

24.11.2000 - Multiple Alignment

Problem: Zu zwei gegebenen Sequenzen und gegebener Score-Funktion, berechne das score-optimale Alignment der beiden Sequenzen.

Beobachtung: Die Sequenzen seien $s_1 = xX$, $s_2 = yY$, dabei seien x und y die Anfangsstuecke (Praefixe) von s_1 und s_2 , und X bzw. Y sei jeweils der letzte Buchstabe von s_1 bzw. s_2 .

Jedes Alignment (insbesondere das optimale) muss auf eine von drei Arten Enden:

1. ...X dabei bedeuten die Punkte ein Alignment von x mit yY

...-

2. ...- dabei bedeuten die Punkte ein Alignment von xX mit y

...Y

3. ...X dabei bedeuten die Punkte ein Alignment von x mit y

...Y

Behauptung: Das OPTIMALE Alignment von xX und yY ergibt sich, wenn man die BESTE Moeglichkeit aus folgenden dreien auswählt:

1. Verlaengere das OPTIMALE Alignment von x und yY mit $(X,-)$

2. Verlaengere das OPTIMALE Alignment von xX und y mit $(-,Y)$

3. Verlaengere das OPTIMALE Alignment von x und y mit (X,Y)

Beweisskizze: (Indirekt.) Angenommen, es gaebe noch ein besseres Alignment von xX und yY . Das muesste auf eine der drei oben genannten Arten enden.

Jetzt Fallunterscheidung... Dann haette man auch schon ein besseres Alignment von mindestens einem der drei Vorgaenger-Faelle gehabt.

Widerspruch mit Induktion.

D.h. man hat das optimale Alignment-Problem von xX und yY auf das von kuerzeren Sequenzen zurueckgefuehrt. Man kann die optimale Score von s_1 und s_2 also mit Hilfe einer Tabelle berechnen, in der man die optimalen Scores aller Praefixkombinationen von s_1 und s_2 speichert:

Sei $T[i,j]$ (0