

Improving Translation to Morphologically Rich Languages

Alexander Fraser

Center for Information and Language Processing, LMU, Munich

fraser@cis.uni-muenchen.de

RÉSUMÉ

Si les techniques statistiques pour la traduction automatique ont fait des progrès significatifs au cours des 20 dernières années, les résultats pour la traduction de langues morphologiquement riches sont toujours mitigés par rapport aux précédentes générations de systèmes à base de règles. Les recherches actuelles en traduction statistique de langues morphologiquement riches varient grandement en fonction de la quantité de connaissances linguistiques utilisées et de la nature de ces connaissances. Cette variation est plus importante en langue cible (par exemple, les ressources utilisées en traduction automatique statistique respectueuse de linguistique en arabe, en français et en allemand sont très différentes). La conférence portera sur les techniques état de l'art dédiées à la tâche de traduction statistique pour une langue cible qui est morphologiquement plus riche que la langue source.

ABSTRACT

Améliorer la traduction des langages morphologiquement riches

While statistical techniques for machine translation have made significant progress in the last 20 years, results for translating to morphologically rich languages are still mixed versus previous generation rule-based systems. Current research in statistical techniques for translating to morphologically rich languages varies greatly in the amount of linguistic knowledge used and the form of this linguistic knowledge. This varies most strongly by target language (e.g., the resources used for linguistically-aware statistical machine translation to Arabic, French, German are very different). The talk will discuss state-of-the-art techniques for statistical translation tasks involving translating to a target language which is morphologically richer than the source language.

MOTS-CLÉS : traduction statistique, langages morphologiquement riches, connaissances linguistiques.

KEYWORDS: statistical translation, morphologically rich languages, linguistic knowledge.
