

CA446

Statistical Machine Translation



Week 10: Exam

Lecturer: Qun Liu

Lab Tutor: Xiaofeng Wu, Iacer Calixto

2nd Semester, 2014-2015 Academic Year

<http://computing.dcu.ie/~qliu/CA446>

The Exam

- Two hours
- The first question is compulsory and is worth 40 marks:
 - Answer 8 out of 10 short questions (5 marks each) on any topic that we've covered
- Answer 2 out of 3 questions (30 marks each):
 - Evaluating machine translation
 - Translation model
 - Language model
 - Decoder

Exam Topics

- Machine Translation Evaluation
- Language modelling
- Word-based translation models
- Phrase-based translation models
- Decoding
- General Topics

Machine Translation Evaluation

- What are the three types of machine translation evaluation?
- What are the benefits of each type?
- How does each type work?
- The inner workings of BLEU
- The importance of multiple reference translations
- Weaknesses of BLEU

Language Modelling

- What is the purpose of a language model?
- Why can the probability of a sentence not be calculated directly?
- What is an n-gram language model?
- How are language models evaluated?
- What is the trade-off between training set size and the order of the language model?
- What is meant by language model smoothing and why is it necessary? How can counts be re-adjusted using different smoothing methods?

Word-based Translation Models

- What is the purpose of a translation model?
- What is a word-based translation model?
- What are the IBM word alignment models and how do the higher models improve over IBM Model 1?
- How are word alignments and lexical probabilities learnt for IBM Model 1?
- How can the learning of word alignments be made more efficient using factoring-out?

Phrase-based Translation Models

- What is the difference between word-based and phrase-based translation models?
- What are the advantages of phrase-based models over word-based models?
- How can phrase pairs be learnt from word alignments?
- How can a symmetric word alignment be built based on IBM alignments?
- What is a log-linear model? What is the advantage of a log-linear model over a noisy channel model?

Decoding

- What is decoding?
- Why is it a difficult problem?
- Strategies used to make the problem tractable
- Histogram pruning versus threshold pruning
- The stack decoding algorithm
- Limitations of the stack decoding algorithm

General Topics

- Why is machine translation difficult?
- Explain the noisy channel model.
- What is the difference between adequacy and fluency?
- How can linguistic knowledge be used in statistical machine translation?
- What are the properties of a good machine translation system?



Discussion

Acknowledgement



Parts of the content of this lecture are taken from previous lectures and presentations given by Jennifer Foster, Declan Groves, Yvette Graham, Kevin Knight, Josef van Genabith, Andy Way.