

研究終了報告書

「新しい学びの形態を実現するための問題自動解説技術の開発」

研究期間： 2017 年 10 月～2021 年 3 月

(新型コロナウイルス感染症の影響を受け 2021 年 9 月まで延長)

研究者： 永田 亮

1. 研究のねらい

教育の場面において、学習者は、演習問題、試験問題など、「何らかの問題を解く」という活動を常に行っている。一方で、答案の添削には多くの時間と労力を要するため、学習者に詳細なフィードバックが与えられないことが多い。単に、問題を解かせるだけということもある。より高い学習効果をもたらすためには、個々の答案に合わせて、なぜ間違っているのか、どのように考えたら良いのかを解説することが重要である。

そこで本研究では、問題とその答案に対する解説(以下、解説文と表記する)を自動生成する手法の開発を行う。特に、英語ライティング学習に焦点を絞り、正しい英文を書くためのアドバイスやヒントを与えることで、学習者自身に考えさせる自律的な学びの形態を作り出すことを大きな目的とする。図 1 に、解説文の生成例を示す。図 1 では、正しい表現“has been raining”を直接提示するのではなく、時制に問題があることを指摘している。また、関連する情報(例:文法書の完了形、継続用法の項)へ自動的にリンクを貼り、学習者にこの問題を解くために必要な知識の習得を促す。これらの機能が、単に答えのみを示す従来の e-learning システムと大きく異なる。

問題: 次の文を英訳しなさい。

昨夜から雨が降り続けている。

学習者の答案:
It **is raining** since last night.

教科書, 参考書, 辞書
などの外部知識

「完了形」, 「継続用法」へのリンク

時制に注意して書き直してみよう。また、文法書の「完了形」, 「継続用法」をチェックしてみよう。

図1: 本研究で目指す解説文の例。赤い部分が生成される情報である。学習者答案に対して、なにが悪いのかを自然言語で解説する。また、関連する教材へのリンクも自動的に作成する。これらの情報により、学習者は自ら考え、問題を解きなおす。

2. 研究成果

(1) 概要

本研究では、

《課題 1》解説文生成研究のための基盤データの整備

《課題 2》解説する箇所を同定し、解説文を自動生成する手法の開発

《課題 3》教科書や辞書などの外部知識との連携による解説文の精緻化

の三つの課題に取り組んだ。《課題 1》では、世界で初めて、解説文生成研究のためのデータセットを構築し、一般公開した。今後、このデータセットは、解説文生成研究のための標準デー

タセットになると期待できる。《課題 2》と《課題 3》では、検索に基づく手法、検索-編集手法、クラスタリングを利用する手法など複数を開発し、その性能評価を行った。更に、解説文データの分析や自動評価手法の開発など、解説文生成研究で必要となる一連の技術の開発に取り組んだ。以上の研究成果により、データ、手法、評価という研究基盤が作成され、今後の解説文生成研究の大きな方向性を示したといえる。産業界からの注目度も高く、実用化も期待される。

新型コロナウイルス感染症の影響を受け六ヶ月間研究期間を延長し、前年度実施が難しかった国際学会におけるシェアードタスク(コンテスト)の開催および学習効果の測定について研究を進めた。シェアードタスクについては、データの準備、開催ルールの策定、評価用システムの策定など開催に必要な実際的な準備を全て終えることができた。一方で、新型コロナウイルスの影響もあり、開催の場となる国際学会の選定は難航した。自然言語処理分野のトップであるACL系の国際学会では、開催を打診したものの採択に至らなかった。最終的に、International Conference on Natural Language Generationという言語生成を専門とする国際学会にプロポーザルを提出した。審査の結果、来年度開催のシェアードタスクとして採択された。したがって、さきがけの研究期間外とはなるが、その国際学会のスケジュールに合わせて、シェアードタスクの実施、結果の発表会をおこなう。更に、さきがけ研究期間内に研究した解説文生成手法を整理し、各手法のメリットデメリットをまとめた論文を発表した。学習効果の測定についてもデータ収集システムなど必要な準備を全て終えることができた。一方で、学習者を集めた収集実験は依然難しい部分があり実施できていない。現在、企業での解説文生成サービスが始まっており、そこでの評価を検討している。

(2) 詳細

上述の目的を達成するために、本研究では、具体的に次の三つの課題に取り組んだ：

《課題 1》解説文生成研究のための基盤データの整備

《課題 2》解説する箇所を同定し、解説文を自動生成する手法の開発

《課題 3》教科書や辞書などの外部知識との連携による解説文の精緻化

それぞれの成果は次のとおりである。

《課題 1》解説文生成研究のための基盤データの整備

解説文生成の従来研究が非常に少ない理由に、利用可能なデータ(解説文が付与された英文)が存在しないということがあった。そもそも、解説として何をどのようにして付与したらよいかの知見もほとんどなかった。

そこで、本研究では、研究者(永田)と専門家(英語教材作成や英語教育の経験がある専門家)で議論を重ね解説文付与のためのガイドライン作りを行った。また、そのガイドラインに基づき、実際に解説文を付与するというパイロットスタディを行った。この過程を数度経て、最終的なガイドラインを作成した。更に、その結果を用いて、解説文生成研究のためのデータセット

を構築した。

構築の際の工夫として、文法用語や学習上重要となる用語を特殊なタグで表すようにした。例えば、〈間接目的語〉や〈現在完了形〉のように教科書や文法書の見出しとなる語句をタグで表す。このような特殊なタグを含む解説文を訓練データとすることで、解説文生成の結果にも同様なタグが含まれることになる。タグを含む語句は、教科書や文法書への見出しであるので、自動的に教科書や文法書へのリンク付けが行われることになる。

構築したデータセットの一部を言語資源協会より公開した (<https://www.gsk.or.jp/catalog/gsk2019-b>; <https://www.gsk.or.jp/catalog/gsk2019-a>)。解説文が付与されたデータの構築及び公開は世界初であり、今後同テーマの研究に活用されると期待できる(公開から1年半ほど経過するが、現時点で前者は8、後者は30の個人/団体に配布している)。なお、この成果により、2019年言語処理学会言語資源賞を受賞した。また、解説文データセットの構築に関する一連の成果は、体系的な論文発表2にて報告した。

《課題2》解説する箇所を同定し、解説文を自動生成する手法の開発

《課題3》教科書や辞書などの外部知識との連携による解説文の精緻化

上で述べたように解説文データの構築方法を工夫することにより、《課題2》と《課題3》を同時に解くこととした。すなわち、《課題2》の解説箇所の同定と解説文生成の性能が向上すると正確なタグ情報も生成されることになり、《課題3》の外部知識への連携も精緻化される。

解説文生成手法の考案に当たり、まず《課題1》で得られた解説文データセットを分析したところ、英文及びそれに付随した解説文には表層的なバリエーションが豊富にあるが、実質的な内容としてはかなりの重複があることが分かった。すなわち、表層的な見た目(単語列)としては異なるが、同種の誤りと同一内容の解説文が多数存在するということが明らかになった。

そこで、あいまい検索を応用した解説文生成手法の考案に取り組んだ(概要を下図に示す)。この手法では、解説文の表層的なバリエーションを吸収するため、深層学習ベースの言語モデルを用いて解説文を解説文ベクトルに変換する(図2のI)。これにより、内容が類似した解説文は、値が似た解説文ベクトルとなる。同じような枠組みで、解説対象の英文もベクトル化し、解説箇所の推定する(図2のII)。更に、推定された解説箇所に対応する文脈ベクトルを用いて、解説文ベクトルを検索するというアーキテクチャとした(図2のIIIとIV)。これにより解説箇所推定と類似解説文ベクトルの検索を同時に解く手法となる。

II. 解説箇所推定

III. 変換

IV. 解説文生成(検索)

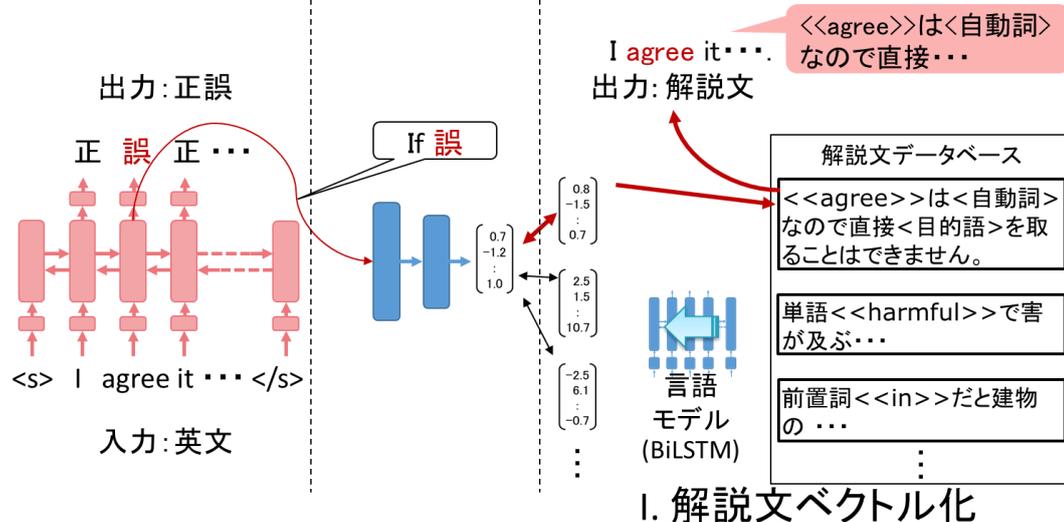


図 2 : あいまい検索を応用した解説文生成

この手法を前置詞誤りの解説に適用したところ、解説文のカバー率(再現率)20~30%、解説文の正確性(適合率)50~70%という性能が得られた。この性能は、従来のアプローチ(規則に基づく手法)と同程度の適合率を達成しつつ、再現率を2~5倍程度改善したことになる。依然、性能の改善の余地は大きい。従来、アプローチさえ知られていなかった問題(誤りの種類)に対しても解説文が性能可能となった意義は大きい。以上の成果は、代表的な論文発表1にて発表した。

上述の検索に基づいた手法は、一定の性能を示したものの、訓練データ(解説文データセット)に存在しない解説文を生成することはできないという制限がある。より柔軟な解説文生成を実現するためには、訓練データにない解説文を生成できる必要がある。

そこで、検索に基づいた手法を拡張し、より柔軟な解説文生成手法を実現することに取り組んだ。基本的なアプローチとしては、検索に基づいた手法で類似する解説文を訓練データより取得し、その結果をエディタと呼ばれる深層学習ベースの文生成器で編集するという方法を取った。この手法は、検索と編集からなるため検索-編集手法と呼ぶ。

検索-編集手法の性能を評価したところ、理論的な予想に反して、基本手法(検索に基づいた手法)より性能が悪くなることが明らかになった。生成結果を分析したところ、検索された解説文と目的の解説文が同じような内容であっても、表層的なバリエーションが大きい。そのため、不必要な編集(単語の追加および削除)をエディタが学習しているということが明らかになった。この問題を低減するため、編集に制限を加える機構を考案した。その結果、検索に基づいた手法より高い性能を達成することに成功した。これ以外にも、いくつかの改善点を加え、性能

の向上, 各手法の特徴を明らかにした。この成果は, 現在, 国際学会 European Chapter of the Association for Computational Linguistics に投稿中である。

以上の生成手法の考案は, 前置詞誤り解説文を対象に行ったが, その成果を一般の解説文に対しても拡張した。改良した検索編集手法の性能は, 再現率 20%前後, 適合率 20%前後と前置詞を対象にした場合よりかなり低くなった。一般解説文の内容は多岐にわたるため, この性能低下は予想されたものであるが, 生成結果を分析すると同手法が有効である事例とそうでない事例に分かれることが明らかとなった。また, 解説箇所推定の性能がボトルネックになるという傾向が顕著にみられた(同じ問題は前置詞解説文でも起こる)。

この解決策として, 解説文をあらかじめクラスタリングし, 解説箇所推定と解説文生成, 両方の性能を向上させるということに取り組んでいる。予備調査の段階で, 解説文の表層情報を用いてクラスタリングを行うと, 同一種類の誤り(すなわち同じような解説文が付与される誤り)を自動的にグループ化でき, 解説箇所推定の性能が向上するとい良好な結果が得られている。残りの時間を使用して, クラスタリングによる解説箇所推定と解説文生成の性能を向上させる予定である。

更なる発展として, 実用を見据えた応用研究にも取り組んだ。解説文生成の失敗には, 重大なもの(学習者が誤った知識を獲得する可能性のある誤生成)とそうでないもの(学習者自ら誤生成と判断できるもの)がある。特に, 生成性能が向上すると表層的には適切に見える(言い換えれば, 文としては自然な)誤生成が増える。重大な誤生成は極力減らすべきである。そこで, 生成結果の信頼度を推定し, 重大な誤生成を取り除くことに取り組んだ。小規模であるが実際に学習者に解説文生成システムを使ってもらったユーザ実験も実施している。

別の応用研究として, 生成結果の自動評価にも取り組んだ。現状では, 生成結果の評価は, 専門家が人手で行っている。その信頼性は高い一方で, 時間とコストがかかるため, 多種多様な手法の比較を行うことは容易でない。そこで, 解説文生成の自動評価手法も考案し, その評価を行った。完全な自動評価の実現には至らなかったが, 人手評価の補助として使えるまでには性能が向上している。

以上の通り, 本研究の成果により, 解説文生成研究に必要となる基盤(データ, 手法, 評価)が出来上がった。この基盤により, 今後の解説文生成研究の大きな方向性を示したといえる。また, 産業界からの問い合わせも多く, 実用化も期待される。

3. 今後の展開

今回構築公開した解説文データセットは, 今後, 解説文生成研究の研究基盤となると期待できる。実際, 公開後 1 年ほどであるが, 既に 10 以上の団体, 個人から利用申請の実績がある。

また, 教育関係の企業からの問い合わせが多く, 当該技術への注目の高さがうかがえる。

教育的倫理(例えば, 誤った解説を学習者に提示しにくい)など周縁的な条件での難しさもあるが, 数年以内に何らかの形で実用化される技術があると予想している. 特に, 人が行う添削を支援するシステムは, 間に人手が入ることにより精度保証できるため, 実用化に向いているといえる. 実用化にむけて, 企業との連携を模索していく予定である.

4. 自己評価

概ね当初の計画通り研究活動を遂行した. 上述の通り, 従来知見の少なかった解説文生成の研究において, データセット, 手法, 評価手法で基盤となる知見を示すことができたのは一定の成果であると自己評価している. 基盤データについては, 著作権および新型ウィルスの影響で当初予定していたものの一部が構築できなかった. 一方で, さきがけの講演がきっかけになり, 新たな研究連携が始まり, 新たなデータの入手, 構築ができた. 全体としては, 予定以上の成果が得られたと評価している. 解説文生成手法についても, 予定していた手法の開発および評価は実施できた. 従来手法では, アプローチ方法さえ知られていなかった事例に対しても解説文が生成できることを示した意義は大きいと自己評価している. 以上の成果により, 新しい学びの形態が近い将来実現すると予想している. 特に, 開発した技術を用いた添削支援システムは, 十分に実用に耐えうるという感触を持っている. また, 使い方によっては, 学習支援システムとしても実用化可能であると予想している.

5. 主な研究成果リスト

(1) 代表的な論文(原著論文)発表

研究期間累積件数: 4件(うち2件はコロナ延長時の成果)

1. Ryo Nagata. Toward a Task of Feedback Comment Generation for Writing Learning. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing. 2019, pp.3206-3215.
解説文生成の重要性を問題提起し, 解説文生成タスクの定式化を行った. 更に, 規則に基づいた方法, 単純な sequence-to-sequence を利用した手法, 検索ベースの手法の3種類の手法を比較し, 検索ベースの手法が解説文生成タスクに有効であることを示した. 解説文生成を自然言語処理の新たに提案し, 研究の方向性を示したマイルストーン的な研究と位置付けられる.
2. Ryo Nagata, Kentaro Inui, Shin'ichiro Ishikawa. Creating Corpora for Research in Feedback Comment Generation. Proceedings of International Conference on Language Resources and Evaluation. 2020, pp340-345.
解説文生成研究のためのデータセットを構築, 公開した. 解説文として何をどのようにアノテーションすべきかをデザインした. また, その結果に基づき実際に学習者が書いた英文に解説文を付与した. 解説文付き学習者英文の公開は世界初であり, 今後の同分野の発展に大きく寄与すると期待できる.
3. Ryo Nagata, Masato Hagiwara, Kazuaki Hanawa, and Masato Mita. A Proposal of Feedback Comment Generation for Language Learners. Proceedings of 14th International Conference on Natural Language Generation. 2021, (to appear).

解説文生成を題材にした shared task を世界で初めて提案した。言語生成専門会議である International Conference on Natural Language Generation で同 shared task を行うことで、幅広い知見が得られるとともに、解説文生成タスクの普及にもつながると期待できる(コロナ延長時の成果)。

4. Kazuaki Hanawa, and Ryo Nagata. Exploring Methods for Generating Feedback Comments for Writing Learning, Proceedings of 2021 Conference on Empirical Methods in Natural Language Processing, 2021, (to appear).

(2) 特許出願

研究期間累積件数: 1 件(特許公開前のものも含む)

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

・永田亮, 英語ライティング学習支援のための前置詞誤り解説生成, 言語処理学会第 25 回年次大会発表論文集, 2019(言語処理学会第 25 回年次大会優秀賞受賞)

・永田亮, 石川慎一郎, 乾健太郎, 解説文生成研究のためのライティング技術解説付き学習者コーパス, 言語処理学会第 25 回年次大会発表論文集, 2019(言語処理学会 2019 年言語資源賞受賞)

・解説文データセットの公開

GSK2019-A Konan-JIEM Learner Corpus Sixth Edition

(<https://www.gsk.or.jp/catalog/gsk2019-a/>)

GSK2019-B ICNALE Learner Essays with Feedback Comments

(<https://www.gsk.or.jp/catalog/gsk2019-b/>)