צביר הרשול
ברק רפקין

עבודה מספר 3- מודלים חישוביים

1. הוכחת המשפט:

L רגולרית אם ורק אם קיימת במטריצה מספר סופי של שורות ייחודיות

$\Leftarrow$ נניח ש-L רגולרית. על פי משפט שנלמד בהרצאה שפה L היא רגולרית אם"ם Rank(L) סופי. ניתן לראות על פי הגדרת המטריצה M(L) כי כל שורה ~~[illegible]~~ ייחודית (שאינה מלאה באפסים) מייצגת מילים שהשפה מקבלת. כל מילה שונה מהשנייה ✓ וקיימת לה נוספת ~~[illegible]~~ מחלקת שקילות ~~[illegible]~~ בשפה. זאת מכיוון שקיימת מילה שמוכלת בה סיפא מפרידה עם מילים אחרות ומכאן שמכיוון שהנחנו ש-L רגולרית, על מנת שהרגולריות תתקיים חייב להיות מספר סופי ומוגבל של שורות ייחודיות, אחרת Rank(L) יהיה אינסופי עקב קבלה של מחלקות שקילות חדשות ככל שיורדים מטה במטריצה כגון קבלה של סיפות מפרידות חדשות בין זוגי מילים ספציפיים.

$\Rightarrow$ נניח ~~שקיימות במטריצה~~ בשלילה. שלא קיימות במטריצה מספר סופי של שורות ייחודיות, כלומר דרגת המטריצה היא אינסופית. ~~[illegible]~~ מכל שורה נמצא לפחות נציג שקילות אחד ממחלקת שקילות שלא היו לה נציגים במטריצה קודם לכן ~~[illegible]~~ זאת מכיוון שככל שמתקדמים למטה בשורות המטריצה ~~[illegible]~~ מיוצגות מילים חדשות ~~שלא~~ מהשרשרת

לכל $i \neq j$ במקום ה $[i,j]$ M(L) במטריצה ולכן מספר הסיפות המפרידות בין זוגי מילים שונות שהוא גדל ומספר מחלקות השקילות יהיה אינסופי ומכאן ש-L לא תהיה רגולרית. ∎

.ב

1. L' רגולרית. נוכיח בעזרת טענת עזר שנלמדה בכיתה.

טענת עזר: שפה L' היא רגולרית אם״ם $Rank(L')$ סופי.

מכיוון ש-A קבוצת מחלקות שקילות שהינה סופית והגדרנו כי $A' \subseteq A$ קבוצה כלשהי של מחלקות שקילות מתוך A ו-L' הוגדרה כך $L' = \bigcup_{A_j \in A'} A_j$, אזי שניתן להסיק כי $Rank(L')$ סופי מכיוון שהוא כולל חלק מן מחלקות השקילות מתוך קבוצת מחלקות השקילות הסופית A.

2. הוכחה עבור הטענה: $rank(L') \leq rank(L) = n$

ראשית כל כי, הגדרנו השאלה $A = \{A_1, A_2, \dots, A_n\}$ היא קבוצת מחלקות השקילות תחת היחס $\sim_L$.
משמע, שדרגת השפה L שווה ערך בדיוק ל-n.
בנוסף לכך כפי שהוסבר בסעיף קודם לכן, השפה L' מוגדרת כאיחוד אונרי של מחלקות השקילות של A' כאשר $A' \subseteq A$. מכיון שכך שדרגת L' תהי״ה שווה לדרגת L במידה ו-$A' = A$ או שהיא תהי״ה קטנה מדרגת L' במידה ו-$A' \subsetneq A$.

3. הוכחה עבור המשפט : $|A^-| \geq rank(L^-)$.

נניח בשלילה ש- $|A^-| \not\geq rank(L^-)$, מכאן
שגודל קבוצת מחלקות השקילות של השפה L^-
צריך להיות גדול שווה מגודל $|A^-|$, אך
לפי הגדרת נתוני השאלה ניתן לראות כי השפה
L^- מוגדרת כך:

$$L^- = \bigcup_{A_j \in A^-} A_j$$

כלומר ה-rank של השפה L^- ~~מוגדר~~ חייב להיות שווה
לגודל של $|A^-|$ ואינו יכול להיות גדול מ-$|A^-|$
משום שאז זה סותר את ההגדרה שהצגתי לעיל.

ח. הוכחה עבור המשפט: קיימת שפה L ובחירה של
מחלקות שקילות A^- כך ש-$rank(L^-) < rank(L)$.

לאפשר כל בחירה של $\acute{A}$ כך ש: $A^- \subsetneq A$ ויצירת שפה
L^- בצורה שהגדירו:

$$L^- = \bigcup_{A_j \in A^-} A_j$$

תיצור שפה L^- שה rank שלה קטן יותר מה-rank
של השפה L. על מנת שזה יקרה, כפי שצוין
A^- צריכה להיות מוכלת ב- A אך לא מוכלת ממש
(כלומר אסור ש- $A^- = A$).

1. הפרכה למשפט.

משפט: לכל שפה רגולרית L ולכל אות $\sigma\in\Sigma$ מתקיים $rank(L)=rank(f_\sigma(L))$ כאשר $f_\sigma(L)$ מוגדרת כשפה הבאה:

$$f_\sigma(L)=\{w\in\Sigma^*:\sigma\cdot w\in L\}$$

ניקח את השפה $L=\{abb\}$

נסיק כי $f_\sigma(L)=\{bb\}$

ניתן לראות כי מספר מחלקות השקילות של השפה L הוא 4. שכן יש סיפא מפרידה

בין $w=\varepsilon$ ל- aa שהיא abb.

בין $w=a$ ל- b שהיא bb.

בין $w=ab$ ל- aa שהיא b.

ובנוסף מחלקת שקילות של שאר המילים שלא בשפה.

ב $F_\sigma(L)$ יש מחלקת שקילות אחת פחות מכיוון שהמילה קצרה יותר בתו אחד. דבר הגורם לסיפא מפרידה אחת פחות.

הסיפות המפרידות:

בין $w=\varepsilon$ ל- aa שהיא bb

בין $w=b$ ל- aa שהיא b

ובנוסף מחלקת שקילות של שאר המילים שלא בשפה

$$4\neq 3$$

$$\Downarrow$$

$$rank(L)\underset{(4)}{\neq} rank(f_\sigma(L))$$

ב. הוכחת המשפט: לכל שפה רגולרית מתקיים

$$rank(L) = rank(L^R)$$

תהייה $A = \{A_1, A_2, \dots, A_n\}$ קבוצת מחלקות השקילות של השפה L. ניקח מילה השייכת לאחת מן מחלקות השקילות שבקבוצה A. ונהי $x \in A_i$ אותה המילה. ניתן לראות כי עבור הסיפא המשלימה של למילה בשפה נקיימת גם סיפא v משלימה למילה ההפוכה בגודל $|x|$ v בשפה L^R (לאחר הפיכת הסיפא). כלומר לכל מילה במחלקת שקילות שבקבוצה A תהיה בהתאמה מילה הפוכה באותו הגודל במחלקות שקילות של השפה L^R. ומכאן ניתן להסיק שמספר מחלקות השקילות של השפה L חייב להיות שווה למספר מחלקות השקילות של השפה L^R.

$$\Downarrow$$

$$rank(L) = rank(L^R)$$

3. הטענה נכונה.

טענה: עבור שפה רגולרית לא ריקה L נסמן ב-k
את אורך המילה הקצרה ביותר ב-L. אז $k < rank(L)$.

מטענה שנלמדה בכיתה עבור שפה רגולרית L שבה
אורך המילה הארוכה ביותר היא n כל אוטומט אי
דטרמיניסטי שמקבל את השפה יכיל לפחות
$n+1$ מצבים. כלומר כגודל של מספר מחלקות
השקילות של השפה לכל הפחות. בלא ובטח שהמשפט
יתקיים עבור המילה הקצרה ביותר שכן המילה
הארוכה ביותר גדולה שווה לאורך המילה הקצרה ביותר.

לכן ~~[illegible]~~

$rank(L) > n > k$

לכבי אוטומט דטרמיניסטי נבמידה ואורך המילה הקצרה
ביותר המתקבלת בשפה הוא k, אזי שאנו צריכים
לעבור לפחות $k+1$ מצבים כדי להגיע לאותו
מצב מקבל כאשר אנו מתחשבים במצב ההתחלתי.
לכל רישא של המילה הקצרה ביותר החל
מ-ε ~~[illegible]~~ ועד לאות האחת לפני האחרונה (אורך $k-1$)
תהייה סיפא מפרידה עם מילה אחרת. ומכאן
שקיימות k מחלקות שקילות כמספר הסיפות
המפרידות של כל רישא ~~[illegible]~~ של המילה הקצרה
ביותר עם איזשהי מילה אחרת. בנוסף קיימת מחלקת
שקילות של כל המילים שאינן מתקבלות בשפה.
ומכאן שמדובר ב- $k+1$ מצבים כמינימום.

■

ח. הפרכה למשפט.

משפט: קיימות שתי שפות רגולריות L_1 ו-L_2 כך ש $rank(L_1\setminus L_2) > rank(L_1)\cdot rank(L_2)$.

כלומר אנו נוכיח שלכל שתי שפות רגולריות L_1 ו-L_2 מתקיים

$$rank(L_1\setminus L_2) \leq rank(L_1)\cdot rank(L_2).$$

נסתכל על הביטוי שמצד שמאל $rank(L_1\setminus L_2)$ אותו הביטוי שבתוך הסוגריים מייצג בפועל חיסור כלשהו של מילים הקיימות ב-L_2 ממילים אשר נמצאות ב-L_1. לכן אם $x = rank(L_1)$ עבור x כלשהו, אזי שחיסור המילים יכול להוביל רק להפחתה של ערך x זה היות ומספר מחלקות השקילות יכול לקטון כתוצאה מהסרה של מילים מהשפה L_1.

לכן $x = rank(L_1) \geq rank(L_1\setminus L_2)$

$$\Downarrow$$

$$rank(L_2)\cdot rank(L_1) \geq rank(L_1\setminus L_2)$$

וזה נכון מכיוון ש $rank(L_2) \geq 1$.

ומכאן שהגענו להפרכת המשפט ובכך סיימנו את ההוכחה.

1. השפה $L = AA^*U^nA^nUU^*A^+$

עבור $n \in N$ אינה רגולרית.

ניקח לדוגמא את המילה $w_1 = AU^i$

ואת המילה $w_2 = AU^j$

ניתן לראות כי על מנת שתתקבל מילה בשפה
מ- w_1 צריך להוסיף ל- w_1 סיפא
שהיא $U^{n-i}A^n$. זוהי תהייה סיפא
מפרידה בין w_1 ל- w_2 שכן w_2 לא תהייה
בשפה לאחר הוספת הסיפא כי $j \neq i$. בגלל
ש $n \in N$ נוכל לתת אינסוף דוגמאות לזוגי
מילים בעלי סיפות מפרידות המשויכות למחלקות
שקילות שונות. ומכאן שלשפה זו אינסוף מחלקות
שקילות, כלומר ש- $rank(L) = \infty$. על פי משפט
שנלמד בכיתה אם דרגת השפה היא אינסופית
אזי שהשפה אינה רגולרית.

2. השפה $L=\{w: |w|_{GC}=|w|_{CG} \wedge |w|_{GU}=|w|_{GA}=|w|_{CU}=|w|_{CA}=0\}$

אינה רגולרית. ניקח לדוגמא את המילים

$w_1=(GC)^i(CG)^{i-k}$

$w_2=(GC)^i(CG)^{i-z}$ כך ש $k\neq z$

קיימת סיפא מפרידה $(CG)^k$ אשר גורמת למילה w_1 להשתייך לשפה L ול-w_2 לא להשתייך אליה. מכיוון ומספר המופעים של תתי המחרוזות אינו מוגבל אז נוכל לבחור אינסוף מילים מ-Σ^* כך שכל זוג מילים ניתן להפרדה לפי היחס $\equiv_L$ אזי L אינה רגולרית מכיוון שקיימות אינסוף מחלקות שקילות.

3. השפה $L=\{w: |w|=n^2+2n, n\in N\}$

אינה רגולרית.

נסתכל על סדרת אורכי המילים בשפה.

$0, 3, 8, 15, 24 \ldots \infty$

סדרה זו מונוטונית עולה וכן גם סדרת ההפרשים:

$3, 5, 7, 9 \ldots \infty$

$(n+1)^2+2(n+1)-(n^2+2n)$

$= \cancel{n^2}+\cancel{2n}+1+2n+2-\cancel{n^2}-\cancel{2n} = 2n+3$

מכיוון ש $n\in N$ והולך ושואף לאינסוף אזי שסדרת ההפרשים אינה חסומה.

לפי מסקנה 2 מהתרגול אם סדרת אורכי המילים בשפה L הינה בעלת סדרת הפרשים שאינה חסומה

אזי $L\notin L_{reg}$ (9)

4. $L=\{G^n \cdot U^m \cdot C^{n\cdot m}, n,m \ge 0\}$

נניח בשלילה כי $L \in L_{reg}$

תהי המילה w $\quad w=G^N \cdot U^m \cdot C^{N\cdot m}$

אם כן קיימת חלוקה של w לשלוש מילים

$x,y,z \in \Sigma^*$ כך ש- $w=xyz$ כאשר מתקיימים תנאי הלמה.

מאחר ו- $|xy| \le N$ הרי ש- y מורכבת

מ- G בלבד. כלומר נגדיר $y=G^k$, $0<k\le N$.

נתבונן במילה הבאה $xy^0z=G^{n-k} U^m C^{n\cdot m}$.

ניתן לראות כי מילה זו לא שייכת לשפה L

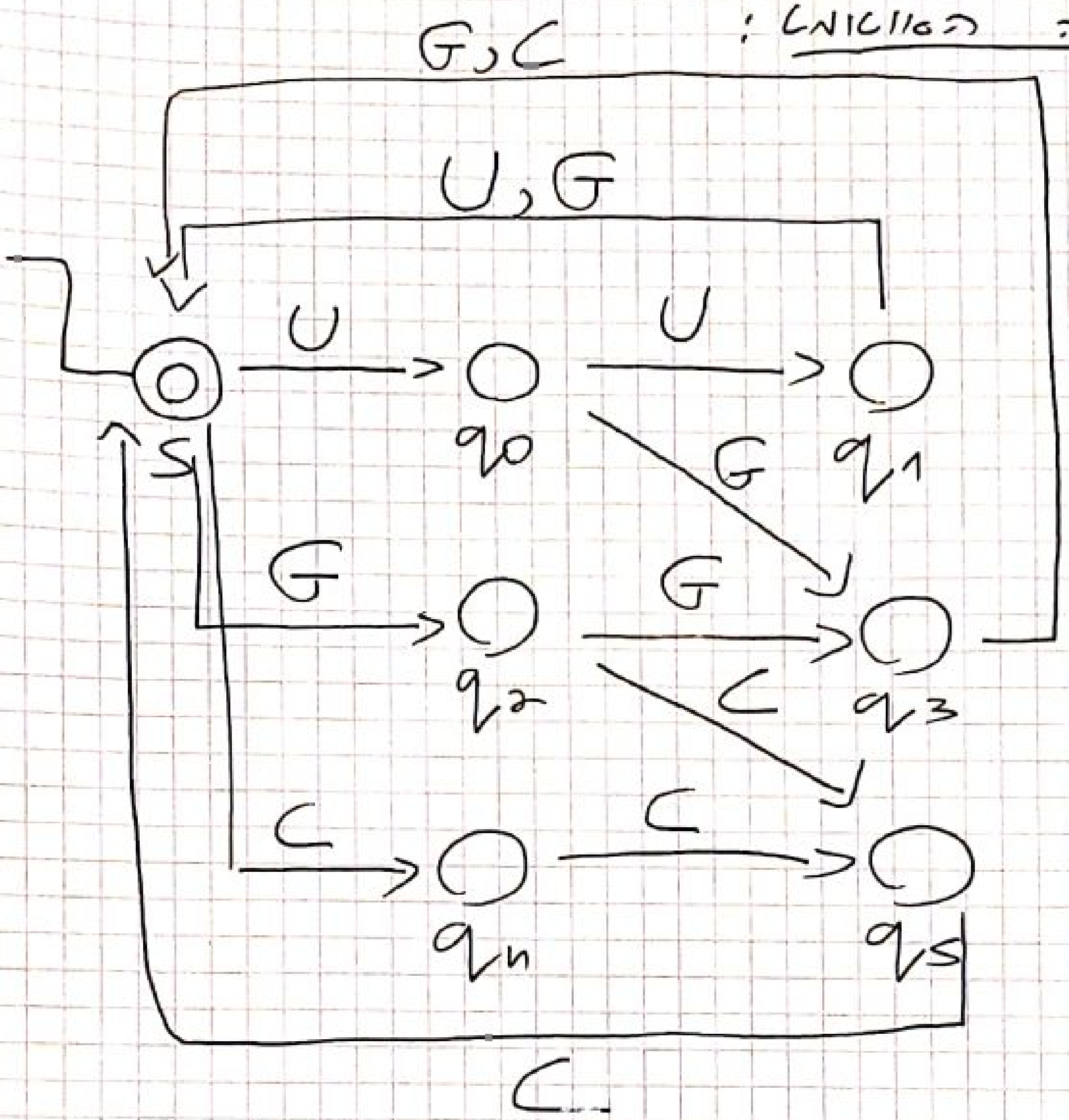
לכן קבלנו סתירה ללמת הניפוח ומכאן $L \notin L_{reg}$.

5. השפה $L=\{U^i \cdot G^j \cdot C^k : (i+j+k) \bmod 3 = 0, i,j,k \ge 0\}$

היא רגולרית. נגדיר את האוטומט.

והיה זה אוטומט בעל ~~7~~ 3 מצבים:

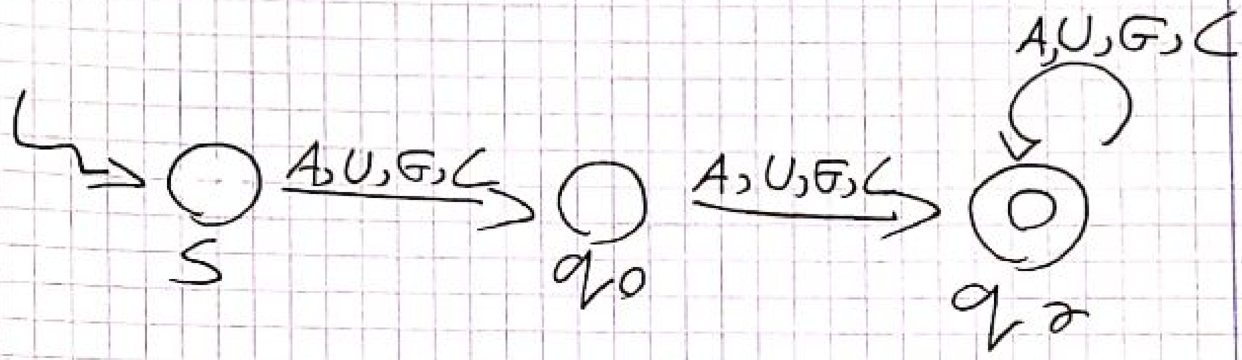
מבנה האוטומט:



ניתן לראות כי אוטומט זה עונה על הגדרת
השפה שכן מתחיל ממצב התחלתי S,
שומר על סדר התווים הנדרש בשפה.
המבוקשת עקב מבנה המדרגות של
האוטומט. ~~רק לאחר~~ תמיד האוטומט יגיע
למצב המקבל רק לאחר שעבר 3·X
מצבים כאשר X נע מ-0 ועד אינסוף.
* לשפה זו קיים אוטומט בעל מספר מצבים
סופי ולכן שפה זו רגולרית.

6. השפה $L=\{w : |w|=n+2 , n\in N\}$

רגולרית. נבנה אוטומט בעל 3 מצבים



כפי שניתן לראות באיור האוטומט, המצב המקבל q_f מתקבל רק לאחר מעבר של לפחות 2 אותיות כלשהן. לאחר מעבר של 2 אותיות הוא חוזר לעצמו לאחר קבלה של כל אות נוספת (התחייה ואשר הנחייה). מכאן שהאוטומט זה עונה בדיוק על הגדרת השפה ולכן שפה זו רגולרית.
* הוא מקבל את כל המילים שאורכן 2 או יותר.

5.

ט. טענה: אם $L=L_1\cup L_2$ אז $n\le \max(n_1,n_2)$

כמו כן קיימות שפות L_1,L_2 כך ש $n<\max(n_1,n_2)$

<u>הוכחה</u>

תהא $L=L_1\cup L_2$ נראה ש $n\le\max(n_1,n_2)$

$w\in L=L_1\cup L_2$ כך שמתקיים $|w|\ge\max(n_1,n_2)$

אז ש- $w\in L_1$ או ש $w\in L_2$. נחלק ל2 מקרים.

מקרה א- $w\in L_1$ וגם $|w|\ge\max(n_1,n_2)$:

$|w|\ge\max(n_1,n_2)\ge n_1$, אז $w=xyz$.

ופירוק w המובטח מלמת הניפוח עבור L_1, מקיים את שלושת התנאים.

מקרה ב- $w\in L_2$ וגם $|w|\ge\max(n_1,n_2)$:

$|w|\ge\max(n_1,n_2)\ge n_2$, אזי $w=xyz$

ופירוק w המובטח מלמת הניפוח עבור L_2 מקיים את שלושת התנאים.

הראינו כי קבוע הניפוח n של השפה החדשה L מקיים $n\le\max(n_1,n_2)$.

ב. עכשיו נוכיח שקיימות שפות L_1,L_2 כך ש $n<\max(n_1,n_2)$,

יהיו L_1,L_2 שפות מעל $\Sigma=\{a,b\}$

$L_1 = \{b\}$ $L_2 \{b\}$ $L_1 \cup L_2 = \{b\}$

שלושתן השפות סופיות ומקיימות את תנאי
למת הניפוח באופן רק כאשר אנו נבחר
h אשר גדול מגודלה של המילה הארוכה בשפה.
נבחר עבור L_1 $h_1=3$ ועבור L_2
$h_2=2$ ועבור $L_1 \cup L_2$ $n=2$

אז מתקיים:

$$2 < \max(n_1, n_2) = 3$$

2) טענה: אמת $L = L_1 \cap L_2$ אם קיימות
שפות L_1, L_2 כך ש $n < \min(n_1, n_2)$
נסתכל על שתי השפות L_1, L_2 מעל $\Sigma = \{x, y\}$

$$L_1 = \{xy, xxy^*x\} \quad L_2 = \{x \cdot y, xxy^*x\}$$

L_1 ו- L_2 שפות רגולריות

$$r_1 = (xy) \cup (xxy^*x) \quad r_2 = (xy) \cup (yyx^*y)$$

$$L(r_1) = L_1 \quad L(r_2) = L_2$$

קבוע הניפוח עבור השפה L_1 הוא $n=n_1$.
תהא $w \in L_1$ מילה כלשהי כך ש $n \leq |w|$.
$w = x^*y^*z^*$ הפירוק של w המובטח מלמת הניפוח
כך ש: $x^* = xx$ $y^* = y$ $z = x$

שלושת תנאי הלמה מתקיימים:

1) $y \neq \varepsilon$

2) $|xy| = 3 < n$

3) לכל $i \in N$ מתקיים $x^* y^{*i} z^* = xxy^i x \in L_1$

הדבר מתקיים בדיוק באותו אופן לקבוע הניפוח עבור השפה L_2 רק שצריך להחליף את x ב-y במילה w שהרכבנו.

בשני השפות הטענה נכונה גם עבור xy כי למת הניפוח מתקיימת באופן ריק יותר ($|xy|=2<n$).

$\min(n_1, n_2) = n$. נסתכל על שפת החיתוך

$$L = L_1 \cap L_2$$

$$L = L_1 \cap L_2 = \{xy, xxy^*x\} \cap \{xy, yyx^*y\}$$

$$= \{xy\}$$

ניתן להוכיח כי את השפה L לא ניתן לנפח כי קיימת בה רק מילה אחת. נבחר את קבוע הניפוח להיות $n=2$ ולכן למת הניפוח תתקיים באופן ריק.

וקיבלנו $2 = n < \min(n_1, n_2) = n$

■

טענה:

3) אם $L=L_1\cdot L_2$ אז $n\le n_1+n_2$.

כמו כן קיימות שפות L_1, L_2 כך ש- $n<n_1+n_2$.

הוכחה

תהא $L=L_1\cdot L_2$ נוכיח ש $n\le n_1+n_2$

תהא $w\in L=L_1\cdot L_2$ כך ש $|w|\ge n_1+n_2$ אז

$w=w_1w_2$ כך ש $w_2\in L_2$ ו- $w_1\in L_1$. כמו כן

מתקיים $|w_1|\ge n_1$ או $|w_2|\ge n_2$

(לא ייתכן שגם $|w_1|<n_1$ וגם $|w_2|<n_2$ כי אז:

$|w|=|w_1|+|w_2|<n_1+n_2$ בסתירה להנחה $|w|\ge n_1+n_2$.

נחלק למקרים:

מקרה א: $w_1\in L_1$ וגם $|w_1|\ge n_1$

אז $w_1=x_1y_1z_1$ הפירוק של w_1 המובטח

מלמת הניפוח. תהא $w=xyz$ כך ש

$z=z_1w_2$ $y=y_1$ $x=x_1$, נראה

ששלושת התנאים מתקיימים:

1) $y=y_1\neq\varepsilon$

2) $|xy|=|x_1y_1|\le n_1\le n_1+n_2$

3) לכל $i\in N$ $xy^iz=x_1y_1^iz_1w_2$

מתקיים ש $w_2\in L_2$ וגם

$x_1y_1^iz_1\in L_1$ ולכן $x_1y_1^iz_1w_2\in L_1\cdot L_2\in L$

מקרה ב: $w_2 \in L_2$ וגם $|w_2| \ge n_2$

אז $w_2 = x_2 y_2 z_2$. פירוק המילה w_2

מובטח מלמת הניפוח עבור L. תהא $w = xyz$

כך ש $x = w_1 x_2$, $y = y_2$, $z = z_2$

נראה כי 3 תנאי הלמה מתקיימים:

1) $y = y_2 \neq \varepsilon$

2) $|xy| = |w_1 x_2 y_2|$, $|x_2 y_2| \le n_2$

וכן $|w_1| < n_1$ ולכן מתקיים $|xy| < n_1 + n_2$

3) לכל $i \in N$ $xy^i z = w_1 x_2 y_2^i z_2$

מתקיים $w_1 \in L_1$ ו- $x_2 y_2^i z_2 \in L_2$

ולכן $w_1 x_2 y_2^i z_2 \in L_1 \cdot L_2 = L$

ובסה"כ הראינו שקבוע הניפוח n של השפה

מקיים $n \le n_1 + n_2$.

$\blacksquare$

Scanned with CamScanner

דוגמא לכך שקיימות שפות L_1,L_2 כך ש$n<n_1+n_2$

$$L_1 = \{ab^i : i \geq 0, i \in N\}$$

קבוע הניפוח של שפה זו יהיה 2. מאחר
ש למילים באורך 1 לא נקבל מילה בשפה
אז עי' $xy^nz \in L$. המינימום הוא:

$w = ab$

$|xy| \leq 2$ $\quad x = a \quad y = b \quad z = \varepsilon$

המילה ו- $xy^nz = ab^n$ לכל $n \geq 0$

$$L_2 = \{b^i a : i \geq 0, i \in N\}$$

קבוע הניפוח של שפה זו יהיה 2.

$w = ba \quad z = a \quad y = b \quad x = \varepsilon$

ו- $xy^nz = b^n a$ לכל $n \geq 0$

כעת אם נגדיר את השפה $L=L_1\cdot L_2$
אז היא תיראה כך:

$$L=\{a\,b^i a : i\geq 0, i\in N\}$$

קבוע הניפוח של שפה זו יהיה 3. מאחר ולמילים באורך 3 ומעלה נקבל מילים בשפה ש־$xy^nz\in L$ לכל $n\geq 0$.

המינימום הנו: מתקיים $|w|=|aba|\geq 3$

$w=aba$

$x=a \quad y=b \quad z=a$

ו־ $xy^nz=ab^na$ לכל $n\geq 0$

ומכאן שקבוע הניפוח הוא 3 ומתקיים

$n=3 \quad n_1=2 \quad n_2=2$

$3<2+2=n$

∎