

## Übungsblatt 5

### Aufgabe 5.1: Offene Hashverfahren

Bestimmen Sie für die folgenden Sondierstrategien jeweils die Belegung einer Hashtabelle der Größe 19, wenn die Schlüsselwerte 17, 1, 41, 22, 26, 16, 31, 42, 2, 24, 7, 6 in dieser Reihenfolge in die anfangs leere Tabelle eingefügt werden und offenes Hashen mit der Funktion  $h(k) = k \bmod 19$  verwendet wird:

- Lineares Sondieren
- Quadratisches Sondieren
- Doppeltes Hashen mit  $h'(k) = 1 + (k \bmod 17)$ .

### Aufgabe 5.2: Quadratisches Sondieren

Es sei eine Hashtabelle  $t$  der Größe 7 mit Belegung  $(t[0], t[1], \dots, t[6]) = (1, 164, 8, 21, 73, 22, 98)$  gegeben. Zum Hashen wurde die Funktion  $h(k) = (\text{Summe der Dezimalziffern von } k) \bmod 7$  gewählt und zur Kollisionsbehandlung wurde quadratisches Sondieren verwendet. Bestimmen Sie alle Reihenfolgen, in denen die Schlüssel in die anfangs leere Hashtabelle eingefügt worden sein könnten.

### Aufgabe 5.3: Hashfunktion

Sei  $W = \bigcup_{3 \leq i \leq 9} A^i$  die Menge der Worte über dem Alphabet  $A = [a..z]$  mit einer Länge zwischen 3 und 9 Buchstaben und  $M$  eine Menge von Worten:

$M = \{\text{jaenner, februar, maerz, april, mai, juni, juli, august, september, oktober, november, dezember}\}$

Gegeben ist die Hashfunktion  $h : W \rightarrow [0, \dots, m-1]$ . Der Hashwert eines Wortes wird berechnet, indem die einzelnen Buchstaben  $[a..z]$  auf die Zahlen  $[0..25]$  abgebildet und aufsummiert werden. Abschließend erfolgt eine Reduktion modulo  $m$ .

- Bestimmen Sie den kleinsten Wert  $m$ , so dass  $h$  bezüglich der Schlüsselmenge  $M$  eine perfekte Hashfunktion darstellt und geben Sie die Hashwerte für alle Worte aus  $M$  an.
- Finden Sie zwei deutschsprachige Worte aus  $W$ , die auf den selben Hashwert abgebildet werden.

### Aufgabe 5.4: Analyse offener Hashverfahren

Zeigen Sie, dass unter den beiden folgenden Bedingungen der durchschnittliche Aufwand (Average-Case) zum Einfügen eines Elementes in die Hashtabelle in  $O(1)$  liegt.

- Die Hashtabelle besteht aus  $n$  Speicherzellen und ist zu weniger als 50% gefüllt (Belegungsfaktor  $\alpha \leq 0.5$ ).
- Die Hashfunktion streut ideal: Für alle Schlüssel  $k$  gilt:  $P[H(k) = i] = 1/n$  (dh. die Wahrscheinlichkeit, dass der Schlüssel  $k$  auf die  $i$ -te Speicherstelle abgebildet wird ist  $1/n$ ).

### Aufgabe 5.5: Suchen von Duplikaten

Entwerfen Sie einen möglichst schnellen Algorithmus (Average-Case sowie Worst-Case), der ein Array auf Duplikate (gleiche Arrayelemente) überprüft.