Preface

It is our pleasure to welcome you to the First International Workshop on Free/Open-Source Rule-Based Machine Translation to be held in Alacant, Spain. This workshop has been inspired, on one hand, by the successful OSMaTran workshop on open-source machine translation —held on September, 2005 at the Tenth Machine Translation Summit— where 3 of the 4 presented papers where about rule-based machine translation; and, on the other, by the successful participation of the Apertium project in the Google Summer of Code 2009.

The free/open-source software movement has arrived into the field of machine translation. Machine translation is special in that, in addition to specific algorithms, it heavily depends on extensive language-dependent data. Therefore, not only the engine or the tools used to manage these data have to be free/open-source, but also the data themselves. There are many machine translation packages of this type available; however most of them are corpus-based, and, in particular, statistical machine translation systems. Rule-based machine translation systems built on these principles are still little known and little used.

There are distinct advantages of having free/open-source licences for rule-based machine translation: the linguistic knowledge for the translation of one language into another, which is explicitly encoded in the form of linguistic data so that both humans and the machine translation engine can process it, can be reused to build data for other language pairs or even for other human language technologies besides machine translation, and, conversely, linguistic knowledge from other sources may be reused to build machine translation systems. The free and open scenario makes this reuse easier, and, if *copyleft* licences are used, builds a commons of knowledge and resources that benefits all the language communities involved. These advantages are even clearer for less-resourced languages, for which large bilingual corpora are not available, and for morphologically rich languages, which even with large corpora suffer from data sparseness.

This workshop aims at bringing together the experience of researchers and developers in the field of rule-based machine translation who have decided to board the free/open-source train and are effectively contributing to create that commons of explicit knowledge: machine translation rules and dictionaries, and machine translation systems whose behaviour is transparent and clearly traceable through their explicit logic. Each of the 16 papers submitted from 11 countries was peer-reviewed by two independent reviewers from the program committee, resulting in a selection of the 10 papers included in these proceedings.

Our special thanks goes to Dr. Amba Kulkarni, from the University of Hyderabad, and Dr. Kepa Sarasola, from Euskal Herriko Unibertsitatea, who kindly agreed to give keynote addresses at the workshop. We also acknowledge all the reviewers, whose names are subsequently listed, for their detailed extensive reviews and useful recommendations which where vital in helping authors to improve their papers. We also appreciate the dedicated efforts of the local organising committee who worked hard to plan this workshop. We are also very grateful to the sponsors who supported this workshop: Institut Universitari d'Investigació Informàtica

of Universitat d'Alacant. Departament de Llenguatges i Sistemes Informàtics of Universitat d'Alacant, Prompsit Language Engineering, and also to the people responsible for the Google Summer of Code 2009, who let us invest the grant received by the Apertium project in planning this workshop. Finally, a big thank you goes to all the authors who made this workshop both possible and successful.

All the papers included in this proceedings can be found at the Open Access Repository of Universitat d'Alacant on `http://rua.ua.es/dspace/handle/10045/11809`.

We look forward to an exciting and productive workshop.

Alacant, November 2009

Juan Antonio Pérez-Ortiz and **Felipe Sánchez-Martínez**
FreeRBMT09 Programme Committee co-chairs