アルゴリズムとデータ構造　第三回レポート課題

2020/06/28

US162039：梶田悠

2020

目次

[課題内容 1](#_Toc15502695)

[処理手順 2](#_Toc15502696)

[ソースコード 2](#_Toc15502697)

[実行結果 2](#_Toc15502698)

[考察 3](#_Toc15502699)

# 課題内容

Wikipediaのデータを分析し、カテゴリに属するページタイトルを得るためのプログラムを作成する。 Wikipediaの各ページには，複数のカテゴリが付与してある。例えば、「file長良川」のページは「愛知県の河川」、「木曽川水系」、「名水百選」等といったカテゴリが付与されている。Wikipediaのページタイトル（例：「長良川」）と、そのページに付与されたカテゴリを記録したファイル（data.txt）と、全てのカテゴリ一覧記録したファイル（category.list）からは、ページに付与されているカテゴリは分かるが、あるカテゴリに属しているページが分からない。そこで，データを分析し、カテゴリに属するページタイトルを表示できるプログラムを作成した（sample.cpp）。このようなデータを作成しておくことで、カテゴリに属するページを高速に検索することが可能となる。 しかし、sample.cpp は非常に動作が遅く、処理が完了するまで約１時間程度の時間がかかる。そこで、データ構造にハッシュを採用することで処理を高速化させ、短時間（全て表示が終わるまで約１０秒）で処理が終了できるようにプログラムを改良せよ。

# 処理手順

category.listファイルからカテゴリ名を走査し、カテゴリ名をキーにハッシュを

## クイックソート

　あ

## ハッシュ構造

　あ

# ソースコード

作成したソースコードは[github](https://github.com/HarukaKajita/DataStructure_report03)にアップロードしています。見づらい場合や情報に不足がある場合は適宜参照して下さい。

|  |
| --- |
|  |

# 実行結果

実行結果 1　report3.cppの実行結果の抜粋

|  |
| --- |
|  |

# 考察

　あ