

Übung 07: Stringsuche

Abgabetermin: 22.06.2010 12:00

Name: _____

Matrikelnummer: _____

Gruppe: G1 (Wolfinger) G2 (Wolfinger)

Aufgabe	Punkte	Gelöst	abzugeben schriftlich	abzugeben elektro- nisch	Korr.	Pkte
Aufgabe 07.1	12	<input type="checkbox"/>	Java-Programm Testfälle und Ergebnis- se	Java-Programm	<input type="checkbox"/>	

Aufgabe 07.1: Stringsuche nach Boyer-Moore (12 Punkte)

a) Implementieren Sie zwei Algorithmen zur Stringsuche: In der Klasse *BruteForceSearch* nach dem Brute-Force-Ansatz und in der Klasse *BoyerMooreSearch* nach Boyer-Moore. Beide Klassen sollen die Schnittstelle *StringSearch* bereitstellen.

```
public interface StringSearch {
    public enum SearchOption { NONE, WHOLE_WORD };
    // returns the number of occurrences of pat in text
    // e.g., search("abcdsaoab", "ab") returns 2.
    int search(String text, String pat);
    // e.g., search("abc ab", "ab", StringSearch.SearchOption.WHOLE_WORD) returns 1
    int search(String text, String pat, SearchOption option);
}
```

- *int search(String text, String pat)* liefert die Häufigkeit mit der das Musters *pat* im Text *text* vorkommt. Mit der *SearchOption.WHOLE_WORD* findet *search()* das Muster nur dann, wenn es als eigenes Wort vorkommt.

b) Definieren Sie Testfälle und testen Sie Ihr Programm.

c) Vergleichen Sie die Suchzeiten der beiden Suchalgorithmen. Lesen Sie den zu durchsuchenden Text (*alice.txt*) und die zu suchenden Worte (*search.txt*) aus den Vorgabedateien ein. Jede Zeile in *search.txt* enthält einen Suchbegriff.

alice.txt:

```
Alice's Adventures in Wonderland
Lewis Carroll
CHAPTER I
Down the Rabbit-Hole
Alice was beginning to get very tired of sitting
by her sister on the bank, and of having nothing
to do: once or twice she had peeped into the book
her sister was reading, but it had no pictures...
```

search.txt:

```
a
affectionately
alarm
alarmed
alas
alice
alive
all
allow
almost...
```

Wiederholen Sie die Suche jedes Wortes 1.000 Mal um Messungenauigkeiten zu reduzieren. Errechnen Sie pro Wort die Gesamtsuchzeit und stellen Sie diese gemeinsam mit der Häufigkeit in einer Tabelle mit folgendem Aufbau dar:

Search	#	Brute	Boyer
a	634	1233	921
affectionately	1	1355	113
alarm	2	1376	211

Implementierungshinweis: Wenn Sie die Testdatei bei jeder Suche neu lesen, verfälscht das die Messergebnisse. Lesen Sie daher statt dessen die Datei *alice.txt* bereits beim Programmstart *In.readFile()* in einen String. Lesen Sie außerdem die Datei *search.txt* in eine *ArrayList*.

Abzugeben ist:

- Das Java-Programm
- Die Testausgabe
- Die Tabelle mit den Messergebnissen