

2019 年 11 月 18 日

(報道発表資料)

日本電信電話株式会社
大学共同利用機関法人 情報・システム研究機構 国立情報学研究所

2019 年大学入試センター試験英語筆記科目において AI が 185 点を獲得！

日本電信電話株式会社(代表取締役社長:澤田 純、以下、NTT)は、NTT グループの AI 関連技術 corevo®^{※1}の研究開発、および大学共同利用機関法人 情報・システム研究機構 国立情報学研究所(所長:喜連川優、以下、NII)の人工知能プロジェクト「ロボットは東大に入れるか」^{※2}(以下、東ロボ)の一環として、大学入試センター試験(以下、センター試験)の英語筆記科目に挑戦した結果、2019 年センター試験の英語筆記本試験において、185 点(偏差値 64.1)の極めて高い成績を達成しました。NTT がこれまで継続的に取り組んできた、センター試験の英語科目の自動解答技術や、深層学習を用いた自然言語処理に関する知見を適用することで、最新の深層学習による機械読解技術を単純に適用した場合の成績に対し、30 点以上(偏差値約 7 相当)の改善を実現しました。

【研究の背景・意義】

NTT のコミュニケーション科学基礎研究所^{※3}(以下、NTT CS 研)では、機械翻訳や情報検索、対話処理などに応用可能なコンピュータによる自然言語処理、知識処理の基礎研究に取り組んでいます。

人間は社会の実際的な問題を解くために、さまざまな自然言語処理を統合的に行っています。たとえば、ある特定の話題について、文献を調査しレポートにまとめるといった問題は、文書の内容を理解し、それらの要約をまとめ、文の流れを整えるという自然言語処理が統合的になされる必要があります。今後、自然言語処理の技術が、より人間に役立つものとなるためには、人間が扱うような社会の実際的な問題に対して、有効性を評価していく必要があります。

「東ロボ」プロジェクト(プロジェクトディレクター:新井紀子教授)は、NII が中心となって 1980 年以降細分化された人工知能分野を再統合することで新たな地平を切り拓くことを目的に、若い人たちに夢を与えるプロジェクトとして発足しました。具体的には、センター試験や東京大学の第 2 次学力試験を用い、人間が実際に解く問題を人工知能がどこまで解けるのかを明らかにすべく研究活動を進めています。この中で、英語問題は、自然

言語処理、知識処理の統合的な問題を多く含みます。NTT CS 研では、「東ロボ」プロジェクトを自然言語処理、知識処理の基礎研究を進めるベンチマークと捉え、センター試験に含まれる多様な英語問題に対する自動解答に関する知見を積み重ねてきました。

一方で、近年深層学習に基づく文書読解技術が急速に進展しています。その最新技術である XLNet^{※4}は、大規模テキストによる事前学習を行ったベースモデルに、問題の性質に合わせた転移学習^{※5}を施すことで、異なる種類の問題を比較的少量のデータから効率的に解くことを可能にしています。しかしながら、学習に利用できるデータが大きく不足している問題や、解答に辞書的な情報が不可欠な問題では、十分な精度で解答することが困難でした。

今回、XLNet では解答が困難であった、不要文除去・段落タイトル付与・発音問題について、NTT CS 研を中心とした東ロボ英語チームの独自技術を適用しました。その結果、2019 年本試験では点数が大きく伸び 185 点（200 点満点、適用前 154 点）、受験者中の偏差値は 64.1（独自技術適用前 57.0）を達成しました。また、同じ技術を過去 3 年間のセンター本試験・追試験に対して適用した結果も、安定して偏差値 60 以上を達成できました（図 1）。

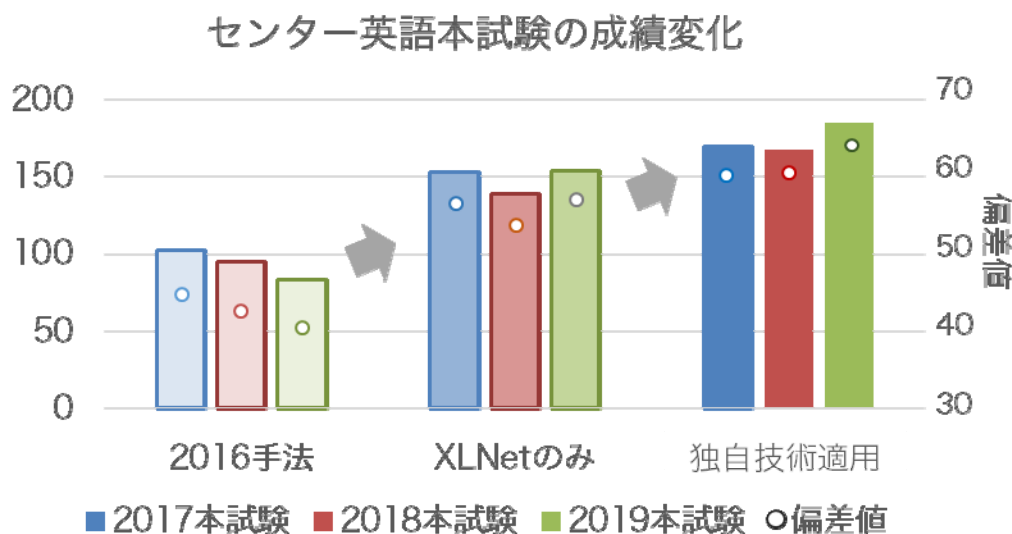


図 1：センター英語本試験の成績変化

【技術のポイント】

(1) 自動的に作成した疑似問題を用いた不要文除去問題の高精度化

機械学習では、正解・不正解となるデータの双方を入力することで、適切にそれらを分類するモデルを学習します。しかしながら、文章から不要な文を見つける不要文除去問題（図 2）は、比較的新しい問題であり、通常の文章には不要な文が含まれないことから、学習に用いるデータを集めることが難しいという問題点がありました。そのため、深層学習でも 2017-2019 年のセンター試験の本試験・追試験の 15 問中 6 問のみ正解と、十分な精

度が得られないことが課題でした。それに対し、不要文を含まない通常の文章から、文の順序を組み換えて擬似的に不自然な流れの文章を作成することで、大量の不要文除去問題を自動作成する手法を考案しました。こうして得られた問題を機械学習に利用することで、15 問全てに正答することができました。また、過去問・独自作成問題を集めた評価用のベンチマークデータに対しても、従来手法による正答率 60%を 86%まで高めることができました。

(2) 段落タイトル付与問題の自動解答

段落タイトル付与問題は、文章中の各段落の内容を表すタイトルを選択肢から選ぶ問題です。この問題を解くには、段落中の各文それぞれを理解するだけでなく、文章全体が表す内容を大まかに把握し、文章全体の構造を理解することが重要です。全ての段落のタイトルを同時に正答しないと正解と認められない完答型の問題のため、ランダムに解答した場合の正答率は 4%程度です。また、構造が特殊な問題であるため、XLNet 等の文書表現技術をそのまま適用することはできません。NTT CS 研を含む東ロボチームでは、各段落と選択肢の類似度を計算し、最適な段落・選択肢の組み合わせを導く手法を考案しました。その結果、2017-2019 年のセンター試験の本試験・追試験の 5 問全てについて正答することができました。また、類似度計算に文書表現技術の 1 つである BERT^{*4}を適用することで、ベンチマークデータに対しても 80%と高い正答率を得ることができました。

(3) アクセント・発音問題に対する辞書を活用した自動解答

センター試験では、実際の受験生の能力を測るため、アクセント・発音問題、穴埋め問題、長文読解と非常に幅広い種類の出題がなされます。NTT CS 研ではこれらの問題に 2014 年より継続的に取り組んでおり、各問題を解答する上で最良の手法の組み合わせの検討や、独自の問題作成を進めてきました。例えば発音の正しさを選ぶアクセント・発音問題は、辞書の発音記号と対応付けて解答する必要があるため、現在の深層学習による文書読解技術では対応できない問題です。この問題には、深層学習ではなく、あえて発音辞書を地道に調べる方法を適用し、表記ゆれを抑える工夫や問題解析器の精度を高めた（図 3）ことで、ほぼ満点の成績を得ています。

これらの英語問題における技術的な前進は、岡山県立大学 菊井玄一郎教授、秋田県立大学 堂坂浩二教授、大阪工業大学 平博順准教授、電気通信大学 南泰浩教授、工学院大学 大和淳司教授らと NTT CS 研との共同研究によるものです。NTT CS 研からは、杉山弘晃研究主任、成松宏美研究主任、東中竜一郎上席特別研究員が共同研究に参画しています。

A 次の問い(問1～3)のパラグラフ(段落)には、まとまりをよくするために取り除いた方がよい文が一つある。取り除く文として最も適当なものを、それぞれ下線部①～④のうちから一つずつ選べ。

問1 27

When flying across the United States, you may see giant arrows made of concrete on the ground. Although nowadays these arrows are basically places of curiosity, in the past, pilots absolutely needed them when flying from one side of the country to the other. ① The arrows were seen as being so successful that some people even suggested floating arrows on the Atlantic Ocean. ② Pilots used the arrows as guides on the flights between New York and San Francisco. ③ Every 16 kilometers, pilots would pass a 21-meter-long arrow that was painted bright yellow. ④ A rotating light in the middle and one light at each end made the arrow visible at night. Since the 1940s, other navigation methods have been introduced and the arrows are generally not used today. Pilots flying through mountainous areas in Montana, however, do still rely on some of them.

図2：不要文除去問題の例（センター2019 本試験 3A）

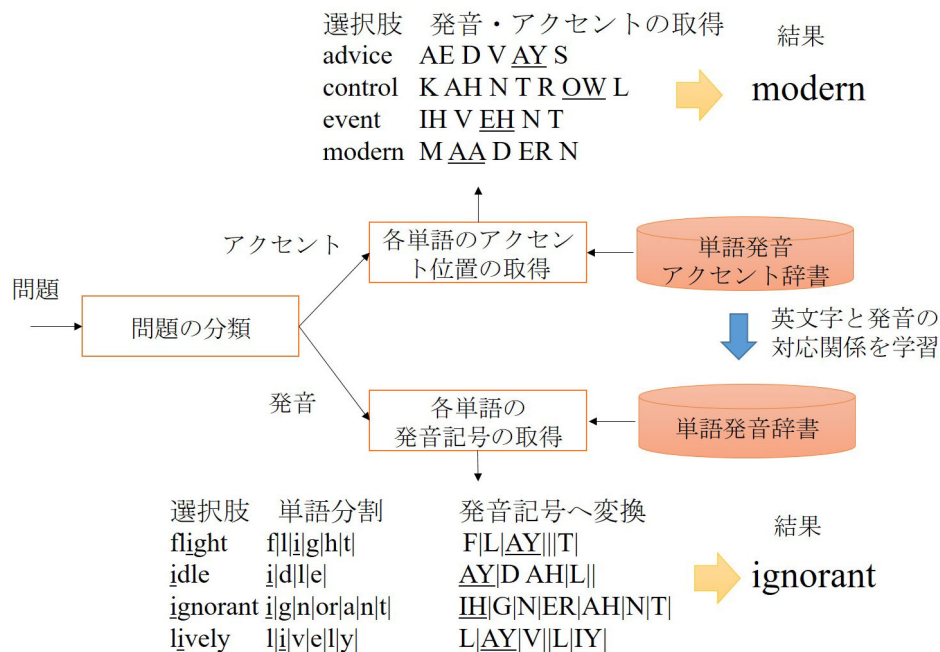


図3：アクセント・発音問題の解答手法（アクセント箇所や単語内の指定箇所の発音を調べる手法）

【今後の展開】

これまでの研究を通して様々な課題が明らかになりました。例えば、生活資料（チラシや広告）などの複数の情報からなる文書の理解や、グラフや表の読解、会話の流れの理解については、未だ安定した自動解答は実現できていません。今後はこのような、言語以外の情報や実世界の常識的知識が強く関わるタイプの問題に対応するため、関連する基礎研究の推進とその統合を進めていきます。

さらに、東ロボプロジェクトへの取り組みを通じて、文脈を理解し常識・専門知識の双方を備えた対話や質問応答を実現し、様々なサービスに展開していきます。

【用語解説】

※1 「corevo」 は日本電信電話株式会社の商標です。



※2 「ロボットは東大に入れるか」

NII の新井紀子教授を中心に、1980 年以降細分化された人工知能分野を再統合することで新たな地平を切り拓くことを目的に、若い人たちに夢を与えるプロジェクトとして 2011 年にスタートしたもの。具体的なベンチマークとして、2016 年までに大学入試センター試験で高得点をマークし、2021 年に東京大学入試を突破することを目標としている。

関連リンク

東ロボプロジェクトホームページ

<http://2lrobot.org/>

※3 NTT コミュニケーション科学基礎研究所

コミュニケーションの本質に応える情報通信の未来に向かって、人間科学と情報科学を融合した学際的アプローチにより新しい原理や概念を創出し、それらを革新的な情報通信サービスにつなげる基礎研究を行う研究所。

※4 XLNet、BERT

深層学習による文書表現技術の一種。極めて大規模なテキストデータを用いた単語の並び情報をもとに文書の的確な表現を獲得できる。対象とする課題のデータが少量であっても、これらの表現をもとに学習を行うことで、数多くの自然言語処理の課題で最高レベルの性能が達成されている。

※5 転移学習

あるデータで学習したモデルを他のデータでも利用可能とするような学習のこと。ドメイン適応やファインチューニングともいう。XLNet など大規模なテキストデータから得られた文書表現のモデルをセンター試験などの特定の問題を解くために利用することは転移学習の好例である。

＜本件に関する問い合わせ先＞

日本電信電話株式会社

先端技術総合研究所 広報担当

science_coretech-pr-m1@hco.ntt.co.jp

TEL046-240-5157

大学共同利用機関法人 情報・システム研究機構 国立情報学研究所

総務部企画課 広報チーム

media@nii.ac.jp

TEL03-4212-2164