OOA

OOA1

Finding Classes and Objects

- Document

- Stoplist
- CLI
- SearchEngine

Finding Attributes (these hold state information)

- name, content, stoplist, invertedIndex, display

Finding Services

- readContent(), iterator(), hasStopWord(), buildIndex(), makeClean(), findQuery(), display(), findMultiWord()

Finding Inheritance

- Document implements Iterator

Finding Delegation

- SearchEngine has a collection of documents

OOA2

class Document

attributes

- String name - stores the name of the Document
- String content - stores content of document

sevices

- void readContent(String filename) - reads each line of file and stores it in content
- Iterator iterator() - Iterates through String

class Stoplist

attributes

- Set stoplist - stores words in stoplist file

services

- boolean hasStopWord(String w) - returns true if word is a word in set

class SearchEngine

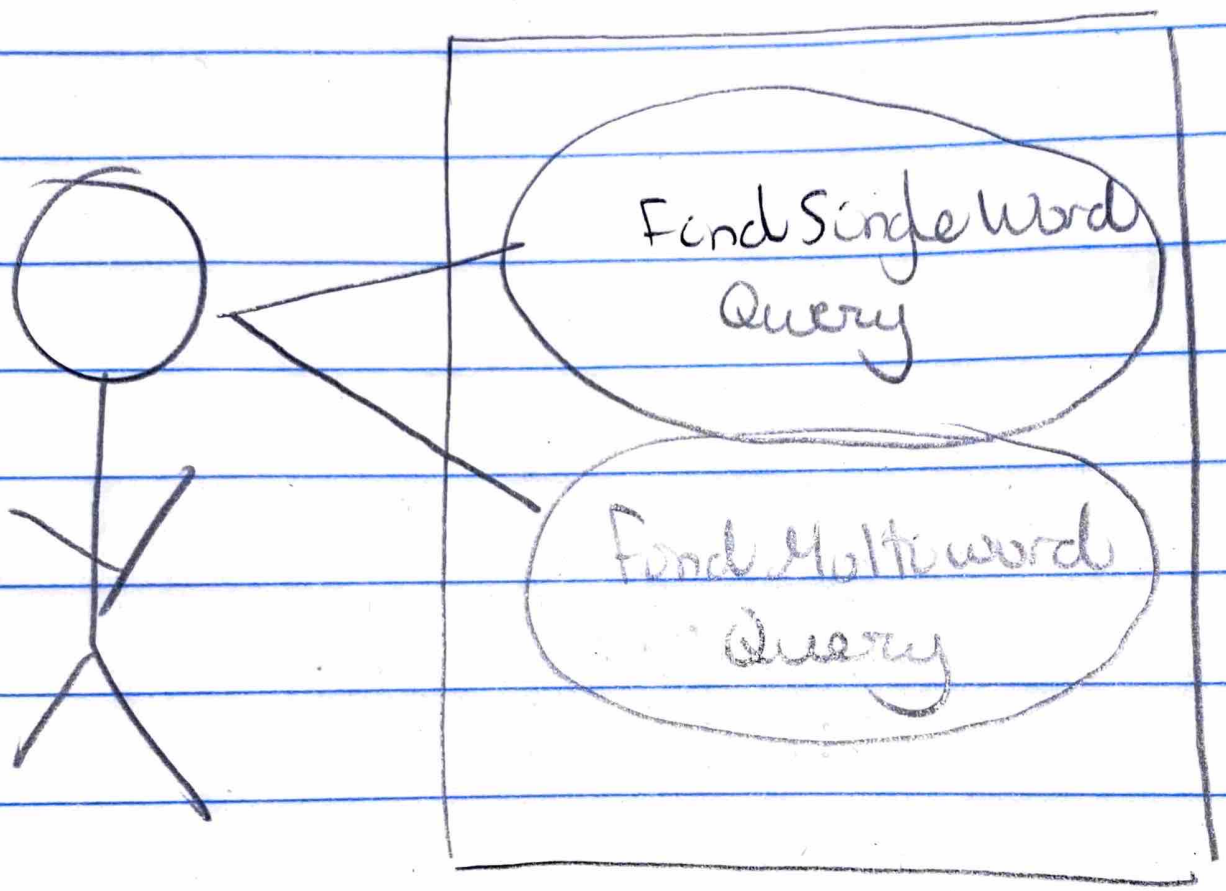
attributes

- HashMap> invertedIndex - stored word key with documents associated to it
- Stoplist stoplist - stoplist collection
- boolean display - the way to display

services

- void buildIndex(Document doc) - buildes invertedIndex
- Set makeClean(Document doc) - cleans document content
- void findQuery(String query) - Finds query(the key) in hashmap
- void displayInvertedIndex() - displays InvertedIndex
- void display(String query, Set querydocs) - displays output
- void findMultiWord(String [] query, String querytofind) - finds query of multiple words

Use Case diagram



Use cases

Use case 1

- Name: FindSingleWordQuery
- Actors: User
- Purpose: finds all the documents related to the word query
- Main success scenario: All documents where you can find the query are outputted by program
- Alternate scenario: No documens contain the word. Word is in Stoplist

Use case 2

- Name: FindMultiWordQuery
- Actors: User
- Purpose: finds all the documents that contain all query words in them
- Main success scenario: All documents where you can find all words in query are outputted by program
- Alternate scenario: No documens contain all the words. Words are in Stoplist.

UML - diagram - retrieval - masosco