שיעור 8
אנטרופיה ומידע

8.1 המושג של מידע

נניח נניח ש- X משתנה מקרי אשר יכול לקבל אחת מארבע אפשריות:

$$X \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{a}\} \ .$$

X ידוע לבוב (B) אבל לא ידוע לאליס (A). כל שאליס יודעת הוא ש- X יכול להיות אחת האותיות $\{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{a}\}$ בהסתברות שווה. אנחנו אומרים כי לאליס יש אי-ודאות על הערך של X. כדי שאליס תמצא את הערך של X אליס שואלת סדרת שאלות בינאריות (שאלות כן/לא) לבוב כדי לקבל מידע על המ"מ X עד שהיא תדע את הערך של X עם אי-ודאות אפס.

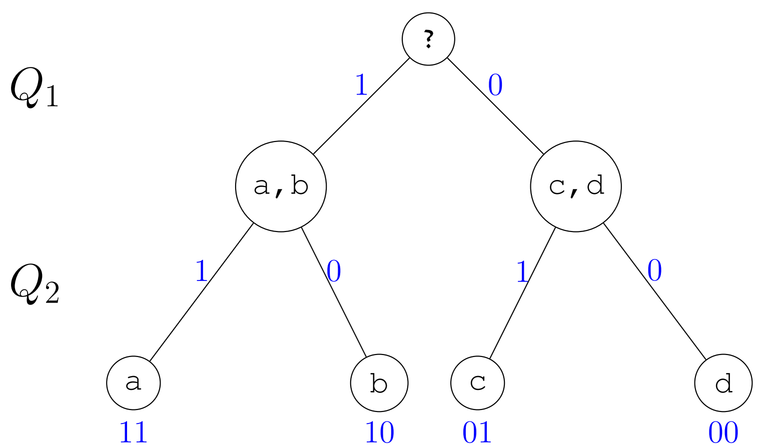
אפשרות אחת לסדרת שאלות היא כך:

Q_1: האם $X \in \{\mathtt{a}, \mathtt{b}\}$?

לפי התשובה אחר כך אליס שואלת

Q_2: אם $X \in \{\mathtt{a}, \mathtt{b}\}$ האם $X = \mathtt{a}$?

אחרת אם $X \notin \{\mathtt{a}, \mathtt{b}\}$ האם $X = \mathtt{c}$?



הסדרה של שאלות בינאריות שמאפשרת לאליס למצוא את X ללא שום אי-ודאות מתוארת בעץ-שאלות למעלה. מספר השאלות הבינאריות $N_Q[X]$, שנדרשות כדי למצוא X ללא אי-ודאות הוא $N_Q[X] = 2$.

כל שאלה היא בינארית, כלומר התשובה היא כן או לא אנחנו מצפינים תשובה כן עם "1" ותשובה לא עם "0". לפי התשובות אנחנו מצפינים את האותיות כך:

$$a \to 11 \ , \quad b \to 10 \ , \quad c \to 01 \ , \quad d \to 00 \ .$$

מכיוון ששתי תשובות בינאריות נדרשות כדי למצוא את X, אנחנו אורמים כי נדרש שני ביטים (bits) של מידע נדרשים כדי למצוא את X.

במילים אחרות, שתי ספרות ביניאריות $X = d_1 d_2$ נדרשות כדי להצפין את X, שערכן הן התשובות לשתי שאלות ביניאריות,

לכן המידע המתקבל על מציאת הערך של X הוא 2 bit.

אליס הייתה יכולה לשנות את הסדרת שאלות שלה כך:

Q'_1 האם $X = \mathtt{a}$?

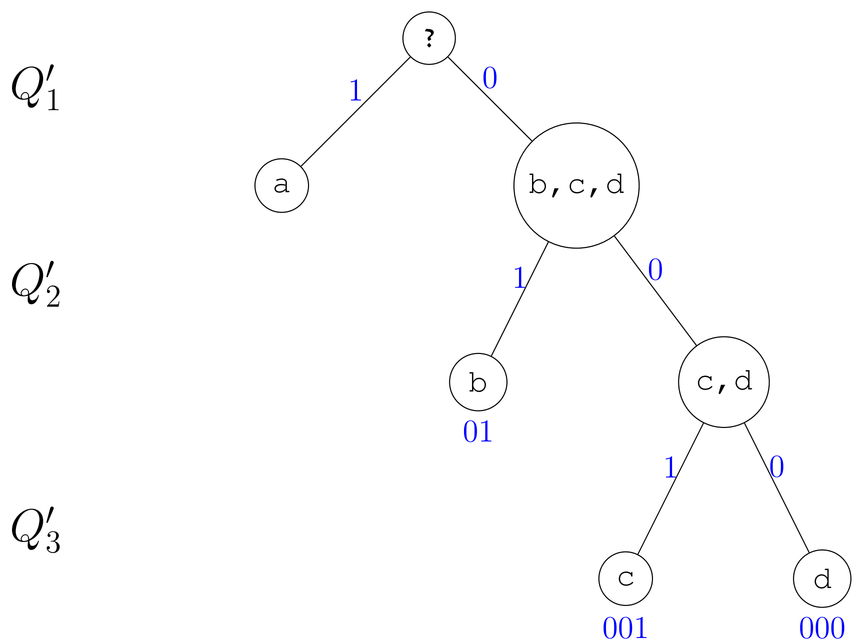
רק אם התשובה היא "לא" אז היא צריכה לשאול שאלה נוספת:

Q'_2 האם $X = \mathtt{b}$?

ורק אם השתובה היא "לא" אז היא צריכה לשאול שאלה נוספת:

Q'_3 האם $X = \mathtt{c}$?

מספר השאלות הבינאריות הנדרשות למצוא את X תלוי על הערך של X: $N_Q(\mathtt{a}) = 1$, $N_Q(\mathtt{b}) = 2$ או $N_Q(\mathtt{c}) = N_Q(\mathtt{d}) = 3$.



X הוא משתנה מקרי בדיד ולכן בהינתן מערכת שאלות, $N_Q(X)$ הוא פונקציה של משתנה מקרי בדיד, ולכן $N_Q[X]$ הוא בעצמו משתנה מקרי בדיד.

כעת נשאל שאלה. נניח כי אליס מעוניינת למצוא מערכת שאלות Q, אשר נותנת את מספר השאלות הממוצע המינימלי. כלומר, כיצד נמצא מערכת שאלות $N_Q[X]$ עבורה התוחלת

$$E\left[N_Q[X]\right] = \sum_{k\in X} P_X(k) N_Q[k]$$

תהיה מינימלית.

לפני שנענה על שאלה הזאת נתן דוגמה.

נתון המשתנה מקרי $X = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}$ בעל הפונקצית הסתברות

$$P_X(\mathtt{a}) = \frac{1}{2}\ , \quad P_X(\mathtt{b}) = \frac{1}{4}\ , \quad P_X(\mathtt{c}) = P_X(\mathtt{d}) = \frac{1}{8}\ .$$

עם ההצפנה הראשונה $N_Q[k] = 2$ לכל $k \in X$ אז התוחלת תהיה $\frac{1}{2}(2) + \frac{1}{4}(2) + \frac{1}{8}(2) + \frac{1}{8}(2) = 2$ כלומר תוחלת מספר השאלות הוא 2.
התוחלת עבור ההצפנה השנייה היא

$$E\left[N_{Q'}[X]\right] = \frac{1}{2}(1) + \frac{1}{4}(2) + \frac{1}{8}(3) + \frac{1}{8}(3) = \frac{7}{4}\ .$$

אשר פחות מהתוחלת עבור ההצפנה הקודמת.

אליס שואלת סדרת שאלות ולכל שאלה נשים ערך ביניארי 0 אם התשובה לא ו- 1 אם התשובה כן. כך אנחנו נשים לכל ערך של X מספר בינארי $d_1 \ldots d_k$ המורכב מספרות ביניאריות $d_i = 0, 1$. טרנספורמציה כזרת בין ערכים של X לבין מספרים בינארים נקראת הצפנה. שימו לב כי אורך ההצפנה $\ell_Q[X]$ של כל ערך של X שווה למספר השאלות ביניאריות הנדרשותת כדי למצוא את X ללא אי-ודאות:

$$\ell_Q[X] = N_Q[X] \ .$$

התוחלת המינימלית מתקבלת באמצעות מערכת שאלות שבה מספר השאלות שמובילות לערך כלשהו ביחס הפוך להסתברות שלו. במילים פשוטות, ככל שההסתברויות של ערך של X גבוהה מספר השאלות המובילות לערך זה יותר קטן, ולהפך.

אקסיום 1) מספר ביטים האופטימלי $\ell_{Q^*}(k)$ הנדרש להצפין את הסימן $X = k$ הוא פונקציה של ההסתברות $P_X(k)$.

אקסיום 2)

$$P_X(k) \geq P_X(k') \quad \Rightarrow \quad \ell_{Q^*}(k) \leq \ell_{Q^*}(k') \ .$$

משפט 8.1 אנטרופיה של שאנון

$$H[X] = -\sum_{k \in X} P_X(k) \log_2 P_X(k) \ .$$

הוכחה: נניח כי $X = Y \cap Z$, כאשר Y, Z משתנים מקרים בלתי תלויים. אז

$$H[X] = H[Y] + H[Z]$$

נסמן $p_x = P_X(x)$. לפי אקסיום 1):

$$\ell_Q(x) = f\left(p_x\right) \ .$$

$$H[X] = \sum p_x \ell_Q(x) = \sum p_x f\left(p_x\right) \ .$$

כעת נניח שיש לנו משתנים מקרים Y ו- Z ושהם בלתי תלויים. יהיו $P_Y(y)$ ו- $P_Z(z)$ פונציות ההסתברות של Y ושל Z בהתאמה. נסמן $p_y = P_Y(y)$ ו- $p_z = P_Z(z)$.

נגדיר את המאורע $X = Y \cap Z$. מכיוון ש- Y ו- Z משתנים בלתי תלויים אז

$$P\left(X = Y \cap Z\right) = P_Y(y) P_Z(z) = p_y p_z \ .$$

ידיעה של Y לא נותנת שום מידע על הערך של Z, אז

$$\ell_Q[Y \cap Z] = \ell_Q[Y] + \ell_Q[Z] \ .$$

לפיכך

$$H[X] = \sum p_x \ell_Q(x) = \sum p_y p_z \left(\ell_Q(y) + \ell_Q(y)\right)$$

$$H[X] = \sum p_x \ell_Q(x) = \sum p_y p_z \left(\ell_Q(y) + \ell_Q(y)\right)$$

מכאן

$$H[X] = \sum p_y p_z f\left(p_y p_z\right) = \sum p_y p_z \left[f\left(p_y\right) + f\left(p_z\right)\right]$$

לכל p_y ו- p_z. לכן

$$f\left(p_y p_z\right) = f\left(p_y\right) + f\left(p_z\right) \ .$$

ז"א $f(p) = C\log(p)$. נדרש כי $f\left(\frac{1}{2}\right) = 1$ ונקבל $f(p) = -\log_2(p)$.

■

8.2 הגדרה של מידע

הגדרה 8.1 מידע של מאורע (שאנון)

נתון משתנה מקרי X. המידע של ערך מסוים של X מסומן $I_X(x)$ ומוגדר ליהות

$$I\left(X=x\right) = \log_2\left(\frac{1}{P_X(x)}\right) = -\log_2\left(P_X(x)\right)$$

כאשר $P_X(x)$ פונקצית ההסתברות של המשתנה מקרי X.

דוגמה 8.1 המידע המתקבל בגילוי תוצאה של הטלת מטבע

נטיל מטבע הוגנת ונגדיר משתנה מקרי X להיות התוצאה של הניסוי. מכאן X מקבל את הערכים

$$X = \{H, T\} \ .$$

מצאו את המידע של המאורע $X = H$.

פתרון:

$P(X=H) = \frac{1}{2}$. לכן

$$I(X=H) = -\log_2\left(\frac{1}{2}\right) = 1 \ .$$

כלומר על קבלת התוצאה "H" אנחנו מקבלים ביט אחד של מידע.

הסבר:

במקום הסימנים "H" ו- "T" בשביל המ"מ X ניתן להצפין את הערכים האפשריים בספרות בינאריות "0" או "1". כלומר

ערך של X	הצפנה בספרות בינאריות
H	0
T	1

ז"א כדי להצפין את הערכים של X אנחנו צריכים ספרה בינארית אחת:

$$d_1 \in \{0, 1\} \ .$$

אשר יכול להחזיק את הערכים 0 או 1.

ספרה ביניארית אחת נדרשת להחזיק את הערך של X לכן המידע של ערך כלשהו של X הוא 1 bit (ביט אחד). ■

דוגמה 8.2 שליפת קלף מחבילת קלפים תיקנית

בניסוי שליפת קלף אחד מחבילת קלפים תיקנית. נגדיר את המשתנה מקרי X להיות הסוג של הקלף (תלתן, עלה לב או יהלום). חשבו את את המידע של המאורע ששלפני קלף מסוג לב.

פתרון:
ההסתברות לשלוף קלף של הסוג לב מחבילת קלפים סטנדרטית היא

$$P(X=\heartsuit)=\frac{13}{52}=\frac{1}{4}\ .$$

לכן

$$I(X=\heartsuit)=-\log_2\left(\frac{1}{4}\right)=2\ \text{bits}$$

הסבר:

יש 4 הערכים האפשריים של X:

$$X=\{\spadesuit,\clubsuit,\heartsuit,\diamondsuit\}$$

כל ספרה בינארית מחזיקה 2 ערכים אפשריים: 0 או 1 לכן ידרש שתי ספרות בינאריות כדי להצפין את ה-4 ערכים האפשריים של X:

$$d_1d_2\ ,\qquad d_1,d_2\in\{0,1\}\ .$$

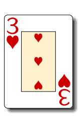
ההצפנה עצמה מתוארת בטבלא למטה:

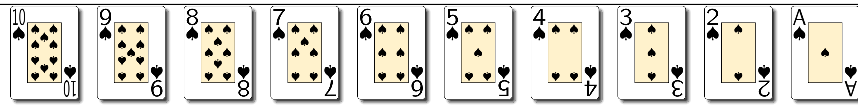
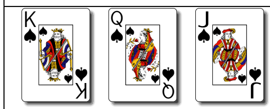
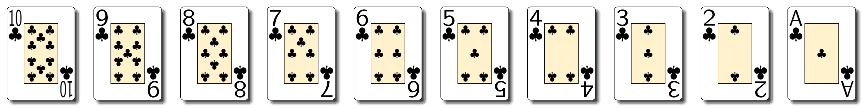
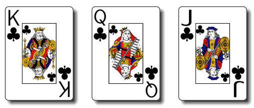
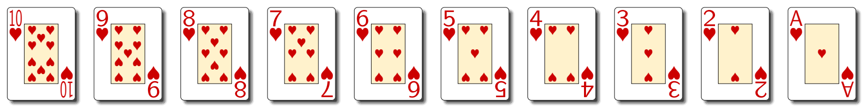
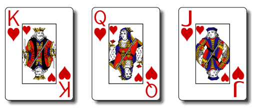
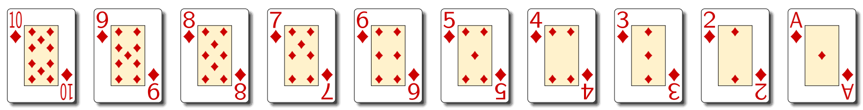
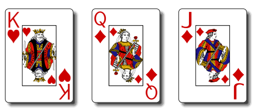
ערך של X	הצפנה בספרות בינאריות
$\spadesuit$	00
$\clubsuit$	01
$\heartsuit$	10
$\diamondsuit$	11

אורך המספר d_1d_2 הוא 2 לכן המידע של המשתנה מקרי X הוא $2\,\text{bits}$ (שני ביטים).

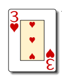
■

דוגמה 8.3 שליפת קלף מחבילת קלפים תיקנית

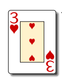
בניסוי שליפת קלף אחד מחבילת קלפים תיקנית. מצאו את המידע המתקבל אם הקלף  נשלף.

צורה	מספרים	תמונות
עלה		
תלתן		
לב		
יהלום		

פתרון:
יהי X המ"מ שמסמן את הקלף הנשלף. ההסתברות לשלוף הקלף שלוש מצורת לב מחבילת קלפים סטנדרטית היא

$P\Big(X =$  $\Big) = \frac{1}{52}$.

לכן

$I\Big(X =$  $\Big) = -\log_2\left(\frac{1}{52}\right) = 5.7 \text{ bits}$

הסבר:

כדי להצפין את כל הערכים האפשריים של X כרצף סיבית, נדרש רצף סיביתחם אשר מקבל לפחות 52 ערכים שונים. רצף עם 5 סיביות לא מספיק כי יש לו רק $2^5 = 32$ ערכים שונים. אבל רצף עם 6 סיביות נותן $2^6 = 64$ ערכים שונים, אשר מספיק להצפין את כל הערכים האפשריים של X.

$$d_1d_2d_3d_4d_5d_6$$

האורך של הרצף סיביות הזה הוא 6 ולכן הרצף סיבית זה נותן 6 bits של מידע.
לכל סיבית יש 2 ערכים אפשריים ולכן 64 ערכים שונים בסה"כ.

רק 52 מתוך ה- 64 צירופים נדרשים כדי להצפין את הערכים האפשריים של X לכן נוריד חלק של הסיביות. הקבוצת סיביות הנשארים מכילה 5.7 bits של מידע.

■

ככל שההסתברות של מאורע יותר קטנה אז המידע המתקבל יותר גבוהה.

כלומר, ככל שהמידע של מאורע יותר גבוהה אז ההסתברות שלו יותר קטנה

8.3 אנטרופיה

הגדרה 8.2 אנטרופיה של מ"מ X

נתון מ"מ בדיד X. נניח כי הערכים האפשריים של X הם

$$X = \{x_1, \ldots, x_N\}\ .$$

האנטרופיה $H(X)$ של מ"מ X מוגדרת להיות התוחלת (הממוצע המשוקלל) של המידע המתקבל על ידי למצוא את הערך של X (כלומר על גילוי התוצאה של הניסוי):

$$H(X) = \sum_{i=1}^{N} P(X = x_i) I(X = x_i) = -\sum_{i=1}^{N} P(X = x_i) \log_2 \left(P(X = x_i)\right)$$

במקרה שההסתברות של כל תוצאה שווה, כלומר

$$P(X = x_i) = \frac{1}{|X|} = \frac{1}{N}$$

אז

$$H(X) = -\sum_{i=1}^{N} \frac{1}{N} \log_2 \left(\frac{1}{N}\right) = \frac{1}{N} \sum_{i=1}^{N} \log_2 N = \log_2 N\ .$$

לכן

$$N = 2^{H(X)}\ .$$

ניתן להוכיח ש- $\log_2 N$ הוא הערך המקסימלי האפשרי של $H(X)$.

משפט 8.2

נתון מ"מ בדיד X אשר מקבל N ערכים שונים:

$$X = \{x_1, \ldots, x_N\}$$

אם ההסתברות של כל ערך שווה, כלומר

$$P(X = x_i) = \frac{1}{N}$$

אז האנטרופיה מקבלת ערך מקסימלי שניתנת על ידי

$$H_{\text{max}}(X) = \log_2 N\ .$$

ערך זה הוא הערך המקסימלי האפשרי של האנטרופיה.

דוגמה 8.4 אנטרופיה בהטלת מטבע

נניח כי נטיל מטבע עם הסתברות p $(0 \leq p \leq 1)$. לקבל H. מצאו את האנטרופיה של המ"מ מקרי X אשר שווה לתוצאת הניסוי.

פתרון:

נסמן $X=\{0,1\}$ כאשר $X=0$ מסמן תוצאת H ו- $X=1$ מסמן תוצאת T. הפונקצית הסתברות היא

$$P_X(0)=p\ , \qquad P_X(1)=1-p\ .$$

לכן המידע של המאורע לקבל תוצאת H הוא

$$I(X=0)=-\log_2\left(P_X(0)\right)=-\log_2\left(p\right)$$

והמידע של המאורע לקבל תוצאת H הוא

$$I(X=1)=-\log_2\left(P_X(1)\right)=-\log_2\left(1-p\right)$$

נשים לב שאם המטבע הוגנת אז $p=\dfrac{1}{2}$ ו- $I(X=0)=I(X=1)=1$. כעת נחשב את האנטרופיה של X:

$$H(X)=-P_X(0)\log_2\left(P_X(0)\right)-P_X(1)\log_2\left(P_X(1)\right)=-p\log_2 p-(1-p)\log_2(1-p)\ .$$

נרשום את האנטרופיה כפונקציה של ההסתברות p:

$$H(X)=--p\log_2 p-(1-p)\log_2(1-p)\ =:h(p).$$

ל- $h(p)$ יש נקודת מקסימום ב- $p=\dfrac{1}{2}$:

$$h'(p)=-\frac{1}{\ln 2}-\log_2 p+\frac{1}{\ln 2}+\log_2(1-p)=-\log_2 p+\log_2(1-p)=\log_2\left(\frac{1}{p}-1\right)\stackrel{!}{=}0\Rightarrow p=\frac{1}{2}\ .$$

ז"א הערך המקסימלי של האנטרופיה מתקבל כאשר לכל הערכים של X יש הסתברות שווה, $P_X(0)=P_X(1)=\dfrac{1}{2}$.
אכן

$$h(p=\tfrac{1}{2})=-\frac{1}{2}\log_2\frac{1}{2}-\frac{1}{2}\log_2\frac{1}{2}=\log_2 2=1\ .$$

■

דוגמה 8.5

בניסוי הטלת מטבע לא מאוזנת, ההסתברות לקבל תוצאה H היא $p=\dfrac{1}{1024}$. מצאו את האנטרופיה של X.

פתרון:
נסמן $X=\{0,1\}$, כאשר $X=0$ מסמן תוצאת H ו- $X=1$ מסמן תוצאת T.

$$I(X=0)=-\log_2\frac{1}{1024}=10\ \text{bits}\ , \qquad I(X=1)=-\log_2\left(1-p\right)=-\log_2\frac{1023}{1024}=0.00141\ \text{bits}\ .$$

לפי זה

$$H(X)=-p\log_2 p-(1-p)\log_2(1-p)=-\frac{1}{1024}\log_2\frac{1}{1024}-\frac{1023}{1024}\log_2\frac{1023}{1024}=0.0112\ \text{bits}\ .$$

■

המשמעות של התשובה לדוגמה הקודמת היא כך. נניח שנטיל אותה מטבע הלא מאוזנת $100,000$ פעמים. בכדי להצפין את כל התוצאות נדרש רצף סיביות של אורך $100,000$, כאשר כל ספרה נותנת התוצאה של ניסוי אחד. ז"א 10^5 bits של מידע נדרש כדי להצפין את כל התוצאות.

מצד שני מצאני כי התוחלת של המידע המתקבל לניסוי (כמות מידע פר ניסוי) הוא 0.0112 bit פר ניסוי. במילים אחרות, ב- 10^5 ניסוים רק 1120 bit של מידע נדרש בממוצע כדי להצפין את כל התוצאות של הרצף ניסויים.

אנטרופיה (בביטים) אומרת לנו את כמות המידע הממוצעת (בביטים) שיש לספק על מנת להעביר את כל התוצאות של המאורע. זהו חסם תחתון על מספר הסיביות שיש להשתמש בהן, בממוצע לקודד (להצפין) את התווים של ההודעה שלנו.

8.4 הצפנת האפמן

נסביר הצפנת האפמן בעזרת הדוגמה הבאה. נתון הטקסט גלוי

$$X = \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}\}$$

ונניח כי הפונקצית הסתברות של X היא לפי הטבלה הבאה:

בחירת אות של $x_i \in X$	$p_i = P_X(x_i)$	$I(X = x_i) = -\log_2(p_i)$
a	$\frac{1}{3}$	1.58 bit
b	$\frac{1}{2}$	1 bit
c	$\frac{1}{12}$	3.58 bit
d	$\frac{1}{12}$	3.58 bit

נשאל את השאלה: כמה ביטים של מידע נדרשים כדי להצפין (בסיביות) רצף של 1000 אותיות של טקסט גלוי X?

יש 4 אותיות ב- X, כלומר 4 ערכים אפשריים של המ"מ בדיד X. לפיכך נדרש רצף של 2 סיביות כדי להצפין טקסט גלוי של תו אחד בהצפנת סיביות קבועה. לדוגמה:

בחירת אות של $x_i \in X$	הצפנה
a	00
b	01
c	10
d	11

ז"א להצפין תו אחד של הטקסט גלוי X נדרש 2 bit. לכן להצפין רצף אותיות של טקטסט גלוי נדרש $2 \times 1000 =$ 2000 bit, כלומר 2000 סיביות.

האנטרופיה של X היא

$$H(X) = -p_1 \log_2(p_1) - p_2 \log_2(p_2) - p_3 \log_2(p_3) - p_4 \log_2(p_4) = 1.62581 \text{ bit} .$$

ז"א לכל ניסוי המידע הממוצע הנדרש כדי להצפין תו אחד של טקסט גלוי הוא 1.62581 bit. לכן המידע הממוצע הנדרש כדי להצפין רצף אותיות של טקסט גלוי הוא

$$1000 \times 1.62581 = 1625.81 \text{ bit} .$$

לכן, רצף סיביות של אורך 1626 בממוצע יהיה מספיק כדי להעביר את ההודעה.

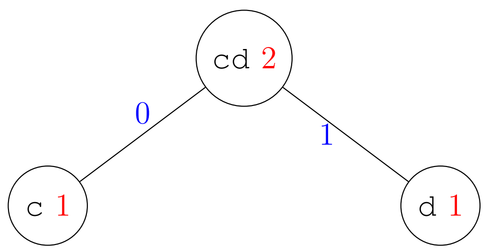
כעת נבנה הצפנה של הטקסט גלוי על ידי האלגוריתם של האפמן.

שלב 1)

	c	d	a	b
	1	1	4	6

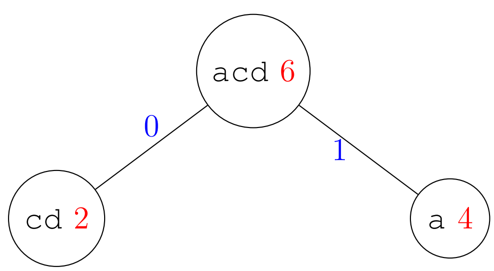
שלב 2)

	c	d	a	b
	1	1	4	6
	0	1		
	2		4	6



שלב 3)

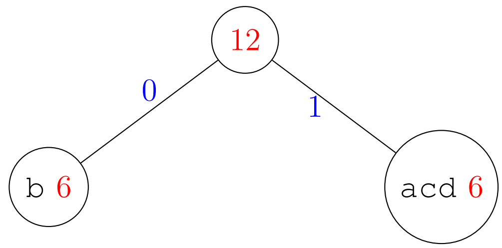
	cd	a	b
	2	4	6
	0	1	
	6		6



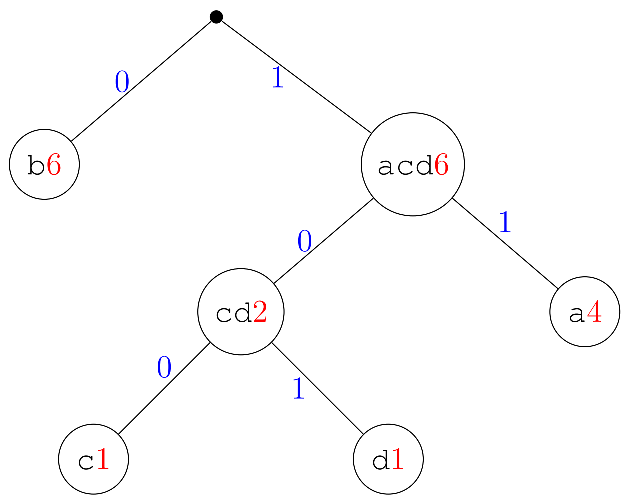
שלב 4)

שלב 5)

	acd	b
	6	6
	0	1
	12	



שלב 6)



בסוף של התהליך האותיות של הטקסט גלוי יהיו בעלים של העץ וההצפנה ניתנת על ידי הרצף סיביות על הענפים במסלול מהנקודת התחלתית של העץ עד העלה בו רשום האות בשאלה.

בחירת אות של $x_i \in X$	הצפנת האפמן
a	11
b	100
c	110
d	101

דוגמה 8.6

נתון הטקסט גלוי הבא

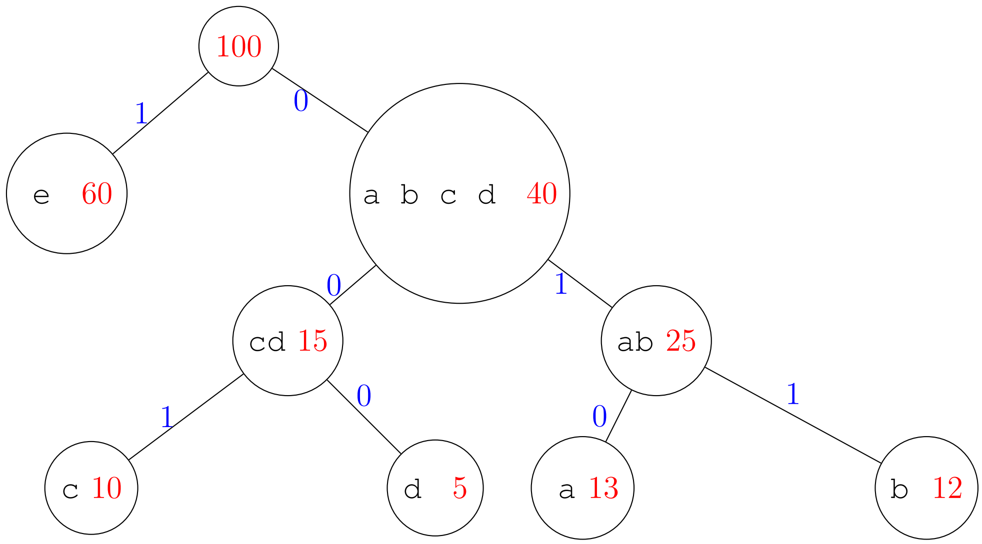
$$X = \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}, \mathrm{e}\}$$

והפונקצית הסתברות

$$P(X=\mathrm{a}) = \frac{13}{100} = 0.13 \ , \quad P(X=\mathrm{b}) = \frac{3}{25} = \frac{12}{100} = 0.12 \ , \quad P(X=\mathrm{c}) = \frac{1}{10} = \frac{10}{100} = 0.1 \ ,$$
$$P(X=\mathrm{d}) = \frac{1}{20} = \frac{5}{100} = 0.05 \ , \quad P(X=\mathrm{e}) = \frac{3}{5} = \frac{60}{100} = 0.6 \ .$$

מצאו את העץ הצפנה וההצפנת האפמן של כל תו של X.

פתרון:



בחירת אות של $x_i \in X$	הצפנת האפמן
a	010
b	011
c	001
d	000
e	1

■

פורמלי הצפנת האפמן מוגדרת לפי ההגדרה הבאה:

הגדרה 8.3 הצפנת האפמן

נתון משתנה מקרי X. נגדיר הצפנת האפמן של X להיות הפונקציה (כלל מצפין)

$$f : X \to \{0,1\}^*$$

כאשר $\{0,1\}^*$ קבוצת רצפים של סיביות סופיים.

נתון רצף מאורעות $x_1, \ldots, x_n$. נגדיר

$$f(x_1 \ldots x_n) = f(x_1)|| \ldots ||f(x_n)$$

כאשר "||" מסמן שרשור (concatenation).

הגדרה 8.4 תוחלת האורך של הצפנת האפמן

נתונה הצפנת האפמן f. תוחלת האורך של ההצפנה מוגדרת

$$l(f) = \sum_{x\in X} P(X = x)\,|f(x)|\ .$$

משפט 8.3 אי שוויון האפמן

נתון קבוצת אותיות של טקסט גלוי X והצפנת האפמן f. נניח כי $l(f)$ תוחלת האורך של ההצפנה ו- $H(X)$ האנטרופיה של הטקסט גלוי. מתקיים

$$H(X) \leq l(f) \leq H(X)+1 \ .$$

דוגמה 8.7 (המשך של דוגמה 8.6)

נתון הטקסט גלוי הבא

$$X = \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}, \mathrm{e}\}$$

והפונקצית הסתברות

$$P(X=\mathrm{a}) = \frac{13}{100} = 0.13 \ , \quad P(X=\mathrm{b}) = \frac{3}{25} = 0.12 \ , \quad P(X=\mathrm{c}) = \frac{1}{10} = 0.1 \ , \quad P(X=\mathrm{d}) = \frac{1}{20} = 0.05 \ ,$$

$$P(X=\mathrm{e}) = \frac{3}{5} = 0.6 \ .$$

1) מצאו את תוחלת האורך של ההצפנת האפמן.

2) מצאו את האנטרופיה.

3) הוכיחו כי אי-שוויון האפמן של ההצפנה שמצאתם בדוגמה 8.6 למעלה מתקיים.

פתרון:

סעיף 1)

$$\begin{aligned} l(f) =& \frac{5}{100}\cdot 3 + \frac{10}{100}\cdot 3 + \frac{12}{100}\cdot 3 + \frac{13}{100}\cdot 3 + \frac{60}{100}\cdot 1 \\ =& \frac{15+30+36+30+60}{100} \\ =& \frac{180}{100} \\ =& 1.8 \end{aligned}$$

סעיף 2)

$$\begin{aligned} H(X) =& -P(X=\mathrm{a})\log_2 P(X=\mathrm{a}) - P(X=\mathrm{b})\log_2 P(X=\mathrm{b}) - P(X=\mathrm{c})\log_2 P(X=\mathrm{c}) \\ & - P(X=\mathrm{d})\log_2 P(X=\mathrm{d}) - P(X=\mathrm{e})\log_2 P(X=\mathrm{e}) \\ =& 1.74018 \ . \end{aligned}$$

סעיף 3) $H(X)=1.74018$, $H(X)+1=1.84018$, $l(f)=1.8$. לכן

$$H(X) \leq l(f) \leq H(X)+1$$

מתקיים.

■