

דריסת נתונים א'

תרגיל מספר 2

ניתן להגיש בזוגות (הגשה אחת), יש להגיש את קובץ הפתרון בפורמט pdf בלבד.

להגשה עד – 2.12.2021 - ניתן להגיש עד שבוע איחור. על כל יום איחור תרדנה שלוש נקודות.

שאלה 1:

מה מספר הצמתים הפנימיים בעץ בינרי מלא בעל n עלים? הוכיחו.

שאלה 2:

נתונה התפלגות $\{p_1, \dots, p_n\}$. הוכיחו ששינוי שגורם להתפלגות להתקרב יותר לאחידות במובן הבא, גורר הגדלת כמות האינפורמציה.

כלומר, נניח $p_1 > p_2$. בחר $0 < \varepsilon < \frac{p_1 - p_2}{2}$, אז עדיין $p_1 - \varepsilon > p_2 + \varepsilon$.

הראו שמתקיים $H(p_1 - \varepsilon, p_2 + \varepsilon, p_3, \dots, p_n) > H(p_1, p_2, p_3, \dots, p_n)$

הדרכה:

ניתן להשתמש במשפט הבא (ללא הוכחתו):

נתונות ההתפלגויות $\{p_1, \dots, p_n\}$ ו- $\{q_1, \dots, q_n\}$.

$$\text{אזי מתקיים ש-} -\sum_{i=1}^n p_i \log p_i \leq -\sum_{i=1}^n p_i \log q_i$$

שאלה 3:

כזכור, קוד נקרא **קוד חסר רישות** או **קוד פרפיקסי**, prefix (free) code, אם אף מילת קוד אינה רישא של מילת קוד אחרת. באופן סימטרי נגדיר **קוד חסר סיפות** או **קוד סופיקסי**, suffix (free) code, בו אף מילת קוד אינה סיפא של מילת קוד אחרת. למעשה, הקוד המתקבל מהפיכת מילות הקוד של קוד חסר רישות, הוא קוד חסר סיפות.

קוד נקרא **קוד אפיקסי**, Affix code, אם הקוד הוא גם קוד פרפיקסי וגם קוד סופיקסי. בפרט, כל קוד בעלי אורך קבוע, הוא קוד אפיקסי.

- א. בנו קוד שלם שהוא גם אפיקסי בעל 9 מילות קוד באורך משתנה. שימו לב שיתכן שזה בלתי אפשרי לא"ב קטן מידי.
- ב. הראו שמספר הקודים השלמים באורך משתנה שהם גם קודים אפיקסים הוא אינסופי.

בהצלחה