



Abstract

Machine translation is a fundamental technology that is gaining more importance each day in our multilingual society. Companies and particulars are turning their attention to machine translation since it dramatically cuts down their expenses on translation and interpreting. However, the output of current machine translation systems is still far from the quality of translations generated by human experts. The overall goal of this thesis is to narrow down this quality gap by developing new methodologies and tools that improve the broader and more efficient deployment of machine translation technology.

We start by proposing a new technique to improve the quality of the translations generated by fully-automatic machine translation systems. The key insight of our approach is that different translation systems, implementing different approaches and technologies, can exhibit different strengths and limitations. Therefore, a proper combination of the outputs of such different systems has the potential to produce translations of improved quality. We present minimum Bayes' risk system combination, an automatic approach that detects the best parts of the candidate translations and combines them to generate a consensus translation that is optimal with respect to a particular performance metric. We thoroughly describe the formalization of our approach as a weighted ensemble of probability distributions and provide efficient algorithms to obtain the optimal consensus translation according to the widespread BLEU score. Empirical results show that the proposed approach is indeed able to generate statistically better translations than the provided candidates. Compared to other state-of-the-art systems combination methods, our approach reports similar performance not requiring any additional data but the candidate translations.

Then, we focus our attention on how to improve the utility of automatic translations for the end-user of the system. Since automatic translations are not perfect, a desirable feature of machine translation systems is the ability to predict at run-time the quality of the generated translations. Quality estimation is usually addressed as a regression problem where a quality score is predicted from a set of features that represents the translation. However,

although the concept of translation quality is intuitively clear, there is no consensus on which are the features that actually account for it. As a consequence, quality estimation systems for machine translation have to utilize a large number of weak features to predict translation quality. This involves several learning problems related to feature collinearity and ambiguity, and due to the “curse” of dimensionality. We address these challenges by adopting a two-step training methodology. First, a dimensionality reduction method computes, from the original features, the reduced set of features that better explains translation quality. Then, a prediction model is built from this reduced set to finally predict the quality score. We study various reduction methods previously used in the literature and propose two new ones based on statistical multivariate analysis techniques. More specifically, the proposed dimensionality reduction methods are based on partial least squares regression. The results of a thorough experimentation show that the quality estimation systems estimated following the proposed two-step methodology obtain better prediction accuracy than systems estimated using all the original features. Moreover, one of the proposed dimensionality reduction methods obtained the best prediction accuracy with only a fraction of the original features. This feature reduction ratio is important because it implies a dramatic reduction of the operating times of the quality estimation system.

An alternative use of current machine translation systems is to embed them within an interactive editing environment where the system and a human expert collaborate to generate error-free translations. This interactive machine translation approach has shown to reduce supervision effort of the user in comparison to the conventional decoupled post-edition approach. However, interactive machine translation considers the translation system as a passive agent in the interaction process. In other words, the system only suggests translations to the user, who then makes the necessary supervision decisions. As a result, the user is bound to exhaustively supervise every suggested translation. This passive approach ensures error-free translations but it also demands a large amount of supervision effort from the user.

Finally, we study different techniques to improve the productivity of current interactive machine translation systems. Specifically, we focus on the development of alternative approaches where the system becomes an active agent in the interaction process. We propose two different active approaches. On the one hand, we describe an active interaction approach where the system informs the user about the reliability of the suggested translations. The hope is that this information may help the user to locate translation errors thus improving the overall translation productivity. We propose different scores to measure

translation reliability at the word and sentence levels and study the influence of such information in the productivity of an interactive machine translation system. Empirical results show that the proposed active interaction protocol is able to achieve a large reduction in supervision effort while still generating translations of very high quality. On the other hand, we study an active learning framework for interactive machine translation. In this case, the system is not only able to inform the user of which suggested translations should be supervised, but it is also able to learn from the user-supervised translations to improve its future suggestions. We develop a value-of-information criterion to select which automatic translations undergo user supervision. However, given its high computational complexity, in practice we study different selection strategies that approximate this optimal criterion. Results of a large scale experimentation show that the proposed active learning framework is able to obtain better compromises between the quality of the generated translations and the human effort required to obtain them. Moreover, in comparison to a conventional interactive machine translation system, our proposal obtained translations of twice the quality with the same supervision effort.