Cross-lingual Event Detection in Discourse

German Rigau
Computer Science Faculty, EHU
`german.rigau@ehu.eus`

Abstract

We describe a system for event extraction across documents and languages. We developed a framework for the interoperable semantic interpretation of mentions of events, participants, locations and time, as well as the relations between them.

Furthermore, we use a common RDF model to represent instances of events and normalised entities and dates. We convert multiple mentions of the same event in English and Spanish to a single representation. We thus resolve cross-document event and entity coreference within a language but also across languages. We tested our system on a Wikinews corpus of 120 English articles that have been manually translated to Spanish.

We report on the cross-lingual cross-document event and entity extraction comparing the Spanish output with respect to English.

1 Speaker's Bio

German Rigau Ph.D. and B.A. in Computer Science by the Universitat Politecnica de Catalunya (UPC). Formerly member of the Computer Science department at the UPC and member of the TALP research group of the UPC, currently, he is teaching at the Computer Science Faculty of the EHU as an Associate Professor. He has published more than hundred-refereed articles and conference papers in the area of Natural Language Processing, and in particular Acquisition of Lexical Knowledge, Word Sense Disambiguation, Semantic Processing and Inferencing.

He has been involved in several European research projects (ESPRIT BRA ACQUILEX, ACQUILEX II, LE EUROWORDNET, LE NAMIC, MEANING, KYOTO, PATHS, OpeNER and NewsReader). He coordinated the 5th Framework MEANING project (IST-2001-34460) and the local groups for NAMIC, KYOTO and OpeNER. Currently, he is coordinating the local group for NewsReader (FP7-ICT-2011-8-316404).

He has been also involved in several Spanish National research projects (ITEM, HERMES, SENSEM, KNOW, KNOW2 and SKaTer). Currently, he is coordinating the local group of the SKaTer project. He served as PC member and reviewer of the main international conferences and workshops in NLP and AI including ACL, EACL, NAACL, COLING, AAAI, ECAI, IJCAI, EMNLP, IJCNLP, CoNLL, TSD, SENSEVAL/SEMEVAL and IWC.

He also served as reviewer of International Journals including: Computers and the Humanities, Journal of Natural Language Engineering, Journal of Artificial Intelligence Research and Artificial Intelligence. He has also participated in all editions of the international competition of SENSEVAL.

Currently, he is member of the Association for Computational Linguistics (ACL) and the Spanish Society for Natural Language Processing (SEPLN).