מודלים חישוביים
תרגול מס' 6

4 במאי 2017

נושאי התרגול:

- אוטומט מחסנית.
- למת הניפוח לשפות חסרות הקשר.
- פעולות סגור נוספות לשפות חסרות הקשר.

1 אוטומט מחסנית

לשפות רגולריות ראינו גם מודל מקבל (אס"ד, למשל) וגם מודל יוצר (ביטויים רגולריים, או דקדוקים רגולריים). לשפות ח"ה ראינו מודל יוצר (דקדוקים חסרי־הקשר). כעת נראה מודל מקבל — אוטומט מחסנית (PDA). אוטומט מחסנית הוא שישייה $M = (Q, \Sigma, \Gamma, \delta, q_0, F)$ כך ש:

- Q – קבוצה סופית של מצבים.
- Σ – א"ב הקלט.
- Γ – א"ב המחסנית.
- $\delta : Q \times \Sigma_\epsilon \times \Gamma_\epsilon \to \mathcal{P}(Q \times \Gamma_\epsilon)$ – פונקצית המעברים (שימו לב לאי־דטרמיניזם).
- $q_0 \in Q$ – מצב התחלתי.
- $F \subseteq Q$ – מצבים מקבלים.

מה המשמעות של $(q_2, c) \in \delta(q_1, a, b)$? אם אנו במצב q_1, קוראים a מהקלט ורואים b במחסנית, אנו מוציאים את b (pop), שמים במקומו c (push) ועוברים ל־ q_2. a, b או c יכולים להיות ϵ. אם $a = \epsilon$ זה אומר שהאוטומט יכול לעשות את המעבר ללא קריאה מהקלט, אם $b = \epsilon$ זה אומר שהאוטומט יכול לעשות את המעבר בלי לקרוא ולהוציא דבר מהמחסנית ואם $c = \epsilon$ זה אומר שהאוטומט לא דוחף דבר למחסנית בעת המעבר ל־ q_2.

הגדרה 1.1 נגדיר ש־ M מקבל את $w \in \Sigma^\star$ אם קיים $q' \in F$ כך ש־ $(q', t) \in \hat{\delta}(q_0, w, \epsilon)$ עבור $t \in \Gamma^\star$ כלשהו. כלומר, אם כאשר מתחילים בקריאת w, M יכול להיות במצב מקבל q' עם תוכן מחסנית t. את $L(M)$ נגדיר, כרגיל, כאוסף המילים ש־ M מקבל.

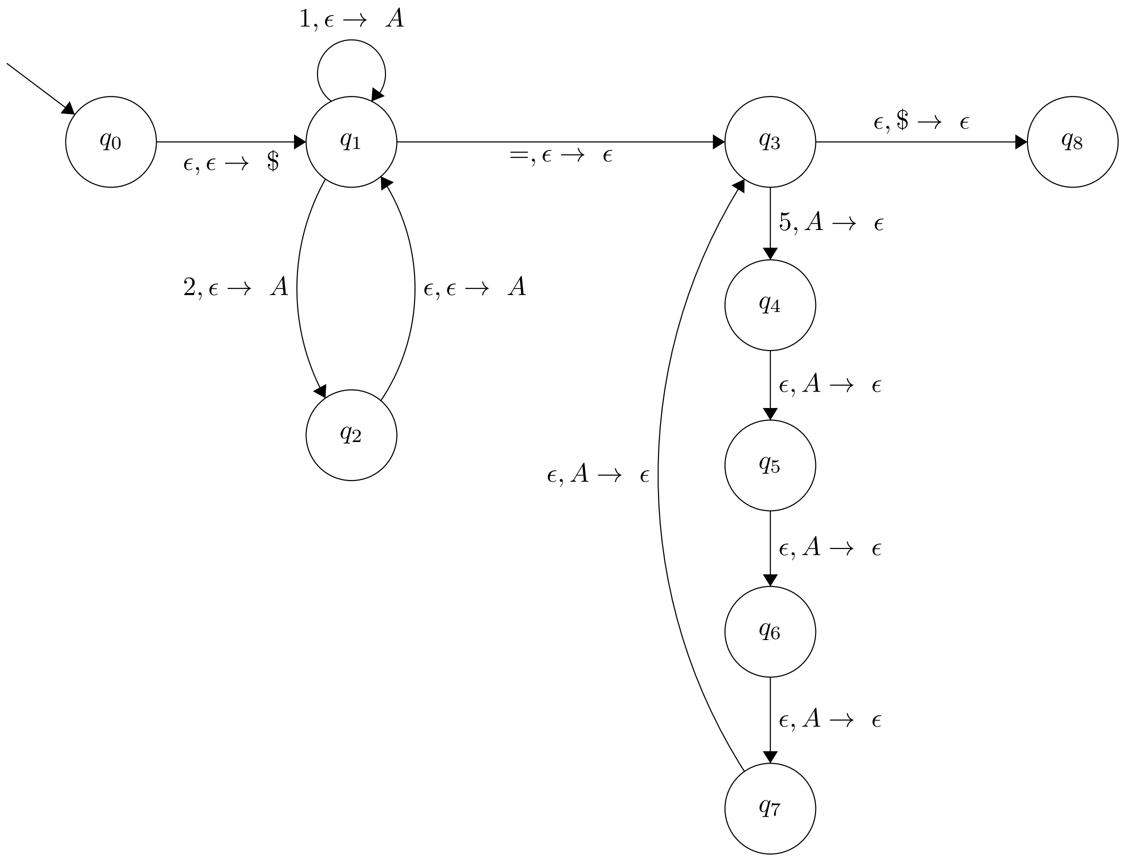
בהרצאה הוכחנו כי שפה מתקבלת ע"י אוטומט מחסנית אם"ם היא שייכת לשפה שיוצר דקדוק ח"ה. אם כך, זהו אכן אפיון נוסף לשפות ח"ה.

דוגמא 1

נבנה אוטומט מחסנית עבור השפה

$$L_2 = \left\{ x = y \mid x \in \{1,2\}^\star \wedge y \in \{5\}^\star \wedge \#_1(x) + 2 \cdot \#_2(x) = 5 \cdot \#_5(y) \right\}$$

לדוגמא, $1\,(121)^4\,21 = 5555 \in L_2$. האוטומט יהיה:



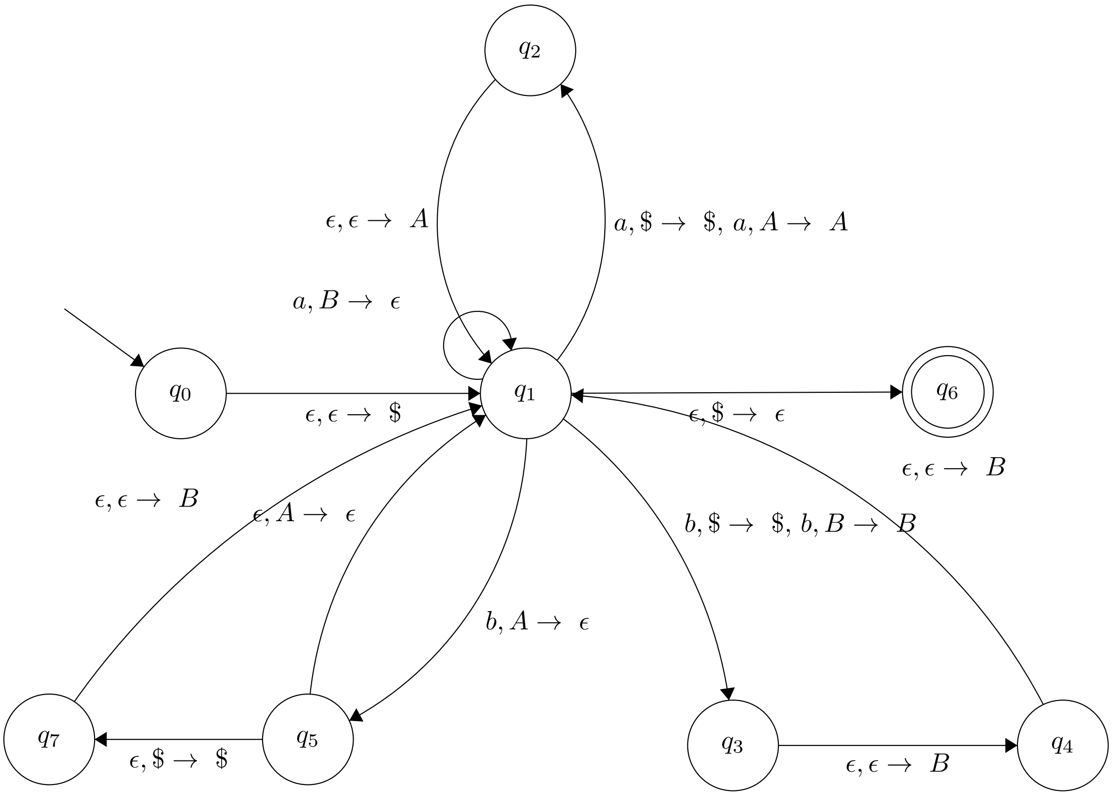
נמצא דקדוק ח"ה עבור L_2, $G = (V, \Sigma, R, S_0)$. נסמן: $V = \{S_0, S_1, S_2, S_3, S_4\}$, $\Sigma = \{1, 2, 5, =\}$ ו־R מכיל את חוקי הגזירה הבאים:

$$\begin{aligned}
S_0 &\to 1S_1 \mid 2S_2 \mid = \\
S_1 &\to 1S_2 \mid 2S_3 \\
S_2 &\to 1S_3 \mid 2S_4 \\
S_3 &\to 1S_4 \mid 2S_0 5 \\
S_4 &\to 1S_0 5 \mid 2S_1 5
\end{aligned}$$

נסו לבנות דקדוק ח"ה בעל מצב יחיד המקבל את L_2.

דוגמא 2

נבנה אוטומט מחסנית עבור השפה $L_1 = \left\{w \in \{a,b\}^\star \mid \#_a(w) = 2 \cdot \#_b(w)\right\}$:



נשים לב שאם בראש המחסנית יש לנו B, קראנו עד כה w' כך ש־ $2\cdot\#_b(w') > \#_a(w')$ ואם יש לנו A אז $2\cdot\#_b(w') < \#_a(w')$. מצאו דקדוק ח"ה ל־ L_1 לפי האלגוריתם שנלמד בכיתה.

תרגיל 1

הוכיחו כי השפות חסרות ההקשר סגורות תחת חיתוך עם שפות רגולריות.

פתרון

תהא L שפה ח"ה ויהא $P=(Q_P,\Sigma,\Gamma,\delta_P,q_P,F_P)$ האוטומט מחסנית (PDA) שמקבל אותה ותהא L' שפה רגולרית כך ש־ $A=(Q_A,\Sigma,\delta_A,q_A,F_A)$ הוא אס"ד (DFA) המקבל אותה. נבנה אוטומט מחסנית P' כך שיתקיים $L(P')=L\cap L'$.

$$P'=(Q_P\times Q_A,\Sigma,\Gamma,\delta',(q_P,q_A),F_P\times F_A)$$

נצטרך להגדיר את δ_A גם על ϵ. לשם כך, נרחיב את δ_A כך: לכל $q\in Q_A$, $\delta_A(q,\epsilon)=q$. כעת, לכל $q_1\in Q_P$, $q_2\in Q_A$, $\sigma\in\Sigma_\epsilon$ ו־ $\gamma\in\Gamma_\epsilon$:

$$\delta'((q_1,q_2),\sigma,\gamma)=\{((r,s),\gamma')\mid (r,\gamma')\in\delta_P(q_1,\sigma,\gamma)\wedge s=\delta_A(q_2,\sigma)\}$$

כדי להוכיח את הטענה, נצטרך להוכיח תחילה שתי למות, אשר הוכחתן תשאר כתרגיל (ההוכחה – באינדוקציה על אורך w):

למה 1.2 קיים מסלול חישוב של P על w שמסתיים במצב מקבל $q\in F_P$ אם"ם קיים מסלול חישוב של P' על w שמסתיים במצב (q,q') עבור $q'\in Q_A$ כלשהו.

למה 1.3 כל מסלול חישוב של A על w מסתיים במצב מקבל $q\in F_A$ אם"ם כל מסלול חישוב של P' על w מסתיים במצב (q',q) עבור $q'\in Q_P$ כלשהו.

כעת, אם $w\in L\cap L'$ אזי מהלמה הראשונה קיים מסלול חישוב של P על w שמסתיים במצב מקבל $f_P\in F_P$. מהלמה השנייה, כל מסלול חישוב של A על w מסתיים במצב מקבל $f_A\in F_A$. מכאן, שקיים מסלול חישוב של P' על w שמתסיים ב־ (f_P,f_A). (f_P,f_A) הוא מצב מקבל של P' ולכן $w\in L(P')$. הכיוון השני סימטרי.

2 למת הניפוח לשפות חסרות הקשר

נזכר בלמת הניפוח לשפות ח"ה, שגם כאן תהווה עבורנו טכניקה להוכחה כי שפה אינה ח"ה:

למה 2.1 לכל שפה ח"ה L קיים $\ell > 0$ כך שלכל $s \in L$ המקיימת $|s| \geq \ell$ קיים פירוק מהצורה $s = uvxyz$ כך ש:

1. לכל $i \geq 0$, $uv^ixy^iz \in L$.
2. $|vy| > 0$.
3. $|vxy| \leq \ell$.

איך נשתמש בלמת הניפוח כדי להוכיח ששפה L כלשהי האינה ח"ה? כמו שעשינו עבור למת הניפוח לשפות רגולריות: נניח בשלילה ש־ L ח"ה ויהא ℓ המובטח לנו. נבחר מילה $s \in L$ באורך גדול מ־ ℓ ונראה שלכל חלוקה $uvxyz$ של s כך ש־ $|vy| > 0$ ו־ $|vxy| \leq \ell$ קיים $i \geq 0$ כך ש־ $uv^ixy^iz \notin L$, בסתירה ללמת הניפוח.

תרגיל 2

הוכיחו כי השפה הבאה אינה ח"ה:

$$L_3 = \left\{a^ib^jc^k \mid j = \max(i, k)\right\}$$

פתרון

ניח בשלילה ש־ L_3 ח"ה ויהא ℓ המובטח לנו. נבחר:

$$s = a^\ell b^\ell c^\ell \in L_3$$

ונשים לב ש־ $|s| \geq \ell$. נסתכל על כל הפירוקים האפשריים של s ל־ $uvxyz$ כך ש־ $|vy| > 0$ ו־ $|vxy| \leq \ell$.
מכך ש־ $|vxy| \leq \ell$, גם v וגם y לא יכולות להכיל יותר משני סוגים שונים של תווים. נפריד למקרים:

- אם vy מכילה את האות b, נסתכל על

$$w = uv^0xy^0z = uxz$$

וקיבלנו שמספר ה־ b־ים בהכרח קטן אך או שמספר ה־ a־ים לא השתנה או שמספר ה־ c־ים לא השתנה (מהאבחנה הקודמת שלנו). מכאן, ש־ $j < \max(i, k) = \ell$ ולכן $w \notin L_3$.

- אם vy אינה מכילה את האות b אז vxy כולה a־ים או כולה c־ים. נסתכל על

$$w = uv^2xy^2z$$

מספר ה־ b־ים נשאר ℓ, אך בהכרח מספר ה־ a־ים גדלו או מספר ה־ c־ים גדלו. מכאן, ש־ $\ell = j < \max(i, k)$ ולכן $w \notin L_3$.

בכל חלוקה אפשרית קיים i כך ש־ $uv^ixy^iz \notin L_3$, בסתירה ללמת הניפוח. לכן, L_3 אינה ח"ה.

תרגיל 3

הוכיחו כי השפה הבאה אינה ח"ה:

$$L_4 = \{1^n \mid n \text{ is prime}\}$$

פתרון

ניח בשלילה ש־ L_4 ח"ה ויהא ℓ המובטח לנו. נבחר $s = 1^m$ כך ש־ m הוא הראשוני הקטן ביותר שהוא לפחות ℓ. מהגדרתינו, $s \in L_4$ ו־ $|s| \geq \ell$. נסתכל על כל הפירוקים האפשריים של s ל־ $uvxyz$ כך ש־ $|vy| > 0$ ו־ $|vxy| \leq \ell$. נסמן $v = 1^k$ ו־ $y = 1^t$. מתקיים אם כך ש־ $t + k > 0$. נבחר $i = m + 1$, ואז:

$$w = uv^ixy^iz = 1^{m+m(k+t)}$$

ומכיוון ש־ $m + m\left(k + t\right) = m\left(1 + k + t\right)$, ברור כי הוא אינו ראשוני ולכן $w \notin L_4$.

תרגיל 4

הוכיחו כי השפה $L_5 = \left\{a^{3j}b^{2^k} \mid j, k > 0\right\} \cup \{b\}^\star$ מקיימת את תנאי למת הניפוח לשפות ח"ה.

פתרון

נראה כי L_5 מקיימת את הלמה עבור קבוע הניפוח $\ell = 3$. תהא $s \in L_5$ כך ש־ $|s| \geq 3$. נפריד למקרים:

- אם $s \in \{b\}^\star$ אז $s = b^m$ עבור $m \geq 3$. נבחר פירוק $u = \epsilon$, $v = b$, $x = y = \epsilon$ ו־ $z = b^{m-1}$. נראה כי הפירוק מקיים את תנאי הלמה:
  - $|vxy| = 1 \leq \ell$.
  - $|vy| = 1 > 0$.
  - לכל $i \geq 0$, $uv^ixy^iz = b^{i+m-1} \in \{b\}^\star \subseteq L$.
- אם $s \notin \{b\}^\star$ הרי ש־ $s \in \left\{a^{3j}b^{2^k} \mid j, k > 0\right\}$. כלומר, קיימים $j, k > 0$ כך ש־ $s = a^{3j}b^{2^k}$. נבחר פירוק $u = \epsilon$, $v = a^3$, $x = y = \epsilon$ ו־ $z = a^{3(j-1)}b^{2^k}$. נראה כי הפירוק מקיים את תנאי הלמה:
  - $|vxy| = 3 \leq \ell$.
  - $|vy| = 3 > 0$.
  - לכל $i \geq 0$,

$$uv^ixy^iz = a^{3i}a^{3(j-1)}b^{2^k} = a^{3(i+j-1)}b^{2^k} \in \left\{a^{3j}b^{2^k} \mid j, k > 0\right\} \subseteq L$$

3 פעולות סגור נוספות לשפות חסרות הקשר

ראינו כי השפות חסרות ההקשר סגורות תחת איחוד, שרשור, סגור קליני, חיתוך עם שפה רגולרית, הומומורפיזם והומומורפיזם הפוך.

תרגיל 5

הוכיחו כי השפה $L_5 = \left\{a^{3j}b^{2^k} \mid j, k > 0\right\} \cup \{b\}^\star$ אינה ח"ה. שימו לב ששפה זו מקיימת את תנאי למת הניפוח לשפות ח"ה.

פתרון

נסמן $\Sigma = \{a, b\}$ ונניח בשלילה כי L_5 ח"ה. נסמן:

$$L' = L_5 \cap L\left(a \cdot \Sigma^\star\right) = \left\{a^{3j}b^{2^k} \mid j, k > 0\right\}$$

מסגירות לחיתוך עם שפות רגולריות, L' ח"ה. כעת, נגדיר $g : \Sigma \to \Sigma^\star$ כך ש־ $g\left(a\right) = \epsilon$ ו־ $g\left(b\right) = b$ ונסמן:

$$L'' = g\left(L'\right) = \left\{b^{2^k} \mid k > 0\right\}$$

מסגירות להומומורפיזם, L'' ח"ה. אבל, נוכיח כי L'' אינה ח"ה תוך שימוש בלמת הניפוח לשפות ח"ה. נניח בשלילה כי L'' ח"ה ויהא ℓ קבוע הניפוח שלה. נבחר $s = b^{2^\ell}$ ומתקיים כי $|s| \geq \ell$. יהא $s = uvxyz$ פירוק כך ש־ $|vy| > 0$ ו־ $|vxy| \leq \ell$. תהא $w = uv^2xy^2z$. אזי, מתקיים כי:

$$2^\ell < 2^\ell + |vy| = |w| \leq 2^\ell + |vxy| \leq 2^\ell + \ell < 2^\ell + 2^\ell = 2^{\ell+1}$$

אם כך, לא קיים k כך ש־ $w = b^{2^k}$ ולכן $w \notin L''$ ו־ L'' אינה ח"ה. זה בסתירה להנחה שלנו ולסגירות לשפות ח"ה ולכן L_5 עצמה אינה ח"ה.

תרגיל 6

הוכיחו כי השפות חסרות ההקשר אינן סגורות תחת

$$C\left(L_1, L_2\right) = \{xyz \mid xy \in L_1 \wedge yz \in L_2 \wedge x, y, z \neq \epsilon\}$$

פתרון

נבחר:

$$\begin{aligned} L_1 &= \{0^n\#1^n\# \mid n \geq 0\} \\ L_2 &= \{\#1^n\#2^n \mid n \geq 0\} \end{aligned}$$

מעל $\Sigma = \{0, 1, 2, \#\}$ ושימו לב כי

$$C\left(L_1, L_2\right) = \{0^n\#1^n\#2^n \mid n \geq 0\} \cup \{0^n\#1^n\#1^m\#2^m \mid n, m \geq 0\}$$

כעת, נניח בשלילה כי $C\left(L_1, L_2\right)$ ח"ה ונסמן $\Delta = \{0, 1, 2\}$. אזי, מסגירות לחיתוך רגולרי, גם

$$L' = C\left(L_1, L_2\right) \cap L\left(\Delta^\star\#\Delta^\star\#\Delta^\star\right) = \{0^n\#1^n\#2^n \mid n \geq 0\}$$

חסרת הקשר. נגדיר הומומורפיזם $h : \Sigma \to \Delta^\star$ כך ש־ $h\left(0\right) = 0$, $h\left(1\right) = 1$, $h\left(2\right) = 2$ ו־ $h\left(\#\right) = \epsilon$. נקבל שגם

$$h\left(L'\right) = \{0^n1^n2^n \mid n \geq 0\}$$

חסרת הקשר, בסתירה למה שראינו בכיתה. אם כך, $C\left(L_1, L_2\right)$ אינה חסרת הקשר.

תרגיל 7

תהא $L \subseteq \Sigma^\star$. הוכיחו כי השפות חסרות ההקשר אינן סגורות חתת

$$DropMiddle\left(L\right) = \{xy \in \Sigma^\star \mid \exists \sigma \in \Sigma . x\sigma y \in L \wedge |x| = |y|\}$$

פתרון

נבחר

$$L = \{0^n1^n42^m3^m \mid m, n \geq 0\}$$

זוהי שפת ח"ה. כדי לראות זאת, נוכל להגדיר $L_1 = \{0^n1^n \mid n \geq 0\}$, $L_2 = \{4\}$ ו־ $L_3 = \{2^n3^n \mid n \geq 0\}$, ששלושתן חסרות הקשר, ואז

$$L = L_1 \circ L_2 \circ L_3$$

גם חסרת הקשר, מסגירות לשרשור. נניח בשלילה ש־ $DropMiddle\left(L\right)$ גם ח"ה. מסגירות לחיתוך עם שפות רגולריות, גם

$$L' = DropMiddle\left(L\right) \cap \{0, 1, 2, 3\}^\star = \{0^n1^n2^n3^n \mid n \geq 0\}$$

חסרת הקשר. אך קל לראות שאם נגדיר הומומורפיזם h הממפה כל תו לעצמו ואת 3 ל־ ϵ נקבל כי גם $h\left(L'\right) = \{0^n1^n2^n \mid n \geq 0\}$ חסרת הקשר, בסתירה למה שראינו בכיתה.

תרגיל 8

הראו כי אם L רגולרית, אזי $Mirror\left(L\right)=\left\{ww^{R}\mid w\in L\right\}$ חסרת הקשר.

פתרון

תהא L שפה רגולרית ו־ $M=\left(Q,\Sigma,\delta,q_{0},F\right)$ אס"ד המקבל אותה. הרעיון: נבנה אוטומט מחסנית המדמה את M ומכניס את אותיות הקלט למחסנית. האוטומט ינחש מתי אנו עוברים לקרוא את w^{R} ויעבור לריקון מחסנית. אם המחסנית ריקה, נקבל. פורמלית, נגדיר:

$$P=\left(Q\cup\left\{s,r,q_{f}\right\},\Sigma,\Sigma\cup\left\{\$\right\},\delta',s,\left\{q_{f}\right\}\right)$$

כך ש־

- s הוא המצב ההתחלתי החדש.
- r הוא המצב שאליו P עוברת כאשר היא מנחשת שקריאת w הסיימה והחלה קריאת w^{R}.
- q_{f} הוא המצב המקבל החדש.
- פונקצית המעברים δ':
  - $\delta'\left(s,\epsilon,\epsilon\right)=\left\{\left(q_{0},\$\right)\right\}$. מתחילים בלדחוף \$ למחסנית ולעבור למצב ההתחלתי של M.
  - $\forall q\in Q,\sigma\in\Sigma.\delta'\left(q,\sigma,\epsilon\right)=\left\{\left(\delta\left(q,\sigma\right),\sigma\right)\right\}$. זהו הסימלוץ של M.
  - $\forall q\in F.\delta'\left(q,\epsilon,\epsilon\right)=\left\{\left(r,\epsilon\right)\right\}$. בהגעה למצב מקבל של M עוברים ל־ r מבלי לשנות את תוכן המחסנית.
  - $\forall\sigma\in\Sigma.\delta'\left(r,\sigma,\sigma\right)=\left\{\left(r,\epsilon\right)\right\}$. מהמצב r, אם מקבלים אות בקלט שגם נמצאת בראש המחסנית, מוציאים את האות מהמחסנית ונשארים ב־ r.
  - $\delta'\left(r,\epsilon,\$\right)=\left\{\left(q_{f},\$\right)\right\}$. אם הצלחנו לרוקן את כל w^{R}, עוברים למצב מקבל.

הוכיחו פורמלית כי $L\left(P\right)=Mirror\left(L\right)$.