University of Applied Sciences
HOCHSCHULE EMDEN·LEER

Theoretische Informatik Einführung

Prof. Dr. Juho Mäkiö

Zuordnung Informatik

- Informatik
  - Technische Informatik (Bindeglied zur Elektrotechnik)
    - Rechnerarchitekturen
    - Hardware
    - Schaltwerke
    - …
  - Praktische Informatik (Konkrete Probleme + deren Umsetzung)
    - SW-Technik
    - Datenstrukturen
    - Algorithmen
    - Datenbanken
    - …
  - Angewandte Informatik
    - Bio-Informatik
    - Med. Informatik
    - …
  - Theoretische Informatik (Theoretisches Fundament)
    - Berechenbarkeit
    - Sprache
    - …

Theoretische Informatik…

- beschäftigt sich mit der Abstraktion, Modellbildung im Zusammenhang mit Problemen, die in einer Verbindung mit Computern sind.
  - Fragestellungen, die mit der Struktur, Verarbeitung, Übertragung und Wiedergabe von Informationen im Zusammenhang stehen (s. Wikipedia)
- ist das älteste Teilgebiet der Informatik – grundlegenden Erkenntnissen wurden vor dem ersten Computer entdeckt.
- verwendet Elemente der
  - Mathematik / Logik,
  - Linguistik und
  - Philosophie.
- ist eine Strukturwissenschaft.
  (Mathematik, Informatik, allgemeine Systemtheorie)
  - Algorithmus, Berechenbarkeit, Datenstruktur, von-Neumann-Architektur

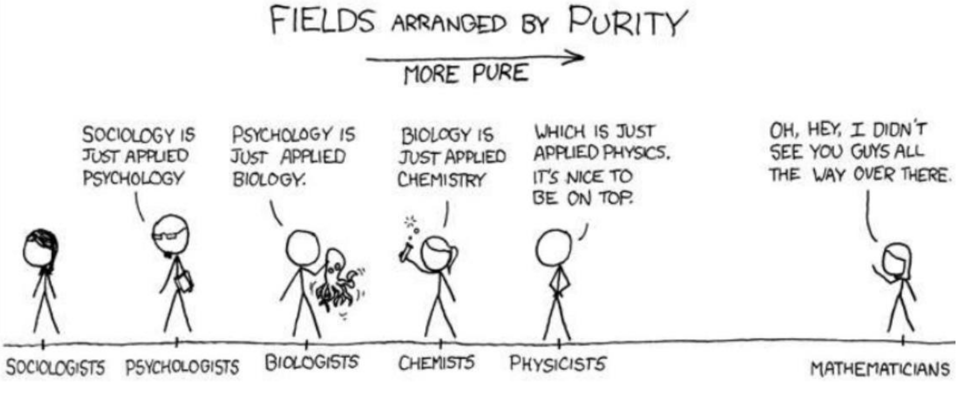
Theoretische Informatik - Themen

- Endliche Automaten
  - DEA,
  - NEA und
  - NEA mit epsilon-Übergängen),
- Kellerautomaten,
- reguläre Ausdrücke,
- Transformationen und Minimierung
  - NEA nach DEA und
  - NEA/eps nach NEA
- regulärer Ausdruck nach NEA/eps),
- reguläre und nicht-reguläre Sprachen,
- Grammatiken und kontextfreie Sprachen

Strukturwissenschaft

- Betrachtet allgemein funktional wirksame Formen
  - keine Allgemeinen oder speziellen Gegenstände der Natur oder der sozialen Wirklichkeit.
- „… nicht nur die reine und angewandte Mathematik …, sondern das in seiner Gliederung noch nicht voll durchschaute Gebiet der Wissenschaften, … Systemanalyse, Informationstheorie, Kybernetik, Spieltheorie … . Sie sind gleichsam die Mathematik zeitlicher Vorgänge, die durch menschliche Entscheidung, durch Planung, durch Strukturen, […] oder schließlich durch Zufall gesteuert werden. Sie sind also Strukturtheorien zeitlicher Veränderung. Ihr wichtigstes praktisches Hilfsmittel ist der Computer, dessen Theorie selbst eine der Strukturwissenschaften ist. Wer in einem Lande den Fortschritt der Wissenschaft fördern will, muss diese Wissenschaften vordringlich fördern, denn sie bezeichnen gleichsam eine neue Bewusstseinsstufe.“

C. F. v. Weizsäcker: Die Einheit der Natur



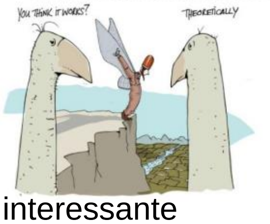
Warum beschäftigen wir uns mit diesem Fach?

- Theoretische Untersuchungen wenig „sinnvoll“, solange Dinge durch blosses Anschauen verstanden werden!
  - Bei komplexeren Zusammenhängen reicht das „blosses Anschauen“ nicht aus!
- Um zuverlässige Aussagen zu erhalten, werden Modelle entwickelt.
  - Modell: eine zutreffende Beschreibung von wesentlichen Aspekten der Wirklichkeit
  - Formale Modelle werden mit der Sprache der Mathematik entwickelt, beschrieben und untersucht.
  - Theorie – eine System von Aussagen, um die Realität zu beschreiben, zu erklären und zu prognostizieren
- Theoretische Informatik befasst sich mit Theorien zur Untersuchung von Phänomenen, die in der Informatik auftreten

Zweck der Theorien in der Informatik

- Theorien dienen der Beschreibung, der Klassifizierung und dem Ordnen des Wissens in einfacher und einheitlicher Weise.
- Theorien helfen Zusammenhänge
  - zu erkennen,
  - zu verstehen und
  - zu bewerten.
- Nützlich insbesondere dann, wenn sie interessante Aussagen über die Wirklichkeit (für die Praxis) liefert, welche einer direkten Einsicht nicht mehr zugänglich sind.

Zwischen Theorie und Praxis



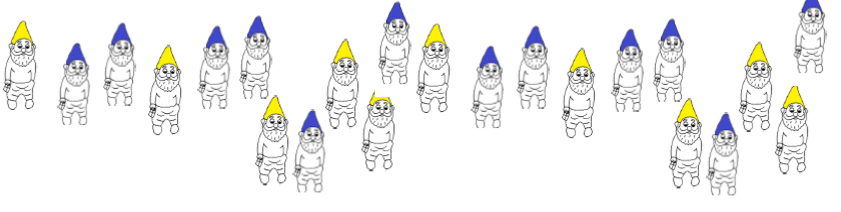
Teilgebiete der theoretischen Informatik

- Berechenbarkeitstheorie
  - Abgrenzung von berechenbaren von unberechenbaren durch Benennung von Problemen, die ein Computer nicht lösen kann
  - Ziel: Nachweis, dass es bestimmmte (wünschenwerte) Algorithmen nicht gibt
- Komplexitätstheorie
  - Untersucht den rechnerischen Aufwand (Zeit, Raum) von Algorithmen
- Automatentheorie
  - Welche einfachen mathematischen Modelle dem Computer zugrunde liegen ← Theorie der endlichen Automaten
- Formale Sprachen
  - Strukturelle Aufbau von Programmiersprachen
    - Kontextfreie Sprachen
    - Kellerautomaten

BERECHENBARKEITSTHEORIE

Aufgabe zum Spaß

- Entwickeln Sie einen Algorithmus für das folgende Problem:
  - Eine Gruppe von Zwergen mit gelben und blauen Mützen sollte nach der Farbe sortiert aufstellen. Es gelten die Annahmen:
    - Kein Zwerg kennt die Farbe der eigenen Mütze nicht.
    - Die Zwerge können nicht kommunizieren
    - Die Zwerge kennen die Farbe der anderen Mützen.



Berechnebarkeitstheorie

- Beschäftigt sich mit den grundsätzlichen Grenzen und Möglichkeiten der Berechenbarkeit
- Aufgabe der BT: Festlegung einer formalen Begriffs für die Berechenbarkeit und Algorithmen
  - Intuitives Verständnis: eine endliche geordnete Folge von Operationen um alle Aufgaben gleichen Typs zu lösen.
  - Das bloßes Erkennen, dass es eine Lösung für ein Problem existiert, liefert noch kein Verfahren zur Lösung selbst!
- Turingmaschine: eine abstrakte Rechenmaschine dient als Grundmodell zur Präzisierung der Begriffe „Berechenbar“ und „Algorithmus“
- Ist ein Problem lösbar? <==> Gibt es zur Lösung des Problems ein Algoritmus.

Berechenbarkeit

- Gibt es Aufgaben, die vom Computern nicht gelöst werden können == Gibt es Probleme, für die es keinen Algorithmus gibt?
  - Hilbert Entscheidungsproblem: gibt es ein Verfahren, das für ein beliebige mathematische Aussage entscheiden kann, ob die Aussge WAHR oder FALSCH sei?
  - Beweis von Gödel: So ein Algorithmus kann nicht existieren (Unvollständigkeitssatz)
- Eine Funktion *f* heisst dann berechenbar, wenn es einen Algorithmus gibt, der für beliebiges Eingabewert *x*, für den *f* definiert ist, nach endlich vielen Schritten anhält und als Ergebnis *f(x)* liefert.

Prominente unlösbare Probleme

- Gesamtverlauf der Evolution
  - Mangels Daten existiert noch keine einheitliche Theorie über die Entwicklung neuer Gene und neuer Arten, sowie über das Zusammenspiel zwischen Ökologie und Evolution
- Radioaktiver Zerfall
  - Wegen spontane änderungen der Eigenschaften des Systems, können keine Zerfallsvorhersagen über den Zerfall eines konkreten Atoms getroffen werden
- Problem des Handelsreisenden
  - Das Problem ist wohl definiert und logisch Lösbar, überschreitet wegen große Anzahl von Kombinationen die zeitlichen und räumlichen Grenzen des Machbaren
- Teilung eines Winkels in drei gleiche Teile mit einem Lineal und Zirkel
  - Ein wohldefiniertes Problem, zu dem es keinen Algorithmus gibt.

Prominente unlösbare Probleme

- Gibt es für jede Zahl *n* einen Funktionswert?

```
while (n!=1){
   if(n!=1){
      n=n/2;
   else
      n=3*n+1
   }
}
```

- Existiert eine ganzzahlige Lösung?
  - Gibt es für eine beliebige Gleichung ein Verfahren, das entscheidet, ob sie eine ganzzahlige Lösung hat.
- Hält ein Programm an?
  - Ein Computer soll entscheiden, ob ein Programm für eine dazugehörige Dateneingabe terminiert.

Einige offene Probleme

Name	Situation	Problem
Kartenproblem	Ein Spediteur hat N Pakete mit unterschiedlichem Gewicht zu transportieren. Jeder Lastwagen faßt maximal ein Gewicht G_{max}.	Ermittlung der minimalen Anzahl von Lieferwagen zum Transport der Pakete.
Verschnitt-problem	Ein Schneider hat eine bestimmte Anzahl von Schnittmustern herzustellen.	Ermittlung des minimalen Stoffverbrauchs
Stundenplan-problem	Gegeben sind eine Liste von Fächern, ein Zeitraster und eine Anzahl von Schulklassen.	Ermittlung der minimalen Anzahl von Leerstunden
Verdrahtungs-problem	Gegeben sind die Anschlußpunkte und die Menge von Verbindungen und Verdrahtungskanälen mit bestimmten Kapazitäten sowie eine Anzahl von Leitungsebenen für eine elektrische Leiterplatte oder das Layout eines Halbleiter-Chips.	Ermittlung des kürzesten, kreuzungsfreien Verdrahtungsplanes
Laufzeit-problem	Gegeben sind eine Anzahl von Programmen mit den dazugehörenden Laufzeiten sowie eine Zeitvorgabe, wonach alle Programme berechnet sein müssen.	Ermittlung der Minimalzahl an Prozessoren zur Bewältigung der Aufgabe
Problem des Handlungs-reisenden	Gegeben seien die Reisezeiten zwischen den verschiedenen Orte eines Landes. Der Handlungsreisende soll davon eine bestimmte Anzahl besuchen.	Ermittlung der minimalen Fahrzeit

KOMPLEXITÄTSTHEORIE

Komplexitätstheorie

- https://www.youtube.com/watch?v=DkE8R-4CKOQ
- KT Befasst sich mit der Komplexität von algorithmisch behandelbaren Problemen auf verschiedenen mathematisch definierten formalen Rechnermodellen.
- Messung der Komplexität anhand von benötigten Ressourcen (Zeit, Speicher)
- Unterschied zur Berechenbarkeitstheorie?

Komplexitätsbetrachtung von Algorithmen

- Die Komplexitätsordnung eines Algorithmus bestimmt maßgeblich die Laufzeit und Speicherplatzbedarf der konkreten Implementierung.
- Weitere Einflussgrößen sind u.a.:
  - Eingabedaten
  - Rechner-Hardware
  - Qualität des Compilers
  - Betriebssystem
- Wir betrachten die abstrakte Laufzeit in Abhängigkeit des Parameters n der Eingabedaten.

Größe der Eingabedaten

- Der Parameter n der Eingabedaten drückt die „Größe“ des Problems aus.
- Diese Größe kann von Fall zu Fall unterschiedlich sein.

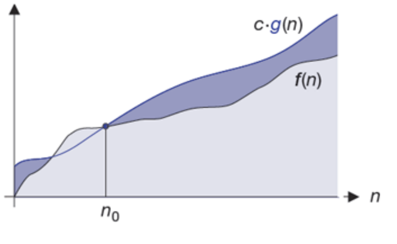
- Beispiele:
  - Beim Sortieren ist *n* die Anzahl der zu sortierenden Werte $a_1, \ldots, a_n$.
  - Bei der Suche nach Primzahlen gibt *n* z.B. die obere Grenze des abzusuchenden Zahlenbereichs an.

Komplexitätsanalyse von Algorithmen

- Asymptotische Notation (O-Notation, Landau-Notation):

  Die Notation *O(g(n))* beschreibt eine Klasse von Funktionen, deren qualitatives Wachstumsverhalten für große n höchstens proportional zu g(n) ist (die Funktion g(n) stellt für große n eine obere Schranke dar):

$$f(n) \in O(g(n)) \quad :\Leftrightarrow \quad \exists\, c, n_0 \; \forall n > n_0 : f(n) < c \cdot g(n)$$



Komplexitätsanalyse von Algorithmen

- Die O-Notation stellt eine obere Schranke für das qualitative Wachstumsverhalten dar. Das bedeutet jedoch nicht, dass es die kleinste obere Schranke ist.
- Beispiel:

  Falls gilt $f(n) \in O(n)$, so gilt z.B. auch $f(n) \in O(n^2)$ oder $f(n) \in O(2^n)$

- Praktisch wählt man im Falle von $f(n) \in O(g(n))$ allerdings häufig ein $g(n)$, das eine kleinste obere Schranke ist.
- Man benutzt dabei $O(g(n))$ im Sinne der Notation $\Theta(g(n))$, wobei $f(n) \in \Theta(g(n))$ nicht nur $f(n) \in O(g(n))$ bedeutet, sondern auch $g(n) \in O(f(n))$.

Betrachtungsweisen der Laufzeit

- Wenn wir die Laufzeit eines Algorithmus betrachten, können verschiedene Fragen interessant sein:
  - Best-Case: Wie lange rechnet das Verfahren mindestens
  - Worst-Case: Wie lange rechnet das Verfahren höchstens
  - Average-Case: Wie lange rechnet das Verfahren durchschnittlich

- Beispiel: Suche eines Elements in einer unsortierten Liste
  - Best-Case: Treffer beim ersten Vergleich → $O(1)$
  - Worst-Case: Treffer beim letzten Vergleich → $O(n)$
  - Average-Case: Treffer nach durchschnittlich $n/2$ Vergleichen

Komplexitätsklassen

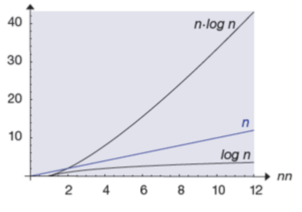
Klasse	Bezeichnung	Beispiel
1	konstant	elementarer Befehl
$\log(\log n)$	doppelt logarithmisch	Interpolationssuche
$\log n$	logarithmisch	binäre Suche
n	linear	lineare Suche, Minimum einer Folge
$n \log n$	überlinear	Divide-and-Conquer-Strategien, effiziente Sortierverfahren, schnelle Fourier-Transformation (FFT)
n^2	quadratisch	einfache Sortierverfahren
n^3	kubisch	Matrizen-Inversion, CYK-Parsing
n^k	polynomiell vom Grad k	lineare Programmierung
2^n	exponentiell	erschöpfende Suche (exhaustive search), Backtracking
$n!$	Fakultät	Zahl der Permutationen (Traveling-Salesman-Problem)
n^n		

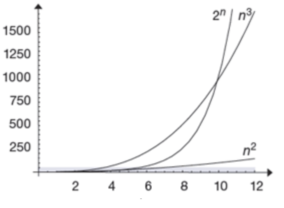
Quelle: Ney (2004): Skript Algorithmen und Datenstrukturen

Wachstumsverhalten

n	$\lg n$	$\lg^2 n$	$\sqrt{n}$	$n \lg n$	$n \lg^2 n$	$n^{\frac{3}{2}}$	n^2
10	3	9	3	30	90	32	100
100	6	36	10	600	3600	1000	10 000
1000	9	81	32	9000	81 000	31 623	1 000 000
10 000	13	169	100	130 000	1 690 000	1 000 000	100 000 000
100 000	16	256	316	1 600 000	25 600 000	31 622 777	10 Milliarden
1 000 000	19	361	1000	19 000 000	361 000 000	1 Milliarde	1 Billion

●Quelle: Herold et. al. (2007): Grundlagen der Informatik





Wachstumsverhalten

Annahme: 1 Schritt dauert 1 *µs = 0.000001 s*

$n =$	10	20	30	40	50	60
n	10 μs	20 μs	30 μs	40 μs	50 μs	60 μs
n^2	100 μs	400 μs	900 μs	1.6 ms	2.5 ms	3.6 ms
n^3	1 ms	8 ms	27 ms	64 ms	125 ms	216 ms
2^n	1 ms	1 s	18 min	13 Tage	36 J	366 Jh
3^n	59 ms	58 min	6.5 J	3855 Jh	10^8 Jh	10^{13} Jh
$n!$	3.62 s	771 Jh	10^{16} Jh	10^{32} Jh	10^{49} Jh	10^{66} Jh

MATHEMATISCHE GRUNDLAGEN

Logik

Aussagen: A, B, C, …

sind entweder wahr (= w, t, 1) oder falsch (= f, 0)

Operationen: ¬ (̅), ∧, ∨, =>, <=>, …

Formeln, Wahrheitstafeln, Tautologien

Aussageform: $A_{(x)}$, $A_{(x,y)}$, …

wird durch konkrete Werte von x zu einer Aussage

Quantoren: ∃, ∀

Beweisverfahren: direkt, indirekt, vollständige Induktion

Menge

- Menge M von Elementen, wird beschrieben als Aufzählung
  - Mengen: A, B, C, …; leere Menge ∅ bzw. { }
    - Beschreibung von Mengen: A = {1, 2, 3}
    - M = {blau, rot, rosa, grün, grau}
- oder als Menge von Elementen mit einer bestimmten Eigenschaft
  - M = {n | n ϵ $\mathbb{N}_0$ und n gerade}
  - B = {n ∈ ℕ | n > 5}
- Allgemeines Format:
  - M = {x | P(x)}
    (M ist Menge aller Elemente x, die die Eigenschaft P erfüllen.)
- Operationen, Operatoren:
  - |A|, ∩, ∪, ¯, \, …

Potenzmenge der Menge A: P(A)

$P_{(A)} = \{B \mid B \subset A\}$

z.B. A = {0, 1} dann ist $P_{(A)} = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$

$|P_{(A)}| = 2^n$ mit $n = |A|$

Beweis?

- Potenzmenge = Menge aller Teilmengen

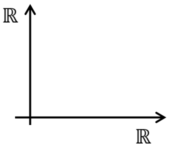
Produktmenge (kartesisches Produkt)

A x B ; A, B Mengen

- zweistellig, binär

$A_1 \times A_2 \times A_3 \times \ldots \times A_n$; $A_1, \ldots, A_n$ Mengen

- n-stellig



~ $\mathbb{R} \times \mathbb{R}$ -> kartesisch

$A \times B = \{(a, b) \mid a \in A, b \in B\}$

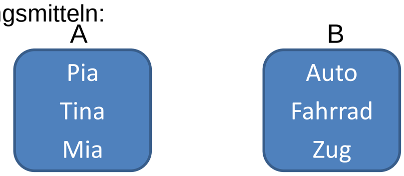
$|A \times B| = n * m$ mit $|A| = n, |B| = m$

Relation

Eine Relation besteht aus zwei Dingen:

1. Eine Aussageform
2. Den Paaren, die diese Aussageform erfüllen

Gegeben Sei die Menge A von drei Frauen und eine Menge B von Fortbewegungsmitteln:



Eine binäre Aussageform zwischen den Elementen der beiden Mengen wäre *a fährt zur Arbeit mit dem Fortbewegungsmittel b*.

Die Paare, die diese Aussageformel füllen sind z.B.

Pia fährt mit dem Auto, Mia fährt mit dem Fahrrad, Tina fährt mit dem Zug, Tina fährt mit dem Auto

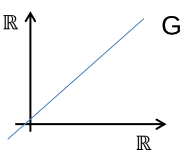
Relation R

- $R \subset A \times B$ ~ zweistellige (binäre) Relation
- [$R \subset A_1 \times A_2 \times \ldots A_n$ ~ n-stellige Relation]

$G \subset \mathbb{R} \times \mathbb{R}$

$G = \{(x, x) \mid x \in \mathbb{R}\}$

Relation ≙ Beziehung



Relation - Beispiel

Es gibt Personen

P ⊂ Nr x, Name

P = {(1, Meyer), (2, Müller), … }   P ist Relation

Aber auch:

Kennt ⊂ P x P   (Beziehung: Relation über Relationen)

Kennt = {((1, Meyer), (2, Müller)), ((3, Schmidt), (7, Schultze)), …}

Funktion

$f : A \rightarrow B$

$a \rightarrow f(a)$

Die Funktion f bildet ein Element $a \in A$ auf ein Element $f(a) \in B$ ab.

A: Definitionsbereich; B: Wertebereich

Beispiel: (Quadratfunktion)

$f: \mathbb{Z} \rightarrow \mathbb{N}_0$, $f(n) = n^2$

…, $-3 \rightarrow 9$, $-2 \rightarrow 4$, $-1 \rightarrow 1$, $0 \rightarrow 0$, $1 \rightarrow 1$, $2 \rightarrow 4$, $3 \rightarrow 9$,…

Dabei ist $\mathbb{N}_0$ die Menge der natürlichen Zahlen (mit der Null) und $\mathbb{Z}$ die Menge der ganzen Zahlen.

Funktion vs. Relation

- Funktionen sind Relationen aber nicht umgekehrt

$$f: \begin{cases} \mathbb{R} \rightarrow \mathbb{R} \\ x \rightarrow f_{(x)} \end{cases}$$

$\rightarrow f \subset \mathbb{R} \times \mathbb{R}$

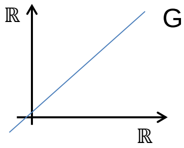
$f = \{(x, f_{(x)}) \mid x \in \mathbb{R}\}$

- z.B. $f_{(x)} = x$

$\rightarrow f = \{(x, x) \mid x \in \mathbb{R}\}$

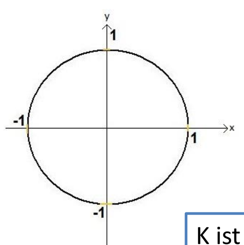
$f = G$

Jede Funktion kann als Relation geschrieben werden!



Funktion vs. Relation

- Gibt es Relationen, die keine Funktionen sind?



$\rightarrow x^2 + y^2 = 1$

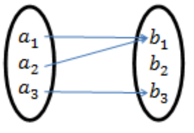
$K = \{(x, y) \mid x^2 + y^2 = 1\} \subset \mathbb{R} \times \mathbb{R}$

K ist Relation, aber keine Fkt., da z.B. x = 0 zwei y-Werte (+1, -1) zugeordnet werden.
Bei einer Fkt. darf es nur eine Zuordnung geben.

Zweistellige Relationen

- $R = \{(a_1, b_1), (a_2, b_1), (a_3, b_3)\} \subset A \times B$

$A = \{a_1, a_2, a_3\}$ $B = \{b_1, b_2, b_3\}$



Andere Notation:

(a, b) ∈ R <=> a R b (Infix-Notation)

~ zwischen a und b besteht die Relation R

~ a steht bzgl. R mit b in Beziehung

Yes we can…