Project in 236379

28.2.2021

Staff:
Associate Professor Eitan Yaakobi
Daniella Bar-Lev

Students:
Raïssa Nataf
Tomer Tsachor

In this project we define an optimal decoder for the 1-Insertion channel. Assumptions on the channel: a bit 0 or 1 is added (each one with probability 0.5) in a random position (with equal probability). We show that among all the decoders that keep the output channel length $(n+1)$ the lazy decoder is preferable over any other decoder. We then define a decoder that is optimal when the output of the decoder may be of any length.

Definition 0. We define the decoder $\mathcal{D}_{Lazy} : (\Sigma_2)^{n+1} \to (\Sigma_2)^{n+1}$:

$\mathcal{D}_{Lazy}(y) = y$.

Lemma 1. The average decoding error probability of the lazy decoder $\mathcal{D}_{Lazy}$ under the 1-insertion channel 1-Ins is $P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}_{Lazy}, d_L) = \frac{1}{n}$.

Proof: The average decoding error probability of the lazy decoder for each codeword c is calculated as follows:

$P_{err}(c, d_L) = \Sigma_{y:\mathcal{D}_{Lazy}(y)\neq c} \frac{d_L(\mathcal{D}_{Lazy}(y),c)}{|c|} p(y|c) = \Sigma_{y\in I_1(c)} \frac{1}{n} p(y|c) = \frac{1}{n}$.

Since this is true for every $c \in \mathcal{C}$, we get that $P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}_{Lazy}, d_L) = \frac{1}{n} \cdot |\mathcal{C}| \cdot \frac{1}{|\mathcal{C}|} = \frac{1}{n} \blacksquare$.

We can now show that the lazy decoder is preferable, with respect to the average decoding error probability, over any decoder that outputs a word of the same length as its input.

Lemma 2. Let $\mathcal{D} : (\Sigma_2)^{n+1} \to (\Sigma_2)^{n+1}$ be a general decoder that preserves the channel's output length. It follows that

$P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}, d_L) \geq P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}_{Lazy}, d_L)$,

and for $\mathcal{C} = (\Sigma_2)^n$ equality is obtained if and only if $\mathcal{D}_{Lazy} = \mathcal{D}$.

Proof: Equality is trivial when $\mathcal{D}_{Lazy} = \mathcal{D}$. Furthermore, since for every $y \in \mathcal{C}$ it holds that $|\mathcal{D}(y)| = n + 1$, it is deduced that $d_L(c, \mathcal{D}(y)) \geq 1$. Hence, similarly to the proof of Lemma 1, it is easy to verify that

$P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}, d_L) \geq \frac{1}{n} = P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}_{Lazy}, d_L)$

where the last equality follows from Lemma 1.

Let us now assume that $\mathcal{D} \neq \mathcal{D}_{Lazy}$, i.e., there exists $z \in (\Sigma_2)^{n+1}$ such that $\mathcal{D}(z) = z' \neq z$.

Since $z' \neq z$ we get that $D_1(z') \neq D_1(z)$, i.e., there exists a word $c \in (\Sigma_2)^n$ such that $c \in D_1(z)$ and $c \notin D_1(z')$. Equivalently, $z \in I_1(c)$ and $z' \notin I_1(c)$ and so $d_L(c, z') \geq 3$ (at least one more insertion and one more deletion are needed in addition to the deletion needed for every word in the insertion ball).

Hence, it is derived that

$$
\begin{aligned}
P_{err}(c, d_L) &= \Sigma_{y\in I_1(c)} \frac{d_L(\mathcal{D}(y), c)}{n} p(y|c) \\
&\geq \Sigma_{y\in I_1(c)\setminus\{z\}} \frac{1}{n} p(y|c) + \Sigma_{y\in I_1(c)} \frac{d_L(\mathcal{D}(z) = z', c)}{n} p(y|c) \\
&> \Sigma_{y\in I_1(c)} \frac{1}{n} p(y|c) = \frac{1}{n}.
\end{aligned}
$$

If $\mathcal{C} = (\Sigma_2)^n$ it must hold that $c \in \mathcal{C}$, and so

$P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}, d_L) \geq \frac{|\mathcal{C}|-1}{|\mathcal{C}|} \cdot \frac{1}{n} + \frac{1}{|\mathcal{C}|} P_{err}(c, d_L) > \frac{1}{n}$.

Combining with Lemma 1 again completes the proof. ■

Before examining the performance of the embedding number decoder, we first discuss its properties over the 1-insertion channel. It is first shown that a decoder that shortens an arbitrary run of maximal length within the input word is equivalent to the embedding number decoder.

Lemma 3. Given $y \in (\Sigma_2)^{n+1}$, the word $\hat{x} \in (\Sigma_2)^n$obtained by reducing a run of maximal length in y satisfies $\text{Emb}(\hat{y}; x) = \max_{x \in \Sigma_2^n} \{\text{Emb}(y; x)\}$.

Proof: Let y be a word with n_r runs of lengths $r_1, r_2, ..., r_{n_r}$. Let y_i $1 \leq i \leq n_r$, be the word obtained from y by reducing the i_{th} run by one, and so $\text{Emb}(y; x_i) = r_i$. Hence, it follows that $\arg\max_{1 \leq i \leq n_r} \{\text{Emb}(y; x_i)\} = \arg\max_{1 \leq i \leq n_r} \{r_i\}$.■

Definition 4. The embedding number decoder $\mathcal{D}_{EN}$ shortens the first run of maximal length in y by one. A decoder $\mathcal{D}$ that shortens one of the runs of maximal length in y by one is said to be equivalent to the embedding number decoder, and is denoted by $\mathcal{D} \equiv \mathcal{D}_{EN}$.

Lemma 5. Let $\mathcal{D} : (\Sigma_2)^{n+1} \to (\Sigma_2)^n$ be a general decoder that reduces the input length by one. It follows that $P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}, d_L) \geq P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}_{EN}, d_L)$.

Proof:

$$
\begin{aligned}
P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}, d_L) &= \frac{1}{|\mathcal{C}|} \Sigma_{c \in \mathcal{C}} \Sigma_{y: \mathcal{D}(y) \neq c} \frac{d_L(\mathcal{D}(y), c)}{|c|} p(y|c) \\
(1) &= \frac{1}{|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \Sigma_{c \in D_1(y)} \frac{d_L(\mathcal{D}(y), c)}{|c|} p(y|c) \\
(2) &\overset{b}{\geq} \frac{1}{|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \frac{2}{n} \left(\left(\Sigma_{c \in D_1(y)} p(y|c) \right) - p\left(y|\mathcal{D}(y)\right) \right) \\
(3) &= \frac{2}{n|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \Sigma_{c \in D_1(y)} p(y|c) - \frac{2}{n|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} p(y|\mathcal{D}(y)) \\
(4) &= \frac{2}{n|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \Sigma_{c \in D_1(y)} p(y|c) - \frac{2}{n|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \frac{\text{Emb}(y, \mathcal{D}(y))}{2(n+1)} \\
(5) &= \frac{2}{n|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \Sigma_{c \in D_1(y)} p(y|c) - \frac{1}{(n+1)n|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \text{Emb}(y, \mathcal{D}(y)) \\
(6) &\overset{d}{\geq} \frac{2}{n|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \Sigma_{c \in D_1(y)} p(y|c) - \frac{1}{(n+1)n|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \max_{c \in \mathcal{C}} \{\text{Emb}(c, \mathcal{D}(y))\} \\
(7) &\overset{e}{\geq} \frac{2}{n|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \Sigma_{c \in D_1(y)} p(y|c) - \frac{1}{(n+1)n|\mathcal{C}|} \Sigma_{y \in (\Sigma_2^{n+1})} \text{Emb}(y; \mathcal{D}_{EN}(y)) \\
(8) &= P_{err}(\text{1-Ins}, \mathcal{C}, \mathcal{D}_{EN}, d_L)
\end{aligned}
$$

(4) results from the fact that $p(y|c) = \underbrace{\tfrac{1}{2}}_{\text{insertion of 0 or 1}} \cdot \underbrace{\tfrac{1}{n+1}}_{\text{possible locations for insertion}} \cdot \underbrace{\text{Emb}(y, c)}_{?}$■

Definition 6: We define $\tau(c)$ being the length of c's maximal run.

Remark 7. Let $y \in \Sigma_2^{n+1}$ with n_R runs of length $r_1, r_2, ..., r_{n_R}$ and $c \in D_1(y)$ when we delete 1 bit of the ith run. $p(c \text{ transmitted } |y \text{ received}) = \frac{r_i}{n+1}$.

Definition 8. Coward-Safe We define the following decoder: $\mathcal{D}_{CS} : \Sigma_2^{n+1} \to \Sigma_2^{n+1} \cup \Sigma_2^n$

$$\mathcal{D}(y) = \begin{cases} \mathcal{D}_{EN}(y) & \tau(y) > \frac{n+1}{2} \\ \mathcal{D}_{Lazy}(y) = y & else \end{cases}$$

Lemma 9. Assume $c \in \mathcal{C} = \Sigma_2^n$ was transmitted, $y \in \Sigma_2^{n+1}$ was received and the result of the decoder is $x' \in \Sigma_2^{n+1} \cup \Sigma_2^n$. Then $Pr(c = x') > \frac{1}{2} \Leftrightarrow x'$ results from shortening by 1 a run whose length is $> \frac{n+1}{2}$.

Proof: $\Leftarrow$: Results directly from remark 7.

$\Rightarrow$ Assume x' does not result from shortening by 1 a run whose length is $> \frac{n+1}{2}$. If shortened then direclty remark A. Else $Pr(c = x') = 0 < \frac{1}{2}$.■

Lemma 10. Assume $\mathcal{C} = \Sigma_2^n$.

$\Sigma_{c \in D_1(y)} \frac{d_L(\mathcal{D}_{CS}(y), c)}{|c|} p(c \text{ is transmitted}|y \text{ received}) \leq \frac{1}{n}$

Proof:

$\Sigma_{c \in D_1(y)} \frac{d_L(\mathcal{D}_{CS}(y), c)}{|c|} p(c|y) \underset{\text{Remark 7}}{\leq} \frac{1}{n} \left(\mathbb{I}\left\{\tau(y) > \frac{n+1}{2}\right\} \cdot 2 \cdot \frac{1}{2} + \mathbb{I}\left\{\tau(y) \leq \frac{n+1}{2}\right\} \cdot 1 \right) = \frac{1}{n}$■

Lemma 11. Let $\mathcal{D} : \Sigma_2^{n+1} \to \Sigma_2^*$ be a general decoder. Then, it holds that

$P_{err}(\text{1-Ins}, \Sigma_2^n, \mathcal{D}, d_L) \geq P_{err}(\text{1-Ins}, \Sigma_2^n, \mathcal{D}_{CS}, d_L)$

Proof: We split Σ_2^{n+1} into 3 disjoint sets: $A = \left\{y \in \Sigma_2^{n+1} : |\mathcal{D}(y)| \neq n\right\}$, $B_{\leq \frac{1}{2}} = \left\{y : |\mathcal{D}(y)| = n \wedge Pr(\mathcal{D}(y)|y$ and $B_{> \frac{1}{2}} = \left\{y : |\mathcal{D}(y)| = n \wedge Pr(\mathcal{D}(y)|y) > \frac{1}{2}\right\}$.

- Let $y \in A$.
  In this case $d_L(c, \mathcal{D}(y)) \geq 1$.
  $\Sigma_{c \in D_1(y)} \frac{d_L(\mathcal{D}(y), c)}{|c|} p(c|y) \geq \frac{1}{n} \cdot \Sigma_{c \in D_1(y)} p(c|y) = \frac{1}{n} \underset{\text{Lemma 10}}{\geq} \Sigma_{c \in D_1(y)} \frac{d_L(\mathcal{D}_{CS}(y), c)}{|c|} p(c|y)$

- Let $y \in B_{\leq \frac{1}{2}}$:

$$\begin{aligned} \Sigma_{c \in D_1(y)} \frac{d_L(\mathcal{D}(y), c)}{|c|} p(c|y) &= \Sigma_{c \in D_1(y) \wedge c \neq \mathcal{D}(y)} \frac{d_L(\mathcal{D}(y), c)}{|c|} p(c|y) \\ &\geq \frac{2}{n} \cdot \Sigma_{c \in D_1(y) \wedge c \neq \mathcal{D}(y)} p(c|y) \\ &= \frac{2}{n} \cdot \Sigma_{c \in D_1(y)} Pr(c \neq \mathcal{D}(y)|y) \\ &\geq \frac{2}{n} \cdot \Sigma_{c \in D_1(y)} (1 - \frac{1}{2}) \\ &\geq \frac{2}{n} \cdot \frac{1}{2} = \frac{1}{n} \\ &\underset{\text{Lemma 10}}{\geq} \Sigma_{c \in D_1(y)} \frac{d_L(\mathcal{D}_{CS}(y), c)}{|c|} p(c|y) \end{aligned}$$

- Let $y \in B_{> \frac{1}{2}}$:
  In this case $\mathcal{D}(y) = \mathcal{D}_{CS}(y)$ according to lemma 9.
  $\Sigma_{c \in D_1(y)} \frac{d_L(\mathcal{D}(y), c)}{|c|} p(y|c) = \Sigma_{c \in D_1(y)} \frac{d_L(\mathcal{D}_{CS}(y), c)}{|c|} p(y|c)$

$$
\begin{aligned}
P_{err}(\text{1-Ins}, \Sigma_2^n, \mathcal{D}, d_L) &= \frac{1}{|\mathcal{C}|}\Sigma_{c\in\Sigma_2^n}\Sigma_{y:\mathcal{D}(y)\neq c}\frac{d_L(\mathcal{D}(y),c)}{|c|}p(y|c)\\
&= \frac{1}{|\mathcal{C}|}\Sigma_{y\in(\Sigma_2^{n+1})}\Sigma_{c\in D_1(y)}\frac{d_L(\mathcal{D}(y),c)}{|c|}p(y|c)\\
&= \frac{1}{|\mathcal{C}|}\Sigma_{y\in A}\Sigma_{c\in D_1(y)}\frac{d_L(\mathcal{D}(y),c)}{|c|}p(c|y)\cdot\frac{p(y)}{p(c)}\\
&+ \frac{1}{|\mathcal{C}|}\Sigma_{y\in B_{\leq\frac{1}{2}}}\Sigma_{c\in D_1(y)}\frac{d_L(\mathcal{D}(y),c)}{|c|}p(c|y)\cdot\frac{p(y)}{p(c)} + \frac{1}{|\mathcal{C}|}\Sigma_{y\in B_{>\frac{1}{2}}}\Sigma_{c\in D_1(y)}\frac{d_L(\mathcal{D}(y),c)}{|c|}p(y|c)\\
&\geq \frac{1}{|\mathcal{C}|}\Sigma_{y\in A}\Sigma_{c\in D_1(y)}\frac{d_L(\mathcal{D}_{CS}(y),c)}{|c|}p(c|y)\cdot\frac{p(y)}{p(c)}\\
&+ \frac{1}{|\mathcal{C}|}\Sigma_{y\in B_{\leq\frac{1}{2}}}\Sigma_{c\in D_1(y)}\frac{d_L(\mathcal{D}_{CS}(y),c)}{|c|}p(c|y)\cdot\frac{p(y)}{p(c)} + \frac{1}{|\mathcal{C}|}\Sigma_{y\in B_{>\frac{1}{2}}}\Sigma_{c\in D_1(y)}\frac{d_L(\mathcal{D}_{CS}(y),c)}{|c|}p(y|c)\\
&= \frac{1}{|\mathcal{C}|}\Sigma_{y\in A}\Sigma_{c\in D_1(y)}\frac{d_L(\mathcal{D}_{CS}(y),c)}{|c|}p(y|c)\\
&+ \frac{1}{|\mathcal{C}|}\Sigma_{y\in B_{\leq\frac{1}{2}}}\Sigma_{c\in D_1(y)}\frac{d_L(\mathcal{D}_{CS}(y),c)}{|c|}p(y|c) + \frac{1}{|\mathcal{C}|}\Sigma_{y\in B_{>\frac{1}{2}}}\Sigma_{c\in D_1(y)}\frac{d_L(\mathcal{D}_{CS}(y),c)}{|c|}p(y|c)\\
&= P_{err}(\text{1-Ins}, \Sigma_2^n, \mathcal{D}_{CS}, d_L)\blacksquare
\end{aligned}
$$