Computer Science 294 Lecture 11 Notes

Daniel Raban

February 21, 2023

1 Proof of Håstad's Switching Lemma

1.1 Argument via encoding

Last time we introduced Håstad's switching lemma.

Lemma 1.1 (Håstad's switching lemma). *Suppose f is a width w DNF, and let $(J, Z) \sim \mathcal{R}_p$. Then for all k,*

$$\mathbb{P}(\textit{Decision Tree Depth}(f_{J,Z}) \geq k) \leq (5pw)^k.$$

We want to think of $p \approx 1/(10w)$ so that we get exponential decay in k. The argument we will give is not Håstad's original argument, and we will only get a bound of $(9pw)^k$. Here is the idea of the argument.

Let the set of bad restrictions be $\mathrm{BAD} = \{(J, z) : \mathrm{DT\ depth}(f_{J,z}) \geq k\}$. We want to show that $\mathbb{P}_{(J,Z)\sim\mathcal{R}_p}((J, Z) \in \mathrm{BAD})$ is small. The naive idea is that to show that $\mathbb{P}(A)$ is small, it suffices to show that there exists some event B such that $\mathbb{P}(B) \geq M\mathbb{P}(A)$ for some large M; therefore,

$$\mathbb{P}(A) \leq \frac{\mathbb{P}(B)}{M} \leq \frac{1}{M}.$$

The main step in the proof will be the prove the following encoding lemma.

Lemma 1.2 (Encoding lemma). *There exists an injective encoding*

$$E : \mathrm{BAD} \to (\textit{all restrictions}) \times [2w]^k \times \{\pm 1\}^k.$$

Moreover,

$$E(\rho) = (\rho \circ \sigma, \beta, \pi),$$

where σ is a restriction that fixes k additional variables.

Here, we should think of the $[2w]^k \times \{\pm 1\}^k$ part as extra information that would allow E to be 1 to 1.

For any fixed restriction $\rho = (J, z)$, we denote by the **weight** $\mathrm{wt}(\rho)$ the probability to sample ρ when sampling from $\mathcal{R}_p$:

$$\mathrm{wt}(\rho) = p^{|J|} \left(\frac{1-p}{2}\right)^{n-|J|}.$$

Example 1.1. If $\rho = (*, *, +1, -1, +1)$, then the weight is

$$\mathrm{wt}(\rho) = p^2 \left(\frac{1-p}{2}\right)^3.$$

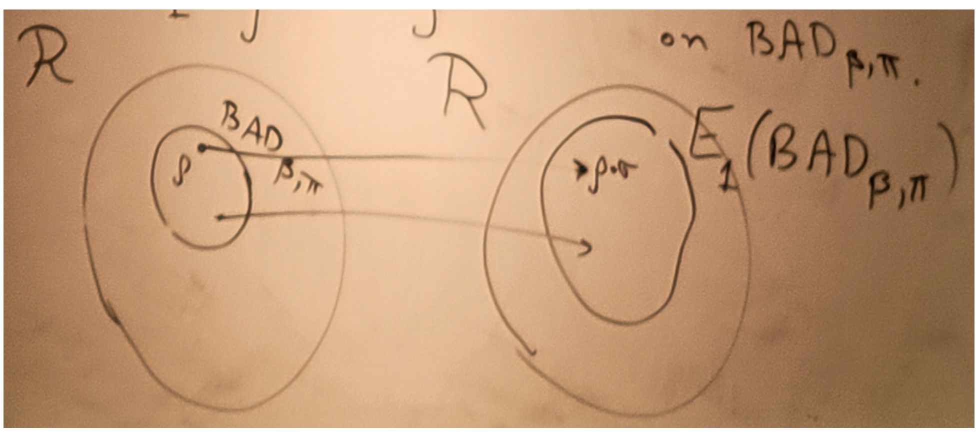
For a set of restrictions S, the **weight** is

$$\mathrm{wt}(S) = \sum_{\rho \in S} \mathrm{wt}(\rho).$$

Proof of switching lemma from encoding lemma. Fix β, π. Consider the set

$$\mathrm{BAD}_{\beta,\pi} = \{\rho \in \mathrm{BAD} : \exists \rho' \text{ such that } E(\rho) = (\rho', \beta, \pi)\}.$$

Then the encoding $E_1 : \rho \mapsto \rho \circ \sigma$ is still 1 to 1 on $\mathrm{BAD}_{\beta,\pi}$. We will show that the probability of $E_1(\mathrm{BAD}_{\beta,\pi})$ is much bigger than the probability of $\mathrm{BAD}_{\beta,\pi}$.



The weight of $\rho = (J, z)$ is

$$\mathrm{wt}(\rho) = p^{|J|} \left(\frac{1-p}{2}\right)^{n-|J|},$$

while the weight of $\rho \circ \sigma = (J', z')$ (with $|J'| = |J| - k$) is

$$\begin{aligned} \mathrm{wt}(\rho \circ \sigma) &= p^{|J'|} \left(\frac{1-p}{2} \right)^{n-|J'|} \\ &= p^{|J|-k} \left(\frac{1-p}{2} \right)^{n-|J|+k} \\ &= \mathrm{wt}(\rho) \left(\frac{1-p}{2p} \right)^{k}. \end{aligned}$$

Therefore,

$$\mathrm{wt}(\mathrm{BAD}_{\beta,\pi}) = \mathrm{wt}(E_1(\mathrm{BAD}_{\beta,\pi})) \left(\frac{2p}{1-p} \right)^{k},$$

and we get

$$\begin{aligned} \mathbb{P}((J,Z) \in \mathrm{BAD}_{\beta,\pi}) &= \mathbb{P}((J,Z) \in E_1(\mathrm{BAD}_{\beta,\pi})) \cdot \left(\frac{2p}{1-p} \right)^{k} \\ &\leq \left(\frac{2p}{1-p} \right)^{k}. \end{aligned}$$

Taking a union bound over (β, π) ($(4w)^k$ options),

$$\begin{aligned} \mathbb{P}_{(J,Z)\sim\mathcal{R}_p}((J,Z) \in \mathrm{BAD}) &\leq (4w)^k \left(\frac{2p}{1-p} \right)^{k} \\ &= \left(\frac{8pw}{1-p} \right)^{k}. \end{aligned}$$

If $p \geq 1/9$, then we get

$$\mathbb{P}_{(J,Z)\sim\mathcal{R}_p}((J,Z) \in \mathrm{BAD}) \leq (9pw)^k.$$

If $p \leq 1/9$, then

$$\mathbb{P}_{(J,Z)\sim\mathcal{R}_p}((J,Z) \in \mathrm{BAD}) \leq \left(\frac{8pw}{8/9} \right)^{k} = (9pw)^k.$$

So we get the desired bound. □

1.2 Proof of the encoding lemma

Here is an example of how the encoding works.

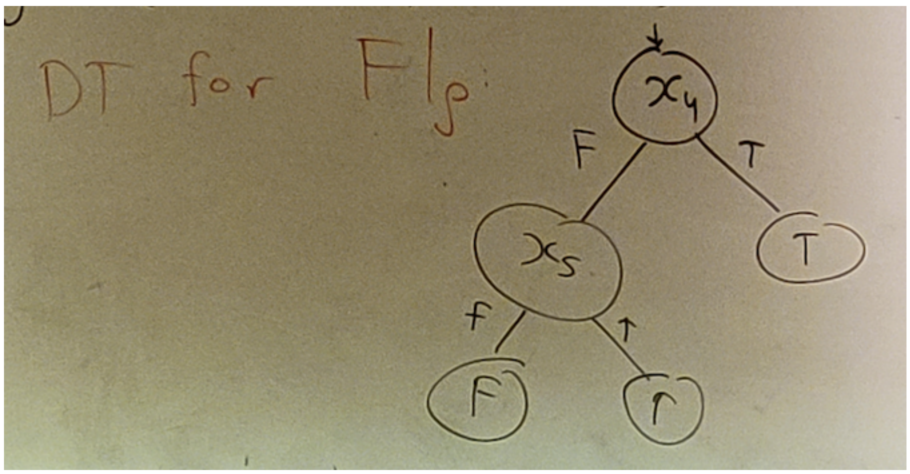
Example 1.2. Suppose we have

$$F = (x_1 \wedge x_2 \wedge x_3) \vee (x_3 \wedge x_4) \vee (x_5 \wedge \overline{x_1} \wedge x_3)$$

and ρ sets x_1 to False and x_3 to True; that is, $\rho = (F, *, T, *, *)$. Then F becomes

$$F|_\rho = x_4 \vee x_5.$$

The decision tree for $F|_\rho$ looks like



If we have $k = 1$ and $\rho' = (F, *, T, T, *)$, then we need to "leave a trail of breadcrumbs" to help us figure out what extra restriction we made and what the original ρ was.

Proof of the encoding lemma. First, we do the case of $k = 1$. The restriction $\rho \in \text{BAD}$ iff DT depth$(f|_\rho) \geq 1$. Equivalently, $f|_\rho$ is not a constant function. Scanning from left t o right, find the first term T_1 such that $T_1|_\rho \not\equiv$ False. Let v_1 be an alive variable in T_1, and let σ_1 assign v_1 so that T_1 is still not falsified. In this case, the mapping should be

$$\rho \mapsto (\rho \circ \sigma_1, \text{location of } v_1 \text{ in } T_1),$$

where the location of v_1 in T_1 is a number in $\{1, 2, \ldots, w\}$.

How do we decode this encoding? Given $\rho' = \rho \circ \sigma$, find the first term T_1' such that $T_1'|_{\rho'} \not\equiv$ False. Then $T_1 = T_1'$. Identify v_1 from the additional information. Make v_1 alive again to recover ρ.

Now we treat the case of $k > 1$. Given a DNF F and a restriction $\rho = (J, z)$, let the **canonical decision tree** of (F, ρ) be

For $i = 1, 2, \ldots,$

Look through F for the first term T_i such that $T_i|_\rho \not\equiv$ False.

If no such term exists, output False.

Otherwise: Let A_i be the set of alive variables in T_i under ρ.

Query all variables in A_i.
Let $\pi_i \in \{\pm 1\}^{|A_i|}$ be the answers.
If $T_i|_\rho$ is satisfied by π_i, then output True.
Else, extend ρ by $\rho \circ (\pi_i \to A_i)$.

Here are two ways to assign variables:

1. π: the "adversarial" strategy that ensures $\mathrm{CDT}(F|_\rho) \geq k$.

2. σ: the "breadcrumbs" strategy that allows decoding.

If $\rho \in \mathrm{BAD}$, then DT depth$(f|_\rho) \geq k$, so there exists a path of length $\geq k$ in any decision tree for $f|_\rho$. In particular, there is a (partial) path of length $= k$ in $\mathrm{CDT}(f|_\rho)$. Here is how we encode (F, ρ):

Let $T_1, T_2, \ldots, T_t$ be the terms considered in this path.

Let $A_1, \ldots, A_t$ be the sets of variables set in each of these terms.

Let $\pi_1 \in \{\pm 1\}^{|A_1|}, \pi_2 \in \{\pm 1\}^{|A_2|}, \ldots, \pi_t \in \{\pm 1\}^{|A_t|}$ be the values assigned to these variables along the path.

In our example, $T_1 = (x_3 \wedge x_4)$, $T_2 = (x_5 \wedge \overline{x_1} \wedge x_3)$, $A_1 = \{x_4\}$, $A_2 = \{x_5\}$, $\pi_1 = F$, and $\pi_2 = T$.

Calculate $T_1, \ldots, T_t, A_1, \ldots, A_t, \pi_1, \ldots, \pi_t$.

For $i = 1, \ldots, t$:

For each variable in A_i, encode as β_i its location in T_i ($\in [w]$) and whether or not it is the last bit.

Set σ_i to be the assignment to A_i that doesn't falsify $T_i|_\rho$ (usually set $T_i|_\rho$ to true)

Replace ρ by $\rho \circ \sigma_i$. $\square$

In our example, we get $\beta_1 = 2$, $\beta_2 = 1$, $\sigma_1 = (x_4 = T)$, and $\sigma_2 = (x_5 = T)$. This gives

$$\rho = (F, *, T, *, *), \qquad \rho \circ \sigma_1 \circ \sigma_2 = (F, *, T, T, T).$$