On background knowledge: The case study of ontology matching

Fausto Giunchiglia

Lightweight ontologies or classifications are pervasive: we use them to classify our messages, our favorite web pages, our files, etc. The problem is that all these classifications are semantically heterogeneous. The most striking consequence is that they classify documents very differently, making therefore very hard and sometimes impossible to find them. Matching classifications is the process which allows us to map those nodes of two classifications which, intuitively, correspond semantically to each other. In the first part of the talk I will show how it is possible to encode this problem into a propositional validity problem, thus allowing for the use of propositional satisfiability (SAT) reasoners. This is done mainly using linguistic resources (e.g., WordNet) and some amount of Natural Language Processing. However, as shown in the second part of the talk, this turns out to be almost useless. In most cases, in fact, linguistic resources do not contain enough of the axioms needed to prove unsatisfiability. We will argue that this is an example of a more general phenomenon, that we call “Lack of Background Knowledge” which arises when using commonsense knowledge. Finally, we will discuss an automatic approach which allows us to discover some of the missing background knowledge in matching tasks.