ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich
Departement Informatik

Open Class – Sieben Wunder der Informatik

Prof. Dr. Juraj Hromkovič

Lösungsvorschläge für die Übungsaufgaben – Blatt 5

Zürich, 7. Dezember 2005

Lösung zu Aufgabe 14

Gegeben seien die Eingaben $x = 1001001$ für den Rechner R_I und $y = 0101011$ für den Rechner R_{II}. Diese Eingaben haben die Länge $n = 7$.

a.) Hiermit ergibt sich die Menge

$$\text{PRIM}(n^2) = \text{PRIM}(49) = \{2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47\} \quad .$$

b.) Wir führen nun das randomisierte Protokoll aus der Vorlesung für die gegebene Eingabe durch. Wir können die Eingaben als Zahlen $x = 2^6 + 2^3 + 2^0 = 64 + 8 + 1 = 73$ und $y = 2^5 + 2^3 + 2^1 + 2^0 = 32 + 8 + 2 + 1 = 43$ interpretieren.

- In unserer ersten Durchführung des Protokolls wählen wir zufällig die Primzahl $p = 3 \in \text{PRIM}(49)$ (Schritt 1 des Protokolls).

  Der Rechner R_I bestimmt jetzt den Rest s bei der Division von x durch p, es ergibt sich $s = 73 \bmod 3 = 1$. Der Rechner R_I schickt jetzt p und s an R_{II} (Schritt 2 des Protokolls).

  Der Rechner R_{II} berechnet jetzt $t = y \bmod p = 43 \bmod 3 = 1$. Da $s = t$ gilt, gibt R_{II} die falsche Antwort „gleich“ aus (Schritt 3 des Protokolls).
- In unserem zweiten Versuch wählen wir zufällig die Primzahl $p = 5$ aus. Damit bestimmt R_I die Zahl $s = x \bmod p = 73 \bmod 5 = 3$ und sendet p und s an R_{II}. Der Rechner R_{II} berechnet nun $t = y \bmod p = 43 \bmod 5 = 3$. Also gilt $s = t$, und R_{II} gibt die falsche Antwort „gleich“ aus.
- Im dritten Versuch wählen wir zufällig die Primzahl $p = 31$ aus. Damit wird $s = x \bmod p = 73 \bmod 31 = 11$ von R_I bestimmt und zusammen mit p an R_{II} geschickt. Der Rechner R_{II} berechnet dann $t = y \bmod p = 43 \bmod 31 = 12$. Also gilt wieder $s \neq t$, und damit gibt R_{II} die korrekte Antwort „ungleich“ aus.

Lösung zu Aufgabe 15

Wie in der Vorlesung gezeigt liefert das Protokoll genau dann ein falsches Ergebnis, wenn $(x - y) \bmod p = 0$ gilt, wenn also die Differenz der zu vergleichenden Eingaben ein Vielfaches der zufällig gewählten Primzahl ist. In dem Beispiel aus Aufgabe 14 ist $x - y = 30$,

und somit liefert nur die Wahl einer der Primzahlen aus $\{2, 3, 5\}$ ein falsches Ergebnis, da dieses die einzigen Primzahlen sind, die $x - y$ teilen. Es gilt $|\,\mathrm{PRIM}(49)| = 15$, also beträgt die Fehlerwahrscheinlichkeit des Protokolls in diesem Beispiel $3/|\,\mathrm{PRIM}(49)| = 0.2 = 20\%$. Wenn wir jetzt beliebige Werte der Länge 7 für x und y zulassen, dann kann die Differenz $x - y$ maximal $2^7 - 1 = 127$ betragen. Damit ist diese Differenz durch maximal drei verschiedene Primzahlen teilbar, da bereits das Produkt der vier kleinsten Primzahlen $2 \cdot 3 \cdot 5 \cdot 7 = 210$ grösser als 127 ist. Also lässt sich die Fehlerwahrscheinlichkeit des Protokolls auch für beliebige Werte (der Länge 7) für x und y durch $3/|\,\mathrm{PRIM}(49)| = 20\%$ beschränken.

Lösung zu Aufgabe 16

Wenn man das Protokoll insoweit modifiziert, dass man die Primzahl zufällig aus $\mathrm{PRIM}(n^3)$ wählt statt aus $\mathrm{PRIM}(n^2)$, dann steigt damit die Kommunikationskomplexität von $2 \cdot \log_2 n^2 = 4 \cdot \log_2 n$ auf $2 \cdot \log_2 n^3 = 6 \cdot \log_2 n$, da die zu übertragenden Zahlen jetzt aus dem Bereich von 0 bis $n^3 - 1$ stammen und somit eine Länge von bis zu $\log_2 n^3$ haben können. Die Menge $\mathrm{PRIM}(7^3) = \mathrm{PRIM}(343)$ enthält 68 Elemente, die Anzahl der schlechten Primzahlen, deren Auswahl zum falschen Ergebnis führt, bleibt aber 3, damit ergibt sich eine Fehlerwahrscheinlichkeit von $3/68 \approx 0.04412 = 4.412\%$.

Lösung zu Bonus-Aufgabe 5

Wir modifizieren das Protokoll insoweit, dass die Primzahl zufällig aus dem Bereich zwischen 2 und n^4 gewählt wird.

a.) Für das so modifizierte Protokoll ergibt sich eine Kommunikationskomplexität von $8 \cdot \log_2 n$, da die beiden zu übertragenden Zahlen p und s jeweils die Länge $\log_2 n^4 = 4 \cdot \log_2 n$ haben können.

b.) Die Fehlerwahrscheinlichkeit für das modifizierte Protokoll lässt sich wie folgt abschätzen: Nach dem Primzahlsatz gibt es in dem Bereich von 1 bis n_4 ungefähr $\frac{n^4}{4 \cdot \ln n}$ verschiedene Primzahlen. Genau wie in dem Beweis der Vorlesung befinden sich darunter höchstens $n-1$ schlechte Primzahlen, denn abermals beträgt die Differenz $x-y$ höchstens 2^n und ist damit durch höchstens $n - 1$ verschiedene Primzahlen teilbar. Nur wenn eine schlechte Primzahl ausgewählt wird, liefert das Protokoll einen Fehler, also beträgt die Fehlerwahrscheinlichkeit höchstens

$$\frac{n-1}{\frac{n^4}{4 \cdot \ln n}} < \frac{n}{\frac{n^4}{4 \cdot \ln n}} = \frac{4 \cdot \ln n}{n^3} \quad .$$

c.) Wir wollen nun das modifizierte Protokoll mit der zweimaligen Ausführung des alten Protokolls vergleichen. Die Kommunikationskomplexität ist in beiden Fällen gleich gross, da im einen Fall zwei Zahlen der Länge $4 \cdot \log_2 n$ übertragen werden und im anderen Fall vier Zahlen der Länge $2 \cdot \log_2 n$.

Wir versuchen nun, die Fehlerwahrscheinlichkeiten beider Protokolle miteinander zu vergleichen. Das modifizierte Protokoll hat eine Fehlerwahrscheinlichkeit von höchstens

$$\frac{4 \cdot \ln n}{n^3} = \frac{1}{n} \cdot \frac{4 \cdot \ln n}{n^2}$$

wie oben gezeigt. Die Fehlerwahrscheinlichkeit für die zweimalige Ausführung des ursprünglichen Protokolls lässt sich berechnen als das Quadrat der Fehlerwahrscheinlichkeit $\frac{2 \cdot \ln n}{n}$ dieses Protokolls, es ergibt sich also eine Fehlerwahrscheinlichkeit von höchstens

$$\left(\frac{2 \cdot \ln n}{n}\right)^2 = \frac{4 \cdot \ln^2 n}{n^2} = \ln n \cdot \frac{4 \cdot \ln n}{n^2} \quad .$$

Ein Vergleich dieser oberen Schranken ergibt, dass die Schranke für das modifizierte Protokoll stets eine kleinere Fehlerwahrscheinlichkeit angibt, da $\frac{1}{n} < \ln n$ gilt (für alle $n > 2$).

Dies bedeutet aber nicht, dass das modifizierte Protokoll tatsächlich in jedem Fall besser ist als die zweimalige Anwendung des alten Protokolls. Dies liegt daran, dass die berechneten oberen Schranken nicht für beide Protokolle eine gleich gute Näherung des tatsächlichen Wertes angeben müssen. Während die Abschätzung des Primzahlsatzes für grosse Werte von n eine sehr gute Näherung angibt, ist die von uns berechnete Abschätzung der Anzahl schlechter Primzahlen unter Umständen sehr ungenau.

Wenn wir allgemein die genaue Anzahl der schlechten Primzahlen im Bereich von 2 bis n^2 mit m bezeichnen und die genaue Anzahl der schlechten Primzahlen im Bereich von 2 bis n^4 mit m' bezeichnen, dann können wir die Fehlerwahrscheinlichkeiten in Abhängigkeit von m und m' genauer abschätzen: Das modifizierte Protokoll hat dann eine Fehlerwahrscheinlichkeit von höchstens

$$m' \cdot \frac{4 \cdot \ln n}{n^4} \quad ,$$

und die zweimalige Anwendung des alten Protokolls hat eine Fehlerwahrscheinlichkeit von höchstens

$$\left(m' \cdot \frac{2 \cdot \ln n}{n^2}\right)^2 = m^2 \cdot \ln n \cdot \frac{4 \cdot \ln n}{n^4} \quad .$$

Unter der Annahme, dass die Abschätzung der Gesamtanzahl von Primzahlen nach dem Primzahlsatz hinreichend genau ist, können wir also folgern, dass das neue Protokoll genau dann besser ist als die zweimalige Anwendung des alten, wenn $m' < m^2 \cdot \ln n$ gilt.

Dies ist sicherlich immer dann der Fall, wenn die schlechten Primzahlen für die betrachtete Differenz $x - y$ alle kleiner als n^2 sind. Wenn wir aber zum Beispiel eine Eingabe betrachten, bei der die Differenz $x - y$ eine Primzahl aus dem Bereich von n^2 bis n^4 ist, dann gibt es für das alte Protokoll gar keine schlechte Primzahl, aber

für das modifizierte Protokoll ist $x - y$ eine schlechte Primzahl. Also ist in diesem Fall das alte Protokoll (sogar ohne Wiederholung) besser als das neue.