

Gedächtnisprotokoll: Diplomprüfung theoretische Informatik

Kurs : 1824 Parallele Algorithmen (Version SS 97)

Prüfer : Dr. Damaschke

Termin : Juni 99

Dauer : 30 Minuten

Themen:

1. Sortieren:

Wie ist die untere Schranke beim Sortieren? Schranke für Laufzeit genannt und Begründung dafür erläutert.

Was ist der Unterschied zwischen Sortieren und lexikografischem Sortieren? Unterschied in Hinblick auf Vergleichbarkeit genannt.

Wieso kann lexikografisch schneller als mit der oben genannten unteren Schranke sortiert werden? Unterschiede in Hinblick auf die Vergleiche noch einmal zu diesem Aspekt erläutert.

Welchen parallelen Sortieralgorithmus haben wir im Kurs kennen gelernt? Mergesort, im besonderen den Merge-Schritt.

Wie arbeitet der Merge-Schritt mit welchen Zeitschranken? Die Herleitungen aus dem Skript der Reihe nach erläutert.

2. Konvexe Hülle

Was ist ein LOP? Genauen Aufbau und die Verbindung zu Schnittmenge von Halbebenen erläutert.

Wie kann man es parallel lösen? Verbindung vom Schnitt der Halbebenen zu konvexen Hüllen dargestellt. Algorithmus und Laufzeit zur Berechnung der konvexen Hülle dargestellt.

3. Polynomoperationen

Wie kann die Polynommultiplikation parallel berechnet werden? Konvolution erläutert und den Zusammenhang zur Polynommultiplikation erklärt. Parallele Berechnung der Polynommultiplikation über FFT erklärt. Dazu auch Aufbau der Fouriermatrix und Begründung, warum hier die Fouriertransformation das gewünschte Ergebnis ergibt.

Wie sind Laufzeit und Aufwand? Werte hergeleitet.

Wie berechnet sich die Polynomdivision? Idee mit Matrixoperationen erklärt. Laufzeit und Aufwand erläutert.

Wie schnell und warum kann man die Matrix aus der Polynomdivision invertieren? Besonderen Aufbau der Matrix beschrieben (untere Dreiecksmatrix und Toeplitzmatrix). Allgemeine Berechnung für A^{-1} für parallele Berechnung genannt. Besonderheiten bei Toeplitzmatrix erläutert und beschrieben, daß eigentlich nur die erste Spalte zu berechnen ist.

Allgemeiner Eindruck:

Dr. Damaschke hat mich teilweise mit seinen Fragen verwirrt, so daß ich nicht genau wußte worauf er eigentlich hinauswill. Rückfragen sind dann erlaubt. Er erklärt dann die Frage noch einmal. Negative Auswirkungen auf die Benotung haben diese Rückfragen allerdings nicht.

Bei der konvexen Hülle wußte ich am Ende ein paar Details nicht. Er hat dann - auf meine Bitte - ein neues Thema begonnen. Das hatte den Vorteil, daß die Lücke zwar bewertet wurde, aber in der Stoffmenge keine so große Rolle mehr spielte.

Dr. Damaschke ist ein netter und wohlgesonnener Prüfer.

Diplomprüfung Theoretische Informatik Teil 1

Kurs : Parallele Algorithmen 1824

Datum : 27.10.1998

Zeit : 11.00 Uhr

Dauer : 30 Minuten

Prüfer : Dr. Damaschke

Beisitzer : Herr Dipl. Inform. Robert Rettinger

Note : 1.3

⚡ Wie kann man Dualzahlen der Länge n effizient multiplizieren?

(Karatsuba Algorithmus)

⚡ Wie ist die Komplexität dieses Algorithmus und wie kommt man darauf?

$O(n^{\log_3 2})$, Abschätzung mittels Rekurrenzgleichung gezeigt (Definition von Satz 2.1.4 war nicht en detail gefragt)

⚡ Geht Multiplikation noch schneller?

Ja, mittels FFT! Definition der Fouriertransformation und den Zusammenhang zu Polynomen erläutert. Konvolutionssatz zitiert und Vorgehensweise geschildert. Komplexität des Verfahrens genannt ($O(\log(n))$ bei paralleler Ausführung der FFT).

⚡ Warum geht FFT so schnell?

Rekursiven FFT geschildert (Zerlegung in gerade und ungerade Ergebniskomponenten, besondere Struktur der Matrix F , Gewinnung der ω^k im voraus mittels Präfixsummenberechnung)

⚡ Funktioniert FFT auf jedem Ring?

Nein, $\omega, \omega^{-1}, n^{-1}$ müssen definiert sein.

⚡ Wie werden Matrizen effizient multipliziert?

Strassen Algorithmus geschildert und Komplexität genannt (Details der Berechnung waren nicht verlangt "das kann man bei Bedarf nachlesen" so Dr. D.).

⚡ Geht das auch mit booleschen Matrizen?

Nein $\{(0,1), ? , ? \}$ ist kein Ring.

⚡ Was muß man tun damit's funktioniert?

Man faßt die Elemente der Matrizen als Elemente einer Restklasse modulo $n+1$ auf!

⚡ Wie berechnet man effizient die konvexe Hülle einer Punktmenge?

Algorithmus bzw. Satz von Atallah Goodrich zitiert.

⚡ Welche vorher eingeführten Algorithmen werden dabei benutzt und warum?

Paralleles Sortieren zum schnellen Sortieren der Punkte nach steigender x-Koordinate. Paralleles Suchen auf den Streckenstücken von $up(S_1)$ und $up(S_2)$ nach dem passenden Steigungsintervall der Tangentensteigung. (deshalb ist der Algorithmus so schnell, $O(1)$ Zeit für die Such-Operation!)

⚡ Wie funktioniert paralleles Sortieren

Mergesort Algorithmus beschrieben (Reduktion auf das ranking Problem, Problematik des parallelen Mergings beschrieben, Lösung durch Zerlegung in $m/\log m$ große Blöcke und sequentielles Mischen dieser Blöcke)

⚡ Wie funktioniert paralleles Suchen?

Man greift mit mehreren Prozessoren gleichzeitig auf die zu durchsuchende Folge zu.

⚡ Beschreiben Sie die parallele Auswertung arithmetischer Ausdrücke!

Voraussetzungen genannt (Operanden in Blättern eines geordneten Bin-Baumes, Operatoren in den Knoten)

Preorder-Numerierung des Bin-Baumes bzw. des Eulerkreises in ihm beschrieben. Zwischenfrage: Was muß man tun, um diesen Preorder Durchlauf mit dem Eulerkreis-Algorithmus zu erzeugen? Knotengewichte: linker Sohn, rechter Sohn, Vater. Blätter mit dem Gewicht 1 versehen, übrige Knoten mit 0. Prefixsums über die Blätter schafft Numerierung in $O(\log n)$ Zeit. Baumkontraktion mit Rake beschrieben und Werterhalt mit Hilfe der Variablen A_v, B_v erläutert (nur verbal, Formalismus war nicht gefordert)

⚡ Was braucht man vor der Präfixsummenberechnung noch?

List-Ranking um überhaupt die Indizes der Listenelemente zu kennen!

⚡ Warum macht man diese aufwendige Operation (Rake etc.), man könnte doch einfach benachbarte Blätter zusammenfassen bzw. auswerten?

Der Baum muß nicht unbedingt balanciert sein sondern kann beliebig entartet sein. Beim Rake erzielt man garantiert bei jedem Schritt eine Halbierung der Anzahl der Blätter (e.g. man hat eine logarithmische Zeitschranke), beim einfachen ‚Baumauswerten‘ verliert man diese Zeitschranke!

Insgesamt war die Prüfung angenehm und verlief in einer ruhigen freundlichen Atmosphäre. Dr. Damaschke hat einen ruhigen Prüfungsstil; er bat fast um Entschuldigung, wenn er mich bei meinen Ausführungen unterbrach, weil ihm die Antwort bis dahin schon genügte oder er noch eine ergänzende Erläuterung machen wollte. Seine Fragen bauten häufig auf meinen Antworten auf. Trotzdem hatte ich vor allem bei der Nachbereitung der Prüfung das Gefühl, daß er sich mit großer Systematik einen Überblick über meinen Wissensstand des gesamten Kurses verschafft hatte (es fehlte eigentlich nichts wesentliches, bis auf Spezialitäten wie Toeplitz Matrizen lineare Rekurrenzen oder lineare Optimierung)

Auch bei dieser Prüfung hat sich kurzes (oder in diesem Fall auch längeres) Zögern beim Antworten oder ‚Verlaufen‘ nicht in – deutlichen – Notenabzügen niedergeschlagen. Dr. Damaschke gibt zum Teil Hilfestellung oder stellte bei meiner Prüfung auch eine Frage zurück um mir Gelegenheit zu geben, durch seine nächsten Fragen den ‚Groschen zum Fallen‘ zu bringen. Wichtig ist es vor allem bei diesem Prüfer einen guten Überblick über den Stoff zu haben und zu den allgemeineren Fragen (‘Was wissen Sie über...‘ oder ‘Erklären Sie ...’) ein paar Sätze im Vortragsstil vorzubereiten. Er stellt gern etwas global gehaltene Fragen die eben auch etwas längere Antworten erfordern.

Dr. Damaschke ist auf alle Fälle ein empfehlenswerter Prüfer.

Diplomprüfung
Theoretische Informatik
Gedächtnisprotokoll

Prüfer: Dr. Damaschke
Termin: 04.11.97
Dauer: 25 min
Note: 1,0
Prüfungsinhalt: Kurs 1824 - Parallele Algorithmen

Prüfungseindruck:

Allgemein ist der gesamte Kursumfang prüfungsrelevant. An den Lehrzielen kann man sich orientieren (steht ja alles drin ;) Besonders wichtig sind das Berechnungsmodell (PRAM mit Varianten) und der Zusammenhang zu den Komplexitäten der Algorithmen. Manchmal werden Algorithmen (z.B. list ranking, sortieren) innerhalb von anderen Algorithmen wieder eingesetzt; solche Zusammenhänge sollten dann bekannt sein. Man sollte zu den Algorithmen das Prinzip parat haben.

Ein globales Verständnis, insbesondere der Zusammenhänge, ist wesentlich wichtiger als technische Details der einzelnen Algorithmen/Sätze/Definitionen zu kennen! Bei Beweisen ist die Idee wichtig, nicht die technische Ausführung. In diesem Kurs scheinen die Beweise nicht so relevant zu sein. Aber Achtung: in einigen Beweisen wird erst der Algo so richtig erklärt (konstruktive Beweise).

Prüfungsverlauf:

1. Welches Rechen-Modell wurde im Kurs eingeführt und welche Unterscheidungen gibt es?
 - PRAM mit den verschiedenen Ausprägungen bzgl. des Zugriffs auf den Speicher.
2. Was ist list ranking?
 - Ich habe erläutert: Initialisierung, parallele Ausführung der WHILE-Schleife und das (und warum) eine EREW-PRAM ausreicht.
3. Wie kann man parallel arithmetische Ausdrücke auswerten?
 - Verfahren erläutern: Eulerkreis, Baumkontraktion, Wertehaltung beim RAKE, etc.
4. Warum kann man den Baum nicht auswerten, indem man die benachbarten Blätter parallel auswertet?
 - Kann man!!! Das klappt natürlich - ABER man hat nicht mehr den Effekt, daß die Anzahl der Knoten immer halbiert wird!!! Damit ist dann die logarithmische Zeite weg :(
5. Wie kann man parallel effizient Polynome multiplizieren?
 - Zusammenhänge erläutern: Konvolution, Fourier-Transformation, FFT (hier die Aufteilung in Teilprobleme) und die Komplexität.

Parallele Algorithmen

11. September 1997

1 PRAMs

Was braucht man für das Grundmodell der PRAM? Welche Arten von Registerzugriffen gibt es: bei welchen gibt es Schreibkonflikte und was für Lösungen gibt es dafür? Ist das Modell der PRAM aus dem Kurs eigentlich realistisch? Nein, die Prozessorkommunikation hat entscheidenden Einfluß.

2 Arithmetische Ausdrücke

Auswertung arithmetischer Ausdrücke: was braucht man dafür? Rake, Blätter numerieren: wie macht man das? Baumkontraktion: Blättern Gewichte zuordnen → Eulerkreis → Präfixsummen. Das mit den (A_v, B_v) wollte er gar nicht so genau wissen, aber: mit welchen Operationen geht das überhaupt?

3 Sortieren

Das Problem liegt im Merge-Schritt: B muss in $n \log n$ Teile zerlegt werden, dann wird A entsprechend zerlegt und die entsprechenden Teile werden zusammengefügt.

4 LOP

LOP: Erklären, was das eigentlich ist. Was ist eine Halbebene? Man braucht eine *lineare* Funktion zum Optimieren, damit man nur die Eckpunkte betrachten muss. Übersetzung in Bestimmung der linearen Hülle. Wie schnell geht komplexe Hülle und warum geht das so schnell?

5 Eindruck

Eigentlich ein sehr netter Prüfer, allerdings stellt er die Fragen so, daß man gleich den ganzen Kurs rekapitulieren könnte, aber es reicht, dann erst mal Stichworte zu nennen. Wie üblich ist der Überblick wichtiger als Detailkenntnisse. Man sollte allerdings ganz gut mit der O-Notation umgehen können, allerdings haben mir leichte Unsicherheiten dabei auch nicht geschadet, immerhin gab's eine glatte 1.

Diplomhauptprüfung Informatik C1857 - Theoretische Informatik Teil 2

Prüfungsinhalt 1824 - Parallele Algorithmen

Prüfer Dr. Damaschke

Datum 9. September 1997

Dauer ca. 25 min

Note 1,0

Fragen

- ✎ **Wie werden Präfixsummen bestimmt?**
 - ✎ Parallele Berechnung genau erklärt
- ✎ **Wo wird dieser Algorithmus verwendet?**
 - ✎ Mir ist zunächst nur die Bestimmung von Abständen in Bäumen eingefallen; aber er wird z.B. auch bei der Berechnung der Potenzen von φ bei der FFT benutzt.
- ✎ **Wie werden die Abstände in Bäumen bestimmt?**
 - ✎ Gerichteten Baum per Eulerpfad bestimmen
 - ✎ Kanten des Eulerpfads mit List ranking durchnummerieren
 - ✎ Den Kanten die Gewichte $+1 [p(v) \leq v]$ bzw. $-1 [v < p(v)]$ zuordnen
 - ✎ Präfixsummen der Kantengewichte bestimmen
- ✎ **Wie werden Matrizen effizient multipliziert?**
 - ✎ Strassen-Algorithmus erläutert
 - ✎ Komplexität: $M(n) = O(n^{\log_2 7}) = O(n^{2.82})$
 - ✎ Zerlegungsidee zur Verbesserung auf $O(n^{2.38})$ [steht nicht im Kurs]
- ✎ **Geht das auch für Matrizen, die als Elemente nur Null und Eins enthalten, mit der Addition modulo 2?**
 - ✎ Ja
- ✎ **Ist der Algorithmus für boolesche Matrizen parallelisierbar?**
 - ✎ Ich hatte schon im Zusammenhang mit der vorhergehenden Frage die Kernaussagen von Satz 4.1.2 erwähnt und wußte daher erst einmal nicht, worauf die Frage hinauslaufen sollte. Die Antwort ist natürlich **nein**, da die Struktur $(\{0,1\}, \oplus, \otimes)$ kein Ring ist und insbesondere keine "Subtraktion" definiert ist.
- ✎ **Multiplikation von Polynomen**
 - ✎ Fourier-Transformation erläutert
 - ✎ Zusammenhang mit Konvolutionssatz beschrieben

- ⚡ **Geht das auf jedem Ring?**
 - ⚡ Nein, $?$, $?^{-1}$ und n^{-1} müssen Elemente des Rings sein.
- ⚡ **Ist das lineare Optimierungsproblem (LOP) parallelisierbar?**
 - ⚡ Rückführung des LOP auf Durchschnitt von Halbräumen
 - ⚡ Einfach parallelisierbar im Fall $n=2$ (Durchschnitt von Halbebenen)
- ⚡ **Durchschnitt von Halbebenen**
 - ⚡ Duale Transformation auf Problem der konvexen Hülle
 - ⚡ Vorgehensweise zur Bestimmung der konvexen Hülle erläutert

Eindruck

Dies war meine zweite Prüfung bei Dr. Damaschke, daher war ich etwas besser auf seinen Fragestil eingestellt. Die Prüfung verlief in ruhiger Atmosphäre. Die klar gestellten Fragen gingen mehr in die Breite als in die Tiefe. Details wurden kaum nachgefragt. Andererseits läßt Dr. Damaschke einem Zeit, auch Details zu erzählen, so man sie kennt. Genaue Angaben zur Komplexität der Algorithmen wurden nicht gefordert. Eine Darstellung der grundlegenden Sachverhalte reichte in der Regel aus. Kleine Aussetzer bei weniger wichtigen Gebieten (siehe *boolesche Matrizen*) führten nicht zu Punktabzug.

Viel Erfolg!