Packing it all up in Search for a Language Independent MT Quality Measure Tool – Part Two

Kimmo Kettunen

Kymenlaakso University of Applied Sciences
Paraatikenttä 7, FIN-045 100 Kouvola, Finland
`kimmo.kettunen@kyamk.fi`

Abstract. This study describes first usage of a particular implementation of Normalized Compression Distance (NCD) as a machine translation quality evaluation tool. NCD has been introduced and tested for clustering and classification of different types of data and found a reliable and general tool. As far as we know NCD in its Complearn implementation has not been evaluated as a MT quality tool yet, and we wish to show that it can also be used for this purpose. We show that NCD scores given for MT outputs in different languages correlate highly with scores of a state-of-the-art MT evaluation metrics, METEOR 0.6. Our experiments are based on translations between one source and three target languages with a smallish sample that has available reference translations, UN's Universal Declaration of Human Rights. Secondly we shall also briefly describe and discuss results of a larger scale evaluation of NCD as an MT metric with WMT08 Shared Task Evaluation Data. These evaluations confirm further that NCD is a noteworthy MT metric both in itself and also enriched with basic language tools, stemming and Wordnet.

Keywords: MT metrics, Normalized Compression Distance, machine translation, evaluation of machine translation systems