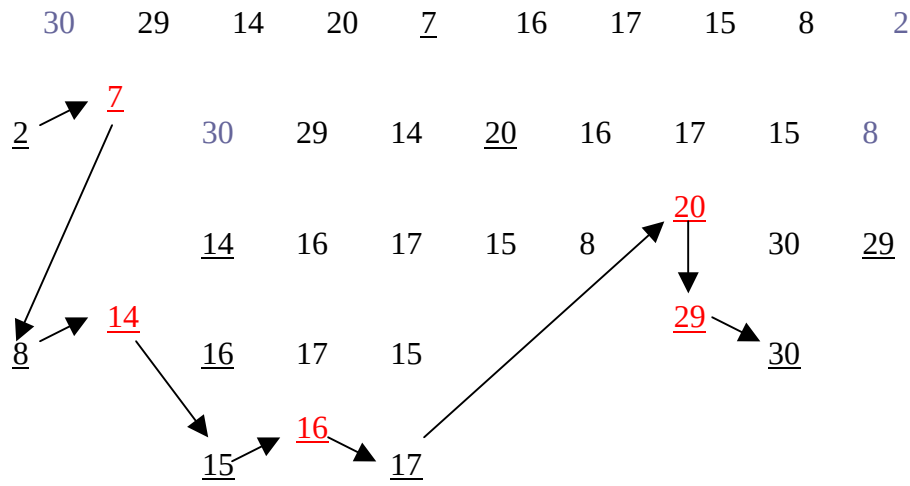
Klausur Datenverarbeitung II (9. Informatik-Lehrgang)

Dozent: Dipl. Ing. Sorber

1) Beschreiben Sie das Verfahren *Quicksort* und realisieren Sie die aufsteigende Sortierung exemplarisch durch die Schlüsselfolge: 30, 29, 14, 20, 7, 16, 17, 15, 8, 2
Verwenden Sie die *3-Median-Strategie* für die Wahl des Pivotelements. **(6P)**

Antwort:
Eine Schlüsselfolge wird mittels Pivot-Elements in 2 Folgen geteilt. Die linke Folge enthält Elemente k < Pivot und die recht k > Pivot. Dies wird solang wiederholt bis nicht mehr geteilt werden kann. Die Pivot-Elemente werden von links nach rechts gesammelt -> sortierte Folge



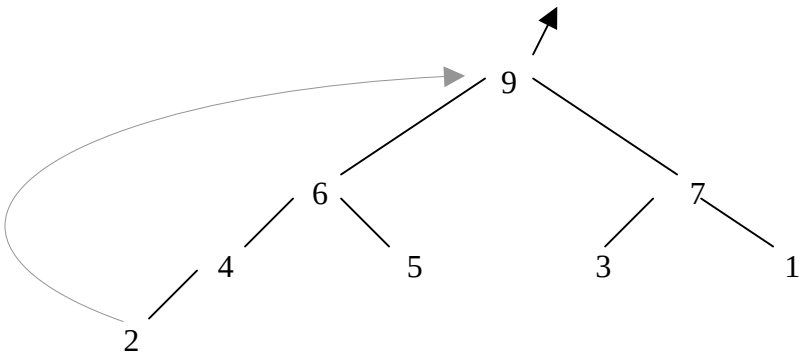
Sortierung: 2 , 7 , 8 , 14 , 15 , 16 , 17 , 20 , 29 , 30

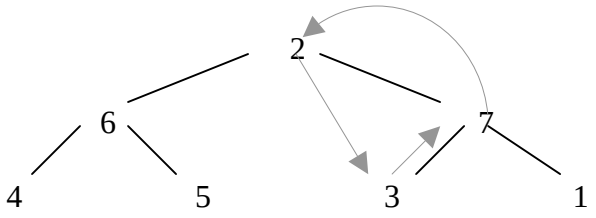
2) Erläutern Sie, unter welcher Bedingung ein Binärbaum ein Heap ist, und demonstrieren Sie die absteigende Sortierung mit Heapsort am Beispiel der Schlüsselfolge:

9 6 7 4 5 3 1 2

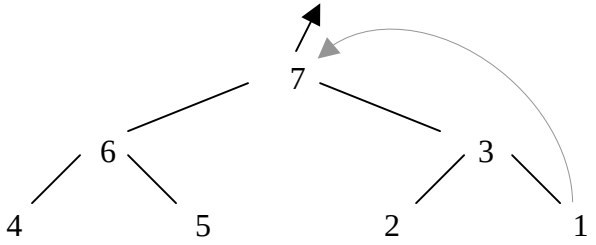
(6P)

Antwort:
Bedingung: Die Nachfolgenden eines Knotens müssen kleiner gleich dem Knoten sein

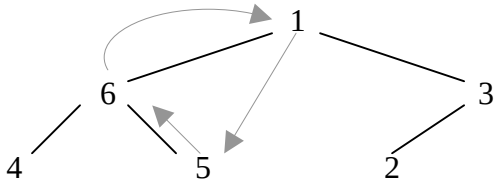


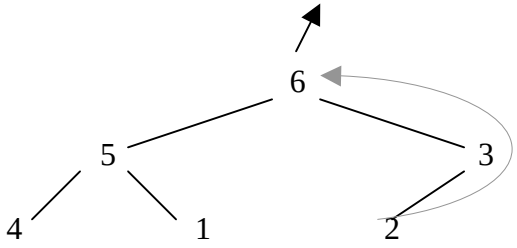


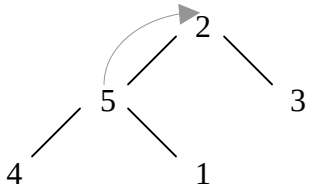
als nächstes wird nach Größe sortiert, so dass alle Nachfolger kleiner sind

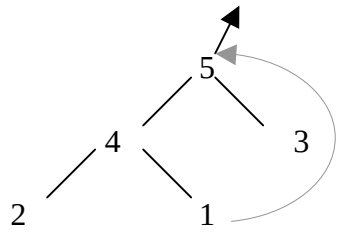


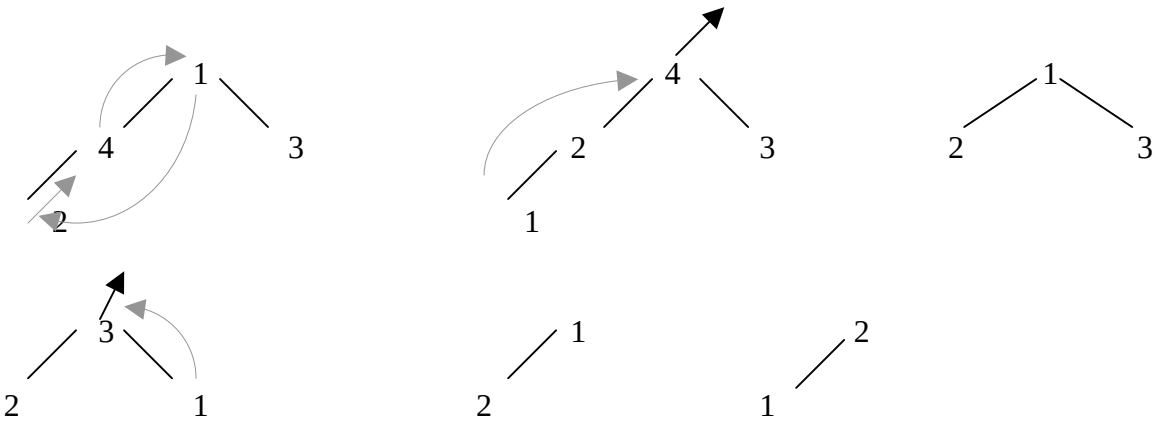
höchster Index (liegt hier bei 1) wird an die Spitze gesetzt (ist somit Wurzel)











Sortierung: 9 7 6 5 4 3 2 1

3) Erklären Sie das Sortierverhalten Binsort (Bucketsort) und sortieren Sie die Schlüsselfolge
28 3 17 8 12 18 20 4 13 38

(6P)

Antwort:
Charakteristisch für das Verfahren ist der Wechsel zwischen Verteilungsphase und Sammelphase.
In der Verteilungsphase werden die Schlüssel auf m Fächer verteilt. D. h. das Fach i nimmt alle Schlüssel, die an der jeweiligen Position t die "Ziffer" i besitzen.
Der jeweils nächste Schlüssel wir im Fach "oben" aufgelegt (auf den bereits vorhanden Schlüssel)
In der Sammelphase werden die Schlüssel aller Fächer m (F0 ... Fm-1) so eingesammelt, dass die Schlüssel (Datensätze) des Fachs i+1 auf die Schlüssel des Fachs i gelegt werden. Anordnung der Schlüssel eines Fachs bleibt dabei unverändert.

Eine Schlüsselfolge wird zuerst nach 1er Stellen in eine Tabelle von 0 – 9 eingetragen. Es folgt eine Sammelphase. Danach wird die neue Folge nach 10er Stellen in eine neue Tabelle eingetragen. Nach erneuter Sammelphase liegt die sortierte Reihenfolge vor.

0	1	2	3	4	5	6	7	8	9
20		12	3	4			17	28	
			13					8	
								18	
								38	

1. Sammelphase: 20, 12, 3, 13, 4, 17, 28, 8, 18, 38

0	1	2	3	4	5	6	7	8	9
3	12	20	38						
4	13	28							
8	17								
	18								

2. Sammelphase (sortierte Folge): 3, 4, 8, 12, 13, 17, 18, 20, 28, 38

4) Erläutern Sie das Prinzip des Hashing **(4P)**

Antwort:

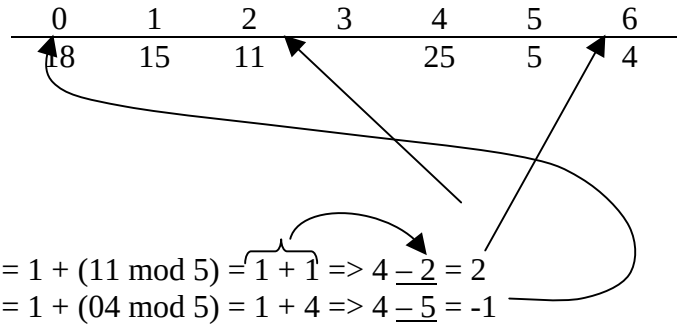
Hashing ist ein Such- und Speicherverfahren
Adressen werden von Datensätzen aus den zugehörigen Schlüsseln berechnet
(aus jedem Schlüssel wird mittels einer Hash-Funktion selbst seine Adresse erzeugt)
Verfahren eignet sich zur Speicherung auf Massenspeichern, wenn Datensätze eingefügt und nicht gelöscht werden
Ziel: möglichst einfache Operationen zu realisieren
Eine gute Hash-Funktion zeichnet aus:

- Sie darf möglichst wenig Kollisionen erzeugen
- Adresskollisionen müssen möglichst effizient aufgelöst ewrden
- sie soll möglichst leicht und schnell berechenbar sein
- sollte die zu speichernden Schlüssel möglichst gleichmäßig auf die Hashtabelle verteilen, um Adresskollisionen zu vermeiden

5) Unter Verwendung von Double Hashing sind die Schlüssel 25, 11, 4, 15, 18, 5 in eine anfangs leere Hashtabelle einzugügen.
Größe der Hashtabelle: m=7
Hash-Funktionen: h(k) = k mod m
h' (k) 0 1+k mod (m-2) **(6P)**

Antwort:

25 mod 7 = 3 Rest 4
11 mod 7 = 1 Rest 4
04 mod 7 = 0 Rest 4
15 mod 7 = 2 Rest 1
18 mod 7 = 2 Rest 4
05 mod 7 = 0 Rest 5



h' (11) = 1 + (11 mod (7-2)) => h' (11) = 1 + (11 mod 5) = 1 + 1 => 4 – 2 = 2
h' (04) = 1 + (04 mod (7-2)) => h' (04) = 1 + (04 mod 5) = 1 + 4 => 4 – 5 = -1
h' (18) = 1 + (18 mod (7-2)) => h' (18) = 1 + (18 mod 5) = 1 + **3** => 4 – 4 = 0

Rest

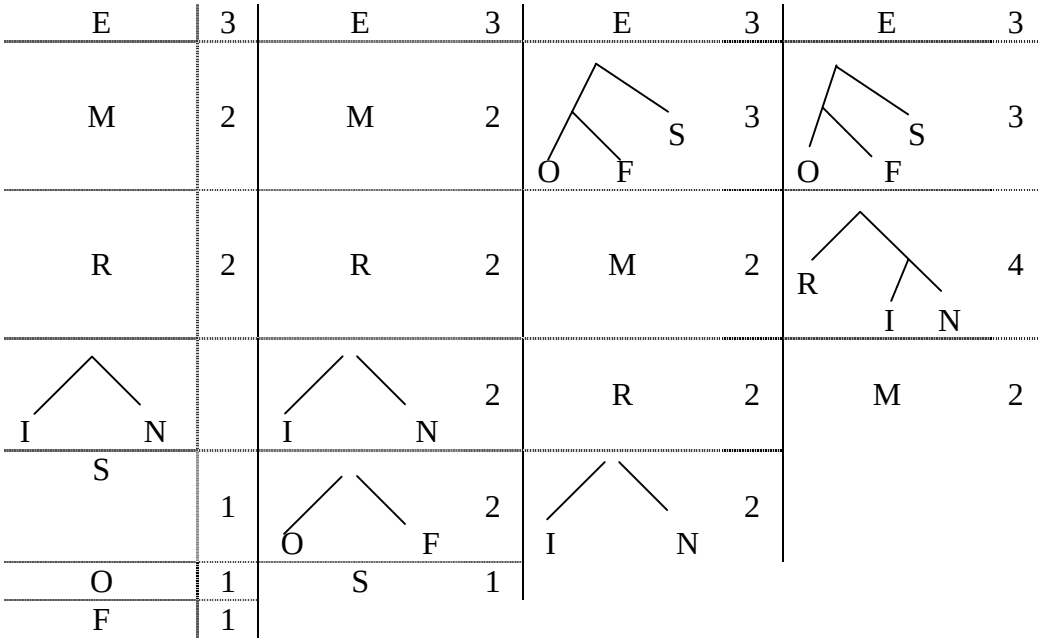
bei k=11: 2 Schritte nach links (auf 2)
bei k=04: 5 Schritte nach links (auf 6)
bei k=18: 4 Schritte nach links (auf 0)

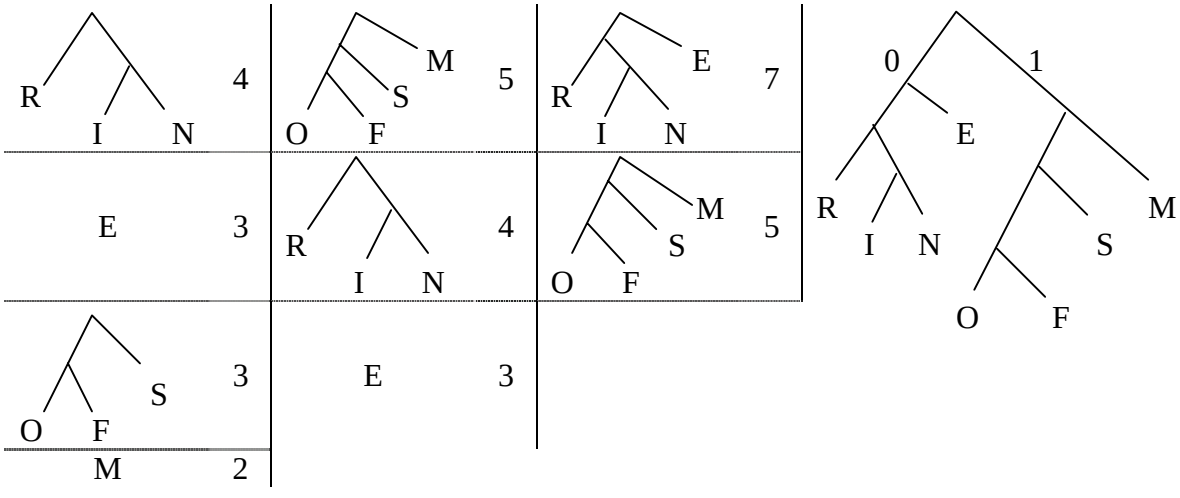
6) Was verstehen Sie unter dem Begriff Huffman-Code?
Demonstrieren Sie seine Erzeugung für eine zu komprimierende Zeichenfolge am Beispiel des Strings "SOMMERFERIEN". **(8P)**

Antwort:

Dient zur Komprimierung von Zeichen/Zeichenketten. Die Häufigkeit jedes Zeichens wird ermittelt und anschließend daraus ein sog. Trie gebildet. Zur Erzeugung des Codes wird der Trie nach links mit 0 durchlaufen und rechts mit 1.

Buchstabe	Häufigkeit	Sortierung nach Häufigkeit	
S	1	E	3
O	1	M	2
M	2	R	2
E	3	S	1
R	2	O	1
F	1	F	1
I	1	I	1
N	1	N	1





Code:

E 0 1	O 1 0 0 0
M 1 1	E 1 0 0 1
R 0 0 0	I 0 0 1 0
S 1 0 1	N 0 0 1 1

7) Erläutern Sie die Bedeutung von Prüfziffern und zeigen Sie eine Möglichkeit zur Bildung von Prüfziffern auf, die eine hohe Sicherheit gewährleistet! Geben Sie ein Beispiel. **(6P)**

Antwort:
Numerische Schlüssel werden durch Prüfziffern auf Konsistenz (Fehler-/Widerspruchsfreiheit) kontrolliert
beim Umgang mit Schlüsseln werden Fehler mehr oder weniger vollständig aufgedeckt
Idee des Prüfziffernverfahrens:
Es wird durch Kodierung der zu erfüllenden Bedingungen eine weitere Ziffer erzeugt (sog. Prüfziffer) und an/in einen Schlüssel an- oder eingefügt

Beispiel:
Zahlenfolge: 2 4 7 9 3
Verfahren: mod 11 -> Gewichtung: 2^{i-1}
Rechnung:

$(2*2^1 + 4*2^2 + 7*2^3 + 9*2^4 + 3*2^5)$	mod 11	= PZ
(4 + 16 + 56 + 144 + 96)	mod 11	= PZ
316	mod 11	= PZ
		= 28 Rest 8
		PZ = 8

8) Beschreiben Sie die prinzipielle Funktionsweise von Kryptosystemen mit öffentlichen Schlüsseln! **(8P)**

Antwort:
Es gibt öffentliche Schlüssel (P), geheime Schlüssel (S), eine zu verschlüsselnde Botschaft (M) sowie den Chiffretext (C)
Für das Kryptosystem müssen folgende Bedingungen gelten:
1. S (P (M)) = M grundlegende Eigenschaft
2. Alle Paare von S und P sind verschieden
3. Ableitung S aus P muss genauso schwer sein, wie Entschlüsseln von C ohne Kenntnis von S
4. P und S müssen sich leicht berechnen lassen

Methode beruht auf Algorithmen mit sehr großen Zahlen.
es werden drei ca. 200-stellige Primzahlen erzeugt (5, x, y)
Bildung von N aus x*y
Bildung von P aus p*s mod (x-1) * (y-1) = 1
es gilt stets: M^{ps} mod N = M

Vorgehensweise:
1. Verschlüsselung der Botschaft M durch Indizes der Buchstaben im Alphabet
2. Verschlüsselung:
   Botschaft M wird in 4-Ziffernlange Teilstücke zerlegt und die p-te Potenz mod N gebildet
3. Zur Entschlüsselung wird s anstelle p verwendet

Jeder Teilnehmer eines Kommunikationsabschnittes erhält einen öffentlich und einen privaten Schlüssel.

Will Teilnehmer A an Teilnehmer B eine Nachricht senden, verschlüsselt er die Nachricht mit dem öffentlichen Schlüssel von B und verschickt diese. Entschlüsselt kann die Nachricht nur mit dem privaten Schlüssel von Teilnehmer A werden.

9) Was verstehen Sie unter dem Begriff der AVL-Ausgeglichenheit von Binärbäumen. Erläutern Sie, warum eine solche Eigenschaft angestrebt werden muss. **(3P)**

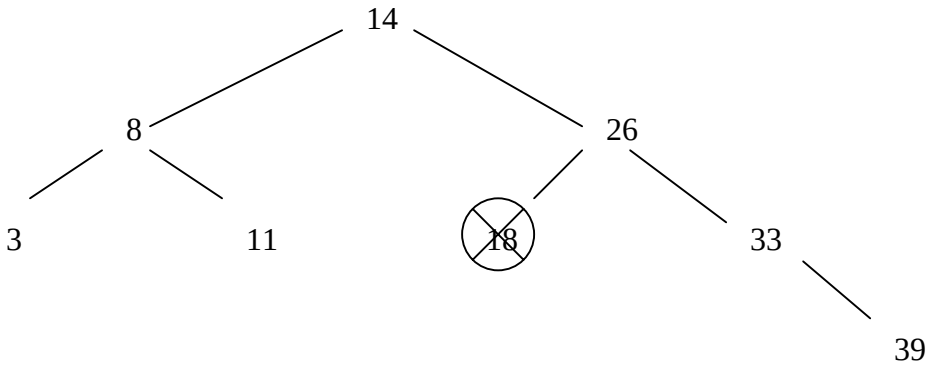
Antwort:

Bei Bäumen existiert eine sog. Balance. Die Tiefe des linken Teilbaumes darf sich höchstens um 1 von der Tiefe des rechten unterscheiden.

Das Suchen, Einfügen und Entfernen eines Knoten in einem zufällig erzeugten Baum mit n Knoten ist immer ausführbar, aber unter Umständen sehr aufwendig. Nämlich dann, wenn der Baum zu einer linearen Liste degeneriert ist.

Durch eine zusätzliche Bedingung an die Struktur eines Baumes, soll so ein Degenerieren verhindert werden -> AVL – Baum (stellt Forderungen an Höhendifferenz beider Teilbäume)

10) Zeigen Sie am folgenden Beispielbaum, wie man nach Löschung des Schlüssels 18 die AVL – Ausgeglichenheit wiederherstellen kann.



Antwort:

Linksrotation des rechten Teilbaums