Prof. R. Wattenhofer

Fictional Language Generator

Sindarin, Klingon, Dothraki. These are some of the few fully developed fictional languages that you could actually learn to speak. There are some more "half-developed" languages (see "Artistic/Fictional languages" here), but it is a relatively short list.

Image you want to make a movie or create a computer game, and you want some of your characters to speak a foreign/alien language to make it more believable for the audience (what kind of alien speaks English after all?!?). The problem is that creating a new language is a non-trivial task that requires deep knowledge of linguistics and a high amount of work. So in the end, it is either something for the gifted and passionate (e.g. Tolkien) or for the deep-pocketed (HBO with Game of Thrones, Hollywood with Avatar).

This is where you come in! We want you to automatically create believable fictional languages. We already have some ideas on how to achieve this. We could for example train machine learning classifiers such as Random Forest on different languages, and then start using those classifiers to mix languages according to our wishes. Depending on what we mix and how we do it, the results will be different. We could also employ evolutionary algorithms to combine languages and then further process them with other techniques (real languages evolve as well, after all). We do not necessarily aim at creating perfect new languages, but something more believable than this is certainly doable. These are just some early ideas, but the area is completely open for exploration. If you are interested, do not hesitate to contact us so that we can have a chat!

You will: Familiarize yourself with the necessary basics of linguistics and Natural Language Generation (NLG). Considerably improve your knowledge of Machine Learning techniques. Read papers on state-of-the-art generative learning algorithms. Implement different algorithms and evaluate them. Meet with your supervisors once a week to discuss progress and open questions.

Requirements: Interest in languages and Machine Learning. Prior knowledge about Machine learning is an advantage, but not a must.

Contacts

- Yuyi Wang: yuyi.wang@tik.ee.ethz.ch, ETZ G94

- Gino Brunner: gino.brunner@tik.ee.ethz.ch, ETZ G63