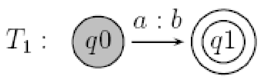
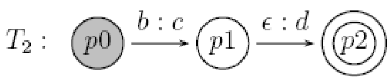
שיעור חזרה למבחן

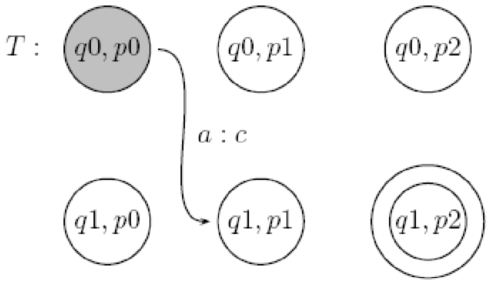
שאלה 1 :

The construction of composition given in the lecture slides is flawed. Consider the relations $R_1 = \{\langle a : b\rangle\}$ and $R_2 = \{\langle b : cd\rangle\}$, trivially defined by the following transducers T_1 and T_2, respectively:





Applying the definition given in class would yield the following transducer T for $R_1 \circ R_2$:



Clearly, the relation defined by T is empty, rather than the desired relation $\{\langle a : cd\rangle\}$.

1. Characterize the problem. Define precisely in what cases the given definition will not produced the correct results.

2. Solve the problem. Give a correct construction for transducer composition.

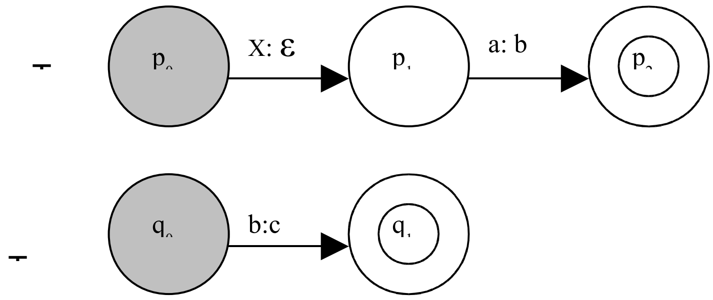
פתרון :

.1 הבעיה היא מעברי אפסילון מימין ב- T1 ומעברי אפסילון משמאל ב- T2 מעברים אלו לא ישתקפו ב composition transducer עפ"י ההגדרה שניתנה בכיתה.

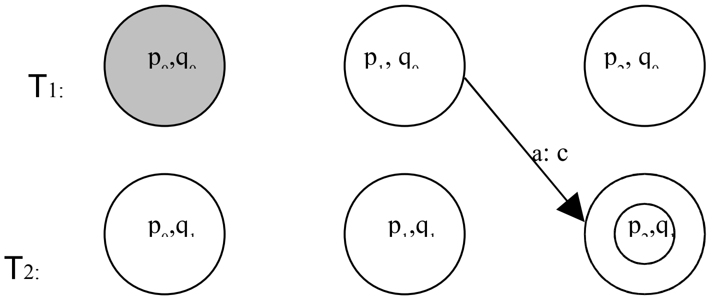
בדוגמא (בתרגיל), קיים מעבר אפסילון משמאל ב- T2 ולכן יש בעיה.

גם מעבר אפסילון מימין ב- T1 יהווה בעיה כלהלן:

Consider the relations R_1 = {<Xa : b>} and R_2 = {<b : c>}, trivially defined by the following transducers T_1 and T_2, respectively:



Applying the definition given in class would yield the following transducer T for $R_1 \circ R_2$:



בדוגמא הנ"ל קיימת אותה בעיה, T הוא ריק, כאשר צריך להיות R= {<Xa : c>}
לכן, אפיון הבעיה היא אם קיימים הבעיה היא מעברי אפסילון מימין ב- T_1 ומעברי אפסילון משמאל ב-T_2, ההגדרה שניתנה בכיתה לא תביא תוצאות נכונות.
(נשים לב כי מעברי אפסילון מימין ב- T_1 ומעברי אפסילון משמאל ב-T_2 כוללים גם מעברי אפסילון:אפסילון)

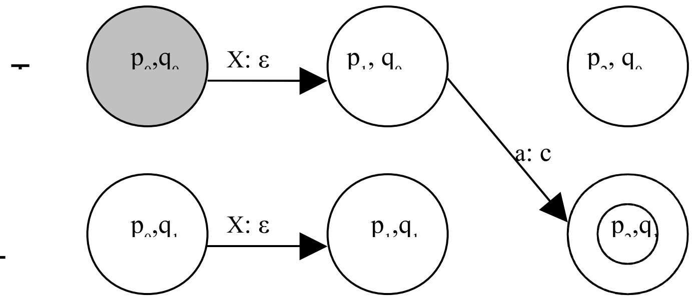
2. להלן ההגדרה שניתנה בכיתה:

```
.(Let T1 = (Q1, q1,∑1,∑2, δ1, F1) and T2 = (Q2, q2, ∑2, ∑3, δ2, F2
Then
T1∘ T2 = T = (Q1 × Q2, <q1, q2>,∑1, ∑3, δ, F1 × F2),
where δ (<s1, s2>, a, b) =
{<t1, t2> | for some c Є ∑2 U {ε}, t1 Є δ 1(s1, a, c) and
           t2 Є δ 2(s2, c, b)}
```

התיקון הנדרש:

```
.(Let T1 = (Q1, q1,∑1,∑2, δ1, F1) and T2 = (Q2, q2, ∑2, ∑3, δ2, F2
Then
T1∘ T2 = T = (Q1 × Q2, <q1, q2>,∑1, ∑3, δ, F1 × F2),
where δ (<s1, s2>, a, b) =
{<t1, t2> | for some c Є ∑2U{ε}, t1 Є δ 1(s1, a, c) and
           t2 Є δ 2(s2, c, b) }
and where δ (<s1, s2>, a, ε) =
{<t1, t2> | t1 Є δ 1(s1, a, ε), a Є ∑1U{ε} and t2 = s2 }
and where δ (<s1, s2>, ε , b) =
{<t1, t2> | t2 Є δ 2(s2, ε , b), b Є ∑3U{ε} and t1 = s1 }
```

התיקון הוא הוספת מעברי אפסילון כנדרש עפ"י מבנה T_2 , T_1.
בדוגמא הנ"ל הבניה החדשה תהיה:



והרלציה החדשה תהיה {<Xa:c>} כנדרש.

שאלה 2 :

Below are some sentences in Swahili:

/mtoto amefika/	"The child has arrived."
/mtoto anafika/	"The child is arriving."
/mtoto atafika/	"The child will arrive."
/watoto wamefika/	"The children have arrived."
/watoto wanafika/	"The children are arriving."
/watoto watafika/	"The children will arrive."
/mtu amelala/	"The man has slept."
/mtu analala/	"The man is sleeping."
/mtu atalala/	"The man will sleep."
/watu wamelala/	"The men have slept."
/watu wanalala/	"The men are sleeping."
/watu watalala/	"The men will sleep."
/kisu kimeanguka/	"The knife has fallen."
/kisu kinaanguka/	"The knife is falling."
/kisu kitaanguka/	"The knife will fall."
/visu vimeanguka/	"The knives have fallen."
/visu vinaanguka/	"The knives are falling."
/visu vitaanguka/	"The knives will fall."
/kikapu kimeanguka/	"The basket has fallen."
/kikapu kinaanguka/	"The basket is falling."
/kikapu kitaanguka/	"The basket will fall."
/vikapu vimeanguka/	"The baskets have fallen."
/vikapu vinaanguka/	"The baskets are falling."
/vikapu vitaanguka/	"The baskets will fall."

Based on the above examples,

1. Construct a basic dictionary for Swahili, including entries for "child", "man", "knife", "basket", "arrive", "sleep" and "fall".

2. Identify bound morphemes in Swahili. For each morpheme provide both the form and the meaning, and determine whether it is a prefix, a suffix, an infix or a circumfix. State also whether the morpheme is inflectional or derivational.

3. Specify the rules that govern the combination of the bound morphemes you identified with the entries of the dictionary.

4. Implement the rules in XFST. You grammar should generate all the Swahili forms above, assigning them an adequate morphological structure. For example, the lexical string associated with the surface form "mtoto amefika" can be "`child-sg+arrive-sg-perfect`", whereas "vikapu vitaanguka" can be associated with "`basket-pl+fall-future`".

Submit an XFST script called `swahili` with your XFST code, where the answers for 1–3 are clearly specified (use comments if needed).

שאלה 3 :

Let L be a regular language, and let w^r denote the reverse of the word w. Define R to be the relation:

$$R = \{\langle w, w^r\rangle \mid w \in L\}$$

Is R a regular relation? If it is, provide an XFST expression which defines it; if it isn't, prove.

Submit a text file called `r` with your answer.

שאלה 4 :

Design and implement a unification grammar for simple Hebrew sentences such as the following (use a straight-forward transliteration):

הילדים הלכו לגן
הילדים הגדולים הלכו לגן

ילדים גדולים משחקים עם ילדות
כל הילדים אוהבים גלידה
רוב הילדים הגדולים משחקים
כל ילד שיחק עם ילדה גדולה

Begin by designing a context-free grammar for these and similar sentences. Then, augment the grammar with feature structures and unification equations to enforce further constraints. Explain and justify your choice of features and values.

Implement the grammar in PC-PATR; the software and its documentation are available from `http://www.sil.org/pcpatr/`. Submit:

- The full code of the grammar, including the lexicon;
- Printouts of the trees assigned by the grammar to the example sentences above;
- A list of sentences you used to check the grammar, for which the grammar produces plausible trees (no need to submit printouts);
- Any additional documentation recording the considerations you applied when designing the grammars.

שאלה 5 :

Design a unification grammar for the language $\{a^{2^n} \mid n > 0\}$. Implement the grammar in PC-PATR. Submit only the grammar and lexicon files.

PC-PATR - ניתן להוריד באופן חופשי מאת : http://www.sil.org/pcpatr

מנתח תחבירי עבור דקדוקים חסרי הקשר ודקדוקי האחדה.

התוכנה קולטת דקדוק חסר הקשר בשני קבצים – קובץ הלקסיקון וקובץ הדקדוק, ולאחר מכן מבצעת ניתוח למחרוזות שונות לבחירתנו כולל מתן כל עצי הגזירה האפשריים.