



DUBLIN CITY UNIVERSITY

SEMESTER 2 EXAMINATIONS 2014/2015

MODULE: CA446 – Statistical Machine Translation

PROGRAMME(S):
CASE BSc in Computer Applications (Sft.Eng.)
MTT MSc in Translation Technology

YEAR OF STUDY: 1,4

EXAMINERS:
Professor Qun Liu (Ext:5644)
Dr. Ian Pitt

TIME ALLOWED: 2 Hours

INSTRUCTIONS: Answer Question One and two other questions.

PLEASE DO NOT TURN OVER THIS PAGE UNTIL YOU ARE INSTRUCTED TO DO SO

The use of programmable or text storing calculators is expressly forbidden.

Please note that where a candidate answers more than the required number of questions, the examiner will mark all questions attempted and then select the highest scoring ones.

Requirements for this paper (Please mark (X) as appropriate)

<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>

Log Tables
Graph Paper
Dictionaries
Statistical Tables

<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>

Thermodynamic Tables
Actuarial Tables
MCQ Only – Do not publish
Attached Answer Sheet

QUESTION 1**[TOTAL MARKS: 40]**

Answer EIGHT of the following ten short questions. Each question is worth 5 marks.

1. Briefly explain the following two uses of machine translation: 1) assimilation, 2) dissemination, and the requirement of the speed and quality of machine translation systems in these two use cases.
2. Explain why monolingual corpus of target language is useful for statistical machine translation.
3. Explain the advantages of log-linear model over noisy channel model.
4. Why should n-gram matching be clipped when we calculate the n-gram precision in BLEU score?
5. Give the minimal edit operations which are needed to edit the sentence "John gave Mary a book" to "John handed a book to Mary"
6. Explain interpolation smoothing of an n-gram language model.
7. Explain the motivation of language model smoothing. What's the advantage of add-alpha smoothing over add-one smoothing?
8. Give two techniques which is used to manage very large data for language modelling
9. Name the three components of the noisy channel model and tell which components are related to fluency and which to adequacy.
10. Why do IBM models not support two or more source words being aligned to the same target word?

[End of Question 1]

QUESTION 2

[TOTAL MARKS: 30]

Q 2(a)

[16 Marks]

Given the following pairs:

S_1	S_2
<i>el perro</i> <i>dog</i>	<i>la perro</i> <i>the dog</i>

State what the following translation probabilities will be after two iterations of the Expectation Maximisation algorithm and show all the steps followed to arrive at these values:

$t(the|el)$
 $t(dog|el)$
 $t(the|la)$
 $t(dog|la)$
 $t(the|perro)$
 $t(dog|perro)$

Assume that only one-to-one and one-to-zero patterns are allowed for source-to-target word alignments in these sentence pairs.

Q 2(b)

[8 Marks]

Considering the same sentence pairs as above, assuming each target word is exactly aligned with one source word, please list all the possible word alignments for the two sentence pairs.

Q 2(c)

[6 Marks]

List all phrase pairs that are consistent with the following word alignment:

	A	B	C
x			
y			
z			

[End of Question 2]

QUESTION 3**[TOTAL MARKS: 30]****Q 3(a)****[15 Marks]**

Given the following sentences:

<s> John loves Mary </s>

<s> Mary loves Tom </s>

<s> Mary hates John </s>

List all the parameters of the unigram model and the bigram language models trained with these sentences without smoothing.

Q 3(b)**[15 Marks]**

Calculate the probabilities of the following sentences:

<s> Tom loves Mary </s>

<s> Tom hates Mary </s>

Using add-alpha smoothing where: $\alpha = 0.2$.

[End of Question 3]

QUESTION 4**[TOTAL MARKS: 30]****Q 4(a)****[14 Marks]**

Given:

Source sentence:

Wo yao mai yizhang xingqitian qu shanghai de jipian

Hypotheses:

C1: Sunday I want to buy a ticket to Shanghai

C2: I would like to buy a Sunday go to Shanghai air ticket

References:

R1: I want to buy an air ticket to Shanghai on Sunday

R2: I would like to buy a ticket to Shanghai on Sunday

R3: A Sunday ticket to Shanghai please

Calculate the BLEU scores of the two candidate translations (hypotheses).

Q 4(b)

[10 Marks]

Prove that the noisy channel model is a special case of the log-linear model.

Q 4(c)

[6 Marks]

Explain the difference between a model error and a search error in statistical machine translation.

[End of Question 4]

[END OF EXAM]