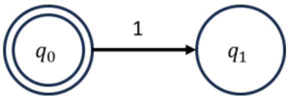
תרגיל בית 3
מגיש: אור דינר
ת.ז: 207035809

<u>שאלה 1</u>
<u>סעיף א</u>

נוכיח שקיים אסל״ד בעל $n+1$.

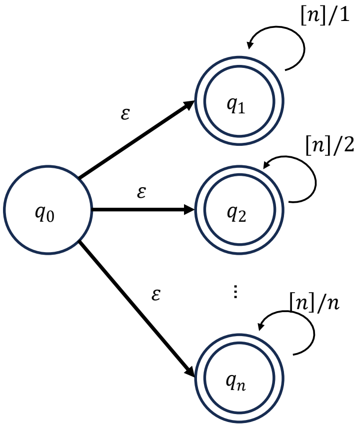
עבור $n=1$:



עבור $n>1$ רעיון הבנייה הוא להלן:

1. ניצור מצב התחלתי q_0.
2. ניצור n מצבים **מקבלים** $q_1, q_2, \dots, q_n$ כאשר מכל אחד נוציא חצים לעצמו עם כל האותיות מלבד האות שמייצגת את המספר מצב שלו – כלומר כל מצב q_i מייצג את המצב בו במילה לא תופיע האות i.
3. מהמצב ההתחלתי נוציא מעברי אפסילון לכל אחד מהמצבים $q_1, q_2, \dots, q_n$.

<u>המחשה</u>



נסמן את האסל״ד ב-N ונוכיח $L=L(N)$ באמצעות הכלה דו כיוונית.

א. $L(N) \subseteq L$

יהי $w \in L(N)$, לפי הגדרת האסל״ד N, כל מסלול חישוב מסתיים במצב q_i כלשהו, ועל פי הגדרת האסל״ד המצב ה-q_i מקבל מילים שלא מופיע בהן האות ה-i. יותר מכך, נניח ש-wלא תתקבל בשפה L, כלומר כל האותיות מופיעות בה לפחות פעם אחת, באסל״ד N אין מצב שמקבל את כל האותיות ואין מעברים בין המצבים בכלל ואז המילה w בהכרח תתקבל בשפה של N.

ב. $L(N) \supseteq L$

יהי $w \in L$, לפי הגדרת L קיימת $a \in [n]$ כך שאינה נכללת במילה w. כאשר נריץ את המילה w באסל״ד N נגיע למצב q_i, במידה ורק a חסרה במילה שלנו אז המצב $q_i = q_a$, אבל אם יש יותר מאות אחת חסרה אז **המסלולים המקבלים** של המילה w באסל״ד N בהכרח יסיימו במצב מקבל כמו שהגדרנו באסל״ד, ובמידה ואיכשהו הגענו למצב שאפשר לעבור ממנו לעצמו בעזרת אות שלא במילה w, אז האסל״ד ייתקע כי במצב שהיא תהיה אין אפשרות לקבל את האות החסרה כקלט. לכן כל מילה שתתקבל על ידי L תתקבל גם על ידי שפת האסל״ד N.

<u>סעיף ב</u>

כדי לחשב באופן דטרמיניסטי האם במילה יש לפחות אות אחת חסרה, נצטרך באס״ד להתייחס לכל מחלקות השקילות שיהיו עבור $[n]$, אשר כל אחת מהן מייצגת מילה שחסרים בה i תווים מהא״ב. כמות מחלקות השקילות היא כמספר תתי הקבוצות של $[n]$, כלומר - 2^n. משום שהראינו שיש לשפה L אסל״ד שיכול לייצג אותה היא רגולרית, ומשום שהיא רגולרית נובע ממשפט מייהיל נרוד כמות המצבים המינימאלי יהיה כמספר מחלקות השקילות, שהוא בדיוק 2^n.

<u>שאלה 2</u>

<u>סעיף א</u>

נבחר את: $L = \{a^n b^m c^m : n > m\}$.

השפה הזו אינה רגולרית אבל מקיימת את למת הניפוח, עבור כל w שמקיים ש- $|w| \geq n_0$, קיים פירוק למילה $w = xyz$, כך ש:

1. $|xy| \leq n_0$
2. $y \neq \varepsilon$
3. ולכל $k > 0$ מתקיים $xy^k z \in L$

כל k שנבחר ישאיר את המילה מנופחת בשפה L כיוון שהיא תמשיך לענות על הקריטריונים של השפה, זאת משום שהחלק שישתנה הוא החלק שיש בו רק a ולכן זה לא יפגע בכך שכמות ה-b שווה לכמות ה-c.

<u>סעיף ב</u>

הטענה **לא נכונה**.

נבחר $L = \{w \in \{0,1\}^* : w\ has\ the\ same\ amount\ of\ 0's\ and\ 1's\}$, אז השפה $Prefix(L)$ תכיל פשוט את כל הרישות של מילה שיש בה כמות זהה של 0 ו-1, וזה פשוט קבוצת כל המילים (Σ^*), וזו שפה רגולרית הניתנת לייצוג על ידי אס"ד.

<u>שאלה 3</u>

<u>סעיף א</u>

משום ש-L ח״ה, קיים $G = (V, \Sigma, R, S)$ דח״ה שמקבל אותה. נגדיר $G^* = (V', \Sigma, R', S)$ כאשר החוקים של הדח״ה הזה זהים לזה של G עם תוספת חוק נוסף שהוא $S \to SS$ שיאפשר שרשור של כל מילה השייכת ל-L.

נוכיח בעזרת הכלה דו כיוונית שמתקיים $L^* = L(G^*)$.

א. $L^* \subseteq L(G^*)$

תהי $w \in L^*$, אז w היא שרשור של כמה מילים מהשפה L כך ש- $w = a_1 a_2 \dots a_n$, ולכל מילה a_i יש סדרת גזירות בדח״ה G שיכול להגיע למילה. לכן כשנפעיל את החוק החדש $S \to SS$ במשך $n-1$ פעמים נקבל n נונטרמינלים התחלתיים שהם גם הנונטרמינלים ההתחלתיים בדח״ה G, וזה אומר שנוכל להגיע לכל אחת מהמילים a_i, כלומר $w \in L(G^*)$.

ב. $L^* \supseteq L(G^*)$

לפי איך שהגדרנו את הדח״ה G^*, כל מילה השייכת ל-$L(G^*)$ היא שרשור של כמה מילים מהשפה $L(G) = L$, לכן כל מילה בשפה $L(G^*)$ תהיה ב-L^* .

הראינו שקיים דח״ה שגוזר את -L^* ולכן היא ח״ה.

<u>סעיף ב</u>

משום ש-L ח״ה, קיים $G = (V, \Sigma, R, S)$ דח״ה שמקבל אותה. נגדיר $G^R = (V', \Sigma, R', S)$ כך שעבור כל חוק ב-R, אנחנו נהפוך את הסדר של הטרמינלים והנונטרמינלים אליו הגזירה מובילה (לדוגמא החוק $A \to aAbBc$ יהפוך לחוק $A \to cBbAa$), בצורה הזו נבטיח שכל המילים שמתקבלות הן הפוכות למילים הנגזרות מהדח״ה G.

נוכיח בעזרת הכלה דו כיוונית שמתקיים $L^R = L(G^R)$.

א. $L^R \subseteq L(G^R)$

תהי $w \in L^R$, כלומר עבור $a_1, a_2, \dots, a_n \in \Sigma$, מתקיים ש-$w = a_n a_{n-1} \dots a_1$, כך שקיימת $u \in L$ כך ש- $u = a_1 a_2 \dots a_n$. בגלל שמתקיים $L = L(G)$, המילה u היא תוצאה של סדרת גזירות בעזרת החוקים של הדח״ה G, בגלל שהפכנו בחוקים של G^R את סדר הטרמינלים והנונטרמינלים נוכל לעשות בדיוק את אותה סדרת גזירות ולהגיע למילה ההפוכה של u, שהיא כאמור w. ולכן המילה w נגזרת על ידי הדח״ה G^R.

ב. $L^R \supseteq L(G^R)$

תהי $w \in L(G^R)$, משום ש-w מורכב מטרמינלים של G^R, ומשום שהפכנו את סדר כל החוקים (טרמינלים ונונטרמינלים), קיימת גזירה כלשהי של G כך שמתקבלת המילה u כך ש-w בדיוק הפוכה לה. בגלל שאותה u מתקבלת בשפה L קיבלנו שהמילה w היא מילה הפוכה למילה המתקבלת על ידי הדח״ה של L ולכן $w \in L^R$.

הראינו שקיים דח״ה הגוזר את L^R ולכן היא ח״ה.

<u>סעיף ג</u>

משום ש-L שפה ח״ה קיים א״מ $M = (Q, \Sigma, \Gamma, \delta, q_0, F)$ המקבל אותה, באותו עיקרון קיים אסל״ד $N = (\hat{Q}, \hat{\Sigma}, \hat{\delta}, \hat{q}_0, \hat{F})$ עבור R שמקבל אותה. נבחין שהחיתוך של שתי השפות $L \cap R$ יכיל את כל המילים אשר מתקבלות בשתי השפות.

נגדיר א״מ $M' = (Q', \Sigma', \Gamma', \delta', q', F')$ אשר יקבל את השפה $L \cap R$. הבנייה תהיה דומה ברעיון לאוטומט מכפלה, המצבים יהיו המכפלה בין שתי האוטומטים המקוריים $Q' = Q \times \hat{Q}$, הא״ב יהיה הא״ב שנמצא בשתי האוטומטים $\Sigma' = \Sigma \times \hat{\Sigma}$, א״ב המחסנית יהיה כמו של הא״מ של L, קבוצת המצבים המקבלים יהיו קבוצת המכפלה של שתי קבוצות המצבים המקבלים של האוטומטים - $F' = F \times \hat{F}$, ופונקצית המעברים δ', תוגדר באופן הבא:

???????????

<u>שאלה 4</u>

ניתן לומר שהשפה L מתארת את המילים בהן כמות ה-0 מתחלקת ב-3 ויש 3 יותר אחדים מה-0.

דוגמא:

$n = 0, w = 111$
$n = 3, w = 000111111$
$n = 6, w = 000000111111111$
$\dots$

נתבונן בדח״ה G עם החוקים הבאים:

1. $S \to A111$
2. $A \to 000A111$
3. $A \to \varepsilon$

הדח״ה הזה מתאר את השפה L, נוכיח ש- $L(G) = L$ באמצעות הכלה דו כיוונית:

א. $L(G) \subseteq L$

יהי $w \in L(G)$, לפי הגדרת הדח״ה w תסתיים בשלושה 1-ים לאחר כמות שווה ומתחלקת ב-3 של 0 ואז 1, וזו בדיוק ההגדרה של המילים המתקבלות בשפה L.

ב. $L(G) \supseteq L$

יהי $w \in L$, אז מתקיים עבור $i \in \mathbb{N}$ $w = 0^i 1^{i+3}$ לפי הגדרת L, ניתן להבחין כי $w = 0^i 1^i 111$, כלומר כשנפעיל את החוק הראשון ואז את החוק השני במשך i פעמים, נקבל את המילה w, ולכן היא גם חלק משפת הדח״ה G.

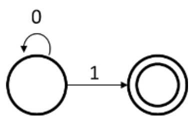
<u>שאלה 5</u>

<u>סעיף א</u>

השפה רגולרית, נראה שהשפה L היא המשלים של השפה $L_U = \{0^n 1 : n \geq 0\} \cup \{1^n 0 : n \geq 0\}$.

נוכיח ש- $L_1 = \{0^n 1 : n \in \mathbb{N}\}$ רגולרית, ובאופן דומה גם $L_2 = \{1^n 0 : n \in \mathbb{N}\}$ תהא רגולרית.

האסל״ד הבא מתאר את השפה L_1:



ההוכחה טריוויאלית ואני לא מבזבז את הזמן על זה.

משום שהראנו שיש ל-L_1 אסל״ד היא רגולרית, באופן דומה גם L_2 רגולרית, ומשום ששפות רגולריות סגורות תחת איחוד גם L_U רגולרית.

נניח בשלילה שיש $w \notin L$ וגם לא שייכת ל-L_U, אז המילה $w = a_1 a_2 \dots a_n$ כאשר $a_1, a_2 \dots, a_n$ אותיות של הא״ב. $w \notin L$ אז התו האחרון $a_n = {a_n}^R$ הוא לא תת מחרוזת של $a_1 a_2 \dots a_{n-1}$, משום שהא״ב הוא רק 0 ו-1, אז אם $a_n = 1$ החלק $a_1 a_2 \dots a_{n-1}$ מהווה רק אפסים, ואם $a_n = 0$ אז החלק $a_1 a_2 \dots a_{n-1}$ מהווה רק אחדים, וזוהי סתירה לכך שהמילה w לא שייכת ל-L_U.

נוכיח שכל מילה ב-L_U אינה ב-L, נניח בה״כ שהמילה היא מהצורה $0^n 1$, במצב הזה כל סיפא של המילה היא $0^k 1$ ושאר המילה היא רצף של $n - k$ אפסים, ובמצב הזה אין תת מחרוזת בתוך $0^k 1$ שמוכלת בתוך רצף אפסים.

משום ששפות רגולריות סגורות תחת משלים, והוכחנו שהשפה המשלימה של L היא L_U, קיבלנו ש-L רגולרית.

הדח״ה G הבא מתאר את השפה L:

1. $A \to 0A \mid 1A \mid \varepsilon$
2. $B \to 0B0 \mid 1B1 \mid C \mid \varepsilon$
3. $C \to 0C \mid 1C \mid \varepsilon$

נוכיח בעזרת הכלה דו כיוונית שמתקיים $L = L(G)$.

א. $L \subseteq L(G)$

תהי $w \in L$, כלומר עבור $a_1, a_2, \dots, a_n \in \Sigma$, מתקיים ש-$w = a_1 a_2 \dots a_n a_j a_{j-1} \dots a_i$ עבור $i < j$ טבעיים. מהדח"ה אפשר להגיע ל-w על ידי הפעלת החוק הראשון למשך $i - 1$ פעמים (A מייצג את החלק הראשון בו התת מחרוזת לא מוכלת), אח"כ הפעלת החוק השני למשך $j - i$ פעמים, ולבסוף הפעלת החוק השלישי למשך $n - j + 1$ פעמים. המילה שתצא תהא בדיוק $w = a_1 a_2 \dots a_n a_j a_{j-1} \dots a_i$ ולכן היא חלק משפת הדח"ה $L(G)$.

ב. $L \supseteq L(G)$

תהי $w \in L(G)$, בכל סדרת גזירות של w נצטרך לגזור כל נונטרמינל לפחות פעם אחת, החוקים על הטרמינל A מחייבים את המילה להתחיל ברצף כלשהו של 0 ו-1, המילה יכולה להסתיים שם כי המילה הריקה היא גם תת מחרוזת של כל מילה. אחר כך תהיה סדרת גזירות שלB שבעצם מבטאת את החלק בו x מופיע, על כל אות שאני מוסיף ל-w אני מוסיף את אותה האות "מהסוף" של המילה, וזה בעצם אומר שעד כה יהיה לנו רצף של אותיות עד אות מסויימת שממנה יהיה פלינדרום באורך זוגי uu^R עבור מילה כלשהי u תחת הא"ב. המילה יכולה להסתיים פה והיא עדיין תהא חלק מהשפה L כיוון שעד כה המילה שלנו מתחילה ברצף 0 ו-1 ואז יש פלינדרום באורך זוגי – וזה עונה על ההגדרה של L, נניח והמשכנו לגזור לפי החוק השלישי אז זה אומר שהוספנו עוד אותיות לחלק השמאלי של המילה וזה גם לא יפגע לנו בקריטריון של L, אז סה"כ המילה תהא גם חלק מהשפה L.

<u>סעיף ב</u>

השפה רגולרית, ההוכחה זהה להוכחה בסעיף הקודם, למעט החלק בון אנחנו משווים את התו האחרון $a_n = {a_n}^R$, אין לזה חשיבות אז ההוכחה זהה.

<u>סעיף ג</u>

השפה אינה חסרת הקשר, נוכיח בעזרת למת הניפוח לשפות חסרות הקשר.

נניח ש$L = \{a^n b^m c^{m \cdot n}\}$ חסרת הקשר, לפי למת הניפוח לשפות חסרות הקשר קיים n_0 כך שלכל $w \in L$ ו- $|w| \geq n_0$ קיימים 5 מילים כך ש- $w = uvxyz$ ומתקיימים שלושת תנאי הלמה:

1. $|vxy| \leq n_0$
2. $vy \neq \varepsilon$
3. עבור כל $k \geq 0$ מתקיים $uv^k xy^k z \in L$

נבחר $w = ab^{n_0}c^{n_0}$ אשר היא מילה בשפה L. נבחר $k = 0$ ונחלק את השרשור של vxy שני מקרים:

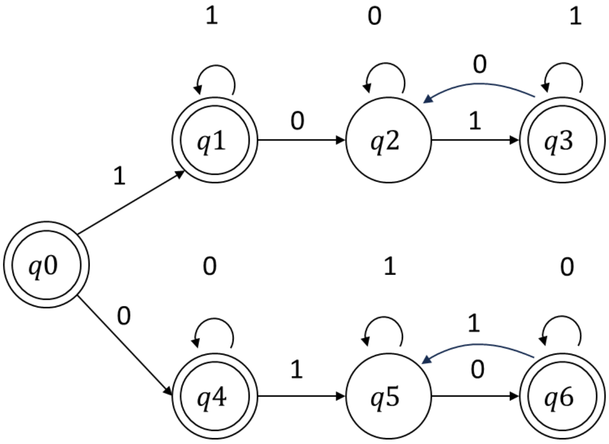
א. ב-vxy יש רק את האות b, בכזה מצב כשנפעיל את תנאי 3 עבור $k = 0$ נקבל $w' = ab^{n_0 - |y| - |v|}c^{n_0}$, והתנאי $n_0 - |y| - |v| = n_0$ צריך להתקיים על פי חוקי השפה ותנאי הלמה, אבל בגלל תנאי 2 של הלמה, השוויון לעולם לא יכול לקרות.

ב. ב-vxy יש את האות b ואת האות a הראשונה, בכזה מצב בבעת הניפוח כמות ה-a במילה תהיה 0, ולכן המכפלה אמורה להיות 0 ואז לא אמורות להיות c בכלל, אבל בגלל שבניפוח אנחנו לא נוגעים בחלק עם c יישארו כל ה-c והמילה לא תתקבל ב-L.

הגענו לסתירה ולכן L לא יכולה להיות חסרת הקשר.

<u>סעיף ד</u>

השפה רגולרית, נראה אס"ד שמקבל אותה.



הסבר:

בשפה L יהיו 2 מקרים בהם לא נרצה לקבל את המילה, כאשר מתישהו נקלט 10 ולא נקלט עדיין 01, או שמתישהו נקלט 01 ועוד לא נקלט 10.

במצב ההתחלתי כמות ה-10 וה-01 שווה 0 ולכן נקבל.

אם קיבלנו 1 נלך ל-q_1 (מצב מקבל, עדיין כמות ה-10 ו-01 היא 0) שמסמל את כך שהמילה מתחילה ברצף אחדים, ושאנחנו מחכים לראות אם נקבל 0 שיגרום לכך שיהיה לנו 10 במילה.

כשקיבלנו 0 לאחר רצף האחדים נלך למצב הבא q_2 (זה אומר שיש 10 בקלט עד כה) ונמשיך לקבל אפס עד שנראה 1 כדי להשלים ל01.

כשנקבל 1 ממצב q_2 נלך למצב q_3 שהוא מצב מקבל, והוא מסמל את כך שהיה לנו 10 ואחרי רצף 0-ים היה לנו גם 01, אז המילה תתקבלך במצב הזה. והחל מהמצב הזה תהיה לולאה בין q_2 ל-q_3 כדי להמשיך לקיים את התנאי שכמות ה-10 וה-01 יהיו שווים.

בצורה סימטרית ננהג כאשר המילה התחילה ברצף של 0-ים, המקרה הזה מיוצג בחלק התחתון של האס״ד.

שאלה 6

נבחר את: $L = \{a^m b^n c^l d^j e^m \mid n, m, l, j \in \mathbb{N}\}$.

עבור כל w שמקיים ש- $|w| \geq n_0$, קיים שרשור של מילים מהא״ב כך ש- $w = uvxyz$ ומתקיימים כל התנאים הבאים:

1. $|vxy| \leq n_0$
2. $vy \neq \varepsilon$
3. ולכל $k > 0$ מתקיים $uv^kxy^kz \in L$

לא משנה באיזה k נבחר, המילה המנופחת תהיה בשפה L כיוון שהיא תמשיך לענות על הקריטריונים של השפה.