8. The k-Insertion Channel

In the k-insertion channel $S_{i,k}$, k arbitrary bits are inserted to the transmitted word uniformly at random. Assume the transmitted word $\boldsymbol{x}$ is arbitrary in $\{0,1\}^n$. A decoder $\mathcal{D}$ for this channel receives the noisy word of length $n+k$ and returns a decoded word $\widehat{\boldsymbol{x}}$. Its average decoding error probably is given by $P_{err}(S_{i,k}, \mathcal{C}, \mathcal{D}) = \frac{\sum_{c\in\mathcal{C}} P_{err}(c,d)}{|\mathcal{C}|}$, where

$$P_{err}(c) = \sum_{y:\mathcal{D}(y)\neq c} \frac{d_L(\mathcal{D}(y), c)}{|c|} \cdot \Pr_{S_{d,k}}\{y\ rec.|c\ trans.\}$$

The goal is to find a decoder that minimizes the average decoding error probability for $k=1$.

בפרויקט זה נעבוד על שאלה 8 עבור המקרה הלא בינארי מעל א״ב בגודל q. נוכיח כי המפענח האופטימלי $\mathcal{D}_{opt}$ הוא מפענח ״היברדי״ שמחזיר עבור מילים בעלות $run\text{-}length$ שגדול ממחצית אורך המילה, מילה בקוד ועבור שאר המילים את המילה שנשלחה, ללא תלות ב-q. יתרה מזאת, נראה נוסחה סגורה לחישוב P_{err} עבור מפענח אופטימלי.

סימונים:

Σ_q – קבוצת התווים בא״ב

Σ_q^n – קבוצת כל המילים באורך n מעל א״ב Σ_q

$\Sigma_{q,r}^n$ – קבוצת כל המילים ב-Σ_q^n כך שאורך הרצף הארוך ביותר של תו בודד (run) במילה הוא r

$r_{c,y}$ – אורך הרצף ב-y ממנו הוסר ביט בכדי לקבל את $c\in D_1(y)$

$r_{\max}^y$ – אורך רצף ארוך ביותר של תו בודד ב-y

$c_{\max}^y$ – מילה המתקבלת ע״י הסרת ביט מרצף ארוך ביותר ב-y

$\mathcal{D}_{max}(y)$ - מפענח המקיים $\mathcal{D}_{max}(y) = c_{\max}^y$

נשים לב כי אם $r_{\max}^y \le \frac{|y|}{2}$ אזי $c_{\max}^y$ לא בהכרח יחיד שכן יתכן יותר מרצף אחד באורך מקסימלי.

טענות עזר:

<u>טענה #1</u>:

נגדיר $r_{c,y}$ אורך הרצף ב-y שממנו הוסר ביט בכדי לקבל את $c\in D_1(y)$, אם לא ניתן לקבל את c, קרי $c\notin D_1(y)$, אז נגדיר $r_{c,y}=0$.

עבור $y\in\Sigma_q^{n+1}$ ועבור $c\in D_1(y)$ מתקיים כי $\Pr_{S_{d,k=1}}\{y\ rec.|c\ trans.\} = \frac{r_{c,y}}{q\cdot|y|}$.

הוכחה:

ההסתברות לקבל את המילה y ע״י הוספת ביט אקראי ל-$c\in D_1(y)$ היא $\frac{m}{q|y|}$ כאשר:

המונה m הוא מספר האינדקסים שבהם אם נוסיף ביט ספציפי מתוך $[q]$ ל-c נקבל את y, שזה שקול לאורך הרצף בו חסר ביט מתוך $[q]$ ביחס ל-y.

המכנה הוא מספר האפשרויות השונות להוספת ביט מתוך $[q]$ ל-c.

<u>טענה #1.1 :</u>

נגדיר $c_{\max}^{y}$ מילה המתקבלת ע״י הסרת ביט מרצף ארוך ביותר ב-y.

עבור $y \in \Sigma_q^{n+1}$ כלשהי, מתקיים $\underset{c}{\operatorname{argmax}} \Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\} = c_{\max}^{y}$

הוכחה :

לפי טענה 1, $\Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\} = \frac{r_{c,y}}{q\cdot|y|}$, מכיוון שהמכנה קבוע לכל c, אזי תוצאה מקסימלית תתקבל עבור $r_{c,y}$ מקסימלי.

מציאת מפענח $\mathcal{D}$ אשר מביא למינימום את $P_{err}(S_{i,k=1},\mathcal{C},\mathcal{D})$

תחילה ננתח את $P_{err}(S_{i,k},\mathcal{C},\mathcal{D})$:

$$P_{err}(S_{i,k},\mathcal{C},\mathcal{D}) = \frac{\Sigma_{c\in\mathcal{C}} P_{err}(c,d)}{|\mathcal{C}|} = \frac{1}{|\mathcal{C}|}\cdot\sum_{c\in\mathcal{C}}\sum_{y:\mathcal{D}(y)\neq c}\frac{d_L(\mathcal{D}(y),c)}{|c|}\cdot\Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\} \overset{(1)}{=}$$

$$\frac{1}{|\mathcal{C}|}\cdot\sum_{c\in\mathcal{C}}\sum_{y\in I_1(c)}\frac{d_L(\mathcal{D}(y),c)}{|c|}\cdot\Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\} \overset{(2)}{=}$$

$$\frac{1}{|\mathcal{C}|}\sum_{y\in\Sigma_q^{n+1}}\sum_{c\in D_1(y)}\frac{d_L(\mathcal{D}(y),c)}{|c|}\cdot\Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\} =$$

$$\frac{1}{n\cdot|\mathcal{C}|}\sum_{y\in\Sigma_q^{n+1}}\sum_{c\in D_1(y)} d_L(\mathcal{D}(y),c)\cdot\Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\}$$

נימוקים :

(1) נראה שעבור $y \notin \{y:\mathcal{D}(y)\neq c\}\cap I_1(c)$ מתקיים כי $\frac{d_L(\mathcal{D}(y),c)}{|c|}\cdot\Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\} = 0$:
- מכיוון שכל $y \in I_1(c)$ שמקיים $\mathcal{D}(y) = c$ מקיים $d_L(\mathcal{D}(y),c) = 0$.
- מכיוון שכל $y \notin I_1(c)$ שמקיים $\mathcal{D}(y) \neq c$ מקיים $\Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\} = 0$.

לכן הסכום לא משתנה.

(2) החלפת סדר הסכימה. אנחנו יודעים ש- $c \in D_1(y)$ אם ורק אם $y \in I_1(c)$.

מטרתנו היא למצוא $\mathcal{D}: \Sigma_q^{n+1} \to \Sigma_q^*$ שמביא למינימום את $P_{err}(S_{i,k=1},\mathcal{C},\mathcal{D})$.

נשים לב כי עבור מפענח $\mathcal{D}_0$ כלשהו, ההגדרה של $\mathcal{D}_0(y_1)$ אינה תלויה בהגדרה של $\mathcal{D}_0(y_2)$ עבור $y_1 \neq y_2$, לכן מאי תלות ומלינאריות החיבור, בכדי להביא את $P_{err}(S_{i,k},\mathcal{C},\mathcal{D})$ למינימום, נרצה לכל $y \in \Sigma_q^{n+1}$ להביא למינימום כל מחובר מהצורה $\sum_{c\in D_1(y)} d_L(\mathcal{D}(y),c)\cdot\Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\}$, כלומר :

$$\min_{\mathcal{D}} P_{err}(S_{i,k},\mathcal{C},\mathcal{D}) = \frac{1}{n\cdot|\mathcal{C}|}\sum_{y\in\Sigma_q^{n+1}}\min_{\mathcal{D}(y)}\sum_{c\in D_1(y)} d_L(\mathcal{D}(y),c)\cdot\Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\}$$

נראה כי $\mathcal{D}_{opt}(y) \in D_1(y) \cup y$

נבחן את $\sum_{c \in D_1(y)} d_L(\mathcal{D}(y), c) \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\}$.

עבור y כלשהו, קלט של המפענח, נחלק למקרים ונבחן את התנהגות המפענח. נשים לב שעבור c, y כלשהן, הגדרת $\mathcal{D}(y)$ כלל לא משפיעה על $\Pr_{S_{d,k}}\{y\ rec. | c\ trans.\}$ שנשאר קבוע:

א. <u>$\mathcal{D}(y) \in \Sigma_q^{n+1}$</u>:

$c \in \Sigma_q^n$, לכן $d_L(\mathcal{D}(y), c) \geq 1$ לכל c, שכן על מנת להגיע למחרוזות זהות נצטרך להוסיף לפחות תו אחד ל-c.
מכיוון שאנחנו מתעניינים ב-$c \in D_1(y)$, אזי $d_L(y, c) = 1$. לכן $\mathcal{D}(y) = y$ הוא אופטימלי מבין כל המילים באורך $n+1$ כיוון שהמרחק שלו מכל $c \in D_1(y)$ הוא המינימלי האפשרי מבין כל המילים באורך $n+1$.

$$\sum_{c \in D_1(y)} \overbrace{d_L(\mathcal{D}(y), c)}^{\geq 1} \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\} \geq \sum_{c \in D_1(y)} \overbrace{d_L(y, c)}^{1} \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\}$$

ב. <u>$\mathcal{D}(y) \in \Sigma_q^n$</u>:

יהיו $c_2 \neq c_1$ כך ש-$|c_1| = |c_2|$, אזי $d_L(c_1, c_2) \geq 2$, כיוון שאם 2 המחרוזות שונות הן שונות בלפחות תו אחד שנדרש להחליף אותו, והחלפה היא החסרה והוספה של תו.
מכיוון שאנחנו מתעניינים ב-$c \in D_1(y)$, אזי אם נגדיר $\mathcal{D}(y) = c_0 \in D_1(y)$ מתקיים שאם $c \in D_1(y)$ אז $d_L(c_0, c) \leq 2$ מאי שיוויון המשולש (מרחק לווינשטיין הוא מטריקה) מכיוון ש:

$$d_L(\mathcal{D}(y), c) \leq d_L(\mathcal{D}(y), y) + d_L(y, c) = 1 + 1 = 2$$

נשים לב ש-$d_L(\mathcal{D}(y), y), d_L(y, c) = 1$ מכיוון ש-$\mathcal{D}(y), c \in D_1(y)$.
משמע, מאיחוד 2 טענות אלו, נקבל כי עבור $\mathcal{D}(y) = c_0$, אם $c_0 \in D_1(y)$ אז כל המרחקים שלנו יהיו 2 פרט למרחק מ-c_0 עצמה שיהיה 0. לעומת זאת אם $c_0 \notin D_1(y)$ אזי כל המרחקים שלנו יהיו לפחות 2.
ולכן:

$$\sum_{c \in D_1(y)} \overbrace{d_L(\mathcal{D}(y) \notin D_1(y), c)}^{\geq 2} \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\}$$
$$\geq \sum_{c \in D_1(y)} \overbrace{d_L(\mathcal{D}(y) \in D_1(y), c)}^{\leq 2} \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\}$$

ג. <u>$\mathcal{D}(y) \in \Sigma_q^{n-1}$</u>:

כל מילה באורך $n-1$ תהיה רחוקה ממילה $c \in \Sigma_q^n$ לפחות ב-1, לכן לא נוכל להשיג תוצאות טובות מאפשרות (א), $\mathcal{D}(y) = y$.

ד. <u>$\mathcal{D}(y) \in \Sigma_q^m, m \geq n+2\ or\ m \leq n-2$</u>:

כל מילה שאורכה רחוק ב-2 מאורך מילה $c \in \Sigma_q^n$, תהיה רחוקה במרחק לווינשטיין של לפחות 2 מאותה מילה, לכן לא נוכל להשיג תוצאות טובות מאפשרות (ב), $\mathcal{D}(y) = c_0 \in D_1(y)$.

כיוון שזו הייתה חלוקה של כלל האפשרויות, ניתן להסיק כי בהינתן y, נרצה להגדיר את $\mathcal{D}(y)$ להיות שווה ל-y עצמו או ל-$c \in D_1(y)$.

נראה כי עבור מקרה ב', $\mathcal{D}(y) = c_0 \in D_1(y)$, מפענח אופטימלי הוא $\mathcal{D}_{max}(y)$, כלומר $c_0 = c_{max}^y$, מילה המתקבלת ע"י הסרת ביט מרצף ארוך ביותר ב-y:

תחת התנאי ש-$\mathcal{D}(y) \in D_1(y)$ נרצה להביא למינימום את $\sum_{c \in D_1(y)} d_L(\mathcal{D}(y), c) \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\}$.

נזכור כי לפי הניתוח של מקרה ב', לכל זוג מילים $c_1, c_2 \in D_1(y)$ מתקיים:

$$d_L(c_1, c_2) = \begin{cases} 0 & c_1 = c_2 \\ 2 & c_1 \neq c_2 \end{cases}$$

$$\sum_{c \in D_1(y)} d_L(\mathcal{D}(y), c) \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\}$$
$$= 2 \cdot \sum_{c \in D_1(y) \setminus \{c_0\}} \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\} + 0 \cdot \Pr_{S_{d,k}}\{y\ rec. | c_0\ trans.\}$$
$$= 2 \cdot \left(\left(\sum_{c \in D_1(y)} \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\} \right) - \Pr_{S_{d,k}}\{y\ rec. | c_0\ trans.\} \right)$$

ניתן לראות כי התוצאה המינימלית תתקבל עבור $c_0 = c_{\max}^y$, כיוון שלפי טענה 1.1 $\Pr_{S_{d,k}}\{y\ rec. | c_0\ trans.\}$ מקסימלי עבור $c_0 = c_{\max}^y$, ובכדי להקטין את הסכום הכולל נרצה להחסיר מחובר גדול ככל האפשר.

נראה כי:

א. עבור $\mathcal{D}(y) = y$ מתקיים כי $\sum_{c \in D_1(y)} d_L(\mathcal{D}(y), c) \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\} = \frac{1}{q}$

ב. עבור $\mathcal{D}_{max}(y)$ מתקיים כי $\sum_{c \in D_1(y)} d_L(\mathcal{D}(y), c) \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\} = 2 \cdot \left(\frac{n+1-r_{max}^y}{q(n+1)}\right)$ כאשר r_{max}^y הוא אורך הרצף הארוך ביותר של תו בודד (run) ב-y

<u>$\mathcal{D}(y) = y$</u>:

נגדיר $r_{c,y}$ אורך הרצף ב-y שממנו הוסר ביט בכדי לקבל את $c \in D_1(y)$

$$\sum_{c \in D_1(y)} d_L(\mathcal{D}(y), c) \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\} \overset{(1)}{=}$$
$$\sum_{c \in D_1(y)} 1 \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\} \overset{(2)}{=} \sum_{c \in D_1(y)} \frac{r_{c,y}}{q(n+1)} \overset{(3)}{=} \frac{|y|}{q(n+1)} = \frac{n+1}{q(n+1)} = \frac{1}{q}$$

(1) לכל $c \in D_1(y)$ מתקיים כי $d_L(c, y) = 1$.
(2) לפי טענה 1.
(3) מתקיים כי $\sum_{c \in D_1(y)} r_{c,y} = |y|$, כיוון שייתקבלו מילים שונות ע"י הסרת ביט מרצפים שונים (הסרת ביטים באינדקסים שונים מאותו הרצף תתן את אותה המילה), וסך אורכי הרצפים הוא $|y|$.

$\underline{\mathcal{D}_{\max}(y)}$:

$$\sum_{c\in D_1(y)} d_L(\mathcal{D}_{\max}(y),c)\cdot \Pr_{\mathrm{S_{d,k}}}\{y\,rec.\,|c\,trans.\} \overset{(1)}{=}$$
$$2\cdot \sum_{c\in D_1(y)\setminus\{c_{\max}^y\}} \Pr_{\mathrm{S_{d,k}}}\{y\,rec.\,|c\,trans.\}+0\cdot \Pr_{\mathrm{S_{d,k}}}\{y\,rec.\,|c_{\max}^y\,trans.\}=$$
$$2\cdot\left(\left(\sum_{c\in D_1(y)} \Pr_{\mathrm{S_{d,k}}}\{y\,rec.\,|c\,trans.\}\right)-\Pr_{S_{d,k}}\{y\,rec.\,|c_{\max}^y\,trans.\}\right)\overset{(2)}{=}$$
$$2\cdot\left(\frac{1}{q}-\frac{r_{\max}^y}{q(n+1)}\right)=2\cdot\left(\frac{n+1-r_{\max}^y}{q(n+1)}\right)$$

(1) לכל זוג מילים $c_1,c_2\in D_1(y)$ מתקיים:

$$d_L(c_1,c_2)=\begin{cases}0 & c_1=c_2\\ 2 & c_1\neq c_2\end{cases}$$

(2) עבור $\mathcal{D}(y)=y$ הראנו כי $\sum_{c\in D_1(y)} \Pr_{\mathrm{S_{d,k}}}\{y\,rec.\,|c\,trans.\}=\frac{1}{q}$ ולפי טענה 1:

$$\Pr_{S_{d,k}}\{y\,rec.\,|c_{\max}^y\,trans.\}=\frac{r_{\max}^y}{q(n+1)}$$

עד כה הראינו כי עבור $y\in\Sigma_q^{n+1}$ כלשהו קיימות 2 אפשרויות "מתחרות" כאופציות אופטימליות להביא למינימום את:

$$\sum_{c\in D_1(y)} d_L(\mathcal{D}(y),c)\cdot \Pr_{\mathrm{S_{d,k}}}\{y\,rec.\,|c\,trans.\}$$

וחישבנו גודל זה עבור כל אחת מן האפשרויות. כעת נותר להשוות בין 2 האפשרויות ולהחליט איזו עדיפה לכל y.

נראה כי עבור $r\geq\frac{n+1}{2}$ נעדיף להשתמש ב-$\mathcal{D}_{max}$ על פני $\mathcal{D}(y)=y$

$$\sum_{c\in D_1(y)} d_L(\mathcal{D}(y)=y,c)\cdot \Pr_{\mathrm{S_{d,k}}}\{y\,rec.\,|c\,trans.\}\geq \sum_{c\in D_1(y)} d_L(\mathcal{D}_{\max}(y),c)\cdot \Pr_{\mathrm{S_{d,k}}}\{y\,rec.\,|c\,trans.\}$$
$$\frac{1}{q}\geq 2\cdot\left(\frac{n+1-r_{\max}^y}{q(n+1)}\right)$$
$$\frac{1}{2}\geq\left(\frac{n+1-r_{\max}^y}{(n+1)}\right)=1-\frac{r_{\max}^y}{n+1}$$
$$\frac{r_{\max}^y}{n+1}\geq\frac{1}{2}$$
$$r_{\max}^y\geq\frac{n+1}{2}$$

לכן, $\mathcal{D}_{opt}$ המוגדר באופן הבא הוא המפענח האופטימלי:

$$\mathcal{D}_{opt}(y)=\begin{cases}c_{\max}^y & r_{\max}^y\geq\frac{n+1}{2}\\ y & r_{\max}^y<\frac{n+1}{2}\end{cases}$$

ראינו כי מפענח המחזיר $\mathcal{D}(y)=c_{\max}^y\ or\ y$ מחזיר אחת מ-2 אפשרויות אופטימליות, וראינו כיצד לבחור באופן מיטבי מבין 2 האפשרויות.

נוסחה סגורה למציאת $P_{err}(S_{i,k=1}, \mathcal{C}, \mathcal{D}_{opt})$

לשם נוחות עתידית, תחילה נבצע מספר חישובי עזר ונוסיף הגדרות:

נגדיר $\Sigma_{q,r}^{n+1}$ כל המילים באורך $n+1$ מעל א״ב בגודל q בעלות רצף מקסימלי באורך r.

נשים לב כי עבור $r \geq \frac{n+1}{2}$ חישוב $\left|\Sigma_{q,r}^{n+1}\right|$ אפשרי באופן הבא:

- עבור $r = n+1$ מתקיים $\left|\Sigma_{q,r}^{n+1}\right| = q$ כיוון שכל הביטים נדרשים עבור הרצף, לכן הם כולם זהים ונותר לקבוע את ערכם.
- עבור $r = n$ מתקיים $\left|\Sigma_{q,r}^{n+1}\right| = 2 \cdot q \cdot (q-1)$ כיוון שיש לנו 2 אפשרויות להחליט איזה ביט אחד לא ישתתף ברצף, הראשון / האחרון, ולאחר מכן יש לנו q אפשרויות לבחירת ערך הרצף, ועוד $q-1$ אפשרויות לבחירת הביט הנותר.
- עבור $n-1 \geq r \geq \frac{n+1}{2}$ מתקיים:

$$\left|\Sigma_{q,r}^{n+1}\right| = (n-r) \cdot q \cdot (q-1)^2 \cdot q^{n-r-1} + 2 \cdot q \cdot (q-1) \cdot q^{n-r}$$

מכיוון ש:

- ישנן $\left((n+1) - r - 2 + 1\right) = (n-r)$ אפשרויות לתחילת הרצף כך שלא יגיע בקצוות. ישנן q אפשרויות לבחירת ערך הביט של הרצף. ישנן $(q-1)^2$ אפשרויות ל״ריפוד״ הרצף, משמע הביטים הצמודים לרצף מימין ומשמאל שמשאירים אותו באורך זה. ישנן q^{n-r-1} אפשרויות למילוי שאר התווים.
- ישנן 2 אפשרויות לרצפים שנוגעים בקצוות. ישנן q אפשרויות לבחירת ערך הביט של הרצף. נדרש ״לרפד״ את הרצף מצד אחד בלבד, לכך יש $(q-1)$ אפשרויות. ישנן q^{n-r} אפשרויות למילוי שאר הביטים.
- נשים לב שכיוון שאורך הרצף שבחרנו גדול ממחצית אורך המילה, כל רצף שיווצר במילוי שאר הביטים בהכרח קטן ממחצית אורך המילה, ועל כן r הוא אורך הרצף המקסימלי.

נשים לב כי עבור $r < \frac{n+1}{2}$, חישוב $\Sigma_{q,r}^{n+1}$ הוא קשה יותר, מכיוון שיתכן ובמילה מסוימת ישנו יותר מרצף ארוך ביותר באורך r אחד ונצטרך להתמודד עם ספירות כפולות. עם זאת, חישוב הכמות הכוללת של מילים באורך $n+1$ מעל א״ב בגודל q בעלות אורך מקסימלי לכל היותר $\frac{n+1}{2}$ (ללא חלוקה פנימית כמה בכל r) הוא פשוט יחסית שכן הוא המשלים של כלל המרחב (q^{n+1}) לחישוב שעשינו קודם לכן. פורמלית:

$$\sum_{r=1}^{\frac{n+1}{2}-1} \left|\Sigma_{q,r}^{n+1}\right| = q^{n+1} - \sum_{r=\frac{n+1}{2}}^{n+1} \left|\Sigma_{q,r}^{n+1}\right|$$

לשמחתנו זהו הגודל שנצטרך בהמשך.

נגדיר $Pen(y,r)$ אשר מייצג את ״קנס״ הטעות בהינתן המפענח $\mathcal{D}_{opt}$, מילה y ואורך הרצף המקסימלי בה r, פורמלית:

$$Pen(y,r) = \begin{cases} 2 \cdot \left(\dfrac{n+1-r}{q(n+1)}\right) & r \geq \dfrac{|y|}{2} \\ \dfrac{1}{q} & r < \dfrac{|y|}{2} \end{cases}$$

משמע $Pen(y,r)$ מייצג את $\sum_{c \in D_1(y)} d_L\left(\mathcal{D}_{opt}(y), c\right) \cdot \Pr_{S_{d,k}}\{y\ rec. | c\ trans.\}$ כפונקציה של y ושל r (שהוא עצמו פונקציה של y)

נפתח את $P_{err}\left(S_{i,k},\mathcal{C},\mathcal{D}_{opt}\right)$:

$$P_{err}\left(S_{i,k},\mathcal{C},\mathcal{D}_{opt}\right)=\frac{1}{n\cdot|\mathcal{C}|}\sum_{y\in\Sigma_q^{n+1}}\sum_{c\in D_1(y)}d_L(\mathcal{D}(y),c)\cdot\Pr_{S_{d,k}}\{y\,rec.\,|c\,trans.\}\overset{(1)}{=}$$

$$\frac{1}{n\cdot|\mathcal{C}|}\sum_{y\in\Sigma_q^{n+1}}Pen\left(y,r_{\max}^{y}\right)=\frac{1}{n\cdot|\mathcal{C}|}\sum_{r=1}^{n+1}\sum_{y\in\Sigma_{q,r}^{n+1}}Pen(y,r)\overset{(2)}{=}$$

$$\frac{1}{n\cdot|\mathcal{C}|}\left(\sum_{r=1}^{\frac{n+1}{2}-1}\sum_{y\in\Sigma_{q,r}^{n+1}}Pen(y,r)+\sum_{r=\frac{n+1}{2}}^{n+1}\sum_{y\in\Sigma_{q,r}^{n+1}}Pen(y,r)\right)$$

$$=\frac{1}{n\cdot|\mathcal{C}|}\left(\sum_{r=1}^{\frac{n+1}{2}-1}\sum_{y\in\Sigma_{q,r}^{n+1}}\frac{1}{q}+\sum_{r=\frac{n+1}{2}}^{n+1}\sum_{y\in\Sigma_{q,r}^{n+1}}2\cdot\left(\frac{n+1-r}{q(n+1)}\right)\right)\overset{(3)}{=}$$

$$\frac{1}{n\cdot|\mathcal{C}|}\left(\frac{1}{q}\left(q^{n+1}-\sum_{r=\frac{n+1}{2}}^{n+1}\left|\Sigma_{q,r}^{n+1}\right|\right)+\sum_{r=\frac{n+1}{2}}^{n+1}2\cdot\left(\frac{n+1-r}{q(n+1)}\right)\cdot\left|\Sigma_{q,r}^{n+1}\right|\right)$$

נימוקים:

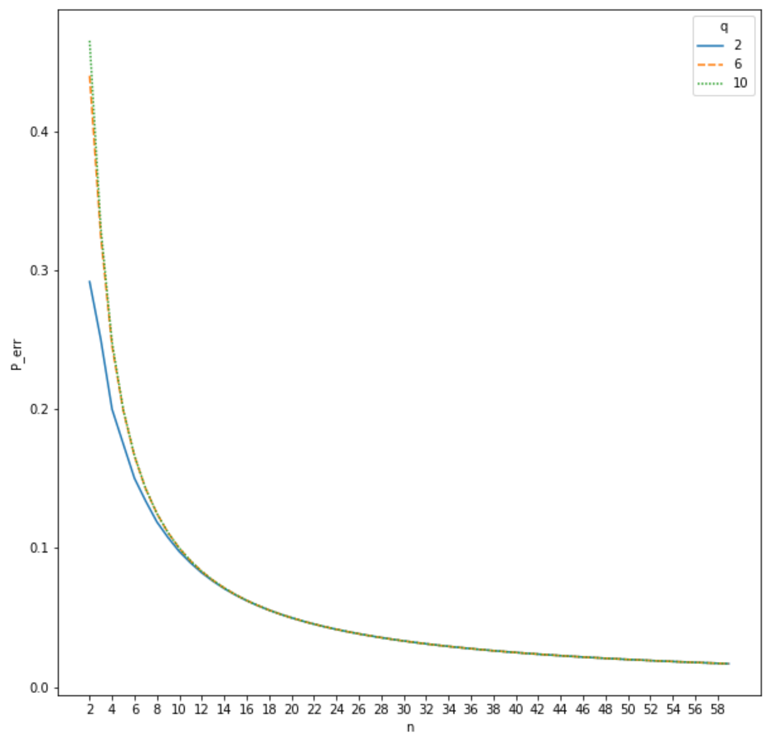
(1) לפי הגדרת $Pen(y,r)$
(2) חלוקת הסכום למקרים לפי המקרים ב-$Pen(y,r)$
(3) לפי חישובי העזר קודם לכן

על מנת לוודא כי לא התבלבלנו בחישובים, ביצענו השוואה בין תוצאות הנוסחה הסגורה לחישוב P_{err} לבין חישוב ישיר של P_{err} ע״י שימוש ב-$\mathcal{D}_{opt}$ וסכימת הטעויות בסימולציה. להלן התוצאות לכל $q\in\{2,3,4\}$ ולכל $n\in\{2\ldots7\}$

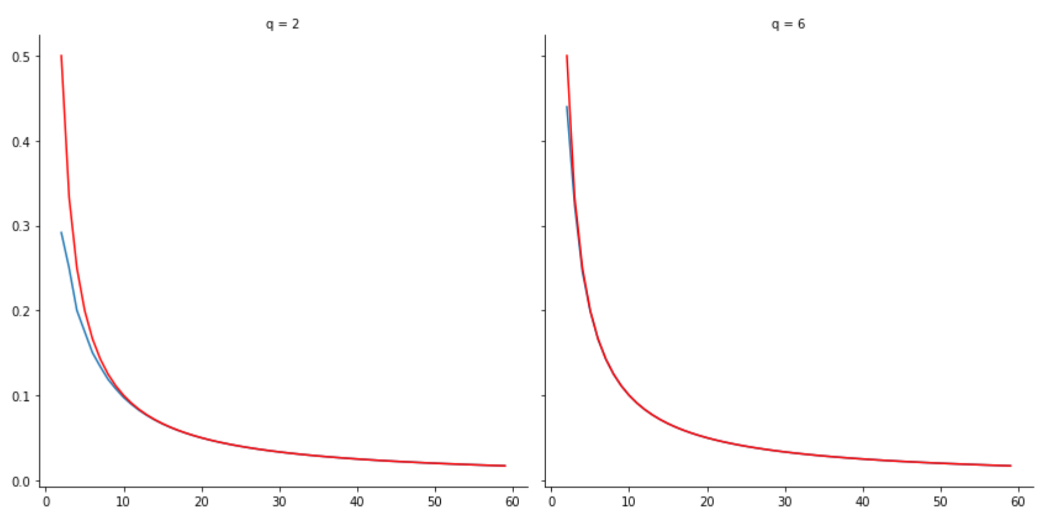
n	q	*Direct*	*Closed form*
2	2	0.291667	0.291667
3	2	0.250000	0.250000
4	2	0.200000	0.200000
5	2	0.175000	0.175000
6	2	0.150298	0.150298
7	2	0.133929	0.133929
2	3	0.370370	0.370370
3	3	0.296296	0.296296
4	3	0.229630	0.229630
5	3	0.192593	0.192593
6	3	0.162257	0.162257
7	3	0.141093	0.141093
2	4	0.406250	0.406250
3	4	0.312500	0.312500
4	4	0.239063	0.239063
5	4	0.196875	0.196875
6	4	0.164900	0.164900
7	4	0.142299	0.142299

ניתן להבחין כי התוצאות המתקבלות זהות (עד כדי שגיאות נומריות).
נוסחה סגורה מאפשרת לנו לבחון את התנהגות P_{err} גם עבור n, q גדולים ולהשוות אותה לתוצאת מפענח "עצלן", $\mathcal{D}_{lazy}(y) = y$, אשר מחזיר תמיד את y, ומגיע ל-P_{err} שהינו $\frac{1}{n}$ (הוכחה בנספח).

ניתן לראות כי ככל ש-q גדול יותר כך P_{err} גדול יותר, אבל בכל מקרה הערך הדומיננטי הוא n, אורך המילה, וכש-n גדל P_{err} קטן מאוד מהר. בהשוואת $\mathcal{D}_{opt}$ ו-$\mathcal{D}_{lazy}$ ניתן לראות שמאוד מהר ההפרש הביצועים נהיה זניח.



השוואה בין q שונים ככל ש-n גדל



השוואה בין $\mathcal{D}_{lazy}$ (באדום) לבין $\mathcal{D}_{opt}$ בכחול עבור $q = 6$ (מימין) ו-$q = 2$ (משמאל)

נספח

נרצה לחשב את $P_{err}\left(S_{i,k=1},\mathcal{C},\mathcal{D}_{lazy}\right)$, נזכור שכבר קודם לכן הראנו את קיום השוויון הבא:

$$\sum_{c\in D_1(y)} d_L\left(\mathcal{D}_{lazy}(y),c\right)\cdot \Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\}=\frac{1}{q}$$

נציב ונחשב:

$$P_{err}\left(S_{i,k=1},\mathcal{C},\mathcal{D}_{lazy}\right)=\frac{1}{n\cdot|\mathcal{C}|}\sum_{y\in\Sigma_q^{n+1}}\sum_{c\in D_1(y)} d_L\left(\mathcal{D}_{lazy}(y),c\right)\cdot \Pr_{\mathrm{S_{d,k}}}\{y\ rec.|c\ trans.\}=\frac{1}{n\cdot|\mathcal{C}|}\sum_{y\in\Sigma_q^{n+1}}\frac{1}{q}$$
$$=\frac{1}{n\cdot q^n}\left(q^{n+1}\cdot\frac{1}{q}\right)=\frac{1}{n}$$

עזרים

(1) במהלך העבודה על הפרויקט קראנו את המאמר:
Daniella Bar-Lev, Yotam Gershon, Omer Sabary, and Eitan Yaakobi (2021). The Error Probability of Maximum-Likelihood Decoding for the t-Deletion Channel.
https://www.omersabary.com/files/ML-ISIT2021EXT.pdf

(2) קוד הפונקציה לחישוב מרחק לווינשטיין לקוח מהבלוג הבא:
Implementing The Levenshtein Distance in Python | Paperspace Blog
עם התאמות קלות למרחק לווינשטיין כפי שהוא מוגדר במסמך הפרויקט (מחיקה והוספה של תו בלבד, ללא החלפת תווים)