An information extraction approach to english-vietnamese weather bulletins machine translation

Pham S.B., Tran G.B., Pham D.D., Phung K.C., Nguyen K.T.

College of Technology, Vietnam National University, Hanoi, Viet Nam; Control Automation in Production and Improvement of Technology Institute, Hanoi, Viet Nam

Abstract: In this paper, we present our method of using Information Extraction techniques to tackle the task of automatically translating English weather bulletins to Vietnamese. It is simple yet effective in satisfying the constraints of low processing power and storage space for the deployment on an embedded system. Experimental results are very promising with the F-measure going up to 96% for extracting relevant information from the weather bulletins. ?? 2009 IEEE.

Index Keywords: F-measure; Information Extraction; Information extraction techniques; Machine translations; Processing power; Storage spaces; Computer aided language translation; Information analysis; Information theory; Speech transmission; Database systems

Year: 2009

Source title: Proceedings - 2009 1st Asian Conference on Intelligent Information and Database Systems, ACIIDS 2009

Art. No.: 5175986

Page : 161-166

Link: Scorpus Link

Correspondence Address: Pham, S. B.; College of Technology, Vietnam National University, Hanoi, Viet Nam; email: Sonpb@vnu.edu.vn

Conference name: 2009 1st Asian Conference on Intelligent Information and Database Systems, ACIIDS 2009

Conference date: 1 April 2009 through 3 April 2009

Conference location: Dong Hoi

Conference code: 78395

ISBN: 9.78E+12

DOI: 10.1109/ACIIDS.2009.90

Language of Original Document: English

Abbreviated Source Title: Proceedings - 2009 1st Asian Conference on Intelligent Information and Database Systems, ACIIDS 2009

Document Type: Conference Paper

Source: Scopus

Authors with affiliations:

1. Pham, S.B., College of Technology, Vietnam National University, Hanoi, Viet Nam
2. Tran, G.B., College of Technology, Vietnam National University, Hanoi, Viet Nam
3. Pham, D.D., College of Technology, Vietnam National University, Hanoi, Viet Nam

4. Phung, K.C., Control Automation in Production and Improvement of Technology Institute, Hanoi, Viet Nam
5. Nguyen, K.T., Control Automation in Production and Improvement of Technology Institute, Hanoi, Viet Nam

References:

1. Agichtein, E., Gravano, L., Snowball: Extracting relations from large plain text collections (2000) Proceedings of the 5th ACM International Conference on Digital Libraries
2. Califf, M.E., Mooney, R.J., Relational learning of patternmatch rules for information extraction (1999) Proceedings of the 16th National Conference on Artificial Intelligence (AAAI), , pages 328-334, Orlando, FL
3. Ayuso, H.F.H.G.R.I.D., Boisen, S., Weischedel, R., Bbn: Description of the plum system as used for muc-4 (1992) Proceedings of the Fourth Message Understanding Conference (MUC-4), , pages 169-176
4. Krupka, L.L.C.G., Jacobs, P., Sider, I., Ge nltoolset: Description of the system as used for muc-4 (1992) Proceedings of the Fourth Message Understanding Conference (MUC-4), , pages 177-185
5. Huffman, S.B., Learning information extraction patterns from examples (1996) Lecture Notes in Computer Science. Connectionist, Statistical, and Symbolic Approaches to Learning for Natural Language Processing, , pages 246-260
6. Hobbs, M.T.J.J.R., Appelt, D., Islael, D., Sri international: Description of the fastus system used for muc-4 (1992) Proceedings of the Fourth Message Understanding Conference (MUC-4), , pages 268-275
7. Kaiser, K., Miksch, S., (2005) Information Extraction - A Survey
8. Kim, J.T., Moldovan, D.I., Acquisition of linguistic patterns for knowledge-based information extraction (1995) IEEE Transactions on Knowledge and Data Engineering, , pages 713-724
9. Yangarber, P.T.R., Grishman, R., Huttunen, S., Automatic acquisition of domain knowledge for information extraction (2000) Proceedings of the 18th International Conference on Computational Linguistics (COLING), , Saarbrucken, Germany
10. Riloff, E., Automatically constructing a dictionary for information extraction tasks (1993) Proc. of the 11th National Conference on Artificial Intelligence, , pages 811-816
11. Riloff, E., Automatically generating extraction patterns from untagged text (1996) Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI), , pages 1044-1049
12. Soderland, C.C.S., Mooney, R.J., Learning information extraction rules for semi-structured and free text (1999) Machine Learning, , pages 233-272
13. Soderland, J.A.S., Fisher, D., Lehnert, W., Crystal: Inducing a conceptual dictionary (1995) Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI), , pages 1314-1319
14. Scolum, J., (1985) Survey of Machine Translation: Its History, Current Status, and Future Prospects
15. Sudo, K., (2004) Unsupervised Discovery of Extraction Patterns for Information Extraction, , New York University New York
16. Yangarber, R., Grishman, R., Nyu: Description of the proteus/pet system as used for muc-7 st (1992) Proceedings of the Seventh Message Understanding Conference (MUC-4), , pages 169-176

Download Full Text: 0283.pdf