

## Musterlösung zur Klausur vom 22. Juli 2002

### Aufgabe 1

- a) Bestimmen Sie die Wahrscheinlichkeit  $p_n$  dafür, daß beim gleichverteilten Hashing von  $n$  Schlüsseln in  $n$  Zellen in jeder Hash-Zelle genau ein Schlüssel zu liegen kommt.

$$\begin{aligned} p_n &= p(\text{“günstige Einfüge-Folge”}) \\ &= \frac{\text{“Anzahl der günstigen Folgen”}}{\text{“Anzahl aller möglichen Folgen”}} \\ &= \frac{n!}{n^n} \end{aligned}$$

(5 Punkte)

- b) Wie groß ist dieser Wert für  $n = 4$  ?

$$p_4 = \frac{4!}{4^4} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{4 \cdot 4 \cdot 4 \cdot 4} = \frac{1 \cdot 1 \cdot 3}{4 \cdot 2 \cdot 4} = \frac{3}{32}$$

(1 Punkt)

6 Punkte

## Aufgabe 2

Beweisen Sie, daß die eindimensionale Punktlokalisierung in  $n$  aufeinanderfolgenden Intervallen die untere Schranke  $\Omega(\log n)$  hat.

Die Intervallgrenzen seien durch die aufsteigend geordneten Werte  $x_0, x_1, \dots, x_n \in \mathbb{R}$  gegeben. Die Punktlokalisierung für ein  $y \in \mathbb{R}$  berechnet denjenigen Index  $i$ , für den  $x_i \leq y < x_{i+1}$  gilt. Der Definitionsbereich ist also  $y \in [x_0, x_n)$ .

**Hinweis:** Die Aufgabe kann durch einen ausformulierten Reduktionsbeweis oder durch Anwendung von Satz A der Vorlesung gelöst werden.

### Lösungsvariante mit Reduktion:

Wir reduzieren die Berechnung der floor-Funktion auf das gegebene Problem der Punktlokalisierung.

Das heißt, ein Algorithmus zur Punktlokalisierung wird dazu “mißbraucht”, um  $\text{floor}(y)$  für  $0 \leq y < n$  zu berechnen. Hierzu werden die Intervallgrenzen auf

$$x_i := i$$

gesetzt. Nach Anwendung der Punktlokalisierung wird ein Wert  $i$  geliefert, für den gilt  $i \leq y < i + 1$ ; das heißt  $i = \lfloor y \rfloor = \text{floor}(y)$ .

Wird also angenommen, daß die Punktlokalisierung schneller als  $\Omega(\log n)$  ginge, so wäre auch die floor-Funktion schneller zu berechnen. *Widerspruch (!)*, da wir wissen, daß floor eine untere Schranke von  $\Omega(\log n)$  hat.

Lösungsvariante mit Satz A (Lower-Bound-Theorem):

Wähle  $q = n$ .

Wähle die  $q$  Punkte (in deren  $\varepsilon$ -Umgebungen paarweise verschiedene rationale Teilfunktionen  $Q_i$  angenommen werden müssen) inmitten der Intervalle als

$$X_i = \frac{x_{i-1} + x_i}{2} \quad \text{für } i = 1, \dots, n$$

und

$$\varepsilon = \min_{i=1, \dots, n} \left\{ \frac{x_{i-1} - x_i}{2} \right\} > 0$$

Mit den rationalen Funktionen  $Q_i(X) \equiv i$  ist dann für  $i = 1, \dots, n$

$$Q_i(X) \equiv \text{Punktlokalisierung}(X) \quad \text{für } |X_i - X| < \varepsilon$$

Das bedeutet, daß mit  $\log_2 q$  notwendigen Vergleichen eine untere Schranke  $\Omega(\log n)$  für den Aufwand der Punktlokalisierung gegeben ist.

*9 Punkte*

### Aufgabe 3

Beschreiben Sie möglichst strukturiert das Vorgehen beim Löschen eines Schlüssels  $s$  in einem B-Baum mit maximal 4 und minimal 2 Schlüsseln pro Knoten, wobei der Wurzelknoten die übliche Ausnahme darstellt.

Ohne Beschränkung der Allgemeinheit seien die Schlüsselwerte positive ganze Zahlen.

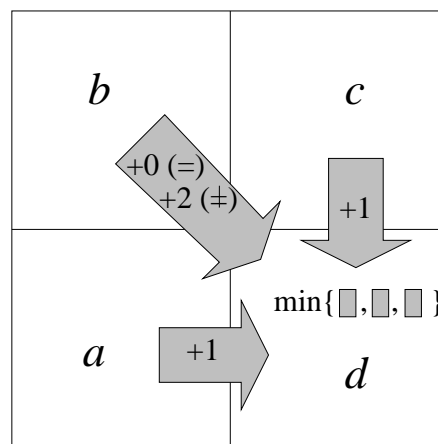
- Suche  $s$  im Baum, beginnend bei der Wurzel:
  - Wenn  $s$  nicht im aktuellen Knoten ist, dann steige jeweils in den passenden Kindknoten ab.
- Wenn  $s$  gefunden wurde:
  - Wenn  $s$  in einem inneren Knoten ist, dann tausche  $s$  mit dem kleinsten Schlüssel im rechten Teilbaum (oder mit dem größten Schlüssel im linken Teilbaum).
  - Lösche  $s$  im Blattknoten.
  - Falls das Blatt weniger als zwei Schlüssel hat (*Unterlauf!*):
    - \* Falls ein Nachbarknoten mehr als zwei Schlüssel hat: Leihe einen Schlüssel von diesem Knoten.
    - \* Falls Nachbarknoten auch nur zwei Schlüssel hat: Verschmelze die zwei Knoten. Bei Unterlauf des Vaterknotens verfare dort analog (Verschmelzen, ggf. bis hinauf zur Wurzel).

6 Punkte

### Aufgabe 4

Betrachten Sie den Editierabstand von Zeichenketten, wenn als Elementaroperationen nur das Zeichen-Einfügen (*insert*) und das Zeichen-Löschen (*delete*) möglich sind.

- a) Mit Hilfe einer  $m \times n$ -Matrix soll der Editierabstand zweier Zeichenketten mit  $m$  bzw.  $n$  Buchstaben bestimmt werden. Geben Sie ein Schema an, nach dem ein Matrixelement aus benachbarten Matrixelementen ermittelt werden kann.



Mit  $a = M_{i,j-1}$ ,  $b = M_{i-1,j-1}$ ,  $c = M_{i-1,j}$  und

$$b' = \begin{cases} b & \Leftarrow \text{zeichen}[i] = \text{zeichen}[j] \\ b + 2 & \Leftarrow \text{sonst} \end{cases}$$

ist  $d = M_{i,j} = \min \{ a + 1, b', c + 1 \}$

(2 Punkte)

b) Bestimmen Sie den Editierabstand zwischen DEBORAH und BARBARA:

		B	A	R	B	A	R	A
	0	1	2	3	4	5	6	7
D	1	2	3	4	5	6	7	8
E	2	3	4	5	6	7	8	9
B	3	<b>2</b>	3	4	<b>5</b>	6	7	8
O	4	3	4	5	6	7	8	9
R	5	4	5	<b>4</b>	5	6	<b>7</b>	8
A	6	5	<b>4</b>	5	6	<b>5</b>	6	<b>7</b>
H	7	6	5	6	7	6	7	8

Das Endergebnis lautet also: 8 Editieroperationen.

(5 Punkte)

c) Beweisen Sie, daß auch nach dieser Definition der Editierabstand eine Metrik auf Zeichenketten ist.

- Es ist  $d(x, y) \geq 0$  [Nichtnegativität]:  
 $d(x, y) > 0 \Leftrightarrow x \neq y$ , da hier editiert werden muß.  
 $d(x, y) = 0 \Leftrightarrow x = y$ , da hier nicht editiert werden muß.
- Es ist  $d(x, y) = d(y, x)$  [Symmetrie]:  
da Löschen und Einfügen gleich teuer sind (vgl. Schema bei (a)).
- Es ist  $d(x, y) + d(y, z) \geq d(x, z)$  [Dreiecksungleichung]:  
Wäre  $d(x, y) + d(y, z)$  kleiner als  $d(x, z)$ , so könnte  $d(x, z)$  nicht der Editierabstand zwischen  $x$  und  $z$  sein, da über den "Umweg"  $y$  das Wort  $x$  schneller in  $z$  zu überführen ist.

(4 Punkte)

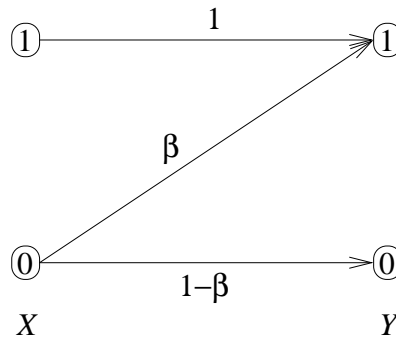
11 Punkte

## Aufgabe 5

Betrachten Sie den folgenden binären asymmetrischen Kanal:

- Eingabe- wie Ausgabealphabet sei  $X = Y = \{0, 1\}$ .
- Die Übertragungswahrscheinlichkeiten  $P(y_i | x_j)$  seien gegeben durch  
 $P(0|0) = 1 - \beta$ ,  $P(1|0) = \beta$ ,  $P(0|1) = 0$ ,  $P(1|1) = 1$ .

a) Für welches  $\beta$  wird die Kanalkapazität maximal? Begründen Sie Ihre Antwort.



$\beta$  ist offenbar eine Wahrscheinlichkeit für Übertragungsfehler. Wenn  $\beta \rightarrow 0$  wird (bzw.  $\beta = 0$  ist), wird ein Maximum an Information übertragen: Kanalkapazität 1 bit pro Zeichen. Der Kanal ist dann deterministisch und störungsfrei (fehlerfrei) und die ganze in  $X$  enthaltene Information kommt ungestört, also identisch, bei  $Y$  an.

(3 Punkte)

b) Für welches  $H(X)$  wird bei gegebener Fehlerwahrscheinlichkeit  $\beta$  der Wert von  $H(Y)$  maximal?

Der Informationsgehalt  $H(Y)$  wird dann maximal, wenn die Fälle “ $Y=0$ ” und “ $Y=1$ ” gleich wahrscheinlich sind, denn es ist ja nach Shannon

$$H(Y) = H(p(Y=0), p(Y=1))$$

Dies ist für  $p(Y=0) = p(Y=1) = \frac{1}{2}$  der Fall.

Die Wahrscheinlichkeiten für dem Empfang  $Y$  berechnen sich aber nach dem obigen Diagramm einfach zu

$$\begin{aligned}
p(Y=0) &= (1-\beta) \cdot p(X=0) \\
p(Y=1) &= p(X=1) + \beta \cdot p(X=0)
\end{aligned}$$

Z. B. nach der ersten Gleichung ergibt sich  $p(X=0)$  als

$$p(X=0) = \frac{p(Y=0)}{1-\beta} = \frac{1/2}{1-\beta} = \frac{1}{2-2\beta}$$

Dann muß die komplementäre Wahrscheinlichkeit  $p(X=1) = 1 - \frac{1}{2-2\beta} = \frac{1-2\beta}{2-2\beta}$  sein, und der Informationsgehalt der Quelle berechnet sich zu

$$\begin{aligned}
H(X) &= H(p(X=0), p(X=1)) = H\left(\frac{1}{2-2\beta}, \frac{1-2\beta}{2-2\beta}\right) \\
&= -\frac{1}{2-2\beta} \log \frac{1}{2-2\beta} - \frac{1-2\beta}{2-2\beta} \log \frac{1-2\beta}{2-2\beta}
\end{aligned}$$

(Dies macht nur für  $\beta \leq 1/2$  Sinn, da sonst keine Quellenverteilung  $P(X)$  die Überlast an Einsen ausgleichen könnte. Diese Fallunterscheidung war aber nicht verlangt.)

(10 Punkte)

13 Punkte

## Aufgabe 6

Kreuzen Sie für folgende Aussagen an, ob diese wahr oder falsch sind.

**Hinweis:** Für jede richtige Antwort gibt es einen Punkt, für jede falsche Antwort wird ein Punkt abgezogen. Die gesamte Aufgabe wird mit einer nichtnegativen Punktezahl gewertet.

<i>Aussage</i>	<i>Antwort</i>
Bei allen Elementaroperationen auf einem AVL-Baum reicht höchstens eine Rotation oder Doppelrotation immer aus, um die AVL-Eigenschaft wieder herzustellen.	<b>falsch</b>
Lösungen linearer Gleichungssysteme lassen sich mithilfe des Gradientenabstiegs beliebig gut approximieren.	<b>wahr</b>
Wenn Transinformation und empfangene Information gleich groß sind, sind keine Übertragungsfehler aufgetreten.	<b>falsch</b>
Für komplexe Einheitswurzeln $v$ und $w$ und natürliche Zahlen $k, m, n$ gilt: $((w^n = 1 \wedge v^m = 1) \Rightarrow \exists k : (v/w)^k = 1)$	<b>wahr</b>
Wenn die Laufzeit $T(P)$ eines Algorithmus durch Umprogrammierung des Algorithmus für alle möglichen Parametersätze $P$ halbiert werden kann, so ist der Algorithmus nicht optimal gewesen.	<b>falsch</b>
Das Problem der Elementeindeutigkeit kann für $n$ -Tupel $(x_1, x_2, \dots, x_n)$ mit im Intervall $[0, 1]$ gleichverteilten Komponenten $x_i$ in der Zeit $T_{\text{mittel}}(n) \in O(n)$ gelöst werden, wenn arithmetische Operationen und Vergleiche sowie die floor-Operation mit dem Zeitaufwand $O(1)$ gewertet werden.	<b>wahr</b>
Wenn der Gradient einer zu optimierenden Funktion mehrerer Variablen die Länge Null hat, so ist ein lokales Minimum oder Maximum erreicht.	<b>falsch</b>
Die Straffunktion kann als Teil der Bewertungsfunktion verstanden werden.	<b>wahr</b>
Die Suche eines Maximums für eine Bewertungsfunktion $BWF(X)$ ist äquivalent zu der Suche eines Minimums für $C - BWF(X)$ , wobei $C$ eine beliebige Konstante ist.	<b>wahr</b>
Bei der Speicherung eines Wörterbuchs inklusive Index in einer großen Datei kann wegen des Hashings nur ca. 75% des Speichers ausgenutzt werden.	<b>falsch</b>
Das klassische Information Retrieval auf der Basis invertierter Dateien hat bei Anfragen im Mittel längere Antwortzeiten als Udi Manbers <i>Glimpse</i> .	<b>falsch</b>
Die schnelle Fouriertransformation (FFT) läßt sich durch parallel arbeitende Hardware so beschleunigen, daß $T_{\text{max}}(n) \in O(\log n)$ gilt.	<b>wahr</b>
Alle bekannten Beweise für Claude Shannons Fundamentalsatz der Informationstheorie sind reine Existenzbeweise; sie erklären nicht, wie man die Codes praktisch konstruieren kann.	<b>falsch</b>
Das Traveling-Salesperson-Problem ( $n$ Städte in der euklidischen Ebene) hat eine untere Schranke von $T_{\text{max}}(n) \in \Omega(n \cdot \log n)$ .	<b>wahr</b>
Je stärker Bilder mit dem JPEG-Verfahren komprimiert werden, umso geringer ist der Rechenaufwand im Zuge der Dekompression.	<b>wahr</b>

15 Punkte