

## Web 上の実世界行動検索システム

ウェブ上には人々が実世界で行った行動が様々な形で掲載されている。これらの行動に関する情報は、主に、最も単純なものから非常に複雑なものまでHowToと呼ばれるウェブサイト上に掲載されている。これらのHowTo サイトでは、人が想像可能なあらゆる目標に対し、それを実際に遂行する方法が詳述されている。人々は検索エンジンを使用してこれらの活動を探索可能である。ここで、行動を表現するためには名詞に加え動詞も必要である。そのため、ユーザの要望にあった行動を探すためには、名詞と動詞をうまく組み合わせるクエリを構成する必要があるが、初心者のユーザはそのような組み合わせを構築するのが難しい。

このような背景のもと本研究では、直接検索クエリを含まないが、概念的にそれと関係がある行動を検索可能なシステムを提案する。具体的には、ConceptNetと呼ばれるMITによって作られた知識ベースを使用して、検索クエリ上で質問拡張を行なうシステムを構築する。質問拡張は、オリジナル質問に含まれる単語とConceptNet上で結ばれる複数の関連語を用いる。関連語には無関係のものも含まれるため、単語と関連語間の関係に紐付けられた重みを計算し、重み順にランキングすることで関連語の候補を抽出する。なお、単語-関連語の重み付けは、遺伝的アルゴリズムを使用して、最適化を実施した。その後、HowToウェブサイトから関連語を用いて関連するページを検索しユーザに応答する。

最後に、最適化されたシステムと最適化されていない単純なウェブ探索と比較し、最適化されたシステムのほうが一方の手法に比べ優れていたことが分かった。将来、多人数を集めシステム全体の評価を行う予定である。

**Keywords:** Data mining, Query expansion, Optimization, Genetic Algorithm

### Reference

- [1] Karapetsas E. Fukazawa Y. and Jun O. 2011. Retrieving Information about Real World Activities from the Web. IPSJ SIG technical reports 2011-MBL-59(2), 1-7, 2011-08-29

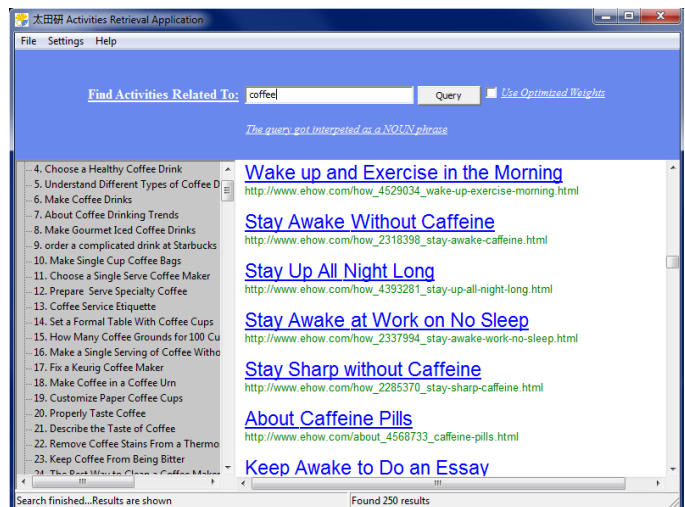


Fig 2 - The GUI of the system.

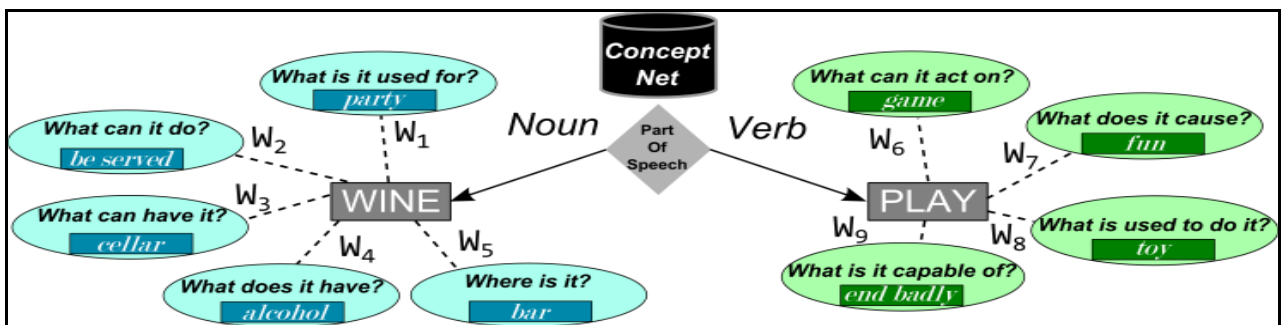


Fig. 1 - ConceptNet and relation weights