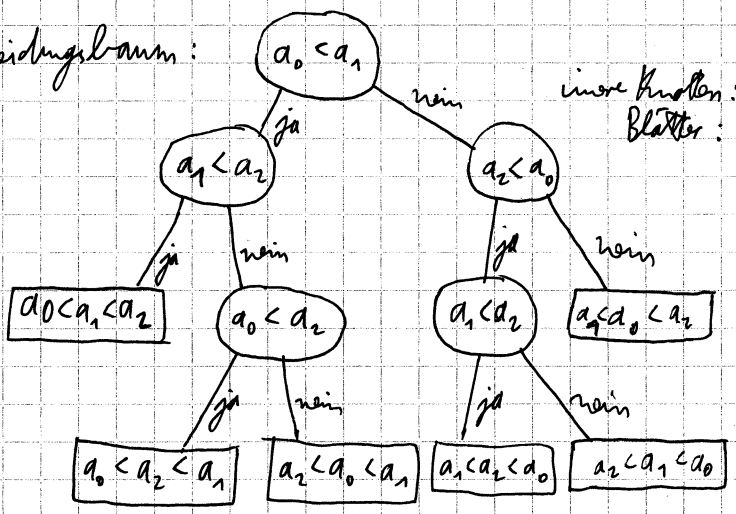
2012-12-05
DSEA-1

Informationstheoretische Untere Schranke für vergleichsbasiertes Sortieren

$a[0], a[1], a[2]$ n = drei Elemente $\rightarrow n! = 6$ Permutationen $a_i \neq a_j$ für $i \neq j$

Entscheidungsbaum:



innere Knoten: Entscheidungen
Blätter: alle $n!$ Permutationen

Worst case ~ maximale Tiefe des Entscheidungsbaumes
~~Mittlere Laufzeit~~ ~ mittlere Tiefe des Entscheidungsbaumes

Blätter im Entscheidungsbaum $\geq n!$ (wenn keine redundanten Entscheidungen =)

Sei t die maximale Tiefe.

$\Rightarrow 2^{t_{max}} \geq n!$

$\Rightarrow t_{max} \geq \log_2 n!$

Betrachte $\left(\frac{2}{1}\right)^1 \left(\frac{3}{2}\right)^2 \left(\frac{4}{3}\right)^3 \left(\frac{5}{4}\right)^4 \cdots \underbrace{\left(\frac{n+1}{n}\right)^n}_{= \left(1+\frac{1}{n}\right)^n \leq e \atop \xrightarrow{n \to \infty} e}$

$= \frac{(n+1)^n}{n!}$

$\leq e^n$

$\Rightarrow n! \geq \left(\frac{n+1}{e}\right)^n$

$\Rightarrow t_{max} \geq \log_2 n! = \Omega(n \log n)$

$$\ln n! = \ln 1 \cdot 2 \cdot 3 \cdots n$$
$$= \sum_{i=1}^{n} \ln i$$
$$\approx \int_1^n \ln x \, dx$$
$$= \left[x \ln x - x \right]_1^n$$
$$= n \ln n - n + 1$$

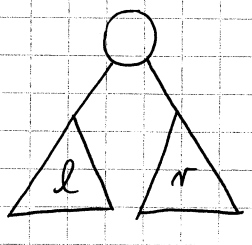
Stirling-Formel

$$\lim_{n \to \infty} \frac{n!}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} = 1$$

2012-12-05

DSEA. 2

Sei $\bar{T}$ die mittlere Tiefe der Blätter im Entscheidungsbaum



$l + r = m = n!$ Anzahl der Blätter, nicht Knoten.

$m \cdot \bar{T}(m) =$ Summe aller Tiefen, da $\bar{T} = \frac{\Sigma \ldots}{m}$

$= l \cdot \left(\bar{T}(l) + 1\right) + r \cdot \left(\bar{T}(r) + 1\right)$

$\bar{T}(m) = \frac{l}{m}\bar{T}(l) + \frac{r}{m}\bar{T}(r) + 1$ wird minimal für $l = r = \frac{m}{2}$

$\vdots$

$\bar{T}(m) \geq \bar{T}\left(\frac{m}{2}\right) + 1 \Rightarrow \bar{T}(m) \geq \log_2 m$

$\Rightarrow \bar{T}(n!) = \Omega(n \log n)$

ÜB ← ausrechnen

Induktion

Ind.: $\bar{T}(k) = \ln k$

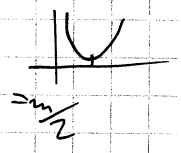
$\bar{T}(m) = \frac{l}{m} \ln l + \frac{r}{m} \cdot \ln r + 1$

$\left(= \ldots \overset{!}{=} \ln m\right)$

$\geq \min_{l+r=m} \frac{1}{m}\left(l \ln l + r \ln r\right) + 1$

$\ldots$ ÜB: $f(x) = x \ln x + (m-x)\ln(m-x)$

$\min f(x)$

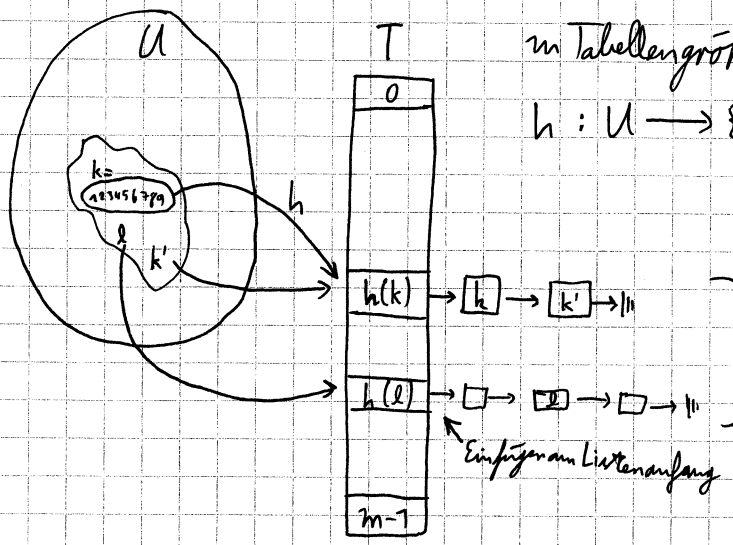


$= m/2$

2012-12-05

DSEA. 3

Hashing



m Tabellengröße, $m \ll |U|$, $m \sim n =: \#$ verwaltete Elemente

$h : U \longrightarrow \{0, 1, 2, \dots, m-1\}$ Hashfunktion

ist nicht injektiv,

Hashkollisionen möglich.

} Hashing mit Verkettung.

Verwaltung einer "unsortierten" Menge

- search
- insert
- delete

Belegungsfaktor der Hashtabelle $\alpha := \frac{n}{m}$

Idealisierte Annahme: Hashfunktion h verteilt die n Elemente zufällig gleichverteilt auf die m Tabelleneinträge.

n_i sei die erwartete Anzahl von Elementen, die auf Tabelleneintrag i gehasht werden.

$n_i := \alpha = \frac{n}{m} \Rightarrow$ Suchzeit $= O(\alpha + 1)$ unter der Voraussetzung, dass sich die Hashfkt in konstanter Zeit auswerten lässt.

Beispiel für Hashfunktion:

$h(k) = (a \cdot k + b) \bmod m$, a, b konst $< m$ ist eine gute Wahl, insbesondere, wenn m eine Primzahl ist.