

Title:

Large-Scale Discriminative Training for Machine Translation into Morphologically-Rich Languages

Author:

Miloš Stanojević

Department:

Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics, Charles University in Prague, Czech Republic

Supervisor:

RNDr. Ondřej Bojar Ph.D.

Abstract:

We present the theoretical foundations for large-scale discriminative training and their application on the machine translation into MRL languages. The research is concentrated mostly on the two aspects of large-scale discriminative training. The first aspect is the effect of sparse features on the choice of word form and word order. The second aspect is application of different sentence-level metrics as an objective function in the parameter optimization.