

STUDIEN

Effiziente Algorithmen

für Studenten der Mathematik und Informatik
an der Freien Universität Berlin

Semesterheft Sommer 2001

STUDIEN
SEMESTERHEFT
SOMMER 2001

Allgemeines

Das Gebiet *Effiziente Algorithmen* ist eines der Bindeglieder zwischen Informatik und Mathematik. Einerseits gehören Algorithmen und Datenstrukturen zum Kern der praktischen Informatik, andererseits bezieht die zugrundeliegende Theorie ihre Methoden im wesentlichen aus der diskreten Mathematik. Die Anwendungen reichen in zahlreiche Gebiete wie Computer-Grafik, Mustererkennung, Robotik, Computer Aided Design, Bioinformatik, Kartographie, Bildverarbeitung usw. Einige konkrete Beispiele werden anhand von Diplomarbeiten, Dissertationen und Projekten am Ende dieser Broschüre vorgestellt.

Das Gebiet ist in Berlin an allen drei Universitäten und am Konrad-Zuse-Zentrum stark vertreten. Diese Institutionen tragen gemeinsam das von der Deutschen Forschungsgemeinschaft geförderte Europäische Graduiertenkolleg *Combinatorics, Geometry, and Computation*, das in Zusammenarbeit mit der ETH Zürich durchgeführt wird. (Siehe auch die WWW-Seite:

<http://www.inf.fu-berlin.de/gk-cgc.>)

Neben zahlreichen Lehrveranstaltungen werden auch andere Möglichkeiten zur Weiterbildung angeboten. So finden zum Beispiel der Berliner Algorithmen-Tag oder mehrtägige Spezialschulen regelmäßig statt.

Der Studienschwerpunkt ist ein Vorschlag zur Vertiefung in dieses Fachgebiet im Rahmen der Studiengänge Mathematik und Informatik an der Freien Universität. Zum Beispiel wird eine abgestimmte Folge von Lehrveranstaltungen von den Grundlagen bis zu den Anwendungen angeboten. Darüberhinaus sollen den Studenten die zahlreichen Angebote in dieser Richtung in Berlin besser zugänglich gemacht werden. Dazu wird diese Informationsbroschüre jedes Semester aktualisiert. Neben Vorschlägen zur Studienplanung werden hier allgemeinere Informationen zum Gebiet *Effiziente Algorithmen* zusammengefaßt. Die Broschüre gibt einen Überblick über die Lehrveranstaltungen zum Gebiet – auch an den anderen Berliner Universitäten – und die Lehrveranstaltungsplanung für die folgenden Semester. Zusätzlich gibt sie Informationen zu Tagungen und ähnlichen Veranstaltungen, zu den in der Arbeitsgruppe *Effiziente Algorithmen* tätigen Mitarbeitern und ihren Arbeitsgebieten sowie zu aktuell im Fachbereich behandelten Forschungsthemen (Diplomarbeiten, Dissertationen etc.).

Interessenten können sich im Sekretariat der Arbeitsgruppe Theoretische Informatik (Takustr. 9, Raum 111) in den Studienschwerpunkt einschreiben. Sie

erhalten dann regelmäßig das Semesterheft und werden laufend über Veranstaltungen wie etwa Vorträge, Spezialschulen und Tagungen informiert.

1 Vorschlag zur Studienplanung

Für die Teilnahme am Studienschwerpunkt sind gewisse Grundkenntnisse aus dem Grundstudium unerlässlich. Es wird empfohlen, Vorlesungen und Kurse zu Kombinatorik, Wahrscheinlichkeitstheorie, Graphentheorie, Algorithmen und Datenstrukturen sowie Programmierung frühzeitig zu besuchen. Innerhalb des Hauptstudiums ist sowohl eine Orientierung in mathematisch-theoretischer Richtung als auch eine zu Anwendungsgebieten hin möglich. Im folgenden geben wir Empfehlungen zur Organisation des Studiums für beide Richtungen. Natürlich sind auch Mischformen und eine andere Organisation möglich.

Die Angebote des Hauptstudiums werden im Studienschwerpunkt mit einem Kürzel ([EA 1,2] Entwurf und Analyse von Algorithmen, [ADM] Angewandte Diskrete Mathematik, [ANW] Anwendungen, [PR] Praktikum) gekennzeichnet. Das Kürzel zeigt an, für welche Phase des Studiums die Veranstaltungen geeignet sind. Der Vorschlag sollte individuell durch begleitende Lehrveranstaltungen aus Mathematik und Informatik ergänzt werden.

Vertiefung in theoretischer Richtung.

- [EA1] (3V + 2 Ü) im 5. Semester
Entwurf und Analyse von Algorithmen.
 - [EA2] (3V + 2 Ü) im 6. Semester
Themenbeispiele: *Algorithmische Geometrie* oder *Graphenalgorithmen* oder *Optimierung* oder *Parallele Algorithmen* oder ...
 - [ADM] Zumindest eine Vorlesung und ein Seminar im 7. & 8. Semester.
Themenbeispiele: *Polytope* oder *Pattern Matching* oder *Randomisierte Algorithmen* oder *Graphenalgorithmen* oder ...
- ★ anschließend *Diplomarbeit*.

Vertiefung in anwendungsbezogener Richtung.

- [EA1] (3V + 2 Ü) im 5. Semester
Entwurf und Analyse von Algorithmen.
- [EA2] (3V + 2 Ü) im 6. Semester
Themenbeispiele: *Algorithmische Geometrie* oder *Graphenalgorithmen* oder *Optimierung* oder *Parallele Algorithmen* oder ...
- [ANW] Zumindest eine Vorlesung und ein Seminar im 7. & 8. Semester.
Themenbeispiele aus den Anwendungsgebieten: *Mustererkennung* oder *Computer-Grafik* oder *Computer Aided Design* oder *Robotik* oder *Bildverarbeitung* oder *Bioinformatik* oder ...
- [PR] (4 PR) Praktikum im 8. Semester.
★ anschließend *Diplomarbeit.*

Entsprechend den vorgehenden Vorschlägen sollen Zyklen von einander ergänzenden Veranstaltungen angeboten werden. Am Beginn jedes Zyklus werden den behandelten Themen Kürzel zugewiesen und der Plan im Semesterheft angekündigt.

Die Hauptvorlesung "Entwurf und Analyse von Algorithmen aus dem Wintersemester wird durch "Algorithmen für Fortgeschrittene" weitergeführt. Es gibt eine Reihe von speziellen Vorlesungen über verschiedene geometrische Gebiete: Dynamische Geometrie, Offene Probleme der diskreten Geometrie, und geometrische Approximationsalgorithmen.

Zwei Seminare befassen sich ebenfalls mit geometrischen Fragestellungen: Diskrete und algorithmische Geometrie, sowie Rekonstruktion und Vereinfachung. Ein weiteres Seminar behandelt Algorithmen für das Weltnetz.

Realisierung im Studienplan.

Die Realisierung innerhalb der bestehenden Studiengänge wird für Mathematik und Informatik getrennt behandelt.

Diplomstudiengang Mathematik.

Bei Teilnahme am Studienschwerpunkt wird Studenten der Mathematik die Belegung des Nebenfachs Informatik empfohlen.

- Grundstudium.

Auf jeden Fall sollte die *Elementare Wahrscheinlichkeitstheorie* besucht werden. Ein *Programmierkurs* ist Pflicht im Grundstudium. Die Vorlesungen *Informatik A* und *B* sind Teil der Nebenfachausbildung. Grundkenntnisse in diskreter Mathematik können durch einen möglichst frühen Besuch von *Kombinatorik* und/oder *Graphentheorie* erworben werden.

- Hauptstudium.

[EA1] Anrechnung in A (Angew. Mathematik).

[EA2] & [ADM] Anrechnung in B (Einarbeitung in Spezialgebiet).

[ANW] & [PR] Anrechnung im Nebenfach Informatik.

Seminare: Anrechnung im Rahmen der mindestens 2 Seminare.

Bei der Diplomprüfung aus Angew. Mathematik wird neben Algorithmentheorie auch ein weiteres Gebiet (Wahrscheinlichkeitstheorie oder Numerik) geprüft.

Diplomstudiengang Informatik.

- Grundstudium.

Die Grundkenntnisse werden durch die Lehrveranstaltungen *Algorithmen und Programmierung III* und *Einführung in die Diskrete Mathematik* abgedeckt. Zum Studienschwerpunkt verwandte Veranstaltungen sind auch *Logik für Informatiker* und *Grundlagen der Theoretischen Informatik*.

- Hauptstudium.

[EA1] & [EA2] Anrechnung in Theoretische Informatik.

[ADM] & [ANW] & [PR] Anrechnung im Rahmen der 14 SWS im Studienschwerpunkt.

Seminare: Anrechnung im Rahmen der mindestens 2 Seminare.

2 Lehrveranstaltungen im Sommer 2001

Algorithmen für Fortgeschrittene

[EA2]

Dozent: Alt, Meißner; Vorlesungszeit: Di, Do 10.00 - 12.00 Uhr, 4-stündig.
Veranstaltungsort: Takustraße 9, SR 055.

INHALT: weiterführende Themen aus der Algorithmentheorie, insbesondere Approximationsalgorithmen für NP-schwere Probleme, arithmetische Algorithmen für ggT, Multiplikation und Primzahltest mit Anwendungen in der Kryptographie, RSA-Kryptosystem, probabilistische Algorithmen

Literatur: T. H. Cormen, C. E. Leiserson und R. L. Rivest. Introduction to Algorithms. MIT Press/McGraw-Hill, 1990 und Originalarbeiten

Dynamische Geometrie

[ADM,ANW]

Dozent: Kortenkamp, Knauer; Vorlesungszeit: Di 10.00 - 12.00 Uhr, 2-stündig.
Veranstaltungsort: Takustraße 9, SR 037.

INHALT: In der Vorlesung werden die mathematischen und informatischen Grundlagen dynamischer Geometriesoftware (siehe z.B. Cinderella, <http://www.cinderella.de>) vermittelt und durch eigene Programmierübungen vertieft. Kenntnisse in Linearer Algebra und Funktionentheorie sind vorteilhaft, aber nicht Voraussetzung. Programmierkenntnisse in Java (z.B. durch den Blockkurs Programmieren in Java", s. 19611, oder einer anderen Programmiersprache sind für die Bearbeitung der Übungen praktisch.

Offene Probleme der diskreten Geometrie

[ADM]

Dozent: Braß; Vorlesungszeit: Di 16.00 - 18.00 Uhr, 2-stündig.
Veranstaltungsort: Takustraße 9, SR 055.

INHALT: In dieser Veranstaltung werde ich eine Reihe offener Probleme im Bereich der Diskreten Geometrie vorstellen, die unter anderem Polygone, Polyeder, Zellzerlegungen, Packungen, Überdeckungen, Gitterpunkte und ähnliche Objekte betreffen. Falls Interesse besteht, werde ich auch eine Arbeitsgruppe anbieten, in der wir gemeinsam an einigen dieser Probleme arbeiten werden.

Damit wird exemplarisch eine Einführung in die wissenschaftliche Arbeit gegeben.

Graphentheorie

Dozent: Rote; Vorlesungszeit: Di, Do 12.00 - 14.00 Uhr, 3-stündig.

Veranstaltungsort: Arnimallee 3, SR 119.

INHALT: Für die im WS 2001/02 geplante weiterführende Vorlesung über Graphentheoretische Algorithmen ist der Besuch der Graphentheorie - Vorlesung vorteilhaft. Graphentheorie ist neben der "Einführung in die Diskrete Mathematik" die zweite Grundvorlesung im Bereich Diskrete Mathematik. Viele kombinatorische Aufgaben, Optimierungsaufgaben und algorithmische Probleme lassen sich am besten graphentheoretisch formulieren. In der Vorlesung werden zunächst die grundlegenden Begriffe wie Zusammenhang, Flüsse, Paarungen, Touren, Planarität, Färbungen, vorgestellt. Wo es sich ergibt, wird der algorithmische Aspekt hervorgehoben. Anschließend werde ich einige speziellere ausgewählte Themen behandeln, etwa Zufallsgraphen, Irrfahrten und elektrische Netze.

Literatur: Diestel, Graphentheorie. West, Introduction to Graph Theory. Aigner, Graphentheorie. Bollobás, Modern graph theory: Springer 1998.

Geometrische Approximationsalgorithmen [EA2,ADM]

Dozent: Hoffmann; Vorlesungszeit: Mi 10.00 - 12.00 Uhr, 2-stündig, Beginn 18.04.

Veranstaltungsort: Takustraße 9, SR 039.

INHALT: Die Vorlesung bietet einen Überblick über den Entwurf und die Analyse von Approximationsalgorithmen für geometrische Optimierungsprobleme. Viele dieser auch aus Anwendersicht interessanten Probleme sind algorithmisch schwer (NP-schwer), so dass man nicht auf effiziente algorithmische Lösungen hoffen kann. Ausweg sind dann meistens heuristische Verfahren oder aber Approximationslösungen. Der Vorteil von letzteren ist, dass sie für alle Eingabeinstanzen Lösungen mit nachweisbarem Gütefaktor liefern. Im Mittelpunkt werden neuere Approximationsergebnisse für geometrische Tourenprobleme (wie das bekannte euklidische Travelling Salesman Problem) stehen, an denen auch klar wird, wie algorithmischer 'Fortschritt' von neuen Einsichten über die zu

Grunde liegenden geometrischen Strukturen abhängt. Literatur: D. Hochbaum (ed.), Approximation Problems for NP-Complete Problems, PWS Publishing Company, 1997

Seminar Diskrete und algorithmische Geometrie [EA2,ADM]

Dozent: Felsner; Vorlesungszeit: n.V., 2-stündig. Beginn 23.04.
Veranstaltungsort: Takustraße 9, SR 051.

INHALT: Im Seminar werden klassische und neuere Arbeiten aus dem Bereich der geometrischen Graphen und der Geradenarrangements behandelt. Geometrische Graphen sind Graphen die mit einer nicht notwendig kreuzungsfreien Zeichnung in der Ebene gegeben sind. Ausgehend von der klassischen Extremaltheorie für Graphen werden die in den letzten Jahren entwickelten Grundzüge einer Extremaltheorie für geometrische Graphen nachvollzogen. Und Anwendungen der Theorie bei Untersuchungen von Geradenanordnungen besprochen.

Seminar Rekonstruktion und Vereinfachung von Oberflächen [ADM,ANW]

Dozent: Rote; Vorlesungszeit: Do. 16.00 - 18.00 Uhr, 2-stündig.
Veranstaltungsort: Arnimallee 2-6, Raum 009.

INHALT: Bei der Aufnahme von dreidimensionalen Objekten durch Laserscanner oder ähnliche Geräte entsteht zunächst als Rohdaten ein Punkthaufen oder eine Serie von Punkthaufen. Die ersten Schritte vor der Weiterverarbeitung dieser Daten ist die Rekonstruktion einer Fläche, die durch die Punkte verläuft, sowie gegebenenfalls die Vereinfachung und Reduktion der Datenmenge auf ein vernünftiges Maß. Für diese Aufgaben sind verschiedene Verfahren vorgeschlagen worden, die wir in diesem Seminar kennenlernen wollen.

Literatur: Originalarbeiten.

Seminar Algorithmen für das WWW [ANW]

Dozent: Alt, Knauer; Vorlesungszeit: Di 18.00 - 16.00 Uhr, 2-stündig.
Veranstaltungsort: Takustraße 9, SR 046.

INHALT: Das Internet und insbesondere das WWW wirft eine Vielzahl von neuen algorithmischen Fragen auf. Wir wollen uns in diesem Seminar einigen

davon zuwenden und einen Einblick in aktuelle Forschungsarbeiten in diesem Umfeld gewinnen. Im Einzelnen sollen folgende Themen angesprochen werden:

- o Datenstrukturen für Suchmaschinen
- o Qualitätsbeurteilung für Webseiten
- o Datenkompression
- o Sicherheit im Internet
- o Der WWW-Graph und seine Eigenschaften
- o Routing- und Schedulingprobleme im Internet

Diplomanden- und Doktorandenseminar

Theoretische Informatik

[EA2, ADM]

Dozent: Alt, Rote, Felsner, Braß, Kriegel; Vorlesungszeit: Di, Do, Fr 12-13 Uhr; Veranstaltungsort: Takustraße 9, SR 055;

INHALT: Vorträge über eigene Forschung und Originalarbeiten aus der Theoretischen Informatik, insbesondere Algorithmen. Die Ankündigungen werden jeweils gesondert gegenüber Raum 111 in der Takustraße 9 ausgehängt.

Vorlesung des Europäischen Graduiertenkollegs

Combinatorics, Geometry and Computation

[ADM]

Dozent: Alt, Rote u. Doz. des Kollegs, Vorlesungszeit: Mo 14-16 Uhr, 2-stündig;

Veranstaltungsort: Takustraße 9, SR 005.

INHALT: Die Dozenten und Gäste des Kollegs halten einführende Vorlesungen (in Blöcken von etwa 2-4 Stunden) zu speziellen Themen des Kollegs. Dazu gehören insbesondere algorithmische und diskrete Geometrie, algorithmische Kombinatorik, Codierungstheorie, Graphentheorie und Graphenalgorithmen, Gruppentheorie, kombinatorische Optimierung, konstruktive Approximation, Mustererkennung und zufällige diskrete Strukturen. Die Themen der Vorlesungen werden auch neben Raum 111 in der Takustraße 9 ausgehängt. (Interessenten können sich bei der Koordinatorin des Kollegs, Frau B. Felsner, auf einen Verteiler für das Verschicken der Ankündigungen setzen lassen.)

Kolloquium des Europäischen Graduiertenkollegs
Combinatorics, Geometry, and Computation

[ADM]

Dozent: Alt u. Doz. des Kollegs; Vorlesungszeit: Mo 16-18 Uhr, 2stündig;
Veranstaltungsort: Takustraße 9, SR 005;

INHALT: Mitglieder und Dozenten des Kollegs sowie Gäste halten Vorträge zum Thema des Kollegs. Die einzelnen Vorträge im Kolloquium werden gesondert angekündigt. (Interessenten können sich bei der Koordinatorin des Kollegs, Frau B. Felsner, auf einen Verteiler für das Verschicken der Ankündigungen setzen lassen.) Die Ankündigungen werden auch neben Raum 111 in der Takustraße 9 ausgehängt.

Die Vorlesungen des Graduiertenkollegs im Sommer 2001

- 23. April 2001
DORIT HOCHBAUM, UNIVERSITY OF CALIFORNIA, BERKELEY:
The Pseudoflow Algorithm for the Maximum Flow Problem.
- 07. Mai 2001
JOHANNES BUCHMANN, UNIVERSITÄT DARMSTADT:
How secure is public key cryptography?.
- 21. Mai 2001
ALEXANDER SCOTT, UNIVERSITY COLLEGE LONDON:
Reconstructing subsets of the plane.
- 28. Mai 2001
MICHIEL SMID, UNIVERSITÄT MAGDEBURG:
Translating a planar object to maximize point containment: Exact and approximation algorithms.
- 11. Juni 2001
CLIFFORD STEIN, DARTMOUTH COLLEGE:
Scheduling to simultaneously optimize two criteria.
- 18. Juni 2001
VOLKER STRASSEN, UNIVERSITÄT KONSTANZ:
Degeneration of linear and bilinear maps.

- 25. Juni 2001

GIL KALAI, JERUSALEM:
tba.

- 02. Juli 2001

LEEN STOUGIE, TECHNISCHE UNIVERSITEIT EINDHOVEN:
Approximation and exact algorithms for the minimum test set problem.

- 09. Juli 2001

JOHANNES BLÖMER, UNIVERSITÄT PADERBORN:
Grids and Cryptography.

- 16. Juli 2001

JEONG HAN KIM, MICROSOFT RESEARCH SEATTLE:
Survey on the nibble method.

Die Kolloquiumsvorträge des Graduiertenkollegs im Sommer 2001

- 23. April 2001

MARC UETZ, TECHNISCHE UNIVERSITÄT BERLIN:
Project Scheduling and Minimum Cuts

- 07. Mai 2001

BRITTA BROSER, FREIE UNIVERSITÄT BERLIN:
Time Series Analysis.

- 21. Mai 2001

SHI LINGSHEN, HUMBOLDT-UNIVERSITÄT ZU BERLIN:
Cube Ramsey Numbers are Polynomial.

- 28. Mai 2001

ILEANA STREINU, SMITH COLLEGE NORTHAMPTON:
Non-Colliding Robot Arm Motions.

- 11. Juni 2001

VANESSA KÄÄB, TECHNISCHE UNIVERSITÄT BERLIN
Critical Sets in AND/OR-Networks.

- 18. Juni 2001

LISA FLEISCHER, CARNEGIE MELLON UNIVERSITY, PITTSBURGH:
tba.

- 25. Juni 2001

JÜRGEN RICHTER-GEBERT, ETH ZÜRICH:
Pappos Theorem — ten proofs and three variations.

- 02. Juli 2001

DIANA POENSGEN, KONRAD-ZUSE-ZENTRUM FÜR INFORMATIONSTECHNIK
BERLIN:
Online Call Admission in Optical Networks.

- 09. Juli 2001

OTFRIED CHEONG, UNIVERSITEIT UTRECHT:
Paths with bounded curvature.

- 16. Juli 2001

MARTIN THIMM, HUMBOLDT-UNIVERSITÄT ZU BERLIN:
On the Approximability of the Steiner Tree Problem.

Weitere Veranstaltungen an der Freien Universität

- Parallele Algorithmen in der Bioinformatik (VL); Dozent: Reinefeld, Steinke.
- Spatial Databases (VL); Dozent: Voissard.
- Bildverarbeitung (VL); Dozent: Rojas

Europäisches Graduiertenkolleg
Berlin ————— **Zürich**
Combinatorics, Geometry, and Computation

In Berlin und Zürich sind Stellen für

Doktorand(inn)en

mit überdurchschnittlichem Studienabschluss in Mathematik, Informatik oder einem verwandtem Gebiet für zwei Jahre zu vergeben.

Das Kolleg ist eine gemeinsame Initiative der ETH Zürich, der drei Berliner Universitäten – Freie Universität, Technische Universität und Humboldt-Universität –, und dem Konrad-Zuse-Zentrum für Informatik Berlin. Die deutschen Partner finanzieren sich aus Mitteln der Deutschen Forschungsgemeinschaft (DFG).

Die Stipendien in Berlin werden nach den Richtlinien der DFG bemessen und betragen monatlich bis zu DM 2.870,- steuerfrei (Familienzuschlag DM 400,-). Das wissenschaftliche Programm reicht von theoretischen Grundlagen bis hin zu Anwendungen. Die Forschungsgebiete sind Kombinatorik, Geometrie, Optimierung, Algorithmen und Berechnung, sowie Computer Graphik und Vision.

Am Studienort Berlin werden die Stipendiaten von den Professoren Aigner, Alt, Rote und Schulz (FU), Möhring, Ziegler (TU), Prömel (HU) und Grötschel (ZIB) betreut.

Zu den Bewerbungsunterlagen der Doktorand(inn)en gehören Lebenslauf, Zeugniskopien, Examensarbeit, Gutachten des letzten Betreuers und Vorstellungen zum gewünschten Promotionsvorhaben. Alle Unterlagen sind bei dem Sprecher des Kollegs in Berlin einzureichen. Sprecher des Kollegs

Prof. Dr. Helmut Alt
Institut für Informatik
Freie Universität Berlin
Takustraße 9
D-14195 Berlin

b.w.

Weitere Informationen sind erhältlich bei
Bettina Felsner
Tel. ++49-30-838 75 104
bfelsner@inf.fu-berlin.de

Internet: <http://www.inf.fu-berlin.de/gk-cgc>

3 Diplomarbeiten, Dissertationen, Projekte

Die von der Arbeitsgruppe Theoretische Informatik behandelten Forschungsthemen werden an Diplomarbeiten, Dissertationen und Projekten beispielhaft vorgestellt.

Diplomarbeit: Zufällige Erzeugung von Catalan-Objekten

Diplomandin: *Dorothea Rochusch*, Betreuer: *Stefan Felsner*

An vielen Stellen wird immer wieder festgestellt, daß deterministische Algorithmen kompliziert sind und lange brauchen, um zu einem Ergebnis zu kommen, während randomisierte Algorithmen oft mit einem sehr einfachen Programm in kurzer Zeit hinreichend gute Ergebnisse liefern. Ein Beispiel dafür ist der oft betrachtete Primzahltest. Aus diesem Grund ist man daran interessiert, immer kompliziertere Strukturen zufällig erzeugen zu können, wenn nur von einem Zufallsgenerator für ganzzahlige Werte bzw. rationale Zahlen in einem festen Bereich ausgegangen werden kann.

In dieser Arbeit werden verschiedene Algorithmen zur zufälligen Erzeugung von Catalan-Objekten vorgestellt und analysiert. Das sind zum Beispiel die binären Bäume mit n inneren Knoten oder die Triangulierungen eines konvexen $(n + 2)$ -gons. Die betrachteten Algorithmen bauen auf sehr verschiedenen Ansätzen auf, wie Markov-Ketten, Ranking-Funktionen oder schrittweise Erzeugung der gewünschten Struktur. Vor allem die Analysetechniken für Markov-Ketten sind allgemeinerer Natur und lassen sich nicht nur bei diesem Problem anwenden.

Sicherlich können nicht alle gleichermaßen in der Praxis eingesetzt werden, den-

noch liefern sie einen guten Überblick über die Möglichkeiten, Algorithmen zur zufälligen Erzeugung festgelegter Strukturen aufzubauen, und zu analysieren.

Diplomarbeit: Stückweise lineare Approximation von Kurven

Diplomand: *Oliver Timm*, Betreuer: *Helmut Alt*

In der Computergraphik werden häufig graphische Objekte durch Kurven

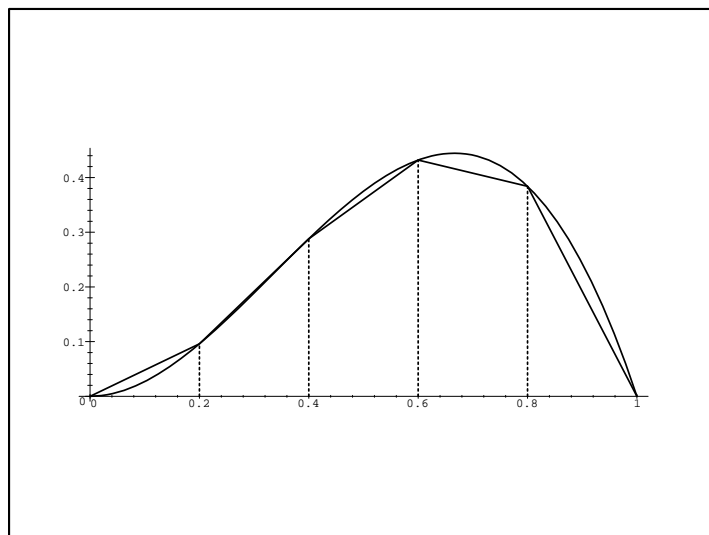


Abbildung 1: Stückweise lineare Approximation eines x-monotonen Polygonzuges in einer vorgegebenen Fehlertoleranz

(z.B. in parametrisierter Form) repräsentiert. Meist erwarten viele graphische Applikationen auch die Darstellung in stückweiser linearer Form (Polygonzug). Daher ist es wichtig, die erste Form durch die zweite effizient approximieren zu können.

In einem anderem Fall liegt eine Kurve in stückweiser linearer Form und mit einer großen Anzahl von Kanten bereits vor. Um die Operationen auf dieser Kurve zu beschleunigen und um Speicherplatz zu sparen, soll diese durch eine weitere stückweise lineare Kurve mit weniger Kanten approximiert werden, ohne daß dabei die geometrische Form der Kurve verloren geht.

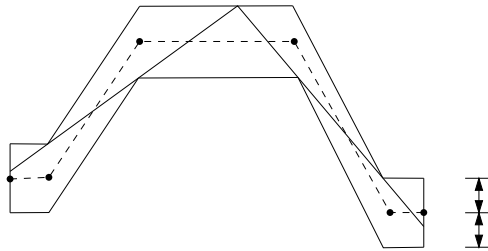


Abbildung 2: Stückweise lineare Approximation einer parametrisierten Kurve

Im allgemeinen geht es um die Lösung folgender zwei Probleme:

1. Minimalisierung der Kantenanzahl, ohne dabei eine vorgegebene Fehler-toleranz zu überschreiten
2. Minimalisierung des Approximationsfehlers, bei vorgegebener Kanten-zahl

In der Diplomarbeit sollen die verschiedenen Methoden der stückweisen linearen Approximation vorgestellt und analysiert werden. Hier soll bezüglich der Art (parametrisiert/stückweise linear) und der Eigenschaft (x-monoton/konvex/ offen/ geschlossen) der Kurven unterschieden und verschiedene Abstandsmaße zur Fehlerermittlung herangezogen werden.

Dissertation: Nächste-Nachbar-Suche in hohen Dimensionen und Externe Algorithmen

Doktorandin: *Laura Heinrich-Litan*, Betreuer: *Helmut Alt*

Für viele Anwendungen, wie zum Beispiel Ähnlichkeitsanfragen in multimedialen Datenbanken, Mustererkennung, Data Mining und Video Kompression, werden Methoden benötigt, die zu einer gegebenen Menge von Punkten $P \subset \mathbb{R}^d$ und einem spezifizierten Punkt $q \in \mathbb{R}^d$, effizient den oder die nächsten Nachbarn aus P zu q finden. Der Abstand wird in einer der Minkowski-Metriken $L_1, L_2, \dots, L_\infty$ definiert. Diese Suche nennt man Nächste-Nachbar-Suche, und das Problem, dafür effiziente Datenstrukturen und Algorithmen zu entwickeln, ist das Nächste-Nachbar-Problem. Bei vielen Methoden oder Datenstrukturen, die für das Nächste-Nachbar-Problem entwickelt wurden, wird angenommen, daß die Dimension d eine kleine Konstante ist. Die Laufzeit, der Speicherbedarf oder die Vorverarbeitungszeit dieser Methoden sind exponentiell in der Dimension d . In vielen Anwendungen ist aber die Dimension d des Suchraumes sehr groß und deswegen verbietet sich eine in d exponentielle Laufzeit. Im Prinzip konkurrieren alle Algorithmen mit der naiven *Bruteforce-methode*, die $\Theta(nd)$ Zeit für Minkowski-Metriken kostet, keine Vorverarbeitung und nur Speicherplatz für P benötigt. Die *Boxmethode* (entwickelt von Hoffmann und Alt 1998) bestimmt aus einer gegebenen Menge P von Punkten im reellen Einheitswürfel $[0, 1]^d$ den nächsten Nachbarn zu einem Punkt $q \in [0, 1]^d$ bezüglich der durch die Maximumnorm induzierten Metrik L_∞ . Der Algorithmus hat unter der Voraussetzung, daß die Punkte aus P und der spezifizierte Punkt q gleichverteilt aus $[0, 1]^d$ sind, die erwartete Laufzeit $O(\frac{nd}{\ln n})$. Varianten der Boxmethode habe ich für das k -Nächste-Nachbar-Problem, zur Bestimmung der k nächsten Nachbarn zu q , erweitert und analysiert. Zukünftig werde ich die Boxmethode bei Zugrundelegung von anderen Wahrscheinlichkeitsverteilungen untersuchen. Von Interesse bleibt auch eine experimentelle Untersuchung, ob diese Methoden in praktischen Anwendungen, wie zum Beispiel bei der Datenbanksuche, sinnvoll einsetzbar sind.

Bei der Entwicklung von Algorithmen für Anwendungen mit sehr großen Datenmengen (large-scale applications), die nicht in den Hauptspeicher passen, ist es wichtig die Input/Output (I/O) Kommunikation zwischen Hauptspeicher und dem langsameren, sekundären Speicher zu minimieren. Geographische In-

formationssysteme (GIS) und Datenbanksysteme sind gute Beispiele für solche Anwendungen, die immense Datenmengen verarbeiten. Externe Algorithmen (I/O Algorithmen) werden für das theoretische 2-Schichten-Berechnungsmodell entwickelt, wobei das Speichersystem aus einem Hauptspeicher und einer Anzahl von externen Speichergeräten (Disks) besteht. Die Kommunikation erfolgt blockweise, d.h. die Daten werden in Blöcken von zusammenhängenden Daten übertragen. Das von Aggarwal und Vitter eingeführte I/O-Berechnungsmodell ist durch folgende Parameter spezifiziert : M die Anzahl von Elementen (items), die in den Hauptspeicher passen, N die Anzahl der Elemente der Problem Instanz, B die Anzahl von Elementen pro Block und D die Anzahl der Disks, wobei $B \geq 1$ und $B \leq M < N$. Eine I/O-Operation ist eine Lese- oder Schreiboperation eines Blockes aus dem oder in den sekundären Speicher. Ziel ist die Entwicklung solcher Algorithmen, welche sowohl die Anzahl von I/O-Operationen als auch die Anzahl von internen Operationen (die CPU-Zeit) minimieren. Mein Interesse liegt in der Entwicklung eines effizienten, externen Algorithmus für das Nächste-Nachbar Problem in hohen Dimensionen.

Dissertation: Methoden der algorithmischen Geometrie in der Mustererkennung

Doktorand: *Christian Knauer*, Betreuer: *Helmut Alt*.

Die algorithmische Geometrie beschäftigt sich mit dem Entwurf und der Analyse von effizienten Datenstrukturen und Algorithmen zur Lösung von geometrisch motivierten Problemstellungen. Typische Beispiele für solche Problemstellungen sind etwa die Berechnung der konvexen Hülle einer Punktmenge oder die Berechnung aller Schnittpunkte einer Menge von Liniensegmenten.

Auch in der Mustererkennung trifft man auf solche geometrischen Fragestellungen: Zu zwei Mengen von Liniensegmenten in der Ebene, M (dem *Muster*) und B (dem *Bild*) der Kardinalität $m = \#M$ und $b = \#B$ suchen wir in einer Menge von *zulässigen Transformationen* \mathcal{T} (z.B. Translationen oder starre Bewegungen) eine Abbildung t , die $\tilde{\delta}_H(t(M), B)$, den *einseitigen Hausdorff-Abstand* von $t(M)$ nach B , minimiert. Dabei ist $\tilde{\delta}_H(t(M), B)$ der größte Abstand, den ein Punkt aus $t(M)$ zu seinem nächsten Nachbarn in B hat

$$\tilde{\delta}_H(t(M), B) = \max_{p \in t(M)} \min_{q \in B} \|p - q\|,$$

also ein Maß dafür, wie „ähnlich“ $t(M)$ zu einem Teilmuster von B ist.

Mit zunehmender Anzahl der Freiheitsgrade der zulässigen Transformationen wird auch die Lösung des Problems aufwendiger. So kann der einseitige Hausdorff-Abstand von M nach B auf einer arithmetischen RAM noch in $\mathcal{O}((b+m)\log(b+m))$ Schritten (uniformes Kostenmaß) berechnet werden ($\mathcal{T} = \emptyset$), während die Laufzeit des derzeit asymptotisch besten Algorithmus zum Auffinden einer optimalen Translation bereits von der Größenordnung $\mathcal{O}((bm)^2 \log^3 bm)$ ist.

Obwohl die Fragestellung sehr stark aus der Anwendung heraus motiviert ist, und die theoretischen Grundlagen bereits weitgehend vorhanden sind, fehlen Implementierungen der entwickelten Verfahren, mit deren Hilfe die Praktikabilität dieser Algorithmen beurteilt werden könnte. Die Realisierung der Algorithmen wird im wesentlichen durch die folgenden Probleme erschwert:

- Oftmals kommen komplexe Datenstrukturen (z.B. Voronoi-Diagramme von Liniensegmenten) und Methoden (z.B. parametrische Suche) zum Einsatz.
- Geometrische Daten müssen durch Zahlentypen beschränkter Genauigkeit im Rechner repräsentiert werden; dem gegenüber steht das idealisierte Rechnen mit reellen Zahlen - eine Annahme, die beim Entwurf der meisten Algorithmen zugrunde gelegt wird.

Dies wirft eine Reihe von Fragen auf, denen man bei der Implementierung von geometrischen Algorithmen immer wieder begegnet, die also nicht unbedingt spezifisch für aus der Mustererkennung motivierte Anwendungen sind:

- Können asymptotisch gute Verfahren auch effizient in die Praxis umgesetzt werden?
- Lohnt es sich, an Stelle von asymptotisch optimalen Verfahren, „schlechtere“, aber konzeptionell einfachere - und damit auch leichter zu implementierende - Methoden zu verwenden?
- Lohnt es sich, an Stelle von exakten Verfahren, approximative, aber schnellere und konzeptionell einfachere Methoden zu verwenden (z.B. Referenzpunktmethoden)?

- Mit welcher Genauigkeit muß gerechnet werden, um die topologische Korrektheit des Endresultates garantieren zu können, bzw. um eine vorgegebene Fehlerschranke nicht zu überschreiten?

Dissertation: Datenstrukturen zum Auffinden von Formen

Doktorand: *Lutz Meißner*, Betreuer: *Helmut Alt*.

Die Menge aller Polygonzüge in der Ebene wird, etwa durch den Hausdorff- oder den Fréchet-Abstand, zu einem metrischen Raum. Von besonderem Interesse sind die Räume, bei denen zur Abstandsmessung die einzelnen Polygonzüge “verschoben” werden können:

$$\delta(P, Q) = \min_{t \in \mathbf{R}^2} \tilde{\delta}(P, Q + t)$$

Es läßt sich nun, bei gegebenen Polygonzügen P_1, P_2, \dots, P_n , die Frage stellen, welches dieser P_i einem weiteren Polygonzug P am “ähnlichsten” ist:

$$\text{NN}(P) = \{P_i | \delta(P, P_i) \leq \delta(P, P_j) \text{ für } j = 1, \dots, n\}$$

Gesucht ist also der nächste Nachbar (oder die nächsten Nachbarn) von P .

Ist man nicht nur an der Bestimmung des nächsten Nachbarn eines, sondern vieler Polygonzüge interessiert, ist es sinnvoll, P_1, P_2, \dots, P_n in einer Datenstruktur zusammenzufassen, um eine effiziente Abfrage zu ermöglichen. Aber wie muß diese Datenstruktur aussehen?

Für die Nächste-Nachbar-Suche von Punkten im \mathbf{R}^d sind effiziente Datenstrukturen bekannt. Diese nutzen jedoch Eigenschaften, etwa die endliche Dimensionalität oder die Vektorraum-Eigenschaften, aus, die bei den Polygonzügen nicht vorhanden sind.

Bei den bekannten Methoden in allgemeinen “großen metrischen Räumen” werden Voraussetzungen sowohl an die zugrundeliegende Metrik als auch an die Verteilung der Datenmenge gestellt, die hier nicht oder nur bedingt zutreffen. Weiter gibt es kaum Aussagen zu den erwarteten Laufzeiten.

Die Berechnung des Abstands zweier Polygonzüge ist zeitaufwendig. Es stellt sich die Frage, ob zur Bestimmung von $NN(P)$ tatsächlich P mit einzelnen P_i verglichen werden muß, oder ob ein effizienteres Vorgehen möglich ist.

Obwohl zunächst die theoretischen Aspekte untersucht werden, wird auch an die Implementierung von entwickelten Ansätzen gedacht.

Dissertation: Geometrische Mustererkennung in höheren Dimensionen

Doktorandin: *Carola Wenk* Betreuer: *Helmut Alt*.

Seien A und B zwei triangulierte Flächen im Raum. Wie ähnlich sind sich A und B ? Als erstes muß ein Abstandsmaß gewählt werden. Dabei kommt die Fréchet-Metrik nicht in Frage, da, wie Michael Godau in seiner Dissertation gezeigt hat, schon das Entscheidungsproblem NP-schwer ist. Deshalb betrachten wir zunächst den Hausdorff-Abstand, der für $n = \max(|A|, |B|)$ in $O(n^3 \log^3 n)$ berechnet werden kann.

In der Regel soll es aber erlaubt sein, aus einer Menge von Abbildungen, z.B. Translationen, diejenige herauszusuchen, die den Hausdorff-Abstand minimiert. Eine derartige Problemstellung gestaltet sich in höheren Dimensionen, d.h. in Dimensionen ≥ 3 , weitaus schwieriger. Betrachten wir das Entscheidungsproblem, ob es für ein gegebenes ϵ eine Translation t gibt, so daß der gerichtete Hausdorff-Abstand $d_H(A + t, B) \leq \epsilon$ ist. Das ist aber äquivalent dazu, daß $t \in S_\epsilon := \bigcap_{a \in A} \overline{B^\epsilon} \oplus (-a)$, wobei B^ϵ die ϵ -Umgebung von B , ein waagerechter Strich das Komplement und \oplus die Minkowskisumme bezeichnen. Die Berechnung von S_ϵ hängt unter anderem davon ab, $B^\epsilon = \bigcup_{b \in B} b^\epsilon$ effizient zu berechnen. Da B^ϵ als ein ebener Schnitt der Voronoi-Fläche von B an Höhe ϵ aufgefaßt werden kann, hängt eine effiziente Berechnung von B^ϵ direkt mit der (kombinatorischen) Komplexität der Voronoi-Fläche, und damit auch mit dem Voronoi-Diagramm von B zusammen. Für Voronoi-Diagramme von Mengen, die statt Punkten Strecken oder Dreiecke enthalten, gibt es bisher wenige nicht-triviale Ergebnisse. Die triviale Annahme, daß die Komplexität von B^ϵ in d Dimensionen $O(n^d)$ ist, beschränkt die von S_ϵ auf $O(n^{d^2+d})$. Für Mengen von Objekten, für die sich die Vermutung, daß B^ϵ eine Komplexität von $O(n^{\lceil \frac{d}{2} \rceil})$ hat bestätigt, kann S_ϵ eine Komplexität von $O(n^{\lceil \frac{d}{2} \rceil d+d})$ haben.

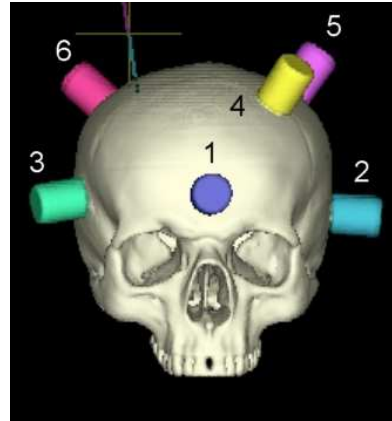
Wenn A nur aus Punkten besteht, geht die Dimension nicht mehr quadratisch sondern nur noch linear im Exponenten ein. Untere Schranken gibt es in höheren Dimensionen bisher nur für Spezialfälle.

Von theoretischem Interesse ist es nun herauszufinden, ob es doch eine Möglichkeit gibt, weniger grobe Abschätzungen zu verwenden. So wurde in einer neuen Arbeit von Agarwal und Sharir für Strecken und sich nicht überschneidende Dreiecke im \mathbb{R}^3 unter Verwendung der euklidischen Metrik gezeigt, daß B^ϵ eine Komplexität von $O(n^{2+\delta})$ für beliebiges $\delta > 0$ hat, was fast der Vermutung $O(n^2)$ entspricht. Weiterhin sind auch andere Klassen von Abbildungen, wie z.B. starre Bewegungen, von Interesse, jedoch aufgrund der vielen Freiheitsgrade gerade in höheren Dimensionen schwierig zu handhaben. Für die Praxis stellt sich die Frage nach Approximationsmöglichkeiten oder einfacheren Spezialfällen, in denen die Laufzeit der Algorithmen überschaubarer ist.

Projekt: Registrierung in der computerunterstützten Chirurgie

Frank Hoffmann, Klaus Kriegel, Sven Schönherr, Carola Wenk.

Dieses Projekt wird in Kooperation mit der Firma Functional Imaging Technologies GmbH bearbeitet. Es geht dabei um Probleme der computergestützten Navigation bei neurochirurgischen Eingriffen. Ein präoperativ erstelltes 3-dimensionales Modellbild zeigt dem Operateur nicht nur die Position des zu entfernenden Tumors, sondern auch mögliche Wege dorthin an, die funktionelle Zentren im Gehirn umgehen. Die Hauptaufgabe des Navigationssystems besteht darin, während der Operation die jeweils aktuelle Position des Eingriffs in dieses Modellbild zu projizieren. Die Berechnung dieser Transformation basiert auf Markierungen, die auf der Kopfhaut befestigt sind. Diese Markierungen sind zum einen im Modellbild zu erkennen, zum anderen können sie unmittelbar vor der Operation durch ein sogenanntes Trackingsystem eingemessen werden. Mathematisch führt dies zu einem geometrischen Matching-Problem (Registrierung). Erschwert wird diese zunächst leicht erscheinende Aufgabe durch Messfehler, größere Probleme entstehen durch Verschiebungen oder das Fehlen von Markierungen. Das entwickelte Verfahren wird erfolgreich am Universitätsklinikum Benjamin Franklin bei Operationen eingesetzt.



Projekt: Point-Pattern-Matching zur Analyse von Gelbildern

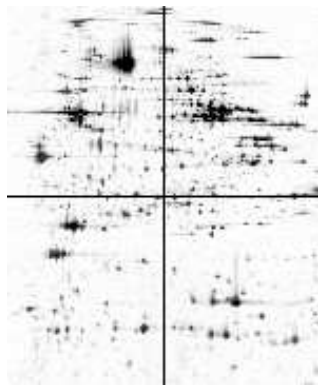
Helmut Alt, Frank Hoffmann, Klaus Kriegel, Christof Schultz, Carola Wenk.

Dieses von der Deutschen Forschungsgemeinschaft geförderte Projekt wird gemeinsam vom Institut für Informatik der FU Berlin und dem Deutschen Herz-

zentrum Berlin bearbeitet. Gegenstand der Untersuchung sind 2-dimensionale Gelbilder, die durch hochauflösende Gelelektrophorese-Techniken erzeugt werden.

Die 1975 von O'Farrell eingeführte Gelelektrophorese hat sich als die zentrale molekularbiologische Methode zur Analyse der Protein/DNA-Zusammensetzung von Gewebeproben etabliert. Jeder Punkt ("Spot") in einem so erzeugten Gelbild repräsentiert ein in der Probe auftretendes Protein. Das Original des hier verkleinert abgebildeten Herz-Ventrikel-Gels enthält ca. 2000 Spots. Ziel der Analyse der Bilder ist es insbesondere, krankheitsassoziierte Proteinausprägungen zu erkennen. Zwar ist es möglich, einzelne Proteine durch Sequenzierung zu bestimmen, dies ist aber sehr teuer und aufwendig und bei der Menge der Daten nicht realistisch. Deshalb basiert bisher die Auswertung der Gelbilder zu großen Teilen auf der genauen (und zeitaufwendigen) Betrachtung durch erfahrene Spezialisten.

Zielstellung des Projektes ist es, Algorithmen zur effektiven computergestützten Gelanalyse zu entwerfen, zu implementieren und im Internet zur Verfügung zu stellen. Dabei stellt das Matching eine der wesentlichen und auch zeitaufwendigsten Voraussetzungen für die quantitative und qualitative Datenanalyse von Protein-Gelbildern dar.



Die entsprechende algorithmische Problemstellung ist eine Variante der 2-dimensionalen Mustererkennung, wobei die besondere Schwierigkeit durch die geometrischen Verzerrungen in den Bildern gegeben ist. Durch die Komplexität der Gelelektrophorese an sich, sind auch von ein und derselben Probe in einem Labor keine zwei identischen Bilder zu erwarten. Die zu entwickelnden Verfahren müssen daher sehr robust sein, um auch den Vergleich von Bildern aus verschiedenen Datenbanken zu ermöglichen. Mit

dem Programmsystem CAROL wurde eine erste Version dieses Analysetools realisiert (vgl. <http://gelmatching.inf.fu-berlin.de>). Im Rahmen einer Diplomarbeit ist vor kurzem die Phase der Spotdetektion, d.h. die Gewinnung einer geometrischen Bildbeschreibung aus den Pixel-Daten algorithmisch neu gelöst und ins CAROL-System eingebunden worden.

4 Die Arbeitsgruppe Theoretische Informatik

Leiter

- PROF. DR. HELMUT ALT
Entwurf und Analyse von Algorithmen, insbesondere Algorithmische Geometrie mit Schwerpunkt Formanalyse.
- PROF. DR. GÜNTER ROTE
Algorithmische Geometrie, kombinatorische Optimierung.

Mitglieder der Arbeitsgruppe

- PD DR. PETER BRASS
Kombinatorische Geometrie, Diskrete Geometrie, Konvexgeometrie, extremale Kombinatorik.
- DARKO DIMITROV
Bildverarbeitung, Computersehen, Flächenrekonstruktion aus dreidimensionalen Punktdaten.
- PD DR. STEFAN FELSNER
Algorithmen für Halbordnungen und Graphen, Algorithmische Geometrie, Kombinatorik.
- LAURA HEINRICH-LITAN
Algorithmische Geometrie, Externe Algorithmen, Nächste-Nachbar- Suche.
- DR. FRANK HOFFMANN
Algorithmische Geometrie, geometrische Online-Probleme, angewandte Matching-Probleme.
- ASTRID KAFFANKE
Algorithmische Geometrie, Graphentheorie.
- CHRISTIAN KNAUER
Algorithmische Geometrie, Implementierung von geometrischen Algorithmen, Ähnlichkeitsbestimmung von polygonalen Figuren.
- DR. ULRICH KORTENKAMP

Dynamische Geometrie, Orientierte Matroide, Nachbarschaftliche Polytope, Java.

- PD DR. KLAUS KRIEGEL

Graphalgorithmen und graphentheoretische Methoden für geometrische Probleme.

- LUTZ MEISSNER

Algorithmische Geometrie, Implementierung von geometrischen Algorithmen.

- CAROLA WENK

Algorithmische Geometrie, Shape Matching.

Weitere Informationen

Prof. Dr. Helmut Alt

Takustr. 9

Raum 112

Tel.: 838-75160

alt@inf.fu-berlin.de

Prof. Dr. Günter Rote

Takustr. 9

Raum 110

Tel.: 838-75150

rote@inf.fu-berlin.de

Dr. Stefan Felsner

Takustr. 9

Raum 117

Tel.: 838-75161

felsner@inf.fu-berlin.de