Satz von Levin

Existenz eines `distNP`-vollständigen Problems

Diese Ausarbeitung basiert auf Kapitel 18 von
Computational Complexity: A Modern Approach
Sanjeev Arora und Boaz Barak (2009)

1 Definition: Verteiltes Problem

Ein verteiltes Problem ist ein Tupel $\langle L, \mathcal{D} \rangle$ mit einer Sprache $L \subseteq \{0,1\}^*$ und einer Familie von Wahrscheinlichkeitsverteilungen $\mathcal{D} = \mathcal{D}_n$, wobei $\mathcal{D}_n$ eine Verteilung über $\{0,1\}^n$ ist.

2 Definition: $\mathbf{P}$-berechenbare Verteilung

Eine Verteilung $\mathcal{D}_n$ heißt $\mathbf{P}$-berechenbar (polynomiell berechenbar), wenn es eine deterministische Turingmaschine gibt, welche zu einer Eingabe $x \in \{0,1\}^n$ die kumulierte Wahrscheinlichkeit $\mu_{\mathcal{D}_n}(x)$ von x bzgl. $\mathcal{D}_n$ berechnet:

$$\mu_{\mathcal{D}_n}(x) = \sum_{y \in \{0,1\}^n : y \leq x} \mathbb{P}_{\mathcal{D}_n}(y)$$

$y \leq x$ bedeutet hierbei, dass y in der lexikographischen Ordnung vor x steht (bzw. identisch mit x ist). $\mathbb{P}_{\mathcal{D}_n}$ bezeichne hier, wie auch im folgenden, das Wahrscheinlichkeitsmaß bezüglich $\mathcal{D}_n$.

3 Definition: Die Komplexitätsklasse `distNP`

`distNP` ist die Klasse aller verteilten Probleme $\langle L, \mathcal{D} \rangle$, für die $L \in \mathbf{NP}$ und deren Verteilung $\mathcal{D}$ $\mathbf{P}$-berechenbar ist.

4 Definition: Average-Case-Reduktion

Eine Average-Case-Reduktion von einem verteilten Problem $\langle L, \mathcal{D} \rangle$ auf ein weiteres verteiltes Problem $\langle L', \mathcal{D}' \rangle$ ist eine polynomiell berechenbare Funktion $f\colon \{0,1\}^n \to \{0,1\}^n$, welche für zwei Polynome p, q die folgenden Eigenschaften erfüllt:

1. Korrektheit: $\forall x \in \{0,1\}^* : x \in L \Leftrightarrow f(x) \in L'$
2. Längentreue: $\forall x \in \{0,1\}^n : |f(x)| = p(n) \quad (n = |x|)$
3. Dominanz: $\forall n \in \mathbb{N} \quad x \in \{0,1\}^n : \mathbb{P}_{f \circ \mathcal{D}_n}(f(x)) \leq q(n) \mathbb{P}_{\mathcal{D}'_{p(n)}}(f(x))$

Gibt es eine Average-Case-Reduktion von $\langle L, \mathcal{D} \rangle$ auf $\langle L', \mathcal{D}' \rangle$, so heißt $\langle L, \mathcal{D} \rangle$ average-case-reduzierbar auf $\langle L', \mathcal{D}' \rangle$, schreibe $\langle L, \mathcal{D} \rangle \leq_{avg} \langle L', \mathcal{D}' \rangle$.

5 Definition: $\mathtt{dist}\mathbf{NP}$-Vollständigkeit

Ein verteiltes Problem $\langle L', \mathcal{D}' \rangle$ ist $\mathtt{dist}\mathbf{NP}$-vollständig, wenn $\langle L', \mathcal{D}' \rangle$ in $\mathtt{dist}\mathbf{NP}$ ist und $\langle L, \mathcal{D} \rangle \leq_{avg} \langle L', \mathcal{D}' \rangle$ für jedes $\langle L, \mathcal{D} \rangle \in \mathtt{dist}\mathbf{NP}$ gilt.

6 Satz von Levin

Es sei U die Sprache aller Tupel $\langle M, x, 1^t \rangle$, wobei M eine nichtdeterministischen Turingmaschine ist, welche die Eingabe x nach höchstens t Schritten akzeptiert.
Für jedes $n \in \mathbb{N}$ sei $\mathcal{U}_n$ die folgende Gleichverteilung von Tupeln $\langle M, x, 1^t \rangle$ der Länge n: Die Kodierung von M wird zufällig aus allen Wörtern der Länge höchstens $\log n$ ausgewählt. t ist eine zufällig gewählte Zahl aus $\{0, \ldots, n - |M|\}$ (unär kodiert) und x ein zufällig gewähltes Wort aus $\{0,1\}^{n-t-|M|}$.
$\mathcal{U} := (\mathcal{U}_n)_{n \in \mathbb{N}}$. Dann ist $\langle U, \mathcal{U} \rangle$ $\mathtt{dist}\mathbf{NP}$-vollständig.

7 Lemma: Spitzenwert-Eliminierung

Es sei $\mathcal{D}_n$ eine polynomiell (**P**-)berechenbare Verteilung. Dann existiert eine polynomiell berechenbare Funktion $g\colon \{0,1\}^* \to \{0,1\}^*$ mit folgenden Eigenschaften:

1. g ist injektiv: $g(x) = g(y) \Leftrightarrow x = y$,
2. $\forall x \in \{0,1\}^n : |g(x)| \leq n + 1$,
3. $\forall x \in \{0,1\}^n : \mathbb{P}_{g \circ \mathcal{D}_n}(g(x)) \leq 2^{-|g(x)|+1} \quad (\text{und } \forall y, g^{-1}(y) = \emptyset : \mathbb{P}_{g \circ \mathcal{D}}(y) = 0)$.

8 Beweis des Lemmas

Für ein Wort $x \in \{0,1\}^*$ mit $\mathbb{P}_{\mathcal{D}}(x) > 2^{-|x|}$ sei $h(x)$ das längste gemeinsame Präfix der Binärrepräsentationen von $\mu_{\mathcal{D}}(x)$ und $\mu_{\mathcal{D}}(x-1)$. Da $\mu_{\mathcal{D}}(x) - \mu_{\mathcal{D}}(x-1) = \mathbb{P}_{\mathcal{D}}(x) > 2^{-|x|}$, müssen sich die Binärzahlen $\mu_{\mathcal{D}}(x)$, $\mu_{\mathcal{D}}(x-1)$ in den ersten $|x|$ Bits unterscheiden. Also ist h polynomiell berechenbar, da $\mathcal{D}$ **P**-berechenbar ist und es gilt:

$$2^{-|x|} \;<\; \mathbb{P}_{\mathcal{D}}(x) \;=\; \underbrace{\mu_{\mathcal{D}}(x) \quad - \quad \mu_{\mathcal{D}}(x-1)}_{\text{(unterscheiden sich erst ab der } |h(x)|-ten \text{ Binärziffer)}} \;<\; 2^{-|h(x)|}.$$

Darüber hinaus ist h injektiv, denn für die Bestimmung von $h(x)$ und $h(y)$ mit $x \neq y$ $(x, y \in \{0,1\}^*)$ werden mindestens drei Binärzahlen verglichen. Haben aber zwei Worte w_1, w_2 das längste gemeinsame Präfix p, so muss für ein drittes Wort $w_3 = p\ldots$ mit gleichem Präfix entweder $w_3 = p\,0\ldots$ oder $w_3 = p\,1\ldots$ gelten. Dann hat w_3 entweder mit w_1 oder mit w_2 ein längeres gemeinsames Präfix.
Definiere nun $g\colon \{0,1\}^* \to \{0,1\}^*$ wie folgt:

$$g(x) = \begin{cases} 0x & \text{wenn } \mathbb{P}_{\mathcal{D}}(x) \leq 2^{-|x|} \\ 1h(x) & \text{sonst} \end{cases}.$$

g ist offenbar polynomiell berechenbar, da h polynomiell berechenbar ist. Per Konstruktion von g und h folgt, dass g injektiv ist und für $x \in \{0,1\}^*$ gilt: $|g(x)| \leq |x| + 1$, da $|h(x)| \leq |x|$. Bleibt noch zu zeigen, dass für $y \in \{0,1\}^*$ gilt: $\mathbb{P}_{g\circ\mathcal{D}}(y) \;\leq\; 2^{-|y|+1}$.
Fall 1: $y \neq g(x)$ für alle $x \in \{0,1\}^*$. Dann ist $\mathbb{P}_{g\circ\mathcal{D}}(y) = 0$.
Fall 2: $y = g(x) = 0x$ für ein $x \in \{0,1\}^*$. Dann gilt nach Konstruktion, dass $\mathbb{P}_{\mathcal{D}}(x) \;\leq\; 2^{-|x|}$, also:

$$\mathbb{P}_{g\circ\mathcal{D}}(y) \;=\; \mathbb{P}_{g\circ\mathcal{D}}(0x) = \mathbb{P}_{\mathcal{D}}(x) \;\leq\; 2^{-|x|} = \leq\; 2^{-|y|+1}.$$

Fall 3: $y = g(x) = 1h(x)$ für ein $x \in \{0,1\}^*$. Dann gilt, wie oben gezeigt, dass $\mathbb{P}_{\mathcal{D}}(x) \;\leq\; 2^{-|h(x)|}$, also:

$$\mathbb{P}_{g\circ\mathcal{D}}(y) \;=\; \mathbb{P}_{g\circ\mathcal{D}}(1h(x)) \overset{(*)}{=} \mathbb{P}_{\mathcal{D}}(x) \;\leq\; 2^{-|h(x)|} = 2^{-|y|+1}.$$

Die Gleichheit bei $(*)$ gilt, weil g injektiv ist. Damit erfüllt g alle geforderten Eigenschaften.

9 Beweis des Satzes

Offenbar ist $U \in \mathbf{NP}$, denn ein Orakelmodul einer weiteren NDTM kann die nichtdeterministischen Zustandsübergänge von M raten und dann M mit diesen Übergängen für eine Dauer von t Schritten simulieren. Hält M binnen dieser Zeit, so ist die Ausgabe identisch mit der Ausgabe von M; anderenfalls wird das Tupel nicht akzeptiert.
Ferner ist $\mu_{\mathcal{U}_n}(x) = \sum_{y\in\{0,1\}^n : y\leq x} \mathbb{P}_{\mathcal{U}_n}(y) = \sum_{y\in\{0,1\}^n : y\leq x} 2^{-n} = (x+1)2^{-n}$ offenbar **P**-berechenbar (x hier interpretiert als natürliche Zahl). Damit ist $\langle U, \mathcal{U}\rangle \in \texttt{distNP}$.

Es sei nun $\langle L, \mathcal{D}\rangle$ mit $\mathcal{D} = (\mathcal{D}_n)_{n\in\mathbb{N}}$ ein Verteilungsproblem aus $\mathtt{distNP}$ und M eine nichtdeterministische Turingmaschine, die L in polynomieller Zeit entscheidet.
Wir konstruieren eine neue NDTM M' wie folgt: Bei Eingabe y rät und überprüft M' das eindeutig bestimmte Wort x, sodass $y = g(x)$, wobei g die im Lemma konstruierte injektive Funktion ist. Danach wird die NDTM M auf der Eingabe x simuliert und deren Ergebnis ausgegeben. Da gemäß Lemma g in polynomieller Zeit berechnet werden kann, hat M' eine polynomielle Laufzeit, die durch ein Polynom t nach oben beschränkt ist.
Für die Average-Case-Reduktion f von $\langle L, \mathcal{D}\rangle$ nach $\langle U, \mathcal{U}\rangle$ wird dann die Eingabe x ($n := |x|$) auf das Tupel $\langle M', g(x), 1^k\rangle$ abgebildet. Hierbei wird $k = t(n) + n + 1 - |g(x)|$ gewählt. Offenbar gilt die zweite Bedingung einer Average-Case-Reduktion, denn $|f(x)| = |\langle M', g(x), 1^k\rangle| = |M'| + |g(x)| + k = t(n) + n + 1 + |M'| =: p(n)$ ist ein Polynom, weil die Länge der Kodierung von M' als konstant angenommen wird.
Nach dem Lemma gilt $|g(x)| \leq |x| + 1 = n + 1$. Daher ist $k \geq t(n)$. Dies bedeutet, dass das Tupel $\langle M', g(x), 1^k\rangle$ genau dann in U liegt, wenn M' die Eingabe $g(x)$ akzeptiert (denn binnen $t(n)$ Schritten hält M' in jedem Fall). Da g injektiv ist, ist dies ferner genau dann der Fall, wenn x von M akzeptiert wird, also x ein Wort aus L ist. Deshalb gilt $x \in L \Leftrightarrow f(x) \in U$, die erste Bedingung einer Average-Case-Reduktion.
Um schließlich die Dominanz-Bedingung zu zeigen, muss nur $\mathbb{P}_{f\circ\mathcal{D}_n}(f(x)) \leq q(n)\mathbb{P}_{\mathcal{U}_{p(n)}}(f(x))$, also $\mathbb{P}_{f\circ\mathcal{D}_n}(\langle M', y, 1^k\rangle) \leq q(n)\mathbb{P}_{\mathcal{U}_{p(n)}}(\langle M', y, 1^k\rangle)$ gezeigt werden, denn falls es zu einem z kein x gibt, sodass $f(x) = z$, so ist $\mathbb{P}_{f\circ\mathcal{D}_n}(z) = 0$ und die Ungleichung trivialerweise erfüllt.
Wir schätzen dafür zunächst die linke Seite nach oben ab. Nach Konstruktion und wegen der Injektivität von g hängt $\mathbb{P}_{f\circ\mathcal{D}_n}(\langle M', y, 1^k\rangle)$ nur von $y = g(x)$ ab und für dieses liefert das Lemma die Ungleichung

$$\mathbb{P}_{f\circ\mathcal{D}_n}(f(x)) = \mathbb{P}_{f\circ\mathcal{D}_n}(\langle M', g(x), 1^k\rangle) = \mathbb{P}_{g\circ\mathcal{D}_n}(g(x)) \leq 2^{-|g(x)|+1} = 2^{-|y|+1}.$$

Um die rechte Seite nach unten abzuschätzen, sei $m := p(n)$ die Länge des Tupels $\langle M', y, 1^k\rangle$. Für große n (und damit für große m) ist $|M'| \leq \log m$, da wieder die Kodierung von M' als konstant angenommen wird. $\mathcal{U}_n$ ist damit so definiert, dass das Wort M' mindestens mit der Wahrscheinlichkeit $2^{-\log m}$ auftritt. Die Wahrscheinlichkeit für k Einsen ist mindestens $1/m$ und für das Wort y genau $2^{-|y|}$. Dies gilt alles, weil $\mathcal{U}$ gleichverteilt ist. Also ergibt sich:

$$\mathbb{P}_{\mathcal{U}_{p(n)}}(f(x)) = \mathbb{P}_{\mathcal{U}_m}(\langle M', y, 1^k\rangle) \geq \frac{1}{2^{\log m}} \cdot \frac{1}{2^{|y|}} \cdot \frac{1}{m} = \frac{1}{m^2} \cdot \frac{1}{2^{|y|}}.$$

Jetzt sei q ein Polynom, sodass

$$2^{-|y|+1} = \frac{2}{2^{|y|}} \leq q(n) \cdot \frac{1}{m^2} \cdot \frac{1}{2^{|y|}} \quad \Leftrightarrow \quad 2 \leq q(n)\frac{1}{m^2} \quad \Leftrightarrow \quad 2m^2 \leq q(n).$$

Dieses existiert, da $m = p(n)$ wiederum ein Polynom ist. Für ein solches q gilt dann insbesondere die gewünschte Ungleichung

$$\mathbb{P}_{f\circ\mathcal{D}_n}(f(x)) = \mathbb{P}_{f\circ\mathcal{D}_n}(\langle M', y, 1^k\rangle) \leq q(n)\mathbb{P}_{\mathcal{U}_{p(n)}}(\langle M', y, 1^k\rangle) = q(n)\mathbb{P}_{\mathcal{U}_{p(n)}}(f(x)),$$

was den Beweis vervollständigt.