Abstract:

This work focuses on two problems in machine translation: lexical choice and target-side morphology. The first problem is the correct transfer of meaning from the source language to the target language. The second problem, which is mainly relevant for morphologically rich target languages, is then the choice of the correct surface form of each target lexeme. We work with these problems within the framework of phrase-based machine translation and we propose a discriminative model of translation which utilizes both source and target context information and which uses rich linguistically motivated features. We show how our model addresses specific weaknesses of standard phrase-based systems and that it provides consistent improvements of translation quality across a broad range of experiments. Apart from our main contribution, we also provide a number of experimental evaluations, analyses and manual annotation experiments, mostly related to English-Czech translation.