要旨として最も相応しいものを選択

するという問題である。この問題に対して、word2vec を用いる手法、CNN を使う手法、attentive reader を使う手法、Word Mover's Distance を使う手法、end-to-end memory networks を使う手法の 5 種類の DNN を使った手法で、その有効性を調べた。この中では、word2vec による手法が最も正解率が高く、その正解率は 40% であった。以上のように、各種の複雑な DNN による手法を用いても、高い正解率を示すことはなかった。

以上より、単語の並びに関する問題は精度良く回答できるようになっているものの、意味理解を必要とする問題ではランダムより多少良い程度にとどまっており、DNN を用いても簡単には精度が向上しないことが分かった。

### 参考文献

- 1) 東中, 杉山, 成松, 磯崎, 菊井, 堂坂, 平, 南, 大和: 「ロボットは東大に入れるか」プロジェクトにおける英語科目の到達点と今後の課題, 人工知能学会全国大会 (2017). (2017 年 3 月 31 日受付)

#### ■杉山弘晃 (正会員) sugiyama.hiroaki@lab.ntt.co.jp

2007 年東京大学工学部機械情報工学科卒業。2009 年同大学院情報理工学系研究科知能機械情報学専攻修士課程修了。同年日本電信電話 (株) 入社。人と自然な対話を行う雑談対話システムの研究に従事。博士 (工学)。人工知能学会, 言語処理学会各会員。

#### ■堂坂浩二 (正会員) dohsaka@akita-pu.ac.jp

1984 年大阪大学基礎工学部情報工学科卒業。1986 年同大学院修士課程修了。同年日本電信電話 (株) 入社。2012 年より秋田県立大学システム科学技術学部教授。博士 (情報科学)。人工知能学会, ACM 各会員。

#### ■南 泰浩 (正会員) minami.yasuhiro@is.uec.ac.jp

1986 年慶應義塾大学理工学部電気工学科卒業。1988 年同大学院修士課程修了。1991 年同大学院博士課程修了。同年日本電信電話 (株) 入社。2014 年より電気通信大学情報システム学研究科教授。工学博士。言語処理学会, 電子情報通信学会, 音響学会, IEEE 各会員。

#### ■成松宏美 (正会員) narimatsu.hiromi@lab.ntt.co.jp

2009 年津田塾大学学芸学部情報数理工学科卒業。2011 年電気通信大学大学院情報システム学研究科修士課程修了。同年日本電信電話 (株) 入社。現在電気通信大学大学院博士課程在学中。電子情報通信学会, IEEE 各会員。

## センター試験英語問題における長文読解問題の自動解答手法

磯崎秀樹 菊井玄一郎 (岡山県立大学情報工学部)

平 博順 (大阪工業大学情報科学部) 大和淳司 (工学院大学情報学部)

次に、読解問題とリスニング問題について解説する。統計資料や生活資料を見て答える読解問題では、図表の解析が必要である。そこで、オープンソースの画像処理ソフト OpenCV と、文字認識ソフト Tesseract OCR を使い、表を読むツールとグラフを読むツールをルールベースで作成した。折れ線グラフ、棒グラフ、円グラフに対応したが、棒グラフと折れ線グラフが重なっているものもあり、多様性への対応が課題である。図表の情報と文章の情報から答を導くところも課題である。

論説の読解問題のうち、内容一致問題は長文の内容に関して最も合致する選択肢を解答する問題である。述語項構造解析や、word2vec などを用いて、本文と選択肢の文のペアで最も意味が似ている選択肢を選ぶ手法<sup>1)</sup>を用いたところ、過去問 15 問中 9 問 (60%) と比較的高い正答率が得られていた。しかし、大きく言い換えられている問題や、常識判断が必要な問題が含まれていた 2016 年度の模試では 5 問中全問不正解になってしまった。今後、構文が大幅に異なる言い換えや常識判断への対応を検討する必要がある。

タイトル付与問題は、長文中の指定された 4 ~ 5 個の段落に対して、それぞれに適切なタイトルを、段落と同数の選択肢から重複なく選ぶ問題である。東ロボ英語ソルバーは、選択肢と段落をそれぞれベクトルにして、その間のコサイン類似度が最大のものを選ぶ。ベクトルは、word2vec によるベクトルに Inverse Document Frequency (IDF, 文書頻度に基づく、単語の珍しさを表す値) をかけたものとした。段落中の全単語ではなく、一部の文だけを使う方法もいくつか試したところ、段落中の選択肢と類似度が一番高い文を用いたときが最も成績が良く、42 問中 19 問完答 (45%) だった<sup>1)</sup>。

リスニング問題には、例年 2 問程度、選択肢が

イラストの問題がある。イラスト選択肢を英文テキストに変換し、ほかのリスニング問題と同様に解くアプローチを試した。画像キャプションの手法と選択肢の差異に着目した説明文生成を組み合わせた。イラスト理解の学習データには、東京大学相澤・山崎研究室による Manga109 データベース<sup>☆1</sup>を利用した。現状イラスト認識については、30 程度のクラスへの分類精度でも 40% 程度であり、解答を導き出す上での貢献には至っていない。今後、学習データの追加によるイラスト認識単体での精度向上と、問題文語彙制約の利用によるキャプション生成高精度化の両面から改良に取り組みたい。なお、技術的な詳細については、文献 1) を参照されたい。

### 参考文献

- 1) 東中, 杉山, 成松, 磯崎, 菊井, 堂坂, 平, 南, 大和: 「ロボットは東大に入れるか」プロジェクトにおける英語科目の到達点と今後の課題, 人工知能学会全国大会 (2017). (2017 年 3 月 31 日受付)

■磯崎秀樹 (正会員) isozaki@cse.oka-pu.ac.jp

1983 年東京大学工学部計数工学科卒業。1986 年同大学院修士課程修了。同年日本電信電話 (株) (NTT) 入社。基礎研究所, ソフトウェア研究所, コミュニケーション科学基礎研究所を経て, 2011 年より岡山県立大学情報工学部教授, 博士 (工学)。

■菊井玄一郎 (正会員) kikui@cse.oka-pu.ac.jp

1984 年京都大学工学部電気工学科卒業。1986 年同大学院修士課程修了。同年日本電信電話 (株) (NTT) 入社。情報通信処理研究所, ATR 音声翻訳通信研究所などを経て, 2011 年より岡山県立大学情報工学部教授, 博士 (情報学)。

■平 博順 (正会員) hirotoshi.taira@oit.ac.jp

1994 年東京大学理学部卒業。1996 年同大学院修士課程修了。同年日本電信電話 (株) (NTT) 入社。1996 年 NTT コミュニケーション科学研究所, 2005 年 NTT データ 技術開発本部などを経て, 2014 年より大阪工業大学情報科学部准教授, 博士 (工学)。

■大和淳司 yamato@cc.kogakuin.ac.jp

1988 年東京大学工学部精密機械工学科卒業。1990 年同大学院修士課程修了。同年日本電信電話 (株) (NTT) 入社。1998 年マサチューセッツ工科大電気工学およびコンピュータ科学科修士課程修了。2016 年より工学院大学情報学部教授, 博士 (工学)。

☆1 [www.manga109.org](http://www.manga109.org)