Inverted Index
Profesor Heider Sanchez

Se le pide desarrollar un programa [preferiblemente en Python] para permitir la recuperación de documentos usando el Índice Invertido.

Probar el correcto funcionamiento del índice usando una colección de resúmenes de 6 libros de "El Señor de los Anillos". Considerar el siguiente proceso:

1. Preprocesamiento
    a. Filtrar los stopwords de cada uno de los textos. Para ello debe usar un stoplist estándar (countwordsfree)
    b. Retirar signos innecesarios.
    c. Reemplazar cada palabra por su raíz (Stemming).

    Puede ayudarse usando las librerías de nltk de Python solo para esta etapa.

2. Construcción del índice invertido:
    a. Construir el índice con los 500 términos más frecuentes de toda la colección.
    b. Guardar el índice en un archivo de texto. Por fines comparativos, ordenar el índice alfabéticamente. Formato:

```
W1:1,3
W2:2,5,6
W3:1,2
W4:2
```

3. Aplicar Consultas Booleanas:
    a. Implemente la función de recuperación booleana y los operadores AND, OR y NOT. Ejemplo:
    Query:

    "*(Comunidad AND Frodo) AND NOT Gondor"*

    Ejecución:

    result = recovery(AND(AND(L("*Calpurnia*"), L("*Brutus*")), L("Cesar"))

    En donde:

    - L() retorna la lista de publicaciones asociadas al termino
    - AND() retorna los documentos que contienen a ambos términos de manera conjunta.
    - OR() retorna los documentos que contienen a al menos uno de los términos.
    - AND-NOT() retorna los documentos que contienen al primer termino pero no al segundo.

    b. Probar el programa con al menos 3 consultas y al menos 3 términos.
    c. (opcional) Elabore un parser que transforme la consulta textual booleana en instrucciones de ejecución.

Entregable: elabore un informe detallando la implementación del índice invertido y los resultados obtenidos.