# Chapter 15
# Improving Ontology Alignment through Genetic Algorithms

**José Manuel Vázquez Naya**
*University of A Coruña, Spain*

**Marcos Martínez Romero**
*University of A Coruña, Spain*

**Javier Pereira Loureiro**
*University of A Coruña, Spain*

**Cristian R. Munteanu**
*University of A Coruña, Spain*

**Alejandro Pazos Sierra**
*University of A Coruña, Spain*

## ABSTRACT

*Ontology alignment is recognized as a fundamental process to achieve an adequate interoperability between people or systems that use different, overlapping ontologies to represent common knowledge. This process consists of finding the semantic relations between different ontologies. There are different techniques conceived to measure the semantic similarity of elements from separate ontologies, which must be adequately combined in order to obtain precise and complete results. Nevertheless, combining multiple measures into a single similarity metric is a complex problem, which has been traditionally solved using weights determined manually by an expert, or calculated through general methods that does not provide optimal results. In this chapter, a genetic algorithm based approach to find out how to aggregate different similarity metrics into a single measure is presented. Starting from an initial population of individuals, each one representing a specific combination of measures, the algorithm finds the combination that provides the best alignment quality.*

DOI: 10.4018/978-1-61520-893-7.ch015

## INTRODUCTION

At present, the role of ontologies for allowing a more effective data and knowledge sharing and reusing is widely recognized (Neches et al., 1991; Guarino, 1998) and a variety of public ontologies exist for different application domains. This innovative technology is considered to be an appropriate solution to the problem of heterogeneity in data, since ontological methods make it possible to reach a common understanding of concepts in a particular domain, supporting the exchange of information between people (or between systems) that utilize different representations for the same or similar knowledge (Ashburner et al., 2000; Gómez-Pérez et al., 2004; Gruber, 1995).

Nevertheless, given that different tasks or different points of view usually require different conceptualizations, utilizing a single ontology is neither always possible nor advisable. This can lead to the usage of different ontologies, although in some cases the different ontologies collectively might contain information that could be overlapping. This, in turn, represents another type of heterogeneity that can result in inefficient processing or misinterpretation of data, information, and knowledge. Having into account that ontologies can interoperate only if correspondences between their elements have been identified and established (Kalfoglou and Schorlemmer, 2003), addressing this problem while at the same time insure an appropriate level of interoperability between heterogeneous systems, requires to find correspondences or mappings that exist between the elements of the different ontologies being used. This process is commonly known as *ontology alignment*, *ontology matching* or *ontology mapping*. The resulting set of inter-ontology relations can be used to adequately exchange information between people, systems and organizations.

As such, ontology engineers face the problem of how to map various ontologies to enable a common understanding (Ding et al., 2002), and several methods for identifying the relationships or correspondences between elements associated with different ontologies have been conceived during the last years. Collectively these methods are called *ontology alignment techniques*. Each one of these techniques provides a numerical value of similarity between elements from separate ontologies that can be used to decide (frequently, according to a predefined threshold) if those elements are semantically similar or not.

Ontology alignment is considered to be a complex, tedious, and time-consuming task, especially when working with ontologies of considerable size (containing, for instance, thousands of elements). In addition, the true potential of ontology alignment is realized when different information-exchange processes are achieved automatically and on real time, thereby providing the framework for reaching a suitable level of efficient interoperability between heterogeneous systems. The importance of automatically aligning ontologies has therefore been a topic of major interest in recent years, and there has been a surge in a variety of software tools dedicated to aligning ontologies in either a fully or partially automated fashion. These tools are commonly referred to as *ontology alignment systems*.

Due to the complex nature of the ontology alignment problem, the application of a single ontology alignment technique normally is not enough to provide satisfactory alignment results. Because of this reason, most ontology alignment systems combine a set of different alignment techniques, which can be based on different approaches (e.g. lexical, structural, etc.), into a single value of aggregated similarity between ontological elements, and subsequently select the semantic mappings on the basis of the aggregated similarity function. Finding the optimal combination of similarity measures is a complex and hard process, which should be achieved automatically in order to get optimal results. Nevertheless, current similarity aggregation approaches generally use weights