

# Similarity measures over refinement graphs

Santiago Ontañón · Enric Plaza

Received: 28 September 2009 / Accepted: 23 November 2011 / Published online: 17 December 2011  
© The Author(s) 2011

**Abstract** Similarity also plays a crucial role in support vector machines. Similarity assessment plays a key role in lazy learning methods such as  $k$ -nearest neighbor or case-based reasoning. In this paper we will show how refinement graphs, that were originally introduced for inductive learning, can be employed to assess and reason about similarity. We will define and analyze two similarity measures,  $S_\lambda$  and  $S_\pi$ , based on refinement graphs. The *anti-unification-based similarity*,  $S_\lambda$ , assesses similarity by finding the anti-unification of two instances, which is a description capturing all the information common to these two instances. The *property-based similarity*,  $S_\pi$ , is based on a process of disintegrating the instances into a set of *properties*, and then analyzing these property sets. Moreover these similarity measures are applicable to any representation language for which a refinement graph that satisfies the requirements we identify can be defined. Specifically, we present a refinement graph for feature terms, in which several languages of increasing expressiveness can be defined. The similarity measures are empirically evaluated on relational data sets belonging to languages of different expressiveness.

**Keywords** Similarity measures · Refinement graphs · Case-based reasoning · Feature terms · Lazy learning

## 1 Introduction

Similarity assessment plays a key role in lazy learning methods such as  $k$ -nearest neighbor (Cover and Hart 1967) or case-based reasoning (Aamodt and Plaza 1994), where new problems are solved typically by selecting, adapting or interpolating the solutions of the

---

Editor: David Page.

S. Ontañón (✉) · E. Plaza  
IIIA, Artificial Intelligence Research Institute, CSIC, Spanish Council for Scientific Research, Campus  
UAB, 08193, Bellaterra, Catalonia, Spain  
e-mail: [santi@iiia.csic.es](mailto:santi@iiia.csic.es)

E. Plaza  
e-mail: [enric@iiia.csic.es](mailto:enric@iiia.csic.es)