

プログラム作成時の検索能力の差異に関する分析

中才 恵太朗 角田 雅照
近畿大学理工学部情報学科

1はじめに

プログラム作成に当たって検索エンジンを利用することにより効率を高めることができる。しかし、検索エンジンを利用し、必要な情報を取り出す能力は、個人により差異があると考えられる。例えば、C言語で文字列を操作したい場合、”c”だけで検索してC言語入門のサイトのメニューから必要な情報を得る人や”c 文字列操作”といったAND検索をしてすぐに必要な情報を得る人などの差異がある。

そこで、プログラミング熟練者はプログラミング初心者より効率のよい検索ができるのではないかと仮定した。本稿では、プログラム作成時の検索方法を知ることを目的として、どのような検索をすることで必要な情報を得られやすいのかを分析した。

本稿では、以下のリサーチクエスチョンを設定した。

- RQ1:時間あたりの検索結果参照数が多い場合、正解に到達しやすいのか？
- RQ2:1回あたりの検索ワード数の平均が多い場合、正解に到達しやすいのか？
- RQ3:問題あたりの検索の種類が少ない場合、正解に到達しやすいのか？

検索結果参照数とは検索結果ページのウィンドウをアクティブにした数であり、検索ワード数とは空白で区切られた単語の数である。

2実験

この実験の目的は、プログラマによって検索エンジンの使い方にどのような違いがあるかを調査することである。そのため、被験者に著者らが用意したプログラムの問題を解答させ、その時の行動を[1][2]のソフトを使用しデータを収集した。また、問題の正誤判定と、解答時間を測定するため正誤判定ツールを作成した。

リサーチクエスチョンに答えるため、Java言語の問題を解く実験を行った。実験対象は学部4年生3人と学部3年生1人である。以下の2種類の問題を実際に解いてもらった。

問1: 浮動小数点の誤差を解決する問題

問2: Streamのライブラリの使い方の問題

Javaで浮動小数点を扱う際には誤差が生じることがある。問1はその誤差を解決するためのライブラリを探す問題である。問2は文字列を読み込むライブラリの使い方についての問題である。また、実験中にわからないことがあれば、検索エンジンでそれを調べてもらった。問題を解く時間は特に定めなかったが、各問題を解き初めて少なくとも20分間は問題を解いた。

てもらった。

3分析結果

被験者4人をABCDと表す。問1ではA, B, Cが正解し、問2ではA, Bが正解した。正解のグループをTG、不正解のグループをFGとする。問1の結果は、被験者がまだ問題に慣れていない状態のため、今回は問2に対する結果について分析した。分析結果を表1に示す。

表 1: 問 2 に対するメトリクス

被験者	A	B	C	D
検索結果参照数/min	0.57	0.66	0.49	0.19
検索ワード数平均	3.88	2.38	2.5	2.78
検索の種類	3	4	6	9

検索結果参照数がともにFGよりTGの方が高い、したがってRQ1に対する答えはYESである。また、検索ワード数平均は、FGよりAは高いがBはCより低いのでRQ2に対する答えはNOである。検索の種類に関してはTGとFGではTGの方が少ないのでRQ3に対する答えはYESである。

ただし、RQ1, RQ3の分析については以下に述べる問題がある。検索結果参照数はウィンドウをアクティブにした数であり検索結果ページのタブを頻繁に切り替えることで急激に増加する。また、不正解の場合だと正解の場合より多くの時間を使用するため、検索の種類が増えることは当然である。なお、RQ3からは、数回検索してわからない場合、正解にたどり着きにくいということがいえる。したがって、数回検索して結果がわからなければ、検索ワードを大きく変えるか人に聞くなどで検索戦略を変えたほうが良いと考えられる。

4おわりに

さらに信頼性の高い分析をするために被験者を増やすこと、より適切なメトリクスを測ることが今後の課題である。

謝辞 本研究の一部は、文部科学省科学研究補助費（挑戦的萌芽：課題番号 26540029、基盤C：課題番号 25330090）による助成を受けた。

参考文献

- [1] きいろがあ
(<http://keylog.web.fc2.com/keyfree/keyfree.html>)
- [2] ManicTime
(<http://www.manictime.com/>)