

語の出現パターンと意味関係分析を用いた Webからのタスク検索

加藤 龍^{a)} 大島 裕明^{1,b)} 山本 岳洋^{1,†1,c)} 田中 克己^{1,†1,d)} 加藤 誠^{1,†1,e)}

概要：本研究では、クエリとしてタスクが与えられた際に、そのタスクを達成するために必要なサブタスク群を Web から発見する手法を提案する。提案手法では、動詞の含意関係や逆意関係にもとづいたルールによりクエリの変換を行う。変換したクエリで Web 検索を行う。得られたページから、タスク特有の言語パターンを用いて、サブタスク候補となるフレーズを発見する。

キーワード：情報処理学会論文誌ジャーナル， \LaTeX ，スタイルファイル，べからず集

Subtask search

RYO KATO^{a)} HIROAKI OHSHIMA^{1,b)} TAKEHIRO YAMAMOTO^{1,†1,c)} KATSUMI TANAKA^{1,†1,d)}
MAKOTO P. KATO^{1,†1,e)}

Abstract: This document is a guide to prepare a draft for submitting to IPSJ Journal, and the final camera-ready manuscript of a paper to appear in IPSJ Journal, using \LaTeX and special style files. Since this document itself is produced with the style files, it will help you to refer its source file which is distributed with the style files.

Keywords: IPSJ Journal, \LaTeX , style files, “Dos and Dont’s” list

1. 動機

Web 検索

Google 目的

誕生以来、Web には多様な資料が集積され続けている。情報をどう走査、発見し、取捨選択するか、様々な方向からのアプローチがなされている。現在、いくつかの検索エンジンが実用化され、Web を舞台に情報発見や集約を行っている。だが、Web における情報検索は、いまだ多くの問題を抱えている。そのうちの一つとして、目的を達成す

る方法を

なにかを成し遂げたいが、どうすれば成功するかわからないとき、Web 検索を使って実現方法を考える行為がよく行われている。たとえば「花粉症対策をする方法」をクエリに検索することで、立体マスクをつける」や「アレロク錠剤を飲む」を発見することができる。だが、そうして発見できた「花粉対策をする方法」を採用すべきなのか、簡単にはわからない。まだ発見できていない「花粉対策をする方法」のほうがよいかもしれないからだ。

こうした状況では、ユーザーは「タスクを遂行するためにどんな方法があるか」をできるだけ多く発見するサブタスク検索を求める。多様な答えが得られることで、安心して各方法を比較したり、自分にとって最適な方法を考えることができるようになる。

サブタスク検索の例を説明する。たとえば「花粉症対策をする」というクエリを入力すると、出力として

¹ 情報処理学会
IPSJ, Chiyoda, Tokyo 101-0062, Japan
^{†1} 現在、京都大学
Presently with Kyoto University
^{a)} r.kato@dl.kuis.kyoto-u.ac.jp
^{b)} ohshima@dl.kuis.kyoto-u.ac.jp
^{c)} yamamoto@dl.kuis.kyoto-u.ac.jp
^{d)} ktanaka@i.kyoto-u.ac.jp
^{e)} kato@dl.kuis.kyoto-u.ac.jp

- 立体マスクをつける
- アレロック錠剤を飲む
- 医師の診断を受ける
- 植物に近寄らない

といったように、複数の異なった選択肢を網羅的に発見する。これがサブタスク検索の例である。

このような、網羅的に手法を探す検索は現在の一般的な検索エンジンでは困難である。「iPhone のゲームを作る方法」で Google 検索すると、iPhone アプリの作り方講座まとめや cocos2d の紹介などのページはヒットする。しかしその検索結果は iPhone のゲームを作る方法を網羅しているわけではない。「Titanium を使う」や「CoreAnimation を使う」といった手法は「iPhone のゲームを作る方法」での検索では容易に発見できない。

2. 関連研究

?において、広告を用いたクエリクラスタリングが試みられている。タスク実行のための、複雑な検索の手順を研究したものとして hassan がある。

3. タスク関係の定義

3.1 タスクの定義

「遂行すると、目的の一部あるいは全部を達成したことになる行動」をタスクと定義する。

「目的を達成した状態をゴールと呼ぶ」

「ゴールに至る行動がタスクといえる」

3.2 スーパータスクとサブタスクの定義

4. タスク構造の概要

「目的を達成すること」は「目的を達成した状態になる」と等価であり、言い換えられる。

目的をゴールと呼ぶ。

目的を達成する、つまり目的を達成している状態になる行動がタスクといえる。

4.1 ゴールの定義

状態が、求める状態になっていること

5. Web ページ発見の手法

6. サブタスクタスク発見の手法

7. 評価

実験を行った。

7.1 実験の手法

ベースラインとして

7.2 実験の結果

8. 考察

9. 結論

参考文献

- [1] The Widom of Advertisers: Mining Subgoals via Query Clustering 入手先 (<http://research.microsoft.com/en-us/people/tesakai/cikm2012yamamoto.pdf>) (2012.11.02).
- [2] Task Tours: Helping Users Tackle Complex Search Tasks: Ahmed Hassan, Ryen W. White 入手先 (<http://research.microsoft.com/pubs/178868/HassanCIKM2012.pdf>)