

# Constructing Cross-language Ontological and Lexical Resources

Shu-Kai HSIEH

Graduate Institute of Linguistics, National Taiwan University

shukaihsieh@ntu.edu.tw

This paper introduces our on-going efforts in constructing the platform for harmonizing cross-language ontological and lexical resources. In the recent years, a large amount of large-scale lexical-semantic resources, such as WordNet (Fellbaum, 1998) and FrameNet (Baker et al, 1998), have been made available and gained great success both in the fields of Natural Language Processing and Information Retrieval.

Following the similar line, EuroWordnet (Vossen, 1998) extends the architecture into a multilingual scenario, in which seven European languages are connected via the proposed interlingual index (ILI) which assumes a set of language-independent concepts, and the ILI is manually linked to the EWN Top Ontology. In addition, these lexical resources are also mapped to other upper level ontologies like SUMO (Suggested Upper Merged Ontology, Niles and Pease, 2001) and DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering, Gangemi et al., 2003).

Among the context of Sinosphere, also known as CCCS (Chinese Character Cultural Sphere), which denotes a grouping of regions and countries where Chinese characters were adopted and integrated to their languages, we also proposed a cross-cultural and language platform called *HanziGrid* (Chou et al., 2007). The underlying assumption is that *Radicals* are relative stable and attested ontology over thousands of years, so they are proposed as ILI-like structure in interfacing multi-lingual lexicon and ontologies. This paper will address the extension of the platform, including the mappings to SUMO, DOLCE on the one side, and Chinese-English bilingual WordNet on the other side, and its potentials in serving as cross-language and cross-cultural linguistic and conceptual knowledge processing infrastructure.