5.4 Bit-Komplexität

Die Ergebnisse in diesem Abschnitt sind aus (Raghavendra, Weitz '17).

Das folgende Beispiel illustriert, dass SOS-Ableitungen vom Grad d (welche dann nur $O(n^d)$ Monome enthalten) möglicherweise Koeffizienten der Größe $2^{\Omega(2^n)}$ haben, wodurch die Darstellung der SOS-Ableitung als Bit-String die Größe $\Omega(2^n)$ hat.

▷ Diese untere Schranke gilt für SOS-Ableitungen eines bestimmten Polynoms.

- Gibt es ein ähnliches Argument für SOS-Widerlegungen?

▷ Diese untere Schranke gilt für bestimmte Systeme von Polynomgleichungen, die keine KNF kodieren.

- Gibt es ein ähnliches Argument für Polynomgleichungen der Form $P_F \cup \{x_i^2 - x_i \mid 1 \le i \le n\}$?

Sei $f_i = z_i^2 - z_{i+1}$ für $1 \le i \le n-1$ und $f_n = z_n^2$.

D.h. wir betrachten das System

$$z_1^2 = z_2,\ z_2^2 = z_3,\ \dots,\ z_{n-1}^2 = z_n,\ z_n^2 = 0$$

dessen einzige Lösung $(0, \dots, 0)$ ist.

$\Rightarrow$ für jede Lösung und jedes $\varepsilon > 0$ gilt

$$f = \varepsilon - z_1 > 0$$ (im Folgenden kann man sich $\varepsilon = \frac{1}{3}$ denken.)

Lemma 5.7

Es gibt $g_i = c_i \in \mathbb{R}$ und p_ℓ mit $\deg(p_\ell) = 1$ sodass

$$\sum_{i=1}^{n} g_i f_i + \sum_{\ell=1}^{n} p_\ell^2 = f.$$

(Nach der Terminologie aus dem Beweis von Satz 5.4 heißt das, dass $-f$ eine Grad-2 Ableitung aus $\{f_1, \dots, f_n\}$ besitzt. Hier haben wir allerdings nicht die Axiome $z_i^2 = z_i$ zur Verfügung!)

Beweis (Lemma 5.7)

- Wir nutzen die c_i nun in folgender Manier $b_i^2 z_i^2$ durch $b_i^2 z_{i+1}$ ersetzen zu können.

- Wir definieren $p_\ell = (a - b_\ell z_\ell)$ für geeignete $a, b_\ell \in \mathbb{R}$.

$$\Rightarrow \sum_{\ell=1}^{n} p_\ell^2 = \sum_{\ell=1}^{n} (a^2 - 2ab_\ell z_\ell + b_\ell^2 z_\ell^2) = \sum_{\ell=1}^{n-1} (a^2 - 2ab_\ell z_\ell + b_\ell^2 z_{\ell+1})$$
$$+ (a^2 - 2ab_n z_n)$$
$$+ \sum_{i=1}^{n} (b_i^2) \cdot f_i$$

Wir wählen $c_i := -b_i^2$ und müssen nun zeigen, dass $f = \varepsilon - z_1 \overset{!}{=} \sum_{\ell=1}^{n-1} (a^2 - 2ab_\ell z_\ell + b_\ell^2 z_{\ell+1}) + (a^2 - 2ab_n z_n)$

$$= n \cdot a^2 + (-2ab_1) z_1 + \sum_{\ell=2}^{n} (b_{\ell-1}^2 - 2ab_\ell) z_\ell$$

Dies erreichen wir dadurch, dass wir

$$a := \sqrt{\frac{\varepsilon}{n}}, \quad b_1 := \frac{1}{2a}, \quad b_\ell := \frac{b_{\ell-1}^2}{2a} \text{ wählen.}$$

Daraus folgt allerdings, dass $b_\ell = \left(\frac{n}{4\varepsilon}\right)^{\frac{2^\ell - 1}{2}}$ und damit

$b_n = 2^{\Omega(2^n)}$ für alle $\varepsilon < \frac{n}{4}$. $\square$

Das nächste Lemma zeigt, dass keine Ableitung mit kleinen Koeffizienten gibt.

Lemma 5.8

Seien $g_i, p_\ell \in \mathbb{R}[z_1, \dots, z_n]$ mit $\deg(g_i) \le d-2$, $\deg(p_\ell) \le \lfloor \frac{d}{2} \rfloor$ sodass $\sum_{i=1}^{n} g_i f_i + \sum_{\ell} p_\ell^2 = f$. Dann hat g_n einen Koeffizienten der Größe $\Omega\left(\frac{1}{n_d}\left(\frac{1}{2\varepsilon}\right)^{2^{n-1}}\right)$.

$n_d = \binom{n+d}{d}$ = # Terme vom Grad d

Beweis

Sei $\|g_n\|$ der Betrag des größten Koeffizienten von g_n.

Für ein Monom ist $\prod_{i=1}^{n} z_i^{\beta_i} \equiv z_1^{\sum_{j=1}^{n} 2^{j-1}\beta_j}$ (durch Substitution von $z_{i+1} \leftarrow z_i^2$) Bsp.: $z_1 z_2 z_3 \equiv z_1^7$.

Wir definieren eine lineare Abbildung $\phi : \mathbb{R}[z_1, \dots, z_n] \to \mathbb{R}$ mit $\phi\left(\prod_{i=1}^{n} z_i^{\beta_i}\right) := (2\varepsilon)^{\sum_i 2^{i-1}\beta_i}$.

Wir zeigen die folgenden Eigenschaften:

a) $\phi(\varepsilon - z_n) = -\varepsilon$

b) $\phi(p^2) \geq 0 \qquad \deg(p) \leq d$

c) $\phi(g_i f_i) = 0 \qquad 1 \leq i \leq n-1 \quad \deg(g_i) \leq d-2$

d) $|\phi(g_n z_n^2)| \leq (2\varepsilon)^{2^{n-1}} \cdot n_d \cdot \|g_n\|$

zu a) einfach

zu b) gilt, da $\phi(P_1 \cdot P_2) = \phi(P_1) \cdot \phi(P_2)$ für alle Polynome P_1, P_2

zu c) $\phi(g_i f_i) = \phi(g_i(z_i^2 - z_{i+1}))$

$= \phi(g_i z_i^2) - \phi(g_i \cdot z_{i+1})$ $\qquad \phi(z_i^2)$

$= \phi(g_i) \cdot \phi(z_i^2) - \phi(g_i) \cdot \phi(z_{i+1}) = 0$

zu d) $|\phi(g_n z_n^2)| = |\phi(z_n^2) \cdot \phi(g_n)|$

$= (2\varepsilon)^{2^{n-1}} \cdot \underbrace{|\phi(g_n)|}_{\leq n_d \cdot \|g_n\|}$ $\qquad$ (da $\phi(a \prod z_i^{k_i}) < a$)

$\Rightarrow -\varepsilon = \phi(f) = \phi\left(\sum_{i=1}^{n-1} g_i f_i + \sum_e p_e^2 + g_n z_n^2\right) \leq \phi(g_n z_n^2)$

$\Rightarrow \varepsilon \leq |\phi(g_n z_n^2)| \leq (2\varepsilon)^{2^{n-1}} \cdot n_d \cdot \|g_n\|$

$\Rightarrow \|g_n\| \geq \varepsilon \left(\frac{1}{2\varepsilon}\right)^{2^{n-1}} \cdot \frac{1}{n_d}$ $\qquad \square$

5.5 Sum-of-Squares simuliert Polynomkalkül

Lemma 5.9.

Wenn f eine Ableitung aus $P \cup \{x_j^2 - x_j \mid 1 \le j \le n\}$ von Grad d im Polynomkalkül $(\{0,1\}, \text{über } \mathbb{R})$ dann gibt es eine Ableitung von f^2 aus $P \cup \{x_j^2 - x_j \mid 1 \le j \le n\}$ von Grad $2d$ im SOS-Kalkül.

Beweis

Induktionsanfang: $f \in P \cup \{x_j^2 - x_j \mid 1 \le j \le n\}$ ($= \{f_i \mid 1 \le i \le m\}$) ✓

Induktionsschritt:

$$\frac{f \quad g}{af + bg}$$

Nach Induktionsannahme:

$$\sum_i g_i' f_i + \sum_j h_j' (x_j^2 - x_j) - \sum_\ell p_\ell'^2 = f^2$$

$$\sum_i g_i'' f_i + \sum_j h_j'' (x_j^2 - x_j) - \sum_\ell p_\ell''^2 = g^2$$

$q := (af - bg)$

$$\sum_i (2a^2 g_i' + 2b^2 g_i'') f_i + \sum_j (2a^2 h_j' + 2b^2 h_j'')(x_j^2 - x_j) - \sum_\ell (\sqrt{2} a p_\ell')^2 - \sum_\ell (\sqrt{2} b p_\ell'')^2 - q^2$$

$$= 2a^2 f^2 + 2b^2 g^2 - q^2$$

$$= 2a^2 f^2 + 2b^2 g^2 - a^2 f^2 + 2ab fg - b^2 g^2$$

$$= a^2 f^2 + 2ab fg + b^2 g^2 = (af + bg)^2$$

$$\frac{f}{xf}$$

Nach Induktionsannahme:

$$\sum_i g_i f_i + \sum_j h_j (x_j^2 - x_j) - \sum_\ell p_\ell^2 = f^2$$

▷ Wir könnte beide Seiten mit x^2 multiplizieren und erhielte eine SOS-Abschätzung von $x^2 f^2$, aber: das würde den Grad erhöhen.

Daher:

$$q := f - \tfrac{1}{2} xf \quad \Rightarrow \quad q^2 = f^2 - xf^2 + \tfrac{1}{4} x^2 f^2$$

$$= f^2 - \tfrac{3}{4} x^2 f^2 + f^2 \cdot (x^2 - x)$$

$\Rightarrow$

$$\sum_i \left(\tfrac{4}{3} g_i\right) f_i + \sum_j \left(\tfrac{4}{3} h_j\right)\left(x_j^2 - x_j\right) + \tfrac{4}{3} f^2 (x^2 - x) - \sum_\ell \left(\sqrt{\tfrac{4}{3}}\, p_\ell\right)^2 - \left(\sqrt{\tfrac{4}{3}}\, q\right)^2 = x^2 f^2 = (xf)^2$$

$\square$

Aus Lemma 5.3 folgt weiterhin, dass SOS das Polynomkalkül p-simuliert.

6. Frege Systeme

6.1 Frege und erweiterte Frege

Ein Frege-Kalkül ist ein vollständiger und korrekter Beweiskalkül über der Menge aller aussagenlogischen Formeln.
(Der Einfachheit halber betrachtet man adäquate Operatormenge wie $\{\vee, \neg\}$ anstelle der Menge $\{0, 1, \vee, \wedge, \neg\}$)

Bsp.: Das Schema sieht so:

$$\frac{}{\varphi \vee \neg\varphi} \quad , \quad \frac{\varphi}{\varphi \vee \psi} \quad , \quad \frac{\varphi \vee \varphi}{\varphi} \quad , \quad \frac{\varphi \vee (\varphi' \vee \varphi'')}{(\varphi \vee \varphi') \vee \varphi''}$$

$$\frac{\psi \vee \varphi \qquad \neg\varphi \vee \psi'}{\psi \vee \psi'}$$

wobei $\varphi, \varphi', \varphi'', \psi, \psi'$ aussagenlogische Formeln über $\{\vee, \neg\}$ sind.

Aus der VL "Logik in der Informatik" ist der Sequenzenkalkül bekannt (für die Prädikatenlogik erster Stufe).
Betrachtet man die Einschränkung dieses Kalküls auf aussagenlogische Formeln und fasst jede Sequenz $\Gamma \vdash \varphi$ als Formel $\bigwedge_{\psi \in \Gamma} \psi \to \varphi$ auf, kann auch dieser Kalkül als Frege-Kalkül betrachtet werden.

Satz 6.1 [Reckhow 75]

Alle Frege-Kalküle p-simulieren einander.

Außerdem sind Frege-Kalküle stärker als alle in der VL betrachtete Kalküle.
Es ist ein zentrales offenes Problem, ob Frege-Kalküle p-beschränkt sind.

Eine Erweiterung von Frege-Kalkülen ist "extended-Frege". Es enthält die zusätzliche Regel

$$\frac{\varphi}{x \leftrightarrow \varphi}$$

Abk. für: $\neg(\neg x \vee \neg\varphi) \vee \neg(x \vee \varphi)$

Wobei x eine "frische" Variable ist, die weder bisherigen im Beweis noch in den Axiomen vorkam.

Diese Regel erlaubt Formeln "abzukürzen" und damit kompakter darzustellen. Dies ist im Besonderen dann wichtig, wenn eine Teilformel mehrfach vorkommt. Es ist offen, ob extended Frege echt stärker ist als Frege.

Es wird vermutet, dass dies der Fall ist.

Eine weitere zentrale offene Frage ist, ob es ein "optimales" Beweissystem gibt, das alle anderen Systeme p-simuliert.

6.2. Bounded-Depth Frege

Die Tiefe einer Formel ist die Anzahl der Alternierungen von $\vee$ und $\wedge$ im Syntaxbaum.

(Klauseln haben Tiefe 1, KNF und DNF haben Tiefe 2, ...)

Depth-d Frege ist ein Frege Kalkül, in dem jede Formel Tiefe $\leq d$ hat.

<u>Beispiel</u>: Depth-1 Frege entspricht Resolution.

▷ Depth-d Frege ist nicht p-beschränkt für

- $d = O(1)$ Ajtai 1994
- $d = O(\log^* n)$ Bellantoni, Pitassi, Urquhart 1992
- $d = O(\log \log n)$ Pitassi, Beame, Impagliazzo 1993
  Krajíček, Pudlák, Woods 1995
  Urquhart, Fu 1996
  Ben-Sasson 2002

} PHP_n^{n+1}

} Zeitz

- $d = O(\sqrt{\log n})$ Pitassi, Rossman, Servedio, Tan 2016