

612 Lab 2 (3 points)

Positional index construction and phrase query processing

Overview

This lab consists of two major tasks:

- Building a positional index, and
- Use the positional index to process phrase queries with any number of query terms

Resources

- You should have read Chapter 2 of Introduction to Information Retrieval.
- Carefully read the lecture examples of weeks 3 and 4 to understand the technical details.
- Go over the lecture notes of weeks 3 and 4.

Note: Make JavaDoc comments in your Java programs including Course #, Lab #, Your name, and main functional description of each method with @param & @return if applicable at the minimum.

Ref. <http://www.oracle.com/technetwork/articles/java/index-137868.html>

Tasks to be completed

- Task1: Complete the constructor of the class that builds the positional index.
- Task2: Implement the intersect method that takes in two postings and output a merged posting.
- Task3: Implement the phraseQuery method that takes in a phrase query with multiple terms and return a list of Doc objects (the Doc class is given in the Appendix).
- Task4: Design and test phrase queries with 2-5 terms.

Submit your programs to a lab drop box in MyCourses by 11:59PM March 15th.

Appendix

```
class Doc{
    int docId;
    ArrayList<Integer> positionList;
    public Doc(int did)
    {
        docId = did;
        positionList = new ArrayList<Integer>();
    }
    public Doc(int did, int position)
    {
        docId = did;
        positionList = new ArrayList<Integer>();
        positionList.add(new Integer(position));
    }

    public void insertPosition(int position)
    {
        positionList.add(new Integer(position));
    }

    public String toString()
    {
        String docIdString = ""+docId + "<";
        for(Integer pos:positionList)
            docIdString += pos + ",";
        docIdString = docIdString.substring(0,docIdString.length()-1) + ">";
        return docIdString;
    }
}
```