

Informatik-Propädeutikum

Dozentin: Dr. Claudia Ermel

Betreuer: Sepp Hartung, André Nichterlein, Clemens Hoffmann

Sekretariat: Christlinde Thielcke (TEL 509b)

TU Berlin

Institut für Softwaretechnik und Theoretische Informatik

Prof. Niedermeier

Fachgruppe Algorithmik und Komplexitätstheorie

<http://www.akt.tu-berlin.de>

Wintersemester 2013/2014

Gliederung

12 Computational Thinking

- Begriff und Definitionsversuch

- Trend Automatisierung

- Trend Kommunikation

- Trend (massive) Datenanalyse: „Big Data“

- Aspekte des Computational Thinkings

- Computational Thinking im Alltag

- Algorithmische Biologie

- Computational Social Choice

- Algorithmische Spieltheorie

Computational Thinking

- Begriff geprägt von **Jeannette M. Wing** (2006 Professorin an der Carnegie Mellon University, Pittsburgh; seit 2012 Chefin von Microsoft Research International)
- Artikel „Computational Thinking“ (Communications of the ACM 2006)
- → Nicht nur Informatiker profitieren davon, informatisch (algorithmisch) zu denken.



Ein Definitionsversuch: „*Computational Thinking* is the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent.“ [CunySnyderWing 2010]

⇒ Paradigmenwechsel von den „*metal tools*“ (Transistoren und Kabel) zu unseren „*mental tools*“ (Abstraktionen und Methoden)

Motivation Computational Thinking

Informatikprodukte beeinflussen alle Lebensbereiche:

Internet, Suchmaschinen, Mobiltelefonie, Electronic Banking, Einkaufen im Internet, Entzifferung des menschlichen Genoms, Klimavorhersage, Navigationssysteme, soziale Netzwerke, Wikipedia, Digitalkameras, ...

Oft vollziehen sich damit einhergehende Veränderungen in rapidem, bislang nicht gekanntem Tempo... Z.B.:

- Wieviele Buchläden wird es in zehn Jahren noch geben? Was ist die Zukunft von Bibliotheken?
- Wer kauft noch Musik-CDs im Laden?
- Wieviele von Menschen verrichtete Arbeiten werden bald durch Roboter oder Informatiksysteme ersetzt (inkl. Pilot, U-Bahnfahrer, Verwaltungsbeamte, Lehrer, Vorleser, ...)?
- Wer bucht seine Reise noch im Reisebüro?
- Wer merkt sich noch Telefonnummern?

Zukunftstrends (dank Informatik)

- Automatisierung.
- Kommunikation.
- Datenanalyse.
- Informatik schafft Werkzeuge zur „Intelligenzverstärkung“, nicht nur zur „Kraftverstärkung“.

Credo: Wer bei Zukunftstrends mitreden und vielleicht auch mitgestalten will, braucht ein solides Grundverständnis, braucht *Computational Thinking*!

Beispiele: Einsicht in

- Strategien des Problemlösens,
- Prinzipien des Systementwurfs und der Systemanalyse,
- menschliches Verhalten.

Trend Automatisierung

Beispiel Smart Home (Prof. Möller, Quality & Usability Lab)

„**Früher**“: Wecker von Hand stellen → Fällt am nächsten Tag die Vorlesung aus: Wecker umstellen → Wecker klingelt immer zur eingestellten Zeit → Wecker von Hand ausstellen.

„**Heute**“: Wecker synchronisiert sich mit dem Online-Kalender und dem BVG-Fahrplan (kennt Termine und Fahrzeiten) → Wecker synchronisiert sich mit Messdaten des Bettes (kennt Biorhythmus) → Wecker wählt passende Musik und steuert Beleuchtung in passender Stärke und Farbe.

Weitere Beispiele:

- Reiseplanung: Internet statt Reisebüro.
- Fotos: Speicherchip statt Film.
- Hochindividuell gestaltbare technische Produkte wie Autos.
- Navigation statt Kartenlesen.
- Telemedizin (Überwachung von Symptomen von zu Hause aus)

Trend Kommunikation

Beispiel Briefe (Prof. Feldmann, Intelligente Netze)

„**Früher**“: Brief in Briefkasten werfen → Einsammeln der Briefe und Transport zur Postzentrale → Sortieren → Transport zur passenden Zielortzentrale → Sortieren → Brief austragen.

„**Heute**“: Email: Click...

Weitere Beispiele: Radikale Verhaltensänderungen durch Dinge wie

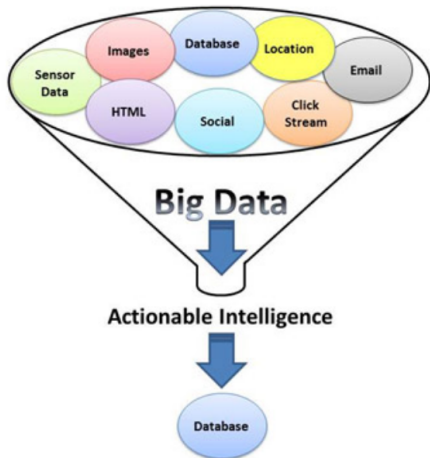
- World Wide Web,
- Smartphones,
- Soziale Netzwerke,
- Internetbasierter Einkauf,
- Online-Banking, ...

Geschwindigkeit und Umfang von damit verknüpften Diensten steigt,
Preis sinkt!

Trend (massive) Datenanalyse: „Big Data“

Massive Datenanalyse (Prof. Markl, DB & Informationsmanagement) ermöglicht zum Beispiel

- Maschinelle Sprachübersetzung
- Maschinensteuerung per Gedanken
- Entzifferung des menschlichen Genoms (Informationsgehalt 6 Milliarden bit)
- Personalisierte Werbung, aber auch personalisierte Medizin;
- Aufdecken von Zusammenhängen in der Literatur; auch: Plagiate finden;
- Empfehlungssysteme („Wer dieses Buch gekauft hat, der hat auch...“)
- Terror-Bekämpfung durch Geheimdienste (?)



Aspekte des Computational Thinkings

Computational Thinking umfasst bzw. verbindet

- algorithmisches Denken;
- mathematisches Denken;
- logisches Denken;
- Denken in Systemen;
- paralleles und verteiltes Denken;
- ingenieurmäßiges Denken.

Eine wesentliche Besonderheit: Mit Software entstehen virtuelle Welten, die nicht durch die physikalische Realität beschränkt sind.

Wichtigster Aspekt des Computational Thinkings: **Abstraktion!**
Hinzu kommen weiterhin Rekursion und Hierarchisierung,
Automatisierung, Analyse.

Thesen zum Computational Thinking

Computational Thinking beansprucht für sich folgende Charakteristiken:

- Denken in Konzepten und Abstraktionen, nicht bloß Programmieren.
- Ergänzt und kombiniert mathematisches und ingenieurmäßiges Denken.
- Ideen zählen, nicht bloß Artefakte (Gegenstände).
- Grundlegende Fähigkeiten statt Routine sind nötig.
- Mit Computational Thinking (und Informatik) kann man in *allen* Bereichen Erfolg haben, von Medizin über Politik bis Kunst.
- Ist Grundlage zur Lösung vieler intellektuell herausfordernder, zentraler wissenschaftlicher Probleme, die noch ungelöst sind.

Computational Thinking im Alltag: Beispiele

Sortieren: Wenn Sie Ihre Krimis nach Autoren alphabetisch sortieren sollen, werden Sie die Vorzüge von „Merge Sort“ oder „Quicksort“ zu schätzen lernen...

Backtracking: Wenn Sie Ihren Handschuh verloren haben, müssen Sie den Weg zurückgehen, bis Sie ihn wiederfinden.

Online Shopping Algorithmus: Welche Gründe bewegt Sie zu der Entscheidung, lieber selber ein Paar Skier zu kaufen statt immer welche zu mieten?

Redundanz und Fehler-Toleranz: Wenn Sie Festnetz-Telefon haben, dann können Sie auch bei Stromausfall telefonieren.

Server-Scheduling: An welche Schlange an der Kasse stellen Sie sich?

Pipelining: Wer hat sich nicht schon mal über die „dumme“ Anordnung von einzelnen Stationen in einer Mensa aufgeregt (z.B. ist es „vernünftig“ (wie in der TU Mensa), dass man das Besteck erst nach der Kasse nimmt...)

Computational X

Beleg für den allumfassenden Anspruch des Computational Thinkings:
„Bindestrichdisziplinen“ der Informatik

(im englischsprachigen Raum oft mit dem Präfix „Computational“ versehen – was im Deutschen manchmal einschränkend mit „algorithmisch“ übertragen wird)

Das X in „Computational X“, steht dabei für Gebiete wie Biologie, Chemie, Geologie, Medizin, Wirtschaft, Spieltheorie, Geometrie, Recht, Politik und Sozialwissenschaft („Computational Social Choice“), Linguistik, Archäologie, ...

Nachfolgend gehen wir kurz näher ein auf

- Algorithmische Biologie,
- Computational Social Choice und
- Algorithmische Spieltheorie.

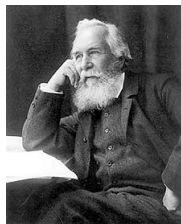
Algorithmische Biologie

Algorithmische Biologie betrachtet die Berechnungskomplexität und die algorithmischen Eigenschaften von Berechnungsproblemen aus der Bioinformatik.

Beispiel: Rekonstruktion von phylogenetischen Bäumen

Ernst Heinrich Philipp August Haeckel (1834-1919)
prägte den Begriff „Stamm“.

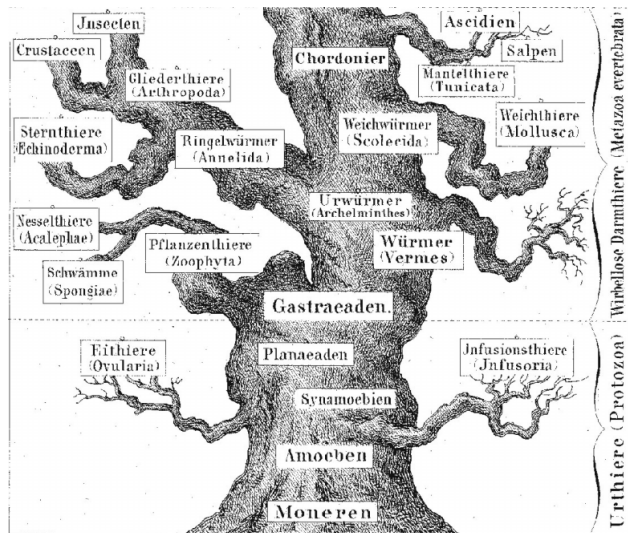
- Deutscher Philosoph und Zoologe;
- verbreitete die Lehren Charles Darwins in Deutschland;
- maßgebliche Arbeiten zur Abstammungslehre.



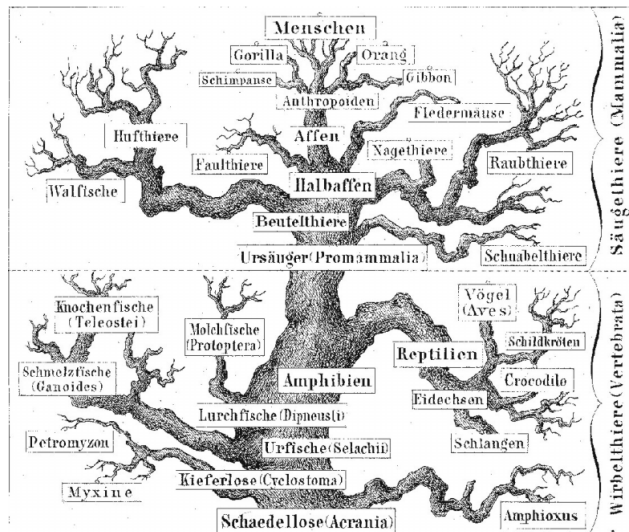
*... the great Tree of Life fills with its dead and broken branches
the crust of the earth, and covers the surface with its
ever-branching and beautiful ramifications.*

— Charles Darwin

Stammbaum des Menschen nach Haeckel I



Stammbaum des Menschen nach Haeckel II

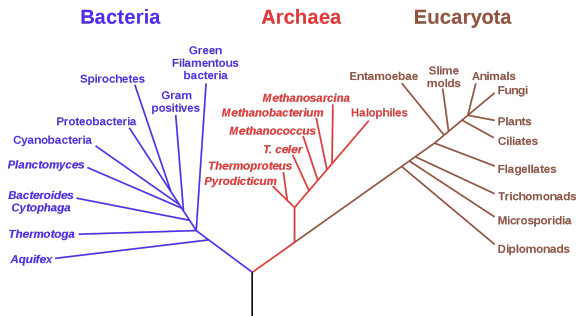


Rekonstruktion von phylogenetischen Bäumen I

Gegeben eine Menge von verschiedenen biologischen Objekten und Beobachtungen über deren evolutionäre Zusammenhänge.

Ziel: Modell in Form eines Baumes (Graph) der evolutionären Relation zwischen den biologischen Objekten. Blätter sind die Objekte und innere Knoten stellen die evolutionären Zusammenhänge dar.

Phylogenetic Tree of Life



Evolutionärer Zusammenhang
wird z. B. über „Editierdistanz“ zwischen den entsprechenden DNS/RNS Sequenzen oder Proteinen gemessen.

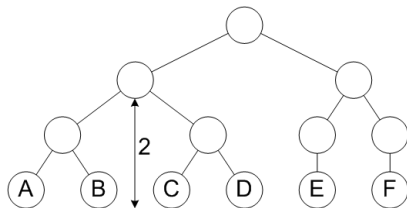
Rekonstruktion von phylogenetischen Bäumen II

Formalisieren der Problemstellung:

Eingabe: Symmetrische Distanzfunktion $d : V \times V \rightarrow \mathbb{N}$ für Objektmenge V .

Aufgabe: Finde Darstellung von d als **ultrametrischer Baum**.

	A	B	C	D	E	F
A	-	1	2	2	3	3
B	-	-	2	2	3	3
C	-	-	-	1	3	3
D	-	-	-	-	3	3
E	-	-	-	-	-	2
F	-	-	-	-	-	-



Distanz im Baum: Höhe des „tiefsten gemeinsamen Vorfahren“

Distanzfunktion d kann als ultrametrischer Baum dargestellt werden gdw. d eine **Ultrametrik** ist, d. h.

$$d(a, b) \leq \max\{d(a, c), d(b, c)\} \text{ für alle } a, b, c \in V$$

Computational Social Choice

Social Choice-Theorie studiert die Mechanismen des kollektiven Entscheidungsfindung wie Wählen, Aggregieren von Präferenzen, faires Teilen, Zuordnungsprobleme (Matching), ...

- Vorläufer: Condorcet, Borda (18. Jahrhundert),...
- „Ernsthafte“ wissenschaftliche Disziplin seit den 1950er Jahren.

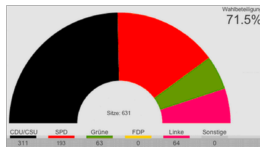
Computational Social Choice fügt dem ganzen eine algorithmische Perspektive hinzu, und untersucht auch die Anwendung von Konzepten aus dem Bereich „Social Choice“ in der Informatik und angrenzenden Gebieten.

- „Klassische“ Veröffentlichungen beginnen ~ 1990.
- Aktives Forschungsgebiet mit regelmäßigen Veröffentlichungen seit etwa 2002.
- Name „COMSOC“ und alle zwei Jahre stattfindende Spezialtagung gleichen Namens.

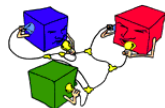
Wir konzentrieren uns hier auf „Wahlprobleme“.

Wo Wahlprobleme auftreten...

Politische Wahlen. Viele Wähler und wenige Kandidaten.



Multiagentensysteme. Einigung auf einen gemeinsamen Arbeitsplan.



Suchmaschinen im Internet. Rank Aggregation mehrerer Suchmaschinen;

bing + Google + YAHOO! + ...

Empfehlungsdienste. Empfehlung von Produkten, z.B. Bücher oder Videos, basierend auf Rankings anderer Benutzer;

amazon.com

Wahlen und Wahlregeln

In einer **Wahl** handelt es sich um eine gemeinsame Entscheidung mehrerer **Wähler** zwischen verschiedenen **Kandidaten** gemäß einer **Wahlregel**.

Beispiel (Wissenschaftlerranking an der TU Berlin)

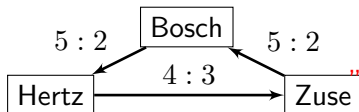
2 Studenten: Bosch \succ Hertz \succ Zuse

2 Studenten: Hertz \succ Zuse \succ Bosch

3 Studenten: Zuse \succ Bosch \succ Hertz

Zwei Wahlregeln:

- 1 **Mehrheitswahl:** Derjenige Kandidat gewinnt, der am häufigsten an erster Stelle in den Stimmabgaben steht. **Gewinner: Zuse.**
- 2 **Condorcet-Verfahren:** Derjenige Kandidat gewinnt, der gegen jeden anderen Kandidaten die Mehrheit im direkten Vergleich erhält.



„Condorcet-Zykel“ \Rightarrow Kein Gewinner!

Wünschenswerte Eigenschaften von Wahlregeln

- **Condorcet-Eigenschaft.** Wenn es einen Condorcet-Gewinner gibt, dann soll dieser auch die Wahl gewinnen.
- **Neutralität.** Alle Kandidaten werden gleich behandelt.
- **Unabhängigkeit von unbedeutenden Kandidaten,** engl. **Independence of Irrelevant Alternatives (IIA).** Wenn Kandidat x Gewinner einer Wahl W_1 ist und ein neuer Kandidat y zusätzlich bei einer Wahl W_2 kandidiert (mit sonst den gleichen Kandidaten wie vorher), dann wird entweder x oder y Gewinner der Wahl W_2 sein.
- **Konsistenz.** Wenn ein Kandidat in zwei Wahlen gewinnt, dann muss dieser auch der Gewinner in der Vereinigung beider Wahlen sein.

Im folgenden betrachten wir die einzige Wahlregel, dass alle diese Eigenschaften besitzt.

Das Kemeny-Verfahren I

Mindestens dreimal unabhängig voneinander beschrieben:

- Condorcet (1785) für 3 Wähler
- Kemeny (1959), auch Erfinder der Programmiersprache Basic
- Fobes (1991) www.votefair.org

Beim **Kemeny-Verfahren** sucht man eine „Konsensliste“, die die „Gesamtmeinung“ aller Präferenzlisten am besten wiedergibt.

Anwendungen:

- Metasuchmaschinen,
- Suche in Datenbanken,
- Bewerberauswahl/Auswahl von Studierenden,
- Bioinformatik,
- ...

Das Kemeny-Verfahren II

Der **Kendall-Tau-Abstand** zwischen zwei Präferenzlisten ist die Anzahl der unterschiedlich geordneten Kandidatenpaare.

Beispiel: Bosch \succ Hertz \succ Zuse
 Hertz \succ Zuse \succ Bosch

Konfliktpaare: {Bosch, Hertz} und {Bosch, Zuse}.

Der Abstand ist also 2.

Die **Kemeny-Punktezahl** einer Liste ist die Summe der Abstände zu allen Präferenzlisten.

Eine **Kemeny-Konsensliste** ist eine Liste mit kleinstmöglicher Punktezahl.

Eingangsbeispiel: 2 Studenten: Bosch \succ Hertz \succ Zuse
 2 Studenten: Hertz \succ Zuse \succ Bosch
 3 Studenten: Zuse \succ Bosch \succ Hertz

Die einzige Kemeny-Konsensliste ist „Zuse \succ Bosch \succ Hertz“ mit 8 Punkten.

Das Kemeny-Verfahren III

Frage: Wie findet man eine Liste mit kleinster Kemeny-Punktzahl?

Naive Lösung: Alle $m!$ möglichen Permutationen der m Kandidaten durchprobieren.

↪ Exponentielle Laufzeit!

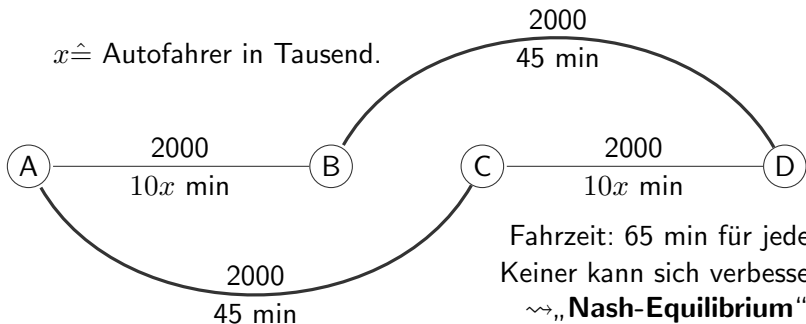
Mitteilung:

- NP-schwer, allerdings noch effizient lösbar für wenige Kandidaten;
- Hilft es, wenn man nur wenige Wähler hat (Bsp. Bewerbungen)?
Nein, auch dann immer noch NP-schwer, wenn es nur vier Wähler gibt.

Algorithmische Spieltheorie I - Nash-Equilibrium

Wikipedia: In der Spieltheorie werden Entscheidungssituationen modelliert, in denen sich mehrere Beteiligte gegenseitig beeinflussen.

Bsp. Verteilung von Autofahrern in einem Straßennetz:



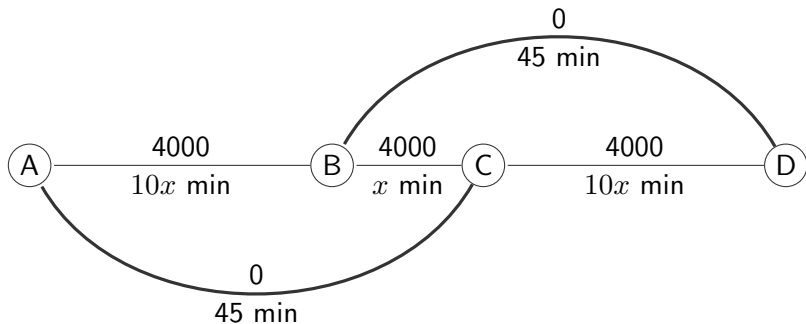
Zur Feierabendzeit wollen 4000 Pendler von A nach D fahren.

\rightsquigarrow was ist der schnellste Weg?

Klar: Fahrtzeit hängt von Verkehrsdichte ab.

Algorithmische Spieltheorie II - Braess-Paradoxon

fortgesetztes Beispiel: Zwischen B und C wird eine große Brücke gebaut.



Aus Pendlersicht: Autobahn ist nicht der schnellste Weg zwischen A und C bzw. B und D.

⇝ Keiner nutzt die Autobahn.

⇝ Fahrzeit erhöht sich auf 84 min für jeden Pendler.

Ausbau des Verkehrsnetzes erhöht Fahrzeit! ⇝ **Braess-Paradoxon!**

Algorithmische Spieltheorie III - Fazit

Vorheriges Beispiel zeigt: Optimiert jeder Fahrer nur die eigene Fahrzeit (**rational egoistisches Vorgehen**), muss dies nicht zu einem **globalem Optimum** führen!

Algorithmische Spieltheorie befasst sich mit der Fragen wie z.B.:

- Wie können solche Nash-Equilibria oder globalen Optima (effizient) berechnet werden?
- Was ist eine optimale Strategie für einen Agenten?
- Ist in einem gegebenen Szenario jedes Nash-Equilibrium ein globales Optimum?

Anwendungen:

- Auktionen, Elektronischer Handel
- Routing (z.B. im Internet)
- Verkehrsplanung
- ...

Letzte Folie von Jeannette Wing (CT for Everyone, 2008)

Spread the Word!

Help make computational thinking commonplace!

To fellow faculty, students, researchers,
administrators, teachers, parents, principals,
guidance counselors, school boards, teachers' unions,
congressmen, policy makers, ...