28.02.2017

METL – TP1

Martino Ferrari

Exercise 1

In this first exercise we were asked to translate 3 sentences from Arcturan to Centauri, two fictionary languages.
To do so we were provided of a simple Arcturan-Centauri “rosetta stone”, where few sentences were written in both languages, as well as a monolingual Centauri text.
To translate the sentneces I write a simple and naive Python script that translate the Arcturan in Centauri using both monogram and bigrams.
However I did not programmed a full dynamic (Viterbi-like) algorithm but a much more basic logic, where I take in account only the more probable translation, and so I do not use any back-pointer, you can find the code in *translator.py*.
The results of the script are the following:

1 A: iat lat pippat eneat hilat oloat at-yurp
C: lalok brok anok enemok ghirok kantok ok-yurp

2 A: totat nnat forat arrat mat bat
C: wiwok rarok nok crrrok yorok ghirok

3 A: wat dat quat cat uskrat at-drubel
C: lalok sprok izok stok ? ok-drubel

Note that in sentence 3 there is a word never translated in the “Rosetta stone”.
The more difficult part of the exercise is the verification as the sense and the grammar of the two languages are unknown, the only verification I did was to pass a translated sentence to the “translator” and look at the resut:

4 A: wat nnat forat arrat vat ga
C: lalok rarok nok izok hihok mok

The translated sentence match the original translation.

Exercise 2

In this second exercise we were asked to implement the EM algorithm for a simple head or tails problem. You can find the code in *em_coins.py*.