Emil Djupfeldt
Chalmers University of Technology
2012-01-20

Master's Thesis specification:
Machine translation of trekking and climbing route descriptions using Grammatical Framework

The goal is to create a system that can automatically translate descriptions and travel guides for trekking and climbing routes. Such texts are often found in the language of the country the route is situated in, which may not be the prospective climber's native tongue. It would be useful if the route descriptions could in some way be translated without the climber having to resort to dictionaries and their own sometimes limited knowledge of the foreign language.

There are some pitfalls however. In both climbing and trekking, terms can be found that are not as common in everyday language. In order to handle this the system will need to be aware of these terms. It would be inconvenient and sometimes even dangerous not to translate them properly.

There are multiple approaches to translating text. This work will concentrate on grammatically aware translation using Grammatical Framework (GF), a programming language for multilingual grammar applications. This is different from some other systems, e.g. Google Translate, that instead use statistical models to create translations.

During the course of this project there are a few tasks that need to be addressed:

- A sufficient amount of text in the target languages needs to be gathered. The texts should include enough domain specific jargon that the GF resource grammar library doesn't already cover it.

- A domain lexicon will be created to cover the jargon. The lexicon should have instances for at least Swedish, English and German. Support for Italian and/or French will be added later if there is time.

- An application that uses the domain lexicon will be developed. The application should be an iPhone app based on the existing GF Android application. Due to differences in the two platforms the existing code most likely cannot be reused, and instead only used as a guideline. The new application should be easy to repurpose to use other lexicons so that it can be reused or extended for other projects.

- Translations generated by the system need to be validated to ensure that the system works as intended.