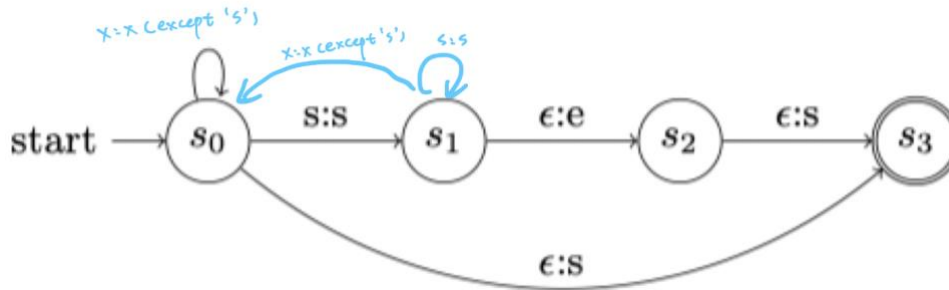


EX1: Finite State Transducers

Jing Huang (933-039-277)

huanjing@oregonstate.edu

1. An FST for Pluralization



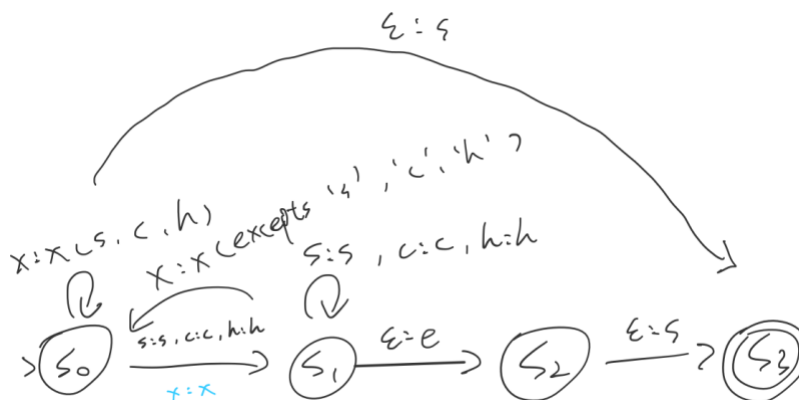
Test Case	Result
bus	Input line 1: b u s (9 states / 8 arcs reduce-> 6/5) b u s e s 1
bass	Input line 1: b a s s (12 states / 11 arcs reduce-> 7/6) b a s s e s 1
sass	Input line 1: s a s s (13 states / 12 arcs reduce-> 7/6) s a s s e s 1
rise	Input line 1: r i s e (11 states / 10 arcs reduce-> 6/5) r i s e s 1

I fixed it by adding a loop at state s1 which is triggered by 's'. And also created a path to state s0 from s1 to let through all characters with the exception of 's'. Besides, the loop at state s0 is also changed to pass all characters except 's'. All test results show that this FST can handle inputs ending with 's'.

2. Other Pluralizations

For regular plurals as shown in the table below, all of them can be covered by the above FST except ‘-sh’ and ‘-ch’. As a result, the loop at state s_1 changed to be triggered by ‘s’, ‘c’ and ‘h’, the loop at state s_0 changed to not be triggered by them, and the path to state s_0 from s_1 changed to all characters with the exception of these three characters.

Regular plurals		
A singular noun ends in a sibilant sound—/s/, /z/, /ʒ/, /ʒ/, /tʃ/ or /dʒ/		
	kiss	kisses
	phase	phases
	dish	dishes
	massage	massages
	witch	witches
A singular form ends in a voiceless consonant (other than a sibilant)—/p/, /t/, /k/, /f/ (sometimes) or /θ/		
	lap	laps
	cat	cats
	clock	clocks
	cuff	cuffs
	death	deaths
For all other words (i.e. words ending in vowels or voiced non-sibilants) the regular plural adds /z/, represented orthographically by -s:		
	boy	boys
	girl	girls
	chair	chairs



Plurals of nouns in -o

With nouns ending in o preceded by a consonant, the plural in many cases is spelled by adding -es (pronounced /z/):

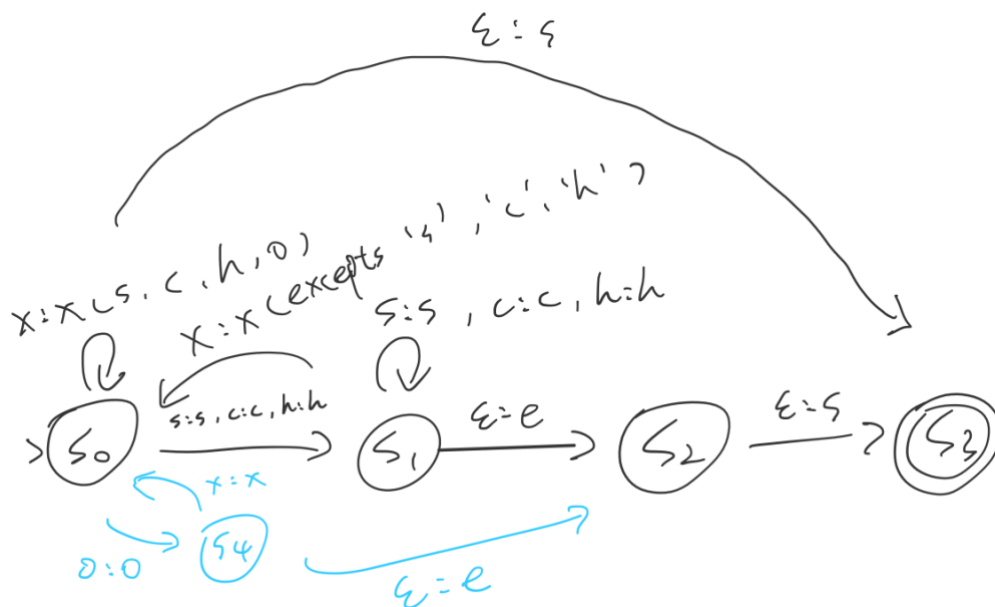
hero	heroes
potato	potatoes
volcano	volcanoes or volcanos

However many nouns of foreign origin, including almost all Italian loanwords, add only -s:

canto	cantos
hetero	heteros
photo	photos

For plurals of nouns in -o as shown in the table above, the nouns ending in o preceded by a consonant, I added a path as the picture below.

However, the nouns of foreign origin might need individual path for each noun. I'm looking forward to finding a better solution than this.



Plurals of nouns in -y

Nouns ending in a vocalic y (that is, used as a vowel) preceded by a consonant usually drop the y and add -ies (pronounced /iz/, or /aiz/ in words where the y is pronounced /ai/):

cherry	cherries
lady	ladies
sky	skies

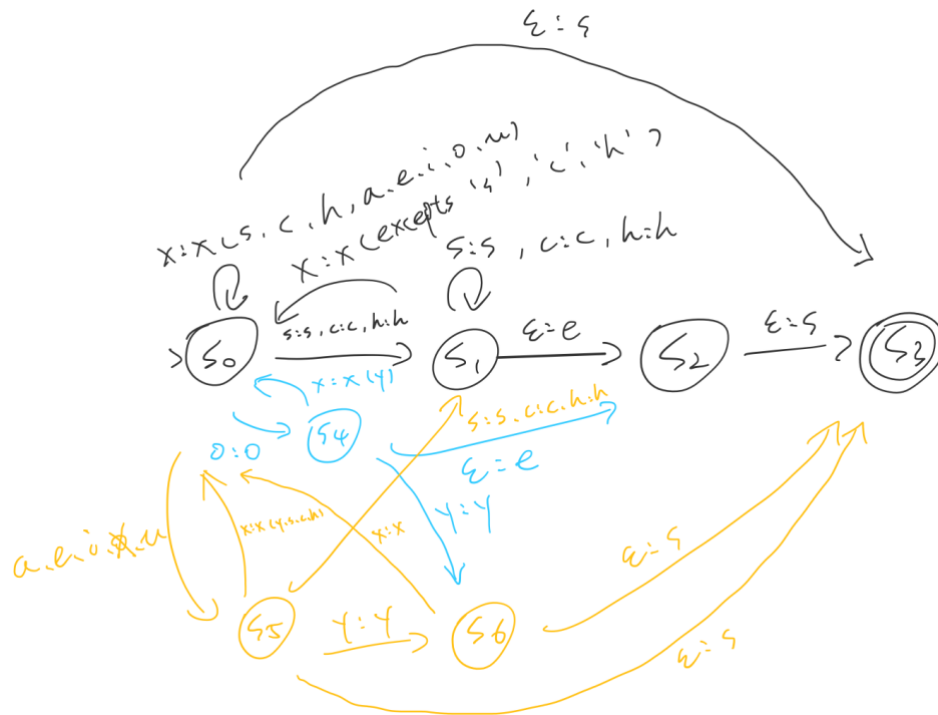
Words ending in quy also follow this pattern:

soliloquy	soliloquies
-----------	-------------

Words ending in a y preceded by a vowel form their plurals by adding -s:

day	days
monkey	monkeys

For words ending in a y preceded by a vowel, I add a path to s5 which is triggered by vowels excepts 'o'. Besides, if 'y' or a vowel is not the end of the word, the FST needs to go back to the initial state, and to compatible with 's', 'c', 'h', the FST needs move to s1 from s5 if the character is 's', 'c' or 'h'. Otherwise, it ends with 's'. And if the input of s4 is 'y', then the FST needs to jump to s6.



As for the noun ending in a vocalic y, I think it could be written by adding a non-vowel path and replacing 'y' with 'ies' if y is the ending. But I need to find out how to write backspace in FST.

Test Case	Result
boy	<pre> Input line 1: b o y (8 states / 7 arcs reduce-> 5/4) b o y s 1 0 </pre>
cat	<pre> Input line 1: c a t (8 states / 7 arcs reduce-> 5/4) c a t s 1 0 </pre>
hero	<pre> Input line 1: h e r o (12 states / 11 arcs reduce-> 7/6) h e r o e s 1 0 </pre>
day	<pre> Input line 1: d a y (8 states / 7 arcs reduce-> 5/4) d a y s 1 0 </pre>
apple	<pre> Input line 1: a p p l e (12 states / 11 arcs reduce-> 7/6) a p p l e s 1 0 </pre>