UNIVERSITÀ DEGLI STUDI DI SALERNO

DIPARTIMENTO DI INFORMATICA

DOTTORATO DI RICERCA IN TEORIE, METODOLOGIE E APPLICAZIONI AVANZATE PER LA COMUNICAZIONE, L'INFORMATICA E LA FISICA

XI CICLO - NUOVA SERIE

Abstract

Memetic Algorithms for Ontology Alignment

Autilia Vitiello

Ph.D. Program Chair

Prof. Giuseppe Persiano

Supervisors

Prof. Vincenzo Loia

Dott. Giovanni Acampora

NOVEMBER 2012

Semantic interoperability represents the capability of two or more systems to meaningfully and accurately interpret the exchanged data so as to produce useful results. It is an essential feature of all distributed and open knowledge based systems designed for both e-government and private businesses, since it enables machine interpretation, inferencing and computable logic. Unfortunately, the task of achieving semantic interoperability is very difficult because it requires that the meanings of any data must be specified in an appropriate detail in order to resolve any potential ambiguity.

Currently, the best technology recognized for achieving such level of precision in specification of meaning is represented by ontologies. According to the most frequently referenced definition [1], an *ontology* is an explicit specification of a conceptualization, i.e., the formal specification of the objects, concepts, and other entities that are presumed to exist in some area of interest and the relationships that hold them [2]. However, different tasks or different points of view lead ontology designers to produce different conceptualizations of the same domain of interest. This means that the subjectivity of the ontology modeling results in the creation of *heterogeneous ontologies* characterized by terminological and conceptual discrepancies. Examples of these discrepancies are the use of different words to name the same concept, the use of the same word to name different concepts, the creation of hierarchies for a specific domain region with different levels of detail and so on. The arising so-called *semantic heterogeneity problem* represents, in turn, an obstacle for achieving semantic interoperability.

In order to overcome this problem and really taking advantage of the ontological representation, the most solid solution is to perform a so-called *ontology alignment process* or simply *matching*. This process leads two heterogeneous ontologies into a mutual agreement by detecting a set of correspondences, called *alignment*, between semantically related ontology entities [3]. The increasing relevance of performing an ontology alignment process in several domains of application such as knowledge management, information retrieval, medical diagnosis, e-Commerce, knowledge acquisition, search engines, bioinformatics, the emerging Semantic Web and so on, has led to develop in years numerous tools, named *ontology alignment systems* [4][5]. Among all exploited techniques, due to the complex and time-consuming nature of the ontology alignment process, approximate methods have emerged as a successfully methodology for computing sub-optimal alignments [6]. From this point of view, evolutionary optimization methods [7][8] could represent an efficient approach for facing the problem, and, indeed, genetic

algorithms have been already applied to solve the ontology alignment problem as shown in [9][10] by reaching acceptable results. However, classical genetic algorithms suffer from some drawbacks such as premature convergence that makes them incapable of searching numerous solutions of the problem area.

Starting from these considerations, this research work investigates an emergent class of evolutionary algorithms, named Memetic Algorithms (MAs), to efficiently face the ontology alignment problem. MAs are population-based search methods which combine genetic algorithms and local refinements. This marriage between global and local search allows keeping high population diversity and reducing the likelihood premature convergence. Several different works demonstrate how MAs converge to high quality solutions more efficiently than their conventional evolutionary counterparts. In detail, the contribution of this thesis is to propose two ontology alignment systems, named *MemeOptiMap* and *MemeMetaMap*, which exploit MAs to produce an ontology alignment by following two different strategies. In particular, *MemeOptiMap* uses MAs to directly solve the ontology alignment problem as a minimum optimization problem. Instead, *MemeMetaMap* follows a meta-optimization approach by using Mas to tune the parameters necessary for performing an ontology alignment process. During the evaluation phase, both systems have been compared with the state of the art by means of a statistical multiple comparison procedure. The test results show that both approaches are competitive, and, in particular, *MemeMetaMap* improves the capabilities of the current ontology alignment processes by working regardless of the user involvement, data availability and the need of *a priori* knowledge about ontology features, and, yielding high performance in terms of alignment quality with respect to top-performers of well-known Ontology Alignment Evaluation Initiative[1] (OAEI), i.e., a coordinated international initiative aimed at providing means to compare and evaluate different ontology alignment systems.

[1] http://oaei.ontologymatching.org/

References

[1] T.R. Gruber. Towards principles for the design of ontologies used for knowledge sharing. Formal Ontology in Conceptual Analysis and Knowledge Representation. Kluwer Academic Publishers, Deventer, The Netherlands, 1993.

[2] M.R. Genesereth and N. J. Nilsson. Logical foundations of artificial intelligence. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987.

[3] S. Pavel and J. Euzenat. Ontology Matching: State of the Art and Future Challenges. IEEE Transactions on Knowledge and Data Engineering, 99[PrePrints], 2012.

[4] Y. Kalfoglou and M. Schorlemmer. Ontology mapping: the state of the art. Knowl. Eng. Rev., 18[1]:1–31, January 2003.

[5] H. Wache, T. Vogele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, and S. Hubner. Ontology-based integration of information - a survey of existing approaches. In Proceedings of the workshop on Ontologies and Information Sharing at the International Joint Conference on Artificial Intelligence (IJCAI), pages 108–117, 2001.

[6] T.C. Hughes and B. C. Ashpole. The Semantics of Ontology Alignment. In I3CON. Information Interpretation and Integration Conference, 2004.

[7] T. Back. Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms. Oxford University Press, Oxford, UK, 1996.

[8] J. H. Holland. Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, MI, USA, 1975.

[9] J. Wang, Z. Ding, and C. Jiang. Gaom: Genetic algorithm based ontology matching. In Proceedings of the 2006 IEEE Asia-Pacific Conference on Services Computing, APSCC ’06, pages 617–620, Washington, DC, USA, 2006. IEEE Computer Society.

[10] J. Martinez-Gil, E. Alba, and J. F. A. Montes. Optimizing Ontology Alignments by Using Genetic Algorithms. In Christophe Gueret, Pascal Hitzler, and Stefan Schlobach, editors, Nature inspired Reasoning for the Semantic Web (NatuReS), 419. CEUR Workshop Proceedings, October 2008.