

פרויקט מסכם

עץ הופמן

קוד הופמן הוא אלגוריתם דחיסה ללא אובדן נתונים. האלגוריתם מקצה קוד באורך שונה לתווי הקלט. ככל ששכיחות התו גבוהה יותר אורך הקוד המתאים לו יהיה קצר יותר. באופן זה אורך הטקסט המקודד יהיה קצר יותר מטקסט שבו קידוד התווים יהיה באורך אחיד.

הקלט לבניית קוד הופמן הוא אוסף של תווים ושכיחות הופעתם בטקסט.

השלבים ליצירת קוד הופמן הם:

- בניית עץ הופמן מאוסף התווים ושכיחותם.
- סריקה של עץ הופמן והתאמת קוד לתווים.

השלבים לבניית עץ הופמן הם:

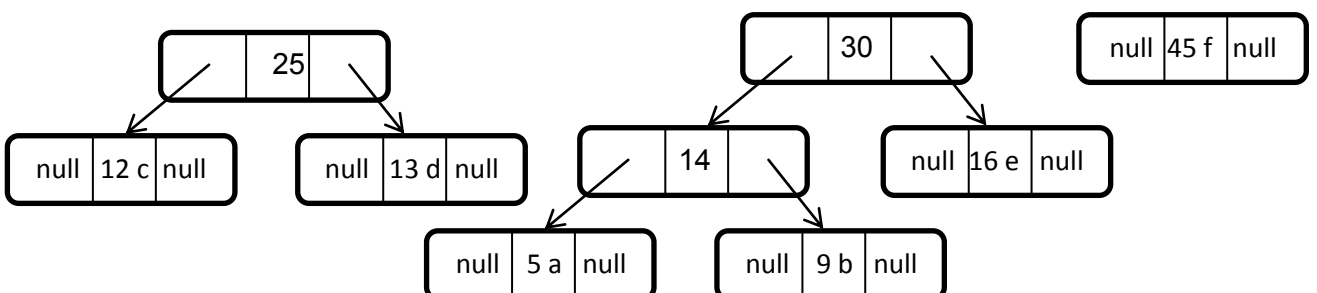
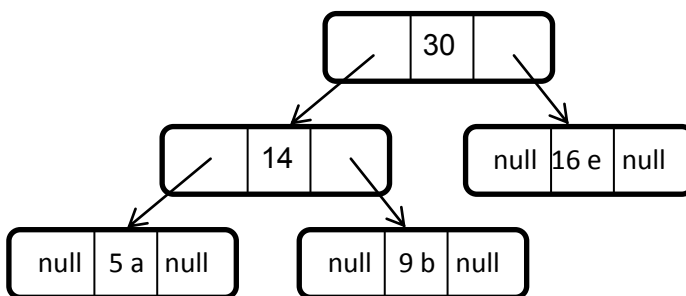
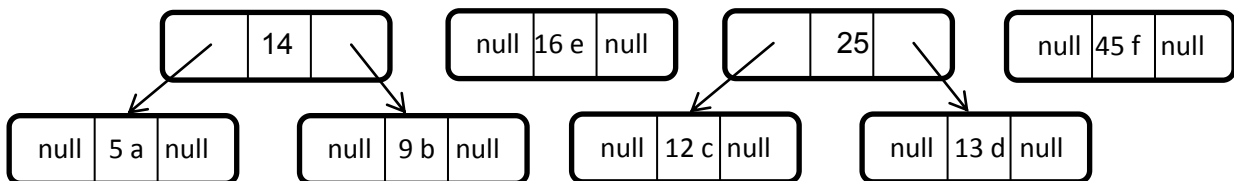
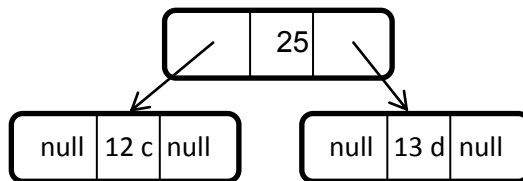
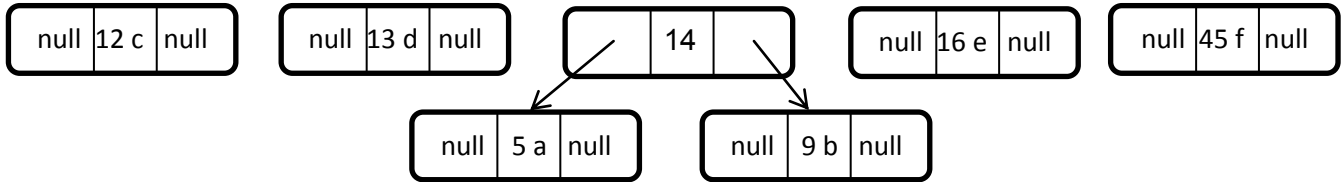
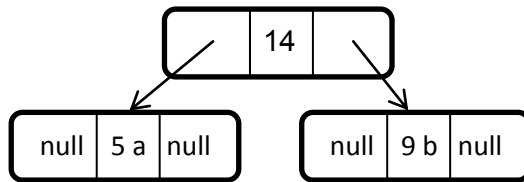
1. המרת התווים ושכיחותם לחוליות בינריות, והכנסתן למבנה ממוין, משכיחות נמוכה לגבוה.
2. הוצאת שתי החוליות בעלות השכיחות המינימליות מהמבנה ובניית חוליה חדשה, ששכיחותה היא סכום השכיחות של שתי החוליות שהוצאו, כך שהחוליה ששכיחותה קטנה יותר תושם כבן שמאלי.
3. הכנסת החוליה החדשה באופן ממוין למבנה.
4. חזרה על שלבים 2 ו-3 עד שהמבנה הממוין יכיל צומת אחד בלבד. צומת זה הוא השורש של עץ הופמן.

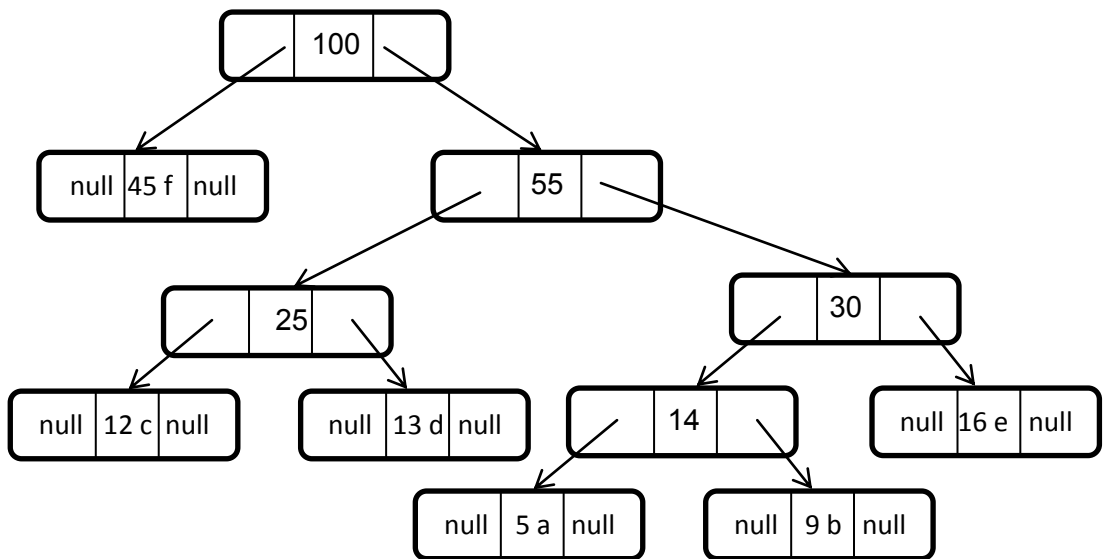
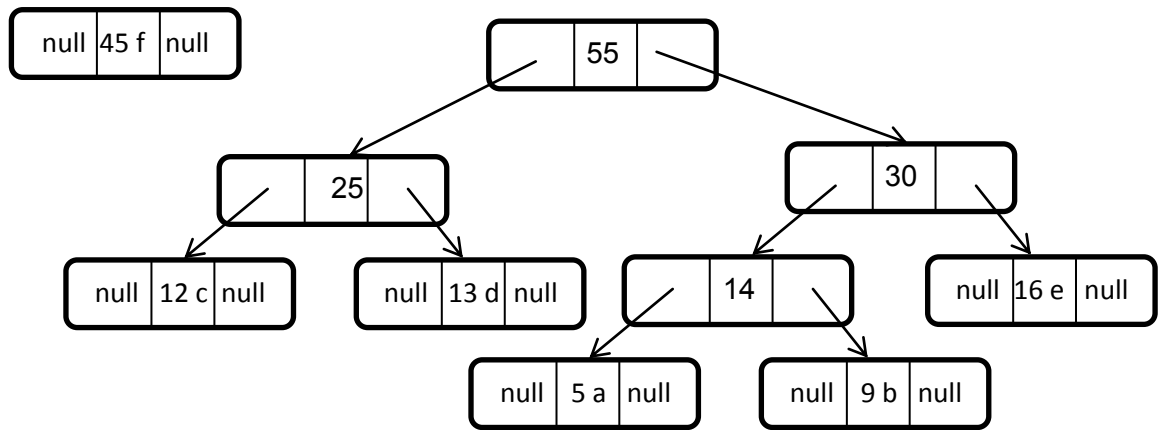
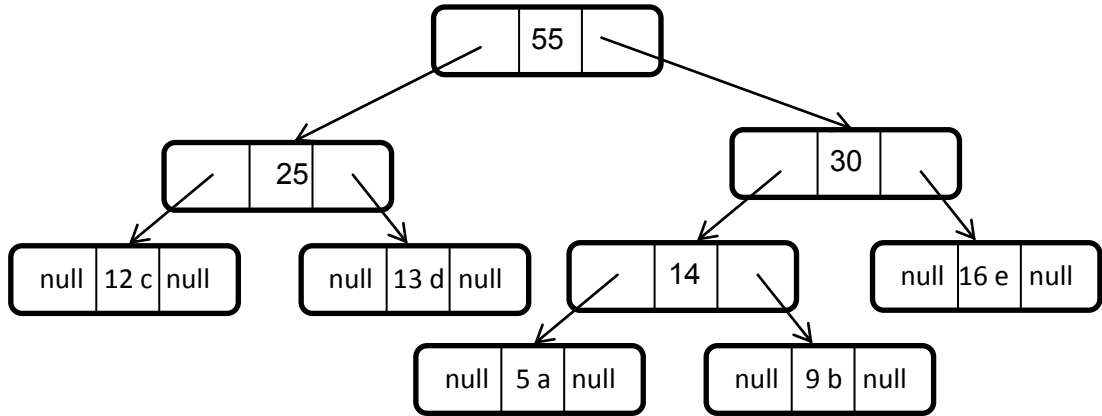
נמחיש את השלבים באמצעות דוגמא:

נתון אוסף תווים ושכיחותם

שכיחות	תו
5	a
9	b
12	c
13	d
16	e
45	f

5 a	9 b	12 c	13 d	16 e	45 f
-----	-----	------	------	------	------





לאחר בניית עץ הופמן יש לסרוק אותו על מנת להתאים לכל תו את הקידוד שלו.
בן שמאלי מיוצג ע"י 0 ובן ימני ע"י 1.
תתקבל הטבלה הבאה:

קידוד	תו
1100	a
1101	b
100	c
101	d
111	e
0	f

גש למורה לקבלת משימה ראשונה.