FAUST: Feedback Analysis for User adaptive Statistical Translation -- First Year Progress

Seventh Framework Programme
Theme FP7-ICT-2009-4, Objective 2.2: Language-based interaction.
STREP
Grant agreement no. 247762
http://faust-fp7.eu

List of partners
Department of Engineering, University of Cambridge, UK
Computer Laboratory, University of Cambridge, UK
Center for Language and Speech Technologies and Applications (TALP), Universitat Politecnica de Catalunya, Spain
Institute of Formal and Applied Linguistics, Charles University, Czech Republic
Language Weaver Inc., USA Language Weaver SRL, Romania
Softissimo, Inc., France

Project duration: February 2010 — January 2013

Summary

The FAUST project is developing fluent MT systems which respond to user feedback. In doing so we will mitigate long-standing causes for user frustration with MT. One problem is that current MT systems do not respond to suggestions for improvement. There are technical reasons for this, including: User feedback tends to be very noisy; no research published to date makes explicit how statistical translation and language models can be adapted to benefit from feedback provided by web users; no mechanisms exist to identify user feedback of value and for immediately affecting the behavior of a statistical MT system so that subsequent users do not run into the same problem; Users are also frustrated by the general lack of fluency. In any NLP system, basic mistakes in grammar or word sense suggest that the technology is not reliable – systems must be fluent if they are to be accepted and trusted by casual users. With these problems in mind, our objectives are to: Enhance the high-volume Reverso.net translation website with an experimental infrastructure for the study of instantaneous user feedback; Deploy novel web-oriented, feedback collection mechanisms that reduce noise and increase the utility of the web contributions; Automatically acquire novel data collections to study translation as informed by user feedback; Develop mechanisms for instantaneously incorporating user feedback into the MT engines; Create novel automatic metrics of translation quality which reflect user feed-back; Develop translation models based on user feedback data and develop approaches to integrate natural language generation directly into MT to improve translation fluency and reduce negative feedback. We will describe our progress in the first project year in the deployment of web-based MT research systems for the collection and analysis of user feedback.

Mikel L. Forcada, Heidi Depraetere, Vincent Vandeghinste (eds.)
Proceedings of the 15th Conference of the European Association for Machine Translation, p. 333
Leuven, Belgium, May 2011