# מודלים חישוביים – חזרה:

## מילים ושפות פורמליות – הגדרות ופעולות:

אורך של מילה – אורך של מילה w הוא מספר האותיות במילה , מסמנים |w| למשל: |abc|=3.
מילה ריקה – מילה שאין בה אותיות. באוטומט המקבל את המילה הריקה המצב ההתחלתי חייב להיות מצב מקבל. מסמנים את המילה הריקה: $\varepsilon$. (אפסילון), . אורך המילה הריקה = 0 אותיות.
שפה ריקה – שפה המכילה אף לא מילה אחת. שפה ריקה יכולה להתקבל למשל בחיתוך 2 שפות זרות זו לזו כלומר שאין אף לא מילה אחת משותפת. מסמנים: $\Phi$.
חשוב: מילה ריקה $\neq$ שפה ריקה.

## פעולות על מילים:

**שרשור שתי מילים** – מילה הנוצרת כאשר כותבים את שתי המילים בזו אחר זו משמאל לימין. דוגמא לשרשור: w1=ab , w2=abba המילה המשורשרת: ababba. מסמנים $w1 \cdot w2$.
**חשוב:** 1. שרשור עם מילה ריקה יוצרת מילה השווה למילה הלא ריקה. 2. יש חשיבות לסדר השרשור - $w1 \cdot w2$ =! $w2 \cdot w1$.
**חזקה** –סימון החזקה $w^n$ – שרשור המילה w n פעמים. שאלה: מהי המילה: $w^0$? המילה הריקה.
**היפוך של מילה** – מילה אשר נוצרת מהיפוך סדר האותיות במילה המקורית. דוגמא: w1=aaba המילה ההפוכה: abaa. מסמנים: R(w). היפוך של מילה ריקה = מילה ריקה.

## פעולות על שפות:

**שרשור על שפות** – שפה המכילה את כל אפשרויות השרשור של מילה מהשפה הראשונה עם מילה מהשפה השנייה. יש חשיבות לסדר השרשור. אם בשפה L1 יש x מילים ובשפה L2 יש y מילים אז סך הכל יכולים להיות לכל היותר: x*y מילים – אבל באופן כללי מתקיים: |L1L2|<x*y. כאשר באחת השפות יש אינסוף מילים אז שפת השרשור תהיה אינסופית גם כן, אלא אם כן אחת השפות היא ריקה.
**חזקה של שפה** – שפת השרשור של שפה L לעצמה n פעמים מסומנת כ - $L^n$.
חשוב: $L^0 = \{\varepsilon\}$. שפה בחזקת 0 היא לא השפה הריקה אלא מכילה מילה אחת-המילה הריקה וגם מתקיים: $L^{n+m}=L^nL^m$.
**היפוך של שפה** – שפה הנוצרת מהיפוך כל המילים בשפה. הסימון: R(L).

## תכונות של משפחת השפות הרגולריות:

משפחת השפות הרגולריות – אוסף כל השפות הרגולריות.
**שפה חלקית - הגדרה** – שפה L2 תיקרא שפה חלקית לשפה L1 אם כל המילים ב – L2 נמצאות גם ב – L1.
מסמנים: $L2 \leq L1$ .
**שפת המשלים - הגדרה** – L היא שפה מעל א"ב אז שפת המשלים ל – L היא השפה המכילה את כל המילים שלא נמצאות ב-L.טענה: נתון א"ב. אם L היא שפה רגולרית מעל הא"ב אז שפת כל המילים מעל א"ב שאינן ב-L (כלומר שפת המשלים) אף היא שפה רגולרית. מסמנים $\bar{L}$.
**שפת החיתוך – הגדרה** – בעבור שתי שפות L1 ו – L2 שפת החיתוך המסומנת ב - $L1 \cap L2$ היא השפה המכילה את כל המלים המשותפות לשתי השפות, כלומר היא מכילה את כל המילים שנמצאות גם ב- L1 וגם ב- L2.
דוגמאות:

1. L1 – כל המילים שמתחילות ב- a ומסתיימות ב- b . L2 – כל המילים שמתחילות ב- a. שפת החיתוך: כל המילים שמתחילות ב –a ומסתיימות ב – b.
2. $L1 = \{a^n b^n \mid n \geq 0\}$ ו - $L2 = \{b^n a^n \mid n > 1\}$ אז שפת החיתוך היא קבוצה ריקה.

דגשים לבניית אוטומט לקבלת שפת חיתוך (וגם)

1. מצב התחלתי יהיה שילוב של שני המצבים ההתחלתיים של שני האוטומטים.
2. מצב באוטומט שפת החיתוך יהיה מצב משולב של מצב מאוטומט A ומאוטומט B.
3. מצב $q_i p_j$ יהיה מצב מקבל רק אם $q_i$ ו$p_j$ היו שניהם מצבים מקבלים באוטומטים.
4. במידה והאות x עוברת ממצב $q_i$ למצב $p_j$ באוטומט A והאות x עוברת ממצב $q_n$ למצב $q_m$ באוטומט B אז באוטומט החדש האות x תעבור מהמצב $p_iq_n$ למצב $p_jq_m$.

**<u>שפת האיחוד – הגדרה</u>** – בעבור שתי שפות L1 ו – L2 שפת האיחוד המסומנת ב - $L1 \cup L2$ היא השפה המכילה את כל המלים הנמצאות בשתי השפות. כלומר, היא מכילה את כל המילים שנמצאות ב- L1 ואת כל המילים שנמצאות ב- L2. דוגמאות: L1 – כל המילים שמתחילות ב- a ומסתיימות ב- b . L2 – כל המילים שמתחילות ב- a. שפת האיחוד: כל המילים שמתחילות ב –a .

<u>דגשים לבניית אוטומט לקבלת שפת איחוד (או):</u>

1. מצב $q_i p_j$ יהיה מצב מקבל אם $q_i$ ו/או $p_j$ היו מצבים מקבלים.
2. לכל זוג מצבים משני האוטומטים האחרים יהיה מצב משולב.

**<u>טענות:</u>**

אם גם $L1$ וגם $L2$ הן שפות רגולריות אז גם :

$L1 \cap L2$ (חיתוך), $L1 \cup L2$ (איחוד), $L1 \bullet L2$ (שרשור), $R(L1)$ (היפוך), $\overline{L_1}$ (משלים) היא שפה רגולרית.

כלומר כל התכונות הללו הן תכונות סגירות.

**<u>תכונת סגירות</u>** – תכונה אשר מפעילים על שפה רגולרית ולאחריה עדיין קיימת שפה רגולרית.

**<u>שפה רגולרית</u>** – שפה שניתן לתאר אותה באמצעות אוטומט סופי דטרמיניסטי מלא.

<u>טענה א':</u> שפה סופית היא בהכרח שפה רגולרית. נימוק: ניתן לבנות לכל מילה מסלול משלה באוטומט ומאחר ומספר המילים בשפה סופי כך גם מספר המצבים באוטומט יהיה סופי.

<u>מסקנה</u>: שפה לא רגולרית תהיה בהכרח שפה אינסופית אך לא כל שפה אינסופית היא לא רגולרית.

<u>הוכחת אי רגולריות: </u>בכדי להוכיח כי שפה היא לא רגולרית יש להוכיח על פי תבנית ההוכחה.

<u>הוכחת רגולריות:</u> בכדי להוכיח כי שפה היא רגולרית ניתן לבנות לה אוטומט או להוכיח על ידי תכונות סגירות של משפחת השפות הרגולריות (שפת משלים, איחוד, חיתוך, שרשור, היפוך).

**<u>תרגילי חזרה:</u>**

**<u>שאלה 1: הוכחה אי רגולרית:</u>**

לפניך שפה מעל הא"ב {a,b,c}: $L = \{a^n b^m c^k \mid n,m,k > 0, k = 2n+3\}$ הוכח אם היא רגולרית או לא.

<u>הוכחה:</u>

השפה אינה רגולרית. נניח בשלילה שהיא רגולרית ולכן קיים אוטומט סופי A שמקבל אותה. נתבונן בקבוצת המילים האינסופית $W=\{a^nb \mid n>0\}$. נראה שעל כל מילה בקבוצה W האוטומט מגיע למצב שונה. נניח בשלילה שקיימות בקבוצה שתי מילים $w_i=a^ib$ ו- $w_j=a^jb$ כך ש- i≠j, i,j>0 והאוטומט A מגיע על $w_j$ ו- $w_i$ לאותו מצב q. נתבונן במילה $a^ibc^{2i+3}$. המילה שייכת לשפה ולכן האוטומט A מגיע עליה למצב מקבל. כלומר, קריאת הסיפא $c^{2i+3}$ מובילה את האוטומט ממצב q למצב מקבל. אבל, מכאן נובע שהאוטומט A מקבל גם את המילה $a^jbc^{2i+3}$ שאינה שייכת לשפה. הגענו לסתירה ולכן הנחת השלילה השניה אינה נכונה – כלומר, לא קיימות מילים $w_i$ ו-$w_j$ כמפורט לעיל ועל כל מילה בקבוצה W האוטומט מגיע למצב שונה. אך W אינסופית, ומכאן נובע של- A אינסוף מצבים בסתירה להיותו אוטומט סופי. לכן, גם הנחת השלילה הראשונה שגויה והשפה אינה רגולרית.

**<u>שאלה 2: הוכחת רגולריות:</u>**

תהי L שפת כל המילים מעל הא"ב {a,b,c} שאורכן לפחות 6, ובין 3 האותיות בהן מתחילה המילה אין שתי אותיות זהות ובין 3 האותיות בהן מסתיימת המילה אין שתי אותיות זהות. האם שפה זו רגולרית? הוכיחו את תשובתכם.

<u>פתרון</u>

ניתן להגדיר את L בעזרת שתי השפות הבאות:

$L_1$ היא שפת כל המילים מעל הא"ב {a,b,c} שאורכן בדיוק 3 ואין בהן שתי אותיות זהות.

$L_2$ היא שפת כל המילים מעל הא"ב {a,b,c}.

$L_1$ היא סופית ולכן רגולרית. $L_2$ אף היא רגולרית (אוטומט בעל מצב אחד מקבל בעל לולאה עצמית עבור אותיות הקלט {a,b,c}. השפה L: $L=L_1 \cdot L_2 \cdot L_1$. מסגירות משפחת השפות הרגולריות לשרשור גם $L_1 \cdot L_2$ רגולרית ולכן גם $L=(L_1 \cdot L_2) \cdot L_1$ רגולרית.

**<u>שאלה 3: הוכחת רגולריות:</u>**

תהי L שפת כל המילים מעל הא"ב {0,1,2}, שמתחילות במספר כלשהו (גדול ממש מ-0) של אפסים ואין בהן שתי אותיות 1 צמודות. האם L רגולרית? הוכח את תשובתך.

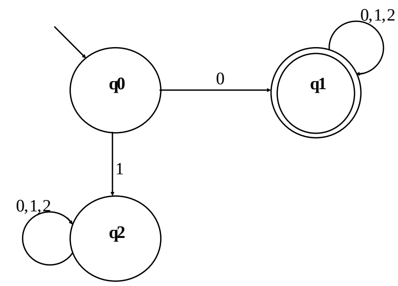
<u>פיתרון:</u>

נפתור בעיה זו בעזרת פירוק לשפות פשוטות יותר ושימוש בתכונות סגירות, אך גם פתרון ישיר, על ידי בניית אוטומט הוא אפשרי. ניתן לפרק את L כך: $L = L_1 \cap \overline{L_2}$, $L_1$ ו-$L_2$ מעל הא"ב {0,1,2}:

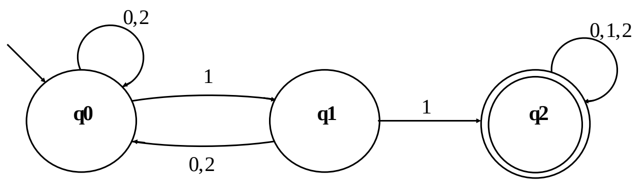
$L_1$={w | w מתחילה ברצף לא ריק של אפסים}

$L_2$={w | w מכילה רצף 11}

אוטומט עבור $L_1$:




אוטומט עבור $L_2$:




מסגירות משפחת השפות הרגולריות לפעולת המשלים גם $\overline{L_2}$ רגולרית,

ומסגירות משפחת השפות הרגולריות לפעולת החיתוך גם $L = L_1 \cap \overline{L_2}$ רגולרית.

**<u>שאלה 4:</u>**

נתונות שתי שפות מעל הא"ב {a,b,c}:

$L_1$ היא שפת כל המילים המתחילות ב-a.

$L_2$ היא שפת כל המילים המתחילות ב-b.

נתון כי $L_1$ ו-$L_2$ רגולריות. הוכח, תוך שימוש בנתון ובתכונות סגירות, כי $L_3$ אף היא רגולרית, כאשר $L_3$ היא שפת כל המילים מעל הא"ב {a,b,c} המתחילות ב-c.

**<u>פיתרון:</u>**

שפת כל המילים המתחילות ב-c מכילה בדיוק את כל המילים שאינן מתחילות ב-a וגם אינן מתחילות ב-b, וגם אינן המילה הריקה. כלומר $L_3 = \overline{L_1} \cap \overline{L_2} \cap \overline{\{\varepsilon\}}$. מאחר שנתון ש-$L_1$ ו-$L_2$ רגולריות, ומאחר ש-{ε} סופית, ולכן גם כן רגולרית, ומאחר שהשפות הרגולריות סגורות לפעולת המשלים הרי שגם $\overline{L_1}$, $\overline{L_2}$ ו- $\overline{\{\varepsilon\}}$ רגולריות ומסגירות משפחת השפות הרגולריות לפעולת החיתוך גם $\overline{L_1} \cap \overline{L_2}$ רגולרית וגם $L_3 = (\overline{L_1} \cap \overline{L_2}) \cap \overline{\{\varepsilon\}}$ רגולרית.

## <u>שאלה 5:</u>

נתונות שתי השפות T,L מעל הא"ב {a,b}.

L - אוסף המילים המתחילות ב- a.

T - אוסף המילים המתחילות ומסתיימות באותה אות, וכן המילה הריקה.

לפניך שלוש טענות. קבע לכל אחת מהן אם היא נכונה או לא <u>**ונמק**</u> קביעתך.

1. $ababaab \in (T \cdot \bar{T}) \cap L$
2. $T = T \cdot T$
3. $(L \cap T) \subseteq R(L)$

<u>**פיתרון:**</u>

1. $\bar{T}$ - אוסף כל המילים שמתחילות ומסתיימות באות שונה ללא המילה הריקה.

ולכן - $T \cdot \bar{T}$ - אוסף כל המילים או שמתחילות ומסתיימות באותה אות ובאמצע יש רצף של האות ההתחלתית והאות הסופית או כל המילים שמתחילות ומסתיימות באותיות שונות ובאמצע יש רצף של שתי אותיות זהות השוות לאות ההתחלתית. ולכן שפת החיתוך היא שפת כל המילים שמתחילות ב-a ונגמרות ב- a עם רצף ab באמצע או כל המילים שמתחילות ב – a ומסתיימות ב – b עם הרצף aa באמצע. לכן לפי ההגדרה המילה שייכת לחיתוך השפות. המילה ababaab שייכת ל- L מאחר והיא מתחילה באות a. המילה גם שייכת ל - $T \cdot \bar{T}$ כאשר: ababa שייכת ל – T והמילה ab שייכת ל - $\bar{T}$.

2. הטענה לא נכונה. דוגמא נגדית: נסתכל על המילה aba השייכת ל – T ועל המילה bab השייכת ל – T שרשור שתי המילים יוצרת מילה חדשה: ababab שלא שייכת על פי ההגדרה ל - T.
3. $L \cap T$ - אוסף כל המילים שמתחילות ב – a ומסתיימות ב – a. $R(L)$ - כל המילים שמסתיימות ב-a . לכן שפת החיתוך מוכלת בתוך שפת ההחיתוך מאחר וכל המילים שמתחילות ב-a ומסתיימות ב –a מהוות חלק מכל הקבוצה של המילים שמסתיימות ב –a. (כל מילה בשפת החיתוך נמצאת גם בשפת ההיפוך).

<u>**בהצלחה**</u>