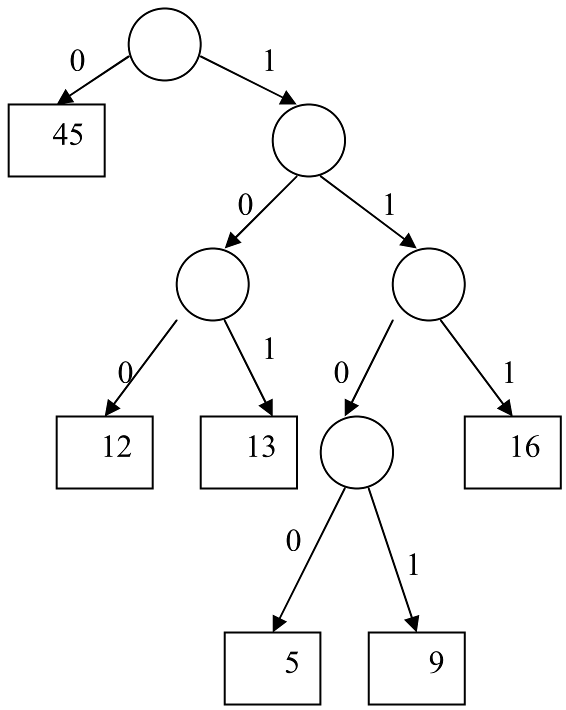
<u>קוד חד פענח (קוד בינארי)</u>
א"ב בן $n$ אותיות. כל אות ממופה למילת קוד $c$.
הקוד הוא חד פענח אם מתקיים התנאי הבא:
$C=\{c_1,c_2,...,c_n\}$ אוסף המילים.
קיים השוויון הבא: $c_{i1}...c_{ik}=c_{j1}...c_{jl}$ אז $k=l$ וגם לכל $m$ מתקיים: $c_{im}=c_{jm}$.

<u>קוד פרפיקסי</u>: קוד שאף מילה איננה רישא של מילה אחרת.
קוד פרפיקסי הוא תמיד חד פענח, אולם קוד חד פענח הוא לא תמיד פרפיקסי.

נניח שלכל מילת קוד $c_i$ יש תדירות $L(c_i)$ והאורך של $c_i$ הוא $l(c_i)$.
קוד חד פענח הוא <u>אופטימאלי</u> אם $\sum_{c_i} f(c_i)\cdot l(c_i)$ הוא מינימאלי.

אלגוריתם <u>הופמן</u> מוצא קוד פרפיקסי אופטימאלי.
קוד פרפיקסי מתאים לעץ (בינארי):
לכל מילת קוד $c_i$ האלגוריתם מתאים עלה בעץ שעומקו הוא $d_T(c_i)$ (T הוא העץ).
העץ של אלגוריתם הופמן מקיים: $\sum_{c_i} f(c_i)\cdot d_T(c_i)$ מינימאלי אפשרי.




שאלה: האם יתכן שקיים קוד חד פענח שאיננו פרפיקסי שמחירו נמוך יותר מהמחיר של הקוד הפרפיקסי האופטימאלי:
תשובה: לא! (נוכיח בהמשך)

יהיו $l_1,...,l_n$ אורכי המילים של קוד חד פענח, אזי קיים קוד פרפיקסי שאלו אורכי מילותיו.
קוד $\sigma$-רי הוא קוד מעל $\sigma$ אותיות.
כל ההגדרות שניתנו עד כה ניתנות להכללה מיידית לקוד $\sigma$-רי.

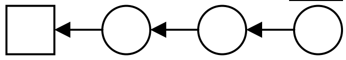
<u>טענה:</u> סדרת אורכים $(l_1,...,l_n)$ היא סדרה של אורכים של מילים של קוד פרפיקסי מעל א"ב בן $\sigma$ אותיות אם: הסכום האופייני $=\sum_{i=1}^{n}\sigma^{-l_i}\le 1$.

<u>הוכחה (כיוון ראשון):</u> נשתמש בעובדה שהקוד הפרפיקסי ניתן לייצוג על ידי עלים בעץ $\sigma$-רי (עץ שבו לכל צומת פנימי יש לכל היותר $\sigma$ בנים).

אם קיים עץ $\sigma$-רי שעומקי עליו הם $(l_1,...,l_n)$ אז ע"י הילוך אקראי המתחיל מהשורש ובכל צומת פנימי בהסתברות $\frac{1}{\sigma}$ הולך לכל אחד מבניו, ההסתברות לעצור בעלה בעומק $l$ היא $\sigma^{-l}$. ההסתברות לעצור בעלה כלשהו $\le 1$.

נוכיח באינדוקציה (על $n=$ מספר האורכים) שאם $\sum_{i=1}^{n}\sigma^{-l_i}\le 1$ אז קיים עץ $\sigma$-רי עם $n$ עלים שעומקיהם $l_1,...,l_n$:

<u>בסיס:</u> עבור $n=1$ טריביאלי.



<u>צעד:</u> נניח שהוכחנו לכל הסדרות בעלות $n$ אורכים, ונוכיח עבור הסדרה $l_1,...,l_{n+1}$

<u>מקרה א':</u>

ריבוי של מספר $\lambda$ בסדרה $(l_1,...,l_{n+1})$ הוא מספר ה-$l_i$-ים שאורכיהם $\lambda$.

נניח שאין מספר שהריבוי שלו $> \sigma-1$.

$(\lambda_i,k_i) \quad \lambda_1<\lambda_2<...<\lambda_n$

בנייה של עת שצמתיו הפנימיים מהווים מסלול ולצומת בעומק $\lambda_i$ מוסיפים $k_i$ בנים.

<u>מקרה ב':</u>

קיים $m$ כך של $\lambda_m$ יש ריבוי $\sigma$.

נבנה מהסדרה הנתונה סדרה חדשה באופן הבאL

נסלק $\sigma$ מספרים שערכם $\lambda_m$. נוסיף מספר אחד שאורכו $\lambda_m-1$.

הסכום האופייני ירד ב: $\sigma\cdot\sigma^{-\lambda_m}=\lambda^{-(\lambda_n-1)}$ ולאחר מכן עלה ב $\sigma^{-(\lambda_m-1)}$.

קיים לפי הנחת האינדוקציה עץ שעומקי עליו מתאימים לסדרה החדשה.

נוכיח כעת שאם $(l_1,...,l_n)$ הם סדרת אורכים של קוד חד פענח אז מתקיים אי השוויון $\sum_{i=1}^{n}\sigma^{-l_i}\le 1$ ולכן קיים קוד פרפיקסי בעל אותם אורכי מילים.

<u>הוכחה (כיוון שני):</u> (carush):

נסתכל על כל השרשורים של $k$ מילים מ $C$.

$l_{\max}=$ ערך מקסימאלי של $\{l_i\}$.

כל המילים יהיו באורך $\le l_{\max}$.

לכל ערך $1\le L\le l_{\max}$ יהיו לכל היותר $\sigma^L$ שרשורים באורך $L$.

$$\left(\sum_{i=1}^{n}\sigma^{-l_i}\right)^k=\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_k=1}^{n}\sigma^{\underbrace{-(l_{i_1}+...+l_{i_k})}_{j}}$$

נגדיר: $N(k,j)$ הוא מספר הפעמים שיתקבל הסכום $l_{i_1}+...+l_{i_k}=j$.

$N(k,j)\le \sigma^{j}$,

$$\sum_{i=1}^{n}\sigma^{-l_i}\le (k\cdot l_{\max})^{\frac{1}{k}} \xrightarrow{k\to\infty} 1$$

לכל $\varepsilon>0$ מתקיים: $\sum_{i=1}^{n}\sigma^{-l_i}<1+\varepsilon$ ולכן $\sum_{i=1}^{n}\sigma^{-l_i}\le 1$. מש"ל

<u>תכנות דינאמי:</u>
דוגמה: מציאת כל המרחקים בגרף משוקלל ללא מעגלים שליליים.
קלט: גרף מכוון $G=(V,E)$ עם פונקצית משקל: $w:E\to\mathbb{R}$, ללא מעגלים שליליים.
פלט: לכל זוג צמתים, $i,j$ המרחק ביניהם: $d(i,j)$
בה"כ: $V=\{1,...,n\}$.
נכליל את פונקצית המרחק $w$ לפונקציה $\bar{w}$ באופן הבא:

$$\bar{w}(i,j)=\begin{cases}0 & i=j\\ \infty & (i,j)\in E, i\neq j\\ w(i,j) & (i,j)\in E\end{cases}$$

<u>כלל שיפור מקומי מוכלל:</u> אם $d(i,k)+d(k,j)<d(i,j)$ אז: $d(i,j)\leftarrow d(i,k)+d(k,j)$

פונקציה $d:v\times v\to\mathbb{R}$ היא <u>פונקצית חסם עליון מוכללת</u> אם לכל זוג צמתים $i,j$ מתקיים:
$dist(i,j)\le d(i,j)\le \bar{w}(i,j)$.

<u>טענה:</u> תהי $d:v\times v\to\mathbb{R}$ פחע"מ (פונקצית חסם עליון מוכללת), אזי:
א. $d$ נשארת פחע"מ גם לאחר ביצוע שיפור מקומי מוכלל.
ב. אם לכל שלושה צמתים $i,j,k$ מתקיים: $d(i,k)+d(k,j)\ge d(i,j)$ אז לכל זוג צמתים $i,j$ מתקיים: $d(i,j)=dist(i,j)$.

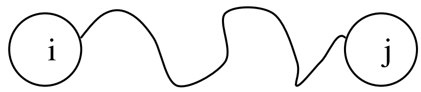
<u>אלגוריתם של פלויד וורשל:</u>
חלוקת המסלולים לרמות:
רמה 0: מסלולים בעלי קשת אחת (או פחות).
רמה 1: מסלולים שקבוצת הצמתים הפנימיים שלהם מוכלת ב $\{1\}$
רמה 2: מסלולים שקבוצת הצמתים הפנימיים שלהם מוכלת ב $\{1,2\}$
רמה $i$: מסלולים שקבוצת הצמתים הפנימיים שלהם מוכלת ב $\{1,2,...,i\}$
רמה $n$: מסלולים כלשהם.

<u>אתחול:</u>
לכל קשת $(i,j)$, $d^0(i,j)=w(i,j)$. אם $(i,j)\notin E$ אז $d^0(i,j)=\infty$ ו: $d^0(i,i)=0$.
$\left(d^0(i,j)\leftarrow \bar{w}(i,j)\right)$
עבור $k=1$ עד $k=n$ בצע:
לכל זוג צמתים $(i,j)$ בצע:

$$d^k(i,j)\leftarrow \min\left\{d^{k-1}(i,j), d^{k-1}(i,k)+d^{k-1}(k,j)\right\}$$

לכל זוג צמתים $(i,j)$ החזר את $d^n(i,j)$.

<u>טענה:</u>
עבור כל $k$ כך ש: $0 \le k \le n$ יתקיים ש: $d^k\left(i,j\right)$ הוא המשקל הקל ביותר מרמה $k$ בין $i$ ל $j$
(אם אין מסלול כזה $\infty$)
ההוכחה באינדוקציה על $k$:
<u>בסיס:</u>
עבור $k=0$ טריביאלי (נובע ישירות מההגדרה).
<u>צעד:</u>
נניח נכונות עבור $k-1$ ונוכיח עבור $k$.




מסלול קל ביותר מרמה k

אם k לא נמצא במסלול קל ביותר ברמה k אז $d^k(i,j)=d^{k-1}(i,j)$ ולכן הטענה מתקיימת ע"פ ה"א
אחרת: $d^k(i,j)=d^{k-1}(i,k)+d^{k-1}(k,j)$ ושוב הטענה מתקיימת ע"פ ה"א.

<u>סיבוכיות:</u> $O\left(n^3\right)$