מודלים חישוביים – פתרון תרגיל 6

1. כתוב דקדוק חסר הקשר לשפות הבאות.

א. $L_1 = \{ a^i b^j c^k \mid i,j,k \geq 0,\ i+k=j \}$

נגדיר דקדוק חסר הקשר $G=(\{S,A,C\},\{a,b,c\},S,P)$
כאשר P הם כללי הדקדוק הבאים:

$S \to AC$

$A \longrightarrow Aab \mid \varepsilon$

$C \longrightarrow bcC \mid \varepsilon$

ב. $L_2 = \{ a^i b^j \mid i,j \geq 0,\ i \neq j \}$

נגדיר דקדוק חסר הקשר $G=(\{S,A,C\},\{a,b\},S,P)$
כאשר P הם כללי הדקדוק הבאים:

$S \to aSb \mid aA \mid Bb$

$A \to aA \mid \varepsilon$

$B \to Bb \mid \varepsilon$

ג. $L_3 = \{ a^i b^j c^k \mid i,j,k \geq 0,\ i \neq j \text{ or } j \neq k \}$

נגדיר דקדוק חסר הקשר $G=(\{S,I,K,A,A',B,B',C,C'\},\{a,b,c\},S,P)$
כאשר P הם כללי הדקדוק הבאים:

$S \to I \mid K$

$I \longrightarrow aAC \mid BbC$

$A \longrightarrow aA \mid aAb \mid \varepsilon$

$B \longrightarrow Bb \mid aBb \mid \varepsilon$

$C \longrightarrow cC \mid \varepsilon$

$K \longrightarrow A'bB' \mid A'C'c$

$A' \longrightarrow A'a \mid \varepsilon$

$B' \longrightarrow bB' \mid bB'c \mid \varepsilon$

$C' \longrightarrow C'c \mid bC'c \mid \varepsilon$

2. עבור השפות הבאות, קבע והוכח אם הן חסרות הקשר.

א. $L_1 = \{ a^n b^{2n} c^n \mid n \in \mathbb{N} \}$

פתרון: זו <u>אינה</u> שפה חסרת הקשר.

הוכחה:

נניח בשלילה שהשפה חסרת הקשר ו-p הוא קבוע הניפוח שלה. נבחר $w = a^p b^{2p} c^p$. נשים לב שמתקיים $|w|>p$. עבור כל חלוקה של w=uxvyw כך ש-$|xvy|\leq p$ מתקיים $xvy=a^j$, $xvy=b^j$, $xvy=c^j$, $xvy=a^t b^k$ או $xvy=b^t c^k$. מכאן

u.multinet.co.il

מתקבל כי בכל בחירה של xvy השומרת על הכללים, לא מופיעה לפחות אות אחת. בפרט מתקיים כי x ו-y מכילים לכל היותר 2 אותיות שונות יחד.
כיוון ש-$|xy|\geq 1$, מתקיים כי לפחות אחת מהמחרוזות x או y חייבת להכיל אות אחת לפחות. לכן, אם נבחר, למשל, i=2 נקבל כי מספר ההופעות של לפחות אות אחת לא ישתנה ומספר ההופעות של אות אחת לפחות כן ישתנה. כלומר, נקבל כי $ux^2vy^2w\notin L_1$ ומכאן שלמת הניפוח לא מתקיימת. לכן, זו <u>אינה</u> שפה חסרת הקשר.

ב. $L_2 = \{ a^ib^jc^k \mid i,j,k \geq 0,\ k=\min(i,j) \}$

פתרון: זו <u>אינה</u> שפה חסרת הקשר.

הוכחה:

נניח בשלילה שהשפה חסרת הקשר ו-p הוא קבוע הניפוח שלה. נבחר $w=a^pb^pc^p$. נשים לב שמתקיים $|w|>p$ וכמובן $w\in L_2$. עבור כל חלוקה של $w=uxvyw$ כך ש-$|xvy|\leq p$ מתקיים $xvy=a^j$, $xvy=b^j$, $xvy=c^j$, $xvy=a^tb^k$ או $xvy=b^tc^k$. נתייחס לכל מקרה בנפרד:

i. $xvy=a^j$ (ו-$xvy=b^j$) - במקרה זה, לפחות אחת מהמחרוזות x או y מכילה מופע אחד לפחות של a (b). נבחר, i=0 ונקבל מילה בה מספר ה-a (ה-b) קטן ממש מ-p, ומספר ההופעות של b (במקרה השני - a) ו-c שווה ל-p. לכן, $ux^0vy^0w\notin L_2$.

ii. $xvy=c^j$ - במקרה זה, לפחות אחת מהמחרוזות x או y מכילה מופע אחד לפחות של c. נבחר, i=0 ונקבל מילה בה מספר ה-c קטן ממש מ-p, ומספר ההופעות של a ו-b שווה ל-p. לכן, $ux^0vy^0w\notin L_2$.

iii. $xvy=a^tb^k$ – במקרה כזה, לפחות אחת מהמחרוזות x או y מכילה מופע אחד או יותר של a או b (כיוון ש-$|xy|>0$). נבחר i=0 ונקבל מילה בה מספר ה-a ו/או ה-b קטן אך מספר ההופעות c נשאר זהה. לכן, $ux^0vy^0w\notin L_2$.

iv. $xvy=b^tc^k$ – במקרה כזה, חייב להתקיים אחד משני המצבים הבאים: המחרוזת xy מכילה את מופע של c. במקרה כזה נבחר i=2. מספר ההופעות של c גדול עתה ממש מ-p, מספר ההופעות של b לא השתנה או גדל ובכל מקרה, מספר ההופעות של a נשאר p ולכן $ux^2vy^2w\notin L_2$. אם המחרוזת xy לא מכילה מופע של c, נבחר i=0. במקרה כזה, מספר ההופעות של c נשאר p אך מספר ההופעות של b קטן ממש מ-p (כיוון ש-$|xy|>0$). לכן, $ux^0vy^0w\notin L_2$.

מכאן שלמת הניפוח לא מתקיימת ולכן, זו <u>אינה</u> שפה חסרת הקשר.

ג. $L_3 = \{ a^mb^nc^nd^m \mid n,m \in \mathbb{N} \}$

פתרון: זוהי שפה חסרת הקשר.

הוכחה:

נראה דקדוק חסר הקשר המכריע את השפה:

G_3: $S \to aSd \mid T \mid \varepsilon$

$T \to bTc \mid \varepsilon$

u.multinet.co.il

הוכחת נכונות:

$L_3 \subseteq L(G_3)$ תהי $w \in L_3$ אזי $w=a^m b^n c^n d^m$ עבור n,m כלשהם. נייצר את המילה באמצעות הדקדוק שכתבנו:

i. נפעיל את הכלל $S \to aSd$, m פעמים ונקבל $a^m S d^m$.

ii. נפעיל את הכלל $S \to T$ פעם אחת ונקבל $a^m T d^m$.

iii. נפעיל את הכלל $T \to bTc$, n פעמים ונקבל $a^m b^n T c^n d^m$.

iv. נפעיל את הכלל $T \to \varepsilon$ פעם אחת וקבל $a^m b^n c^n d^m$.

$L(G_3) \subseteq L_3$ יהי $w \in L(G_3)$. נעקוב אחר בניית w.

ע"פ דקדוק השפה, לאחר כל מעבר מסוג $S \to X$ המילה תהיה אחת מהצורות: ε, $a^i d^i$, $a^i S d^i$ או $a^i T d^i$. עבור $a^i d^i$ כללי השפה נשמרים (נבחר m=i, n=0). מקרה בו קיבלנו ε זה מקרה פרטי של $a^i d^i$ ולכן, כללי השפה נשמרים (נבחר n=m=0). אם קיבלנו $a^i S d^i$ אנו למעשה נשארים באותו השלב (זו אינה מילה סופית). לכן, נותר להראות שגם עבור $a^i T d^i$ נקבל בהכרח מילה השייכת לשפה.

לאחר כל מעבר מסוג $T \to X$ המילה תהיה אחת מהצורות: ε, $b^j c^j$, $a^i b^j c^j d^i$, $a^i b^j T c^j d^i$ או $b^j T c^j$. עבור $b^j c^j$ כללי השפה נשמרים (נבחר m=0, n=j). כך גם עבור $a^i b^j c^j d^i$ (נבחר m=i, n=j). במקרה וקיבלנו ε כללי השפה גם כן נשמרים (נבחר n=m=0) ואם קיבלנו $a^i b^j T c^j d^i$ או $b^j T c^j$ אנו למעשה נשארים באותו השלב (זו אינה מילה סופית).
לכן, עבור כל רצף של בחירות, w שייכת לשפה L_3 ומכאן $L(G_3) \subseteq L_3$.

הראנו כי $L_3 \subseteq L(G_3)$ וכן כי $L(G_3) \subseteq L_3$ ומכאן שבהכרח מתקיים $L(G_3) = L_3$ כנדרש.

3. לכל אחת מהפעולות הבאות, קבע והוכח האם שפות רגולריות סגורות עליה והאם שפות חסרות הקשר סגורות עליה.

א. $Reverse(L) = \{ w \mid w^R \in L \}$

פתרון:

i. שפות רגולריות סגורות על פעולה זו.

הוכחה:
נציג אוטומט המקבל את L^R. השפה L היא שפה רגולרית. לכן, קיים אוטומט סופי דטרמיניסטי $A=(Q,\Sigma,\delta,q_0,F)$ המקבל אותה. נשתמש ב-A על מנת להגדיר אוטומט אי-דטרמיניסטי A^R המקבל את L^R.

$A^R=(Q\cup\{q_s\},\Sigma, \delta',q_s,F')$ כאשר:

1. q_s הוא מצב חדש המחובר באמצעות אפסילון לכל אחד ממצבי הסיום (יוגדר ב-'δ).

2. $F'=\{q_0\}$

3. $$\delta'(q,a) = \begin{cases} \{ q' \mid \delta(q',a) = q\} & q \in Q \wedge a \in \Sigma \\ \{ q' \mid q' \in F\} & q = q_s \wedge a = \varepsilon \end{cases}$$
("הפכנו" את החיצים באוטומט המקורי והוספנו מעברי ε מהמצב החדש q_s אל כל אחד ממצבי הסיום המקוריים).

u.multinet.co.il

הוכחת נכונות:
⇐ נניח כי $w\in L(A)$ ונראה כי $w^R\in L(A^R)$. כיוון ש-w שייך ל-L(A) קיים מסלול $q_0\to q_1\to\dots\to q_n$ באוטומט A המקבל את המילה. ע"פ הגדרת 'δ, באוטומט A^R קיים מסלול הפוך $q_n\to q_{n-1}\to\dots\to q_0$ (עבור אותיות w מהסוף להתחלה). כמו כן, כיוון ש-q_n הוא מצב סיום ב-A, קיים מעבר אפסילון $q_s\to q_n$. בנוסף, q_0 הוגדר להיות מצב מקבל ב-A^R. לכן, אם $w\in L(A)$ אזי $w^R\in L(A^R)$.

⇒ נניח כי $w^R\in L(A^R)$ ונראה כי $w\in L(A)$. כיוון ש-w^R שייך ל-$L(A^R)$ קיים מסלול $q_s\to_\varepsilon q_n\to\dots\to q_0$ באוטומט A^R המקבל את המילה (q_0 הוא המצב המקבל היחיד). ע"פ הגדרת 'δ, באוטומט A קיים מסלול הפוך ללא המעבר הראשון $q_s\to_\varepsilon q_n$, כלומר קיים המסלול $q_n\to q_{n-1}\to\dots\to q_0$ (עבור אותיות w מההתחלה לסוף). כמו כן, כיוון ש-q_n הוא מצב המחובר חיבור אפסילון למצב q_s ב-A^R, באוטומט A הוא מוגדר להיות מצב מקבל. לסיום, q_0 הוא אכן מצב ההתחלה ב-A. לכן, אם $w^R\in L(A^R)$ אזי $w\in L(A)$.

ii. שפות חסרות הקשר סגורות על פעולה זו.

הוכחה:
נציג דקדוק חסר הקשר המקבל את L^R. השפה L היא שפה חסרת הקשר. לכן, קיים דקדוק G=(V,Σ,R,S) המקבל אותה. נשתמש ב-G על מנת להגדיר דקדוק המקבל את L^R: $G^R=(V,\Sigma,R',S)$ כאשר
$R' = \{ A\to\alpha \mid A \in V,\ \alpha\in(V\cup\Sigma)^*,\ (A\to\alpha^R)\in R \}$.

הוכחת נכונות:
⇐ נניח כי $w\in L(G)$ ונראה כי $w^R\in L(G^R)$. כיוון ש-$w\in L(G)$ קיים רצף גזירות ב-G המקבל את המילה. בדקדוק G^R, סדר הסמלים (אותיות ומשתנים) הוא הפוך. לכן, בעת הפעלת אותם פעולות גזירה, בכל שלב נקבל היפוך של סדר הסמלים. אם נסתכל על עץ הגזירה של G ועץ הגזירה של G^R, נקבל כי בכל צומת בעץ, סדר הבנים הפוך. לכן, סדר העלים של העץ עבור G^R יהיה היפוך של סדר העלים עבור G. מכאן, כי אם $w\in L(G)$ אזי $w^R\in L(G^R)$ כנדרש.

⇒ נניח כי $w^R\in L(G^R)$ ונראה כי $w\in L(G)$. בחלק הראשון הראנו כי אם $w\in L(G)$ אזי $w^R\in L(G^R)$. היות ו-$w^{R^R} = w$, נהפוך את התפקידים בחלק הראשון של ההוכחה ונקבל כי אם $w^R\in L(G^R)$ אזי $w\in L(G)$ כנדרש.

ב. PP(L) = { w | w∈L and no proper prefix of w is in L}

i. שפות רגולריות סגורות על פעולה זו.

הוכחה:
נציג אוטומט המקבל את PP(L). השפה L היא שפה רגולרית. לכן, קיים אוטומט סופי דטרמיניסטי $A=(Q,\Sigma,\delta,q_0,F)$ המקבל אותה. נשתמש ב-A על מנת להגדיר אוטומט A' המקבל את PP(L).

$A'=(Q\cup\{q_{loop}\},\Sigma,\delta',q_0,F)$ כאשר:

1. q_{loop} הוא מצב חדש.
2. $$\delta'(q,a)=\begin{cases}\delta(q,a) & q\in Q\wedge q\neq q_{loop}\wedge a\in\Sigma\\ q_{loop} & q\in F\cup\{q_{loop}\}\end{cases}$$
(החלפנו את כל המעברים ממצבי הסיום במעברים אל המצב החדש q_{loop} וכן חץ עצמי ב- q_{loop}. יתר המעברים נשארו כשהיו).

u.multinet.co.il

הוכחת נכונות:
$PP(L)\subseteq L(A')$ – תהי $w=w_1\cdots w_n\in PP(L)$. כיוון ש-w שייכת ל-PP(L), קיים מסלול $q_0\to q_1\to\ldots\to q_{n+1}$ באוטומט A כך ש-$\delta(q_{i-1},w_i)=q_i$ לכל $0<i<n+1$ (כך ש-q_{n+1} הוא מצב מסיים). כיוון ש-w שייכת ל-PP(L), וכן היות ו-A הוא אוטומט דטרמיניסטי סופי, מתקיים כי $q_0, q_1,\ldots,q_n\notin F$. (אחרת, הייתה מילה שהיא Proper Prefix של w ואז w לא הייתה ב-PP(L)). לכן, ב- 'A גם קיים המסלול $q_0\to q_1\to\ldots\to q_{n+1}$ וכיוון ש-q_{n+1} שייך ל-F, האוטומט 'A מקבל את המילה. כלומר, $w\in L(A')$ כנדרש.
$PP(L)\supseteq L(A')$ – תהי $w=w_1\cdots w_n\in L(A')$. כיוון ש-w שייכת ל-L('A), קיים מסלול $q_0\to q_1\to\ldots\to q_{n+1}$ באוטומט 'A כך ש-$\delta'(q_{i-1},w_i)=q_i$ לכל $0<i<n+1$ (כך ש-q_{n+1} הוא מצב מסיים). היות ומהמצב החדש q_{loop} לא ניתן "לצאת" וכיוון שזהו אינו מצב מסיים, נסיק כי $q_0,\ldots,q_{n+1}\neq q_{loop}$. לכן, מסלול זה קיים גם באוטומט A המכריע את L. כמו כן, בשל השינויים שביצענו ב-'δ, לפיהם ממצב מסיים ניתן לעבור רק אל מצב q_{loop}, נסיק כי $q_0,\ldots,q_n\notin F$. לכן, לא קיים ל-w מילה השייכת ל-L ומהווה Proper Prefix שלה. ומכאן $w\in PP(L)$ כנדרש.

הראנו הכלה בשני הכיוונים ולכן נסיק כי $PP(L)=L(A')$.

ii. שפות חסרות הקשר לא סגורות על פעולה זו.

הוכחה:

נציג דוגמה נגדית. נגדיר שפה חדשה L באופן הבא:
$L=\{ a^nb^mc^k \mid k\geq n \text{ or } k\geq m \}$. נשים לב כי L היא איחוד של השפות $L_1=\{ a^nb^mc^k \mid k\geq n \}$ ו-$L_2=\{ a^nb^mc^k \mid k\geq m \}$.

נראה כי L_1 ו-L_2 אכן שפות ח"ה ע"י הצגת הדקדוק המכריע אותן:

1. דקדוק G_1 עבור L_1:
$S \to Sc \mid T$
$T \to aTc \mid R$
$R \to bR \mid \varepsilon$

הוכחת נכונות:

$L_1 \subseteq L(G_1)$ יהי $w=a^nb^mc^k$ כך ש-$k\geq n$. נראה סדר גזירה כדי לקבל את המילה: תחילה נבצע k-n פעמים את הגזירה $S \to Sc$. לאחר גזירות אלו נקבל את המילה Sc^{k-n}. עתה, נבצע פעם אחת את הגזירה $S \to T$ ולאחריה n פעמים את הגזירה $T \to aTc$. לאחר גזירות אלו תתקבל המילה $a^nTc^nc^{k-n}= a^nTc^k$. לסיום נבצע פעם אחת את הגזירה $T \to R$, לאחריה m פעמים $R \to bR$ ולבסוף $R \to \varepsilon$. בתום גזירות אלו נקבל $a^nb^mc^k$ כנדרש.

$L(G_1) \subseteq L_1$ יהי $w\in L(G_1)$. נעקוב אחר בניית w. לאחר הפעלת כלל מהצורה $S \to X$, תתקבל מילה עם אחת מהצורות הבאות Sc^i, Tc^i ($i\geq 0$). עבור Sc^i כללי השפה נשמרים. לכן, נראה כי גם עבור Tc^i כללי השפה נשמרים. לאחר גזירות מסוג $T\to X$ נקבל מילה מהצורה $a^jTc^jc^i$, $a^jRc^jc^i$ (כאשר $j\geq 0$). עבור $a^jTc^jc^i$ תנאי השפה נשמרים (מספר ההופעות של c הוא לפחות כשל a וסדר האותיות נשמר עד כה). לכן, נמשיך עבור $a^jRc^jc^i$. לאחר גזירות מהסוג $R \to X$ נקבל מילה מהצורה $a^jb^pc^jc^i$. ניתן להבחין בקלות כי מילה זו שומרת על כללי השפה (מספר ההופעות של c הוא לפחות כשל a וסדר האותיות הוא כנדרש).

הראנו הכלה בשני הכיוונים ולכן נסיק **$L(G_1) = L_1$.**

u.multinet.co.il

2. דקדוק G_2 עבור L_2:
$S \to aS \mid T$
$T \to Tc \mid R$
$R \to bRc \mid \varepsilon$

הוכחת נכונות:

$L_2 \subseteq L(G_2)$ יהי $w=a^nb^mc^k$ כך ש-$k\geq m$. נראה סדר גזירה כדי לקבל את המילה: תחילה נבצע n פעמים את הגזירה $S \to aS$. לאחר גזירות אלו נקבל את המילה a^nS. עתה, נבצע פעם אחת את הגזירה $S \to T$ ולאחריה k-m פעמים את הגזירה $T \to Tc$. לאחר גזירות אלו תתקבל המילה a^nTc^{k-m}. לסיום נבצע פעם אחת את הגזירה $T \to R$, לאחריה m פעמים $R \to bRc$ ולבסוף $R \to \varepsilon$. בתום גזירות אלו נקבל $a^nb^mc^mc^{k-m}= a^nb^mc^k$ כנדרש.

$L(G_2) \subseteq L_2$ יהי $w\in L(G_2)$. נעקוב אחר בניית w. לאחר הפעלת כלל מהצורה $S \to X$, תתקבל מילה עם אחת מהצורות הבאות a^iS, a^iT ($i\geq 0$). עבור a^iS כללי השפה נשמרים. לכן, נראה כי גם עבור a^iT כללי השפה נשמרים. לאחר גזירות מסוג $T\to X$ נקבל מילה מהצורה a^iTc^j, a^iRc^j (כאשר $j\geq 0$). עבור a^iTc^j תנאי השפה נשמרים. לכן, נמשיך עבור a^iRc^j. לאחר גזירות מהסוג $R \to X$ נקבל מילה מהצורה $a^ib^pc^pc^j$. ניתן להבחין בקלות כי מילה זו שומרת על כללי השפה (מספר ההופעות של c הוא לפחות כשל b וסדר האותיות הוא כנדרש).

הראנו הכלה בשני הכיוונים ולכן נסיק **$L(G_2) = L_2$.**

ראינו כי L_1 ו-L_2 הן שפות חסרות הקשר ולכן L הינה גם שפה חסרת הקשר (כאיחוד של שפות אלו).

נשים לב כי $PP(L) = \{ a^nb^mc^k \mid k=\min(a,m) \}$. בשאלה 2, ראינו כי שפה זו אינה חסרת הקשר. קיבלנו סתירה ולכן שפות חסרות הקשר <u>אינן סגורות</u> על פעולה זו.

ג. $Half(L)=\{ x \mid \exists y\in\Sigma^* \text{ s.t. } |x|=|y| \text{ and } xy\in L\}$

i. שפות רגולריות סגורות על פעולה זו.

הוכחה:

נציג אוטומט המקבל את השפה Halt(L). L היא שפה רגולרית ולכן קיים אוטומט סופי דטרמיניסטי $A=(Q,\Sigma,\delta,q_0,F)$ המקבל אותה. נגדיר אוטומט חדש $A'=(QxQxQ\cup\{<q_s,q_s,q_s>\},\Sigma,\delta',<q_s,q_s,q_s>,F')$ כאשר:

$$\delta'(<q_1,q_2,q_3>,a)=\begin{cases}\{<q_0,q,q> \mid q\epsilon Q\} & q_1=q_2=q_3=q_s \wedge a=\varepsilon\\ \{<\delta(q_1,a),q_2,\delta(q_2,b)> \mid b\in\Sigma\} & a\neq\varepsilon\end{cases}$$

$$F'=\{<q_1,q_2,q_3> \mid q_1,q_2\in Q \wedge q_1=q_2 \wedge q_3\in F\}$$

הוכחת נכונות:
$L(A')\subseteq Half(L)$ תהי $w=w_1\cdots w_k\in L(A')$ אזי קיים מסלול $<q_s,q_s,q_s>\to<q_0,q_k,q_k>\to\ldots\to<q_k,q_k,q_n>$ באורך k כך ש-$q_n\in F$. מכאן, ע"פ הגדרת δ' ב-A קיים מסלול $q_0\to\ldots\to q_k$ וכן מסלול $q_k\to\ldots\to q_n$ שניהם באורך k. מכאן, שקיים y כך ש-$|y|=|w|=k$ ו-A מקבל את המילה wy באמצעות המסלול $q_0\to\ldots\to q_k\to\ldots\to q_n$ ומכאן ש-$w\in Half(L)$ כנדרש.

u.multinet.co.il

$L(A')\subseteq Half(L)$ תהי $w=w_1\cdots w_k\in Half(L)$ אזי קיים $y=y_1\cdots y_k$ כך ש- $|y|=|w|=k$ וכן $wy\in L$. מכאן שקיים מסלול $q_0\to\ldots\to q_k$ כך ש- $\delta(q_{i-1},w_i)=q_i$ לכל $0<i\leq k$ וכן קיים מסלול $q_k\to\ldots\to q_n$ כך ש- $\delta(q_{i-1},y_i)=q_i$ לכל $k<i\leq n$. מכאן שקיים מסלול מקבל $<q_s,q_s,q_s>\to<q_0,q_k,q_k>\to\ldots\to<q_k,q_k,q_n>$ ב-A' ולכן $w\in L(A')$ כנדרש.

הראנו הכלה בשני הכיוונים ולכן נסיק $L(A')=Half(L)$. כלומר, האוטומט שבנינו מקבל את השפה ולכן זו שפה רגולרית.

ii. שפות חסרות הקשר <u>אינן</u> סגורות על פעולה זו.

הוכחה:

נציג דוגמה נגדית. נגדיר שפה חדשה L באופן הבא:
$L=\{ 0^n1^n2^m\%\%3^{3m} \mid n,m\in\mathbb{N} \}$. נשים לב כי L היא שרשור של השפות $L_1=\{ 0^n1^n \mid n\in\mathbb{N} \}$ ו- $L_2=\{ 2^m\%\%3^{3m} \mid m\in\mathbb{N} \}$.

1. בכיתה ראינו כי L_1 היא שפה חסרת הקשר.

2. נראה כי גם L_2 היא שפה חסרת הקשר ע"י הצגת דקדוק G_2 המכריע אותה: $S \to 2S333 \mid \%\%$.

הוכחת נכונות:

$L_2 \subseteq L(G_2)$ יהי $w=2^m\%\%3^{3m}$. נראה סדר גזירה כדי לקבל את המילה: תחילה נבצע n פעמים את הגזירה $S \to 2S333$. לאחר גזירות אלו נקבל את המילה 2^mS3^{3m}. עתה, נבצע פעם אחת את הגזירה $S \to \%\%$ ונקבל את המילה $2^m\%\%3^{3m}$ כנדרש.

$L(G_2) \subseteq L_2$ יהי $w\in L(G_2)$. נעקוב אחר בניית w. לאחר הפעלת כלל מהצורה $S \to 2S333$, תתקבל מילה מהצורה 2^iS3^{3i}. ככל שנמשיך לגזור עם כלל זה גודל i יגדל אך המילה תישאר בצורה זו. במידה נגזור $S \to \%\%$ בניית המילה תסתיים ונקבל $2^i\%\%3^{3i}$ כנדרש.

הראנו הכלה בשני הכיוונים ולכן נסיק **$L(G_2) = L_2$.**

ראינו כי L_1 ו-L_2 הן שפות חסרות הקשר ולכן L הינה גם שפה חסרת הקשר (כשרשור של שפות אלו).

נגדיר שפה רגולרית $L_3=0^*1^*2^*\%$.

נניח בשלילה כי שפות חסרות הקשר סגורות על הפעולה Half(L). לכן, השפה $L_4=Half(L)\cap L_3$ הינה שפה חסרת הקשר (כחיתוך של שפה רגולרית ושפה חסרת הקשר). נשים לב כי $L_4=\{ 0^n1^n2^n\% \mid n\in\mathbb{N} \}$ (המילים הנכללות בחיתוך הן מהמקרה בו n=m). אך שפה זו אינה חסרת הקשר. קיבלנו סתירה ולכן שפות חסרות הקשר אינן סגורות על Half.

4. יהי G דקדוק חסר הקשר בצורה נורמאלית של חומסקי. הוכח פורמאלית שלכל $w\in L(G)$, $|w|=n$ $(n>0)$, w נגזר בדיוק ב-2n-1 שלבי גזירה.

הוכחה:

נניח בשלילה שהטענה אינה נכונה. כלומר, קיימת דרך לקבל את w בדקדוק חומסקי ע"י גזירה עם יותר או פחות מ-2n-1 שלבי גזירה. לפי דקדוק חומסקי, גזירה לאות הינה גזירה מהצורה $A\to a$ בלבד. לכן, כדי להגיע למילה בת n תווים, עלינו לבצע n גזירות

u.multinet.co.il

ממצב לאות. בנוסף, על מנת לבצע גזירות אלו, עלינו תחילה לקבל רצף של לפחות n מצבים. היות ובדקדוק חומסקי לא ניתן לבצע גזירה ממצב השונה ממצב ההתחלה לאפסילון, עלינו לקבל רצף של בדיוק n מצבים. כיוון שגזירה של מצב יכולה להיגזר לאות או ל-2 מצבים, אנו זקוקים בדיוק ל-n1 גזירות כאלו (היות ואנו מראש מתחילים עם מצב התחלה S ללא גזירה ולא ניתן לבצע גזירה ממצב אחר לאפסילון). סה"כ קיבלנו כי אם נבצע פחות מ-2n1 גזירות, אורך המילה יהיה קטן מ-n. כמו כן, אם נבצע יותר מ-2n-1 גזירות, אורך המילה יהיה בוודאות גדול מ-n בסתירה לכך שניתן לגזור את w במספר אחר של שלבי גזירה.

5. ברצוננו לבנות אלגוריתם הבודק האם שפה של אוטומט סופי דטרמיניסטי או של דקדוק חסר הקשר, היא אינסופית.

א. הוכח ש$L(A)$ אינסופית, אמ"מ קיימת מילה $w\in L(A)$ כך שמתקיים $|Q|<|w|\le 2|Q|$.

הוכחה:
תחילה, נשים לב שבאוטומט סופי דטרמיניסטי אין מעברי אפסילון ולכן בכל מעבר אורך המילה גדל ב-1.

$\Leftarrow$ נניח כי $L(A)$ אינסופית. כיוון שמספר המצבים באוטומט הוא סופי, ע"מ שהשפה תהיה אינסופית, חייב להיות מעגל באוטומט. גודל המעגל הינו לכל היותר $|Q|$. נגדיר את תת המחרוזת הנוצרת בריצה אחת במעגל להיות y. את המחרוזת עד המעגל נסמן כ-u ואת המסלול מהמעגל עד המצב המסיים להיות v. נגדיר את $w=uy^iv$ (כיוון שאנו יכולים לרוץ במעגל כמה פעמים שנרצה, נסמן את מספר הריצות ב-i). היות וגודל המעגל קטן שווה ל-$|Q|$, נוכל לבחור את i להיות כך ש-$|Q|<|uy^iv|\le 2|Q|$ (נגדיל את i עד אשר לראשונה, $|uy^iv|$ יהיה גדול מ-$|Q|$. גודל המילה שתוצר לא תהיה גדולה מ-$2|Q|$ כיוון שגודל המעגל קטן מ-$|Q|$). מצאנו מילה כזו כנדרש.

$\Rightarrow$ נניח כי קיימת מילה $w\epsilon L(A)$ כך שמתקיים $|Q|<|w|\le 2|Q|$. כיוון ש-A בשפה, קיים רצף מצבים $q_0\to\ldots\to q_n$ (כאשר q_n הוא מצב מסיים). כיוון $|Q|<|w|$, במסלול זה קיים מעגל. לכן, ניתן "ליצר" מספר בן מניה של מילים באמצעות מעגל זה (עבור כל מילה לרוץ i פעמים במעגל לכל $i\epsilon\mathbb{N}$). לכן, מספר המילים בשפה הוא אינסופי, כלומר זוהי שפה אינסופית כנדרש.

ב. הסק כי קיים אלגוריתם המחליט האם $L(A)$ היא אינסופית

הוכחה:
נתאר את האלגוריתם: ניתן לייצג את האוטומט A גרף מכוון. נייצר גרף מכוון כזה H: עבור כל מצב של האוטומט, נוסיף לגרף צומת וכל חוק $\delta(q,a)=q'$ נוכל לייצג כקשת מכוונת (q,q') (ניתן לציין את a בקשת אך לצורך הוכחה זו אין צורך בכך). נזכור שבאוטומט סופי דטרמיניסטי לא קיימים מעברי אפסילון. לכן, אם קיים בגרף שייצרנו מעגל, הרי שהשפה $L(A)$ היא אינסופית. נוכל לבדוק זאת באמצעות אלגוריתם למציאת מעגל בגרף מכוון עבור H. אם קיים ב-H מעגל, הרי שהשפה אינסופית. אחרת, השפה אינה אינסופית.

ג. הוכח שבהינתן G, דקדוק חסר הקשר בצורת חומסקי, $L(G)$ אינסופית אמ"מ קיימת מילה $w\in L(G)$ כך ש- $2^{|V|}<|W|\le 2^{|V|+1}$ (V היא קבוצת non-terminals של הדקדוק G).

הוכחה:
$\Rightarrow$ נניח שקיימת מילה w המקיימת את התנאים. ע"פ דקדוק חומסקי, אורך הסמלים מצד ימין של גזירה הגדול ביותר הוא 2 (במקרה של מעבר של מצב ל-2 מצבים). כמו כן, רצף האותיות הארוך ביותר בגזירה הוא 1 (כלומר, ניתן לגזור כל מצב לאות יחידה בלבד). לכן, על מנת להגיע למילה עם מספר תווים

u.multinet.co.il

הגדול מ-$2^{|V|}$, בעץ הגזירה צריכות להיות $|V|+2$ שכבות כולל שכבת האותיות, כלומר, $2^{|V|+1}$ שכבות המכילות משתנים. לכן, בעץ הגזירה קיים רצף גזירות באורך $|V|+1$ המכיל משתנים בלבד ומכאן שקיים משתנה המופיע פעמיים ברצף. לכן, קיימת מעגליות בשפה ומכאן שנוכל "לנפח" אותה באמצעות מעגל זה. מכאן, שהשפה אינסופית כנדרש.

⇐ נניח ש-L(G) היא שפה אינסופית. מספר המצבים בדקדוק הוא סופי, לכן, היות והשפה היא שפה אינסופית, הרי שקיים "מעגל" בין המצבים. כלומר, קיים רצף של גזירות בו אותו המצב מופיע פעמיים. גודל "מעגל" זה הוא לכל היותר $|v|$. נגדיר את w להיות מילה הכוללת מצב כלשהו מהמעגל הקיים בשפה אך מעגל זה לא הופעל ביצירת w. כלומר, ב-w אין חזרה על אותו מצב פעמיים (כלומר ללא מעגליות - כמו מצב בו i=1). בדקדוק בצורת חומסקי, מספר הסימנים המקסימאלי בכל גזירה הוא 2 וגזירה ממצב אל אות מתבצעת עבור אות בודדת בלבד. לכן, אורך מילה זו הוא לכל היותר $2^{|v|}$. כמו כן, תרומת המעגל היא לכל היותר $2^{|v|}$. לכן, נוכל להפעיל את המעגל כך שהמילה המנופחת w' תקיים $2^{|v|}<|w'|\le 2^{|v|}+2^{|v|}=2*2^{|v|}=2^{|v|+1}$ כנדרש.

ד. הסק כי קיים אלגוריתם המחליט האם L(G) היא אינסופית.

הוכחה:
בסעיף הקודם הראנו כי אם השפה הינה שפה אינסופית, קיימת מילה w כך ש-$2^{|V|}<|w|\le 2^{|V|+1}$. מספר המילים באורך זה הוא מספר סופי בגודל $O(\Sigma^{2^{|V|}})$. כמו כן, ראינו בכיתה כי בעיית השייכות לשפה חסרת הקשר הינה בעיה כריעה. לכן, נוכל לייצר את כל המילים האפשרויות באורך זה ועבור כל אחת לבדוק האם היא שייכת לשפה. אם כן, נחזיר T. אחרת, השפה אינה שפה אינסופית ולכן נחזיר F.

ה. הראה אלגוריתם המחליט האם $|L(A)|=2009$.

תאור האלגוריתם:
תחילה, נשתמש באלגוריתם שהצגנו בסעיף ב' ונבדוק האם זוהי שפה אינסופית. אם כן, נחזיר F.
אחרת, אורך המילים שהאוטומט מייצר הוא לכל היותר $|Q|$. לכן, קיימות לכל היותר $O(|\Sigma|^{|Q|})$ מילים בשפה. עבור כל מילה, נוכל לבדוק באמצעות האוטומט האם היא שייכת לשפה. נעשה זאת באמצעות הרצת סימולציה של האוטומט עבור כל מילה. אם בסוף רצף האותיות של המילה הגענו למצב מסיים הרי שהמילה שייכת לשפה. במקרה כזה, נקדם מונה שייכות ב-1. בתום בדיקת כל המילים, נבדוק האם המונה ספר 2009 מילים. אם כן, נחזיר T, אחרת נחזיר F.

הוכחת נכונות:
אם השפה הינה שפה אינסופית, הרי שבוודאי אינה מורכבת מ-2009 מילים. לכן, יוחזר F כנדרש. אחרת, קיים כאמור מספר סופי של מילים. האלגוריתם בודק שייכות של כל מילה "פוטנציאלית". אם המילה שייכת לשפה היא נספרת במונה. לכן, אם בסוף הבדיקה המונה עומד על 2009, הרי שבשפה יש בדיוק 2009 מילים ויוחזר T. אחרת, בשפה אין בדיוק 2009 מילים ולכן יוחזר F כנדרש.

ו. הראה אלגוריתם המחליט האם $|L(G)|=2009$.

תאור האלגוריתם:
תחילה, נשתמש באלגוריתם המכריע את שאלת אינסופיות של שפה חסרת הקשר ונקבע האם L(G) היא אכן אינסופית. אם כן, נחזיר F.
אחרת, נוכל לתרגם את דקדוק השפה לדקדוק בצורת חומסקי (ראינו בשיעור). כיוון שזו שפה סופית, אורך המילים שהדקדוק מייצר הוא לכל היותר $2^{|v|+1}$ (כאשר V היא קבוצת המצבים). מספר המילים באורך זה הוא מספר סופי. לכן, נייצר את כל המילים האפשרויות ועבור כל אחת נבדוק האם המילה שייכת

u.multinet.co.il

לשפה (זו כאמור בעיה כריעה). אם כן, נקדם מונה שייכות ב-1. בתום בדיקת כל המילים, נבדוק האם המונה ספר 2009 מילים. אם כן, נחזיר T, אחרת נחזיר F.

הוכחת נכונות:
אם השפה הינה שפה אינסופית, הרי שבוודאי אינה מורכבת מ-2009 מילים. לכן, יוחזר F כנדרש. אחרת, קיים כאמור מספר סופי של מילים. האלגוריתם בודק שייכות של כל מילה "פוטנציאלית". אם המילה שייכת לשפה היא נספרת במונה. לכן, אם בסוף הבדיקה המונה עומד על 2009, הרי שבשפה יש בדיוק 2009 מילים ויוחזר T. אחרת, בשפה אין בדיוק 2009 מילים ולכן יוחזר F כנדרש.

6.

א. הראה שלא ניתן להשתמש בלמת הניפוח כדי להוכיח שהשפה הבאה אינה חסרת הקשר $L=\{ 3^n0^i1^j2^k \mid \text{if } n=1 \text{ then } i=j=k \}$

הוכחה:
תהי מילה w כך ש-$|w|>p$ (כאשר p הוא קבוע הניפוח של L). נתייחס ל-3 מצבים:

i. אם ב-w לא קיים 3 (כלומר n=0), נחלק את w=uvxyz כך ש-u=x=y=ε, v יכיל את האות הראשונה ב-w ו-z יכיל את w פחות האות הראשונה. נשים לב כי $|vy|=1>0$ וכן כי $|vxy|=1\le p$. ניפוח חלוקה זו לא תשפיע על סדר האותיות אלא רק תגדיל או תקטין את מספר ההופעות של אחת מהאותיות. במצב זה אין הגבלה על כמות ההופעה של כל אות (אלא על הסדר בלבד) ולכן ניפוח של המילה בחלוקה זו יהיה שייך ל-L.

ii. אם ב-w יש הופעה אחת של 3 (כלומר n=1). במקרה זה נחלק את w=uvxyz כך ש-y=x=u=ε, v=3 ו-z יכיל את יתר האותיות של w. נשים לב כי $|vy|=1>0$ וכן כי $|vxy|=1\le p$. עבור i=1 נקבל את w השייך ל-L. עבור i≠1, נקבל את w עם מספר הופעות מנופח של 3. לכן, כל ניפוח של המילה יהיה שייך ל-L.

iii. אם ב-w יש 2 הופעות של 3 (כלומר n≥2). במקרה ו-$2\le p$, נחלק את w=uvxyz כך ש-y=z=u=ε, $v=3^2$ ו-x יכיל את יתר האותיות של w. נשים לב כי $|vy|=2>0$ וכן כי $|vxy|=2\le p$. עבור i=1 נקבל את w השייך ל-L. עבור i≠1, נקבל את w עם מספר הופעות מנופח של 3 ובכל מקרה, מספר ההופעות של 3 לא יהיה 1. כמו כן, סדר האותיות ב-w ישמר. לכן, כל ניפוח של המילה יהיה שייך ל-L.
במקרה ו-$2>p$, אזי p=1, אם $|w|>2$ נחלק את w=uvxyz כך ש-x=y=ε, $u=3^2$, v יכיל את האות השלישית ב-w ו-z יכיל את יתר האותיות ב-w (האות הרביעית והלאה, אם קיימות). בחלוקה זו, מספר ההופעות של 3 תמיד גדול שווה ל-2 (ובפרט לא 1) וכן סדר האותיות נשאר זהה. לכן, כל ניפוח של המילה יהיה שייך ל-L. אם $|w|=2$, אזי $w=3^2$.
במקרה כזה כל ניפוח יצור מילה המכילה רק הופעות של 3 ולכן שייכת ל-L. לכן, גם כאן ניתן לבצע חלוקה.

ב. הוכח ש-L אינה חסרת הקשר.

הוכחה:

נניח בשלילה כי L היא שפה חסרת הקשר. נגדיר שפה חדשה $G=L\cap 3(0)^*(1)^*2^n$. השפה $3(0)^*(1)^*2^n$ היא שפה רגולרית כיוון שמורכבת משרשור ופעולת כוכבית על שפות רגולריות. לכן, ע"פ המשפט שהוכחנו בכיתה, G היא שפה חסרת הקשר (כיוון שהיא נוצרת מחיתוך של שפה חסרת הקשר ושפה רגולרית). כמו כן, $G=30^n1^n2^n$. נראה כי G <u>אינה</u> שפה חסרת הקשר.

u.multinet.co.il

טענה: G אינה חסרת הקשר.

הוכחה: נניח בשלילה כי G חסרת הקשר ו-p קבוע הניפוח שלה. נבחר מילה $w=30^p1^p2^p$. לכל חלוקה $w=uvxyz$ מתקיים $xyz=30^i$, $xyz=0^i$, $xyz=1^i$, $xyz=2^i$, $xyz=0^i1^j$ או $xyz=1^i2^j$. עבור כל המקרים הראשונים נבחר i=0. נקבל כי לאחר הניפוח מספר ההופעות של לפחות אחת האותיות 0,1,2 לא השתנה ומספר ההופעות של לפחות אחת מהאותיות 0,1,2 קטן ממש. כלומר, מספר ההופעות של 0,1,2 אינו שווה ולכן המילה שהתקבלה אינה ב-G. מכאן, שלמת הניפוח לא מתקיימת ולכן G אינה חסרת הקשר.

הראנו כי G אינה חסרת הקשר בסתירה למשפט שהוכחנו בכיתה. לכן, L אינה שפה חסרת הקשר.

u.multinet.co.il