

Suchen von Mustern in Graphen und Texten

verantwortlich: Dr. André Schulz

Ziel des Seminars ist es grundlegende Techniken und Probleme beim Suchen von Mustern in Graphen und Texten zu behandeln. Als Muster bezeichnet man in diesem Zusammenhang einen Teilgraphen, welchen man im Suchgraphen finden möchte. Dieses Problem (Subgraphisomorphie) ist ein NP-schweres Problem und im Allgemeinen nicht effizient zu lösen. Dennoch existieren für einfache Graphklassen (Bäume, planare Graphen, Graphen mit beschränkter Baumbreite, ...) Algorithmen mit polynomieller Laufzeit. Für den allgemeinen Fall stehen Heuristiken zur Verfügung. Im Seminar werden ausgewählte Algorithmen zum effizienten Lösen dieses Problems besprochen. Zusätzlich wird auf seine Komplexität eingegangen.

Das Suchen in Texten stellt einen Spezialfall dar (in diesem Fall ist der Suchgraph ein Pfad). Einmalige Suchen können mit den Algorithmen von Knuth-Morris-Pratt bzw. von Rabin-Karp effizient gelöst werden. Für das mehrmalige Suchen empfiehlt es sich eine Datenstruktur (Suffixbaum, Suffixarray) zu benutzen. Ein Teil des Seminars widmet sich dem effizienten Aufbau dieser Datenstrukturen.

Das Seminar richtet sich an Studenten der Informatik, Mathematik und Wirtschaftsinformatik im Hauptstudium (bzw. Master-Studenten). Interessierte Studenten im Grundstudium (bzw. Bachelor-Studenten) können nach Absprache teilnehmen. Es handelt sich um ein Master-Ergänzungsmodulseminar, für welches 6 ECTS vergeben werden. Zum erfolgreichen Bestehen ist ein Vortrag und eine Ausarbeitung erforderlich.

Die Vorbesprechung findet am Freitag, dem **7.05.2010 von 12:00 bis 14:00 Uhr im Seminarraum N1** statt. Die Termine für die Vorträge werden nach Absprache festgelegt. Die Teilnehmerzahl des Seminars ist auf 14 Teilnehmer beschränkt. Jeder Teilnehmer sollte sich mit einer kurzen Email an schulz@csail.mit.edu anmelden.

Fragen können jeder Zeit an schulz@csail.mit.edu gerichtet werden. Aktuelle Informationen sind unter <http://people.csail.mit.edu/schulz/teaching.html> verfügbar.

Literatur (Auswahl)

- (1) H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*. MIT Press.
- (2) Jörg Flum, Martin Grohe. *Parameterized Complexity Theory*. Springer, 2006.
- (3) D. Eppstein. Subgraph isomorphism in planar graphs and related problems. *J. Graph Algorithms and Applications* 3(3):1-27, 1999. arXiv:cs.DS/9911003
- (4) J. R. Ullmann. An algorithm for subgraph isomorphism. *J. Assoc. Comput. Mach.* 23:31-42 1976.
- (5) Noga Alon and R. Yuster and U. Zwick. Color-coding: a new method for finding simple paths, cycles and other small subgraphs within large graphs. *Proc. 26th ACM Symp. Theory of Computing*, 1995.
- (6) Juha Kärkkäinen, Peter Sanders, und Stefan Burkhardt. Linear work suffix array construction. *J. ACM* 53(6):918-936, 2006.