Technische Universität München
Fakultät für Informatik
Lehrstuhl für Effiziente Algorithmen
Dr. Volker Heun

Sommersemester 2002
Semestralklausur
16.07.02

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. mit Bleistift schreiben.
- Der Studentenausweis und ein amtlicher Lichtbildausweis sind bereit zu halten.
- Die reine Bearbeitungszeit beträgt **120 Minuten**.
- Es sind insgesamt 40 Punkte zu erreichen, zum Bestehen sind 17 Punkte nötig.

Hinweise zum Multiple-Choice-Teil

- Der rechte Rand ist frei zu lassen. Die Rückseiten und die letzten Seiten dürfen als Schmierpapier verwendet werden.
- Falls ein Kästchen versehentlich angekreuzt wurde, so ist ist dieses vollständig auszufüllen und unmittelbar rechts bzw. links daneben eine neues Kästchen zu malen: ■ ☐. Es gilt dann als nicht angekreuzt.
- Falls dieses Kästchen dennoch wieder angekreuzt werden soll, so ist das rechts bzw. links daneben gemalte Kästchen anzukreuzen: ■ ☒.
- Bei Antworten, die einen Zahlenwert verlangen, ist der konkrete Zahlenwert oder eine kurze Formel einzutragen, also z.B. 24 bzw. 4!.

Viel Glück und Erfolg!

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

Vorzeitig abgegeben um ....

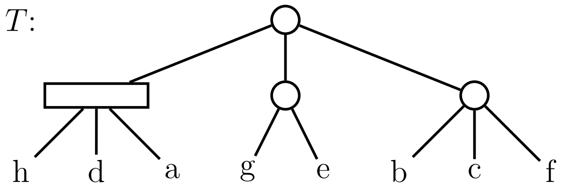
	Seite 2	Seite 3	Seite 4	A1	A2	A3	A4	A5	$\sum$
Erstk.									
Zweitk.									

Meiner Meinung nach war die Klausur . . . leicht ☐ ☐ ☐ ☐ ☐ schwer

Vorname: ______________ Name: ______________ Matrikelnummer: ______________

Aufgabe M1 (2 Punkte)

Sei im Folgenden T der unten angegebene echte PQ-Baum über $\Sigma = \{a, b, c, d, e, f, g, h\}$.



Kreuze zu jeder Aussage an, ob diese richtig [R] oder falsch [F] ist.

- $egadhcfb \in \text{consistent}(T)$. . . . . ☐ R | F ☐
- Es gilt $\text{consistent}(T) \cap \Pi(\Sigma, \{e, f, g, h\}) = \emptyset$. . . . . ☐ R | F ☐
- Wie viele verschiedene zu T äquivalente PQ-Bäume gibt es, deren Frontier mit f beginnen? . . . . ☐

Aufgabe M2 (1 Punkt)

Kreuze zu jeder Aussage an, ob diese richtig [R] oder falsch [F] ist.

- Jeder echte PQ-Baum über Σ besitzt maximal $|\Sigma|$ P-Knoten. . . . . ☐ R | F ☐
- Jeder echte PQ-Baum besitzt mehr P-Knoten als Q-Knoten. . . . . ☐ R | F ☐

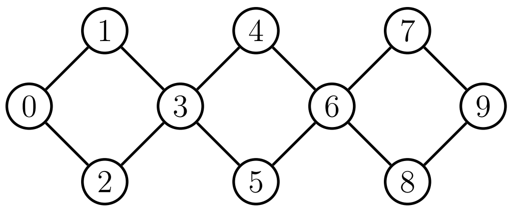
Aufgabe M3 (1 Punkt)

Kreuze zu jeder Aussage an, ob diese richtig [R] oder falsch [F] ist.

- Jeder Baum ist ein Intervall-Graph. . . . . ☐ R | F ☐
- Jeder Intervall-Graph ist ein Einheits-Intervall-Graph. . . . . ☐ R | F ☐

Aufgabe M4 (1 Punkt)

Betrachte folgende Eingabe (V, M, F) für das allgemeine Intervall Sandwich Problem. Hierbei ist $V = [0 : 9]$, der Graph (V, M) ist unten angegeben und F wird in den Teilaufgaben spezifiziert.



Kreuze zu jeder Aussage an, ob diese richtig [R] oder falsch [F] ist.

- Für die Menge $F = \emptyset$ ist die Menge $\{0, 1, 8, 9\}$ ein zulässiger Kern für das allgemeine Intervall Sandwich Problem. . . . . ☐ R | F ☐
- Für die Menge $F = \{\{0, 3\}, \{1, 2\}, \{3, 6\}, \{4, 5\}\}$ ist die Menge $\{6, 7, 8, 9\}$ ein zulässiger Kern für das allgemeine Intervall Sandwich Problem. ☐ R | F ☐

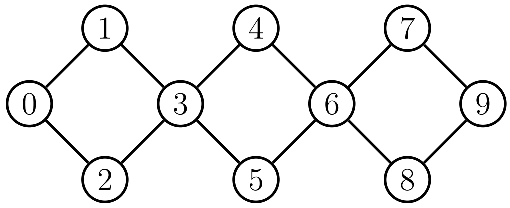
1 ☐

2 ☐

Vorname: __________ Name: __________ Matrikelnummer: __________

Aufgabe M5 (1 Punkt)

Betrachte die folgende Eingabe (V, M, F) für das Bounded Degree Interval Sandwich Problem, wobei der Ausgabegraph maximalen Grad 4 besitzen soll. Hierbei ist $V = [0 : 9]$, $F = \emptyset$ und der Graph (V, M) ist unten angegeben.



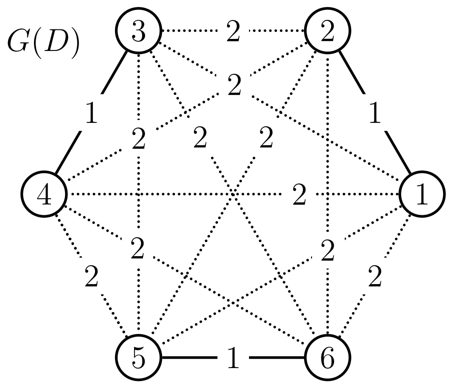
Kreuze zu jeder Aussage an, ob diese richtig [R] oder falsch [F] ist.

- Die Menge $V = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ ist ein zulässiger Kern für das Bounded Degree Interval Sandwich Problem mit maximalem Grad 4. .... ☐ R | F ☐
- Die Menge $\{3, 4, 6\}$ ist ein zulässiger Kern für das Bounded Degree Interval Sandwich Problem mit maximalem Grad 4. .... ☐ R | F ☐

Aufgabe M6 (1 Punkt)

Wir betrachten die folgende symmetrische Distanzmatrix D, die unten angegeben ist und rechts daneben als gewichteter Graph $G(D)$ illustriert ist.

D	1	2	3	4	5	6
1	0	1	2	2	2	2
2		0	2	2	2	2
3			0	1	2	2
4				0	2	2
5					0	1
6						0



Kreuze zu jeder Aussage an, ob diese richtig [R] oder falsch [F] ist.

- Es existiert ein kompakter additiver Baum für D. .... ☐ R | F ☐
- Es existiert ein ultrametrischer Baum für D. .... ☐ R | F ☐

1 ☐
2 ☐

Vorname: ______________ Name: ______________ Matrikelnummer: ______________

Aufgabe M7 (1 Punkt)

Im Folgenden sind zwei binäre 4×5-Charaktermatrizen M_1 und M_2 angegeben.

M_1	a	b	c	d	e
1	0	1	1	1	1
2	0	0	0	1	0
3	0	1	1	1	0
4	1	1	0	0	0

M_2	a	b	c	d	e
1	0	1	1	1	1
2	0	0	0	0	0
3	0	1	1	1	0
4	1	1	0	0	0

Kreuze zu jeder Aussage an, ob diese richtig [R] oder falsch [F] ist.

- M_1 besitzt eine perfekte Phylogenie. .... ☐ R | F ☐
- M_2 besitzt eine perfekte Phylogenie. .... ☐ R | F ☐

Aufgabe M8 (2 Punkte)

Im Folgenden ist eine binäre 4×5-Charaktermatrizen M_3 angegeben.

M_3	a	b	c	d	e
1	0	1	0	0	0
2	1	0	1	0	1
3	0	0	1	1	1
4	0	0	1	0	0

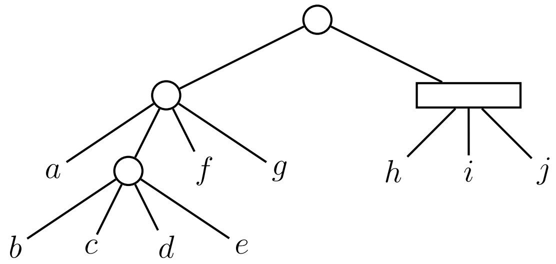
Konstruiere mit Hilfe des in der Vorlesung angegebenen Algorithmus einen phylogenetische Baum für M_3:

1 ☐
2 ☐

Vorname: ________________ Name: ________________ Matrikelnummer: ______________

Aufgabe 1 (6 Punkte)

Betrachte den rechts abgebildeten PQ-Baum. Modifiziere den PQ-Baum gemäß der in der Vorlesung angegebenen Schablonen, so dass zusätzlich auch die Restriktion $\{a, c, e, g, i, j\}$ erfüllt wird. Dabei sind alle Zwischenschritte anzugeben und zu dokumentieren.



Aufgabe 2 (6 Punkte)

Die Distanzmatrizen D_1 und D_2 seien durch die folgenden Tabellen definiert.

D_1	1	2	3	4	5	6	7	8
1	0	5	8	8	1	5	7	7
2		0	8	8	5	2	7	7
3			0	6	7	7	3	6
4				0	7	7	6	4
5					0	5	7	7
6						0	7	7
7							0	6
8								0

D_2	1	2	3	4	5	6	7	8
1	0	4	4	7	7	1	7	7
2		0	3	7	7	4	7	7
3			0	7	7	4	7	7
4				0	6	7	6	2
5					0	7	5	6
6						0	7	7
7							0	6
8								0

Entscheide, welche der Matrizen D_1 und D_2 ultrametrisch ist oder nicht. Im Falle, dass eine Matrix ultrametrisch ist, ist die explizite Konstruktion des zugehörigen ultrametrischen Baumes mit dem in der Vorlesung vorgestellten Algorithmus ausführlich mit allen Schritten anzugeben. Im Falle, dass eine Matrix nicht ultrametrisch ist, ist dies zu begründen.

Vorname: ______________ Name: ______________ Matrikelnummer: ____________

Aufgabe 3 (6 Punkte)

Gegeben seien die beiden folgenden 9×9-Matrizen D_ℓ und D_h. Entscheide, ob es eine ultrametrische Matrix $D \in [D_\ell, D_h]$ gibt oder nicht. Konstruiere dazu entweder einen ultrametrischen Baum für M nach einer der in der Vorlesung angegebenen Algorithmen, oder gib eine Begründung, dass dies nicht möglich ist. Alle Zwischenschritte müssen dabei mit angegeben werden.

D_ℓ	1	2	3	4	5	6
1	0	4	4	1	6	8
2		0	4	3	6	6
3			0	7	2	6
4				0	6	6
5					0	5
6						0

D_h	1	2	3	4	5	6
1	0	6	9	3	9	9
2		0	9	5	8	9
3			0	8	5	9
4				0	8	8
5					0	7
6						0

Vorname: ________________ Name: ________________ Matrikelnummer: ______________

Aufgabe 4 (6 Punkte)

Sei M eine binäre $n \times m$-Charaktermatrix, die eine perfekte Phylogenie besitzt und deren Zeilen paarweise verschieden sind (d.h. es gibt keine zwei Objekte, die genau dieselben Charaktere besitzt). Zeige, dass die $n \times n$-Matrix D_M additiv ist, wobei

$$D_M(i,j) = |\{c \in [1:m] \mid M(c,i) \neq M(j,c)\}|.$$

Vorname: ______________ Name: ______________ Matrikelnummer: ____________

Aufgabe 5 (6 Punkte)

Sei M eine binäre $n \times m$-Charaktermatrix.

a) Zeige, dass M die *Consecutive Ones Property* für die Spalten besitzt, wenn M einen perfekten phylogenetischen Baum besitzt.

b) Konstruiere mit der Aussage aus Teil a) einen effizienten Algorithmus zur Konstruktion phylogenetischer Bäume für binäre Charaktermatrizen.

Erinnerung: Eine Matrix besitzt die *Consecutive Ones Property* für die Spalten, wenn es eine Permutation der Zeilen gibt, so dass in jeder Spalte der Matrix die 1en aufeinander folgend auftreten.