Progetto intermedio 1

Nella cartella del progetto è possibile trovare i seguenti file:

* DataCounter.java
* MyDataCounter\_Hashtable.java
* MyDataCounter\_TreeMap.java
* Applicazione.java
* Test\_per\_implementazione.java
* Test\_per\_applicazione.java

E i seguenti testcase:

* Prova1.txt
* Prova2.txt
* Prova3.txt
* Prova4.txt
* Prova5.txt
* Prova6.txt

I primi due test verificano il corretto funzionamento della classe Applicazione su dei paragrafi estratti dal testo del progetto, gli altri test sono per verificare che la classe Applicazione restituisca dei risultati corretti anche con file di testo più particolari, prendendo infatti in input un file di testo vuoto, uno contenete molti invii, uno contenente molti spazzi. Le classi MyDataCounter\_TreeMap e MyDataCounter\_HashTable implementano come richiesto l’interfaccia DataCounter utilizzando rispettivamente come variabili di stato una treemap e hashtable. Per entrambe le due classi si mette a disposizione due metodi costruttori, uno che crea un oggetto vuoto (gli elementi possono essere aggiunti tramite il metodo incCount), e un altro che prende come parametro una map e crea un oggetto con gli stessi elementi della map passata come parametro. La classe Applicazione contiene il metodo applicazione che prende in input un file di testo esistente, e utilizzando un oggetto di tipo MyDataCounter\_Hashtable<String> (si noti che è possibile anche utilizzare un oggetto di tipo MyDataCounter\_Treemap<String> modificando una sola istruzione del codice di Applicazione) prende una stringa, la normalizza e poi la inserisce nell’oggetto di tipo DataCounter. Infine si stampa a video le occorrenze delle stringhe, il numero di parole, e il numero di caratteri con e senza spazi presenti nel file di testo. La normalizzazione consiste nel rimuovere tutti i caratteri non alfanumerici e rimpiazzarli con uno spazio, e nel sostituire tutte le lettere maiuscole con la rispettiva lettera minuscola. Ho deciso di non sostituire le lettere accentate, per evitare ad esempio che la lettera “è” che un predicato sia conteggiata come lettera “e” di congiunzione. Nella classe Test\_per\_implementazione si fanno una serie di test sull’implementazione delle classe MyDataCounter\_TreeMap, anche in questo caso se vogliamo utilizzare la classe MyDataCounter\_Hashtable è sufficiente fare una piccola modifica al costruttore utilizzato. Per l’esecuzione della classe Test\_per\_implementazione si consiglia di abilitare le asserzioni, in quanto ne faccio uso per i test. Ho deciso di utilizzare MyDataCounter\_Hashtable nella classe applicazione perché al caso medio il tempo di inserimento in una tabella hash è costante mentre in una treemap il tempo di inserimento è logaritmico.