

Präsenzübungen zur Vorlesung Sequenzanalyse II

Universität Bielefeld, SoSe 2012

Dr. Alexander Sczyrba · Nina Luhmann · Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2012summer/SequenzAnalyse>

Präsenzübungsblatt 7, Woche 26/2012

Aufgabe 1 (Center-Star Approximation)

Im folgenden findest du den Beweis, warum die Center-Star Approximation eine 2-Approximation ist. Versuche, die Umformungen nachzuvollziehen und schreibe hin, was in jedem Schritt gemacht wird.

$$\begin{aligned} D_{SP}(A_c) &= \sum_{p < q} D_2(\pi_{\{p,q\}}(A_c)) \\ &= \frac{1}{2} \sum_{(p,q)} D_2(\pi_{\{p,q\}}(A_c)) \\ &\leq \frac{1}{2} \sum_{(p,q)} (D_2(\pi_{\{p,c\}}(A_c)) + D_2(\pi_{\{c,q\}}(A_c))) \\ &= \frac{1}{2} \sum_{(p,q)} (d(s_p, s_c) + d(s_c, s_q)) \\ &= k \cdot \sum_q d(s_c, s_q) \\ &\leq \sum_{(p,q)} d(s_p, s_q) \\ &\leq \sum_{(p,q)} D_2(\pi_{\{p,q\}}(A^{opt})) \\ &= 2 \cdot \sum_{p < q} D_2(\pi_{\{p,q\}}(A^{opt})) \\ &= 2 \cdot D_{SP}(A^{opt}) \end{aligned}$$

Aufgabe 2 (Center-Star mit κ Sequenzen)

Wie gut wird die Approximation des Center-Star Algorithmus für insgesamt k Sequenzen, wenn zur Berechnung der Center Sequenz s_c nicht die Summe über alle paarweisen Sequenzvergleiche mit s_c berechnet wird, sondern die Summe über alle Alignments mit jeweils κ Sequenzen, in denen s_c involviert ist? Wie groß muss κ sein, damit der Algorithmus exakt wird?