מגישים: אלמוג יעקב, ליאור עטיה

דחיסת נתונים – מטלה 1

שאלה 1:

תחילה נוכיח כי עבור כל 2 התפלגויות מתקיים:

$$\sum_{i=1}^{I} p_i \cdot \log_2(\frac{1}{p_i}) \le \sum_{i=1}^{I} p_i \cdot \log_2(\frac{1}{q_i})$$

עבור X > 1 מתקיים $\ln x \le x - 1$.

ע"פ נוסחת החלפת בסיס לוגים $\log_b x = \frac{\log_c x}{\log_c b}$

לכן:

$$\ln x \le x - 1 \;\rightarrow\; \frac{\log_2 x}{\log_2 e} \le x - 1 \;\rightarrow\; \log_2 x \le \log_2 e \cdot (x-1) \;\rightarrow$$
$$\log_2 x \le \frac{\log_e e}{\log_e 2} \cdot (x-1) \;\rightarrow\; \log_2 x \le \frac{(x-1)}{\ln 2}$$

מכאן:

$$\sum_{i=1}^{I} p_i \cdot [\log_2\left(\frac{1}{p_i}\right) - \log_2(\frac{1}{q_i})] = \sum_{i=1}^{I} p_i \cdot \log_2\left(\frac{q_i}{p_i}\right)$$
$$\le \frac{1}{\ln 2} \cdot \sum_{i=1}^{I} p_i \cdot \left(\frac{q_i}{p_i} - 1\right) = \frac{1}{\ln 2} \cdot \left(\sum_{i=1}^{I} q_i - \sum_{i=1}^{I} p_i\right) = 0$$

נזכיר כי $K(C) = \sum_{i=1}^{I} 2^{-|c_i|}$

ונציב $q_i = \frac{2^{-|c_i|}}{\sum_{i=1}^{I} 2^{-|c_i|}}$ (ניתן לראות כי q_i התפלגות).

כעת, נוכיח $H(C) \le E(C,P)$

$$H(C) = \sum_{i=1}^{I} p_i \cdot \log_2(\frac{1}{p_i}) \le \sum_{i=1}^{I} p_i \cdot \log_2(\frac{1}{q_i})$$
$$= \sum_{i=1}^{I} p_i \cdot \log_2(2^{|c_i|} \cdot \sum_{i=1}^{I} 2^{-|c_i|}) = \sum_{i=1}^{I} p_i \cdot (|c_i| + \log_2(\sum_{i=1}^{I} 2^{-|c_i|}))$$
$$\le \sum_{i=1}^{I} p_i \cdot |c_i| = E(C,P)$$

כאשר האי שוויון האחרון נובע מכך שהקוד C הוא UD ע"פ הנתון, ומכאן $\sum_{i=1}^{I} 2^{-|c_i|} \le 1$ ולכן $\log_2(\sum_{i=1}^{I} 2^{-|c_i|}) \le 0$.

שאלה 2:

נוכיח כי C הוא קוד חסר רישות אמ"מ הוא קוד מיידי.

<u>כיוון ראשון</u>: C הוא קוד חסר רישות ← C הוא קוד מיידי

מכיוון ש-C הוא קוד חסר רישות בסיום פענוח של מילת קוד X נדע בהכרח שסיימנו לקרוא מילה זו. שאם אינו כן, ובהנחה שבהמשך הקריאה נוכל לקרוא מילת קוד אחרת Y, אזי נקבל כי X הינה רישא של Y בסתירה להנחה.

לכן, בסיום קריאה של מילת קוד אנו יודעים כי הסימן הבא שייך לקריאת המילה הבאה.

כלומר, C הוא קוד מיידי.

<u>כיוון שני</u>: C הוא קוד חסר רישות → C הוא קוד מיידי

ע"פ חוקי דה מורגן $B \rightarrow A$ שקול אל $!B$ ← $!A$!

לכן, נוכיח באופן שקול C הוא קוד שאינו חסר רישות ← C אינו קוד מיידי

מאחר ש-C הינו קוד שאינו חסר רישות, קיימת מילת קוד A שהינה רישא של מילה אחרת B.

כלומר, לאחר שקראנו קוד כלשהו ובפענות מילת הקוד A לא נוכל לדעת אם סיימנו את הקריאה של A או שאנו צריכים להמשיך לקרוא עבור מילה B. לכן, נסיק כי C אינו קוד מיידי.

מגישים: אלמוג יעקב, ליאור עטיה

<u>שאלה 3:</u>

נתון קוד C עם מילות קוד באורכים $l_1, l_2, \dots, l_n$
נוכיח כי אם C הוא קוד שלם אזי $\sum_{i=1}^{n} 2^{-l_i} = 1$

נוכיח באינדוקציה חזקה.
בסיס: n = 2. ומכיוון שאנו עוסקים בקודים בינאריים, ניקח קוד בינארי מינימלי ולכן: $l_1 = l_2 = 1$.

$$\sum_{i=1}^{2} 2^{-l_i} = 2^{-1} + 2^{-1} = 1$$

נניח שהטענה נכונה עד n ונוכיח עבור n+1.
כלומר, ע"פ ההנחה מתקיים:

$$\sum_{i=1}^{n} 2^{-l_i} = 1$$

כעת, נוכיח:

$$\sum_{i=1}^{n+1} 2^{-l_i} = 1$$

נוכיח ע"י בניה של קוד כלשהו בעל n+1.

נניח קוד כלשהו עם n מילות קוד.
נניח כי c_k הינה מילת קוד עם האורך המקסימלי. כלומר, l_k הינו האורך שלה (בהתאמה) ולכן הוא האורך המקסימלי.
כעת, מכיוון שאנו עוסקים בקודים בינאריים אזי נוכל להוסיף סיבית בסוף מילת הקוד c_k על מנת לקבל מילת קוד חדשה בצורה בה תכונת קוד שלם תשמר עבור קוד C.
מכאן, נקבל $2 \cdot (l_k + 1)$ אורכים במקום האורך היחיד l_k.
לכן עבור קוד C הנ"ל:

$$\sum_{i=1}^{n+1} 2^{-l_i} = 2^{-l_1} + \dots + \mathbf{2^{-(l_k+1)}} + \mathbf{2^{-(l_k+1)}} + \dots + 2^{-l_n} =$$
$$2^{-l_1} + \dots + \mathbf{2^{-l_k}} + \dots + 2^{-l_n} = \sum_{i=1}^{n} 2^{-l_i} = 1$$

כאשר השוויון השני נובע מהמשוואה: $2^{-(l_k+1)} + 2^{-(l_k+1)} = \frac{1}{2^{(l_k+1)}} + \frac{1}{2^{(l_k+1)}} = \frac{2}{2^{(l_k+1)}} = \frac{1}{2^{l_k}}$
והשוויון הרביעי נובע מההנחה.

* במידה ונרצה לבנות קוד שלם שאינו חסר רישות עבור בחירת c_k הנ"ל (כלומר, נקבל $2 \cdot (l_k + 1)$ אורכים **בנוסף** לאורך היחיד l_k) נקבל את השוויון הבא:

$$\sum_{i=1}^{n+1} 2^{-l_i} = 2^{-l_1} + \dots + \mathbf{2^{-l_k}} + \mathbf{2^{-(l_k+1)}} + \mathbf{2^{-(l_k+1)}} + \dots + 2^{-l_n} =$$
$$2^{-l_1} + \dots + \mathbf{2^{-l_k}} + \mathbf{2^{-l_k}} + \dots + 2^{-l_n} =$$
$$2^{-l_1} + \dots + \mathbf{2^{-(l_k-1)}} + \dots + 2^{-l_n} = \sum_{i=1}^{n-1} 2^{-l_i} = 1$$

מ.ש.ל.