

STATISTICAL MACHINE TRANSLATION CHALLENGES, SOLUTIONS AND TRENDS

Adrià de Gispert



Department of Engineering
University of Cambridge

Vigo, 4-8 May 2009

Contents

- Day 1** INTRODUCTION TO SMT
 What is MT? Some History. Why is MT difficult? What is quality?
 SMT Fundamentals. General Pipeline. Sentence Alignment
- Day 2** WORD ALIGNMENT. N-GRAM- AND PHRASE-BASED SMT
 Word Alignment models. N-gram-based SMT. Phrase-based SMT.
 Weighted Finite-State Transducers for Translation.
- Day 3** INTRODUCING LINGUISTIC KNOWLEDGE
 Data Transformation and Context-dependent models of Word Alignment
 Verb form Classification and Morphology-simplified Translation models
- Day 4** REORDERING MODELS. RESCORING
 Reordering in N-gram-based SMT and Phrase-based SMT.
 Hierarchical Phrase-based Models. Syntax-based SMT.
 Rescoring or Hypothesis Re-ranking
- Day 5** OTHER APPLICATIONS. CONCLUSIONS.
 Speech Translation, System Combination,...