Informatik II: Algorithmen und Datenstrukturen SS 2017

Vorlesung 3b, Dienstag, 10. Mai 2017
(O-Notation, Teil 2)

Prof. Dr. Hannah Bast
Lehrstuhl für Algorithmen und Datenstrukturen
Institut für Informatik
Universität Freiburg

Blick über die Vorlesung heute

- Drumherum
  - Klausurtermin — **29. August** … 14 – 17 Uhr
  - Astrologie — Was ist dran?
  - Fragestunde — oder auch nur halbe Stunde
- Inhalt
  - O-Notation und Grenzwerte — Bestimmung via $\lim_{n \to \infty}$
  - Diskussion — Sinn und Grenzen, mehrere Variablen

Klausurtermin

Dienstag

- Der Termin steht jetzt fest
  - Es ist der **29. August 2017**, von 14 Uhr bis max. 17 Uhr
  - Es wird geschrieben im **HS 026 + 036** und bei Bedarf auch im SR 00 10/14 und SR 01 09/13 (alles hier in 101)
  - Zur Aufteilung sagen wir dann noch rechtzeitig was, siehe die Seite zur Klausur auf dem Wiki der Veranstaltung
  - Das Horoskop für den Termin: Sonne in Jungfrau im 9. Haus, Aszendent in Skorpion, Mond in Schütze im 1. Haus

    "Fleiß + nicht zur Ruhe kommen, bis das Soll erreicht ist"

    Achtung: Aszendent wechselt kurz vor 15 Uhr zu Schütze !

    "Suche nach Wahrheit + von den eigenen Ideen überzeugt"

Was ist dran an Astrologie 1/4

- Auszug aus Ihren Rückmeldungen
  - "Da muss ich erst mal meine Glaskugel befragen"
  - "Astrologie ist wahr: meine Katze ist Sternzeichen Fisch und sie isst Fisch auch sehr gerne (Aszendent Löwe)"
  - "Ich glaube nicht an Astrologie, wir Widder sind misstrauisch!"
  - "Die Sterne versprechen mir volle Punktzahl für dieses Blatt"
  - "Gutes Gesprächsthema, wenn man Gothic-Mädchen daten will"
  - "Was ist dran an der Realität?"
  - "Laut Horoskop bin ich als Skorpion bekannt für meine sexuelle Ausdauer. Das stimmt eindeutig nicht ..."
  - The positions of the stars and planets will ...

Was ist dran an Astrologie 2/4

- Studie aus dem Jahr 1968 ... von Michel Gauquelin
  - 150 Personen (zufällig ausgewählt über ein Zeitungsinserat) bekommen ihr ganz persönliches Horoskop
  - Sie werden gefragt, wie sehr sie sich darin wiedererkennen

    94% sagen: ja, es passt; 90% sagen: sogar **sehr** passend

    Alle Personen erhielten denselben Text, erstellt von einem Profi-Astrologen für eine (ihm nicht bekannte) Person X

    Die Person X war der Serienmörder Marcel Petiot
  - Das ist der sogenannte "Barnum-Effekt":

    Die Neigung vage bzw. allgemeingültige Aussagen über die eigene Person als zutreffend zu interpretieren

Was ist dran an Astrologie 3/4

- Weitere Studien
  - S. Carlson: A Double-Blind Test in Astrology, Nature 1985

    Detaillierte Persönlichkeitsprofile von 128 Testpersonen

    Ein Team von 26 Profi-Astrologen, die von jedem der 128 nur das Geburtsdatum (und damit das Horoskop) kannten

    Aufgabe für die Astrologen: die Personen auf der Grundlage der Horoskope den Persönlichkeitsprofilen zuordnen

    Das Design des Experimentes wurde vorher im Detail mit dem Team von Astrologen abgestimmt

    Ergebnis: die "Trefferquote" war nicht signifikant höher als bei einer zufälligen Zuordnung

Was ist dran an Astrologie 4/4

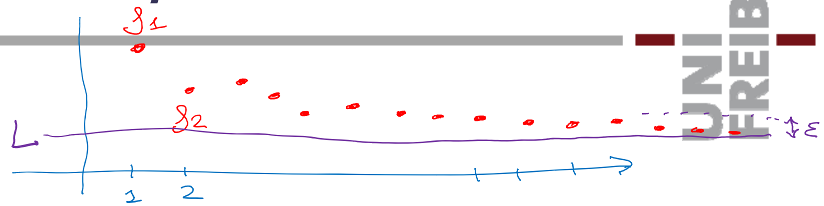
- Pro Astrologie
  - Die ersten Lebensmonate sind sehr prägend, insbesondere für die Gehirnentwicklung
  - Die Geburtsdaten und der Ort sind ein Hinweis auf die zu dieser Zeit vorherrschenden Witterungsverhältnisse
  - Man kann sich schon vorstellen, dass das einen Einfluss hat
  - Allerdings: Persönlichkeitsprofile von Menschen können sich über die Spanne ihres Lebens **sehr** stark verändern

    Personality Stability from Age 14 to Age 77, Psych & Aging '16

    174 Testpersonen, 6 Persönlichkeitsmerkmale, starke Veränderung über einen Zeitraum von 63 Lebensjahren

    etwas Stabilität nur bei: "Stetigkeit" und "Gewissenhaftigkeit"

O-Notation – Grenzwerte 1/7



- Grenzwertbegriff
  - Die Definitionen von der Vorlesung 3a erinnern sehr stark an den **Grenzwertbegriff** aus der **Analysis**
  - **Definition:** Eine unendliche Folge $f_1, f_2, f_3, \ldots$ hat einen Grenzwert L, wenn für alle $\varepsilon > 0$ ein $n_0 \in \mathbf{N}$ existiert so dass für alle $n \geq n_0$ gilt dass $|f_n - L| \leq \varepsilon$
  - In Symbolen schreibt man dann $\lim_{n\to\infty} f_n = L$
  - Eine Funktion $f : \mathbf{N} \to \mathbf{R}$ kann man genauso gut als Folge $f(1), f(2), f(3), \ldots$ auffassen und schreibt $\lim_{n\to\infty} f(n) = L$

O-Notation – Grenzwerte 2/7

- Beispiel für einen Beweis von einem Grenzwert
  (sollten Sie eigentlich in Mathe 1 schon mal gesehen haben)
  - Es gilt: $\lim_{n\to\infty} 1/n = \overset{=L}{0}$

zu zeigen: $\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 \quad \underbrace{\left| \frac{1}{n} - \overset{=L}{0} \right|}_{=\frac{1}{n}} \leq \varepsilon$

Beweis: sei $\varepsilon > 0$ beliebig

Finde ein $n_0 \in \mathbb{N}$: $\forall n \geq n_0 \quad \frac{1}{n} \leq \varepsilon$

z.B. $\varepsilon = \frac{1}{1000}$ $\quad n_0 = 1000$? $\quad n \geq n_0 \Rightarrow \frac{1}{n} \leq \frac{1}{n_0} \leq \frac{1}{1000}$

Allgemein $n_0 = \lceil 1/\varepsilon \rceil$? $\quad n \geq n_0 \Rightarrow \frac{1}{n} \overset{n \geq n_0}{\leq} \frac{1}{n_0} \leq \frac{1}{\underbrace{\lceil \frac{1}{\varepsilon} \rceil}_{\geq \frac{1}{\varepsilon}}} \leq \varepsilon$

$\leq \frac{1}{\frac{1}{\varepsilon}}$ $\blacksquare$

O-Notation – Grenzwerte 3/7

- Satz
  - Seien $f, g : \mathbf{N} \to \mathbf{R}$ und der Grenzwert $\lim_{n\to\infty} f(n)/g(n)$ existiert (evtl. ist er ∞) … dann gelten:

    (1) $f = O(g) \Leftrightarrow \lim_{n\to\infty} f(n)/g(n) < \infty$

    (2) $f = \Omega(g) \Leftrightarrow \lim_{n\to\infty} f(n)/g(n) > 0$

    (3) $f = \Theta(g) \Leftrightarrow \lim_{n\to\infty} f(n)/g(n) > 0$ und $< \infty$

    (4) $f = o(g) \Leftrightarrow \lim_{n\to\infty} f(n)/g(n) = 0$

    (5) $f = \omega(g) \Leftrightarrow \lim_{n\to\infty} f(n)/g(n) = \infty$
  - Wir beweisen auf der nächsten Folie Aussage (1)

    Die Beweise für die anderen Aussagen gehen analog

O-Notation – Grenzwerte 4/7

- Beweis von: $f = O(g) \Leftrightarrow \lim_{n\to\infty} f(n)/g(n) < \infty$

Beweis von „$\Rightarrow$" (Hinrichtung):

$f = O(g) \Rightarrow \exists C>0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 :\ \underbrace{f(n) \leq C \cdot g(n)}_{\Leftrightarrow\ f(n)/g(n) \leq C}$

Annahme: $\lim\limits_{n\to\infty} \frac{f(n)}{g(n)} = \infty$

$\Rightarrow \forall D>0\ \exists n_1 \in \mathbb{N}\ \forall n \geq n_1 :\ \frac{f(n)}{g(n)} \geq D$

insbesondere $\forall n \geq n_1 :\ \frac{f(n)}{g(n)} \geq C+1$

Wir haben also: $\forall n \geq n_0 :\ \frac{f(n)}{g(n)} \leq C$; $\forall n \geq n_1 :\ \frac{f(n)}{g(n)} \geq C+1$ ↯

Beweis von „$\Leftarrow$" (Rückrichtung):

$\lim\limits_{n\to\infty} f(n)/g(n) = C < \infty \Rightarrow \forall \varepsilon>0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 :\ \frac{f(n)}{g(n)} \leq C+\varepsilon$

zum Beispiel für $\varepsilon = 1$: $\exists n_0 \in \mathbb{N}\ \forall n \geq n_0 :\ f(n) \leq (C+1) \cdot g(n)$

$\xrightarrow[\text{Def. von } O(\ldots)]{\text{nach}}\ f = O(g)$ ∎

O-Notation – Grenzwerte 5/7

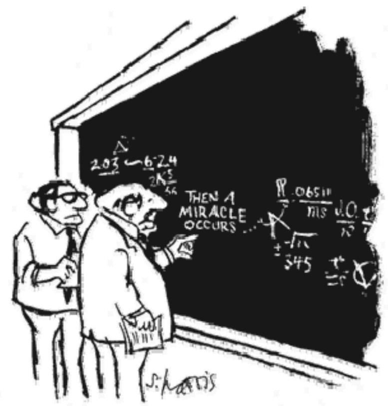
$1' = 0$
$n' = 1$

$\lim_{n \to \infty} \frac{1}{n} = \lim_{n \to \infty} \frac{0}{1} = 0$

- Variante 1: "zu Fuß"
  - Dafür hatten wir gerade das Beispiel

    $\lim_{n \to \infty} 1/n = 0$
- Variante 2: Regel von L'Hôpital
  - Seien $f, g : \mathbf{N} \to \mathbf{R}$ wie gehabt
  - Es existieren die ersten Ableitungen f' und g', sowie der Grenzwert $\lim_{n \to \infty} f'(n)/g'(n)$ ... dann gilt

    $\lim_{n \to \infty} f(n)/g(n) = \lim_{n \to \infty} f'(n)/g'(n)$
- Variante 3: "sieht man doch"
  - Erst mit Professur erlaubt …



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."

O-Notation – Grenzwerte 6/7

$\ln n = \log_e n$

$e = 2.71828\ldots$

- Beispiel: Grenzwert mit L'Hôpital
  - Was ist $\lim_{n \to \infty} (\ln n)/n$?

$\lim_{n \to \infty} \ln n = \infty$

$\lim_{n \to \infty} n = \infty$

$\frac{\infty}{\infty} \to ?$

SEMIKOLON

$f(n) = \ln n \;;\; f'(n) = \frac{1}{n}$

$g(n) = n \;;\; g'(n) = 1$

$$\lim_{n \to \infty} \frac{\ln n}{n} \overset{\text{L'Hôpital}}{=} \lim_{n \to \infty} \frac{1/n}{1} = \lim_{n \to \infty} \frac{1}{n} = 0$$

Daraus folgt insbesondere $\ln n = O(n)$

aber nicht $\ln n = \Theta(n)$

O-Notation – Grenzwerte 7/7

- Was darf man ohne Beweis annehmen?
  - Gute Frage !!! Da gibt es keine klare Regel

    Im Zweifelsfall immer mehr beweisen als weniger
  - **Beispiel 1:** $\lim_{n\to\infty} 1/n = 0$

    Brauchen Sie nicht mehr weiter zu beweisen
  - **Beispiel 2:** $\lim_{n\to\infty} 1/n^2 = 0$ $\lim\limits_{n\to\infty} \frac{1}{n^2} = \underbrace{\left(\lim\limits_{n\to\infty} \frac{1}{n}\right)^2}_{=0} = 0$

    Kann man leicht auf Beispiel 1 zurückführen
  - **Beispiel 3:** $\lim_{n\to\infty} (\log n)/n = 0$

    Das sollte man beweisen, zum Beispiel mit L'Hôpital

O-Notation – Diskussion 1/3

- Asymptotische Analyse
  - Die O-Notation schaut sich das Verhalten der Funktionen an, wenn $n \to \infty$ geht (es interessieren nur die $n \geq n_0$)
  - Wenn man Laufzeiten o.ä. als O(...) etc. ausdrückt, spricht man daher von **asymptotischer Analyse**
  - Vorsicht: asymptotische Analyse sagt nichts über das Verhalten bei "kleinen" Eingabegrößen ($n < n_0$) aus
  - Für $n < 2$ oder $n < 10$ ist das egal, da wird schon nichts Schlimmes passieren
  - Aber das n_0 ist nicht immer so klein ... siehe nächste Folie

O-Notation – Diskussion 2/3

irgendwann ist der lineare Alg. zeitlich besser

- Beispiel wo das n_0 nicht ganz so klein ist
  - Algorithmus A hat Laufzeit $f(n) = 80 \cdot n$ „linear" $\Theta(n)$
  - Algorithmus B hat Laufzeit $g(n) = 2 \cdot n \cdot \log_2 n$ $\Theta(n \cdot \log n)$
  - Dann ist $f = O(g)$ und sogar $f = o(g)$ aber nicht $g = \Theta(g)$

    Insbesondere für alle $n \geq$ irgendein n_0 : $f(n) \leq g(n)$

    Das heißt, A ist asymptotisch echt schneller als B
  - Allerdings:

$$n_0 = 2^{40} = (\underbrace{2^{10}}_{\approx 1000})^4 \approx 1000^4 = 10^{12} = 1 \text{ Billionen}$$ „TERA"

$$n < 2^{40}: \quad g(n) = 2 \cdot n \cdot \underbrace{\log_2 n}_{< \log_2 2^{40} = 40} < 80 \cdot n = f(n)$$

erst als $n \geq 2^{40}$ ist $g(n) \geq f(n)$

O-Notation – Diskussion 3/3

auch leicht auf $\mathbb{N}^3$, $\mathbb{N}^4$, ... verallgemeinerbar

- **Mehrere Variablen**
  - Es kommt öfter mal vor, dass die Laufzeit (oder eine andere Größe) von mehr als einer Variablen abhängt
  - Zum Beispiel von der Eingabegröße n und der Anzahl m der verschiedenen Elemente in der Eingabe
  - Dann würden wir auch gerne sowas schreiben wie

    $O(n + m \cdot \log m)$

    $\mathbb{N} \times \mathbb{N}$
  - Wir können die Definitionen aus Vorlesung 3a leicht auf Funktionen $\mathbf{N^2} \to \mathbf{R}$ verallgemeinern, zum Beispiel:

    $$O(f) = \{\, g : \mathbf{N^2} \to \mathbf{R} \mid \exists n_0 \in \mathbf{N} \;\; \exists C > 0 \;\; \forall n, m \geq n_0 \;\; g(n, m) \leq C \cdot f(n, m) \,\}$$

Frageviertelstunde

- Fragen Sie !

Sie dürfen verwenden: $\log_x y = \frac{\ln y}{\ln x} = \frac{\log_2 y}{\log_2 x}$

Literatur / Links

- O-Notation / Ω-Notation / Θ-Notation
  - In Mehlhorn/Sanders:

    2.1 Asymptotic Notation
  - In Wikipedia

    http://en.wikipedia.org/wiki/Big_O_notation

    http://de.wikipedia.org/wiki/Landau-Symbole