KIT
Institut für Theoretische Informatik

Jun.-Prof. Henning Meyerhenke, Jun.-Prof. Dennis Hofheinz
Christian Staudt, Christoph Striecks

4. Übungsblatt zu Algorithmen I im SS 2015

https://crypto.iti.kit.edu/algo-sose15
{staudt,striecks}@kit.edu

Aufgabe 1 (*Hashing,* $2+3+1+1$ *Punkte*)

Sei eine anfangs leere Hashtabelle gegeben, die natürliche Zahlen speichert und 17 Einträge besitzt. Weiterhin sei h mit $h_{a,b}(x) = (ax+b) \bmod 17$ eine Hashfunktion, für Konstanten $a, b \in \mathbb{N}$.

a) Nutzen Sie das Hashing mit verketteten Listen, indem Sie jeweils den Hashwert der Elemente

$$24, 35, 41, 31, 9, 27, 15, 45, 10, 17, 21, 4, 18, 13, 47, 55$$

bilden und in die Hashtabelle einfügen. (Dabei seien $a = 2$ und $b = 7$ die Konstanten der Hashfunktion h.) Geben Sie eine Darstellung der Hashtabelle nach dem Einfügen der Hashwerte von 45 und 55 an.

b) Nutzen Sie das Hashing mit linearer Suche (und Puffergröße $m' = 3$), um jeweils die Hashwerte der Elemente aus a) in eine neue anfangs leere Hashtabelle einzutragen. (Dabei seien $a = 2$ und $b = 7$.) Geben Sie eine Darstellung der Hashtabelle nach dem Einfügen der Hashwerte von 45 und 55 an.

c) Wie viele Hashtabelleneinträge werden beim Einfügen der Elemente in a) und b) jeweils betrachtet?

d) Wie groß ist der Speicherverbrauch der Hashtabellen aus a) und b) in Einheiten, wenn jeder Zeiger und jedes Element der Hashtabelle eine Einheit belegen? (Gehen Sie von verketteten Listen aus der Vorlesung aus.)

Aufgabe 2 (*Plagiatsdetektor,* $2+2$)

Ein Paar von Zeichenketten $\{s_1, s_2\}$ heißt 'k-verdächtig', wenn beide Strings einen Substring der Länge $\geq k$ gemeinsam haben.

a) Entwerfen Sie einen effizienten Algorithmus, der als Eingabe eine Menge von n Strings sowie eine positive Ganzzahl k erhält und als Ausgabe alle k-verdächtigen Paare zurückgibt. Nehmen Sie an, dass die meisten Paare nicht k-verdächtig sind. Verwenden Sie für die Implementierung die aus der Vorlesung bekannte Datenstruktur Hashtabelle.

b) Analysieren Sie die Laufzeit Ihres Algorithmus in Abhängigkeit von n und k. Sie können voraussetzen, dass sich ein String der Länge k in Zeit $O(k)$ hashen lässt.

Aufgabe 3 (*Wahrscheinlichkeitstheorie,* $1+1+1+2$ (Zusatzpunkte))

Der ebenso geniale wie einfallsreiche Wissenschaftler und Superbösewicht Doktor Meta ist geschäftig. Da er sich auf seine menschlichen Gehilfen nicht verlassen kann, hat er beschlossen, einen willenlosen Automaten in Form eines mechanischen Arbeiters herzustellen. Von Donnerschlägen begleitet beginnt er die Arbeit an dieser neuartigen Erfindung, die er Ro-bo-ter nennt. Der Bausatz eines schwedischen Möbelhauses (røbøt) enthält neben einem Inbusschlüssel einen jeweils unerschöpflichen Vorrat an 426 unterschiedlichen Komponenten. Beim Zusammenbau muss Doktor Meta darauf achten, nicht versehentlich zweimal das gleiche Teil zu verbauen, da dies zu einem unbrauchbaren Gehilfen führen würde. Er geht dabei folgendermaßen vor: Er zieht mit Zurücklegen zufällig gleichverteilt n Zahlen aus der Menge $\{1, \ldots, 426\}$. Sei $x = (x_1, x_2, ..., x_n)$ die Folge der gezogenen Zahlen.

a) Wie viele paarweise verschiedene Folgen x kann es geben?

b) Wie viele paarweise verschiedene Folgen x kann es geben, sodass für jede Folge gilt: $\forall i, j \in \{1, \ldots, n\}, i \neq j : x_i \neq x_j$?

c) Berechnen Sie die relative Häufigkeit für die Folgen aus b).

d) Zusatzaufgabe. Wie groß darf n maximal gewählt werden, sodass die relative Häufigkeit für die Folgen aus b) größer oder gleich 0.5 beträgt?

Hinweis: Vereinfachen Sie zunächst Ihr Ergebnis aus c) mithilfe der Näherung für große Zahlen $n \in \mathbb{N}$: $n! \approx \sqrt{2\pi n}(n/e)^n$. Versuchen Sie nicht, Ihre Formel nach n aufzulösen, eine numerische Lösung genügt.

Ausgabe: Mittwoch, 06.05.2014
Abgabe: Freitag, 15.05.2014, 12:45 im Briefkasten im Untergeschoss von Gebäude 50.34