

2.3. 単語 D を取得

名詞 D は、名詞 C の同音異義語とする。同音異義語とは発音が同じだが、意味の異なる単語のことをいう。本システムでは概念ベースの中の acb008 を使用するため、acb008 の中から同音異義語を検索する。同音異義語を判断する条件は 2 つある。1 つ目は、同音の単語を検索する。同音を検索するにあたって本システムでは MeCab を使用し、MeCab から取得した読みが同じだった場合同音と判断する。2 つ目は、関連度計算を行い、意味の異なる単語か判断する。関連度計算を行い 0 以上 1 未満の値が得られた場合、異なる意味をもつ単語として判断する。

この 2 つの条件に当てはまった場合、本システムでは同音異義語として判断される。名詞 C の同音異義語である名詞 D が取得できなかった場合、なぞかけ生成不可能とする。

2.4. 名詞 B を取得

名詞 B は、名詞 D の属性の中で 1 番重みの強い属性を用いる。名詞 D の属性が取得出来なかった場合、なぞかけ生成不可能とする。

2.5. 名詞 C と名詞 D に関する形容詞が一致する

名詞 C と名詞 D に関する形容詞を感覚想起システムから取得する。取得した形容詞が一致する場合、なぞかけの答えの部分成り立つと判断する。取得した形容詞が一致しなかった場合、なぞかけ生成不可能とする。

2.6. なぞかけ生成

取得した単語を用いて「単語 A とかけてまして単語 B ととく、そのところはどちらも(単語 C/単語 D)が形容詞です。」と出力する。例として、ユーザーが「薬」と入力した場合、「薬とかけてまして戦意ととく、そのところはどちらも(医師/意思)が心強いです。」と出力される。この場合、単語 A が「薬」、単語 B が「戦意」、単語 C が「医師」、単語 D が「意思」、形容詞が「心強い」にあたる。

3. 評価と考察

3.1. 評価方法

今回の研究では、2 つの評価方法を用意する。1 つ目は、使用した際のシステムの「実用性」、「満足度」、「使い勝手」を(すごく悪い、悪い、普通、良い、すごく良い)の 5 段階で評価してもらう。2 つ目は、被験者がシステムに入力し、生成されたなぞかけの面白さを(すごく悪い、悪い、普通、良い、すごく良い)の 5 段階で評価してもらう。

3.2. 評価結果

被験者 6 人が 5 回ずつ、計 30 回システムを使用した。そのうち、なぞかけが出力された回数は 9 回だった。よって出力率は 30.0% となった。表 1 に提案システムのアンケート結果を示す。

表 1 提案システムのアンケート結果

| | すごく悪い | 悪い | 普通 | 良い | すごく良い |
|------|-------|-------|-------|-------|-------|
| 使い勝手 | 0 | 0% | 50% | 33.3% | 16.7% |
| 満足度 | 16.7% | 33.3% | 0% | 16.7% | 33.3% |
| 実用性 | 16.7% | 16.7% | 33.3% | 16.7% | 16.7% |

3.2. 考察

今回、概念ベースを用いることで既存システムの問題になっていた、ユーザーへの負担を軽減してなぞかけを

生成することができるようになった。その結果として、使い勝手は、被験者の 50% が「すごく良い」または「良い」と評価した。しかし、満足度は、被験者の 50% が「すごく悪い」または「悪い」と評価し、実用性には被験者の 33.3% が「すごく悪い」または「悪い」と評価した。満足度と実用性の評価の低さの原因に出力率の低さがあると思われる。被験者の中には未登録語や生成できない単語を入力したため、5 回の入力のうち 1 回しかなぞかけが出力されない人がいた。ここで未登録語とは、今回使用した acb008 に収録されていない単語とする。本システムでは、全ての工程で acb008 を用いたため、acb008 に収録されていない単語ではなぞかけを生成することができない。この出力率の低さの原因に、acb008 の概念数の少なさがあると思われる。acb008 の概念数は 87,242 となっていて、地名や人名、日常的に耳にする話題の単語に対応することができず、出力する割合が低くなってしまった。

それらを踏まえて出力率を向上させるために概念ベースの入れ替え、または未登録語への処理を行う必要がある。概念ベースの入れ替えは acb008 から cd2018 に入れ替える必要がある。cd2018 は Wikipedia を知識源とした概念ベースと acb008 を統合したものであり、概念数が 1,068,719 ある。本システムを cd2018 に入れ替えることにより出力率の増加、取得できる属性数も増加すると考えられる。また、直近で話題になっている単語の入力が多かったことから、未登録語に関する情報を取得するシステムを実現するべきだと考える。

4. おわりに

今回、なぞかけを生成するための手間を減らし、ユーザーが気軽になぞかけを楽しめることを目的として、概念ベースを用いたなぞかけ自動生成システムを構築した。概念ベースを使うことで、感覚想起システムから自動的に形容詞を取得することが可能になったため、なぞかけを生成するユーザーの負担を減らすことができた。これにより、さらに気軽になぞかけを楽しめるようになった。しかし、未登録語への処理ができないなどの理由から、出力率が低くなってしまった。

使用するデータベースを入れ替えるなど、未登録語への対応を可能にし、出力率を向上させることが今後の目標である。

謝辞

本研究の一部は、JSPS 科研費 16K00311 の助成を受けて行ったものです。

参考文献

- [1] “なぞかけとは？ビジネスとなぞかけの意外な関係”, TOPSIC, 2021 年 6 月 11 日, 2021 年 8 月 16 日 <https://products.sint.co.jp/topsic/blog/nazokake#toc-7>
- [2] 奥村紀之, 土屋誠司, 渡部広一, 河岡司, “概念間の関連度計算のための大規模概念ベースの構築”, 自然言語処理, Vol. 14, No. 5, pp. 41–64, 2007.
- [3] 井岡大志, 渡部広一, 河岡司, “概念ベースを用いた連想機能実現のための関連度計算方式”, 情報科学技術フォーラム FIT2002, pp. 159–160, 2002.
- [4] 近藤真生, “常識的知覚感覚判断システムの精度向上”, 卒業論文, 2020.
- [5] MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://taku910.github.io/Mecab/>, (2021/01/25 参照)
- [6] 内村圭佑, 灘本明代, “Web コンテンツを用いたなぞかけ自動生成の提案と評価”, 情報処理学会研究報告, Vol.2009-DBS-148 No.25