Technische Universität München
Institut für Informatik
Prof. Dr. Ernst W. Mayr

SS 1998
Aufgabenblatt 10
20. Juli 1998

Praktikum Algorithmen-Entwurf

(Abgabetermin: Montag, den 27. Juli 1998, $24.^{00}$Uhr)

Textprobleme und Dynamisches Programmieren

Aufgabe 1 Edit-Distanz

Implementieren Sie eine Funktion, die für zwei gegebene ASCII-Texte X und Y (übergeben als Parameter vom Typ "`char *`") der Länge m und n in Zeit $O(mn)$ die Edit-Distanz zwischen X und Y berechnet. Dabei können Sie entweder die Lösung verwenden, die mittels dynamischen Programmierens eine Tabelle konstruiert, oder die Lösung, die das Problem auf ein Kürzeste-Wege-Problem in einem gerichteten, azyklischen Graphen zurückführt.
Bei der ersten Alternative (dynamisches Programmieren) muß Ihre Funktion ohne LEDA auskommen und Sie müssen die Tabellenkonstruktion so optimieren, daß nur Speicherplatz $O(\min\{m, n\})$ verwendet wird. Bei der Zurückführung auf ein Kürzeste-Wege-Problem können Sie dagegen LEDA einsetzen und auch Ihren Algorithmus zur Bestimmung kürzester Pfade in azyklischen Graphen (Blatt 4, Aufgabe 2) wiederverwenden.

Aufgabe 2 Längste gemeinsame Teilsequenz

Implementieren Sie eine Funktion, die für zwei gegebene ASCII-Texte X und Y (übergeben als Parameter vom Typ "`char *`") der Länge m und n in Zeit $O(mn)$ die Länge einer längsten gemeinsamen Teilsequenz berechnet. Wieder sind beide Lösungsmöglichkeiten (dynamisches Programmieren, Kürzeste-Wege-Problem) unter denselben Voraussetzungen zugelassen.

Aufgabe 3 Suchen nach Übereinstimmungen mit Fehlern

Implementieren Sie eine Funktion, die für zwei gegebene ASCII-Texte X und Y (übergeben als Parameter vom Typ "`char *`") der Länge m und n mittels dynamischen Programmierens in Zeit $O(mn)$ die minimale Zahl k bestimmt, so daß die Edit-Distanz zwischen Y und einem gewissen Teilwort X' von X genau k beträgt. Es soll auch die erste und letzte Position in X ermittelt werden, an der ein solches Teilwort X' endet.

Hinweise

Geben Sie nur **ein einziges Programm** ab, das die Texte X und Y aus einer Datei liest und dann insgesamt 5 Werte ausgibt: die Edit-Distanz zwischen X und Y; die Länge einer längsten gemeinsamen Teilsequenz von X und Y; das minimale k, so daß in Text X eine Übereinstimmung mit Muster Y mit k Fehlern vorliegt; sowie die erste und (evtl. identische) letzte Position in X, an der eine solche Übereinstimmung mit k Fehlern endet.

Als Eingabe für Ihr Programm können Sie die Dateien `textpair*` aus dem Verzeichnis `/usr/local/lea/praktikum/dynprog` verwenden, die nacheinander zwei Texte, jeweils mit "$\n" abgeschlossen, enthalten ('$' kommt sonst in den beiden Texten nicht vor).