קריפטוגרפיה 2

הרצאה 5: פונקציות פסאודואקראיות (PRF)

Textbook: Katz and Lindell. Introduction to Modern Cryptography.

מרצה: אורי שטמר

- מוטיבציה: נשתמש ב PRF כדי לבנות מע' הצפנה בטוחה כנגד הצפנות מרובות
- לפונק' פסאודואקראית יש 2 קלטים ופלט אחד:

$$F:\underbrace{\{0,1\}^{\ell_{key}(n)}}_{\text{מפתח}}\times\underbrace{\{0,1\}^{\ell_{in}(n)}}_{\text{קלט}}\to\underbrace{\{0,1\}^{\ell_{out}(n)}}_{\text{פלט}}$$

- נניח לשם פשטות כי $n=\ell_{key}(n)=\ell_{in}(n)=\ell_{out}(n)$
- בשימוש נגריל $k\sim U_n$ ונחשב איתו את F על הרבה ערכים, כלומר נחשב $F_k(x)=F(k,x)$.
  שימו לב: המפתח k נבחר באקראי ⟵ התפלגות על הפונ' F_k
- כשדיברנו על גנרטורים פסאודואקראיים שמייצרים מחרוזת "שנראית אקראית" אמרנו שניתן את המחרוזת הזאת למבחין שאמור לפלוט 0/1. עכשיו האוביקט שאנחנו רוצים "שיראה אקראי" זה הפונקציה F_k. כלומר אנחנו נגיד ש- F היא פונק' פסאודואקראית אם כאשר $k\sim U_n$ אז הפונק' F_k לא ניתנת להבחנה בזמן פולינומי מפונק' אקראית.

איך מגדירים שיריב לא יכול להבחין בין פונק' אקראית לפונק' פסאודואקראית?

ננסה להגדיר כמו עבור PRG:

- יריב פולינומי מקבל פונק' פסאודואקראית או פונק' אקראית וצריך לדעת להבחין בין המקרים.

הבעייה עם הניסוח הזה היא שפונק' אקראית אי אפשר לתת למבחין, כי גודל התיאור של פונק' אקראית הוא $n\cdot 2^n$ ביטים.

לכן לא ניתן את תיאור הפונק' ליריב, אלא רק נאפשר לו "גישת אורקל" כלומר גישת input-output לפונקציה. בכל שלב בריצת היריב, הוא יכול לבחור קלט x ולקבל אחרי יחידת זמן אחת את ערך הפונקציה בנקודה x על ידי קריאה לפרוצדורה $O(x)$.

הערה: מכיוון שהיריב רץ בזמן $poly(n)$, לאורך הריצה הוא יכול לשאול רק מספר פולינומי של שאילתות.

נסמן ב- $D^{F_k(\cdot)}$ יריב שמקבל תשובות לפי F_k פסאודואקראית ונסמן ב- $D^{f(\cdot)}$ יריב שמקבל תשובות לפי f אקראית לגמרי.

<u>הגדרה:</u>
תהי $F:\{0,1\}^*\times\{0,1\}^*\to\{0,1\}^*$ פונק' שניתנת לחישוב בזמן פולינומי, כל שלכל n, אם $|x|=|k|=n$ אזי $|F(k,x)|=n$. הפונק' F היא פסאודואקראית אם לכל יריב פולינומי אקראי (לא אחיד) D קיימת פונק' זניחה $negl$ כך ש-

$$\left|\Pr_{\substack{k\sim U_n\\ D}}\left[D^{F_k(\cdot)}(1^n)=1\right]-\Pr_{\substack{f\\ D}}\left[D^{f(\cdot)}(1^n)=1\right]\right|\le negl(n)$$

כאשר ההסתברות השמאלית היא מעל הגרלת $k\in\{0,1\}^n$ באקראי ומעל האקראיות של D וההסתברות הימנית היא מעל הגרלת f באקראי ($n\cdot 2^n$ ביטים אקראיים) ומעל האקראיות של D.

הערות:

* היריב D מקבל רק גישה פונקציונלית ל- $F_k(\cdot)$ או ל- $f(\cdot)$. הוא לא יודע כלום על תהליך החישוב/זמן החישוב/...
* ל- D יש גישה אדאפטיבית לאורקל. הוא יכול לשאול שאלה, על סמך התשובה לשאול עוד שאלה וכו'.

<u>תרגיל לחימום:</u> האם הפונקציה הבאה היא פסאודואקראית:

$$F(k,x) = k \oplus x$$

<u>פתרון</u>: הפונקציה הזאת <u>איננה</u> פסאודואקראית.

נשים לב שאם k אקראי אזי כל פלט בודד של F_k הוא אקראי, אבל יש קשרים חזקים בין פלטים על קלטים שונים:

לכל x, y מתקיים

$$F_k(x) \oplus F_k(y) = (k \oplus x) \oplus (k \oplus y) = x \oplus y$$

לעומת זאת, עבור x, y נתונים, ההסתברות שבהגרלת f אקראית נקבל $f(x) \oplus f(y) = x \oplus y$ היא $1/2^n$.

המבחין D על קלט 1^n:
(1) אם $1^n = O(1^n) \oplus O(0^n)$ החזר 1. אחרת החזר 0.

כמו שהבנו (אינטואיטיבית לפחות) בשיעור הקודם, פונקציה פסאודואקראית זה משהו שיכול לעזור לנו לבנות מע' הצפנה עבור הרבה הודעות. אבל איך בונים פונקציה פסאודואקראית? היום נראה איך בונים כזאת פונקציה מ-PRG ובפעם הבאה נראה איך לבנות מערכת הצפנה מ- PRF.

<u>משפט:</u> קיים גנרטור פסאודואקראי אם ורק אם קיימת פונקציה פסאודואקראית.

<u>כיוון קל:</u> אם קיימת פונקציה פסאודואקראית אזי לכל פולינום $\ell(n)$ קיים גנרטור פסאודואקראי עם הרחבה $\ell(n)$.

<u>הוכחת הכיוון הקל:</u>

בה"כ $\ell(n) = n \cdot \ell'(n)$ עבור $\ell'(n)$ שלם. נגדיר גנרטור:

$$G(s) = F_s(\underbrace{0^{n-1}1}_{1}) \circ F_s(\underbrace{0^{n-2}10}_{2}) \circ \dots \circ F_s(\underbrace{\quad\ell'(n)\quad}_{\substack{\text{בייצוג בינארי} \\ \text{כמחרוזת באורך } \boldsymbol{n}}})$$

נראה ש- G גנרטור פסאודואקראי. יהי D' יריב פולינומי כלשהו עבור G. עלינו להראות ש- D' מבחין בין $G(s)$ לבין מחרוזת אקראית בהסתברות זניחה לכל היותר. לצורך כך, נבנה מבחין D עבור F:

<u>$D(1^n)$:</u>
(1) $x = O(1) \circ O(2) \circ \dots \circ O(\ell'(n))$
(2) החזר $D'(x)$

מתקיים:

* אם O פונקציה אקראית אזי x מורכב מ- n הפעלות של פונק' אקראית ולכן x היא מחרוזת אקראית באורך $\ell(n) = \ell'(n) \cdot n$.
* אם $O = F_s$ אזי $x = G(s)$

כלומר:

$$\left| \Pr_{\substack{r\sim U_{\ell(n)} \\ D'}}[D'(r)=1] - \Pr_{\substack{s\sim U_n \\ D'}}\left[D'\left(G(s)\right)=1\right] \right| =$$

$$= \left| \Pr_{\substack{f \\ D}}\left[D^{f(\cdot)}(1^n)=1\right] - \Pr_{\substack{s\sim U_n \\ D}}\left[D^{F_s(\cdot)}(1^n)=1\right] \right| \underset{\text{כי } \boldsymbol{F} \text{ היא פסאודואקראית}}{\leq} negl(n)$$

מ.ש.ל.

עכשיו אנחנו רוצים להראות את הכיוון השני שאומר שאם קיים *PRG* אז קיימת *PRF*. לפני כן נוכיח טענת עזר.

<u>טענת עזר:</u> יהי G גנרטור פסאודואקראי עם פונק' מתיחה $\ell(n)=2n$. אזי לכל פולינום $t(n)$ מתקיים ש-

$$G'(\underbrace{s_1 \ \dots \ s_{t(n)}}_{\substack{\text{מחרוזות } \boldsymbol{t(n)} \\ \text{כ"א באורך } \boldsymbol{n}}}) = G(s_1)\circ G(s_2)\circ \dots \circ G(s_{t(n)})$$

הוא גנרטור פסאודואקראי.

<u>הוכחת טענת העזר:</u>
ההוכחה היא על ידי היברידים. לכל $0 \leq i \leq t(n)$ נגדיר:

$= G_n^i$	התפלגות על מחרוזות באורך $2n \cdot t(n)$ שבה i הבלוקים הראשונים באורך $2n$ הם אקראיים ושאר הבלוקים $(t(n)-i)$ הם פסאודואקראיים.

כלומר, כדי לדגום מחרוזת מ- G_n^i אנחנו מגרילים i בלוקים באקראי $r_1, r_2, \dots, r_i \in \{0,1\}^{2n}$ כ"א באורך $2n$ ומגרילים עוד $(t(n)-i)$ מחרוזות $s_{i+1}, \dots, s_{t(n)} \in \{0,1\}^n$ כ"א באורך n ומחזירים
$r_1 \circ r_2 \circ \dots \circ r_i \circ G(s_{i+1}) \circ \dots \circ G(S_{t(n)})$

נשים לב: ב- $G_n^{t(n)}$ כל הבלוקים הם אקראיים
ב- G_n^0 כל הבלוקים הם פסאודואקראיים

יהי D' יריב ל- G' ונסמן

$$\left| \Pr_{\substack{\vec{r}\sim U_{2n\cdot t(n)} \\ D'}}[D'(\vec{r})=1] - \Pr_{\substack{\vec{s}\sim U_{(n\cdot t(n))} \\ D'}}\left[D'\left(G'(\vec{s})\right)=1\right] \right| = \epsilon(n)$$

עלינו להראות ש- $\epsilon(n)$ היא פונקציה זניחה. כמו בשיעור הקודם, לפי אי שיוויון המשולש, קיים אינדקס i_n כך ש-

$$\left| \Pr_{\substack{\vec{r}\sim G_n^{i_n} \\ D'}}[D'(\vec{r})=1] - \Pr_{\substack{\vec{r}\sim G_n^{i_n-1} \\ D'}}[D'(\vec{r})=1] \right| \geq \frac{\epsilon(n)}{t(n)}$$

נשתמש ב- D' וב- i_n כדי לבנות מבחין (לא-אחיד) D ל- G.

המבחין D: קלט $w \in \{0,1\}^{2n}$ (D מנסה להבחין האם w אקראי או $w = G(s)$)

1. הגרל $i_n - 1$ בלוקים $r_1, \dots, r_{i_n-1} \in \{0,1\}^{2n}$ אקראיים כ"א באורך $2n$ והגרל $(t(n) - i_n)$ בלוקים אקראיים $s_{i_n+1}, \dots, s_{t(n)} \in \{0,1\}^n$ כ"א באורך n.
2. הרץ $D'\left(r_1 \circ \dots \circ r_{i_n-1} \circ w \circ G\left(s_{i_n+1}\right) \circ \dots \circ G\left(S_{t(n)}\right)\right)$ וענה כמוהו.

מתקיים:

* אם w אקראי אזי מריצים את D' על דגימה מ- $G_n^{i_n}$
* אם w פסאודואקראי אזי מריצים את D' על דגימה מ- $G_n^{i_n-1}$.

לכן:

$$\left| \Pr_{\substack{w \sim U_{2n} \\ D}}[D(w) = 1] - \Pr_{\substack{s \sim U_n \\ w = G(s) \\ D}}[D(w) = 1] \right| = \left| \Pr_{\substack{\vec{r} \sim G_n^{i_n} \\ D'}}[D'(\vec{r}) = 1] - \Pr_{\substack{\vec{r} \sim G_n^{i_n-1} \\ D'}}[D'(\vec{r}) = 1] \right| \geq \frac{\epsilon(n)}{t(t)}$$

ולכן $\epsilon(n)$ היא פונקציה זניחה מכיוון ש- G הוא יצרן פסאודואקראי.

מ.ש.ל.

הערה: אנחנו הראינו כאן מבחין D לא-אחיד עבור G. אפשר "לתקן" את זה ולבנות מבחין אחיד בדומה למה שתראו בתרגיל הבית (על ידי הגרלת i_n באקראי).

עכשיו אנחנו חוזרים להוכחת הכיוון השני של המשפט שמקשר בין PRG לבין PRF ואנחנו רוצים להראות שאם יש PRG אז יש גם PRF.

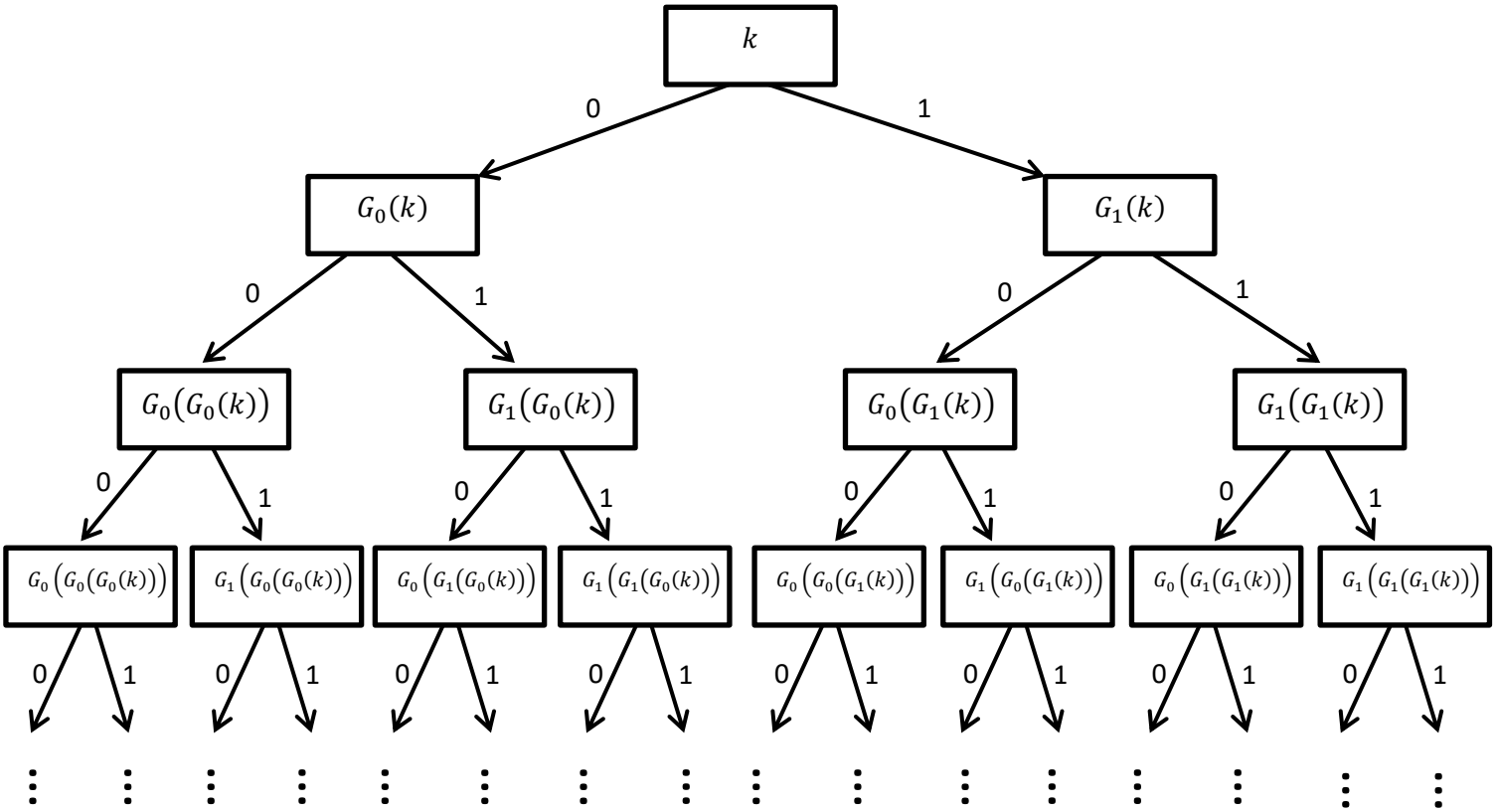
<u>כיוון שני של המשפט:</u> אם קיים גנרטור פסאודואקראי G עם פונק' מתיחה $\ell(n) = 2n$ אזי קיימת פונקציה פסאודואקראית.

<u>הוכחה:</u> עבור מחרוזת s באורך n נסמן

$$G(s) = \underbrace{G_0(s)}_{\boldsymbol{n}\text{ ביטים}} \circ \underbrace{G_1(s)}_{\boldsymbol{n}\text{ ביטים}}$$

כלומר, $G_0(s)$ הם n הביטים השמאליים ב- $G(s)$ ו- $G_1(s)$ הם n הביטים הימניים ב- $G(s)$.

מתוך G אנחנו רוצים להגדיר עכשיו פונקציה $F(k,x) = F_k(x)$. נגדיר אותה באופן הבא. עבור $k \in \{0,1\}^n$, נסתכל על העץ הבינארי הבא (בו יש $n+1$ רמות).



* כל צומת בעץ מסומן על ידי מחרוזת באורך n
* פלט $F_k(x)$ עבור $x = x_1x_2 \dots x_n$:
נלך במסלול בעץ $x_1 \to x_2 \to \cdots \to x_n$

<u>הגדרת F באופן פורמלי:</u>
קלט: $x = x_1x_2 \dots x_n$ ומפתח k
פלט:

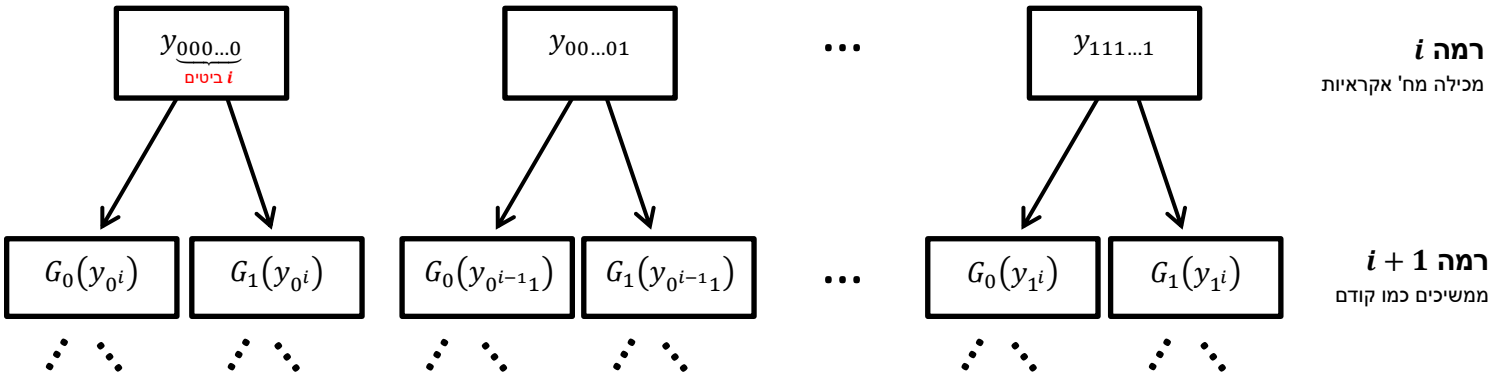
$$F_k(x_1x_2 \dots x_n) = G_{x_n}\left(\dots\left(G_{x_2}\left(G_{x_1}(k)\right)\right)\dots\right)$$

נניח שקיים יריב D^O (פולינומי לא-אחיד) כנגד הפונקציה הפסאודואקראית שבנינו. נסמן ב- $t(n)$ חסם עליון על מספר השאילתות ש- D^O מבצע. נגדיר $n+1$ היברידים (התפלגויות ביניים):

$$H_0, H_1, \dots, H_n$$

כאשר:

* H_0 היא הפונקציה הפסאודואקראית שהגדרנו
* H_n היא פונקציה אקראית
* H_i היא הפונקציה כאשר בעץ שהגדרנו נגריל לכל צומת ברמה ה- i (מלמעלה) מחרוזת אקראית ונחשב משם את העץ כמו בהגדרת הפונקציה הפסאודואקראית. בציור:



נסמן ב- $\epsilon(n)$ את ההסתברות ש- D^O מבחין בין פונקציה אקראית לפונקציה הפסאודואקראית שלנו, כלומר:

$$\left|\Pr\left[D^{f(\cdot)}(1^n)=1\right]-\Pr\left[D^{F_k(\cdot)}(1^n)=1\right]\right|=\epsilon(n)$$

עלינו להראות ש- $\epsilon(n)$ היא פונקציה זניחה. על פי אי שיוויון המשולש, קיים אינדקס i_n כך ש-

$$\left|\Pr\left[D^{H_{i_n}(\cdot)}(1^n)=1\right]-\Pr\left[D^{H_{i_n+1}(\cdot)}(1^n)=1\right]\right|\geq\frac{\epsilon(n)}{n}$$

נשתמש ב- D^O וב- i_n כדי לבנות יריב A שמבחין בין $t(n)$ מחרוזות אקראיות לבין $t(n)$ פלטים פסאודואקראיים של G (ולפי טענת העזר אין כזה מבחין...)

המבחין A: קלט: $t(n)$ מחרוזות, כ"א באורך $2n$ ביטים: $w_1,\ldots,w_{t(n)}$

(1) הרץ את $D(1^n)$. כאשר D מבצע שאילתא $x=x_1\ldots x_n$ נבצע:

(א) נסתכל על התחילית $x_1\ldots x_{i_n}$. ישנם 2 מקרים:

אם D לא שאל שאלה המתחילה ב- $x_1\ldots x_{i_n}$ אזי ניקח את המחרוזת w הבאה שלא השתמשנו בה, ונשים בצומת $x_1\ldots x_{i_n}0$ (ברמה i_n+1) את n הביטים השמאליים של w ונשים בצומת $x_1\ldots x_{i_n}1$ את n הביטים הימניים.

אם D כבר שאל שאלה המתחילה ב- $x_1\ldots x_{i_n}$ אזי הצומת $x_1\ldots x_{i_n}x_{i_n+1}$ כבר מוגדר בעץ.

(ב) נשתמש בערך שבצומת $x_1\ldots x_{i_n}x_{i_n+1}$ כדי לחשב את ערך העלה המתאים ל- $x_1\ldots x_{i_n}$ ונחזיר ערך זה ל- D.

(2) בסוף הריצה נענה כמו ש- D עונה.

ניתוח ריצת A כאשר $w_1 \dots w_{t(n)}$ אקראיים:

בכל צומת בעץ ברמה $i_n + 1$ ש- D התעניין בה שמנו מחרוזת אקראית וברמות היותר נמוכות הפעלנו את הפונקציה. לכן A הפעיל את D כך ש- D מקבל תשובות לפי H_{i_n+1}. כלומר,

$$\Pr_{\substack{w_j \\ \text{אקראיים}}} \left[A(w_1 \dots w_{t(n)}) = 1\right] = \Pr\left[D^{H_{i_n+1}(\cdot)}(1^n) = 1\right]$$

ניתוח ריצת A כאשר $w_1 \dots w_{t(n)}$ פסאודואקראיים, כלומר קיימים $s_1 \dots s_{t(n)}$ אקראיים כך ש- $w_j \leftarrow G(s_j)$ עבור $1 \le j \le t(n)$

A מריץ את D כאשר ברמה $i_n + 1$, בכל צומת ש- D התעניין בה, שמנו מחרוזת פסאודואקראית, כלומר ברמה i_n שמנו מחרוזת אקראית. כלומר D מקבל תשובות לפי H_{i_n}. כלומר

$$\Pr_{\substack{w_j \\ \text{פסאודואקראיים}}} \left[A(w_1 \dots w_{t(n)}) = 1\right] = \Pr\left[D^{H_{i_n}(\cdot)}(1^n) = 1\right]$$

כלומר קיבלנו:

$$negl(n) = \left| \Pr_{\substack{w_j \\ \text{אקראיים}}} \left[A(w_1 \dots w_{t(n)}) = 1\right] - \Pr_{\substack{w_j \\ \text{פסאודואקראיים}}} \left[A(w_1 \dots w_{t(n)}) = 1\right] \right| =$$

$$= \left| \Pr\left[D^{H_{i_n+1}(\cdot)}(1^n) = 1\right] - \Pr\left[D^{H_{i_n}(\cdot)}(1^n) = 1\right] \right| \ \ge \ \frac{\epsilon(n)}{n}$$

ולכן $\epsilon(n)$ צריך להיות זניח.

מ.ש.ל.