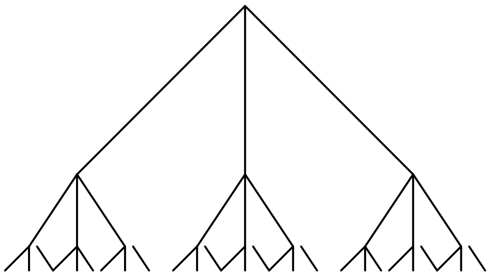
**<u>תזכורת:</u>**



**<u>עצי k- HST</u>** (הגדרה שונה במעט מזו שניתנה בשבוע שעבר)
מרחב מטרי המוגדר על העלים של מושרש.

לכל צומת v משויכת תווית:
$\Delta(v)$ ($\Delta(v) \in \Re^{+}$).

ברמה ה-i-ית בעץ:
$\Delta(v) = \phi \cdot k^{-i}$.

לכל שני עלים x,y:
$d(x,y) = \Delta(lca(x,y))$.

מעצם ההגדרה נוצר היררכי עם קוטר הולך וקטן.

**<u>משפט:</u>**
כל מרחב מטרי על n נקודות ניתן לשיכון ב- k-HST עם עיוות $O(k \log_k n)$.

**<u>חלוקה הסתברותית:</u>** $(\lambda, \Delta)$
חלוקה של המרחב $M=(V,d)$,
לצבירים: $c_1, c_2, ..., c_n$
כאשר:

- $c_i \cap c_j = \varnothing$.
- $\bigcup_i c_i = V$.
- $\dim(c_i) \geq \Delta$.
- $\forall u,v : \Pr[u \in c_i \wedge v \in c_j, i \neq j] \leq \frac{\lambda}{\Delta} d(u,v)$.

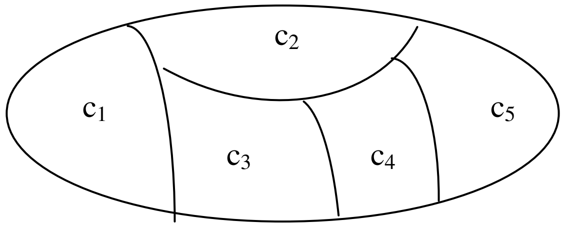
**<u>משפט מוחלש:</u>**
בהינתן $(\lambda, \Delta)$ חלוקות ל- M, לכל $\Delta$.
אז המרחב ניתן לשיכון ב- k-HST עם עיוות:
$O(k\lambda \log_k \phi)$.

| |
|---|
| הקוטר המנורמל של M. (הקוטר המקורי מחולק במרחק המינימאלי).<br>כאשר קוטר המסומן גם ב- daim זהו המרחק הגדול ביותר בין קודקודים במטריקה. |

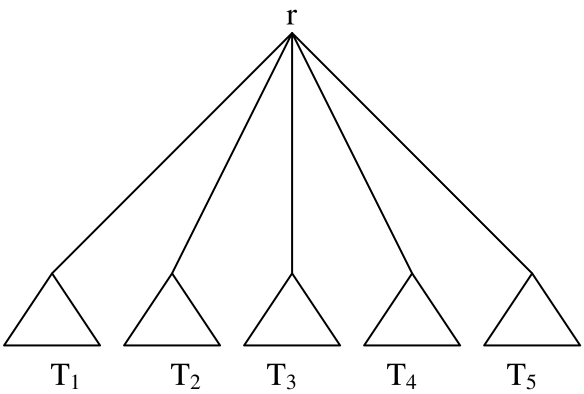
**<u>הוכחת המשפט המוחלש:</u>**
באינדוקציה על גודל המרחב.
**בסיס האינדוקציה:** n=1, הטענה טריוויאלית.
**שלב האינדוקציה:**
נבנה $(\lambda, \Delta)$ חלוקה הסתברותית עם $\Delta = \phi / k$.

לכל צביר נבנה באופן רקורסיבי עץ k-HST:
$T_1, T_2, \ldots, T_n$.
((על פי הנחת האינדוקציה ניתן לעשות כן).

נקבל מצב המוצג בצורה הבאה:




מצב הצבירים.




לכל צביר בנינו עץ k- HST המסומן במשולש.

כעת נחבר את עצי הצבירים לעץ אחד על ידי הוספת שורש r.
נגדיר:

$\Delta(r) = \phi = diam(M)$ .

לפי הבנייה לכל (הרקורסיבית) $i \le t$ :

$\Delta(r_i) = diam(c_i) \le \phi / k$ .

**הערה:** ניתן היה להגיע לשיוויון.

כעת, עלינו להוכיח שני דברים (בכדי להראות שאכן קיבלנו חלוקה הסתברותית):

1. השיכון הוא אכן שיכון שאינו מכווץ.
2. $E[d_T(u,v)] \le [\lambda k \log_k \phi] \cdot d(u,v)$ .

**הוכחת 1:**

צ.ל. $\Delta r = diam(M) \ge d(x,y)$ .

באופן אינדוקטיבי אם (x,y) היו באותו צביר אז המרחק שלהם קטן על פי הנחת האינדוקציה.
אחרת, (הם בצבירים שונים):

$d(x,y) = \Delta r = \phi$ (אבל $\phi$ זה המרחק המקורי הגדול ביותר, ולכן המרחק d(x,y) רק גדל).

**הוכחת 2:**

ע"פ הנחת האינדוקציה:

$E[d_{T_i}(u,v)] \le [\lambda k \log_k \phi / k] \cdot d(u,v)$ .

$$E(d_T(u,v)] \le \Pr[u \in c_i \wedge v \in c_j, i \ne j] \cdot \phi + \Pr[u,v \in c_i] \cdot E[d_{T_i}(u,v)] .$$

$$\le \frac{\lambda}{\Delta} d(u,v) \cdot \phi + [\lambda k \log_k (\phi / k)] d(u,v)$$

$$\le \frac{\lambda}{\phi / k} d(u,v) \cdot \phi + [\lambda k \log_k (\phi / k)] d(u,v)$$

$$= \lambda \cdot k \cdot d(u,v) + \log_k (\phi / k) \cdot \lambda \cdot k \cdot d(u,v)$$

$$= \lambda \cdot k \cdot d(u,v)[1 + \log_k (\phi / k)]$$

$$= [\lambda k \log_k \phi] \cdot d(u,v)$$

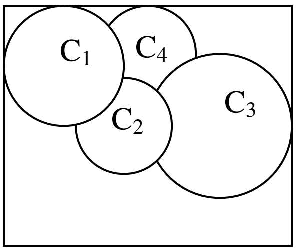
**למה:**

לכל מרחב מטרי על n נקודות, $\forall \Delta > 0$ ,

יש $(\log n, \Delta)$ חלוקה הסתברותית.

**הוכחה:**

בהינתן $\Delta > 0$ :




נבחר נקודה כלשהי $v_1 \in V$ ,

נבנה סביבה צביר (שהוא למעשה כדור) באופן הבא:

נבחר רדיוס בהתפלגות שנקבע בהמשך.

ונגדיר:

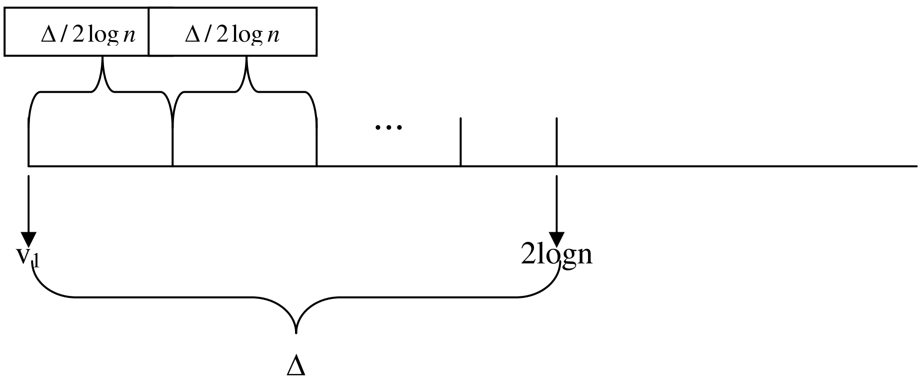
$C_1 = B\ (v_1, r_1)$.

נמשיך באותו אופן על תת המרחב שנותר: $V\text{-}\ C_1$,

(כלומר נבחר נקודה שאינה בכדור, ונבנה סביבה צביר)

עד שנקבל חלוקה מלאה של כל המרחב לצבירים.

כעת נסביר מהי ההתפלגות של r (יתכן כי $r > \Delta$ אח"כ נתקן):




נבחר את r להיות בקטע $[\frac{(i-1)\Delta}{2\log n}, \frac{i\Delta}{2\log n}]$

בהסתברות $1/2^i$.

לאחר שקבענו באיזה קטע r יהיה, נבחר אותו בהסתברות אחידה מקטע זה.

נרצה להוכיח:

$$\forall u, v : \Pr[u \in c_i \wedge v \in c_j, i \neq j] \leq \frac{2\log n}{\Delta} d(u,v)$$

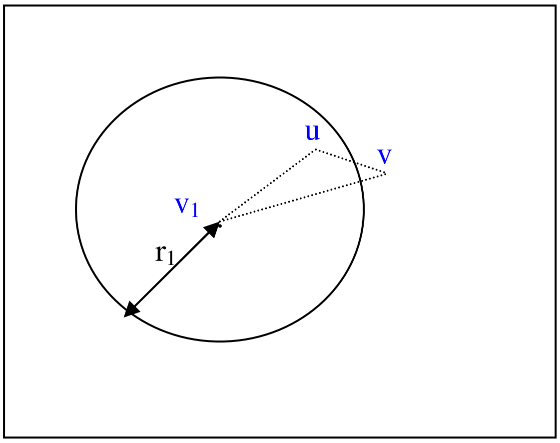
נוכל להסתכל רק על u,v המקיימים:

$$d(u,v) \leq \Delta / 2\log n$$

(על קואורדינאטות אחרות הטענה לא אומרת הרבה).

נניח בה"כ:

$$d(v_1, v) \geq d(v_1, u).$$




נחלק לשני מקרים :

**מקרה רע:**

ההסתברות ש- u,v מופרדים בשלב זה,
זה הסיכוי ש-

$$d(v_1,u) < r \le d(v_1,v)$$

$$d(v_1,v)-d(v_1,u) \le d(u,v) < \frac{n}{2\log n}$$

אי-שיוויון המשולש

**מקרה טוב:**

$. r_1 \ge d(v_1,v)$

Pr[מקרה רע בשלב 1] $\le \frac{1}{2^i}\cdot\frac{d(v_1,v)-d(v,u)}{\Delta/2\log n}$

ההסתברות שנבחר הקטע ה-i-י בו u,v נמצאים

ההסתברות ש-u במעגל ו- v לא

$$\le \frac{1}{2^i}\cdot\frac{2\log n}{\Delta/2\log n}\cdot d(u,v)$$

ניתן להכליל את הטענה באופן לשלב ה-j-י :

Pr[מקרה רע בשלב j] $\le \frac{1}{2^i}\cdot\frac{d(v_j,v)-d(v,u)}{\Delta/2\log n}$

ההסתברות שנבחר הקטע ה-i-י בו u,v נמצאים (לאו דווקא אותו ה-i של שלב 1)

ההסתברות ש-u במעגל ו- v לא

$$\le \frac{1}{2^i}\cdot\frac{2\log n}{\Delta}\cdot d(u,v)$$

Pr[מקרה טוב בשלב j] $\ge \sum_{l>i}\frac{1}{2^l}$

נתעלם מן שייתכן ש-r נופל בקטע ה-i –י, ואנו עדיין במקרה הטוב

Pr [מקרה רע בשלב j] $\le$ Pr [מקרה טוב בשלב j] $\cdot\frac{2\log n}{\Delta}\cdot d(u,v)$

$\Rightarrow$ Pr [מקרה רע ] $\le$ Pr [מקרה טוב] $\cdot\frac{2\log n}{\Delta}\cdot d(u,v)$

אם במהלך כל שלב היחס נשמר, אזי הוא נשמר גם בסופו של התהליך.

$$\Rightarrow \quad \Pr[\text{מקרה רע}] \le \frac{2\log n}{\Delta}\cdot d(u,v)$$

$$\Pr[\text{מקרה טוב}] < 1$$

מדוע ההסתברות לכך שנבחר $r \ge \Delta$ הינה קטנה מספיק?
ההסתברות לכך היא:
נסמן:
: $i_\Delta = 2\log n$

$$\frac{1}{2^{i_\Delta}} \le \frac{1}{n^2}$$

וזאת עבור כל שלב.
וכלן ע״פ חסם האיחוד נקבל שההסתברות שקיבלנו שאחד מן הרדיוסים שנבחר גדול מ- $\Delta$
הינה $\le \frac{1}{n}$. (ולכן בהסתברות קטנה יקרה מצב זה, ואם אכן יקרה נחזור על התהליך מחדש).

דוגמא לשימוש:

## **Metric Labeling**

נתונים:

1. קבוצה O של אובייקטים בגודל n.
2. קבוצה L של תוויות בגודל l.
3. יש מחיר השמה הנתון ע״י:

$$\forall p \in O, \forall a \in : c(p,a)$$

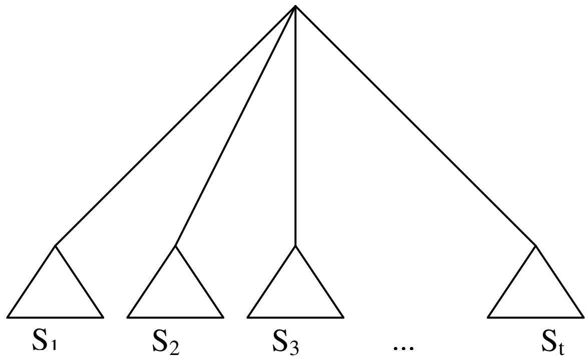
4. מרחב מטרי על התוויות.
5. לכל שני אובייקטים $p,q \in O$ יש משקל $w_{pq}$.

רוצים לתת תווית לכל אובייקט:
.f: O→L
באופן הבא:

$$\min \sum_{p\in O} c(p, f(p)) + \sum_{p,q\in O} w_{pq} d(f(p), f(q))$$

ניתן אלגוריתם 2 קירוב ל- k-HST (k=2).
נקבל אלגוריתם למטריקה כללית בקירוב O(log(l)) . (ע"י שיכון הסתברותי ב- k-HST של מטריקת התוויות).

איזה סט תוויות נבחר
בתוך $S_i$ המרחקים שווים.




קיבלנו רדוקציה לבעיית Metric Labeling במרחב יוניפורמי.
בעיה זו ניתן לפתור בעזרת תוכנית ה- LP.
ובעזרת הפתרון השיברי ניתן לבצע randomized rounding על מנת להגיע לפתרון בשלמים.