

discriminative learning method based on a comprehensive theoretical framework and optimization technique [8][9]. While the evaluation we participated is text translation only, its effectiveness provides an indirect evidence that its extension to speech translation will be promising, which is a more natural task targeted by our theoretical framework presented in [8]. For the method of topic adaptation, with more data available, we expect the adaptation technique will show greater strength than presented in this paper.

6. Acknowledgements

We would like to thank the organization committee of the IWSLT2011 evaluation campaign that makes the evaluation presented in this paper possible.

7. References

- [1] A. Axelrod, X. He, J. Gao. "Domain adaptation via pseudo in-domain data selection". Proceedings of Empirical Methods in Natural Language Processing, 2011.
- [2] D. Blei, A. Ng, M.I. Jordan, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, 2003.
- [3] D. Chiang. "A hierarchical phrase-based model for statistical machine translation." In *Proc. of ACL*, 2005.
- [4] D. Chiang, K. Knight and W. Wang, "11,001 new features for statistical machine translation," in *Proc. NAACL-HLT*, 2009.
- [5] M. Federico, L. Bentivogli, M. Paul, and S. Stueker, "Overview of the IWSLT 2011 Evaluation Campaign", In *Proc. IWSLT*, 2011
- [6] X. He, "Using word-dependent transition models in HMM-based word alignment for statistical machine translation," *Proc. ACL-WMT*, 2007.
- [7] X. He and L. Deng, "Robust speech translation by domain adaptation." in *Proc. Interspeech*, 2011
- [8] X. He and L. Deng. "Speech Recognition, Machine Translation, and Speech Translation – A Unified Discriminative Learning Paradigm" *IEEE Sig. Proc. Mag.*, Sept 2011.
- [9] X. He, L. Deng, W. Chou, "Discriminative learning in sequential pattern recognition." *IEEE Sig. Proc. Mag.*, Sept., 2008.
- [10] B.-H. Juang, W. Chou, and C.-H. Lee, "Minimum classification error rate methods for speech recognition," *IEEE Trans. Speech Audio Processing*, May 1997.
- [11] P. Koehn, F. Och, and D. Marcu. "Statistical phrase-based translation," *Proc. HLT-NAACL*, 2003
- [12] P. Liang, A. Bouchard-Cote, D. Klein and B. Taskar, "An end-to-end discriminative approach to machine translation," in *Proc. COLING-ACL*, 2006
- [13] A. McCallum, "MALLET: A machine learning for language toolkit", <http://mallet.cs.umass.edu>, 2002.
- [14] F. Och, "Minimum error rate training in statistical machine translation." *Proc. ACL*, 2003.
- [15] F. Och and H. Ney. "Discriminative training and maximum entropy models for statistical machine translation." In *Proc. ACL* 2002.
- [16] K. Papineni, S. Roukos, T. Ward, and W. Zhu, "BLEU: A method for automatic evaluation of machine translation." *Proc. ACL*, 2002.
- [17] N. Ruiz, M. Federico. "Topic adaptation for lecture translation through bilingual latent semantic models". In *Proc of WMT* 2011.
- [18] Y.C. Tam, I. Lane, T. Schultz, "Bilingual-LSA based LM adaptation for spoken language translation". In *Proc of ACL*, 2007.
- [19] Y.C. Tam and T. Schultz, "Unsupervised language model adaptation using latent semantic marginals", In *Proc of Interspeech*, 2006.
- [20] X. He, M. Yang, J. Gao, P. Nguyen, R. Moore. "Indirect-HMM-based Hypothesis Alignment for Combining Outputs from Machine Translation Systems." In *Proc. EMNLP*, 2008
- [21] C-H. Li, X. He, Y. Liu and N. Xi. "Incremental HMM Alignment for MT System Combination." In *Proc. ACL*, 2009
- [22] X. He and K. Toutanova. "Joint Optimization for Machine Translation System Combination." In *Proc. EMNLP*, 2009.
- [23] R. Moore and C. Quirk. "Faster Beam-Search Decoding for Phrasal Statistical Machine Translation." In *Proc. of MT Summit XI*, 2007.