Pineapple Distance

Capella Edwards

Group

- Project name: Pineapple Distance
- Group name: Pineapple Fanclub
- Members: Capella Edwards
- Responsibilities: All of them

Introduction

- We need a way to determine the importance (or relevance) of concepts
- In this case, I want to find how closely related a given topic is to a pineapple

- Note: This program is for **Windows**

Motivation

- Pineapples are a very cool fruit
  - multi-fruit: literally a fusion of berries
  - contains a digestive enzyme (that's why it's tingly)
  - produces up to or over 200 flowers
- Therefore, pineapples are used as the basis of importance. Any input word is given a "pineapple distance", where a low score means it is more directly related to pineapples, and a high score means it is more distant.

Method

- Start from the Wikipedia page on pineapples
- Search whether your term is on the page
- Look at other links on the page, then search those pages
- Track how many pages away you've gotten from the Pineapple page: this is "pineapple distance"

Method

- Written in Java
- Overhead "queue" modified by all threads
    - (synchronization avoids critical section problem)
- Queue avoids recursive threads (which would grow exponentially)
- Creates a new thread to look at each given page, searching it for the given term and adding all the links it finds to the queue
- Searches max. 10 pages, BUT data is saved to files so you can reopen the program and continue where you left off with 10 more (if your term is not found by then)

Results

- Distance to given topic from Pineapple (i.e. how many clicks did it take to reach a page mentioning your search term?)

- Input: search term
- Output: Pineapple Distance
    - (and the page your term was found on)

Results

```
PS E:\GSU\class\csc6320\project_j> java Main
Welcome to Pineapple Distance. Please enter a search term (or 'c' to continue from where you left off):
distance
Searching for: distance
Viewing page Pineapple
Viewing page Decades of the New World
Viewing page Wheatgrass
Viewing page Powder (substance)
Viewing page Tincture
Viewing page Syrup
Viewing page Drinkware
Viewing page Cup
Viewing page Cup-bearer
Viewing page Cleveland
Viewing page Cuyahoga River
Term found, stopping threads...
Found distance on page Cuyahoga River, 8 click(s) away from Pineapple.
```

Threads make pages from different search paths appear between one another, but you can still verify the path yourself. In this example, Pineapple (0)'s expandable "Juice" section links to Wheatgrass (1), which links to Powder (2), which links to Tincture (3), which links to Syrup (4), whose expandable "Bartending" section links to Drinkware (5), which links to Cup (6), which links to Cleveland (7), which links to Cuyahoga River (8 clicks!), which mentions the search term "distance"!

- note that Decades of the New World and Cup-Bearer are not part of this path; it's not just counting all pages seen, it really is tracking the "pineapple distance"
- **Your results may vary!** This program uses threads and ejects old links from the queue (to avoid getting stuck too long in one place, & storing an unwieldy amount of links it's never going to get to).

Results

Alternate pathway taken by the same search as the previous slide, this time taking multiple searches with the save/continue function

```
Welcome to Pineapple Distance. Please enter a search term (or 'c' to continue from where you
 left off):
distance
Searching for: distance
Viewing page Pineapple
Viewing page Decades of the New World
Viewing page Wheatgrass
Viewing page Powder (substance)
Viewing page Tincture
Viewing page Syrup
Viewing page Ice cube
Viewing page Insufflation (medicine)
Viewing page Solution (chemistry)
Viewing page Powder
Viewing page Molecular encapsulation
Searched many pages, but could not find your term!
To continue searching another batch, run the program again and enter 'c' as your search term
.
Saving & quitting... (please do not manually quit or data may not save correctly)
PS E:\GSU\class\csc6320\project_j> java Main
Welcome to Pineapple Distance. Please enter a search term (or 'c' to continue from where you
 left off):
c
Searching for: distance
Viewing page Softgel
Viewing page Hydrogel
Viewing page PMID (identifier)
Viewing page PubMed Central
Viewing page PubMed
Viewing page Online patient education
Viewing page Author-level metrics
Term found, stopping threads...
Viewing page Bibliometrics
Term found, stopping threads...
Found distance on page Author-level metrics, 14 click(s) away from Pineapple.
Clearing data & quitting...
Found distance on page Bibliometrics, 14 click(s) away from Pineapple.
Clearing data & quitting...
```

References

- https://en.wikipedia.org/wiki/Pineapple
- https://www.w3schools.com/java
- https://www.oracle.com/java/technologies/downloads/#jdk21-windows
- https://docs.oracle.com/en/java/javase/20/docs/api/