Almog Jakov: https://github.com/AlmogJakov

אוטומטים 2 – דף נוסחאות

יחסי שקילות ומחלקות שקילות:

יחס שקילות: יחס דו מקומי (משווים בין 2 מילים מ Σ^*) שהוא:
- רפלקסיבי - $(w,w) \in R$ לכל $w \in A$
- סימטרי - $(x,y) \in R$ אם $(y,x) \in R$ לכל $x,y \in A$
- טרזיטיבי - אם $(x,y) \in R$ וגם $(y,z) \in R$ אז $(x,z) \in R$ לכל $x,y,z \in A$.

תכונות של יחס שקילות:
מחלקות שקילות:
חלוקה של Σ^* לקבוצות (כל קבוצה היא מחלקה):
1. זרות (אין חיתוך)
2. משלימות סה"כ ל Σ^*
3. בין כל 2 מחברי קבוצה מתקיים היחס (אין סיפא מפרידה).
4. בין כל 2 קבוצות שונות לא מתקיים היחס (יש סיפא מפרידה).

דוגמא:
- עבור: $R = \{(x,y): |x| = |y|\}$. מחלקות השקילות הן: $S_i = \{x: |x| = i\}$ לכל $0 \le i$.

יחס שקילות אינווריאנטי:
<u>אינווריאנטי משמאל</u>: אם 2 מילים מקיימות את היחס אז לכל התחלה שנשרשר לשתיהן, התוצאה תהיה 2 מילים חדשות שעדיין מקיימות את היחס.
<u>אינווריאנטי מימין</u>: אם 2 מילים מקיימות את היחס אז לכל סוף שנשרשר לשתיהן, התוצאה תהיה 2 מילים חדשות שעדיין מקיימות את היחס.

יחס שקילות המעדן יחס שקילות:
נאמר שיחס שקילות S מעדן יחס שקילות R אם:
כל זוג שמקיים את היחס S מקיים גם את היחס R.
במילים אחרות, $S \subseteq R$ (אבל שניהם חייבים להיות יחס שקילות).
ליחס S (היחס המעדן) יהיו יותר (או שווה) מחלקות שקילות מליחס R.

יחס השקילות $R_{\tilde{L}}$:
שפה היא סוג של יחס. עבור שפה L ניתן לדבר על היחס:
$R_{\tilde{L}} = \{(x,y): x \in L \leftrightarrow y \in L\}$.
תמיד ל $R_{\tilde{L}}$ יש 2 מחלקות שקילות: כל המילים שבשפה וכל המילים שלא בשפה.

<u>יחס שקילות R^1 מעדן את R^2 אם מתקיים:</u>
$(x,y) \in R^1 \Rightarrow (x,y) \in R^2$ כאשר $R^1 \subseteq R^2$
ובאופן דומה, יחס שקילות R מעדן שפה אם כל זוג מילים שמקיימות את היחס, או ששתיהן בשפה או ששתיהן לא בשפה.

יחס השקילות R_A: (יחס זה אינווריאנטי מימין ומעדן את $R_{L(A)}$)
עבור אוטומט סופי דטרמיניסטי A.
$(x,y) \in R_A \leftrightarrow [\delta(q_0,x) = \delta(q_0,y)]$
כאשר A הוא אס"ד בעל החמישייה: $A = (\mathbb{F}, \mathbb{Q}, q_0, \Sigma, \delta)$
עבור אוטומט A, יחס השקילות R_A הוא כל זוגות המילים שקריאתן מסתיימת באותו מצב (כאשר מתחילים בכל מילה ב q_0 וקוראים כל מילה בנפרד)
לא חשוב אם המצב מקבל או לא.
ליחס R_A יש מחלקות שקילות כמספר המצבים באוטומט.

יחס השקילות R_L: (יחס זה אינווריאנטי מימין ומעדן את $R_{\tilde{L}}$)
$R_L = \{(x,y) \in (\Sigma^*)^2 | \forall z \in \Sigma^*, xz \in L \Leftrightarrow yz \in L\}$
זוגות המילים (x,y) כך שלכל סייפא z מתקיים: $xz \in L \leftrightarrow yz \in L$.
* סימון: מספר מחלקות השקילות של R_L מסומן ב $index(R_L)$.
* נשים לב שישנו שינוי בין הגדרת R_L לבין יחס אינווריאנטי מימין, מאחר וביחס R_L מדובר על שייכות **לשפה** בעוד שביחס האינווריאנטי מימין מדובר על שייכות **ליחס** עצמו.

<u>**הערה:**</u>
כאשר אנו מחלקים שפה למחלקות שקילות נשים לב כי ϵ נמצאת במחלקת שקילות נפרדת.
לדוגמא, עבור השפה $L = \{0,1,00,10,101\}$
מחלקות השקילות הינן:
- $S_1 = \{\epsilon\}$
- $S_2 = \{0\}$
- $S_3 = \{1\}$
- $S_4 = \{10\}$
- $S_5 = \{00, 101\}$
- $S_6 = \{x \in \Sigma^+ | x \notin L\}$

משפט נרוד:

שפה L היא רגולרית $\leftrightarrow$ מספר מחלקות השקילות של R_L הוא סופי $\leftrightarrow$ קיים יחס R שמעדן את L ואינווריאנטי מימין כך ש$index(R) < \infty$

* לפי מסקנה ממשפט נרוד - מספר המצבים המינימאלי באס"ד המקבל את L שווה למספר מחלקות השקילות של R_L.

<u>דוגמאות לשימוש במשפט נרוד:</u>
- $L = \{0^n 1^n : n \in N\}$.
  הוכחה ש L אינה רגולרית:
  נתבונן בקבוצת המילים: $\{0^i : i \in N\}$.
  לכל $i \ne j$ המילה 0^i מופרדת ביחס R_L מהמילה 0^j ע"י הסייפא: $z = 1^i$ כי $0^i 1^i \in L$ אבל $0^j 1^i \notin L$ ולכן כל מילה בקבוצה הנ"ל היא במחלקת שקילות נפרדת ולכן ליחס R_L יש אינסוף מחלקות שקילות ולפי משפט נרוד, השפה אינה רגולרית.
- $L = \{a^n b^n | n \ge 1\}$
  נראה אינסוף מילים שכל אחת מהן במחלקת שקילות שונה
  $\mathrm{a^i b^1} \quad \forall \mathrm{i} \ge 1$
  עבור כל מילה בצורה זאת, המילה $\mathrm{b^{i-1}}$ היא סיפא מפרידה בין המילה ובין כל שאר המילים בקבוצה, שהן בצורה $\mathrm{a^k b^1}$ כאשר $\mathrm{k} \ne \mathrm{i}$, מכיוון שאם נשרשר את $\mathrm{b^{i-1}}$ ל $\mathrm{a^i b^1}$ נקבל מילה בשפה, ואם נשרשר לכל שאר המילים בקבוצה נקבל מילה שאינה בשפה כי $\mathrm{k} \ne \mathrm{i}$. מכיוון שיש אינסוף מילים שבין כל אחת ואחת יש סיפא מפרידה אז כל אחת במחלקת שקילות שונה ולכן יש אינסוף מחלקות שקילות.

משפט האפיון של R_L

1. R_L מעדן את $R_{\tilde{L}}$ (במילים אחרות R_L מעדן את השפה L).
2. R_L אינווריאנטי מימין.
3. אם R הינו יחס שקילות אינווריאנטי מימין אשר מעדן את השפה L, אז R מעדן את R_L.

(במילים אחרות קיים יחס R כלשהו שמקיים את תכונות (1), (2) אשר מעדן את R_L, כלומר, R_L יהיה עבורנו היחס ה"גס" ביותר - היחס עם מספר השקילות הקטן ביותר).

אלגוריתם Hopcroft:

(אלגוריתם לצמצום אס"ד)

1. מחק מצבים שאינם ישיגים מq_0'.
   לאחר המחיקה, סמן את האוטומט המתקבל בA.
2. אתחל יחס $R \leftarrow \emptyset$. זהו יחס סימטרי (בינארי על Q).
   2.1 עדכן $R \leftarrow \{\{p,q\} | q \in F \wedge p \notin F\}$
   2.2.1 חשב
   $S \leftarrow \{\{p,q\} \notin R | \exists a \in \Sigma: \{\delta(q,a), \delta(p,a)\} \in R\}$
   2.2.2 אם $S \ne \emptyset$ עדכן $R \leftarrow R \cup S$ וחזור ל2.2.1.
3. אם $\bar{R}$ אינו יחס שקילות על Q החזר *FAIL* (לא באמת קורה)
4. בנה *DFA* מצומצם $\tilde{A} = \{\tilde{Q}, \Sigma, \widetilde{q_0}, \tilde{\delta}, \tilde{F}\}$ בצורה הבאה:
   $\tilde{Q}$ = קבוצת מחלקות השקילות של $\bar{R}$.
   $\widetilde{q_0}$ = מחלקת השקילות של q_0 ב$\bar{R}$. נסמן ב $[q_0]_{\bar{R}}$.
   $\tilde{F} = \{\tilde{q} | \tilde{q} \cap F \ne \emptyset\}$
   $\forall \tilde{q} \in \widetilde{Q},\ \forall \alpha \in \Sigma: \tilde{\delta}(\tilde{q}, a) = [\delta(q,a)]_{\bar{R}}$

* $\bar{R}$ הוא יחס שקילות בסוף האלגוריתם וכן בסוף כל איטרציה.

דקדוקים

הדקדוק בנוי מ 4 מרכיבים:
משתני הדקדוק: V - קבוצת משתנים סופית.
א"ב: T - במקום Σ - קבוצת תווים סופית.
משתנה התחלתי S.
קבוצת כללי גזירה P - קבוצה סופית.

<u>ההירכיה של חומסקי</u>

מחלקה 1: דקדוקים רגולריים: יכולים ליצור רק שפות רגולריות.
<u>דקדוקים ליניאריים ימניים</u>: שבהם כללי הגזירה הם רק מהצורה: $A \to a, A \to aB$ כלומר משתנה גוזר אות או משתנה גוזר אות ביחד עם משתנה כאשר המשתנה הוא בצד ימין. בנוסף מותר: $S \to \varepsilon$.
<u>דקדוקים ליניאריים שמאליים</u>: שבהם כללי הגזירה הם רק מהצורה: $A \to a, A \to Ba$ כלומר משתנה גוזר אות או משתנה גוזר אות ביחד עם משתנה כאשר המשתנה הוא בצד שמאל. בנוסף מותר: $S \to \varepsilon$.

<u>הערה חשובה</u>: אסור לערבב ביחד בין הכללים. או רק ימני או רק שמאלי. אם חלק מהכללים הם ליניאריים ימניים וחלק מהכללים הם ליניאריים שמאליים אז זה לא דקדוק רגולרי. (הוא גם יכול ליצור שפה לא רגולרית)

מחלקה 2: דקדוקים חסרי הקשר (ח"ה): יוצרים שפות רגולריות וח"ה
שפה חסרת הקשר היא שפה שיודעת לספור משהו אחד כנגד משהו אחד.
לדוגמא $\{a^n b^n : n \ge 0\}$ היא שפה חסרת הקשר.
שפות חסרות הקשר לא יודעות לדוגמא לבצע ספירה מדויקת (לדוגמא כאשר החזקה היא מספר ראשוני או ריבוע שלם וכו') ואינן יודעות לספור תלות של 3 דברים (כמו $a^n b^n c^n$).

בדקדוקים חסרי הקשר כללי הגזירה הם רק מהצורה: $A \to ..$ כאשר את המילה הריקה מותר לגזור רק מ S. (פרט לזה אין מגבלה על מה שיש בצד ימין של הגזירה)
הדגש כאן הוא שבצד שמאל יש רק משתנה בודד.

מחלקה 3: דקדוקים תלויי הקשר: יכולים ליצור שפות תלויות הקשר (וכמובן את כל מה שהקודמים יכלו ליצור). בדקדוקים אלו גם מותרים כללים מהצורה: $\alpha \to \beta$ (ביטוי גוזר ביטוי). גם בצד שמאל לא חייבים לשים משתנה בודד אלא ביטוי ארוך. דוגמא: $aA \to bAa$.
המגבלה היחידה היא שצד ימין צריך להיות ביטוי ארוך או שווה באורכו לביטוי של צד שמאל. כלומר אסור לקצר את הביטוי.

מחלקה 4: דקדוקים כללים: יכולים ליצור כל שפה כריעה
אין מגבלה כלל על כללי הגזירה.

שפות חסרות הקשר (זיהוי):

שפה שאין לה זיכרון מדויק אלא יודעת לספור משהו כנגד משהו.

איך מזהים ששפה היא לא ח"ה?
1. <u>ספירה משולשת אינסופית:</u>
   כן ח"ה: $L = \{a^n b^n : n \in N\}$.
   לא ח"ה: $L = \{a^n b^n c^n : n \in N\}$. 3 תלויים אחד בשני.
   כן ח"ה: $L = \{a^n b^m c^{n+m}, n,m \in N\}$. אחד תלוי ב2, ולא 3 שתלויים אחד בשני.
2. <u>ספירה מתחלפת אינסופית:</u>
   (באמצע הספירה צריך לזכור משהו נוסף - אז זוכרים כמו במחסנית את הדבר האחרון שהכנסנו)
   לא ח"ה: $L = \{a^n b^m c^n d^m : n,m \in N\}$.
   כן ח"ה: $L = \{a^n c^n b^m d^m : n,m \in N\}$.
   כן ח"ה: $L = \{a^n b^m d^m c^n : n,m \in N\}$.
3. <u>זיכרון אינסופי של תור ולא מחסנית:</u>
   (לא טוב, בשפות ח"ה הזיכרון הוא של מחסנית) - ספירה ישרה ולא הפוכה: (אות ראשונה כנגד אות אחרונה כלפי פנימה)
   לא ח"ה: $L = \{ww : w \in \{a,b\}^*\}$.
   כן ח"ה: $L = \{ww^R : w \in \{a,b\}^*\}$.
4. <u>ספירה מדויקת אינסופית:</u>
   לא ח"ה: $L = \{a^p : p\ is\ prime\}$.
   לא ח"ה: $L = \{a^{n^2} : n \in N\}$.
   לא ח"ה: $L = \{a^n b^m c^{n \cdot m} : n,m \in N\}$.
   כן ח"ה: $L = \{a^{8n+5} : n \in N\}$. (למעשה כל עוד החזקה היא ליניארית - זה ח"ה)

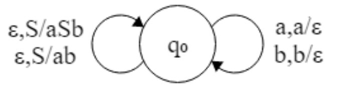
נראה אלגוריתם יצירת אוטומט מחסנית מדקדוק ח"ה והפוך.
נשים לב כי אנו עובדים עם אוטומט מחסנית המקבל ע"י ריקון אך א"מ המקבל ע"י ריקון שקול לא"מ המקבל ע"י מצב מקבל.

יצירת אוטומט מחסנית מדקדוק חסר הקשר

1. האוטומט מתחיל ע"י הכנסה של הסימן ההתחלתי S למחסנית.
2. בכל צעד מוציאים את הסימן שבראש המחסנית:
   - אם הסימן שנשלף מהמחסנית הוא משתנה, אזי מבלי להתקדם בקריאת מילת הקלט (כלומר, קוראים אפסילון) בוחרים באופן לא דטרמיניסטי את אחד מחוקי היצירה שמתאימים למחסנית (לדוגמא אם שלפנו את משתנה A ובחרנו בכלל היצירה $A \to aSbb$ נכניס למחסנית את התווים b,b,S,a בסדר הפוך כך שהתו a יהיה בראש המחסנית).
   - אם הסימן הינו טרמינל, מתקדמים בקריאת מילת הקלט ע"י שמשווים את האות שבראש המחסנית לאות הקלט הנוכחית ומוציאים אותה מהמחסנית (במילים אחרות, קיים מעבר לכל אות קלט עם האות המקבילה בראש המחסנית).

בעצם מה שקורה כאן הוא שבתוך המחסנית נוצרים לנו מילים כלשהם מתוך השפה באופן לא דטרמיניסטי.
<u>דוגמא:</u>
$L = \{a^n b^n | n \ge 1\}$, והדקדוק היוצר אותה הוא:
$G = (\{a,b\}, \{S\}, S, S \to aSb | ab)$
נבנה את האוטומט מחסנית שלה כך:
$M = (\{q_0\}, \{a,b\}, \{a,b,S\}, \delta, q_0, S, \phi)$
$\delta(q_0, \varepsilon, S) = \{(q_0, aSb), (q_0, ab)\}$
$\delta(q_0, a, a) = (q_0, \varepsilon) = \delta(q_0, b, b)$

* עבור $\delta(q_0, \varepsilon, S) = (q_0, aSb)$ הכוונה שכותבים למחסנית aSb כך שהתו a יהיה בראש המחסנית).



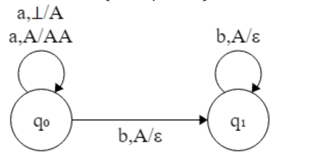
יצירת דקדוק חסר הקשר מאוטומט מחסנית

אינטואיציה: אנו מגדירים הרבה משתנים שכל משתנה מייצג מצב מסוים (המתאפיין ב3 היבטים: מצב המוצא, התו שבראש המחסנית, מצב היעד). לדוגמא: $[q, A, q_1]$. ורק עבור המקרים בהם אנו מרוקנים את המחסנית לגמרי (סעיף 3 למטה) אנו נגזור טרמינל (ואז נקבל מילה ללא משתנים).

1. בשביל להתחיל לדמות את האוטומט ע"י הדקדוק אנו נאפשר גזירה ממשתנה S לכל משתנה בו המוצא הוא q_0.
2. אם בקריאת תו כלשהו a במילת הקלט האוטומט מכניס **תת מילה** כלשהי למחסנית אז אנו נאפשר שהדקדוק יגזור את התו המסוים a בתוספת כך שכל תו בתת המילה יגזור למשתנה אחר וכך ניצור את כל קומבינציות המסלולים איתם ניתן להגיע מהמצב הנוכחי לכל מצב אחר בקריאת תת מילה זו (לכן יהיו הרבה קומבינציות כמו בדוגמא).
3. אם בקריאת תו כלשהו a ממצב מסוים p מגיעים למצב q ומרוקנים תו (לדוגמה, מוציאים מהמחסנית את התו A [כתיבת אפסילון]) אז אנו נאפשר מהמשתנה המייצג את מצב זה (כלומר ע"י השלשה: מוצא p, תו ראש המחסנית A, יעד q) גזירה של האות a.

כלומר, נבנה דקדוק $G = (Q \times \Gamma \times Q \cup \{S\}, \Sigma, P, S\}$ כך:
1. לכל $q \in Q$ נכניס את הכלל $S \to [q_0, \perp, q]$
2. לכל מעבר באוטומט $(q_1, B_1 B_2 \dots B_n) \in \delta(q, a, A)$ כאשר $a \in \Sigma \cup \{\varepsilon\}$ וגם $n > 0$ ולכל בחירת מצבים $q_2, q_3, \dots, q_{n+1}$ נכניס אל P את הכללים: $[q, A, q_{n+1}] \to a[q_1, B_1, q_2][q_2, B_2, q_3] \dots [q_n, B_n, q_{n+1}]$.
3. לכל מעבר $(q_1, \varepsilon) \in \delta(q, a, A)$ כאשר $a \in \Sigma \cup \{\varepsilon\}$ נכניס אל P את הכלל $[q, A, q_1] \to a$.

<u>דוגמא:</u>
$L = \{a^n b^n | n \ge 1\}$



האוטומט הינו:
$M = (\{q_0, q_1\}, \{a,b\}, \{A, \perp\}, \delta, q_0, \perp, \phi)$,
$\delta(q_0, a, \perp) = \{(q_0, A)\}$,
$\delta(q_0, a, A) = \{(q_0, AA)\}$,
$\delta(q_0, b, A) = \{(q_1, \varepsilon)\} = \delta(q_1, b, A)$

תחילה: נגדיר את קבוצת המשתנים (כל השלשות האפשריות):
$$V = \{S, [q_0, A, q_0], [q_0, \perp, q_0], [q_0, A, q_1], [q_0, \perp, q_1], [q_1, A, q_0], [q_1, \perp, q_0], [q_1, A, q_1], [q_1, \perp, q_1]\}$$
כעת נפעל לפי האלגוריתם.
1. נוסיף את הכללים: $S \to [q_0, \perp, q_0], S \to [q_0, \perp, q_1]$
2. עבור החץ $\delta(q_0, a, \perp) = \{(q_0, A)\}$ נוסיף:
   $[q_0, \perp, q_0] \to a[q_0, A, q_0]$
   $[q_0, \perp, q_1] \to a[q_0, A, q_1]$.
   עבור החץ $\delta(q_0, a, A) = \{(q_0, AA)\}$ נוסיף:
   $[q_0, A, q_0] \to a[q_0 A, q_0][q_0 A, q_0]$
   $[q_0, A, q_0] \to a[q_0 A, q_1][q_1 A, q_0]$
   $[q_0, A, q_1] \to a[q_0 A, q_0][q_0 A, q_1]$
   $[q_0, A, q_1] \to a[q_0 A, q_1][q_1 A, q_1]$
3. עבור $\delta(q_0, b, A) = \{(q_1, \varepsilon)\}$ נוסיף $[q_0, A, q_1] \to b$, ועבור $\delta(q_1, b, A) = \{(q_1, \varepsilon)\}$ נוסיף $[q_1, A, q_1] \to b$.

Almog Jakov: https://github.com/AlmogJakov

אוטומטים 2 – דף נוסחאות

שפה חסרת הקשר (בהרחבה)

שפה שאין לה זיכרון מדויק אלא יודעת לספור משהו כנגד משהו.

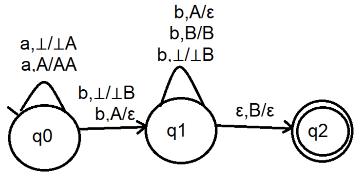
אפשר להוכיח ששפה היא חסרת הקשר ע"י:
א. דקדוק חסר הקשר.
ב. אוטומט מחסנית. (דח"ה ואוטומט מחסנית שקולים!)
ואפשר להוכיח ששפה היא לא ח"ה ע"י למת הניפוח לשפות ח"ה.

<u>דקדוקים חסרי הקשר:</u>
אלו דקדוקים מהצורה: $A \to \alpha$ כאשר α יכולה להכיל כמות (סופית) של משתנים ותווים.

אם הראינו דקדוק ח"ה עבור שפה L, זה מוכיח שהיא ח"ה.

דוגמא: דקדוק ח"ה עבור השפה: $L = \{a^n b^m c^{n+m} : n,m \in N\}$:
הדקדוק: $G = (V,T,S,P)$ כאשר:
אוסף המשתנים $V = \{S,T\}$.
אוסף התווים: $T = \{a,b,c\}$.
משתנה התחלתי: S.
כללי הגזירה: $P: S \to aSc \mid T \mid \varepsilon \;\; , \;\; T \to bTc \mid \varepsilon$

<u>אוטומט מחסנית:</u>
אוטומט אי דטרמיניסטי עם מסעי אפסילון שיש לו בנוסף למצבים גם זיכרון של מחסנית: LIFO.
פורמאלית: $M = (Q,\Sigma,\Gamma,\delta,\perp,q_0,F)$.
Q - קבוצה סופית של מצבים.
Σ - קבוצה סופית של תווי הקלט.
Γ - קבוצה סופית של תווי המחסנית (מה מותר להכניס למחסנית)
q_0 - מצב התחלתי.
$\perp$ - תו תחתית המחסנית (באתחול הוא התו שכבר נמצא בתוך המחסנית וכאשר מוציאים אותו - המחסנית נשארת ריקה)
F - קבוצת המצבים המקבלים.
$\delta: Q \times \Gamma \cup \{\varepsilon\} \to Q \times \Gamma^*$. לדוגמא: $\delta(q_0,A) = (q_1,ABC)$ זה אומר שבמקום התו A שבראש המחסנית, דוחפים אליה את התווים: ABC כאשר A הכי למטה מבניהם, C הכי למעלה.
(דוגמא נוספת לפונקציית מעבר: $\delta(q_0,A,\perp_1) = (q_1,\perp_1 A)$)
דוגמא לאוטומט מחסנית:
נבנה אוטומט מחסנית לשפה: $L = \{a^n b^m : m > n\}$.



סוגי אוטומט מחסנית:
- אוטומט מחסנית המקבל ע"י ריקון: אין מצבים מקבלים. מילה מתקבלת באוטומט רק אם כאשר מסיימים לקרוא אותה - המחסנית ריקה (גם מהתו של תחתית המחסנית שצריך להוציא אותו). (א"מ זה שקול לא"מ המקבל ע"י מצב מקבל!)
- אוטומט מחסנית המקבל ע"י מצבים מקבלים: כאן לא חייב לרוקן את המחסנית. (א"מ זה שקול לא"מ המקבל ע"י ריקון!)
- אוטומט מחסנית דטרמיניסטי: אין מעברי אפסילון ואין מספר מעברים לאותו דבר בדיוק (תו קלט+תו ראש מחסנית) האוטומט הזה חלש יותר ולא יכול לקבל את כל השפות חסרות ההקשר.

* כשניצור אוטומט מחסנית <u>חשוב</u> לציין אם מקבל ע"י ריקון או ע"י מצב מקבל.

שקילות בין אוטומט לדקדוק חסר הקשר:
רעיון ההוכחה: אם יש דקדוק אז ניתן לסמלץ אותו ע"י האוטומט.

<u>סגירויות של שפות חסרות הקשר:</u>
- אם יש לנו שפה ח"ה L_1 ושפה רגולרית L_2 אז: חיתוך, איחוד, שרשור ביניהן הוא שפה ח"ה.
- איחוד ושרשור 2 שפות ח"ה הוא שפה ח"ה.
- שפות ח"ה אינן סגורות לחיתוך ולמשלים!!!
- שפות ח"ה סגורות ל L^*, L^R.

למת הניפוח לשפות חסרות הקשר:

אם L חסרת הקשר אז:
קיים $n \in N$.
כך שלכל $z \in L$ המקיימת: $|z| \geq n$.
קיים פירוק: $z = uvwxy$ המקיים:
1. $|vwx| \leq n$.
2. $|vx| \geq 1$.
3. לכל $i \in N$ מתקיים: $z' = uv^iwx^iy \in L$.

<u>דוגמא:</u>
הוכיחו שהשפה: $L = \{a^{2^n} : n \in N\}$ אינה ח"ה.
פתרון:
הוכחה: נניח בשלילה שהשפה כן ח"ה ולכן מקיימת את למת הניפוח לשפות חסרות הקשר.
יהי $n \in N$ הקבוע המובטח מהלמה.
נבחר: $z = a^{2^n} \in L$ ומתקיים: $|z| = 2^n \geq n$.
יהי $z = uvwxy$ פירוק כלשהו המקיים: $|vx| \geq 1$, $|vwx| \leq n$.
מכאן: $vx = a^k$ כאשר: $1 \leq k \leq n$.
נבחר: $i = 2$ ונקבל: $z' = a^{2^n+k} \notin L$ כיוון ש: $2^n < 2^n + k < 2^n + 2^n = 2^{n+1}$. ולכן הוא לא חזקה שלמה של 2.
קיבלנו סתירה. ולכן השפה לא מקיימת את הלמה ולכן היא לא ח"ה.

הוכחה ששפה של הדקדוק שווה לשפה. כלומר, $L(G) = L$.

בדרך כלל אנו נחלק את ההוכחה לחלקים קטנים שבכל חלק נדבר על התת שפה שמייצר משתנה ספציפי בדקדוק.
לבסוף נוכיח מה מייצר S (המשתנה ההתחלתי).
להקלה על ההוכחה נסביר את החלקים בדקדוק:
1. משתנים סופיים. משתנים שרק גוזרים דברים סופיים. קל לדעת מהי השפה שהם מייצרים. דוגמא: $A \to a \mid abc$.
2. משתנים רקורסיביים על עצמם בלבד. דוגמא: $A \to a \mid aA$. עליהם עושים טענת אינדוקציה על קבוצת המילים שניתן לגזור מהם כאשר הבסיס הוא כללי הגזירה המיידיים (ללא רקורסיה) והצעד הוא הכללים הרקורסיביים.
3. משתנים רקורסיביים הקוראים רק למשתנים אחרים אבל האחרים לא קוראים להם.
   דוגמא: $A \to aB \mid BC$ כאשר B,C לא משתמשים ב A.
   תחילה מוכיחים את השפה של B,C ורק לאחר מכן אומרים מהי השפה של A (ללא אינדוקציה).
   עבור 2 משתנים צמודים (או יותר) או משתנה ואות - מבצעים שרשור של כל המשתנים. עבור מס' כללי גזירה מאותו משתנה זו "או" = איחוד בין השפות.
4. משתנים רקורסיביים שהעצירה שלהם זה לקרוא למשתנה אחר שלא קורא להם יותר. דוגמא: $A \to aA \mid B$.
   תחילה מוכיחים את השפה של B ואז באינדוקציה מוכיחים את השפה של A שבנויה מהשפה של B.
5. משתנים קלועים (אחד קורא לשני):
   דוגמא: $A \to aS \mid bA, S \to Ac \mid b$
   כאן צריך יותר מחשבה על מה השפה והאינדוקציה תהיה על 2 המשתנים בו זמנית.

<u>דוגמא:</u>
$$L = \{a^n c^{2k+1} d^{3k+1} b^m : n \leq m \leq 2n, k \geq 0, n \geq 1\}$$
כאשר L(G):
$$P: S \to aXbb \mid aXb \mid aSbb \mid aSb$$
$$X \to cd \mid X \to ccXddd$$

הוכחה ש: $L(G) = L$:
נוכיח את הטענה הבאה: $L(X) = \{c^{2n+1}d^{3n+1} : n \geq 0\} = L_x$.
<u>כיוון 1:</u> $L(X) \subseteq L_x$.
נוכיח באינדוקציה על אורך הגזירה.
בסיס: צעד גזירה אחד: המילה cd ואכן היא שייכת ל L_x עבור $n = 0$.
צעד: נניח שהטענה נכונה עבור כל מילה הנגזרת באמצעות n צעדי גזירה ונוכיח עבור מילים הנגזרות לאחר $n + 1$ צעדי גזירה.
תהי w מילה כזו כך ש: $X \Rightarrow^{n+1} w$. מכאן צעד הגזירה הראשון בהכרח היה: $X \to c^2Xd^3$.
כאשר מהמשתנה X השני כבר יש עוד n צעדי גזירה עד שנגיע סה"כ ל w ולכן לפי הנחת האינדוקציה: $X \to c^{2n+1}d^{3n+1}$ כאשר $n \geq 0$.
ומכאן: $w = c^2c^{2n+1}d^{3n+1}d^3 = c^{2(n+1)+1}d^{3(n+1)+1} \in L_x$.

<u>כיוון 2:</u> $L_x \subseteq L(X)$.
נוכיח באינדוקציה על אורך המילה
בסיס: המילה הקצרה ביותר ב L_x היא cd ואכן: $X \to cd$.
צעד: נניח שהטענה נכונה לכל מילה בשפה באורך עד n ונוכיח עבור מילים בשפה באורך $n + 1$.
תהי $w \in L_x$ באורך $n + 1$. מכאן: $w = c^{2m+1}d^{3m+1}$ כאשר $m \geq 1$.
נבצע את צעד הגזירה: $X \to c^2Xd^3$ ונשים לב: $w = c^2xd^3$ כאשר: $x = c^{2m-1}d^{3m-2} = c^{2(m-1)+1}d^{3(m-1)+1} \in L_x$ ולכן לפי הנחת האינדוקציה קיימת סדרת גזירה $X \Rightarrow^* x$.
ולכן: $X \to c^2Xd^3 \Rightarrow^* c^2xd^3 = w \in L(X)$.

נוכיח כעת את הטענה שכל מילה הנגזרת מ S היא מהצורה: a^nxb^m כאשר $1 \leq n \leq m \leq 2n$. $x \in L_x$.
<u>כיוון 1:</u> אם מילה נגזרת מ S אז היא מהצורה הנ"ל.
נוכיח באינדוקציה על אורך הגזירה.
בסיס: צעד גזירה אחד: $S \to aXb$ ומכאן לפי הטענה הקודמת: $X \to x \in L_x$ ולכן המילה שנגזרה מהצורה: axb כנדרש.
או $S \to aXbb$ ומכאן לפי הטענה הקודמת: $X \to x \in L_x$ ולכן המילה שנגזרה מהצורה: $axbb$ כנדרש.
צעד: נניח שהטענה נכונה לכל מילה הנגזרת מ S ב n צעדים ומגיעה לתבנית a^nXb^m כאשר $1 \leq n \leq m \leq 2n$ ונוכיח עבור $n + 1$.
מכאן: צעד ראשון היה $S \to aSb$ או $S \to aSbb$ ומכאן נותרו מ S עוד n צעדי גזירה ולפי הנחת האינדוקציה ניתן לגזור ממנה מילה מהצורה הנ"ל וביחד עם ה a הנוספת בהתחלה, ובין b אחת ל 2 b בסוף נקבל מילה מהצורה הנ"ל.

<u>כיוון 2:</u> אם מילה היא מהצורה הנ"ל אז ניתן לגזור אותה מ S.
נוכיח באינדוקציה על אורך המילה.
בסיס: המילה הקצרה ביותר מהצורה הנ"ל: axb ואכן: $S \to aXb$
צעד: נניח שהטענה נכונה לכל מילה בשפה באורך עד n ונוכיח עבור מילים בשפה באורך $n + 1$.
תהי w באורך $n + 1$. מכאן: $w = a^nxb^m$ כאשר $m \geq 1$.
(צריך לטפל במקרי קצה בהם $m = 2n$ וגם $m = n$ ולכן:)
* אם $m = 2n$ נסמן $w = a\tilde{w}bb$ ונבצע את צעד הגזירה: $S \to aSbb$
$|\tilde{w}| < n + 1$ ולכן לפי ההנחה יש סדרת גזירה $S \Rightarrow^* \tilde{w} = a^ixb^j$.
לכן, קיימת סדרת גזירה: $S \to a\tilde{w}bb \Rightarrow^* aa^ixb^jbb$.
כאשר $1 \leq i \leq j \leq 2i$. $x \in L_x$.
ובפרט $1 \leq i + 1 \leq j + 2 \leq 2i + 2 = 2(i + 1)$
* אם $m < 2n$ (ובפרט $m = n$) נסמן $w = a\tilde{w}b$ ונבצע את צעד הגזירה: $S \to aSb$. מתקיים $|\tilde{w}| < n + 1$ ולכן לפי ההנחה קיימת סדרת גזירה $S \Rightarrow^* \tilde{w} = a^ixb^j$.
לכן, קיימת סדרת גזירה: $S \to a\tilde{w}bb \Rightarrow^* aa^ixb^jb$.
כאשר $1 \leq i \leq j \leq 2i$. $x \in L_x$.
ובפרט $1 \leq i + 1 \leq j + 1 \leq 2i + 1 \leq 2(i + 1)$.

תכונות סגור:

בהינתן $L, L1, L2$ שפות חסרות הקשר ושפה רגולרית R, גם השפות הבאות הינן חסרות הקשר:
- $L \cap R$
- L^*
- L^R
- $L_1 \cup L_2$
- $L_1 \cdot L_2$

<u>דוגמאות להבחנת שפות ח"ה ע"י תכונות סגור:</u>
$$L_3 = \{a^nb^n(ab)^n(ba)^n | n \geq 1\}$$
$$L_4 = \{a^nb^na^nb^n | n \geq 1\}$$
$$L_5 = \{a^nb^n(ab)^k(ba)^k | n,k \geq 1\}$$
$$L_6 = \{a^nb^n \cup (ab)^n(ba)^n | n \geq 1\}$$
$$L_7 = \{d^*\} \cup \{a^nb^nc^n | n \geq 1\}$$
L_3 אינה חסרת הקשר (למת הניפוח).
L_4 אינה חסרת הקשר (למת הניפוח).
L_5 חסרת הקשר (שרשור שפות חסרות הקשר).
L_6 חסרת הקשר (איחוד שפות חסרות הקשר).
L_7 אינה חסרת הקשר.
הוכחת L_7: נניח בשלילה כי L_7 חסרת הקשר ותהי $L = \{a^*b^*c^*\}$
נשקול את השפה $L_\cap = L \cap L_7 = \{a^nb^nc^n | n \geq 1\}$
לפי תכונת סגור "חיתוך עם שפה רגולרית" ציפינו לקבל שפה חסרת הקשר, אך קיבלנו שפה שאינה חסרת הקשר! (הוכח באמצעות למת הניפוח) ולכן L_7 אינה חסרת הקשר.

בעיות הכרעה:

<u>בעיית השייכות:</u>
- בהינתן דקדוק G ומילה w האם $w \in L(G)$?
- עבור $w = \epsilon$ נבדוק האם S אפיס. קל.

עבור $w \neq \epsilon$ ניעזר בחומסקי – נמיר את G לצורה הנורמלית של חומסקי. אם $w \in L(G)$ אזי קיימת סדרת גזירה באורך 2|w|-1 שגוזרת אותו (בכל צעד גזירה מקבלים שני משתנים, וצריך גם $|w|$ צעדי גזירה לגזור אותיות). לכן, יש כמות סופית של סדרות גזירה אפשריות באורך זה – נוכל לעבור על כולן ולחפש האם אחת מהן גוזרת את w. אם אף אחת, $w \notin L(G)$.
- האלגוריתם הזה רץ בזמן אקספוננציאלי, לא כל כך נעים. קיים אלגוריתם CYK הרץ בזמן $O(|w|^3)$ ומופיע במודל החל מעמוד 13 של הקובץ למעוניינים.
- קיים אלגוריתם בזמן ליניארי - עבור שפות חסרות הקשר המתקבלות באוטומט מחסנית **דטרמיניסטי** בלבד.

<u>בעיית הריקנות:</u>
הוכחה: קיים אלגוריתם לסילוק משתנים שאינם טרמינליים.
S משתנה טרמינלי $\leftrightarrow L(G) \neq \emptyset$

<u>בעיית הסופיות:</u>
משפט: בהינתן דקדוק ח"ה G קיים אלגוריתם המכריע אם $L(G)$ סופית.
הוכחה:
- נניח בה"כ כי $\epsilon \notin L(G)$ וכי ב G אין כללי יחידה, כללי ϵ וסימנים מיותרים.

(אם יש כאלה, נוכל להפעיל את האלגוריתמים המסלקים אותם)
- נבנה גרף מכוון $D = (V,E)$ בו קבוצת הצמתים V היא קבוצת המשתנים של G והצלעות מוגדרות כך:
  $E \coloneqq \{(A,B) | A \to \alpha B \beta \in P, \alpha,\beta \in (V \cup T)^*\}$
  נקודה חשובה: ב G אין כללי יחידה, ולכן $|\alpha\beta| \geq 1$.
- טענה: $L(G)$ אינסופית $\leftrightarrow$ בגרף D קיים מעגל מכוון.
- כיוון א':
  נניח כי ב D יש מעגל מכוון.
  ויהי צלעותיו $A_1 \to A_2 \to \cdots \to A_n \to A_1$
  לכן, ע"פ הגדרת D נקבל כי
  $\forall i \in [1..n]: \alpha i, \beta i \in (V \cup T)^*$ כך שמתקיים:
  $\forall i: A_i \to \alpha_i A_{i+1} \beta_i \in P$ וגם $A_n \to \alpha_n A \beta_n$
  לכן נקבל גזירה:
  $A_1 \to \alpha_1 A_2 \beta_1 \Rightarrow* \alpha_1 \ldots \alpha_{n-1} A_n \beta_1 \ldots \beta_{n-1} \Rightarrow \alpha_{1\ldots n} A_1 \beta_{1\ldots n}$
  וכפי שהבחנו בשקף הקודם $\forall i: |\alpha_i\beta_i| \geq 1$.
  היות ואין בדקדוק סימנים מיותרים, קיימות מילים $v,w,x \in T^*$ כך שמתקיים:
  $$\alpha_{1\ldots n} \Rightarrow^* u, \qquad B_{1\ldots n} \Rightarrow^* x, \qquad A_1 \Rightarrow^* w$$
  בנוסף A_1 אינו מיותר, ולכן קיימות המילים $u,y \in T^*$ וקיימת סדרת הגזירה $S \Rightarrow^* uA_1y$
  ולסיכום, נוכל לרשום את סדרות הגזירה הבאות:
  $$S \Rightarrow^* uA_1y \Rightarrow^* uwy$$
  $$S \Rightarrow^* uA_1y \Rightarrow^* uvA_1xy \Rightarrow^* uvwxy$$
  ובאופן כללי $\forall i$ נוכל לרשום:
  $$S \Rightarrow^* uv^iwx^iy$$
  ולכן השפה אינסופית.

  כיוון ב':
  נניח כי $L(G)$ אינסופית, ונניח כי ב G יש n משתנים.
  מספר עצי הגזירה ב G שגובהם אינו עולה על n הינו סופי (יש כמות סופית של מילים מעצי גזירה שאורכם חסום), ולכן יש עץ גזירה שאורכו לפחות $n + 1$.
  מסלול כזה מכיל לפחות $n + 2$ סימנים, כלומר, לפחות $n + 1$ משתנים.
  לפי שובך היונים, קיים משתנה המופיע פעמיים. יהא זה A_1.
  לכן קיימים משתנים $A_2, A_3, \ldots A_t$ כך ש:
  $$A_1 \to A_2 \to A_3 \to \cdots A_t \to A_1$$
  הוא מסלול בעץ הגזירה.
  ולכן, מהגדרת D נקבל שזהו מעגל בD.

בעיות לא כריעות:

(בעיות שהוכח כי לא ניתן להכריע אותן)
- בעיית השקילות
- בעיית ההכלה
- בעיית החיתוך הריק
- בעיית הרגולריות

Almog Jakov: https://github.com/AlmogJakov

אוטומטים 2 – דף נוסחאות

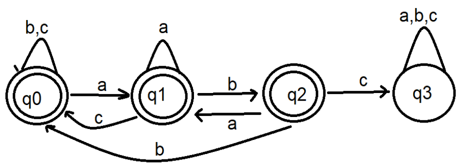
בניית דקדוק רגולרי (ליניארי ימני) לשפה רגולרית

נתונה שפה רגולרית:
- בונים לה אוטומט אס"ד.
- הא"ב של הדקדוק יהיה הא"ב של האוטומט.
- המשתנה ההתחלתי יהיה המצב ההתחלתי.
- קבוצת המשתנים תהיה קבוצת המצבים.
- המעברים יהיו באופן הבא:

אם $\delta(p,a)=q$ אז הכלל שנוסיף יהיה $p\to aq$.
אם בנוסף, $q\in F$ אז נוסיף גם את הכלל: $p\to a$.
לבסוף, אם $q_0\in F$ אז נוסיף גם את הכלל: $q_0\to\varepsilon$.

דוגמא: בנו דקדוק רגולרי (ליניארי ימני) לשפה:
$L=\{w\in\{a,b,c\}^*: abc \text{ is not in } w\}$.

שלב 1: אוטומט עבור השפה:



שלב 2: הדקדוק יהיה:

$G=(V=\{q_0,q_1,q_2,q_3\},T=\{a,b,c\},q_0,P)$
$P: q_0\to aq_1 \mid a \mid bq_0 \mid b \mid cq_0 \mid c \mid \varepsilon$
$q_1\to aq_1 \mid a \mid bq_2 \mid b \mid cq_0 \mid c$
$q_2\to aq_1 \mid a \mid bq_0 \mid b \mid cq_3$

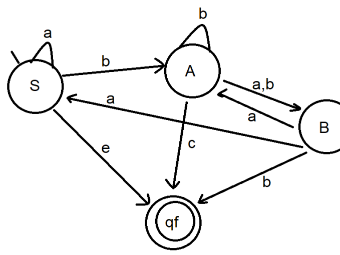
מעבר מדקדוק ליניארי ימני לאוטומט

- יוצרים אוטומט אי דטרמיניסטי.
- הא"ב של האוטומט יהיה הא"ב של הדקדוק.
- המצבים יהיו המשתנים.
- המצב ההתחלתי יהיה המשתנה ההתחלתי.
- המעברים יהיו באופן הבא:

אם $A\to aB$ אז: $\delta(A,a)=B$.
אם $A\to a$ אז $\delta(A,a)=q_f$.
אם $S\to\varepsilon$ אז: $\delta(S,\varepsilon)=q_f$.
- המצב המקבל יהיה רק q_f.

דוגמא: מצאו אוטומט עבור הדקדוק:
$P: S\to aS \mid bA \mid \varepsilon$
$A\to bA \mid aB \mid bB \mid c$
$B\to aA \mid aS \mid b$

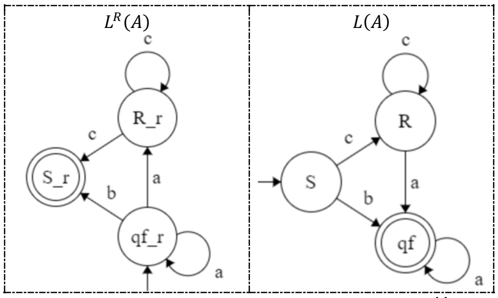
האוטומט המתאים יהיה:



מעבר מדקדוק ליניארי שמאלי לימני

הרעיון: לקחת את שפת הדקדוק, L, לבנות לה את השפה ההפוכה (reverse) - את L^R.
ולשפה ההפוכה לבנות דקדוק ליניארי שמאלי ואז להפוך את הכללים מ: $A\to Ba$ ל $A\to aB$.

דוגמא: (שמאלי לימני)
$S\to Rc|q_f b|b,\quad R\to Rc|q_f a|a,\quad q_f\to q_f a|a$
תחילה נהפוך את כללי הדקדוק:
$S\to cR|bq_f|b,\quad R\to cR|aq_f|a,\quad q_f\to aq_f|a$
כעת נבנה אוטומט רגיל ובעזרתו נבנה את אוטומט הרוורס:
$L(A)=\{x|x\in\{ba^*\}\vee x\in\{c^+a^+\}\}$
$L^R(A)=\{x|x\in\{a^*b\}\vee x\in\{a^+c^+\}\}$



אזי, כללי הדקדוק הם:
$G=(V=\{qf_r,R_r,S_r\},T=\{a,b,c\},P,S=qf_r)$
$P:=qf_r\to aqf_r|aR_r|bS_r,\quad R_r\to cR_r|cS_r,\quad S_r\to\epsilon$
בכדי לקבל דקדוק ימני נצטרך להוריד את S_r אשר גוזר אפסילון.
כלומר, הדקדוק הימיני השקול הינו:
$G=(V=\{qf_r,R_r\},T=\{a,b,c\},P,S=qf_r)$
$P:=qf_r\to aqf_r|aR_r|b,\quad R_r\to cR_r|c$

מעבר בין דקדוק ליניארי ימני לשמאלי

אותו רעיון רק שלשפה ההפוכה בונים דקדוק ליניארי ימני ואז הופכים את הכללים באותו אופן.

הוכחת שקילות אוטומטי מחסנית: (אוטומט ריקון ואוטומט מקבל)

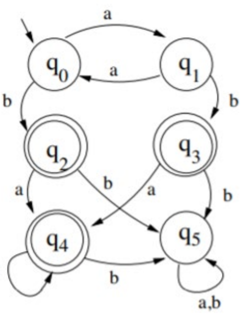
מאוטומט ריקון לאוטומט מצב מקבל: לאחר שהמחסנית התרוקנה נוצא מסע אפסילון למצב יחיד מקבל.
מאוטומט מצב מקבל לאוטומט ריקון: נהפוך מצבים מקבלים ללא מקבלים ונוציא מהם חץ למצב שירוקן את תכולת המחסנית.

אינטואיציית ההוכחה ללמת הניפוח לשפות חסרות הקשר:

חומסקי טען שכל דקדוק ח"ה ניתן להמיר אותו לדקדוק שכל הכללים בו הם מהצורה: $A\to a$ או $A\to BC$ (משתנה גוזר אות או 2 משתנים)
הצורה הזו מייצגת לנו עץ בינארי כי כל משתנה גוזר בדיוק 2 בנים (2 משתנים) או שהוא גוזר אות בודדת ואז זה בן אחד שהוא עלה.
כעת, אם גובה העץ הוא n אז יש מסלול שיש בו $n+1$ משתנים (שורש=משתנה התחלתי + n משתנים נוספים שעל המסלול).
בנוסף, לעץ בגובה n יש 2^n עלים (ולכן המילה שניתן ליצור בעץ כזה היא לכל היותר באורך 2^n)
אם ניקח מילה שהיא אכן באורך לפחות 2^n. אז גובה העץ יהיה לפחות n ואז יהיה בו מסלול עם לפחות $n+1$ משתנים על המסלול.
אבל יש רק n משתנים שונים ולכן מעיקרון שובך היונים - יהיה משתנה שיחזור על עצמו לפחות פעמיים.
כעת נוצר סוג של מעגל - יצאנו ממשתנה A והגענו שוב אליו. ולכן ניתן לנפח ולהגיע אל A שוב ושוב ובינתיים לקבל מילים נוספות בדקדוק.

אלגוריתם לצמצום אוטומטים:

(הסבר ע"י דוגמא)



שלב ראשון – נמחק מצבים בלתי ניתנים להשגה מq_0: אין.

שלב שני ושלישי - ניצור טבלה של כל זוגות המצבים ולכל $q\in F,p\notin F$ נסמן שהזוג $\{q,p\}$ מופרד ב: ϵ:

	q_0	q_1	q_2	q_3	q_4	q_5
q_0	-		ϵ	ϵ	ϵ	
q_1	-	-	ϵ	ϵ	ϵ	
q_2	-	-	-			ϵ
q_3	-	-	-	-		ϵ
q_4	-	-	-	-	-	ϵ
q_5	-	-	-	-	-	-

שלב רביעי – לכל זוג שלא מסומן בטבלה כמופרד נבדוק האם קיים סיפא מפרידה כך:
לכל שני מצבים q_i,q_j שאינם מלאים בטבלה נעבור על כל אות $\alpha\in\Sigma$ בסיגמא וכל פעם נסתכל באוטומט לאן אנו מגיעים לאחר קריאת אות זו. אם הגענו לשני מצבים שמופרדים בטבלה ע"י אות $\beta\in\Sigma$ כלשהי אזי שני המצבים מופרדים ע"י $\alpha\beta$. אחרת, אם הגענו למצבים שאינם מופרדים אז ממשיכים לנסות הלאה בטבלה על שאר התאים הריקים.
לדוגמא: בקריאת האות b מהמצבים q_0,q_5 אנו מגיעים למצבים q_2,q_5 שמופרדים כבר ע"י ϵ ולכן נאמר שהמצבים q_0,q_5 מופרדים ע"י $b\epsilon=b$.

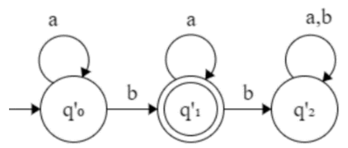
נקבל:

	q_0	q_1	q_2	q_3	q_4	q_5
q_0	-		ϵ	ϵ	ϵ	b
q_1	-	-	ϵ	ϵ	ϵ	b
q_2	-	-	-			ϵ
q_3	-	-	-	-		ϵ
q_4	-	-	-	-	-	ϵ
q_5	-	-	-	-	-	-

שלב חמישי – נגדיר את האוטומט ע"פ הטבלה:
עבור כל קבוצת מצבים אשר אין בניהם מילה מפרידה נגדיר קבוצה חדשה התייצג מצב באוטומט החדש.
כעת, נגדיר את פונקציית המעברים בעזרת אוטומט המקור.

$A=(Q',\Sigma',q_0',\delta',F')$
$Q'=\{q_0'=\{q_0,q_1\},q_1'=\{q_2,q_3,q_4\},q_2'=\{q_5\}\}$
$F'=\{q_1'\}$
$\delta'(q_0',a)=q_0',\delta'(q_0',b)=q_1'$
$\delta'(q_1',a)=q_1',\delta'(q_1',b)=q_2'$
$\delta'(q_2',a)=q_2',\delta'(q_2',b)=q_2'$

כלומר, קיבלנו אוטומט מצומצם:



שיטת הפישוט:

1. סילוק משתנים שאינם טרמינליים

אתחול: מגדירים קבוצה Var של המשתנים שהם טרמינליים. מאותחלת לקבוצה ריקה.
בכל איטרציה, מכניסים לקבוצה את כל המשתנים של הדקדוק (שאינם ב Var) שמגיעים לביטויים המכילים רק תווים או משתנים שכבר נמצאים ב Var.
דוגמא:

איטרציה ראשונה: $Var=\{S\}$. איטרציה שנייה: $Var=\{S,D\}$ איטרציה שלישית: עוצרים	$P: S\to AB \mid a \mid BB \mid ABC \mid SaS \mid b$ $A\to AC \mid AA$ $B\to aaB$ $C\to A \mid C \mid DAD$ $D\to SD \mid S \mid E$

2. סילוק משתנים לא ישיגים

אתחול: מגדירים קבוצה של הישגים: $V'=\{S\},T=\Phi$
(V' מאותחלת ל S והתווים הישגים מאותחלים לקבוצה ריקה).
בכל אטרציה מוסיפים את המשתנים שניתן להגיע אליהם מתוך משתני V' ואת התווים שניתן להגיע אליהם מ V'.
דוגמא:

אתחול: $V'=\{S\},T=\Phi$. איטרציה ראשונה: $V'=\{S,A,B\}$, $T=\{a,b\}$ איטרציה שנייה: עוצרים	$P: S\to AB \mid a \mid BB \mid ABA \mid SaS \mid b$ $A\to AS \mid AA$ $B\to aaB$ $C\to A \mid C \mid DAD$ $D\to SD \mid S \mid E$

3. סילוק כללי אפסילון

אתחול האלגוריתם הראשי:
$N=$ כל המשתנים האפסים (משתנים שניתן לגזור מהם את ε).
תת אלגוריתם למציאת משתנים אפיסים:
אתחול: $N=$ כל המשתנים שיש להם את הכלל: $A\to\varepsilon$.
כעת, בכל איטרציה מוסיפים ל N את המשתנים שגוזרים רק משתנים שכבר היו ב N (ללא תוספות).
חזרה לאלגוריתם הראשי:
מוחקים את כל הכללים מהצורה $A\to\varepsilon$.
עבור כל כלל: $A\to\alpha$. לכל משתנה אפיס בתוך α שכפל לכלל המכיל אותו וכלל שאינו מכיל אותו.
דוגמא:

א. מציאת משתנים אפיסים אתחול: $N=\{S,B,C\}$ איטרציה ראשונה: $N=\{S,B,C,D\}$ איטרציה שנייה: עוצרים	$P: S\to AB \mid a \mid BB \mid ABC \mid SaS \mid b \mid \varepsilon$ $A\to AS \mid AA \mid a$ $B\to aaB \mid \varepsilon$ $C\to A \mid C \mid DAD \mid \varepsilon \mid SB$ $D\to SD \mid S \mid E \mid aa$
ב. לאחר מחיקת כללי ε הערה: S הוא משתנה אפיס ולכן נוסיף אותו בשלב הסופי!	$P: S\to AB \mid a \mid BB \mid ABC \mid SaS \mid b$ $A\to AS \mid AA \mid a$ $B\to aaB$ $C\to A \mid C \mid DAD \mid SB$ $D\to SD \mid S \mid E \mid aa$

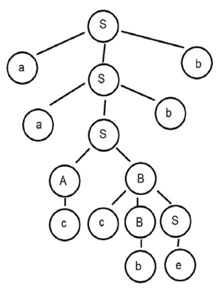
ג. לאחר תיקון כל כלל (והוספת S הגוזר אפסילון [כי S אפיס]):
$P: S\to AB \mid A \mid a \mid BB \mid B \mid ABC \mid AC \mid SaS \mid aS \mid Sa \mid b \mid \varepsilon$
$A\to AS \mid A \mid AA \mid a$
$B\to aa \mid aaB$
$C\to A \mid C \mid DAD \mid AD \mid DA \mid SB \mid B \mid S$
$D\to SD \mid D \mid S \mid E \mid aa$

4. סילוק כללי יחידה (כללים מהצורה: $A\to B$). **(ירד מהחומר)**

5. (נשתמש שוב באלגוריתם של סילוק משתנים לא ישיגים).

עץ גזירה:

עץ גזירה: עבור דקדוק G ומילה w עץ שהשורש שלו הוא S, כל קודקוד פנימי בעץ הוא מייצג משתנה מתוך V וכל עלה מייצג את המילה הריקה או תו מתוך T.
אם $S\to\alpha_1\alpha_2\dots\alpha_n$ אז $\alpha_1,\alpha_2,\dots,\alpha_n$ הם הבנים של S בעץ הגזירה.
דוגמא: הדקדוק:
$S\to aSb \mid AB \mid \varepsilon,\quad A\to cA \mid A \mid c,\quad B\to cBS \mid b$
נתבונן בסדרת הגזירה הבאה: $S\to aSb\to aaSbb\to aaABbb\to aacBbb\to aaccBSbb\to aaccbSbb\to aaccbbb$
עץ הגזירה עבורה יראה כך:



* חזית העץ: אוסף כל העלים משמאל לימין (מייצג את המילה הסופית שנגזרה).
* אם אנו עוצרים באמצע הגזירה ולא ממשיכים עד הסוף אז הביטוי בחזית העץ יכול לכלול משתנים.
* תבנית פסוקית - ביטוי שיש בו גם משתנים וגם אותיות.
* עץ גזירה מלא סופי יוביל רק למילים סופיות (ללא משתנים).
* לכל מילה בדקדוק יש עץ גזירה.

אוטומטים 2 – דף נוסחאות

Almog Jakov: https://github.com/AlmogJakov

בעיות הכרעה (אינטואיטיבי ובפחות פורמליות)

- בעיית הריקנות (האם השפה שמקבל המודל ריקה או לא?)

בעיית הריקנות בדקדוק ח"ה:
- נתון לנו דקדוק ח"ה וצריך לומר true או false האם שפת הדקדוק ריקה

פתרון: שפה של דקדוק היא ריקה אם ורק אם S אינו טרמינאלי.
דוגמא לדקדוק שהשפה שלו ריקה:

$$S \to SA, \qquad A \to a$$

האלגוריתם:
- מצא את המשתנים הטרמינאליים.
- אם S לא נמצא ביניהם - החזר $true$.
- אחרת החזר $false$.

מציאת משתנים טרמינאליים:
אתחול: V': כל המשתנים שגוזרים טרמינלים (אותיות ללא משתנים)
בכל צעד מוסיפים לקבוצה V' את המשתנים שגוזרים רק ביטויים המכילים טרמינאלים ומשתנים מ V'.
דוגמא:

$$S \to aA \mid ABC, \qquad A \to bbB \mid cCc$$
$$B \to aa, \qquad C \to aB \mid B \mid a$$

אז:
אתחול: $V' = \{B, C\}$
לאחר צעד 1: $V' = \{A, B, C\}$
לאחר צעד 2: $V' = \{S, A, B, C\}$

בעיית הריקנות באוטומט מחסנית
- נתון לנו אוטומט מחסנית. האם השפה שהוא מקבל היא ריקה?

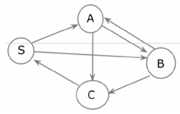
האלגוריתם:
- נמיר את האוטומט למודל המקבל ע"י מצבים מקבלים (ולא ע"י ריקון)
- צריך לבדוק האם יש מסלול שמתחיל במצב ההתחלתי ומגיע למצב מקבל.
- עבור כל מצב, צריך לשמור את כל התווים האפשריים שיכולים להיות בראש המחסנית.
- עבור כל יציאה מהמצב, נבדוק האם התו שמצוין בראש המחסנית יכול להיות אחד מהתווים ששמרנו שיכולים להיות בראש המחסנית.

- בעיית הסופיות (האם השפה שמקבל המודל סופית או לא?)

בעיית הסופיות בדקדוק ח"ה:
האלגוריתם:
- הסר מהדקדוק משתנים שאינם טרמינאליים ושאינם ישיגים.
- בנה גרף מכוון שהקודקודים בו הם המשתנים (שנשארו) ויהיה צלע מ A ל B אם ורק אם B נגזר בין היתר מתוך A. (אפילו אם הוא נגזר בתוך ביטוי גדול)
- בדוק האם יש מעגל החל מ S בגרף שהתקבל.
- אם כן - השפה אינסופית. אם לא - השפה סופית.

דוגמא:

$$S \to AAa \mid B \mid bab$$
$$A \to C \mid Baa$$
$$B \to CCC \mid aabA$$
$$C \to aSa \mid a$$



- בעיית השייכות (האם מילה x שייכת לשפה של המודל)

בעיית השייכות בדקדוק ח"ה:
אלגוריתם:
- הפוך את הדקדוק לצורה הנורמאלית של חומסקי: כל כלל יהיה מהצורה: $A \to XY$ או $A \to a$.
- עבור המילה x שרוצים לבדוק עבורה האם היא בשפה של הדקדוק או לא. נסמן: $n = |x|$.
- נעבור על כל האפשרויות לגזור מילה באורך n:
  - מכיוון שהדקדוק בצורה הנורמאלית של חומסקי, זה מאפשר בדיוק $n-1$ צעדי גזירה של משתנים + n צעדי גזירה של אותיות.

דוגמא:

$$S \to AB \mid BB \mid a, \qquad A \to b \mid SA$$
$$B \to CC \mid b \mid SS, \qquad C \to AS \mid BA$$

נבדוק האם aba שייכת לדקדוק.
אפשרות 1: $S \to AB \to SAB \to abb$
אפשרות 2: $S \to AB \to ACC \to b - -$
אפשרות 3: $S \to AB \to ASS \to baa$
אפשרות 4: $S \to BB \to CCB \to - - b$
אפשרות 5: $S \to BB \to SSB \to aab$
אפשרות 6: $S \to BB \to BCC \to b - -$
אפשרות 7: $S \to BB \to BSS \to baa$
לכן aba לא שייכת לדקדוק.

דו משמעות:

הגדרה: דקדוק הוא דו משמעי אם ניתן לייצר ממנו שני עצי גזירה שונים בעלי חזית זהה.
באופן שקול, במושגים של סדרות גזירה:
דקדוק הוא דו משמעי אם ניתן לגזור ממנו מילה בשני אופנים שונים, עד כדי סדר גזירת משתנים.
כלומר, שתי הסדרות גזיר ה הבאות **לא** מעידות על דו משמעות:

$$S \Rightarrow AB \Rightarrow aB \Rightarrow ab, \qquad S \Rightarrow AB \Rightarrow Ab \Rightarrow ab$$

ואילו שתי הסדרות גזירה הבאות **כן** מעידות על דו משמעות:

$$S \Rightarrow AB \Rightarrow aB \Rightarrow ab, \qquad S \Rightarrow C \Rightarrow ab$$

הגדרה:
שפה היא דו משמעית אם **כל** דקדוק שיוצר אותה הוא דו משמעי.
(אנחנו לא לומדים להוכיח ששפה ח"ה היא דו משמעית לכן סביר להניח שהתשובה היא "לא").

תרגיל: קבעו האם הדקדוק הבא הוא דו משמעי או לא:

$$S \to aSb \mid AB \mid \varepsilon, \qquad A \to cA \mid A \mid c, \qquad B \to cBS \mid b$$

תשובה: הכלל $A \to A$ מייצר דו משמעות כיוון שבכל סדרת גזירה בה A מופיע ניתן להרחיב אותה ע"י הוספת מעבר נוסף של $A \to A$.
דוגמא למילה: cb.
דרך א: $S \to AB \to cB \to cb$.
דרך ב: $S \to AB \to AB \to cB \to cb$.

תרגיל: האם הדקדוק הנ"ל הוא חד משמעי?

$$S \to aSb \mid AB \mid \varepsilon, \qquad A \to cA \mid c, \qquad B \to cBS \mid b$$

פתרון:
דוגמא למילה: ccb.
דרך א: $S \to AB \to cB \to ccBS \to ccbS \to ccb$.
דרך ב: $S \to AB \to cAB \to ccB \to ccb$.

חומסקי

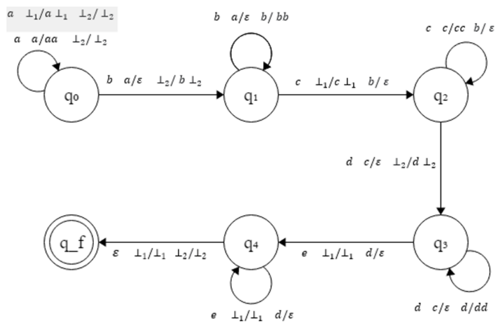
חומסקי טען שכל דקדוק ח" ה ניתן להמיר אותו לדקדוק שכל הכללים בו הם מהצורה $A \to a$ או $A \to BC$.

א"מ עם 2 מחסניות

$$L = \{a^n b^n c^n d^n e^n | n \geq 1\}$$

הסבר אינטואיטיבי:
אנו קוראים בתחילה כמות של a-ים למחסנית הראשונה ולאחר מכן בהופעת b-ים אנו מרוקנים את הa-ים ובמקביל ממלאים את המחסנית השנייה בb-ים על מנת "לשמור" את מספר ההופעות של הa-ים וכך לוודא שישנה כמות זהה של b-ים וכך נעשה באופן זהה עבור קריאת c-ים, נרוקן את הb-ים בהופעת c-ים ובמקביל נמלא את המחסנית הראשונה בc-ים וכך נמשיך בהצלבה. נשים לב כי ע"פ הגדרת פונקציית המעברים סדר שונה של הופעת האותיות ו/או אי שוויון של כמות הופעת האותיות יגרור אי קבלת המילה באוטומט הנ"ל.

$$\delta(q_0, a, \bot_1, \bot_2) \vdash (q_0, a\bot_1, \bot_2), \quad \delta(q_0, a, a, \bot_2) \vdash (q_0, aa, \bot_2)$$
$$\delta(q_0, b, a, \bot_2) \vdash (q_1, \varepsilon, b\bot_2), \quad \delta(q_1, b, a, b) \vdash (q_1, \varepsilon, bb)$$
$$\delta(q_1, c, \bot_1, b) \vdash (q_2, c\bot_1, \varepsilon), \quad \delta(q_2, c, c, b) \vdash (q_2, cc, \varepsilon)$$
$$\delta(q_2, d, c, \bot_2) \vdash (q_3, \varepsilon, d\bot_2), \quad \delta(q_3, d, c, d) \vdash (q_3, \varepsilon, dd)$$
$$\delta(q_3, e, \bot_1, d) \vdash (q_4, \bot_1, \varepsilon), \quad \delta(q_4, e, \bot_1, d) \vdash (q_4, \bot_1, \varepsilon)$$
$$\delta(q_4, \varepsilon, \bot_1, \bot_2) \vdash (q_f, \bot_1, \bot_2)$$



שאלות ממבחנים

שאלה ממבחן 2019 מועד ב: (אוטומט מחסנית)

$$L_3 = \{a^n b^m c^k | n \leq 2m, and\ m - n + k \equiv 0\ mod\ 10\}$$

השפה כן חסרת הקשר.
נבנה אוטומט מחסנית לשפה. x מסמן "חצי תו".

$$\delta(q_i, a, \bot) = (q_{(i+1)(mod\ 10)}, \bot x),$$
$$\delta(q_i, a, x) = (q_{(i+1)(mod\ 10)}, a),$$
$$\boldsymbol{\delta(q_i, a, a) = (q_{(i+1)(mod\ 10)}, ax)}$$
$$\delta(q_i, \varepsilon, *) = (p_i, *)$$
$$\delta(p_i, b, \bot) = (p_{(i-1)(mod\ 10)}, \bot),$$
$$\delta(p_i, b, x) = (p_{(i-1)(mod\ 10)}, \varepsilon),$$
$$\boldsymbol{\delta(p_i, b, a) = (p_{(i-1)(mod\ 10)}, \varepsilon)}$$
$$\delta(p_i, \varepsilon, \bot) = (r_i, \bot)$$
$$\delta(r_i, c, \bot) = (r_{(i-1)(mod\ 10)}, \bot)$$

$$Q = \{q_0, q_1, \dots, q_9, p_0, p_1, \dots, p_9, r_0, r_1, \dots, r_9\}$$
$$\Sigma = \{a, b, c\}$$
$$\Gamma = \{\bot, a, x\}$$
$$F = \{r_0\}$$

שאלה ממבחן 2019 מועד א: (משפט נרוד)

$$L_3 = \{a^n b^m | n, m \geq 0, \exists k \geq 0 (m - n = k^2)\}$$

פתרון:
נתבונן בקבוצה a^i לכל $i \in N$. ונראה שכל מילים בקבוצה ניתנות להפרדה.
יהיו $i > j$ ומכאן: $x = a^i$ ניתנת להפרדה מ $y = a^j$ ע"י הסיפא: $z = b^{i^2+i}$. ומכאן: $xz = a^i b^{i^2+i} \in L$ אבל $yz = a^j b^{i^2+i} \notin L$ כי: $i^2 + i - j$ אינו ריבוע שלם.

שאלה ממבחן 2019 מועד א: (למת הניפוח)

$$L_1 = \{ww^R ww^R | w \in \{0,1\}^*\}$$

פתרון: השפה אינה ח"ה.
נניח שכן ולכן היא מקיימת את למת הניפוח לח"ה.
יהי $n \in N$ הקבוע המובטח מהלמה.
נבחר: $z = 0^n 110^n 0^n 110^n \in L_1$.
יהי $z = uvwxy$ פירוק כלשהו המקיים: $|vx| \geq 1, |vwx| \leq n$.
נבחר $i = 0$ ונחלק למקרים:
1. אם vx מכיל את אחד ה 1 אז הוא לא מכיל את כולם לפי (2) אזי לאחר הניפוח יהיה מספר אחדות שלא מתחלק ב 4 ולכן לא ייתכן שהמילה בשפה כי כל תו משוכפל בדיוק 4 פעמים.
2. אם vx מכיל רק אפסים מאיזשהו רצף אז לאחר הניפוח, יהיו רצפים עם n אפסים ויהיה רצף אחד עם פחות מ n אפסים.

שאלה ממבחן 2019 מועד א:

שקילות א"מ לא"מ עם מחסנית שיכולה להכיל רק את התווים $\{\bot, a\}$
פתרון:
א. נראה שבהינתן אוטומט שבו ה Γ הוא סופי כלשהו. ניתן לבנות אוטומט שקול בו $\Gamma = \{\bot, a\}$.
נצמיד לכל תו ב Γ אינדקס i כאשר האינדקס של $\bot$ הוא 0, ושל כל אחד מהאחרים הוא אינדקס עולה מ 1 עד $|\Gamma| - 1$.
כעת, לכל מעבר שבו מכניסים את התו ה i למחסנית, נכניס במקומו את $a^i \bot$.
לכל מעבר בו קוראים את התו ה i נפצל למסלול מעברים שקוראים את כמות ה a עד שפוגשים ב $\bot$, מחזירים את a^i בחזרה וממשיכים למצב היעד.

תרגיל 2019 מועד א: (הפרכה)

יהיו L_1, L_2 שפות חסרות הקשר האם השפה הבאה חסרת הקשר:

$$Zip_{L_1,L_2} = \{x_1 y_1 \dots x_n y_n | x \in L_1, y \in L_2\ S.T\ x_i, y_i\ letters\}$$

פתרון:
הפרכה: ניקח את $L_1 = \{w: \#_w(a) = \#_w(b)\}, L_2 = \{w: \#_w(c) = \#_w(d)\}$.
ומכאן: Zip_{L_1,L_2} - כל המילים בהן כמות הa זהה לכמות הb וגם כמות הc זהה לכמות הd כאשר כל תו זוגי הוא c או d וכל תו אי זוגי הוא a או b.
וזו שפה לא ח"ה.

תרגיל מגעיל ממבחן תשע"ז ב מועד ב:

עבור כל אחד מהמודלים בסעיפים א' ו ב' הוכח או הפרך שהוא מאפשר לממש בדיוק את שפות חסרות ההקשר.
א. אוטומט מחסנית הזהה לאוטומט מחסנית המקבל באמצעות ריקון, אלא שטבלת המעברים היא כזו, שבכל צעד, מותר לכתוב תו אחד או 2 תווים למחסנית (במקום מחרוזת באורך כלשהו).
ב. אוטומט מחסנית הזהה לאוטומט מחסנית המקבל באמצעות ריקון, אלא שבכל צעד נכתבת מחרוזת באורך גדול מ0 למחסנית.
ג. נתון המודל הבא של אוטומט מחסנית:
אוטומט מחסנית דטרמיניסטי, הזהה לאוטומט מחסנית המקבל ע"י ריקון, אלא שעומק המחסנית הוא לכל היותר $|x|^{0.5}$ ומותר לכתוב למחסנית רק את התו a. כאן, $|x|$ הוא אורך הקלט שנקרא עד כה. אם מנסים לכתוב למחסנית מעבר לגבול שלה, ה"עודף" הולך לאיבוד והראש מצביע לתא בראש המחסנית.
הראו שמחסנית כזו לא יכולה לקבל את השפה $L = \{a^n b^n | n \geq 0\}$.
פתרונות:
א. המודלים שקולים.
המודל החדש הוא מקרה פרטי בתוך המודל הרגיל שבו פשוט מכניסים רק תו אחד למחסנית לכל היותר בכל מעבר.
בכיוון ה 2: אם יש מודל רגיל אז ניתן לבנות אוטומט מחסנית שקול מהמודל החדש באופן הבא: עבור כל מעבר שבו מכניסים מילה באורך לפחות 2 למחסנית, נפרק למספר מצבים חדשים עם מעברי אפסילון ביניהם כך שבכל מעבר נכניס תו אחד מתוך המחרוזת שצריכה להכינס. לבסוף נעבור למצב שהיינו צריכים להגיע אליו.

ב. המודלים שקולים.
המודל החדש הוא מקרה פרטי בתוך המודל הרגיל שבו פשוט מכניסים לפחות תו אחד למחסנית לכל היותר בכל מעבר של הכנסה.
בכיוון ה 2: אם יש מודל רגיל אז ניתן לבנות אוטומט מחסנית שקול מהמודל החדש באופן הבא: עבור כל מעבר בו לא מכניסים כלום למחסנית, ניצור מצב חדש שנעבור אליו ונוציא תו אחד מהמחסנית ואז נעבור למצב שאליו רצינו להגיע ע"י הכנסת התו מחדש.

ג. נניח בשלילה שקיים אוטומט מחסנית מהמודל החדש המקבל את השפה. נסמן ב n את מספר המצבים באוטומט. נתבונן במילים: $a^0, a^1, a^2, \dots, a^m$. עבור: $m = n^4$.
נשים לב שגודל המחסנית עבור קריאת המילים הללו הוא לכל היותר: $\sqrt{m}$. ולכן עבור המילים הללו, לאוטומט הנ"ל יש לכל היותר
$n \cdot (\sqrt{m} + 1) = n^3 + n < n^4$ קונפיגורציות שונות.
לכן לפי עיקרון שובך היונים, קיימות 2 מילים מהרשימה לעיל שסיום קריאתן יגיע לאותה קונפיגורציה.
נניח ש a^i, a^j הגיעו לאותה קונפיגורציה. מכאן, אם נמשיך לקרוא b^i נגיע ב 2 המילים לאותו מצב אבל אז $a^i b^i$ בשפה ולכן זהו מצב מקבל (=המחסנית ריקה) ומצד שני $a^j b^i$ לא בשפה ולכן זהו מצב לא מקבל. סתירה!
בגדול, היות והשפה דורשת זיכרון באורך ,n ואין לה, הא"מ לא יקבל את השפה הזו כי הוא לא יכול להשוות בין האורכים.

תרגיל ממבחן 2020:

$$L_2 = \{a^n b^m c^k | 0 < 2n < m < 3n\ or\ 0 < 5n < m < 7n\}$$

כאשר $m, n, k \geq 0$.
ציינו ערך n עבורו מתקיימת מסקנת למת הניפוח לשפות חסרות הקשר.
הוכיחו שאכן קיים פירוק מתאים לכל מילה בשפה שאורכה לפחות n.
פתרון:
נבחר: $n = 12$.
תהי $z \in L_2$ מילה כלשהי באורך לפחות n.
מכאן: $z = a^t b^m c^k$ כאשר: $2t \leq m \leq 3t$ או $5t \leq m \leq 7t$.
נחלק למקרים:
אם $k > 0$ אז נבחר את הפירוק הבא:

$$u = a^t b^m, v = c, w = \varepsilon, x = \varepsilon, y = c^{k-1}$$

ואכן: $1 \geq 1 = |vx| n \leq 1 = |vwx|$. ולכל $i \in N$ מתקיים: $a^t b^m c^{k-1+i} \in L_2$.
אם $k = 0$ וגם $2t \leq m \leq 3t$ כאשר $t \geq 2$ ואז $m \geq 4$, אז נבחר את הפירוק הבא:
$u = a^{t-1}, v = a, w = \varepsilon, x = bb, y = b^{m-2}$.
ואכן: $1 \geq 3 = |vx| n \leq 3 = |vwx|$ .. ולכל $i \in N$ מתקיים: $uv^i wx^i y =$
$a^{t-1} a^i b^{2i} b^{m-2} \in L_2$
אם $k = 0$ וגם $5t \leq m \leq 7t$ כאשר $t \geq 2$ ואז $m \geq 10$, אז נבחר את הפירוק הבא:
$u = a^{t-1}, v = a, w = \varepsilon, x = bbbbb, y = b^{m-5}$.
ואכן: $1 \geq 6 = |vx| n \leq 6 = |vwx|$ .. ולכל $i \in N$ מתקיים: $uv^i wx^i y =$
$a^{t-1} a^i b^{5i} b^{m-5} \in L_2$.

שאלה ממבחן 2020:

א. הוכח או הפרך: לכל שפה רגולרית L קיים אס"ד A המקבל אותה שמס' מחלקות השקילות של R_A הוא לפחות $index(R_L) + 10$.
ב. הוכח: לכל שפה רגולרית L קיים אס"ד A המקבל אותה שמספר מחלקות השקילות של R_A הוא בדיוק $index(R_L) + 1$.
ד. הוכח: תהי L שפה כך שקיימות זוג מילים $x \neq y \in \Sigma^*$ כך שמתקיים $(x, y) \in R_L$ אזי קיים יחס שקילות אינווריאנטי מימין המעדן את $\tilde{R_L}$ ואינו שווה לR_L.
פתרונות:
א. הוכחה: תהי L רגולרית ויהי A_L אס"ד המקבל אותה עם מספר מצבים מינימאלי.
כלומר: מספר המצבים ב A_L הוא $index(R_L)$. בהכרח יש ב A_L מעגל ולכן ניתן להרחיב את המעגל פי 11 מגודלו.
ג. הוכחה: באותו אופן כמו סעיף א', נמצא מעגל באוטומט, נסתכל על המצב האחרון במעגל (החל מסריקה המתחילה ב q_0) ונחבר אותו לעוד מצב שייתנהג בדיוק כמו המצב הראשון של המעגל ויחובר ישירות למצב השני של המעגל.
קיבלנו אוטומט עם בדיוק $index(R_L) + 1$ מצבים.
ד. הוכחה: נגדיר את היחס: $R = R_L \backslash \{(a, b): a = x\ xor\ b = x\}$.
תחילה נשים לב ש: $R \subseteq R_L$ ולכן R מעדן את R_L וגם R_L מעדן את $\sim R_L$ ולכן גם R מעדן את $\sim R_L$.
נשים לב ש R יחס שקילות כי $(a, a) \in R$ כי $(a, a) \in R_L$ ולא הורדנו אותו.
יהיו $(a, b) \in R$ אזי בפרט: $a \neq x, b \neq x$ ולכן $(b, a) \in R$ כי R_L הוא יחס שקילות.
יהיו $(a, b), (b, c) \in R$ אזי בפרט: $a = b = c = x$ ואז הוא טרנזיטיבי. או ש $a \neq x, b \neq x, c \neq x$ ואז $(a, b) \in R$ כי הוא היה גם ב R_L כי R_L הוא יחס שקילות ולכן טרנזיטיבי.
היחס אינווריאנטי מימין כי אם $(a, b) \in R$ אז לכל סיפא z מתקיים: $(az, bz) \in R$.

שאלה ממבחן:

א. הוכח או הפרך: עבור $L \subseteq \{a\}^*$ אם השפה מקיימת את למת הניפוח לשפות ח"ה אז היא מקיימת גם את למת הניפוח לשפות רגולריות.
ב. הוכח או הפרך: לכל א"ב סופי, אם שפה מקיימת את למת הניפוח לשפות ח"ה אז היא מקיימת את למת הניפוח לשפות רגולריות.
פתרון:
א. הוכחה: נניח ש L מקיימת את למת הניפוח לח"ה. צ"ל: L מקיימת את למת הניפוח לרגולריות. נבחר את n להיות מה שמובטח לנו מהלמה של ח"ה.
תהי $z \in L$ מילה באורך $|z| \geq n$. ולכן לפי הלמה לח"ה, קיים פירוק: $z = uvwxy$ המקיים: $|vx| \geq 1$, $|vwx| \leq n$ ולכל $i \in N$ מתקיים: $z' = uv^i wx^i y \in L$. נבחר את הפירוק: $u' = w, v' = vx, w' = uy$. מכיוון שכל התווים הם a ניתן להחליף בסדר ולבחור את הפירוק הנ"ל. ואכן מתקיים: $|u'v'| = |vwx| \leq n, |v'| = |vx| \geq 1$ ומתקיים: $z' = u'v'^i w' = uv^i wx^i y \in L$ לכל $i \in N$.
ב. הפרכה: דוגמא נגדית: $L = \{a^n b^n : n \in N\}$.
היא לא מקיימת את למת הניפוח לרגולריות (כי היא לא רגולרית).
אבל היא כן מקיימת את למת הניפוח לח"ה: נבחר: $n = 2$ ואכן: תהי $z \in L$ כך ש: $|z| \geq 2$ ולכן $z = a^k b^k$ כאשר $k \geq 1$. נבחר את הפירוק:
$u = a^{k-1}, v = a, w = \varepsilon, x = b, y = b^{k-1}$.
מתקיים: $1 \geq 2 = |vx|, 2 \leq 2 = |vwx|$ ולכל $i \in N$ מתקיים:
$z' = uv^i wx^i y = a^{k-1+i} b^{k-1+i} \in L$.