Assignment 2

L445 / L545

Due Wednesday, February 8

1. Consider the following Japanese data:[1]

tabeta	'X ate Y'
aketa	'X opened Y'
tabesaseta	'X made Y eat Z'
akesaseta	'X made Y open Z'
taberareta	'X was eaten'
akerareta	'X was opened'
tabesaserareta	'X was made to eat Y'
akesaserareta	'X was made to open Y'
tabesasenai	'X doesn't/won't make Y eat Z'
tabenai	'X doesn't/won't eat Y'
tabesaserareru	'X is/will be made to eat Y'

   (a) Based on this data, give the Japanese morphemes for the following English translations:
      i. 'open'
      ii. 'eat'
      iii. passive (+PASS) marker ('... be VERB-ed', e.g., 'They were opened/eaten')
      iv. causative marker (+CAUS) ('... make X VERB', e.g., 'Robin makes Tracey laugh')
      v. non-past marker (-PST) (present or future tense)
      vi. past marker (+PST)
      vii. negative marker (+NEG)
   (b) Provide (i.e., properly define) a finite-state transducer (FST) to recognize and analyze these words

2. Do question 3.3 on p. 81 from Jurafsky & Martin, dealing with K insertion in English. (This refers to the chart on p. 63 in section 3.6, which defines K insertion as "verbs ending with *vowel* + *-c* add *-k*", e.g., *panic/panicked*)

3. Draw a FST which accounts for the following spelling rules, where the input is a morphotactic form and the output is an actual English word:

   (1) y → ie / C _ ^s#
   (2) y → i / C _ ^ed#

   where C stands for any consonant.

   You do not need to worry about interactions with other spelling rules.

4. Do question 3.5 on p. 81-82 of Jurafsky & Martin, involving the Soundex algorithm.

[1] Question adapted from *The Language Files*.