

**01**

E-BUSINESS INTEROPERABILITY THROUGH ONTOLOGY SEMANTIC MAPPING

Nuno Silva, João Rocha

*GECAD - Knowledge Engineering and
Decision Support Research Group
Instituto Superior de Engenharia do Porto
4200-072 Porto – Portugal
{Nuno.Silva, Joao.Rocha}@dei.isep.ipp.pt*

José Cardoso

*Engenharias, Universidade de Trás-os-Montes e Alto Douro
Apartado 202, 5001-911 Vila Real - Portugal
jcardoso@utad.pt*

With the advent of Semantic Web, knowledge-based interoperability in VE faces a new technological shift, in which ontologies and semantic web technologies plays a major role. Exploiting the explicit semantic description of the domain of discourse allows reasoning and automatically acquiring semantic relations between two different domains of discourses. Such semantic relations would be further applied in converting data between such domains. This process is named Ontology Mapping. The MAFRA–Mapping FRamework cover all the phases of the process, including analysis, specification, representation, execution and evolution. The execution strategy and methodology are the focus of this paper. The MAFRA Toolkit is being applied in the European project Harmonise, which aims to provide solutions for (semi-) automatic interoperability between major operators in e-tourism.

1. INTRODUCTION

Knowledge interoperability is becoming a key issue in Virtual Enterprises (VE) due to the limitations of data integration. Data integration lacks important features required in VE, like flexibility, interoperability and automatic configuration of interactions between agents engaged in an specific business activity. Semantic Web, the next WWW trend, suggests the annotation of Web resources with machine-processable metadata, which can provide tools to analyse meaning and semantic relations between documents and their parts. Ontologies as means for conceptualizing and structuring knowledge are seen as the key to the realization of the Semantic Web vision. Ontology allows the explicit specification of a domain of discourse, which permits to access to and reason about an agent knowledge. Semantic Web and ontologies are therefore fully geared as a valuable framework for distinct applications, namely business applications like E-Commerce, B2B and Virtual Enterprises.

However, it is hardly conceivable that a single ontology is applied for all domains and interactions, which raises new interoperability problems. Ontology Mapping intends to solve this problem by defining semantic relations between ontologies (at conceptual level), and apply those relations at data level over in transforming ontology instances between ontology representations.

The work described in this paper is developed in the scope of MAFRA (Mapping Framework), a conceptual framework (Maedche, A. *et al*, 2002) covering all phases of the ontology mapping process. MAFRA approach aims to maximize declarativity, complexity, expressiveness and modularization (Silva, N. and Rocha, J., 2003). This paper focus on the automatic bridging phase of MAFRA process. Semantic bridging is a highly subjective time-consuming task, demanding extensive domain expertise.

In Section 2, the essence of ontology-based interoperability is described. Section 3 describes generically the MAFRA architecture, in which services play a crucial role. Section 4 describes the new automatic bridging methodology. Section 5 will present a short overview of similar or related projects found in literature. This subject is presented at this point in order to argue on benefits and limitations of MAFRA proposals. Finally, Section 6 will provide an overview of the achieved results and point out some current and future efforts. It is considered that the Semantic Bridging Ontology and MAFRA ontology mapping methodology are previously known (Silva, N. and Rocha, J., 2003).

2. ONTOLOGY-BASED INTEROPERABILITY

The ultimate goal of MAFRA is to provide mechanisms to ease and automate knowledge interoperability in agent-based systems, namely concerning Virtual Enterprise and E-Business scenarios. Agents are computer-based entities acting on behalf and according to real world entities goals (Jennings, N. and Wooldridge, M., 1996). Agents are autonomous, reactive, pro-active and socially able. As consequence, agents would require advanced interoperability capabilities, namely concerning communications, interaction languages and knowledge sharing. With the advent and expansion of Semantic Web technologies, ontologies became a more and more frequent approach to describe knowledge in an independent but powerful manner.

Type and extension of heterogeneity between two ontologies depends on different dimensions of the ontologies, specially, described domain, application, granularity (Silva, N. and Rocha, J., 2003). Generally, any ontology is built on three distinct layers: the model layer, the axiomatic layer and lexical layer. The model layer specifies domain and/or application entities, their inter-relations (e.g. `subClassOf`) and properties. The axiomatic layer constraints the interpretation and application of entities through axioms or rules (e.g. parents of an instance of Person are instances of Person). The lexical layer characterizes entities and their properties with natural language lexicons, giving them a real word meaning (e.g. XPTO entity corresponds to real world entity Person or Individual). Besides the obvious benefits in ontology development and evolution, these three layers separation provide independent elements to reason upon, allowing the argumentation and determination of the extension and causes of heterogeneity.

3. SERVICE-ORIENTED APPROACH

One of the ultimate innovations in MAFRA is its service-centric approach, which will be described in this section.

Ontology mapping aims to define semantic relations between source ontology entities and target ontology entities, which will be further applied at instance level to transform source instances into target instances. Semantic relations are realized through the semantic bridges:

```
semanticBridge(TR, SE, TE, SC)
```

where TR is the process to be applied in transforming instances of the source entities into instances of the target entities, SE is a subset of source ontology entities considered to create the TE, the subset of target ontology entities. Finally, SC is the set of condition expressions constraining the execution of the semantic bridge.

Transformations requirements vary enormously from mapping to mapping and from semantic bridge to semantic bridge. Therefore, it is virtually impossible to provide all possible transformation requirements in a centralized static ontology mapping system. This observation lead to the adoption of a modular, decentralized approach, where independent transformation modules are attached to the system functional core modules (e.g. bridging, execution, negotiation, evolution). These modules are called Services and comply with a specific interface, acting as intermediary between services and functional core modules.

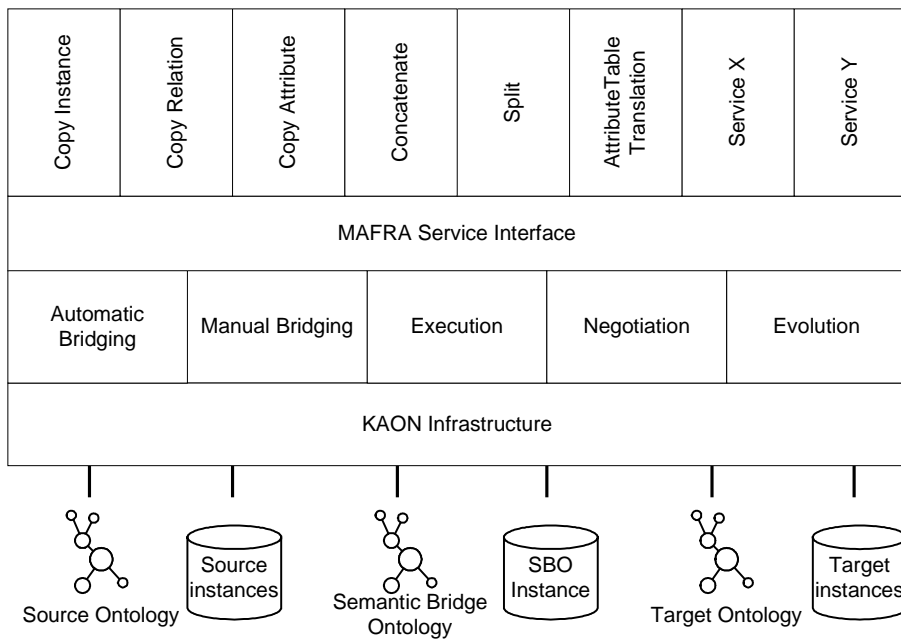


Figure 1 – MAFRA System Architecture

Besides the variety of transformation requirements, it is possible to distinguish between two types of transformations according to the type of the target entity (Silva, N. and Rocha, J., 2003):

- When the resulting instance is a concept instance, a new instance will be created and uniquely identified. A concept instance is a container for property instances;
- When the resulting instance is a property instance than the new property instance will always be attached to a previously created concept instance.

Pragmatically this distinction leads to identification of two types of semantic bridges:

- ConceptBridges, those who create target instances, and always apply the CopyInstance transformation procedure. The CopyInstance service is universal, in the sense that all characteristics of the transformation are known *a priori* and are available in the service;
- PropertyBridges, whose service varies depending on specific transformation requirements.

Each Service is uniquely identified and is characterized by a set of arguments, defining its interface. In the current implementation of MAFRA several transformation services are available, providing sufficient transformation capabilities for a wide range of cases. Table 1 presents the CopyInstance service interface.

Table 1 – Example of standard Services available in MAFRA Toolkit

CopyInstance	Creates target concept instances for each source concept instance that fulfills specified conditions.	
Argument ID	Argument Type	Comment
Source Concept	Concept	Source ontology class whose instances will be transformed
Target Concept	Concept	Target ontology class to create
Extensional Specification	ArrayOfConditions	Extensional definition of source class instances
Generic Conditions	ArrayOfConditions	Constraint of the bridge execution
Minimum Cardinality	Integer	The minimum number of instances to translate
Maximum Cardinality	Integer	The maximum number of instances to translate

Semantic bridges entities are connected to service arguments through a map of argument identifiers and values. For example¹:

```
CB0=semanticBridge( CopyInstance,
                    {sourceConcept=<O1>C0},
                    {targetConcept=<O2>C0},
                    {})
```

4. AUTOMATIC BRIDGING

Semantic bridging is a highly subjective time-consuming task, demanding extensive domain expertise. MAFRA provides support and automation facilities.

Early stages in the semantic bridging phase typically aim to capture lexical similarities between source and target entities. Such task is very subjective since

¹ Finer details about the MAFRA ontology mapping process can be found in (Silva, N. and Rocha, J., 2003).

domain expert deals with natural language description of entities. Recurring to lexical tools, like dictionaries, specific domain thesaurus and WordNet, it is possible to multi-classify similarities between pair of entities. These classifications are then combined into an overall similarity measure. The result of these early steps is a set of pairs in the form of <Source ontology entity, Target ontology entity>, denoting semantic equivalence between source and target ontology entities. According to a pre-defined threshold level, the similarity pair is kept or disregarded. However, the same entity (source or target ontology entity) might be present in several pair forming therefore a complex n:m map.

Semantic bridges shall be instantiated according to such similarities. However, due to those complex relations and multitude of possible transformation procedures, a centralized rule-based system would be too complex and static, because services are very different from each other, have different requirements on the type and number of necessary ontology entities, and finally because services are dynamic in number and type.

Typically, when the problem is too complex, one of the possible solutions is the division of the problem into smaller problems. Such division and modularization is already present in the system (independent modular services), and it seems reasonable to exploit it.

Hence, the proposed process consist in pushing each and all similarity pairs to each and all available services (Figure 2), which in turn determine if the similarity pair:

- should be added to an already existent semantic bridge, in which case entities are attached to the respective arguments;
- is relevant for the creation of a new semantic bridge, in which case a new provisory semantic bridge is created with corresponding service, and entities attached to the respective arguments;
- invalidates an already existent semantic bridge, in which case the provisory semantic bridge is deleted.

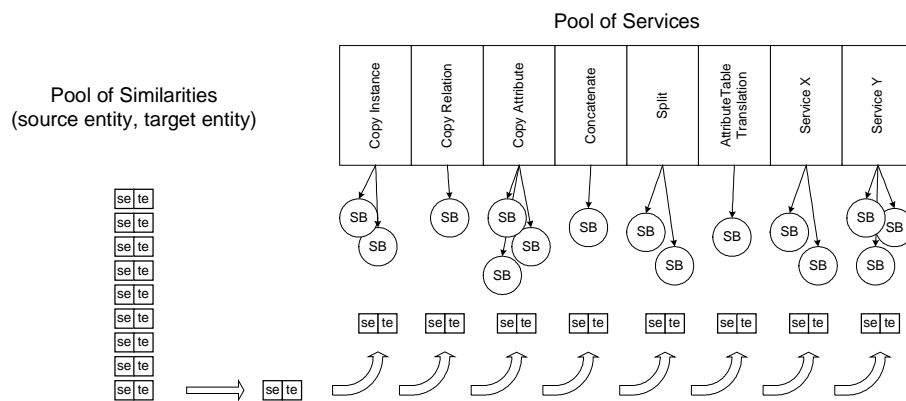


Figure 2 – Figurative representation of the automatic bridging process

In order to clarify the proposed methodology, lets follow the example represented in Figure 3. Four similarity pairs and two services considered.

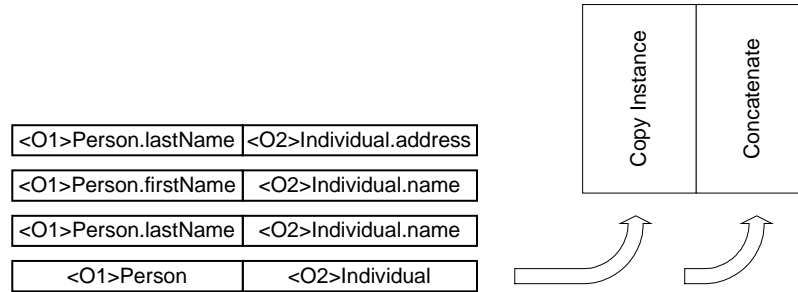


Figure 3 – Automatic Bridging example

To be attached to a semantic bridge, the Copy Instance service requires a similarity pair of concepts. The Concatenate service requires a set (cardinality greater than 1) of similarity pairs, in which the target entity is the same in each and all similarity pairs. The semantic bridge becomes invalid if the target entity is similar to a source entity which in turn is similar to more than one target entity. Source and target entities must be attributes (literals). In this example, the first pair (lowest pair in the heap) will give rise to a provisory semantic bridge with Copy Instance service (CB1). Second similarity pair is irrelevant for Copy Instance service, since the type of entities required by service do not match. Besides, second pair is accepted by the Concatenate service in order to create a provisory semantic bridge (PB1). The third pair is also accepted by Concatenate in order to complement PB1. Like both previous pairs, fourth pair is refused by Copy Instance service. Concatenate service on the contrary considers it, since both entities are attributes. However, the pair's source entity is already mapped in PB1 with different target entity, which violates one of the Concatenate service constraints. In such circumstances, the Concatenate service becomes invalid and PB1 bridge is deleted. However, let's consider that fourth similarity pair was disregarded, and PB1 is valid. The resulting semantic bridges would be:

```

CB1=semanticBridge( CopyInstance,
                    {sourceConcept=<O1>Person},
                    {targetConcept=<O2>Individual},
                    {} )

PB1=semanticBridge( Concatenate,
                    {sourceAttributes=list(
                      <O1>Person.firstName,
                      <O1>Person.lastName) },
                    {targetAttribute=<O2>Individual.name},
                    {} )

```

Finally, it is necessary to inter-relate semantic bridges. This process is named inter-bridging and aims to find a concept bridge whose concepts are root of all properties in a property bridge. If no concept bridge is found, an inferred concept bridge is proposed to the domain expert between root concepts. In this example, root concepts (<O1>Person and <O2>Individual) are already bridged, and therefore the scope is set:

```
hasBridge(CB1, PB1)
```

5. RELATED WORK

Three distinct ontology mapping projects should be pointed as paradigmatic approaches. In (Park, J.Y. *et al*, 1998) an extension to Protégé has been described, consisting of a definition of the mapping between domain ontologies and problem solving methods. This approach defines a valuable set of desiderata and mapping dimensions, but lacks some important features, especially in allowing mapping between multiple concepts. Nevertheless, there is no record of experiments that apply it to the Semantic Web environment. The second approach is RDFT (Omelayenko, B., 2002), a meta-ontology that describes Equivalence and Versioning relations between either an XML DTD or RDFS document and another XML DTD or RDFS document. An RDFT instantiation describes the semantic relations between source and target documents, which will be further applied in the transformation of documents. Thirdly, the Buster project (Stuckenschmidt, H. and Wache, H., 2000) applies information integration to the GIS domain. Two distinct approaches were proposed: rule-based transformation and re-classification. The rule-based approach applies a procedural transformation to instance properties, while classification applies class membership conditions to infer target classification through description-logic tools. However, these two approaches are not integrated, which limits mapping capabilities.

In general MAFRA capabilities go beyond those provided by prior projects (Silva, N. and Rocha, J., 2003). In fact, MAFRA is the first approach to integrate all phases and dimensions of the ontology mapping process into an overall perspective. Semantic Bridge Ontology is some how similar to RDFT, but MAFRA allows a wider range of semantic relations. Graphical User interface and automatic bridging methodology are not mentioned in any of the prior projects. However, MAFRA lacks substantial application and integration experience and do not currently support ontology evolution.

6. CONCLUSIONS AND FUTURE WORK

MAFRA tackles very complex ontology mapping problems with its standard transformation services, but its architecture allows for an easy integration of additional services, which proves to be important on the semantic web's most common applications: e-business and virtual enterprises. The MAFRA Toolkit implements the ideas described in this paper, providing domain expert with an intuitive, easy to use and integrated GUI (Figure 4).

MAFRA was adopted as the representation and transformation engine core technology for the Harmonise project. Harmonise uses an "Interoperability Minimum Harmonisation Ontology" (IMHO) as *lingua franca* between agents. MAFRA is responsible for the acquisition, representation and execution of the ontology mapping between each agent specific ontology and IMHO. Application with real world ontologies and instances confirmed MAFRA natural skills to deal with very complex semantic relations. In fact, only few improvements were necessary to standard services to completely fulfill transformation requirements.

Our efforts are currently focused in improving the automatic bridging process, specially considering the case where several semantic bridge are suggested for the

same similarity pair. We are also interested in the evolution of the ontologies and its consequences to the ontology mapping process. A longer-term project should facilitate the mapping acquisition between different agents using meaning negotiation. Our longer-term research concentrate in negotiation of ontology mapping using argumentation.

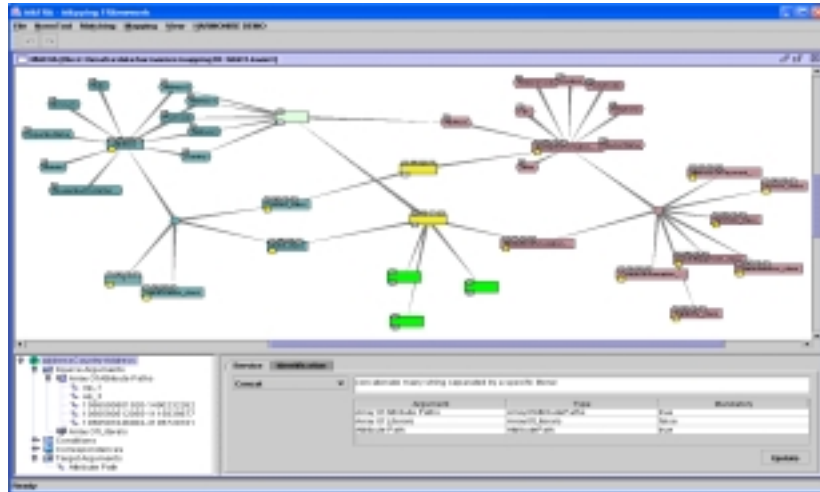


Figure 4 - Screen shot of the Graphical User Interface of MAFRA Toolkit

7. ACKNOWLEDGMENTS

This work is partially supported by the Portuguese MCT-FCT project POCTI/2001/GES/41830. Many thanks to Alexander Maedche for his ideas and support. Thanks to Oliver and Mirella Dell'Erba for their feedback on the application of MAFRA in the Harmonise project. Thanks to Jorge Santos for his opinions and revisions.

8. REFERENCES

1. Maedche A, Motik B, Silva N, and Volz R; MAFRA - An Ontology MAPPING FRAMework in the Context of the Semantic Web; Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02); 2002; 235-250.
2. Silva N and Rocha J; MAFRA - An Ontology MAPPING FRAMework for the Semantic Web; Proceedings of the 6th International Conference on Business Information Systems; Colorado Springs (CO), USA; 2003.
3. Silva N and Rocha J; Semantic Web Complex Ontology Mapping; Proceedings of the Web Intelligence'2003; Halifax, Canada; 2003.
4. Jennings, Nick and Wooldridge, Michael; Software Agents; IEE Review; 1996; 17-20.
5. Park JY, Gennari JH, and Musen MA; Mappings for Reuse in Knowledge-based Systems; 11th Workshop on Knowledge Acquisition, Modelling and Management (KAW 98); Banff, Canada; 1998;
6. Omelayenko B; RDFT: A Mapping Meta-Ontology for Business Integration; Proceedings of the Workshop on Knowledge Transformation for the Semantic Web (KTSW 2002) at ECAI2002; Lyon, France; 2002; 76-83.
7. Stuckenschmidt H and Wache H; Context Modeling and Transformation for Semantic Interoperability; Knowledge Representation Meets Databases; 2000; 115-126.