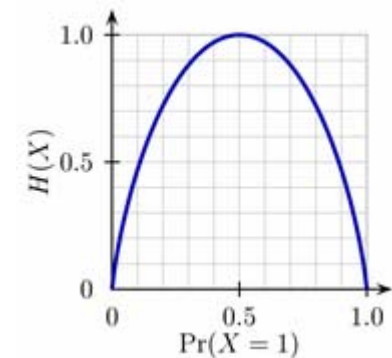


Information und Kommunikation

Hartmut Klauck
Universität Frankfurt
SS 07
13.7.



Freie Partitionen

- Wir haben bisher meist die Berechnung von Funktionen $f:\{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}$ betrachtet
- Dabei ist die Aufteilung der Eingabe an Alice und Bob fest vorgegeben
- Wir betrachten nun den Fall, dass die Aufteilung der Eingabe beliebig gewählt werden kann
 - So, dass jeder Spieler „genug“ Eingaben hat

Freie Partitionen

- Wir betrachten Funktionen $f:\{0,1\}^n \rightarrow \{0,1\}$
- Eine Partition der Eingaben in zwei Mengen ist *zulässig*, wenn jede der Mengen mindestens $n/3$ Elemente hat
- Die deterministische Kommunikationskomplexität unter freier Partition ist über das Minimum über alle zulässigen Partitionen definiert
- Beispiel: $EQ(x)=1$ wenn $x=yy$
- Unter der Partition, die die Eingaben $x_i, x_{n/2+i}$ jeweils bei einem Spieler sind, ist die Komplexität nur $O(1)$

Freie Partitionen

- $SEQ(x,y,i): |x|=|y|=n, i \in \{0, \dots, n-1\}$
- $SEQ(x,y,i)=1$ wenn x gleich y , falls y um i zyklisch verschoben wird.
- D.h. für alle $j: x_j = y_{i+j \bmod n}$
- „shifted equality“

Freie Partitionen

- **Theorem 25.1**
 - SEQ braucht deterministische Kommunikation $\Omega(n)$ unter jeder zulässigen Partition der Eingaben
- **Beweis:**
 - analog zum Einweg Fall

Anwendung

- Wir betrachten VLSI Chips
- Ein VLSI Chip ist ein auf ein Rechteck aufgebrachter Schaltkreis
- Die Länge und Breite des Rechtecks sind Vielfache eines Parameters Δ , der minimalen Breite einer Leitung auf dem Chip
- Es gibt Gatter und Leitungen zwischen den Gattern
- Es gibt m Eingaben, und eine Ausgabe
- Die Eingaben können beliebig auf dem Chip angebracht sein
- Der Chip berechnet eine Funktion $f:\{0,1\}^m \rightarrow \{0,1\}$
- Der Chip arbeitet getaktet, und hat daher eine Rechenzeit T (Anzahl der Takte bis f berechnet ist)
- Die Eingaben werden alle zur Zeit 0 angelegt
- Das Produkt der Seitenlängen des Chips ist die Fläche $A=ab$

VLSI

- Wir sind interessiert an der Größe AT^2
- AT entspricht normalerweise dem Energieverbrauch eines Chips
- Daher misst AT^2 einen Tradeoff zwischen Geschwindigkeit und Energieverbrauch
- **Theorem 25.2**
 - Sei f eine Funktion, die unter jeder Partition der Eingaben Kommunikationskomplexität C hat.
 - Dann gilt $AT^2 \geq \Omega(C^2)$ für alle Chips, die f berechnen

VLSI

- Wir erhalten:
 - Alle VLSI Chips für SEQ haben $AT^2 = \Omega(n^2)$
- Dasselbe kann man für die Multiplikationsfunktion zeigen

Beweis 25.2

- Wir nehmen an der Chip sei ein $a \times b$ Rechteck und $a \leq b$
- Wir nehmen an, dass $a \leq m/3$, sonst gilt bereits $A=ab \geq m^2/9$
- Wir schneiden den Chip in zwei Hälften, so dass jede höchstens $2m/3$ und mind. $m/3$ Eingaben hat
 - Wir schneiden parallel zur kürzeren Seite
 - Jede „Linie“ hat höchstens $m/3$ Eingaben
 - Es gibt also einen Schnitt wie behauptet

Beweis 25.2

- Jetzt kann der Chip durch ein Protokoll simuliert werden:
 - Alice erhält die Eingaben der linken Seite, Bob die Eingaben der rechten Seite
 - Jede Kommunikationsrunde entspricht einem Zeitschritt
 - Jeder Spieler kann seinen Teil des Chips simulieren
 - Es müssen die Daten kommuniziert werden, die über Leitungen zwischen beiden Teilen ausgetauscht werden
 - Es gibt also T Runden und a Kommunikation
 - Damit gilt: $Ta \geq C$
 - Und somit $T^2ab \geq T^2a^2 \geq C^2$
 - Also $T^2A \geq C^2$

Threshold Schaltkreise

- Wir betrachten jetzt Schaltkreise mit Threshold Gattern
- Eine Boolesche Funktion f ist eine Threshold Funktion wenn
 - $f(x_1, \dots, x_n) = 1$ wenn $\sum_i w_i x_i \geq t$
 - $f(x_1, \dots, x_n) = 0$ wenn $\sum_i w_i x_i < t$
 - für irgendwelche reellen Zahlen w_i und t
- Beispiele:
 - $\text{UND}(x_1, \dots, x_n) = 1$ wenn $\sum x_i \geq n$
 - $\text{ODER}(x_1, \dots, x_n) = 1$ wenn $\sum x_i \geq 1$
 - $\text{GT}(x_1, \dots, x_n, y_1, \dots, y_n) = 1$ wenn $\sum_i 2^i x_i - \sum_j 2^j y_j \geq 0$
 - $\text{Majority}(x_1, \dots, x_n) = 1$ wenn $\sum x_i \geq n/2$

Threshold Schaltkreise

- Ein Threshold Schaltkreis ist ein Schaltkreis mit unbeschränktem fan-in, dessen Gatter mit Threshold Funktionen beschriftet sind
- Die Größe eines Threshold Schaltkreises ist die Anzahl der Gatter (ohne die Eingaben zu zählen)
- Die Tiefe ist die Länge eines längsten Pfades von einer Eingabe zur Ausgabe

Threshold Schaltkreise

- Threshold Schaltkreise sind ein einfaches Modell für neuronale Netze
- Ein Threshold Gatter heißt auch ein Perceptron
- Perceptrons/Threshold Funktionen können effizient gelernt werden
- Für flache Threshold Schaltkreise gibt es den Backpropagation Algorithmus für das Lernproblem
- Wie groß muss ein neuronales Netz sein, um die gewünschte Funktion überhaupt darstellen zu können?

Threshold Schaltkreise

- Es gibt Threshold Schaltkreise der Größe 1 für UND, ODER, Majority, GT
- Wir wissen bereits, dass $R(GT) \leq O(\log n)$ ist
- Man kann leicht zeigen, dass tatsächlich für jede Threshold Funktion $f : R_{1/n^2}(f) = O(\log n)$ unter jeder Aufteilung der Eingaben an Alice und Bob (auch solchen, wo Alice und Bob stark unterschiedlich vielen Eingaben erhalten)

Threshold Schaltkreise

- **Theorem 25.3**
 - Für jede Funktion $f:\{0,1\}^n \rightarrow \{0,1\}$ gilt, dass jeder Threshold Schaltkreis für f mindestens $\Omega(C/\log n)$ Gatter benötigt, wenn C die randomisierte Kommunikationskomplexität von f unter irgendeiner Aufteilung der Eingaben ist
- **Beispiel: $IP(x,y)$**
 - $R(IP)=\Omega(n)$ unter der Standard Aufteilung der Eingaben
- D.h. jeder Threshold Schaltkreis für IP braucht mindestens $\Omega(n/\log n)$ Gatter
- Das gleiche gilt für $DISJ$

Beweis 25.3

- Wir zeigen ein Protokoll für f
- Gegeben ist ein Threshold Schaltkreis für f mit Größe $s \leq n$
- Alice erhält eine Menge von Eingaben, Bob eine Menge von Eingaben
- Sie werten den Threshold Schaltkreis Gatter für Gatter aus
- Die Funktion jedes Gatter kann mit $O(\log n)$ Kommunikation und Fehler $1/n^2$ berechnet werden
- Damit ist die Gesamtkommunikation $O(s \log n)$
- Der Fehler ist höchstens $s/n^2 \leq 1/n$
- Also gilt $C \leq O(s \log n)$

Schaltkreise und Kommunikation

- Man kann obige Konstruktion leicht auf andere Schaltkreisklassen anwenden:
 - Z.B. Schaltkreise mit UND, ODER, NICHT Gattern mit unbeschränktem fan-in
 - Die deterministische Kommunikationskomplexität der Gatter ist $O(1)$
 - Daher braucht ein solcher Schaltkreis für eine Funktion f mind. $\Omega(C)$ viele Gatter wenn C die deterministische Kommunikationskomplexität von f unter irgendeiner Aufteilung der Eingaben ist
 - Solche Schaltkreise für Equality oder DISJ brauchen also $\Omega(n)$ Gatter

Threshold Schaltkreise und Gewichte

- Wir betrachten Threshold Schaltkreise
- Man kann zeigen, dass es ausreicht, ganzzahlige Gewichte der Größe $\exp(n \log n)$ zu verwenden
- D.h. für alle Threshold Funktionen $f(x_1, \dots, x_n) = 1$ wenn $\sum x_i w_i \geq t$, können wir eine Darstellung finden, in der die Gewichte ganzzahlig sind und ihr Betrag nicht größer als $\exp(n \log n)$ ist
- Wir sind insbesondere interessiert an polynomiell beschränkten Gewichten

Gewichte

- Jede Threshold Funktion f mit ganzzahligen polynomiell beschränkten Gewichten erfüllt $D(f) = O(\log n)$ unter einer beliebigen Aufteilung der Eingabe
 - Alice berechnet $\sum x_i w_i$ für ihre Eingaben
 - Dies ist eine Zahl, die polynomiell beschränkt ist
 - Alice sendet die Zahl an Bob, der nun $\sum x_i w_i$ über alle Eingaben berechnen kann
 - Bob kann nun die Funktion entscheiden

Gewichte

- Analog zu 25.3 folgt
- **Theorem 25.4**
 - Für jede Funktion f gilt, dass jeder Threshold Schaltkreis mit polynomiellen Gewichten für f mindestens $\Omega(C/\log n)$ Gatter benötigt, wenn C die *deterministische* Kommunikationskomplexität von f unter irgendeiner Aufteilung der Eingaben ist
- **Beispiel:**
 - Equality: Jeder Threshold Schaltkreis mit pol. Gewichten braucht mind. $\Omega(n/\log n)$ Gatter
 - Ebenso GT

Tiefe und Gewichte

- Es gibt also eine Funktion (GT), die bei exponentiellen Gewichten in Tiefe 1 berechnet werden kann, bei polynomiellen Gewichten nur in Tiefe 2
 - Denn Größe >1 impliziert Tiefe >1
- Man kann zeigen, dass jede Funktion, die in Tiefe d berechnet werden kann (bei polynomieller Größe) mit beliebigen Gewichten auch in Tiefe $d+1$ durch einen polynomiell großen Schaltkreis nur aus Majority Gattern berechnet werden