Bug in LICS paper

Aaryan Gupta

June 27 2022

1 Some definitions

The set $S \subseteq \{0,1\}^n$ contains boolean strings of length n, corresponding to solutions of a CNF formula F.

The bound provided in Cyrus's thesis is on $e_{(b,a)}(S)$, which is different the bound we need which is $c_S(w)$. Both the quantities are mathematically defined as follows:

$$e_{(b,a)}(S) = |\{(x,y) \in E_{\leq w}(S) : |x \backslash y| = b, |y \backslash x| = a\}|$$

where $E_{\leq w}(S) = \{(x,y) : x, y \in S, d_H(x,y) \leq w\}$ and d_H represents the hamming distance function.

$$c_S(w) = |\{(x,y) : x, y \in S, d_H(x,y) = w\}|$$

Note the difference in the equality and inequality on the hamming distance. Some more definitions are as stated below:

$$\ell = \lfloor \log |S| \rfloor$$

$$\beta = \left\lfloor \left(\frac{n}{log|S|} \right)^{\frac{1}{2}} l \right\rfloor$$

For $x \in S$, we define

$$\ell_x = |x \cap \{\beta+1, ..., n\}|$$

2 The argument in the LICS paper

Lemma.
For a left-compressed and down set S, $c_S(w) \leq 2 \cdot \left(\frac{8e\sqrt{n \cdot \ell}}{w}\right)^w \cdot |S|$ where $\ell = \lceil \log |S| \rceil$.

Proof.
The proof is based on the bounds derived by Rashtchian in his thesis. We give a few more details as we will need them later when we explain our implementation. More specifically, the proof uses Equations 4.2, 4.5, 4.8, and 4.10 from his thesis. I have briefly summarised and stated the equations below for reference:

Eq 4.2. For a fixed $x \in S$ and for each $p \in [a] \cup \{0\}$, the number of $y \in \{0,1\}^n$ such that $(x,y) \in E_{(b,a)}(S)$ and $\ell_y \leq \ell_x$ and $|(y \backslash x) \cap \{\beta+1, ..., n\}| = p$ is at most

$$\binom{n-\beta-\ell_x}{p}\binom{\ell_x}{p}\binom{\beta-|x|+\ell_x}{a-p}\binom{|x|}{b-p}$$

Eq 4.5. For a fixed $y \in S$ and for each $p \in [a]$, the number of $y \in \{0,1\}^n$ such that $(x,y) \in E_{(b,a)}(S)$ and $\ell_y > \ell_x$ and $|(x \backslash y) \cap \{\beta+1, ..., n\}| = p-1$ is at most

$$\binom{n-\beta-\ell_y}{p-1}\binom{\ell_y}{p}\binom{\beta-|x|+\ell_y}{b-p+1}\binom{|y|}{b-p}$$

Eq 4.8. For even w,

$$e_{(b,a)}(S) \leq \left(\frac{8e\sqrt{nl}}{w}\right)^w$$

Eq 4.10. For odd w,

$$e_{(b,a)}(S) \leq \left(\frac{8e}{w}\right)^w . (n\ell)^{\frac{w-1}{2}} . \log |S|$$

It is crucial to note that these equations hold only for a left-compressed and down set and not for an arbitrary set S. The proof follows by breaking into two cases based on the parity of w.

For even $w = 2t$, Rashtchian upper bounds the expressions obtained in Eq. 4.2 and 4.5 by Eq 4.8 in his theis. We rewrite Eq 4.8 by substituting $2t$ by w to obtain $c_S(w) \leq 2 \cdot \left(\frac{8e\sqrt{n \cdot \ell}}{w}\right)^w$. For odd w, Rashtchian upper bounds the upper bound for $c_S(w)$ obtained in Eq. 4.2 and 4.5 by Eq 4.10. We rewrite Eq 4.10 by noting that $w = 2t + 1$ to obtain $c_S(w) \leq 2 \cdot \left(\frac{8e}{w}\right)^w (\sqrt{n \cdot \ell})^{(w-1)} \ell$. Noting that $\ell \leq \sqrt{n \cdot \ell}$, we have $c_S(w) \leq 2 \cdot \left(\frac{8e\sqrt{n \cdot \ell}}{w}\right)^w$. Thus, combining these cases, we get our lemma.

3 The inconsistency

The statement " We rewrite Eq 4.8 by substituting $2t$ by w to obtain $c_S(w) \leq 2 \cdot \left(\frac{8e\sqrt{n \cdot \ell}}{w}\right)^w$" in the paper is inconsistent because upon substituting the $2t$ by w, we get the bound

$$e_{(b,a)}(S) \leq \left(\frac{8e\sqrt{nl}}{w}\right)^w$$

This bound is on $e_{(b,a)}(S)$, not $c_S(w)$, however intimately they might be connected and the paper does not explain the factor of 2 showing up.

4 The fix

We make use of the fact that for $(x, y) \in E_{(b,a)}(S)$, we have $d_H(x, y) = b + a$ as $|x \backslash y| = b, |y \backslash x| = a$. This in turn implies that

$$c_S(w) = \sum_{(b,a) \in U*} e_{(b,a)}(S)$$

where $U^* = \{(b, a) : b \geq a, b + a = w\}$. Now, we see that $|U^*| \leq \frac{w}{2}$. For even w, this leads to the line of thought

$$c_S(w) = \sum_{(b,a) \in U^*} e_{(b,a)}(S) \leq |U^*|.\max_{\text{over } (b,a) \in U^*} e_{(b,a)}(S) \leq \frac{w}{2} \cdot \left(\frac{8e\sqrt{n \cdot \ell}}{w}\right)^w$$

This new bound has a factor of $\frac{w}{2}$ instead of the old factor of 2. For odd w, we again have :

$$c_S(w) = \sum_{(b,a) \in U^*} e_{(b,a)}(S) \leq |U^*|.\max_{\text{over } (b,a) \in U^*} e_{(b,a)}(S) \leq \frac{w}{2} \cdot \left(\frac{8e}{w}\right)^w .(n\ell)^{\frac{w-1}{2}} . \log |S| = \frac{w}{2} \cdot \left(\frac{8e}{w}\right)^w .(n\ell)^{\frac{w-1}{2}} .\ell$$

We use the inequality $\ell \leq \sqrt{n\ell}$, to conclude that for all $w \leq n$ we have

$$c_S(w) \leq \frac{w}{2} \cdot \left(\frac{8e\sqrt{n \cdot \ell}}{w}\right)^w$$