**CMPE 590 SP.TOP. IN CMPE: MACHINE TRANSLATION**  22.05.2017

**FINAL**

1. We have a phrase-based MT system that uses the standard phrase-based translation model (i.e. the model having three components). The phrase translation table is shown below. The system is translating the sentence “*kitap bir ihtiyaçtır*”.

|  |  |  |
| --- | --- | --- |
| **f** | **e** | **ϕ(f|e)** |
| kitap | book | 0.60 |
|  | books | 0.30 |
|  | this book | 0.10 |
| bir | one | 0.50 |
|  | a | 0.50 |
| ihtiyaçtır | necessity | 0.20 |
|  | is necessity | 0.80 |
| kitap bir | a book | 0.40 |
|  | one book | 0.60 |
| bir ihtiyaçtır | is a necessity | 1.00 |

1. Beginning from the empty hypothesis, show the decoding process on a tree structure. You need to take into account the complete search space without using any heuristics or pruning. Show a large portion of the search space by going 2-3 levels down in the tree. For each node (hypothesis) in the tree, calculate the partial translation score.
2. Instead of generating the complete search space, do the decoding process using beam search. Group the hypotheses wrt the number of input words translated and use histogram pruning with “beam size (max. number of hypothesis in a group) = 3”. Show the process in a proper format.
3. In an example-based MT system, the database contains the following three aligned sentences:

*Otobüste giderken kitap okumayı unuttu – He forgot reading book while traveling in bus*

*Tarih önemlidir – History is important*

*Ahmet futbolu severdi – Ahmet liked football*

The system tries to translate the sentence *Ahmet otobüste giderken tarihi kitaplar okumayı sever*. By using the EBMT method covered in the lectures, show the translation step by step. Show all the necessary information used in the translation (subsentential alignments, dependency trees, matching expressions, etc.). Also calculate a goodness score for your proposed translation.

1. We have the following translation output

*several videos is showing us how to carried out our programme exercises correctly*

and the following two reference translations

*various videos indicate us how to properly perform our workout plan in this manner*

*several videos show how to properly plan your exercise*

1. Calculate precision, recall, and f-measure between system translation and reference translation. (Calculate for each reference separately.)
2. Calculate BLEU-n (for n=1,…,3) assuming that we have only the first reference.
3. Calculate BLEU-n (for n=1,…,3) using both references.
4. Calculate METEOR score.

(The word pairs that are synonyms: several-various, show-indicate, showing-indicate, correctly-properly, programme-plan.)

(continued on next page)

*Notes:*

* Questions 1-3 : 35 points
* Time: 1:50 hours
* Close notes/books

ne olur kim olduğunu bilsem pia’nın

ellerini bir tutsam ölsem

böyle uzak uzak seslenmese

ben bir şehre geldiğim vakit

o başka bir şehre gitmese

otelleri bomboş bulmasam

içlenip buzlu bir kadeh gibi

buğulanıp buğulanıp durmasam

ne olur sabaha karşı rıhtımda

çocuklar pia’yı görseler

bana haber salsalar bilsem

içimi büsbütün yıldız basar

bir hançer gibi çıkıp giderdim

ben bir şehre geldiğim vakit

o başka bir şehre gitmese

singapur yolunda demeseler

bana bunu yapmasalar yorgunum

üstelik parasızım pasaportsuzum

ne olur sabaha karşı rıhtımda

seslendiğini duysam pia’nın

sırtında yoksul bir yağmurluk

çocuk gözleri büyük büyük

üşümüş ürpermiş soluk

ellerini tutabilsem pia’nın

ölsem eksiksiz ölürdüm