Ludwig-Maximilians-Universität München
Institut für Informatik
Prof. Dr. Volker Heun

Wintersemester 2010/11
Semestralklausur
16. Februar 2011

Algorithmische Bioinformatik II

Vorname	Name	Matrikelnummer

Reihe	Platz	Unterschrift

Hiermit stimme ich einer Veröffentlichung meines Klausurergebnisses dieser Semestralklausur unter Verwendung meiner Matrikelnummer im Internet zu. *Ja* ☐ *Nein* ☐

_____
(Unterschrift)

Allgemeine Hinweise zur Semestralklausur

- Vor der Prüfung ist diese Seite mit Vornamen, Namen, Matrikelnummer, Reihe und Platz **leserlich mit Druckbuchstaben** zu versehen und zu unterschreiben.
- Bitte **nicht** in *roter* oder *grüner* Farbe bzw. **nicht** mit Bleistift schreiben.
- Der Studentenausweis und ein amtlicher Lichtbildausweis sind bereit zu halten.
- Die reine Bearbeitungszeit beträgt **150 Minuten**.
- Es sind insgesamt 40 Punkte zu erreichen, zum Bestehen sind 17 Punkte nötig.

Viel Erfolg!

Hörsaal verlassen von ..... bis ..... von ..... bis .....

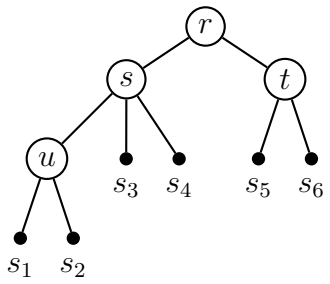
Vorzeitig abgegeben um .....

	Hz	A1	A2	A3	A4	A5	A6	A7	$\sum$
Erstkorrektur									
Nachkorrektur									
Zweitprüfer									

Vorname: _____ Name: _____ Matrikelnummer: _____

Aufgabe 1 (6 Punkte)

Berechne für den rechts angegebenen Baum **alle** optimalen gelifteten Sequenzen gemäß der dynamischen Programmierung.



d	s_1	s_2	s_3	s_4	s_5	s_6
s_1	0	1	1	2	2	3
s_2		0	2	2	3	3
s_3			0	3	3	3
s_4				0	1	1
s_5					0	2
s_6						0

Vorname: _____ Name: _____ Matrikelnummer: _____

Aufgabe 2 (5 Punkte)

Gegeben seien die folgenden vier Sequenzen s_1 = GACTT, s_2 = CGATT, s_3 = GCTTA und s_4 = GCTT. Konstruiere für diese Sequenzen ein mehrfaches Alignment mit Hilfe der Center-Star-Methode.

d	s_1	s_2	s_3	s_4
s_1	0	2	2	1
s_2	2	0	3	2
s_3	2	3	0	1
s_4	1	2	1	0

Hierbei gilt $w(a, b) = 1$ und $w(a, a) = 0$ für alle $a \neq b \in \overline{\Sigma}$.

Lücken dürfen dabei wiederverwendet werden.

Vorname: _____ Name: _____ Matrikelnummer: _____

Aufgabe 3 (5 Punkte)

Bestimme für die folgenden Blöcke von Sequenzen die zugehörigen Häufigkeiten $H(a, b)$ für die BLOSUM60-Matrix.

$$
\begin{array}{ll}
s_1^{(1)} = \text{BBBBC} & s_1^{(2)} = \text{BBBAAACCBB} \\
s_2^{(1)} = \text{ABCBC} & s_2^{(2)} = \text{CBBAABBBBB} \\
s_3^{(1)} = \text{CAABC} & s_3^{(2)} = \text{BCBCAACABA} \\
s_4^{(1)} = \text{AAABA} & s_4^{(2)} = \text{ACACABCABA}
\end{array}
$$

Vorname: _____ Name: _____ Matrikelnummer: _____

Aufgabe 4 (6 Punkte)

Wir betrachten zwei verschiedene Typen von Würfeln namens A und B. Ein Würfel vom Typ A bzw. B zeigt mit Wahrscheinlichkeit $\frac{1}{6}$ bzw. $\frac{1}{2}$ sechs Augen und mit der entsprechenden Gegenwahrscheinlichkeit weniger Augen. Wir wählen einen Würfel und erhalten beim zweimaligen Werfen zweimal sechs Augen.

a) Bestimme den Maximum-Likelihood-Schätzer (ML-Schätzer) für den gewählten Würfel aus dem Parameterraum $\Theta = \{A, B\}$.

b) Bestimme den Maximum-A-Posteriori-Schätzer (MAP-Schätzer) für den gewählten Würfel aus dem Parameterraum $\Theta = \{A, B\}$, wobei die Priori-Wahrscheinlichkeit, dass wir einen Würfel vom Typ A wählen, bei 3/4 und für Typ B bei 1/4 liegt.

Gib nicht nur die Berechnung, sondern auch die Formeln für den Maximum-Likelihood-Schätzer sowie den Maximum-A-Posteriori-Schätzer an.

Vorname: _____ Name: _____ Matrikelnummer: _____

Aufgabe 5 (6 Punkte)

Sei P ein $\mathcal{NPO}$-vollständiges Optimierungsproblem. Zeige, dass aus der Existenz eines Approximationsalgorithmus mit konstanter Approximationsgüte für P die Aussage $\mathcal{NPO} = \mathcal{APX}$ folgt.

Vorname: _____ Name: _____ Matrikelnummer: _____

Aufgabe 6 (6 Punkte)

MinPartition

Eingabe: Ein Folge $x = (x_1, \ldots, x_n) \in \mathbb{N}^n$ mit $x_i \leq \frac{1}{2} \sum_{j=1}^{n} x_j$ für alle $i \in [1:n]$.
Lösung: Ein Teilmenge $I \subseteq [1:n]$.
Optimum: Minimiere $\mu(I, n) := \max\{\sum_{i \in I} x_i, \sum_{i \in [1:n] \setminus I} x_i\}$.

a) Zeige, dass MinPartition $\in \mathcal{NPO}$.

b) Konstruiere eine polynomielle $\frac{3}{2}$-Approximation für MinPartition.

Hinweis: Beweise nicht vergessen. Das polynomielle Approximationsschema aus den Übungen darf dabei **nicht** als Unterprogramm mit $\epsilon = \frac{1}{2}$ verwendet werden.

Vorname: _____ Name: _____ Matrikelnummer: _____

Aufgabe 7 (6 Punkte)

Zeige, dass MinVertexCover $\leq_{\text{PTAS}}$ MinSAT. Gib dazu explizit eine PTAS-Reduktion (f, g, α) an und weise die erforderlichen Eigenschaften einer PTAS-Reduktion nach.

MinVertexCover

Eingabe: Ein ungerichteter Graph $G = (V, E)$.
Lösung: Eine Teilmenge $V' \subseteq V$ der Knoten, so dass von jeder Kante mindestens ein Endpunkt in V' liegt, d.h. $\forall \{v, w\} \in E : \{v, w\} \cap V' \neq \emptyset$.
Optimum: Minimiere $|V'|$.

MinSAT

Eingabe: Eine Boolesche Formel F über $V(F) = X$.
Lösung: Eine erfüllende Belegung $B : X \to \mathbb{B}$, d.h. $B(F) = 1$.
Optimum: Minimiere $\mu(B) = |\{x \in X : B(x) = 1\}|$.

Achtung: Bei MinSAT ist **nicht** die Anzahl erfüllter Klauseln zu minimieren, sondern die Anzahl auf wahr gesetzter Variablen.