TP2: Pronunciation Dictionary

Nov. 12, 2012

**Context:**

Pronunciation dictionary is one of the major components in language processing. It plays an important role in mapping between signal unit and phonetic symbol.

We can create two types of pronunciation dictionary:

1. Phonetic based dictionary
2. Grapheme base dictionary

E.g.: សាលា could be written: *សាលា​ /s aa l aa /* as a phonetic based OR *សាលា /Sa AA Lo AA/* as a grapheme based.​

Phonetic based dictionary is difficult to create since it required many linguistic rules and knowledge and also the variation of the word pronunciation. The better way is to create a grapheme based dictionary which is simple to create.

**To do:**

Write a program to create a grapheme based dictionary. The program needs **vocabulary** and **converted rules** as input and the **grapheme dictionary** as output.

> java generateGraphemDico in\_vocab in\_convert\_rule out\_dico\_grapheme