

単語想起支援システムにおける インターネットを利用したデータ拡充手法

馬場 柁也[†]
† 立命館大学情報理工学部

和田 侑大^{††} 桑原 和宏[†]
†† 立命館大学大学院情報理工学研究科

1. はじめに

失語とは、様々な影響で脳に損傷を負い、一度獲得した、聞く、書く等の言語能力に障害のある状態を指す。このような人を支援する単語想起支援システムを開発している[1]。このシステムでは、複数の質問文を用意し、質問に対し回答していくことで想起したい単語を導く。単語想起支援システムの有用性を向上させるために[2]では対話的に単語データベースを拡充する手法を提案したが、専門用語等のデータを単語想起支援システムに追加するためには専門家が必要となる。これに対し本稿ではインターネット上のデータを利用したデータの拡充手法を提案する。

2. 提案手法

本稿では、地域特有の内容を含む「郷土料理」を例題として取り上げ、インターネットを利用し、図1に示すように追加したい料理と、食材に関するオントロジーを構築することで、データを拡充する。

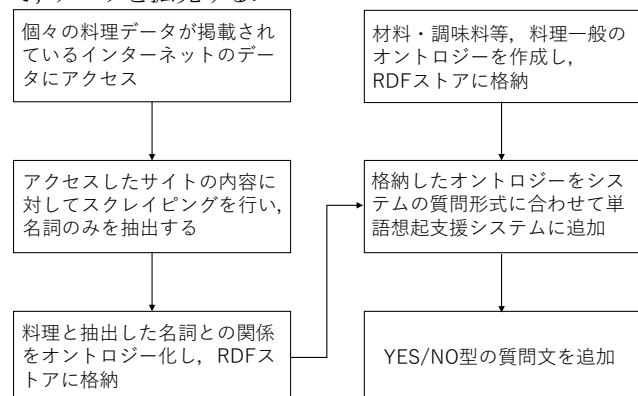


図1. 提案手法

ここでは、個々の料理データが掲載されているサイトとして、Wikipediaの郷土料理一覧のページと、そのリンク先のページのデータを利用する。また、食材、調味料等の料理一般に関するオントロジーとして、まずは手動で構築し、その後インターネット上にある郷土料理レシピのWebサイトに掲載されている112種類の料理にアクセスし、料理一般に関するドメインオントロジーを用意した。さらに、各食材、調味料の使用の有無を問う質問を追加することで、単語想起支援システムの有用性を向上させる。

3. 実験方法

次の3つのパターンについて単語想起支援システム上で検証する。(1)郷土料理のデータに関するオントロジーに対して、手動で構築したドメインオントロジー、(2)インターネットを利用したドメインオントロジー、(3)単語想起支援

システムに食材の使用に関する質問を追加したもの。

単語想起支援システムでは想起する単語に対して、質問文とその回答を行う会話パートナーとシステムの対話を模擬することが可能であり、料理を絞り込むまでの「平均質問数」、「一問当たりの平均選択肢数」、および「平均絞り込み単語数」の3つの観点から検証する。

4. 実験結果

Wikipediaの郷土料理ページに掲載されている料理名と郷土料理レシピのWebサイトに掲載されている料理名のうち、両方のサイトの料理名表記で整合性が取れた47種類の料理に対するシミュレーション結果を図2に示す。

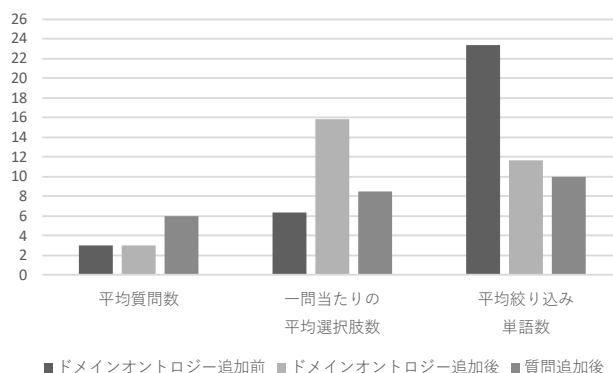


図2. シミュレーション結果

図2からわかるように料理一般のドメインオントロジーに関して、インターネットを利用することで質問一問当たりの平均選択肢数は増加するが、単語の絞り込みは改善された。また、食材、調味料の使用の有無について問う質問を追加することで、さらに単語の絞り込みが可能となった。

5. まとめ

Webサイトのデータを活用して単語想起支援システムのデータを拡充する手法を提案し、シミュレーション実験によりその効果を検証した。インターネット上の2種類のサイトを利用し単語データを拡充することで単語想起支援システムの有用性を高めることが期待できる。今後は、別のWebサイトやSNS上のデータの活用や取得したデータの正確性を高める手法を検討する。

参考文献

- [1] 岩前ほか, 失語のある方に対する単語想起支援手法の提案, 第60回システム制御情報学会研究発表講演会, 2016.
- [2] 和田ほか, 失語のある人を対象にした単語想起支援システムにおけるデータベースの拡充, 電子情報通信学会2017年総合大会「学生ポスターセッション」ISS-P-142, 2017.