

Homework #1: CMPT-825

Anoop Sarkar – anoop@cs.sfu.ca

- (1) NASA's Exploration Rover Opportunity has found some strange tablets during its travels in the Meridiani plains on Mars. Apparently one tablet is a kind of Rashid (Rosetta) stone which has translations from a language we will call MARTIAN-A to another language we will call MARTIAN-B. The ASCII transcription of the alien script on the Rashid tablet based on the photographs transmitted back to Earth is given below:

- | | |
|------------------------------------------------|-----------------------------------------|
| 1a. ok'sifar zvau hu . | 8a. ked bzayr myi pell eoq . |
| 1b. at'sifar somuds geyu . | 8b. gakh up ashi erder kvig . |
| 2a. ok'anko ok'sifar myi pell hu . | 9a. yux eoq qebb zada ok'nefos . |
| 2b. at'anko at'sifar ashi erder geyu . | 9b. diza kvig pai goli at'nefos . |
| 3a. oprashyo hu qebb yuzvo oxloyzo . | 10a. ked amn eoq kin oxloyzo hom . |
| 3b. diza geyu isvat iwla pown . | 10b. dimbe kvig baz iluh ejuo pown . |
| 4a. ok'sifar myi rig bzayr zu . | 11a. ked eoq tazih yuzvo kin dabal'ok . |
| 4b. at'sifar keerat ashi parq up . | 11b. dimbe kvig isvat iluh dabal'at . |
| 5a. yux druh qebb stovokor . | 12a. ked mina eoq qebb yuzvo amn . |
| 5b. diza viodaws pai shun . | 12b. dimbe kvig zeg isvat iwla baz . |
| 6a. ked hu qebb zu stovokor . | |
| 6b. dimbe geyu keerat pai shun . | |
| 7a. ked druh zvau ked hu qebb pnah . | |
| 7b. dimbe viodaws somuds dimbe geyu iwla woq . | |

Due to severe budget cutbacks at NASA, decryption of these tablets has fallen to various Canadian graduate students, namely you.

- a. Your first task is to examine the above translations and produce a translation dictionary. For each word in MARTIAN-A provide an equivalent word in MARTIAN-B.

- b. Using your translation dictionary, provide a word for word translation for the following MARTIAN-B sentences found on a new tablet near the Rashid tablet above.

13b. gakh up ashi woq pown goli at'nefos .

14b. diza kvig zeg isvat iluh ejuo .

15b. dimbe geyu pai shun hunslob at'anko .

The MARTIAN-A sentences you produce will probably appear to be in a different word order from the MARTIAN-A sentences you observed on the Rashid tablet. Some words might be unseen and so seemingly untranslatable.

- c. The word for word translation can be improved with additional knowledge about MARTIAN-A word order. Luckily another tablet containing some MARTIAN-A sentences (untranslated) was found on the dusty plains of Mars. Use these MARTIAN-A sentences in order to find the most plausible word order for the MARTIAN-A sentences translated from MARTIAN-B sentences in (1b).

ok'anko myi oxloyzo druh .

yux mina eoq esky oxloyzo pnah .

ok'anko yolk stovokor koos oprashyo pnah zada ok'nefos yun zu kin hom .

ked hom qebb koos ok'anko .

ok'sifar zvau hu .

ok'anko ok'sifar

myi pell hu .

oprashyo hu qebb yuzvo oxloyzo .

ok'sifar myi rig bzayr zu .

yux druh qebb stovokor .

ked hu qebb zu stovokor .

ked bzayr myi pell eoq .

ked druh zvau ked hu qebb pnah .

yux eoq qebb zada ok'nefos .

ked amn eoq kin oxloyzo hom .

ked eoq tazih yuzvo kin dabal'ok .

ked mina eoq qebb yuzvo amn .

Using this additional MARTIAN-A text you can even find a translation for words that are missing from the translation dictionary.

This exercise is derived from an exposition by a famous computational linguist and statistical MT pioneer. However, some of the details have been changed and the reference is intentionally obfuscated so that the temptation of using google will not spoil the fun of working through this example without external help.