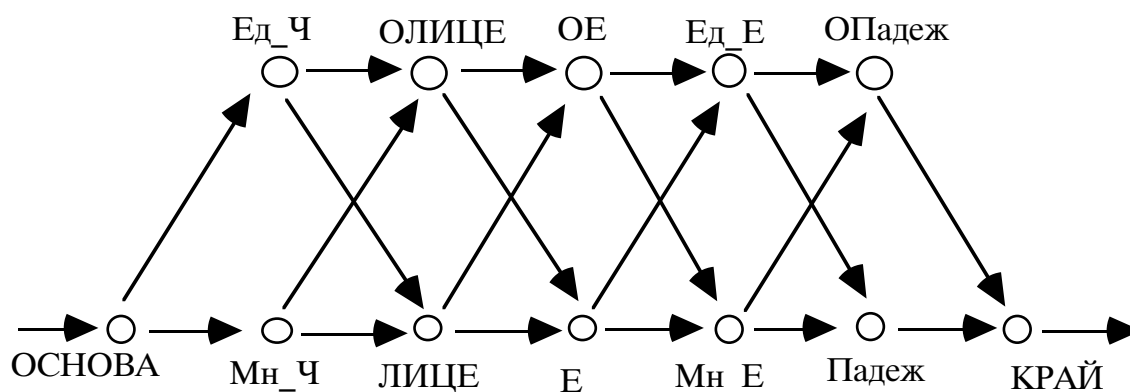


Задача.

Даден е краен автомат (finite state automaton; а когато устройството пише, то се нарича по-точно finite state transducer - понеже променя разпознавания низ или синтезира), който моделира обобщено образуването на:

- формите на съществителните нарицателни в унгарския език и
- формите на някои групи на съществителните нарицателни имена, които се срещат в изреченията на унгарски език.



Състоянията на автомата са наименовани с морфологични категории, които могат да участват в рамките на една “дума” (под дума ще разбираме буквен низ между два интервала, който е форма на съществително нарицателно в изречението). Дъгите показват начина на преход от едно състояние в друго, като при това се оформя “скелета” на словоформата според възможните градивни части. При преход по дъга към подадения на входа низ се долепя някаква морфема. Морфологичните категории са следните:

ОСНОВА - отбелязва началното състояние. На входа се подава основа на съществително нарицателно. На изхода на автомата (състояние КРАЙ) се получава възможна форма.

ЧИСЛО - означение на категорията “число”. В унгарския език има две числа - единствено и множествено, които са отразени като две състояния. При минаване през състоянието Мн_Ч към основата се долепят определени наставки, а при минаване през състоянието Ед_Ч се долепя празният низ. В рамките на една словоформа могат да се споменават неколккратно елементи, означаващи категорията “число”, като те се отнасят винаги към непосредствено предхождащата ги категория. По-долу с ОСНОВА + Ед_Ч ще означаваме единствено число на основата, а с ОСНОВА + Мн_Ч - множественото число на основата.

ЛИЦЕ - категория на лицето. Означава лицето на притежателя, ако той е посочен. Чрез тази категория в унгарския език се реализират формите, които на български език съставяме с притежателни местоимения - мой, твой и т.н. Минаване през състоянието на категорията ЛИЦЕ води до прибаване към подадения низ на наставка, означаваща една от следните шест категории: 1_Ед, 2_Ед, 3_Ед, 1_Мн, 2_Мн, 3_Мн. Състоянието \emptyset ЛИЦЕ означава отсъствие на категорията в конкретната форма. Например за основата *könyv* (книга) имаме:

- *könyv* + Ед_Ч + ЛИЦЕ = *könyv*;
- *könyv* + Мн_Ч + ЛИЦЕ = *könyvek* (книги);
- *könyv* + Ед_Ч + 1_Ед = *könyvem* (моята книга,
-em е наставка за 1_Ед);
- *könyv* + Ед_Ч + 1_Мн = *könyvünk* (нашата книга,
-ünk е наставка за 1_Мн).

ПРИТЕЖАНИЕ - в езика се реализира чрез наставка -é, а в автоматата може да приеме две стойности; със състоянието “Е” означаваме наличие на категория за притежание, а с \emptyset Е - отсъствие на такава категория в словоформата. Например:

- *könyv* + Ед_Ч + \emptyset ЛИЦЕ + Е + Ед_Е = *könyvé* (/това/ на книгата, такава форма може да се използва в изречението “Хартията на списанието е хубава, а тази на книгата - не.”);

Елементът за притежание също може да бъде поставян в множествено число, това са двете категории Ед_Е и Мн_Е; Ед_Е не предизвиква долепяне на наставка, а Мн_Е води до долепяне на наставката -i. Например,

- *könyv* + Ед_Ч + \emptyset ЛИЦЕ + Е + Мн_Е = *könyvéi* (тези на книгата).

ПАДЕЖ - категория за падежа на съществителното нарицателно. Тя може да приеме около 17 типа стойности. Например: *könyvvel* - с книгата, *könyvön* - върху книгата, *könyvre* - в книгата и т.н.

Очевидно автоматът произвежда доста сложни форми. Но на практика това е само “идеалният скелет” за синтез на някои групи на съществителните. В реалния език има изменения на основата при долепяне на различните наставки. Броят на наставките е много голям - от порядъка на 700 за съществителните. Освен това, в отделни форми са възможни цикли между категорията за притежание и нейното число.

Задание 1. Дадени са унгарски думи и техните преводи на български език:

községében, gépünkért, velünk, géped, érted, közséért, tolem, tuztol: с нас, за нашата машина, за тебе, от огъня, в твоето село, твоята машина, от мене, за селото.

Установете какъв превод съответства на всяка дума и обяснете своето решение. Ако някои от тези думи са получени по горния краен автомат, покажете как са синтезирани.

Задание 2. Преведете от български на унгарски език:
в тези от твоите книги, това на книгите.
Обяснете начина на синтез.

Задание 3. Преведете от унгарски на български език:
könyvünké, könyveméi, községekért.
Обяснете как са синтезирани тези форми на унгарски език.

Забележка: Задачата е построена по статията
Ференц Пап. Автоматично генериране на унгарски словоформи. Сп.
"Компютърна лингвистика", Дебрецен, 1976 (на английски език).
(Давана е на лингвистични олимпиади и е използвана в задочната школа
по математическа лингвистика през 1982-84 г).