

NAISTにて取り組む研究について

氏名: 新妻巧朗
試験区分: 情報科学区分
希望研究室: 自然言語処理研究室

1 はじめに

私はNAISTにて、「情報検索システムにおけるファセットの自動生成」について研究をしたいと考えている。

2 研究の背景

2.1 研究の社会的意義

現代社会において情報収集をおこなうためには、検索エンジンを利用することは必要不可欠である。しかし、検索エンジンを適切に活用できずに、目的の情報に至れない人も存在している。それは現代の検索エンジンの仕組みが、利用者に対して情報検索能力を要求しているからだ。この能力は福島らの研究によって個人差が大きく能力差によって情報格差が生じていることがわかっている [1]。こうした課題を解決することによって、情報に辿りつけなくて引き起こされている機会損失を減らすことができる。齋藤らによる教育を通して情報検索能力を向上させる研究 [2] も存在しているが、本研究ではシステムによって解決するアプローチを考えていく。

2.2 研究の背景

言語能力の高さが情報検索能力に高さに関係しているとされている [1]。つまり、言語能力の高低が情報検索における情報格差を生み出していると考えられる。そのため、情報の探索過程にて言語能力を要求する場面で、その能力を補助をするシステムを提案することで解決をしたいと考えている。

2.3 問題の解決法の提案

そこで、利用者が入力した検索キーワードに対して適切なファセットを提供し、インタラクティブに検索意図を読み取るシステムを提案したい。情報の探索行動を掘り下げると、検索意図と検索対象の意味的な距離を縮めていくプロセスと考えられる。そのため、私はファセット検索という手法が役に立つのではないかと考えたからだ。ファセット検索とは、システムの利用者に検索対象を何らかの側面で絞り込むファセットを提示し、検索対象を絞り込む検索手法である。つまり、システムの利用者が、検索意図を掘り下げて言語化をする行動を代行しているとも言える。そのため、検索エンジンが個人の言語能力に依存している課題の緩和が期待できると考えている。

3 研究の展望

先に提案したシステムを実現するためには、実現しなければならない課題が二つある。

- (1) 検索キーワードに対応するファセットを提示する方法
- (2) システムが保有するドキュメントから最適なファセットを推論して生成する方法

(1) ファセット検索は無限に数があるけれど、適切なファセットだけを提示できるといいよね従来のファセット検索は。。。 (ファセット検索は用意するものであるが、自動生成したいよね) 上記の課題を解決するためには、語彙間の関係性を抽出することがが共通した重要なタスクとして考えられる。これまで語彙同

士の関係性を抽出する研究は、情報抽出の領域における banko らによる研究 [3] を皮切りに、OpenIE(Open Information Extraction)という分野にて研究されてきた。これらの

4 まとめ

brabra

参考文献

- [1] 福島健介・小原 格・須原慎太郎・生田 茂 (2005), ”インターネット検索能力の差異に及ぼす 要因の検討 その1”, コンピュータ&エデュケーション VOL.18 2005
- [2] 齋藤ひとみ・三輪和久 (2004), ”Web 情報検索におけるリフレクションの支援”, 人工知能学会論文誌 19 巻 4 号 C(2004 年)
- [3] Michele Banko, Michael J Cafarella, Stephen Soderland, Matt Broadhead and Oren Etzioni(2007), ”Open Information Extraction from the Web”, IJCAI'07 Proceedings of the 20th international joint conference on Artificial intelligence Pages 2670-2676