

# אלגוריתם

⊕ מציאת פתרון מינימלי (בעזרת אלגוריתם דיינסטרוי)

Reverse Delete - צורה של קטענות

Kruskal/Prim - מציאת פתרון מינימלי

הצורה של פתרון מינימלי R.D. מציאת פתרון מינימלי

הצורה של פתרון מינימלי  $(E=V-1)$  R.D. מציאת פתרון מינימלי

כמו כן PRIM מציאת פתרון מינימלי מציאת פתרון מינימלי

⊕ אלגוריתם דיינסטרוי - מציאת פתרון מינימלי

(1) מציאת פתרון מינימלי (אלגוריתם דיינסטרוי)

מציאת פתרון מינימלי (2)

(2) מציאת פתרון מינימלי (אלגוריתם דיינסטרוי)

(3) מציאת פתרון מינימלי (אלגוריתם דיינסטרוי)

אלגוריתם דיינסטרוי - מציאת פתרון מינימלי

אלגוריתם דיינסטרוי

אלגוריתם דיינסטרוי - מציאת פתרון מינימלי

הצורה של פתרון מינימלי  $S$  מציאת פתרון מינימלי

$T$  מציאת פתרון מינימלי

$S_0 = \{i\}$   $T_0 = G - \{i\}$  מציאת פתרון מינימלי

$d(i, i) = 0$   $d(i, j) = \infty$  מציאת פתרון מינימלי

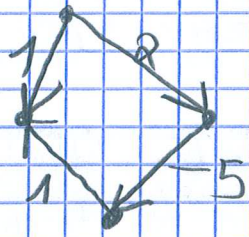
מציאת פתרון מינימלי (אלגוריתם דיינסטרוי)

מציאת פתרון מינימלי

$$d_{L+1}(i, j) = \min(d_L(i, j), d_L(i, k) + w_{k, j})$$

# אלגוריתם Belman-Ford

אלגוריתם פתרון בעיה של חישוב מסלול קצר ביותר (אורך) בין שני נקודות.

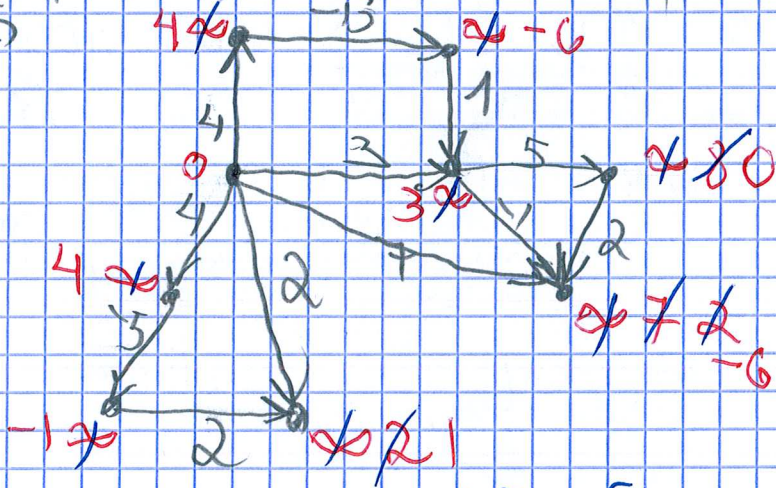


אנחנו רוצים למצוא את המסלול הקצר ביותר.

הבעיה היא שיש לנו מסלולים שונים וצריך למצוא את הקצר ביותר.

$$d_{L+1}(i,j) = \min(d_L(i,j), d_L(i,k) + w_{k,j})$$

ההתקדמות היא שכל פעם אנחנו מחשבים את המרחקים הקטנים ביותר.



כל פעם אנחנו מחשבים את המרחקים הקטנים ביותר (בניסוח אחר).

## מרחק מן המקור

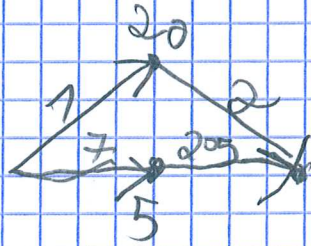
המרחק מן המקור הוא המרחק הקצר ביותר.

המרחק מן המקור הוא המרחק הקצר ביותר בין המקור לנקודה.

$$d^5(1,3) = \infty$$

$$d^{20}(1,3) = d^0(1,3) = 212$$

$$d^{21}(1,3) = 3$$



לפי הבעיה, המרחק הקצר ביותר הוא 3.

$$d^{21}(1,3) = d^{20}(1,20) + d^{20}(20,3)$$

המרחק הקצר ביותר הוא המרחק הקצר ביותר.

$$d^{21}(1,3) = d^{20}(1,3)$$

$$d^{21}(1,3) = \min(d^{20}(1,3), d^{20}(1,20) + d^{20}(20,3))$$

הוא יתכן שיש קיצור, אבל לא

$$d^{k+1}(i,j) = \min(d^k(i,j), d^k(i,k) + d^k(k,j))$$

### Floyd-Warshall

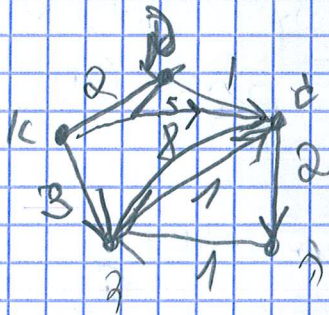
מחשב את המרחקים בין כל זוג צמתים

$[O(n^3)]$

$$d^k(i,j) = \begin{cases} w_{i,j} & \text{יש קצה} \\ \infty & \text{אין קצה} \end{cases}$$

$$d^{k+1}(i,j) = \min(d^k(i,j), d^k(i,k) + d^k(k,j))$$

$O(n^3)$  בזמן;  $O(n^3)$  זיכרון



$d^k$	k	p	c	z	v
k	0	2	5	3	8
p	$\infty$	0	1	$\infty$	$\infty$
c	$\infty$	$\infty$	0	8	2
z	$\infty$	$\infty$	1	0	$\infty$
v	$\infty$	$\infty$	$\infty$	1	0

$d^1$	k	p	c	z	v
k	0	2	3	3	$\infty$
p					
c					
z					
v					

$d^2$	k	p	c	z	v
k	0	2	3	3	$\infty$
p					
c					
z					
v					

$d^3$	k	p	c	z	v
k	0	2	3	3	5
p	$\infty$	0	1	$\infty$	3
c	$\infty$	$\infty$	0	8	2
z	$\infty$	$\infty$	1	0	3
v	$\infty$	$\infty$	$\infty$	1	0

הצורה הזו היא המרחקים בין כל זוג צמתים

$d^0$	k	d	e	z	o
k	0	2	3	3	5
d	8	0	1	8	3
e	11	8	0	8	2
z	8	8	1	0	3
o	8	8	2	1	0

יכנס למטריצת האלה  
(d' - d)

עלות זמן של האלגוריתם הנ"ל  $O(n^2)$ ; יחס  $O(n^3)$

### טבלאות זינג (Hash Table)

נרצה למצוא נתונים רבים בתצורה וצורה קבועה-אם כן  
 בלי ~~לכתוב~~ אצבע זמן זה מאפשר לעלות השאלה  
 האם יש צורך לפרק את המפתח ~~ל~~ שאלו כנס  
 הפרטים השונים הוא ייתכן שאלו היצג של ז"א  
 הוא גוף מצב.

פונקציה זינג - נ"ל שנותנת תוצאות קטנה לשל  
 $n-10^5$  נחיל קטן  $n-10^6$  עומד סוף לתוצר  
 בין  $1-10^6$  ותוצר כמעט-ח"ל הולך יש מצב  
 של התבססות נוצר מנגנון היצג או שאלו מקום  
 נמצא כשיהיה נוצר הם מקום כמה שמתא.

היצג-זינג אולי מנגנון-זינג פתוח



את האלמנטים המתוארים על ידי

$$Z_{ij}^1 = \begin{cases} Z_{ij} & \text{זוגי שמים} \\ 0 & \text{אחרת} \end{cases}$$

האלמנטים אלה (אחסן) וכן אלה אלוילים

### טבלאות גיבוב

אנחנו נבנים טבלאות גיבוב של קטן ושל גדול יותר. מטרתנו היא להבין את המבנה של הטבלאות האלו.

הטבלאות אלו הן מטריצות סימטריות. כלומר,  $Z_{ij} = Z_{ji}$ .

אם  $Z$  היא מטריצה סימטרית, אז  $Z^T = Z$ .

הטבלאות אלו הן מטריצות סימטריות. כלומר,  $Z_{ij} = Z_{ji}$ .

טבלאות גיבוב פתוחים: כלומר, אין להם גבולות.

אם  $Z$  היא מטריצה סימטרית, אז  $Z^T = Z$ .

הטבלאות אלו הן מטריצות סימטריות. כלומר,  $Z_{ij} = Z_{ji}$ .

$$\alpha = \frac{\sum_{i,j} Z_{ij}}{m}$$

הטבלאות אלו הן מטריצות סימטריות.

אם  $f(x) = f(x_1) = f(x_2)$  ויש להם אותו המבנה.

הטבלאות אלו הן מטריצות סימטריות. כלומר,  $Z_{ij} = Z_{ji}$ .

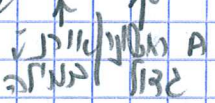
אם  $Z$  היא מטריצה סימטרית, אז  $Z^T = Z$ .

הטבלאות אלו הן מטריצות סימטריות. כלומר,  $Z_{ij} = Z_{ji}$ .

$$E[Z_{ij}] = \frac{\sum_{i,j} Z_{ij}}{m} = \alpha$$

הטבלאות אלו הן מטריצות סימטריות. כלומר,  $Z_{ij} = Z_{ji}$ .

$$f(x) = \sum_{i,j} x_i x_j A_{ij}$$



אם  $A$  היא מטריצה סימטרית, אז  $A^T = A$ .

הטבלאות אלו הן מטריצות סימטריות. כלומר,  $Z_{ij} = Z_{ji}$ .

$$[x^T A - x^T A] B$$

פונקציה  $f$  היא טובה עבור  $x$  שהיא שלמים (הכפול תמיד סגור)  
 אצות אפילו  $u \cdot (xA)$  טובה יותר.

האמות עקב סגירות: נכון החלוקה עבור ישאלו חזקה  
 יש לה פונקציה ג'בוק  $f(x)$  עם סדרת פונקציה ג'בוק  
 חוצ  $f_i(x)$  (rehash), זאת עם התחנה בתוך  
 הטובה לבנה ואי אפשר לשרשר.

כאוב את  $f(x) = f(x)$  עם  $x$  כבר נטבול  $x$  נלק  $f(x)$   
 וכן כליה.

אפשר, עבור  $f(x)$  פונקציה ג'בוק נתונה אפשר לקחת  
 $f_i(x) = [f(x) + c] \cdot u$  - linear rehash  
 $f_i(x) = [f(x) + c^2] \cdot u$  quadratic rehash  
 $f_i(x) = [f(x)^i] \cdot u$  אפשר גם לקחת

בלדה - שני את  $A, B, C$  בטובה  
 ומתקנו את  $B$   
 כוונים  $A$  ב- $A$  ונחשב את  $C$ . נקח איזה שהיה  $B$ ,  
 נבוא יק, כלומר נסין  $C$  או בטובה, אתה בך  
 $A$  נכון, אם במחיקים נשים Deleted במקום  
 המתק, נל ממשיכים לרשף שלעקום מתקום  
 כך אלוטו. אם בטובה יש המכה Deleted, נכנס  
 אולי את  $B$  האיורים מהשל.

צמח גישה איתנו:

$$1 + \underbrace{\alpha}_{\text{קטוב אחר}} + \underbrace{\alpha^2}_{\text{שני קטוב}} + \dots \approx \frac{1}{1-\alpha} \approx \underbrace{\alpha}_{\text{דקלים}}$$

סיבים קוקייה: אר  $x$  קיימת שתי פונקציה ג'בוק  $f_1, f_2$  (התק'אל)  
 (1) אר  $x$ :  $f_1(x) \neq f_2(x)$   
 (2) או מתקיים  $f_1(x) = f_1(y)$  עם  $f_2(x) = f_2(y)$  (אם  $x \neq y$ )  
 צ' שרישה (2) תתקיים נבטול  $am < \text{ג'בוק המחה}$

הפונקציה היא כזו:

נשים את  $x$  ב- $f_1(x)$ . את  $f_1(x)$  נשים ב- $f_2(x)$ .  
את  $f_2(x)$  נשים ב- $f_3(x)$ . וכן הלאה.  
את  $y$  נשים ב- $f_1(y)$ . את  $f_1(y)$  נשים ב- $f_2(y)$ .  
את  $f_2(y)$  נשים ב- $f_3(y)$ . וכן הלאה.  
( $f_2(y) \neq f_1(x)$  וגם  $f_1(y) \neq f_2(x)$ )

כעת, אם יש לנו פונקציה  $f: A \rightarrow B$  ופונקציה  $g: B \rightarrow C$ ,  
אז  $g \circ f: A \rightarrow C$  היא פונקציה מ- $A$  ל- $C$ .  
כלומר, הפונקציה  $g \circ f$  היא הפונקציה שמתקבלת  
באמצעות חיבור הפונקציות  $f$  ו- $g$ .

⊗ הפונקציה  $f: A \rightarrow B$  היא הפונקציה  $f(x) = \frac{1}{x}$ .

(1) הפונקציה  $f: A \rightarrow B$  היא הפונקציה  $f(x) = \frac{1}{x}$ .  
הפונקציה  $g: B \rightarrow C$  היא הפונקציה  $g(x) = x^2$ .  
אז  $g \circ f: A \rightarrow C$  היא הפונקציה  $g \circ f(x) = \left(\frac{1}{x}\right)^2 = \frac{1}{x^2}$ .

(2) הפונקציה  $f: A \rightarrow B$  היא הפונקציה  $f(x) = \frac{1}{x}$ .  
הפונקציה  $g: B \rightarrow C$  היא הפונקציה  $g(x) = \frac{1}{x}$ .  
אז  $g \circ f: A \rightarrow C$  היא הפונקציה  $g \circ f(x) = \frac{1}{\frac{1}{x}} = x$ .

⊗ הפונקציה  $f: A \rightarrow B$  היא הפונקציה  $f(x) = \frac{1}{x}$ .  
הפונקציה  $g: B \rightarrow C$  היא הפונקציה  $g(x) = \frac{1}{x}$ .  
אז  $g \circ f: A \rightarrow C$  היא הפונקציה  $g \circ f(x) = \frac{1}{\frac{1}{x}} = x$ .



(12/12/21) (11)

# עצים ורשתות

Prim: מ"צ רץ מניחיל פורם יש את השיטה D (קובקובים)  
(פ"ק) והשיטה DT (הצ' לא בה"פ).

בידינו הפשוטה ביותר היא שימוש בקבוצת מיונים,  
ומתוך נוצר הקשת היחיד ביותר.

ניצב עלי עליה (יש קשתות אם משקל וק"י

מקצב (השיטה קובקובים עם משתנה בינארי DT/DT).

במצב עם איבר בקבוצה עם משקל אצל קובקובים  
בהם מקצה, נבדוק האם הצלח יותר - מקיים

לחצו ב-D והשני ב-DT.

אם כן - נחילי דהן DT ב-D ונעצב מס' הויבריס ב-D

קמלוק:

בהם נחצו נחצו "ע" (קצבם ימים לאינם חלפים),  
בהם נחצו נחצו את הקשת היחיד ביותר ולאו

סלקת מקצה.

נצטרך מנהי נחצו שמצד צדו עם קובקוב לאינה חב

היו ש"ג, צדו אלוה לאינה חב שיינים וכו'

לחצו בין חובסית - נצב בצהר הצדק ביונה (צ"ו (U)).  
(מצב עם נקט Union-Set).

במצב, הקובקובים יהיו Union-Set (הקשתות בקבוצת

מ"ן נחילי קשתות עם ק מחדשת חוב' קטילות,

אל נחצו Union-Set ונחילי קשת עלי (לחצו, נחצו)  
קשת הכולל.

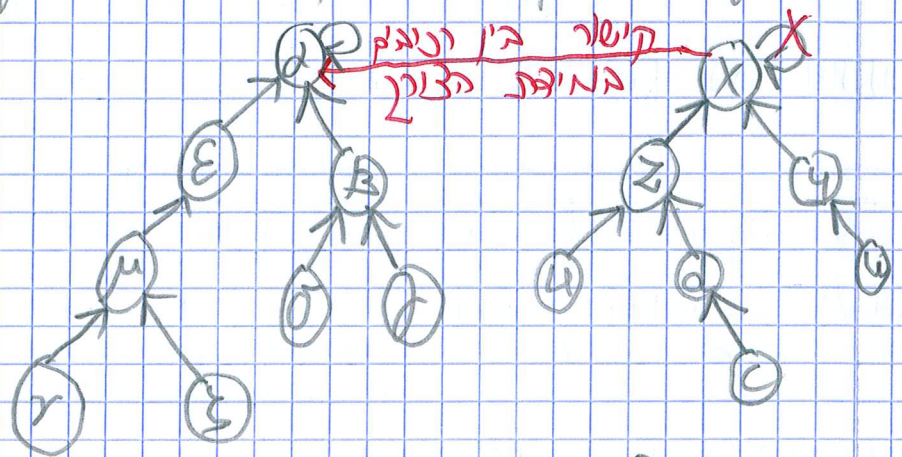
אם צברים בקבוצה  $O(E \log E)$

(הקבוצות חלפות  $O(E \cdot \text{Union-Set})$ )

עם  $O(E \log E)$  חלפות ל Union-Set

על הצדדים, אלו בעל און את נחצו נחצו כצ

קצת נבחרת אחרון נתונים נוספים  
 א כתיבה מה שירות "לבו" ע"י פ  
 מציגה אביו כמה, ואם הפי מציגה אביו



כעת, אמצאת הכיתה הקטילה

Find(y) {

if  $y == y.parent$ : return y;

else: return Find(y, parent);

$h = \log_2 E$  יוצר, הנתן (אם), יוצר

עבור איתור בין הכי קטילה

Union(x, y) {

z = Find(x);

w = Find(y);

w.parent = z;

יש לך בעיות (אם)

קשה, נשים את כל פני הנתן וזה נוסף תמיד פ

כיצד נחמת אחרון אלה פ ישר, במקרה כזה

הן כאשר נתן שני עצים בעלי שמה

לפני: בחיבור כזה נתן נתן הוא לא היות  $\log_2(k)$  כאשר

א א האחרים נתן

אנני, נבחר סטנדרט כי על האחרים נתן את אחרת

אם  $2^m \leq k < 2^{m+1}$ ,  $H \leq \log_2(k)$

$H=0 \leftarrow H=1$  אולי האדם הקוף וכו'  $k=1$  טרנס

כמה, נניח וכוונות  $H-1$ .  
 על פניו  $H$  יוצר את ויק לאותו שני דברים קרובים זה  
 הוציא  $H$  מה עליו הוא אחרת  $H-1$  וכן הוא עליו אחרת  
 ,  $H$  ,  $H-1$

המקרה של סוף העץ שלן תהיה

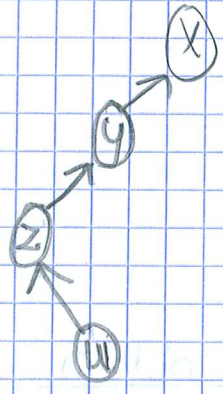
$$O(E \log V + E \log E) = O(E \log E)$$

↑  
 עליות אחר  
 כפולות  
 הנכנסות לזמן  
 קצת מעליהם

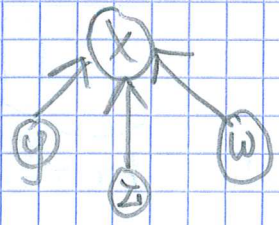
אולי  
 במה  
 זמן  
 כי בסידור קטנה נצטרך אין המטות  
 מקרה.

נקרא, נצטרף איך קנה יתר וקראו  $H=0$ .

ל  
 אולם  
 אולי  
 אחרת  
 אולי  
 אולי  
 אולי  
 אולי



Find(w) ⇒



נקרא, נקרא שטחים  $H = \log^*(n)$   
 $\log^*$  הוא מס' הפעמים שבה  $\log$  חף מ  
 על קראת מס' הקראו  $n-1$ .

נקרא שטחים  $O(E \log E)$  הקלות תהיה  
 מולו זניה.



12 (2/2/4)

# השאלות מחזוריות

ד-סקס ארוך;  $n$ -מחזוריות קצרה ( $1 < |S| < |T|$ )  
 נרצה ארנס את  $S$  ההופכות של  $S$  ב- $T$ .

הרישום הטיוטלי חלק (מחזוריות).

נוס' ארנס אלגוריתמים שיפוי (מחזוריות) ופוי  $(\frac{|T|}{|S|})$ , והאלגוריתמים שיפוי אלו הם:

$O( T )$	RABIN-KARP (RK)	1
$O( T )$	Knuth-Morris-Pratt (KMP)	2
$O( T )$	Boyer-Moore (BM)	3
$O( T )$		4

## אלגוריתם Rk

נרצה סקן' עיקוב  $f$  ונשווה בין לבין:

$$f(S) = f(T(i:i+L-1)); \quad i=0 \dots n-L$$

$L=|S|$        $K=|T|$       טיפוס

אלו נרצה שישם חיות המיילס  $f(T(i:i+L-1))$  ואלו  $O(L)$  ואלו  $O(1)$  ואלו

אלו ק' אלו (rolling hash) ארנס מתחילת

מ' ארנס חזרה  
מ' ארנס קצרה



$$f(V) = \sum_{i=0}^{L-1} V[i] \cdot A^i \cdot B$$

אלו מתחילת  $A$  מתחילת ארנס

$$f(\text{banana}) = 2 \cdot A^0 + 1 \cdot A^1 + 14 \cdot A^2 + 1 \cdot A^3 + 14 \cdot A^4 + 1 \cdot A^5$$

$$f(\text{ananaa}) = \frac{f(\text{banana}) - f(b)}{A} + f(a) \cdot A^{L-1}$$

מ' ארנס חזרה מ' ארנס קצרה מ' ארנס חזרה מ' ארנס קצרה





מאגרי

מאגרי

מאגרי

מאגרי

מאגרי

מאגרי

$S[m(m-1)]$   
מאגרי  
מאגרי

מאגרי

מאגרי

מאגרי



(ב/ב/ו/ו) (ב/ב/ו/ו)

# השאלות למחצית

## אלגוריתם R

נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן  
 אורכו  $n$  ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן

## אלגוריתם M

נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן  
 $prefix_i(S) = S[0:i-1]$

נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן  
 נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן

BANANA (ב) **ANANA**

נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן  
 נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן

ANANAS (ב)

נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן  
 נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן

**ANANA**

BANANABANANA

ANANAS

**ANANAS**

⇒ **ANANAS**

↑  
 נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן

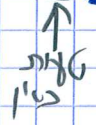
נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן  
 נתון  $S$  אורכו  $n$ , חתוך  $T$ , ונתון  $f(S)$  ונתון  $f(T)$  שונים להלן



ב היתא האבד היות של ש מחוקק א'תא ל היתא ה"ן

פיתוח

S: BANACANABANA



נת המחוצות ANA חוקקת ANA-ת, ונתק ב-נ

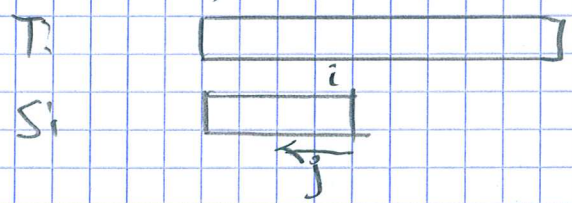
S: BANACANABANA



ANABANA ל היתא חוקקת BANA

האלגוריתם:

ל, א, א-1, א'קום היות ל, א  
 א-0, א'קום א-0



ל א'קום א'קום

ל א'קום א'קום

אם  $S[L-1-j] = T[i-j]$  :  $j++$   
 אחרת:

(1)  $i = i + BST(T[i-j])$  אם  $BST(T[i-j]) > GST(j)$

(2)  $j = 0$

(1)  $i = i + GST(j)$  אחרת:

(2)  $j = 0$

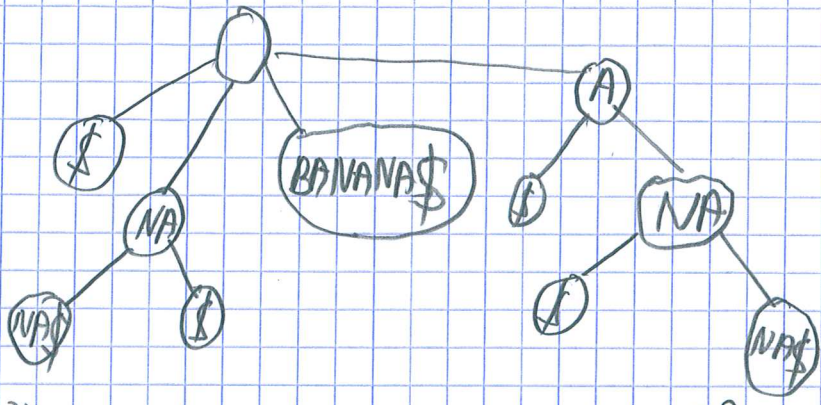
אם אחרת א'קום א'קום א'קום א'קום

BANANA

SUFFIX TRIE

א'קום א'קום א'קום א'קום א'קום א'קום  
 א'קום א'קום א'קום א'קום א'קום א'קום

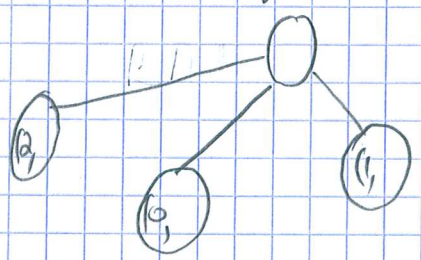


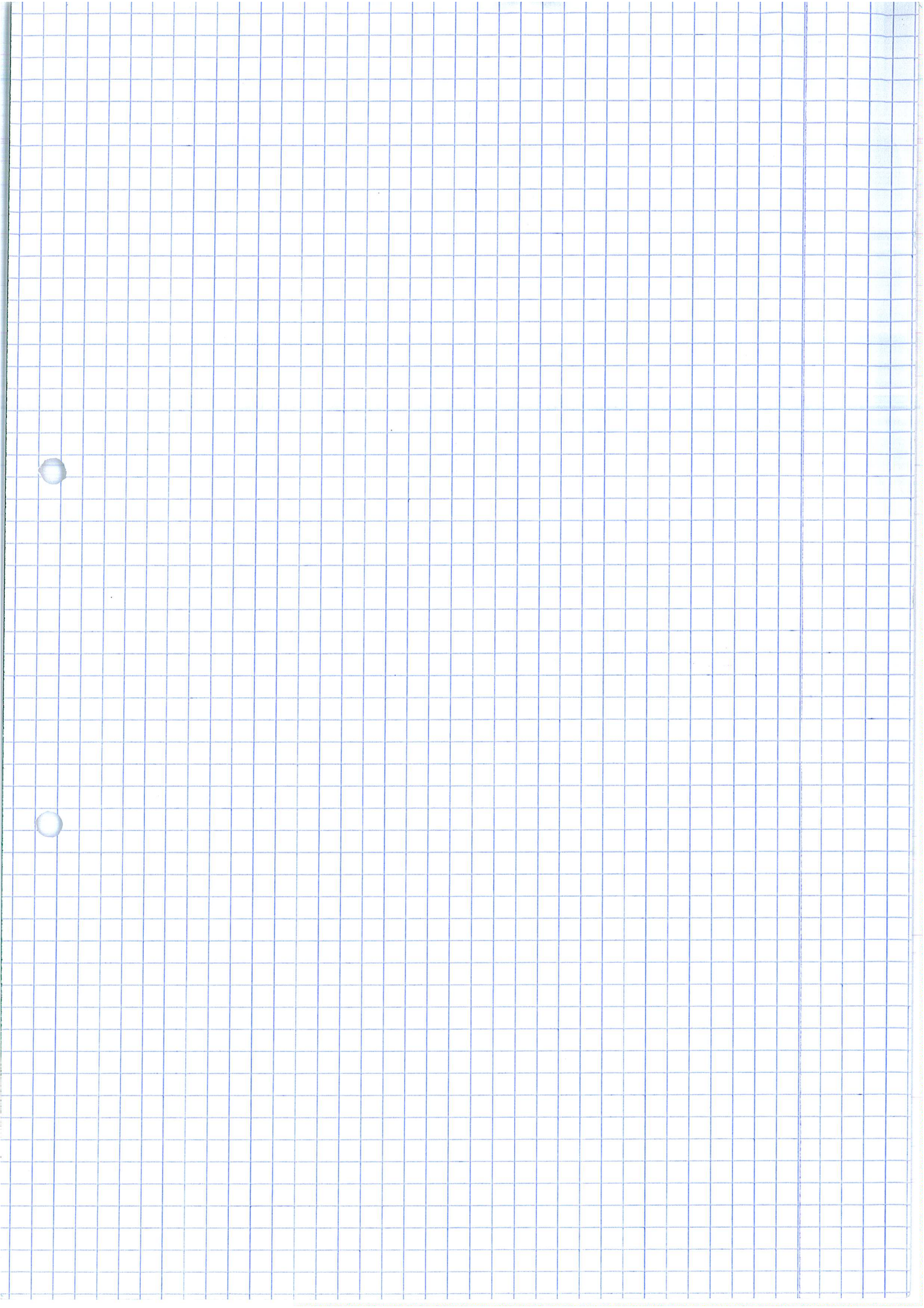


תורת המספרים - המספרים הם  
 אבסטרקט, הם לא פיזיקליים.  
 אבסטרקט, הם לא פיזיקליים.  
 אבסטרקט, הם לא פיזיקליים.

אבסטרקט, הם לא פיזיקליים.  
 אבסטרקט, הם לא פיזיקליים.  
 אבסטרקט, הם לא פיזיקליים.

0 1 2 3 4 5  
 BANANA





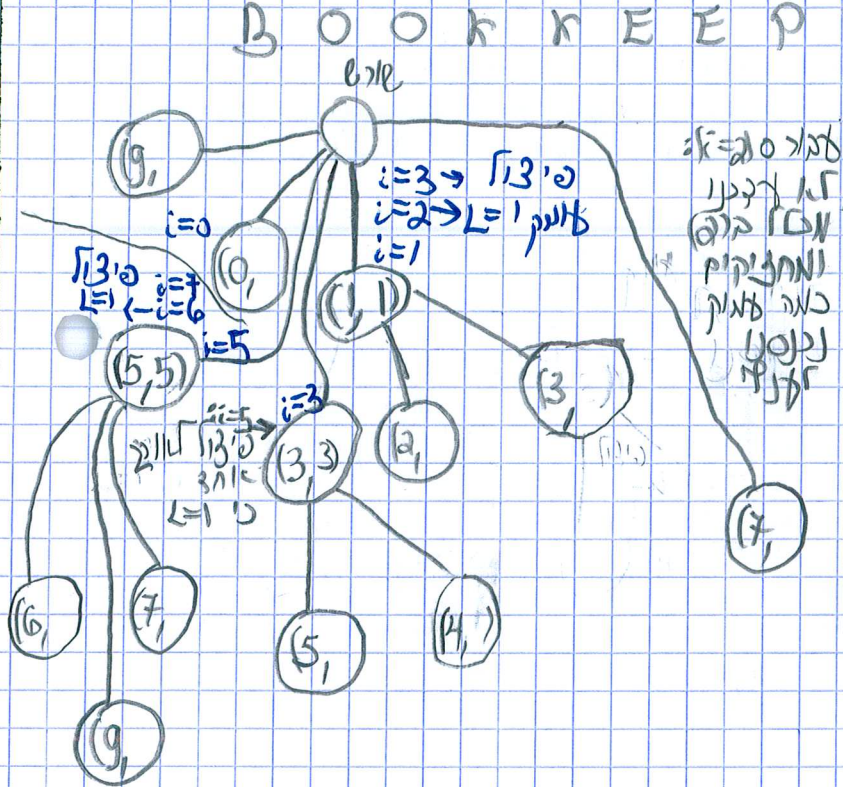


כל העלים מסתיימים ב-6 ולכן אין צורך לבדוק

אותו המכשיר

במסגרת נתונה סיבולת 2 וסיבולת 2. (x איות)  
 כל מילה המכילה את ה-6 היא מילה  
 קבוצה

i = 0 1 2 3 4 5 6 7 8 9  
 B O O K K E E P E R



$j=1$ : B O O K K E E P E R  
 $j=2$ : O O K K E E P E R  
 $j=3$ : ~~O O K~~ K E E P E R  
 $j=4$ : K K E E P E R  
 $j=5$ : K E E P E R  
 $j=6$ : E E P E R  
 $j=7$ : E P E R  
 $j=8$ : P E R  
 $j=9$ : E R  
 $j=10$ : R

כלים והעברה לעץ:

- א. שורשים ואם מתחבנים סוף של חיס
- ב. כאלה מוספים אלו חפשה יש שלוש אפשרויות:
  - (1) ק"מית היא בקף שמתיימה אלוה-
  - פיתחם חזק חפס.
  - (2) ק"מית היא בקף שמתיימה זלוה- סופ.
  - (3) נכנסני חזק וההמשך זה מתיים (טכנית: הפגל
  - שנה טאלס טאל חמש הקף הינוכחי זה מתיים):
  - (4) נכנס חזק ינוכחי
  - (5) אהתקדם ב-7



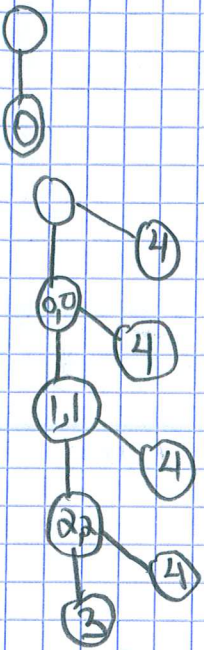




15 (18/12/12)

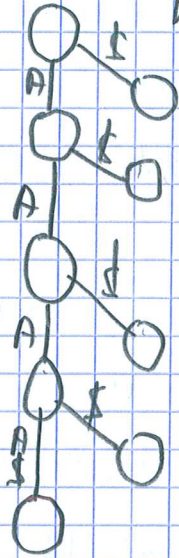
# צבי סוס

פס סוס - מנסה - היסטורי \$ בסוף  
אחת, הפך יקב סוס  
צבי AAAA, על סוס סוס ירה:



פס סוס מנסה ירה:

את הפס אנשי איצבי עם מקומות  
סו אי איתם, כיום הפס שלטון שקול:



באלימיתם אפנית הפס ישלנו  
ג-היות האחורה שהיסטורי  
פס הייסא מוהחחחח

באופן פשוט, כולל מוצאות  
ש האצו בחרים אול  
ובאולגה הפנימית עשויים

אורה אסות הפס (מחרחוס מהחחחחח)

loop  $i=1, \dots, |S|-1$

loop  $j=0, \dots, i$ :

מיס'ים סיסות מהאלגה ביותר ( $|S|+i$ )

אצ אקצרה ביותר ( $|S|+i$ )

פאי:

א	מפנימים	פאי
ב	מחציקים	כצורה מנסה קוקוקים פנימים
ג	מחציקים	suffix-links

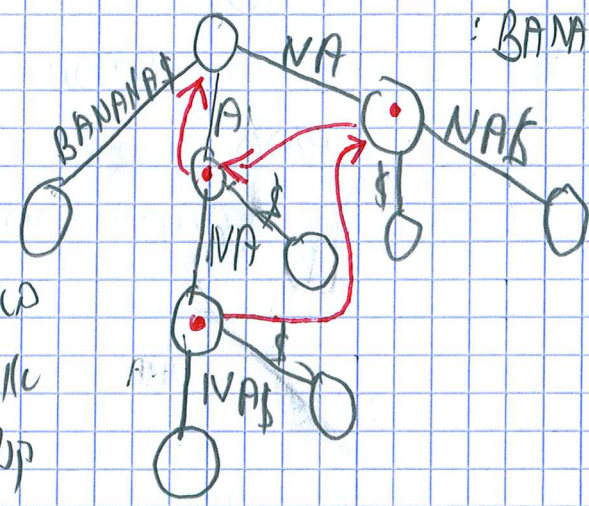
פצרה

אם  $v$  קודקוד פנימי שהחולו מהחולו אליו הוא  $u$

( $x$ -אול,  $z$ -מחחחחח) אם  $w$  הוא קודקוד פנימי שהחולו אליו הוא  $v$

W-F V-N suffix links מוספים גל, 2, 10, 11, 12

BANANA : BANANA סוף



suffix links - מוספים  
 כל קובץ מסוים ו-3  
 קשר בין

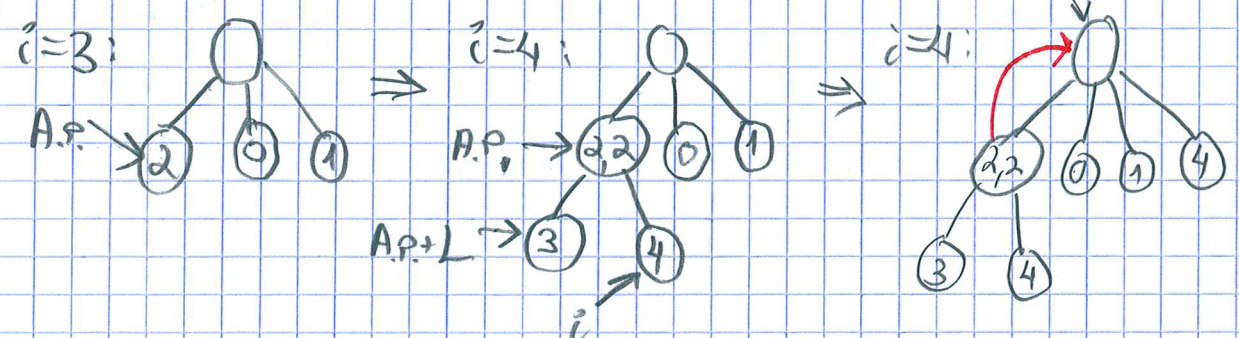
active point - מציין  
 end point - מציין  
 מציין את המיקום של האות  
 מציין את המיקום של האות

סיומי

אם אין את המטא-נתונים סיומי  
 יש את המטא-נתונים סיומי  
 אין את המטא-נתונים סיומי

VIDDAVIDAVID סוף

i:	1	2	3	4	5	6	7	8	9	10	11	12 (L)
j:	0	0	0									
L:	0	0	1	1	1	2	3	1	2	3	4	
D:	א	א	ב	ב	ב	ב	ב	ג	ג	ג	ג	ד



active point - מציין  
 מציין את המיקום של האות  
 מציין את המיקום של האות  
 מציין את המיקום של האות



בנין שני, HW<sub>2</sub> , HW<sub>1</sub> , HW<sub>2</sub>  
BFS , DFS , DFS , DFS







# צפייה ואינפורמציה (25/10)

נניח שנתפס שלנו מתפלג  $1-0$ , וזה אהב זהם  
 סימ'  $P(1), P(0)$

$P(1)$  אמה אמל  $P(1)=1$   $P(0)=0$ , אז הסיק הי'ן

צ'נים בקלות (כנתפס כמה פתמים  $1$  מופל).  
 ב' אה אמל  $P(0)=P(1)=0.5$  יותר קשר אהום.

צ'חיסה לקול אהתפלגות לא אהיה (כל אהתפלגות  
 אי'פ אהיה נ' יותר ק' אהום). אם נצ'ק  
 מצ'פ ל' אהיקול

נק'ת  $P(x) - \text{פוק'}$  ה'תפלגות,  $\sum P(x)=1$ ,  $P(x) \geq 0$   
 $-\infty \leq \log(P(x)) \leq 0$

וק'ת מצ'פ  
 $H = -\sum_x P(x) \log_2(P(x))$   
 א'ם פוק' ח'ים בסי' 2 ב'ה BIT  
 3 TRIT  
 e (NAT)

וק'ת י'פ מ'חיס'ט  
 $H \geq 0$

ש-  $H=0$  (ו'ז' התפ' שלנו י'ה ט'ב)  
 נצ'ק' אם  $P(x)=0$  "  $P(x)=1$

ח פוק' ה'תפלגות ול' ק'ים מ'אול'  $x$  ב'פ' אהום  
 $P(x)=1$  ול' מ'אול'  $y$  אה'  $P(y)=0$  אהיק'ה  
 כל ה'תפלגות מ'אול' לא אהיק'ה, ו'ש א'ו מ'פ'ה

ש' מ'פ'ל מ'אול' (פ'ול' מ'אול' מ'ה ק'רה אה'ו)  
 א'חיס'ט -  $H$ , מ'פ' אה'ר אהיק'ה ל' ה'מ'פ'ל

ס'פ'ט ב'ן א'ם  $H=0$  ש' א'חיס'ט מ'רבות, א'ז א'ם  
 $P(x)$  אה'ר, מ'אול'  $P(x)=\frac{1}{N}$  כ'א'ש מ' א'חיס'ט

א'חיס'ט מ'ק'  $\leftarrow H = -\sum \frac{1}{N} \log\left(\frac{1}{N}\right) = \log N$

בהינתן שיש  $N$  מילים באשר לכל מילה יש סב'  $p_i$  ואורך  $L = \sum_i L_i p_i$  הוא סך הכל הממוצע של  $L$ .

המילה של צמיחה היא לתקוף את  $L$  קוד בינארי - קוד שבו כל מילה היא לא ריבוי של אותיות  
 0, 10, 100, 1000, 10000, 100000, 1000000, 10000000, 100000000, 1000000000  
 אם נקרא את  $L$  ככה פירוט לא ידע קוד בינארי

מילוי בקוד בינארי  $H$ .

⊗ הקבוצה אינפורמציה - אם  $A$  ו- $B$  מקביר ניבוי  $X$  ו- $Y$

וזה  $P(X|Y) = 1/n$

נפרד אקט בולבולות מה קרה (צמיחות מילוי)

אחרת אם  $P(X|Y) = P(X)$

לא נוכל אקט פיר מה קרה

סדרה, ביטוי  $H(X|Y) = - \sum_{x,y} P(x,y) \log(P(X|Y))$  יש לנו אינסוף מילים

כאשר  $M(X,Y) = H(X) - H(X|Y)$  (האינסוף היא)

$H(X,Y) = - \sum P(x,y) \cdot \log(P(x,y))$

$H(X|Y) = - \sum P(x,y) \cdot \log(P(X|Y)) = - \sum P(x,y) \log\left(\frac{P(x,y)}{P(y)}\right) =$

$= - \sum P(x,y) [\log P(x,y) - \log P(y)] = H(X,Y) - H(Y)$

$\Rightarrow M(X,Y) = H(X) + H(Y) - H(X,Y)$

אינסוף - כמה יבד חפש משהו את הבהירות אי-אזנות הולך זיבן הקבוצה שיש לנו

⊗ אם  $p, q$  התפלגות  $P$  אלו מונה איתם אז  $-\sum P \log P \leq -\sum P \log Q$

נסתה עכשיו

$$-\sum p \log p - [-\sum p \log q] = -\sum p [\log p - \log q] = -\sum p \log \left(\frac{p}{q}\right)$$

$$= \sum p \log \left(\frac{q}{p}\right) \leq \sum p (1 - \frac{q}{p}) = \sum p - \sum q = 1 - 1 = 0$$

$\downarrow$   
 $\log(x) \leq 1 - \frac{1}{x}$

מרחק Kulback Leibler

$$D_{KL}(P, Q) = \sum p \log \left(\frac{p}{q}\right)$$

$$D_{KL}(P, Q) + D_{KL}(Q, P)$$

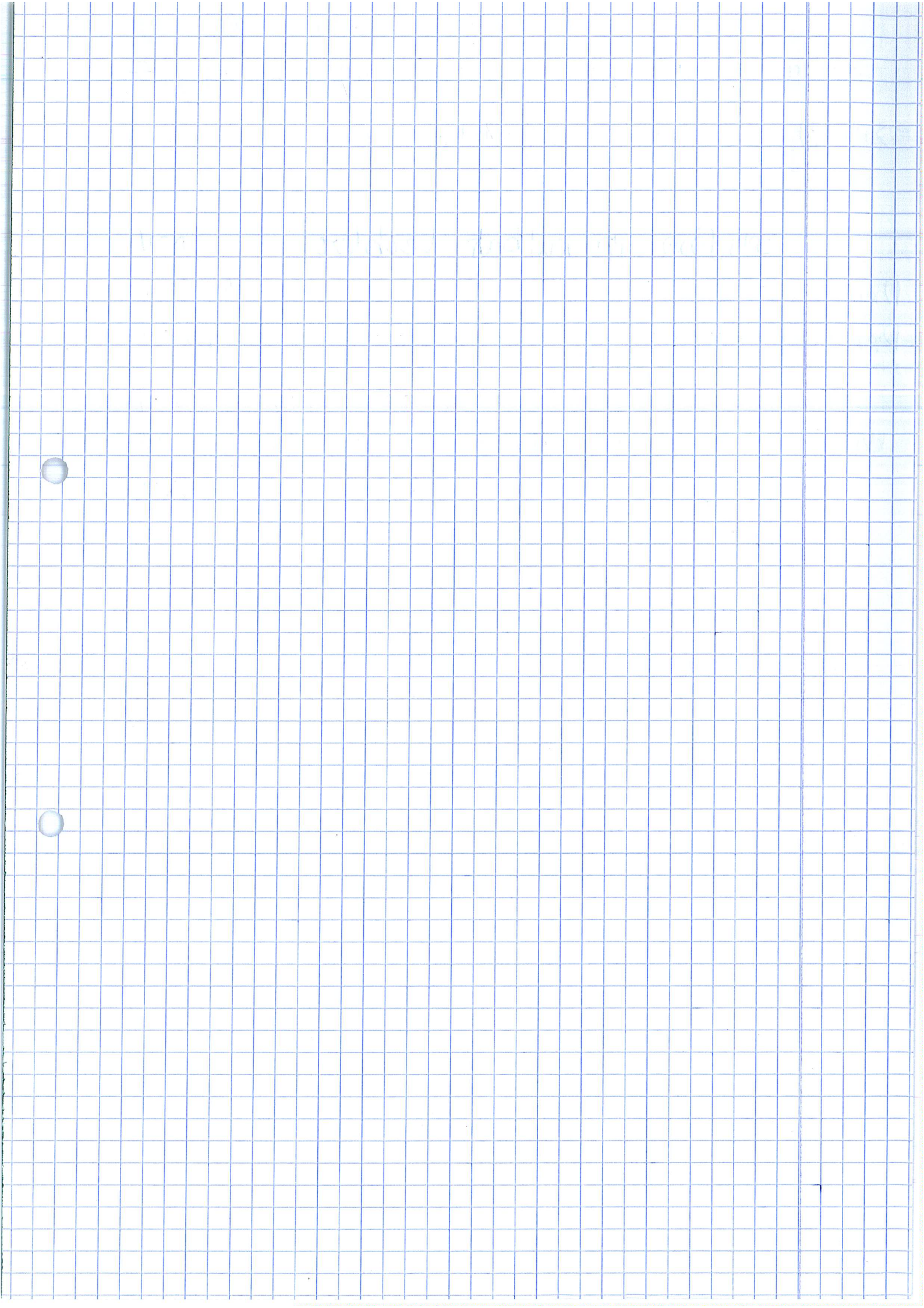
מרחק סימטרי  
מרחק סימטרי

כאן - אי שוויון Kraft

אם נתונות מילים  $\{x_i\}$  אז

$$\sum_x D^{-L(x)} \leq 1$$

סקר הייל





פרטים באמצעות קוד בינארי

מציאת קוד אופטימלי נדרש  $L = \sum p_i l_i$  מינימלי  
 $p_i$  - שכיחות מילה  $i$ ,  $l_i$  אורכה  
 חלוקה יתקיים  $L \geq H$  שומר  $\sum p_i l_i \geq -\sum p_i \log_2(p_i)$   
 משיג אילוות בא"ק

$H = -E[\log(p)]$  - אנטרופי, שומר מיפהת אי וולאיות

$$H(x) = -\sum p(x) \log(p(x))$$

$$H(x,y) = -\sum p(x,y) \log(p(x,y))$$

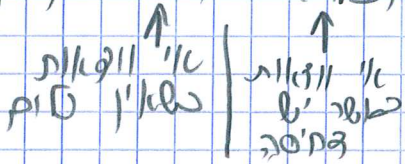
$$H(x|y) = -\sum p(x,y) \log(p(x|y))$$

אינאומציה משותפת:  $MI(x,y) = H(x) - H(x|y) = H(x) + H(y) - H(x,y)$

אם  $A$ ,  $B$  תמונה אלמנטית,  $B$  תמונה פרטים, נרצה

$$H(A|B)$$

$$MI(A,B) = H(A) - H(A|B)$$



⊗ ג'יט  $p, q$  התפלגויות:

$$-\sum p_i \log(p_i) \leq -\sum p_i \log(q_i)$$

⊗ אם משתמשים בקוד בינארי  $\sum D^{-l_i} \leq 1$

הוכחה: נניח את  $f$  הבינארי של קוד הבינארי ונניח  $D^{-L}$

אם באורך הקוד

אם הקוד הוא  $2^k$  יש  $D^k$  ענפים

אם יש  $2^k$  ענפים,  $2^k$  קודם של  $2^k$  מספר  $2^k$  שלם

קודם היתרון יורד

$$\sum D^{-l_i} \leq D^k \Rightarrow \sum D^{-l_i} \leq 1$$

⊗ ענפים "אלמנטים"







0, 1, 3, 0, 2, 5, 7, 4, 4, 7, 7

פתיחה:

↓

הקוויבצ'ה  $AB, S$  ויציאה -  $AB, S$

$AB \Rightarrow$  אחרת תהיה א-1-1 (יובאים את זה) אמיליון  $AB=3$  הוסטני

מילי קובצ'ה + אחרת קובצ'ה אמיליון  $AB=3$  עקב

של מילי נכחית.

של התחילי, נקבא:

0 1 3 0 2 5 7 4 4 7 7

A|B|AB|A|S|ABA|SA|BA|BA|SA|SA

0 A I SA

והמיליון ירדה:

1 B 8 ABAS

2 S 9 SAB

3 AB 10 BAB

4 BA 11 BAS

5 ABA 12 SAS

6 AS

