

情報科学論文の情報量と文章量を維持した自動平易化の試み

大西耕介¹ 菊池英明¹ 藤倉将平²

¹ 早稲田大学人間科学部 ² 株式会社サイシキ (SAI-SHIKI Inc.)

024@asagi.waseda.jp kikuchi@waseda.jp fujikura@sai-shiki.com

概要

専門的な文書の平易化は医療分野を対象としたものが多く、情報科学の分野を対象にしたものは少ない。また、既存の平易化研究では専門用語の意味を大きく削って平易化を行ったり、専門用語を非常に長い文章で平易化したりしている。これは論文の情報を全て簡潔に理解するためには適切とは言えない。本研究では情報科学分野の論文を対象に、専門用語の平易な説明のポップアップを用いて平易化を行うシステムの開発を行った。目標として、内容の理解しやすさの改善、文章量の増加防止、情報量の保持を掲げた。内容の理解しやすさの改善は見られなかったが、先行研究手法と比べ情報量の保持に成功し、文章量の増加防止は達成した。

1 はじめに

テキスト平易化とは文の意味を保ちつつ、語彙平易化や構文平易化などの複数の編集操作を通じて平易に書き換えるタスクである [1]。この技術は原文を読むための十分な知識のない人の文章読解支援に役立つ。論文は専門用語を多く含んでいる。専門用語は、特定の専門分野や業界でのみ使用される特別な用語が多く、特定の概念を厳密に表現するために用いられる [2]。そのため、専門用語を理解するためには、該当専門分野の知識が必要であり、一般の人にとって専門用語はなじみの薄い存在となることが多い [2]。そのため、初学者は論文を読むことが難しい。専門用語を適切に平易化することで初学者でも論文の内容を知ることが出来るようになる。

専門的な文章の平易化研究は主に医療分野を対象に行われており [2][3]、情報科学の分野を対象にしたものは少ない。しかし、情報科学の分野においても価値があると考えられる。情報科学の分野は他の分野と比べて専門的な道具を必要とせず、初学者を幅広く受け入れている。そういった初学者が最先端の研究に触れられるようになれば、情報科学分野の研究

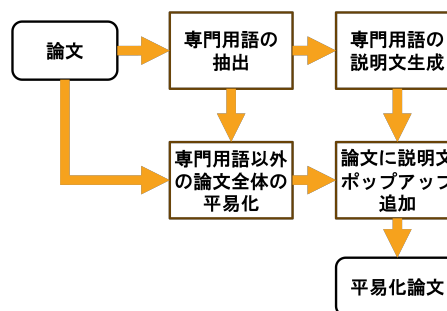


図1 システムフロー図

の発展に繋がる。

また、既存の研究の問題として専門用語の平易化を行う際に専門用語の持つ意味を大きく失っていたり [3]、専門用語の平易化が非常に長かったりするというものがある [2]。これらの問題は専門用語をそのまま残し、ポップアップによって専門用語の平易な説明を表示することで解消できるのではないかと考えた。

以上の背景より、本研究では情報科学論文の平易化システムの構築を行う。また、専門用語は本文中にそのまま残し、システムにより生成した平易な説明をポップアップによって表示する。論文の難しさに関しては、専門用語による難しさ以外にも言い回しの難しさや取り扱うものの自体の難しさ、論理展開の難しさなどが存在するが、本研究では言い回しの難しさと専門用語による難しさのみを取り扱う。

2 システム概要

本システムは React による UI と Python を用いた ChatGPT の API によって実装を行う。本研究では通常の平易化に加えて専門用語の説明文をポップアップとして表示するため、図 1 に示すフローで入力論文の平易化を実現する。

2.1 専門用語抽出モジュール

専門用語抽出モジュールでは論文からの専門用語の抽出を行う。本モジュールでは論文を入力とし、

GPT-4o を用いて python のリスト形式で抽出した専門用語を出力させる。また、その際に前後に不要なテキストが入る場合があるため、正規表現を用いて [] に囲われていない部分のテキストを取り除く。

2.2 専門用語説明文生成モジュール

専門用語説明文生成モジュールでは専門用語抽出モジュールによって抽出された専門用語の平易な説明を GPT-4o を用いて生成する。入力には論文の要約文と専門用語抽出モジュールにより抽出された専門用語とし、JSON 形式で専門用語とその平易な説明を出力させる。入力に論文の要約文が必要な理由は論文の文脈を与えないと文脈に沿った専門用語の説明を行えないためであり、この要約文も別途 GPT-4o を用いて生成する。専門用語の平易な説明を生成する際に内容が不十分になることがあったため「説明は必ず 5 文で行ってください。」という文言をプロンプトに追加した。また、専門用語が勝手に日本語から英語に、あるいは英語から日本語に翻訳されることがあったため、それを禁止するプロンプトを追加した。

2.3 全体平易化モジュール

全体平易化モジュールは専門用語抽出モジュールによって抽出された専門用語以外の論文全体を GPT-4o によって平易化するモジュールである。本モジュールでは 1 文ずつ論文の平易化を行う。また、本モジュールでも論文の要約文を入力することで、文脈に沿った平易化を行う。また、1 文ずつ平易化を行うことにより、予期せぬ情報の削除を防いでいる。文を分割する際、ピリオドや句点によって文を分割するが、文末以外のピリオドでも文を分割してしまう危険性があった。そのため、それらを回避するよう正規表現を用いて処理を行った。また、専門用語抽出モジュールで抽出された専門用語は特殊な記号で囲み、その内部を平易化しないようプロンプティングしている。

3 評価実験

本研究では内容の理解しやすさ、文章量、情報量に着目して研究を行なっているため、評価実験ではそれらの指標を評価する。また、提案手法の比較対象は元論文、先行研究手法 ([4] による Zero-shot 平易化)、提案手法からポップアップによる専門用語説明を除いたもの（以下、ポップアップ無し提案手法と言う）、専門用語も平易化した人手による平易化論文

（以下、ポップアップ無し人手平易化論文と言う）、専門用語は残して説明をポップアップによって表示する人手による平易化論文（以下、ポップアップ有り人手平易化論文と言う）とする。人手による平易化は大学院生に依頼する。

3.1 人手による情報科学論文平易化

3.1.1 依頼先

人手による情報科学論文の平易化は論文に対して一定の理解がある人が行う必要がある。そのため、平易化する論文の分野を研究している大学院生に平易化を依頼する。平易化を行う論文の数は 6 件であり、1 人が複数論文の平易化を行うことを認める。依頼の結果、4 名の大学院生が論文の平易化を行なった。

3.1.2 平易化対象

平易化を行う対象は大学院生が専門分野とする情報科学論文のアブストラクト部分である。大学院生に適した論文の平易化を行うため、平易化を行う論文は平易化を行う大学院生自身が選ぶ。論文を選ぶ際には専門用語が多く含まれ、初学者に読みにくいと思われる情報科学論文であることを条件とした。

3.1.3 平易化方法

平易化の際には専門用語をそのまま残すような平易化と専門用語をほとんど残さない平易化の 2 種類の平易化を行う。どちらの平易化方法でも難解な言い回しの平易化は行う。専門用語をそのまま残すような平易化では本文中の専門用語は書き換えず、別途専門用語の平易な説明を行う。専門用語をほとんど残さない平易化では専門用語を端的に平易化する。なお、評価実験においては専門用語をそのまま残すような平易化はポップアップ有り人手平易化論文に加工し、専門用語をほとんど残さない平易化はポップアップ無し人手平易化論文とする。また、BLEU による評価を行う関係上、平易化前の文と平易化後の文が 1 対 1 の対応になっている必要があるため、1 文ずつ平易化を行い、前の文を参照するような指示語は新たに追加しないこととする。また、依頼協力者の参考として著者による平易化例を与えた。

3.2 内容の理解しやすさの主観評価実験

内容の理解しやすさは人手による 5 件法によって評価を行う。本研究では情報科学論文をほとんど読

んだことがない初学者が情報科学論文の内容を理解出来るようにするのが目的であるため、被験者は情報科学論文にあまり触れていない人 23 名を対象とする。本実験はオンラインで行う。被験者はまず事前アンケートに回答する。その後、評価対象となる 6 つの手法から 1 手法につき 1 論文、それぞれ別の元論文の計 6 本の論文を読み、内容の理解しやすさを評価する。また、被験者は論文を読む際には論文を読むのに掛かった時間を計測する。また、順序効果を低減するために読む順番はランダムにする。

3.3 文章量の評価

文章量の評価は単純に文字数をカウントする。比較対象は元論文、提案手法による平易化論文、先行研究手法による平易化論文、ポップアップ無し人手平易化論文、ポップアップ有り人手平易化論文の 5 つである。ポップアップ無し提案手法による平易化論文は提案手法による平易化論文と同じであるため評価を行わない。

3.4 情報量の評価

情報量の評価は BLEU を用いて行う。BLEU はテキスト平易化に用いられる評価指標の中でも流暢性と意味保持性に相関があるとされている [5]。そのため、BLEU を用いて情報量の評価を行う。BLEU を用いて評価を行う際に用いる正解となる平易化論文はポップアップ無し人手平易化論文とする。評価の際、BLEU は 1 文対 1 文の比較で評価を行うため、ポップアップによる専門用語の説明の評価は難しい。そのため、ポップアップによる専門用語の説明は評価の対象としない。

4 実験結果と考察

4.1 人手による情報科学論文平易化

人手による平易化では計 4 名の大学院生が 6 つの論文の平易化を行なった。その結果、表 1 の平易化結果が得られた。なお、元字数は元論文の文字数、専無字数は専門用語をほとんど残さない平易化での文字数、専有字数は専門用語をそのまま残すような平易化での文字数、専門用語数は専門用語をそのまま残すような平易化で専門用語と判断された単語の数である。どの論文も専門用語を少なからず含んでおり、専門用語をそのまま残すような平易化では文字数の変化が少ないことから、専門用語を多く含み専門用

語をそのまま残すような平易化では大きな言い換えをしないという条件は達成できているものと思われる。

表 1 人手による平易化の結果 (元字数は元論文の文字数、専無字数は専門用語をほとんど残さない平易化での文字数、専有字数は専門用語をそのまま残すような平易化での文字数)

	元字数	専無字数	専有字数	専門用語数
論文 A	454	750	502	15
論文 B	298	412	337	6
論文 C	296	340	290	6
論文 D	404	559	436	9
論文 E	370	518	376	10
論文 F	356	509	365	9

4.2 内容の理解しやすさの主観評価実験

内容の理解しやすさの主観評価実験では合計 23 名の回答より、表 2、表 3 の実験結果が得られた。

また、等分散性を明らかにするために各種手法に対してルビーン検定を行った。その結果、各群の分散に差があるとは言えなかった ($W=0.63$, $p=0.68$)。その後、評価値の分布の正規性を明らかにするためにそれぞれの平易化手法に対してシャピロ・ウィルク検定を行なった結果、元論文を含む全ての平易化手法で正規性がないと判断された。正規性が認められなかったため、クラスカル・ウォリス検定を行って各群の評価に差がないかを調べた。その結果、各群間の差は認められなかった ($H=7.72$, $p=0.17$)。そのため、提案手法によって論文を理解しやすく出来たとは認められなかった。

今回、有意差が現れなかった原因として考えられるのは元論文が平易化を行わずとも理解しやすいものだったということである。本研究は難解な論文を平易化することによって理解しやすくするということを目的としていた。しかし、元論文の理解しやすさの評価は平均が 3.6 と高かった。そのため、各種手法を用いて平易化を行っても大きな差が現れなかったのではないかと考えられる。

また、有意差が現れなかったとして考えられる他の要因は理解しやすいと感じることと実際に理解しやすいことが違うのではないかとということである。元論文とポップアップ無し提案手法の読了時間があまり変わらないのに対して、提案手法は読了時間が伸びていることから、提案手法ではポップアップによる専門用語説明を読んでいると思われる。その場

合、ポップアップの説明を読んだ分だけ理解できなかった専門用語を理解することが可能であり、より論文の内容が理解できるはずである。しかし、それが理解しやすさの評価に繋がっていないのならば理解しやすいと感じることと実際に理解しやすいことに違いがある可能性がある。

表 2 内容の理解しやすさの評価結果

平易化手法	平均	標準偏差
元論文	3.6	1.4
提案手法	3.7	1.5
PU 無し提案手法	3.2	1.6
先行研究手法	4.1	1.1
PU 無し人手平易化論文	3.3	1.1
PU 有り人手平易化論文	3.3	1.5

表 3 内容の理解しやすさ実験における読了時間 (分)

平易化手法	平均	標準偏差
元論文	1.71	1.79
提案手法	2.20	2.23
PU 無し提案手法	1.66	1.69
先行研究手法	1.29	1.66
PU 無し人手平易化論文	1.85	1.13
PU 有り人手平易化論文	2.30	2.19

4.3 文章量の評価

文章量の比較では表 4 の結果が得られた。文章量に関しては提案手法は文字数を大きく増やすことはなく、本研究の目的を達成できたと言える。一方、先行研究手法は文字数を元論文より減らすことに成功していた。また、ポップアップ無し人手平易化論文は元論文と比べて文字数を大きく伸ばしており、ポップアップ無しに情報量を失わせない平易化を行う際は文字数が増えてしまうことが読み取れる。

表 4 文章量 (文字数) の比較結果

平易化手法	平均	標準偏差
元論文	363	61
提案手法	372	60
先行研究手法	335	84
PU 無し人手平易化論文	515	140
PU 有り人手平易化論文	384	75

4.4 情報量の評価

情報量の評価では表 5 の結果が得られた。BLEU を用いた情報量の評価では元論文が最も高い評価を

得ており、先行研究手法が最も低い評価で、提案手法はその中間であった。このことから、情報量の保存に成功したとは言えない。しかし、内容が理解しやすく、文章量を短くすることにも成功していた先行研究手法は情報量を大きく失っており、情報量の保持という点においては提案手法の方が本研究の目的に合致していると言える。提案手法は言い換えが少なく、先行研究手法は言い換えが多かったことを考えると、言い換えた分だけ人手による平易化から外れてゆき BLEU スコアを減らしたのではないかと推測される。言い換えが評価を下げた原因として、BLEU で評価を行う際に与えた正解とする人手平易化論文が 1 種類しかなかったことが考えられる。平易化を行う際には様々な言い換えが考えられ、適切な言い換えは 1 つだけとは限らない。しかし、正解を一つしか用意しなかったためそれ以外の平易化は不適切だとされてしまった。

表 5 情報量の評価結果

平易化手法	BLEU
元論文	46.86
提案手法	25.57
先行研究手法	9.06

5 おわりに

今後の課題として見つかったのは主に以下の 3 点である。

1. 実験の対象となる論文の適切な選定
2. 理解しやすさが適切に評価できるような評価方法の改善
3. BLEU による評価を行う際の正解とする人手による平易化論文の不足

今後の展望としてはまず評価実験を完遂することが必要である。評価実験が計画時の 30 名より少ない 23 名で行われているため、実験を継続する。その後、今後の課題としてあげた項目について適切な実験状況を整える。そのため、まずは実験の対象とする論文の再選定と平易化を行い、再度評価実験を行う。また、理解しやすさの適切な評価のため、主観評価ではないより客観性の高い評価方法を定める。その後、人手による平易化論文を複数に増やした上で BLEU による情報量の保持に関する評価を行う。その後、まだシステムの平易化に問題がある場合はシステムの改善を行っていく。

参考文献

- [1] Fernando Alva-Manchego, Carolina Scarton, and Lucia Specia. The (un)suitability of automatic evaluation metrics for text simplification. **Computational Linguistics**, Vol. 47, No. 4, pp. 861–889, December 2021.
- [2] 黄海洪, 金丸敏幸. 介護分野における専門用語の平易化に向けた語彙リストの構築. 言語資源ワークショップ発表論文集, Vol. 1, pp. 26–39, 2023.
- [3] 堀口航輝, 梶原智之, 二宮崇, 若宮翔子, 荒牧英治. 日本語医療テキスト平易化の訓練用データセットの構築. 人工知能学会全国大会論文集, Vol. JSAI2024, pp. 3S1OS7b04–3S1OS7b04, 2024.
- [4] 樽本空宙, 畠垣光希, 宮田莉奈, 梶原智之, 二宮崇. Chatgpt の日本語生成能力の評価. 自然言語処理, Vol. 31, No. 2, pp. 349–373, 2024.
- [5] Tu Vu, Baotian Hu, Tsendsuren Munkhdalai, and Hong Yu. Sentence simplification with memory-augmented neural networks. In Marilyn Walker, Heng Ji, and Amanda Stent, editors, **Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)**, pp. 79–85, New Orleans, Louisiana, June 2018. Association for Computational Linguistics.