Randomness vs structure

DNF compression and sunflower lemma

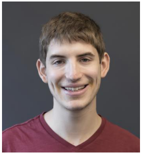
Kewen Wu
Peking University

Joint works with

Ryan Alweiss
Princeton

Shachar Lovett
UCSD

Jiapeng Zhang
Harvard



Overview

- This talk is about two recent works accepted by STOC 2020.
  - Decision list compression by mild random restrictions
    Shachar Lovett, **Kewen Wu**, Jiapeng Zhang
  - Improved bounds for the sunflower lemma
    Ryan Alweiss, Shachar Lovett, **Kewen Wu**, Jiapeng Zhang
- The main approach is to study *mild random restrictions*.

Small-width DNFs simplify under (mild) random restrictions.

DNF (disjunctive normal form)

- literal: x_i or $\neg x_i$
- term: a conjuction of literals
- DNF: a disjunction of terms
- width of a DNF: maximum number of literals in a term
- size of a DNF: number of conjunctions

Example

$(x_1 \wedge x_2) \vee (\neg x_1 \wedge x_3 \wedge x_5)$ is a DNF of width 3 and size 2.

Section 2

DNF compression

What we proved

Definition (ε-approximation)

f is ε-approximated by g if $\Pr_{x\sim\{0,1\}^n}[f(x) \neq g(x)] \leq \varepsilon$.

Theorem (DNF compression)

Width-w DNF can be ε-approximated by a width-w size-s DNF.

- Gopalan, Meka and Reingold 2013: $s = (w \log(1/\varepsilon))^{O(w)}$.
- Lovett and Zhang 2019: $s = (1/\varepsilon)^{O(w)}$.
- Now: $s = \left(2 + \frac{1}{w} \log \frac{1}{\varepsilon}\right)^{O(w)}$ and this is tight.

Applications

- Decision list compression
  Small-width decision lists can be approximated by decision lists of same width and small size.

  If $C_1(x) = \texttt{True}$ **then** output v_1,
  $\ldots$,
  else if $C_{m-1}(x) = \texttt{True}$ **then** output v_{m-1},
  else output default value v_m.
- Junta theorem
  Small-width DNFs can be approximated by DNFs of same width and few input bits.
- Learning small-width DNFs
  Small-width DNFs are efficiently PAC learnable.

How we proved – more definitions

Let $f = C_1 \vee \cdots \vee C_m$ be a DNF.

Definition (Index function Ind f)

Ind $f(x)$ is the index of first satisfied term (or $\perp$ if $f(x) = 0$).

Definition (Useful index)

Index i is useful if there exists x such that Ind $f(x) = i$.
#useful (f) is the number of useful indices.

Example

Assume $f = (x_1 \wedge x_2) \vee (x_2) \vee (x_1 \wedge x_2 \wedge \neg x_3)$. Then Ind $f(1, 1, 0) = 1$ and #useful $(f) = 2$.

How we proved – Step 1: randomness kills structure

We should be able to compress f (in some sense) under restrictions.

Lemma (Håstad's switching lemma 1987)

Let f be a width-w DNF, $\alpha \in (0,1)$, and d be an integer. If ρ randomly restrict each input bit to $0, 1, *$ w.p. $(1-\alpha)/2, (1-\alpha)/2, \alpha$, then

$$\Pr_{\rho}\left[\mathrm{DecisionTree}(f\restriction_{\rho}) \geq d\right] \leq (5\alpha w)^{d}.$$

- Prove by encoding bad ρ.
- Meaningful only when $\alpha \leq O(1/w) \implies$ most bits are fixed.

Mild random restrictions

Let's directly analyze f's size under restrictions.

Lemma

Let f be a width-w DNF, $\alpha \in (0,1)$, and s be an integer. If ρ randomly restrict each input bit to $0, 1, *$ w.p. $(1-\alpha)/2, (1-\alpha)/2, \alpha$, then

$$\Pr_{\rho}\left[\#\textit{useful}(f \restriction_\rho) \geq s\right] \leq \frac{1}{s}\left(\frac{4}{1-\alpha}\right)^w .$$

- Prove by encoding bad ρ, $(\rho, i) \to (\rho', \mathsf{aux})$.
  ρ' activates $*$'s in $C_i \restriction_\rho$ for useful $i \implies \mathsf{Ind}f(\rho') = i$.
- Meaningful for all kinds of α.

How we proved – Step 2: intuition for compression

Let $f = C_1 \vee \cdots \vee C_m$ be a width-w DNF.

- If index i is not useful, we can safely remove C_i.
- Assume $p_i = \Pr_x[\mathsf{Ind}\, f(x) = i]$ is decreasing in i.
  If p_i decrease quickly, we only need to keep the top ones.
- Let $g = C_1 \vee \cdots \vee C_t$. Then

$$\Pr[f(x) \neq g(x)] = \Pr[\mathsf{Ind}\, f(x) > t] = \sum_{i>t} p_i.$$

Now what?

Let $f = C_1 \vee \cdots \vee C_m$ be a width-w DNF.

- What we can do so far?
  We can analyze $q_i(\alpha) = \Pr_\rho[\text{index } i \text{ is useful in } f\restriction_\rho]$, since

$$\sum_i q_i(\alpha) = \mathop{\mathbb{E}}_\rho[\#\text{useful}(f\restriction_\rho)].$$

- What we want to do next?
  We want to bound $p_i = \Pr_x[\mathrm{Ind}f(x) = i]$.

How we proved – Step 3: noise stability

Let's introduce noise stability.

Definition (Noise distribution $\mathcal{N}_\beta$)

$y \sim \mathcal{N}_\beta(x)$ is sampled by taking $\Pr[y_i = x_i] = (1+\beta)/2$.

Then for $x \sim \{0,1\}^n, y \sim \mathcal{N}_\beta(x)$, we can also do it as

1. sample common restriction ρ with $\Pr\left[\rho_i = *\right] = 1 - \beta$;
2. sample x' by uniformly filling out $*$'s in ρ, and set $x = \rho \circ x'$;
3. sample y' by uniformly filling out $*$'s in ρ, and set $y = \rho \circ y'$.

How we proved – Step 4: bridging lemma

Let $f = C_1 \vee \cdots \vee C_m$ be a width-w DNF and fix i.
Sample $x \sim \{0,1\}^n, y \sim \mathcal{N}_\beta(x)$, which can be seen as ρ, x', y'.

Define $\mathsf{Stab}(\beta) = \Pr\left[\mathsf{Ind}f(x) = \mathsf{Ind}f(y) = i\right]$ and
recall $p = \Pr[\mathsf{Ind}f(x) = i]$, $q = \Pr[\text{index } i \text{ is useful in } f\restriction_\rho]$.

Fact (Hypercontractivity)

$\mathit{Stab}(\beta) \le \left(\Pr\left[\mathit{Ind}f(x) = i\right]\right)^{\frac{2}{1+\beta}} = p^{\frac{2}{1+\beta}}$.

We also have

$$
\begin{aligned}
\mathsf{Stab}(\beta) &= \Pr\left[\mathsf{Ind}f(x) = \mathsf{Ind}f(y) = i, \text{index } i \text{ is useful in } f\restriction_\rho\right] \\
&= q \Pr\left[\mathsf{Ind}f(x) = \mathsf{Ind}f(y) = i \mid \text{index } i \text{ is useful in } f\restriction_\rho\right] \\
&\ge q\left(\Pr\left[\mathsf{Ind}f(x) = i \mid \text{index } i \text{ is useful in } f\restriction_\rho\right]\right)^2 \\
&= p^2/q.
\end{aligned}
$$

Section 3

Sunflower lemma

What we proved

Definition (w-set system and r-sunflower)

A w-set system is a family of sets of size at most w.
An r-sunflower is r sets with same pairwise intersection.

Theorem (Erdős-Rado sunflower lemma)

Any w-set system of size s has r-sunflower.

Let's focus on $r = 3$.

- Erdős and Rado 1960: $s = w! \cdot 2^w \approx w^w$.
- Kostochka 2000: $s \approx (w \log\log\log w / \log\log w)^w$.
- Fukuyama 2018: $s \approx w^{0.75w}$.
- Now: $s \approx (\log w)^w$ and this is tight for our approach.

Applications

- Combinatorics
  - Intersecting set systems
  - Erdős-Szemerédi sunflower lemma
  - Alon-Jaeger-Tarsi conjecture
  - Random graph
- Theoretical computer science
  - Circuit lower bounds and data structure lower bounds
  - Matrix multiplication
  - Pseudorandomness: DNF compression
  - Cryptography
  - Property testing
  - Fixed parameter complexity

How we proved – Step 1: make it robust

Assume $\mathcal{F} = \{S_1, \dots, S_m\}$ is a w-set system. Define a width-w DNF $f_{\mathcal{F}}$ as $f_{\mathcal{F}} = \bigvee_{i=1}^{m} \bigwedge_{j \in S_i} x_j$.

Definition (Satisfying system)

$\mathcal{F}$ is satisfying if $\Pr[f_{\mathcal{F}}(x) = 0] < 1/3$ with $\Pr[x_i = 1] = 1/3$, i.e., $\Pr[\forall i \in [m], S_i \not\subset S] < 1/3$ with $\Pr[x_i \in S] = 1/3$.

Lemma

If $\mathcal{F}$ is satisfying, then it has 3 pairwise disjoint sets (3-sunflower).

Prove by randomly 3-coloring x and union bound.

How we proved – Step 2: induction

Assume $\mathcal{F} = \{S_1, \dots, S_m\}, m > \kappa^w$ is a w-set system. Define link $\mathcal{F}_Y = \{S_i \backslash Y \mid Y \subset S_i\}$, which is a $(w - |Y|)$-set system.

If there exists Y such that $|\mathcal{F}_Y| \geq m/\kappa^{|Y|} > \kappa^{w-|Y|}$, then we can apply induction and find 3-sunflower in $\mathcal{F}_Y$.

So induction starts at such $\mathcal{F}$, that $|\mathcal{F}_Y| < m/\kappa^{|Y|}$ holds for any Y. Thus it suffices to prove

Lemma

Let $\kappa \geq (\log w)^{O(1)}$. If $|\mathcal{F}_Y| < m/\kappa^{|Y|}$ holds for any Y, then $\mathcal{F}$ is satisfying, which means there are 3 pairwise disjoint sets in $\mathcal{F}$.

How we proved – Step 3: randomness preserves structure

Assume $\mathcal{F} = \{S_1, \dots, S_m\}, m > \kappa^w$ is a w-(multi-)set system.
Assume $|\mathcal{F}_Y| < m/\kappa^{|Y|}$ holds for any Y. $\Leftarrow \mathcal{F}$ is structured
Take $\approx 1/\sqrt{\kappa}$-fraction of the ground set as W,
and construct a $w/2$-(multi-)set system $\mathcal{F}'$ from each S_i:

- Good: If there exists $|S_j \backslash W| \le w/2$ and $S_j \backslash W \subset S_i \backslash W$, then put $S_j \backslash W$ into $\mathcal{F}'$; (j may equal i)
  To satisfy $\{\{x_1, x_2\}, \{x_1, x_2, x_3, x_5\}, \{x_4\}\}$, it suffices to satisfy $\{\{x_1, x_2\}, \{x_1, x_2\}, \{x_4\}\}$.
- Bad: otherwise, we do nothing for S_i.

Then $|\mathcal{F}'| \approx m$ and $|\mathcal{F}'_Y| \le |\mathcal{F}_Y|, \forall Y$. $\Leftarrow \mathcal{F}'$ is also structured
Prove by encoding bad $(W, i) \to (W' = W \cup S_i, k, \text{aux})$,
where $S_j \backslash W \subset S_i \backslash W$ and S_i ranks $k < m/\kappa^{w/2}$ in $\mathcal{F}_{S_j \cap S_i}$.

Assume $\mathcal{F} = \{S_1, \dots, S_m\}, m > \kappa^w$ is a w-(multi-)set system. Assume $|\mathcal{F}_Y| < m/\kappa^{|Y|}$ holds for any Y. Assume $\kappa = (\log w)^{O(1)}$.

- $\Pr[x_i \in S] = 1/3$
  $\approx$ take $1/3$-fraction of the ground set as S
  $\approx$ view S as $W_1, W_2, \dots, W_{\log w}$, each of $\approx 1/\sqrt{\kappa}$-fraction

Then

$$\mathcal{F} \xrightarrow{W_1} \mathcal{F}' \xrightarrow{W_2} \mathcal{F}'' \xrightarrow{W_3} \cdots \xrightarrow{W_{\log w}} \mathcal{F}^{\text{last}}.$$

- either we stop at W_i when some set is contained in $\bigcup_{j<i} W_j$,
- or, $\mathcal{F}^{\text{last}}$ is a width-0 (multi-)set system of size $\approx m > \kappa^w$ and still $\left|\mathcal{F}^{\text{last}}_Y\right| \lesssim \left|\mathcal{F}^{\text{last}}\right| / \kappa^{|Y|}$ holds for any Y. impossible

Thus, (informally) we proved such $\mathcal{F}$ is satisfying, which means $\mathcal{F}$ has 3-sunflower (3 pairwise disjoint sets).

Section 4

Open problems

Problem 1 – Upper bound compression

Compressed DNF g is constructed by removing several terms from DNF f, thus $g(x) \leq f(x)$.

Problem (Upper bound compression)

Width-w DNF can be ε-approximated by a width-w size-s DNF from above. $(g(x) \geq f(x))$

- Gopalan, Meka and Reingold 2013: $s = (w \log(1/\varepsilon))^{O(w)}$.
- Lovett, Solomon and Zhang 2019: restricted in monotone case, $s = (\log w/\varepsilon)^{O(w)}$ implies improved sunflower lemma.

Now we have the improved sunflower lemma, can we do better?

Problem 2 – Erdős-Rado sunflower

Problem (Erdős-Rado sunflower conjecture)

Any w-set system of size $O_r(1)^w$ has r-sunflower.

- Our robust sunflower cannot overcome $(\log w)^{(1-o(1))w}$. We need new ideas.
- Lift the sunflower size? $r = 3 \implies r = 4$
- Is $(\log w)^{(1-o(1))w}$ actually tight? Counterexamples?

Problem 3 – Erdős-Szemerédi sunflower

Assume $\mathcal{F} = \{S_1, \ldots, S_m\}$ and $S_i \subset \{1, 2. \ldots, n\}$.

Problem (Erdős-Szemerédi sunflower conjecture)

There exists function $\varepsilon = \varepsilon(r) > 0$, such that, if $m > 2^{n(1-\varepsilon)}$, then $\mathcal{F}$ has r-sunflower.

- Now:
  - general r: $\varepsilon = O_r\left(\log n\right)^{-(1+o(1))}$ from ER sunflower.
  - $r = 3$: Naslund proved it using polynomial method.
- ER sunflower conjecture $\implies$ ES sunflower conjecture.

Section 5

Thanks