

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

L'interoperabilità semantica rappresenta la capacità di due o più sistemi di interpretare significativamente l'informazione scambiata allo scopo di produrre risultati utili. Essa è una caratteristica essenziale per tutti i sistemi distribuiti basati su conoscenza progettati sia per applicazioni pubbliche che private. Sfortunatamente, raggiungere una piena interoperabilità semantica è molto difficile perché richiede che il significato di ogni informazione sia specificato con un appropriato livello di dettaglio in modo da risolvere qualsiasi possibile ambiguità.

Attualmente, la miglior metodologia per raggiungere tale livello di precisione nella specifica del significato dell'informazione è rappresentata dalle ontologie. Seguendo la definizione più accreditata [1], un'ontologia è un'esplicita specifica di una concettualizzazione, ossia, la formale specifica degli oggetti e dei concetti, e delle relazioni tra essi, che si presume esistano in un dominio d'interesse [2]. Comunque, differenti scopi e punti di vista conducono i progettisti a produrre diverse concettualizzazioni della stessa area di dominio. In altre parole, la soggettività caratterizzante la modellazione di un'ontologia comporta la creazione di ontologie cosiddette *eterogenee* caratterizzate da diversità terminologiche e concettuali. Esempi di queste diversità sono: usare differenti parole per indicare lo stesso concetto, impiegare la stessa parola per indicare concetti diversi, o ancora, la creazione di gerarchie per uno specifico dominio con differenti livelli di dettaglio, etc.. Il problema che ne deriva, conosciuto come *problema dell'eterogeneità semantica*, rappresenta a sua volta un ostacolo al raggiungimento dell'interoperabilità semantica.

Allo scopo di superare questo problema e prendere davvero vantaggio dall'usare una rappresentazione ontologica, la soluzione più solida è rappresentata dall' eseguire un cosiddetto *processo di allineamento ontologico*. Questo processo conduce due ontologie eterogenee in un mutuo accordo attraverso la rilevazione di un insieme di corrispondenze, chiamato *allineamento*, tra entità semanticamente relate [3]. La crescente importanza di eseguire un processo di allineamento ontologico in svariati domini applicativi come *knowledge management*, *information retrieval*, diagnosi mediche, *e-Commerce*, *knowledge acquisition*, motori di ricerca, bioinformatica, e l'emergente *Semantic Web*, ha portato a sviluppare negli anni numerosi *tools*, denominati *sistemi di allineamento ontologico* [4][5]. Tra tutte le tecniche, dovuta alla complessità del processo di allineamento ontologico, i metodi di approssimazione sono emersi come metodologie di successo per produrre allineamenti sub-ottimali [6]. Da questo punto di vista, i metodi di

ottimizzazione basati su *Computational Intelligence* [7][8] potrebbero rappresentare un approccio efficiente per affrontare il problema e, infatti, gli algoritmi genetici sono stati già investigati in alcuni lavori [9][10] ottenendo accettabili risultati. Comunque, questi algoritmi soffrono di alcune debolezze, come la convergenza prematura, che li rende non pienamente efficienti nella ricerca di soluzioni nello spazio del problema.

A partire da queste considerazioni, questo lavoro di ricerca investiga una classe emergente di algoritmi evolutivi, chiamata Algoritmi Memetici (MAs), per affrontare efficientemente il problema dell'allineamento ontologico. I MAs sono algoritmi basati su popolazione che combinano insieme gli algoritmi genetici e procedure di ricerca locale. Il matrimonio tra ricerca globale e locale permette di mantenere un'alta diversità nella popolazione e ridurre la probabilità di prematura convergenza. Sviziati lavori dimostrano come i MAs convergono a soluzioni con alta qualità più efficientemente delle loro controparti evolutive tradizionali. In dettaglio, il contributo di questa tesi consiste nel proporre due sistemi di allineamento ontologico, chiamati *MemeOptiMap* e *MemeMetaMap*, che sfruttano i MAs per produrre un allineamento ontologico seguendo due differenti direzioni. In particolare, *MemeOptiMap* impiega i MAs per risolvere il problema dell'allineamento ontologico in maniera diretta trasformandolo in un problema di ottimizzazione. Invece, *MemeMetaMap* segue un approccio di meta-ottimizzazione impiegando i MAs per ottimizzare i parametri necessari per l'esecuzione di un tipico processo di allineamento ontologico. Durante la fase di analisi e valutazione, entrambi i sistemi sono stati comparati con lo stato dell'arte attraverso una procedura statistica di multi-comparazione. I risultati del test statistico mostrano che entrambi gli approcci sono competitivi, e, in particolare, *MemeMetaMap* mantiene alte prestazioni in termini di qualità degli allineamenti prodotti riuscendo a competere con il *top-performer* riconosciuto dall'*Ontology Alignment Evaluation Initiative*¹ (OAEI), ossia, una organizzazione internazionale che ha come obiettivo fornire i mezzi per comparare e valutare i differenti sistemi di allineamento ontologico. In più, a differenza di esistenti sistemi di allineamento ontologico, *MemeMetaMap* ha l'enorme vantaggio di produrre allineamenti a prescindere dalla disponibilità di dati e di conoscenza *a priori* circa le caratteristiche delle ontologie sotto allineamento.

¹ <http://oaei.ontologymatching.org/>

Riferimenti

- [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.