

Revealing Granularity of Domain Terminology with Inductive Method of Model Self-Organization

NATALIA PONOMAREVA
Universidad Politécnica de Valencia, Spain,

ABSTRACT

The aim of this work is to suggest a method of domain terms selection for different granularity levels. First, we give a definition of corpus-based term granularity and propose entropy-based and standard deviation-based weighting schemes for its evaluation. A chosen term weighting scheme is a decisive factor of granularity approximation quality. We declare a hypothesis of how to reveal boundaries between different granularity levels using our modified version of the Inductive Method of Model Self-Organization. Although the suggested method demonstrates stability in the framework of our hypothesis some additional study of its reliability must be accomplished and other more precise weighting schemes should be applied.

1. INTRODUCTION

Evaluation of ontology granularity level is still a difficult and poorly reflected problem in literature, although the importance of its solution is indisputable. A notion of granularity is used in a very intuitive way, neither its formal definition nor a mode of its measuring has been proposed up to now. Ontology granularity can be considered from different aspects: either at a lexical level, which refers to granularity of ontology concepts or at a conceptual level where expressiveness of ontology properties and relations is in a center of investigation.

In this paper, we consider a problem of ontology granularity at a lexical level (the lowest level of expressiveness according to Ontology Summit 2007¹). Therefore, we aim at evaluating only granularity of ontology concepts without taking into account other ontology components. In other words, approximation of ontology by a list of its concepts is realized. Although it is a very rude approximation we argue that it is a first step of ontology granularity evaluation.