Science of Computer Programming 95 (2014) 3–25

Contents lists available at ScienceDirect

Science of Computer Programming

www.elsevier.com/locate/scico

An extensible argument-based ontology matching negotiation approach

Paulo Maio *, Nuno Silva

GECAD – Knowledge Engineering and Decision Support Research Group, School of Engineering of Polytechnic of Porto, Rua Dr. Bernardino de Almeida 431, 4200-072 Porto, Portugal

HIGHLIGHTS

- A novel argument-based ontology matching negotiation approach is proposed.
- An explicit, formal, shared and extensible argumentation model is adopted.
- Experiments demonstrate the usefulness and pertinence of the approach.
- Easy to adapt and evolve the approach to support different scenarios' requirements.
- A Software Development Framework for the adoption of the proposed approach.

ARTICLE INFO

Article history:
Received 1 February 2013
Received in revised form 22 October 2013
Accepted 24 January 2014
Available online 31 January 2014

Keywords:
Ontology matching
Argumentation
Negotiation
Systems interoperability

ABSTRACT

Computational systems operating in open, dynamic and decentralized environments are required to share data with previously unknown computational systems. Due to this ill specification and emergent operation the systems are required to share the data's respective schemas and semantics so that the systems can correctly manipulate, understand and reason upon the shared data. The schemas and semantics are typically provided by ontologies using specific semantics provided by the ontology language. Because computational systems adopt different ontologies to describe their domain of discourse, a consistent and compatible communication relies on the ability to reconcile (in run-time) the vocabulary used in their ontologies. Since each computational system might have its own perspective about what are the best correspondences between the adopted ontologies, conflicts can arise. To address such conflicts, computational systems may engage in any kind of negotiation process that is able to lead them to a common and acceptable agreement.

This paper proposes an argumentation-based approach where the computational entities describe their own arguments according to a commonly agreed argumentation meta-model. In order to support autonomy and conceptual differences, the community argumentation model can be individually extended yet maintaining computational effectiveness. Based on the formal specification, a software development framework is proposed.

1. Introduction

More and more computational systems (e.g. agents, web services) operating in open, dynamic and decentralized environments (e.g. semantic web, e-commerce, peer-to-peer, agent-based systems) require information sharing with previously unknown systems. Due to this ill specification and emergent operation, the computational systems are now required to

* Corresponding author. Tel.: +351 22 834 05 00.
E-mail addresses: pam@isep.ipp.pt (P. Maio), nps@isep.ipp.pt (N. Silva).

0167-6423/$ – see front matter
http://dx.doi.org/10.1016/j.scico.2014.01.011

share the data's respective schemas and semantics, so that the systems can correctly manipulate, understand and reason upon the shared data. The schemas and semantics are typically provided by ontologies using specific semantics provided by the ontology language. Nevertheless, computational systems maintain their autonomy and conceptual specificities, leading to different ontologies and thus preventing the direct information sharing. Accordingly, a successful systems interaction relies on the ability to reconcile their ontologies in run-time. In literature the ontology reconciliation problem is usually referred to as Ontology Matching [1]. A reconciliation process consists of establishing a set of correspondences (referred to as alignment) between the system's ontologies, which are further exploited to interpret or translate exchanged messages and their content. Therefore, systems need to autonomously decide on each and all correspondences between the ontologies they adopt in a conversation/interaction. For that purpose, a common approach found in literature consists in providing an ontology matching service such that the interacting computational systems agree (implicitly or explicitly) on using that service and, therefore, an alignment is requested as needed. However, ontology matching is a burdensome and error-prone process due to different factors. Firstly, because of the different applied semantics of the ontology languages and modeling approaches. Secondly, because of the conceptual interpretation of the linguistic dimension of the ontology, which typically grounds the ontology to the domain of knowledge, but unfortunately is a source for multiple interpretations and therefore for matching ambiguities. Consequently, the ontology matching process can lead to different and contradictory results (i.e. alignments) depending on the adopted matching approaches. Thus, considering that distinct computational systems may have different needs and objectives and, therefore, different preferences concerning the matching process, computational systems may be able to exploit the matching services they find more convenient instead of relying on a common matching service. For example, a computational system may prefer alignments having a high recall in disfavor of precision, while the other one may prefer precision instead of recall. In scenarios like the one described above, i.e. where each interacting computational system may adopt its most suitable matching service, it is necessary to provide a mechanism enabling those systems to avoid and/or resolve possible alignment conflicts. In that sense, state-of-the-art literature refers to two negotiation-based approaches: relaxation-based [2] and argument-based approaches [3,4].

This paper proposes a novel argument-based approach where arguments are described according to a state-of-the-art argumentation meta-model that captures general argumentation semantics. Moreover, the adopted meta-model is first instantiated by the negotiating community into a community argumentation model capturing the commonly agreed arguments (types or schemes) regarding the domain application. Further, in order to support autonomy and conceptual differences between individual systems, the community argumentation model can be individually extended, yet maintaining computational effectiveness.

Based on the formal specification (Sections 3 and 4), a software development framework is proposed and its architecture and design are discussed (Section 5). Examples and experiments adopting the proposals are finally presented (Section 6). Yet, in order to introduce the reader to important concepts and terminology, the next section revises important background knowledge.

2. Background knowledge

First, this section concisely surveys the ontology matching domain. Further, it defines the ontology matching negotiation problem and briefly describes current state-of-the-art approaches.

2.1. Ontology matching

Ontology matching is seen as the process of discovering, (semi-) automatically, the correspondences between semantically related entities of two different but overlapping ontologies. Thus, as stated in [1], the matching process is formally defined as a function $f:(O_1, O_2, p, res, A) \rightarrow A'$ which, from a pair of ontologies to match O_1 and O_2, a set of parameters p, a set of oracles and resources res and an input alignment A, it returns an alignment A' between the matched ontologies. Ontologies O_1 and O_2 are often denominated as source and target ontologies, respectively. An alignment is a set of correspondences expressed according to:

- Two entity languages Q_{L_1} and Q_{L_2} associated with the ontology languages L_1 and L_2 of matching ontologies (respectively) defining the matchable entities (e.g. classes, object properties, data properties, individuals);
- A set of relations R that is used to express the relation held between the entities (e.g. equivalence, subsumption, disjoint, concatenation, split);
- A confidence structure ϕ that is used to assign a degree of confidence in a correspondence. It has a greatest element $\top$ and a smallest element $\perp$. The most common structure is the real numbers in the interval $[0, 1]$, where 0 represents the lowest confidence and 1 represents the highest confidence.

Hence, a correspondence (or a match) is a 4-tuple $c = (e, e', r, n)$ where $e \in Q_{L_1}(O_1)$ and $e' \in Q_{L_2}(O_1)$ are the entities between which a relation $r \in R$ is asserted and $n \in \phi$ is the degree of confidence in the correspondence.

Over recent years, research initiatives in ontology matching have developed many systems (e.g. [5]) that rely on the combination of several basic algorithms yielding different and complementary competencies, to achieve better results. A basic algorithm generates correspondences based on a single matching criterion [6]. These algorithms can be multiply classified