Seminar-Themen

„Algorithmen und Komplexität“ und „Komplexität“

Wintersemester 2017/18 — Prof. Dr. Georg Schnitger

Im Folgenden finden Sie Vortragsthemen für das Bachelor-Seminar „Algorithmen und Komplexität“ und das Master-Seminar „Komplexität“ im Wintersemester 17/18. Die Vergabe der Themen erfolgt in der Vorbesprechung. Bachelor-Studenten dürfen jedes Thema wählen, das mit einem Ⓑ markiert ist. Für Master-Studenten kommen nur Themen in Frage, die mit Ⓜ gekennzeichnet sind. Sind Themen sowohl mit Ⓑ als auch mit Ⓜ markiert, stehen sie beiden Studiengängen offen, wobei der geforderte Umfang für Bachelor-Studenten entsprechend reduziert wird.

Bitte machen Sie sich vorab mit den Themen vertraut, sodass Sie in der Vorbesprechung bereits eine Vorstellung haben, welche Themen für Sie in Frage kommen.

Wenn Sie ein eigenes Thema vorschlagen wollen, setzen Sie sich bitte vor der Vorbesprechung mit uns in Verbindung.

Inhaltsverzeichnis

Komplexitätstheorie

① Threshold Computation und Postselection Ⓜ

Yenjo Han, Lane A. Hemaspaandra, Thomas Thierauf. *Threshold Computation and Cryptographic Security*. SIAM J. Comput. 26(1): 59-78 (1997). http://dblp.org/rec/html/journals/siamcomp/HanHT97

Kurzbeschreibung: Während eine probabilistische Turingmaschine (TM) akzeptiert, wenn die akzeptierenden Berechnungen hinreichend *wahrscheinlich* sind, akzeptiert eine Threshold TM, falls der *Anteil akzeptierender Pfade* hinreichend groß ist. Dies entfesselt eine erstaunlich große Rechenkraft, die in klassischen Berechnungen äquivalent zur Fähigkeit der *Postselection* (s. u.) ist.

Scott Aaronson. *Quantum Computing, Postselection, and Probabilistic Polynomial-Time*. Electronic Colloquium on Computational Complexity (ECCC)(003) (2005). http://dblp.org/rec/html/journals/eccc/ECCC-TR05-003

Kurzbeschreibung: Eine TM mit *Postselection* darf nach einer probabilistischen Berechnung auf ein Ereignis (z. B. „das erste Bit ist 1“) bedingen, d. h. alle Berechnungspfade, die das Ereignis nicht erfüllen, eliminieren. Polynomialzeit-Quantenrechner mit Postselection sind äquivalent zur (sehr mächtigen) Klasse PP. Dieses Resultat liefert auch Indizien dafür, warum die Quantenmechanik so ist, wie sie ist. Würde man bestimmte Stellschrauben leicht ändern, ließe sich Postselection effizient simulieren.

② A Personal View of Average-Case Complexity Ⓜ

Russell Impagliazzo
Structure in Complexity Theory Conference 1995: 134-147. http://dblp.org/rec/conf/coco/Impagliazzo95

Kurzbeschreibung: Was wäre, wenn Hypothesen aus der Komplexitätstheorie wie bpsw. „P = NP“ gälte oder Einwegfunktionen existierten? Auf umgangssprachliche Weise werden fünf verschiedene Welten beschrieben: Algorithmica, Heuristica, Pessiland, Minicrypt und Cryptomania.

Bemerkungen: Diese Arbeit ist als Ausgangspunkt für eine Formalisierung der beschriebenen Welten (Definitionen, Sätze, Beweisideen) anhand der dort zitierten Literatur aufzufassen.
Es können bis zu zwei Vorträge zu dieser Arbeit vergeben werden. Dabei sollte der eine im Themengebiet Average-Case-Complexity (die ersten drei Welten) und der andere im Themengebiet One-Way-Functions (die letzten drei Welten) angesiedelt werden.

③ Zwei-Wege-Automaten – wie stark ist Nichtdeterminismus? Ⓑ

Kapitel 3.3.3 im Skript Theoretische Informatik 2 von Georg Schnitger

Kurzbeschreibung: Für konventionelle Automaten ist die Kraft des Nichtdeterminismus geklärt: Ein NFA kann exponentiell weniger Zustände besitzen als ein dazu äquivalenter DFA. Die Situation für Zwei-Wege-Automaten – diese dürfen den Lesekopf nach rechts oder links bewegen – ist bis heute ungeklärt. Resultate gibt es aber für weiter eingeschränkte Automatenklassen.

④ Lower Bounds for Monotone Circuits Ⓑ Ⓜ

Kapitel 14.3 in *Computational Complexity: A Modern Approach*, Cambridge University Press 2009 von Sanjeev Arora und Baoz Barak

Kurzbeschreibung: Ein monotoner Schaltkreis besitzt OR- und AND-Gatter, aber keine NOT-Gatter. Eine Technik zur Herleitung unterer Schranken für die Größe monotoner Schaltkreise für Entscheidungsprobleme wird vorgestellt. Beispielsweise benötigt das Clique-Problem monotone Schaltkreise mindestens exponentieller Größe. Könnte man die Einschränkung der Monotonie weglassen, so wäre $\mathsf{P} \neq \mathsf{NP}$ gezeigt, doch leider benötigen monotone Schaltkreise auch exponentielle Größe für die Lösung des Matching-Problems.

⑤ Large Peg-Army Maneuvers Ⓑ Ⓜ

Luciano Gualà, Stefano Leucci, Emanuele Natale, Roberto Tauraso
FUN 2016. http://dblp.org/rec/html/conf/fun/Guala0NT16

Kurzbeschreibung: Beim Solitaire-Spiel sind die Spielsteine auf einem Gitter angeordnet. Überspringt man mit einem Stein einen anderen, so muss der andere Stein entfernt werden. Ziel des Spiels ist es, am Ende möglichst wenige Steine auf dem Spielfeld zu haben. Die Entscheidungsversion des Spiels ist NP-vollständig.

⑥ Hardness of Easy Problems: Basing Hardness on Popular Conjectures such as the Strong Exponential Time Hypothesis Ⓑ Ⓜ

Virginia V. Williams
IPEC 2015: 17-29. http://dblp.org/rec/conf/iwpec/Williams15

Kurzbeschreibung: Standardmäßig zeigen wir mithilfe von Polynomialzeit-Reduktionen, dass gewisse Probleme NP-schwer sind. Eine solche Unterscheidung in „effizient“ und „vermutlich nicht effizient“ ist leider nur sehr grob. Diese Arbeit beschäftigt sich mit feineren Reduktionen („fine-grained reductions“) für einfache Probleme (in P), um Fragen von der Form „Geht es besser als $\mathcal{O}(n^k)$?“ zu beantworten.

⑦ Turning Evil Regexes Harmless Ⓑ Ⓜ

Brink van der Merwe, Nicolaas Weideman, Martin Berglund
Proceeding SAICSIT '17. https://dl.acm.org/citation.cfm?id=3129440

Kurzbeschreibung: In der Praxis werden reguläre Ausdrücke verwendet, um eingegebene Strings zu überprüfen, beispielsweise um Fragen wie „Ist die E-Mail-Adresse gültig?“ zu beantworten. Dabei wird häufig versucht, die eingegebenen Strings mithilfe von Backtracking-Algorithmen zu matchen. Reguläre Ausdrücke wie etwa $(a^*)^*$ sind „böse“, da sie bei der Eingabe von $a^n b$ exponentielle Laufzeit von Backtracking-Algorithmen auslösen können. Wie erkennt man böse reguläre Ausdrücke und wie nimmt man ihnen ihre Bösartigkeit?

⑧ Das Schubfachprinzip: Untere und obere Schranken für die Resolution Ⓑ Ⓜ

Bemerkung: Beide Quellen sind zu bearbeiten. Zwei Vorträge sind möglich.

Georg Schnitger: *Skript zur Vorlesung Theoretische Informatik 2, Abschnitt 7.2.1 (Das Schubfachprinzip).* http://thi.cs.uni-frankfurt.de/lehre/gl2/sose17/gl2_skript17.pdf

Kurzbeschreibung: Das Schubfachprinzip (auch: Taubenschlagprinzip) besagt, dass es keine injektive Abbildung von $\{1,\dots,m\}$ nach $\{1,\dots,n\}$ gibt, falls $m > n$. Die Widersprüchlichkeit von KNF-Formeln kann mithilfe der Resolution bewiesen werden. Doch wie lang sind solche Beweise? Satz 7.11 etabliert eine untere Schranke von $2^{n/24}$ Resolutionsschritten für das Schubfachprinzip mit n Elementen und $n-1$ Schubfächern.

Sam Buss, Toniann Pitassi: *Resolution and the Weak Pigeonhole Principle*, CSL 1997. Lecture Notes in Computer Science, vol 1414. https://link.springer.com/chapter/10.1007/BFb0028012

Kurzbeschreibung: In dieser Arbeit wird für einen Spezialfall eine polynomielle obere Schranke für Resolutionsbeweise nachgewiesen. Im Vortrag sollte auch gezeigt werden, dass in der erweiterten Resolution stets Beweise polynomieller Länge existieren; siehe Aufgabe 151 im Skript zur Theoretischen Informatik 2.

⑨ On the resolution complexity of graph non-isomorphism Ⓑ Ⓜ

Jacobo Torán
Proceedings of the 16th international conference on Theory and Applications of Satisfiability Testing (2013). https://dl.acm.org/citation.cfm?id=2525722

Kurzbeschreibung: Im Graph-Isomorphie-Problem sind zwei Graphen gegeben und es ist zu entscheiden, ob ein Isomorphismus zwischen ihnen existiert. Hier betrachten wir eine Einschränkung des Problems, welche in Polynomialzeit lösbar ist: Jeder Knoten hat eine Farbe und Isomorphismen müssen farberhaltend sein; dabei darf keine Farbe häufiger als k mal auftreten. Als KNF formuliert lässt sich die Nicht-Isomorphie zweier Graphen mittels Resolution widerlegen. In dieser Arbeit werden aber Graphen konstruiert, für die (leider) nur exponentiell lange Resolutionsbeweise existieren.

⑩ Solving Single-Digit Sudoku Subproblems Ⓑ

David Eppstein
FUN 2012: 142-153. http://dblp.org/rec/html/conf/fun/Eppstein12

Kurzbeschreibung: Eine populäre Technik zum Lösen von Sudokus ist Nishio. Dabei merkt man sich für jede Ziffer die Menge der Felder, in denen die Ziffer potenziell eingetragen werden kann, bis nur noch eine Möglichkeit übrig bleibt. Diese Arbeit beschreibt einen $o(2^n)$-Algorithmus für Nishio und weist nach, dass Nishio NP-schwer ist.

Algorithmen

(11) An Optimal Online Algorithm for Weighted Bipartite Matching and Extensions to Combinatorial Auctions (B) (M)

Thomas Kesselheim, Klaus Radke, Andreas Tönnis, Berthold Vöcking
ESA 2013: 589-600. http://dblp.org/rec/html/conf/esa/KesselheimRTV13

Kurzbeschreibung: In einem Online-Szenario soll ein wertvollstes Matching auf einem bipartiten Graphen gefunden werden. Dabei sind die Knoten auf der rechten Seite vorab bekannt, während die Knoten auf der linken Seite und die inzidenten Kanten nacheinander eintreffen. Eine frisch eingetroffene Kante muss entweder sofort oder gar nicht ins finale Matching aufgenommen werden. Durch eine Verallgemeinerung des *Secretary Problem* (bzw. dessen Lösung) erhält man hierfür einen guten Algorithmus.

(12) Optimal Partitioning for Dual-Pivot Quicksort (B) (M)

Martin Aumüller, Martin Dietzfelbinger
ACM Trans. Algorithms 12(2): 18:1-18:36 (2016). http://dblp.org/rec/html/journals/talg/AumullerD16

Kurzbeschreibung: Wie schlägt sich ein Quicksort-Algorithmus, der mit zwei Pivot-Elementen $p < q$ arbeitet und die Eingabe in drei Teile „kleiner als p“, „zwischen p und q“ und „größer als q“ partitioniert? Ziel ist es, die asymptotische Anzahl an Vergleichen insgesamt zu minimieren. Es zeigt sich, dass es genügt, die Partitionierungsfunktion zu betrachten. Entscheidend ist, nach welcher Strategie die Frage beantwortet wird, ob ein Element zuerst mit p oder mit q verglichen wird. Erstaunlich leicht lässt sich eine (fast) optimale Strategie herleiten, die dem klassischem 1-Pivot Quicksort überlegen ist.

(13) Train Marshalling Is Fixed Parameter Tractable (B) (M)

Leo Brueggeman, Michael Fellows, Rudolf Fleischer, Martin Lackner, Christian Komusiewicz, Yiannis Koutis, Andreas Pfandler, Frances Rosamond
FUN 2012: 51-56. http://dblp.org/rec/html/conf/fun/BrueggemanFFLKKPR12

Kurzbeschreibung: Wenn die n Wagons eines Zuges am Frankfurter Hauptbahnhof zu neuen Zügen mit unterschiedlichen Zielorten zusammengestellt werden sollen, muss auf mehreren Abstellgleisen rangiert werden. Das zugrundeliegende Entscheidungsproblem ist NP-vollständig. Ist k die Zahl der Abstellgleise, so kann das Problem durch einen Algorithmus mit der Laufzeit $\mathcal{O}(2^{\mathcal{O}(k)} \cdot \text{poly}(n))$, d. h. durch einen sogenannten FPT-Algorithmus (fixed parameter tractable), gelöst werden.

(14) SAT-Algorithmen: Schöning-Algorithmus und Moser-Scheder-Algorithmus (B)

Uwe Schöning, Jacobo Torán
Das Erfüllbarkeitsproblem SAT - Algorithmen und Analysen. Mathematik für Anwendungen 1, Lehmann 2012. http://dblp.org/rec/books/daglib/0028796

Kurzbeschreibung: Im bekannten SAT-Problem ist eine KNF-Formel gegeben und auf Erfüllbarkeit zu prüfen. Da das Problem NP-vollständig ist, sind hier keine Wunder zu erwarten, aber man möchte die exponentielle Laufzeit so niedrig wie möglich drücken. Der randomisierte Schöning-WalkSAT-Algorithmus mit Laufzeit für 3-SAT von $O\big(\text{poly}(n) \cdot (\frac{4}{3})^n\big)$ basiert auf einer Markov-Kette, die den Belegungsraum durchsucht. Der Moser-Scheder-Algorithmus erreicht dieselbe Laufzeit sogar deterministisch, hier werden Überdeckungscodes sowie lokale Suche verwendet.

Bemerkungen: Es können bis zu zwei Vorträge zu dem Thema vergeben werden.

⑮ The Byzantine Generals Problem Ⓑ

Leslie Lamport, Robert E. Shostak, Marshall C. Pease
ACM Trans. Program. Lang. Syst. 4(3): 382-401 (1982). http://dblp.org/rec/html/journals/toplas/LamportSP82

Kurzbeschreibung: Beim verteilten Rechnen gilt es byzantinische Fehler zu vermeiden, die auf folgendes Problem zurückzuführen sind: Mehrere örtlich getrennte Generäle müssen sich auf eine einheitliche Strategie (*Angriff* oder *Rückzug*) verständigen, die Kommunikation ist aber nur zwischen je zwei Generälen mittels Boten möglich. Erschwerend kommt hinzu, dass sich unter den Generälen auch Verräter befinden können, die versuchen werden, die Einigung zu sabotieren.

⑯ Matching Is As Easy As Matrix Inversion Ⓑ

Ketan Mulmuley, Umesh V. Vazirani, Vijay V. Vazirani
Combinatorica 7(1): 105-113 (1987). http://dblp.org/rec/html/journals/combinatorica/MulmuleyVV87

Kurzbeschreibung: Mithilfe eines Isolation-Lemmas wird ein randomisierter, parallelisierbarer Algorithmus zur Berechnung eines perfekten Matchings vorgestellt. Der „schwierigste“ Schritt des Algorithmus ist das Invertieren einer ganzzahligen Matrix.

⑰ The Power of a Pebble: Exploring and Mapping Directed Graphs Ⓑ Ⓜ

Michael A. Bender, Antonio Fernández, Dana Ron, Amit Sahai, Salil Vadhan
Inf. Comput. 176(1): 1-21 (2002). http://dblp.org/rec/html/journals/iandc/BenderFRSV02

Kurzbeschreibung: Ein gerichteter Graph soll von einem Roboter erkundet und kartographiert werden, doch alle Knoten sehen gleich aus. Um die Knoten im Verlauf seiner Reise überhaupt unterscheiden zu können, erhält der Roboter genau einen „Pebble“ der nach Belieben beim Besuch eines Knotens abgelegt oder wieder aufgenommen werden darf. So kann der Roboter einen bereits erforschten Knoten wiedererkennen. Reicht dies für die vollständige Erkundung aus? Und wie stellt man sicher, den Pebble stets wiederzufinden?

⑱ Faktorgraphen und Belief Propagation auf Bäumen Ⓑ

Abschnitt 15.2.2 im Skript zur Vorlesung *Computational Learning Theory* von Georg Schnitger. http://thi.cs.uni-frankfurt.de/lehre/clt/sose17/skript_clt17.pdf

Kurzbeschreibung: Sei $g(x, y, z) = f_1(x, y) \cdot f_2(y) \cdot f_3(x, z)$ eine Faktorisierung der Funktion g. Ein Faktorgraph für g ist ein bipartiter Graph mit der Variablenknotenmenge $\{x, y, z\}$, der Faktorknotenmenge $\{f_1, f_2, f_3\}$ und der Kantenmenge $\big\{\{f_1, x\}, \{f_1, y\}, \{f_2, y\}, \{f_3, x\}, \{f_3, z\}\big\}$. Die Baumstruktur des Graphen kann ausgenutzt werden, um die Summe $\sum_{x,y,z} g(x, y, z)$ auf geschickte Weise zu berechnen („Belief Propagation“ bzw. „Sum-product algorithm“). Semiringe und ihr Distributivgesetz spielen dabei eine besondere Rolle und ermöglichen interessante Anwendungen in der Welt der kombinatorischen Optimierung.

⑲ Survey Propagation Ⓑ Ⓜ

Bemerkung: Beide Quellen sind zu bearbeiten. Ziel des Vortrags soll neben einer verständlichen Darstellung der Funktionsweise des Algorithmus auch eine Einordnung des aktuellen Forschungsstands sein.

M. Mézard, G. Parisi, R. Zecchina: *Analytic and Algorithmic Solution of Random Satisfiability Problems*. Science Vol. 297, Issue 5582, pp. 812-815 (2002). http://science.sciencemag.org/content/297/5582/812

Kurzbeschreibung: Survey Propagation ist ein Algorithmus zur Lösung des KNF-SAT-Problems, der (wie auch Belief Propagation) von Methoden der statistischen Physik inspiriert ist. Dieser Artikel bietet eine schöne Übersicht über die physikalischen Hintergründe des Verfahrens.

Braunstein, A., Mézard, M. and Zecchina: *Survey propagation: An algorithm for satisfiability*, Random Struct. Alg., 27: 201–226 (2005). http://onlinelibrary.wiley.com/doi/10.1002/rsa.20057/abstract

Kurzbeschreibung: Diese Arbeit behandelt die technischen Details.

Lerntheorie

(20) The Multi-armed bandit Problem: Stochastic bandits (M)

Kapitel 2 aus *Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems* von Sébastien Bubeck und Nicolò Cesa-Bianchi in *Foundations and Trends in Machine Learning, Volume 5*. http://dblp.org/rec/html/journals/ftml/BubeckC12

Kurzbeschreibung: Ein Spielautomat mit K Armen ist gegeben. Welchen Arm sollten wir ziehen, um unseren Gewinn zu maximieren, wenn die Gewinnausschüttung jedes Armes einer fixen aber uns unbekannten Wahrscheinlichkeitsverteilung folgt? Hier muss eine Abwägung zwischen Exploration und Exploitation getroffen werden: Ziehe ich einen Arm, über den ich noch nicht viel weiß, oder einen der bisher erfolgversprechendsten? Das Problem wird sowohl im Hinblick auf algorithmische Lösungen als auch untere Schranken behandelt.

(21) The Multi-armed bandit Problem: Adversarial bandits (M)

Kapitel 3.1 und 3.3 aus *Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems* von Sébastien Bubeck und Nicolò Cesa-Bianchi in *Foundations and Trends in Machine Learning, Volume 5*. http://dblp.org/rec/html/journals/ftml/BubeckC12

Kurzbeschreibung: Wieder ist ein Spielautomat mit K Armen gegeben, doch hier wird die Gewinnausschüttung jedes Armes von einem gegnerischen Spieler bestimmt. Auch hier werden algorithmische Lösungen und untere Schranken geliefert.

(22) Beat the mean Bandit (M)

Yisong Yue, Thorsten Joachims
ICML 2011: 241-248. http://dblp.org/rec/html/conf/icml/YueJ11

Kurzbeschreibung: Hier werden duellierende Banditen betrachtet: In jedem Schritt werden zwei Arme simultan gezogen und statt eines Gewinns erhält der Spieler die Information, welcher Arm das „Duell“ gewonnen hat. Ziel ist es, den besten Arm zu finden und dabei möglichst wenig schlechte Arme zu ziehen.

(23) A Revealing Introduction to Hidden Markov Models (B)

Mark Stamp
Department of Computer Science San Jose State University (2004). http://www.cs.sjsu.edu/faculty/stamp/RUA/HMM.pdf (aktualisierte Version mit kleinen Verbesserungen und zusätzlichen Übungsaufgaben)

Kurzbeschreibung: Eine Einführung in Hidden Markov Models; wie kann man Rückschlüsse auf „versteckte“ Zustände einer Markov-Kette ziehen, wenn man sie nicht direkt beobachten kann? Beispiel: Klimabedingungen in der Vergangenheit lassen sich als Markov-Kette modellieren, deren Übergangsmatrix wir indirekt anhand von Baumringen oder Fossilienfunden rekonstruieren können.