Ontological Association in the Ground Target Domain

Jenny Lagerlöf, Egils Sviestins, and Klas Wallenius[1]

***Abstract*—This paper addresses the problem of associating reports from sensors and human observers to objects in a ground situation picture. Both reports and known objects are assumed to be expressed according to an ontology, in this case expressed in the OWL (Web Ontology Language) formalism. The association here does not consider spatial information; it is entirely based on the class hierarchies and properties. The technique combines results from Information Retrieval Theory for the Semantic Web with an ontology for the Ground Domain and applies to a scenario demonstrating how the potential associations are evaluated and ranked.**

***Index Terms*—Ontology, association, ground targets, intelligence**

I. Introduction

The main difficulty in tracking ground objects is data association. One may discuss three categories of difficulties. The first category is related to shortcomings in sensor performance, such as insufficient coverage, low dimensionality, and insufficient discrimination between relevant reports and background, especially if the target does not move. The second category relates to the diversity of report and target characteristics, e.g. one report may have information about a four wheel vehicle, while the object data base contains cars and armored vehicles. The third category is about the scarcity of reports; often very long time passes between the observations, and in such cases data association based on locations only can be highly ambiguous or even impossible.

In the air and sea domains one can often get adequate tracking performance just by using kinematic data in combination with more or less sophisticated association techniques – nearest neighbor, maximum likelihood, JPDA etc, perhaps in combination with multiple hypothesis tracking. In the ground domain this will be enough only under very favorable circumstances. A more typical situation is where the operator, e.g. an intelligence analyst, must rely on reconnaissance reports together with snippets of sensor information from areas where sensors do exist. This creates a very complex situation picture that puts high demands on the analyst in matching the right report with the right track.

Data association should thus be based both on kinematics and on extra information. We here assume that the extra information is formalized in terms of an ontology. In this paper we leave the kinematic association aside and focus entirely on the ontological association.

This association technique can be characterized as semi-automatic: Basically it suggests, or ranks, possible associations to the operator. When the association is sufficiently unambiguous, according to yet unspecified criteria, one should allow the association to be effectuated automatically.

We will describe a scenario comprising a number of ground objects and a sequence of reports (without location information), and apply two techniques to obtain a ranking on how associable each report is to each object. The first technique excludes certain potential report-to-object associations based on the ontology at hand. The second technique, inspired by information retrieval theory, ranks the potential associations according to specificity; e.g if you are tracking a Jeep, then a report on a Jeep is more likely to represent this particular Jeep than a report about a vehicle of unknown kind. Finally we will show the results of these association techniques to the scenario.

We are not aware of any (semi-)automatic method that our technique can be evaluated against. It would be interesting though to carry out field experiments to find out the operators' perception, comparing with the traditional highly manual handling of intelligence information. We are however not there yet. The contribution of this work is thus to suggest a way to get a better ground situation picture, by incorporating reports with characteristics other than purely kinematic.

This paper is largely based on a Master's Thesis work by one of the authors [1].

II. Scenario

In the toy scenario that we consider, see Figure 1, there are the following objects:

- A mechanized infantry platoon with three MT-LB tracked armored personnel carriers; armament: machine gun
- A tank platoon with four T-90 battle tanks, armament: a

[1] Authors' affiliation is Saab AB, Järfälla, Sweden. Mail addresses are jenny.lagerlof@saabgroup.com, egils.sviestins@saabgroup.com and klas.wallenius@saabgroup.com

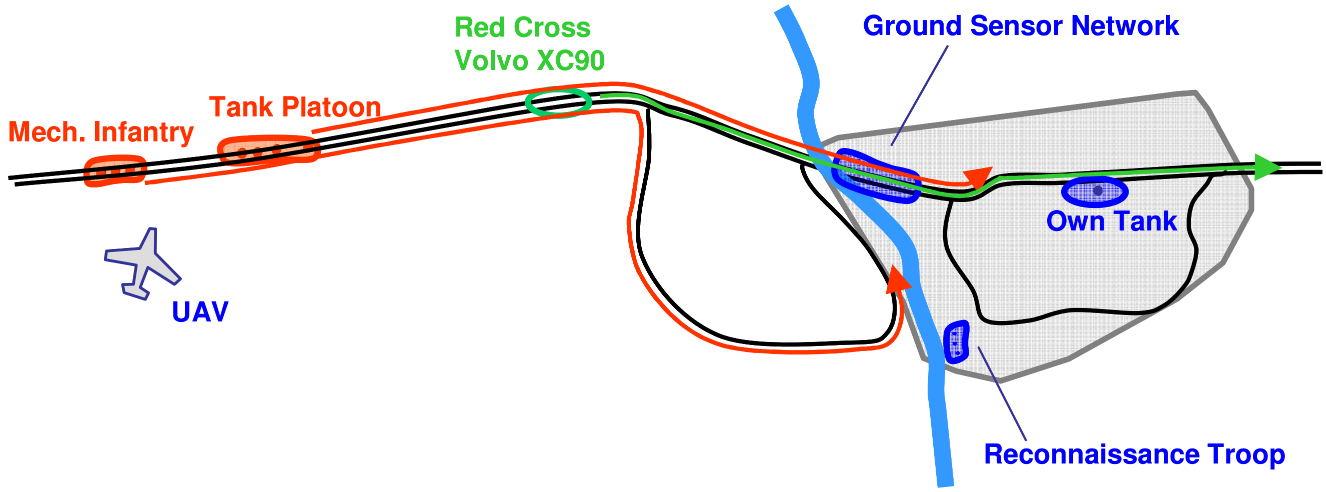

Figure 1 The scenario. A mechanized infantry platoon, a tank platoon and a Red Cross Volvo XC90 move towards a defended area (gray), and are observed by a UAV, a ground sensor network, a reconnaissance troop, and a tank commander.

cannon and two machine guns

- A Red Cross Volvo XC90 SUV

The following sources provide reports

- A UAV observer
- A ground sensor network (GSN)
- A reconnaissance troop (REC)
- A friendly tank commander (TCR)

The following reports are provided, in turn

- UAV1: "Three vehicles with machine gun"
- UAV2: "Four vehicles with machine gun and cannon"
- GSN1: "Light wheeled vehicle"
- GSN2: "Three heavy tracked vehicles"
- REC: "Three armored personnel carriers"
- TCR: "A Red Cross Volvo XC90"

Here e.g. UAV2 labels the second report from the UAV.

The question is now: which of these reports should be associated with which object? Note that we here have omitted the positional information; this is to be handled by the kinematic association part which is not addressed in this paper.

Most of the associations may seem obvious to a human, but for the data association engine, having an ontology as the only means of understanding the situation, this is not the case.

We will return to this scenario in Section IV.

III. Association

A. Ontology

The main efforts today in digitizing the battle field reports rely on extensive formalizing of the information, with the hope to minimize the need for textual information and to maximize the possibilities of automated processing. Probably the most comprehensive work in this direction is made in the Multilateral Interoperability Program, MIP, [2]. The results are given in their C2 Information Exchange Data Model, C2IEDM, for the land domain, now developed further into JC3IEDM for the joint air, land and sea domains.

Ontologies can be expressed in many ways, and in a sense JC3IEDM is an ontology. However, the most popular way of expressing ontologies today is using OWL, The Web Ontology Language [3], endorsed by the World Wide Web Consortium. For the purpose of this study, we started with a tightly JC3IEDM related information model in a Swedish battle management system (known as "SLB"), and expressed part of it simplified in OWL. The resulting ontology is shown in Figure 2.

Information models may look different depending on the intended usage, even though they describe the same domain. This became apparent when we defined our ontology. The JC3IEDM is intended to serve as a model for information exchange between different systems, and its definitions are made to support military commanders in their decisions, while, in our case, we need an ontology that can be used to discriminate between different sensor reports. As an example, a 'Tank' in JC3IEDM is a subclass to 'Weapon', while an 'Armored Personal Carrier' is a subclass to 'Vehicle'. This makes sense from the commander's perspective as the main purpose of a tank is to use it as a weapon while the carrier mainly is used for transportation. However, when a sensor observes a tracked vehicle with a gun (without knowing the weight of the vehicle or the size of the gun) it is difficult to know whether it is a tank or a carrier. Thus it makes more sense from our perspective to let 'Tank' be a subclass to 'Vehicle' instead of 'Weapon'. Considering all such different cases that may occur, there will be a substantial effort to design a useful ontology when developing a full scale system supporting ontological association.

Incoming reports may need to be classified with respect to the ontology, i.e., one identifies the most specific class in the ontology that is supported by the report. For example, a report on something with hasAttribute = Wheeled could describe an Automobile, but is not classified as such unless it also has hasWeight = LightWeight. For this reason we want to keep the ontology relatively simple. If the ontology in addition

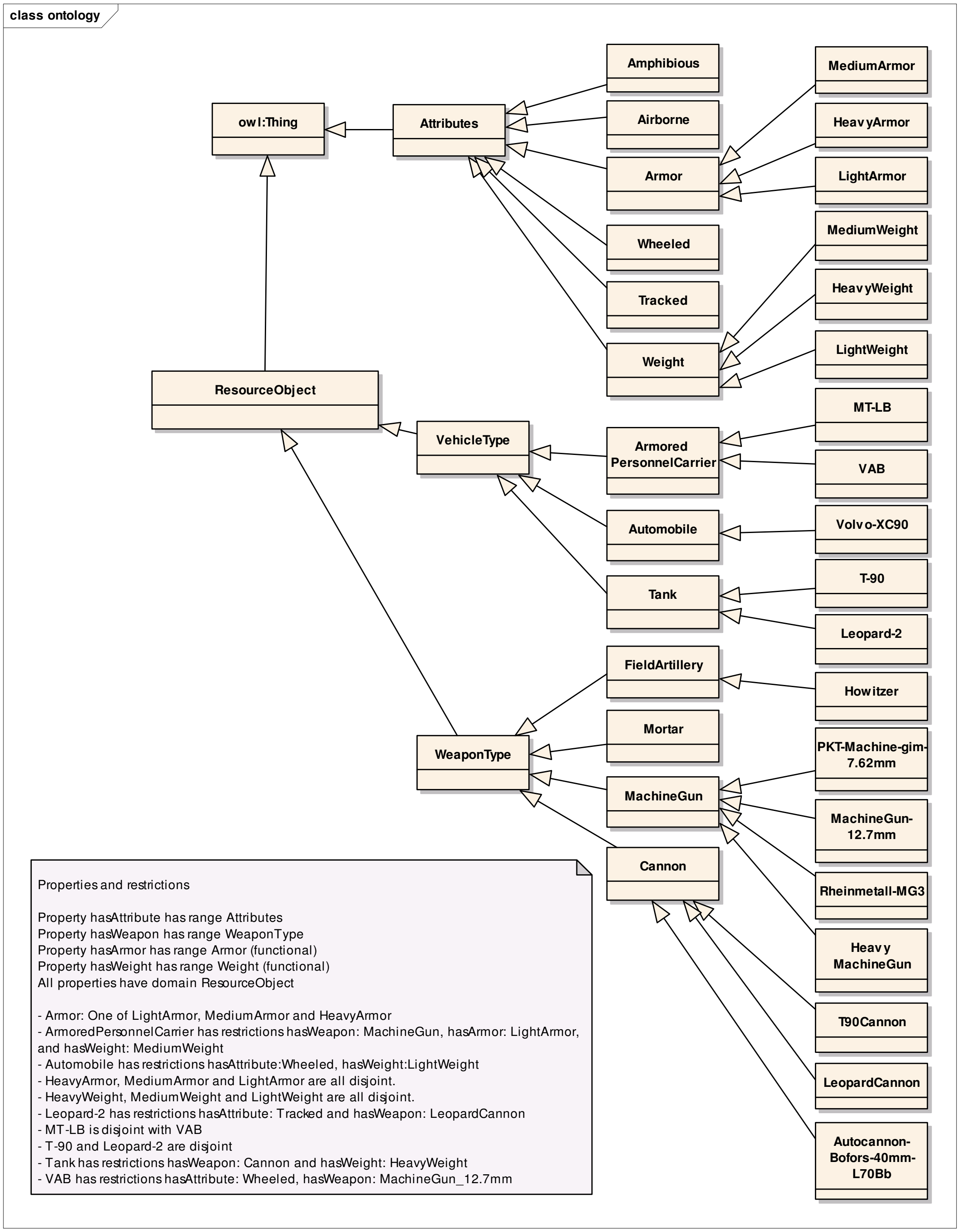

Figure 2 The simplified ontology

had stated that an Automobile has a GloveCompartment, then only rarely would a report on an automobile actually be classified as such. To handle cases like this we actually state explicitly (not shown in Figure 2) that the combination of Wheeled and LightWeight is necessary and sufficient for the classification. A probabilistic approach should be the best way to capture the fact that glove compartments are common and observable but almost never reported.

B. Consistency

Let us now assume that the known existing objects, EOs, are expressed as individuals (instances) of the ontology classes, and that the information sources likewise provide reports, or incoming objects, IO, as individuals. The first question to answer is whether an IO can possibly represent an EO. As a realistic ontology is quite complex, this can be far from easy to determine. However, as OWL (in Light and DL versions) is based on Description Logics [4], there is an unambiguous answer to the question, and there are tools that will do the work. Kokar et al have described the technique in [5] using a tool called ConsVISor; we have relied on the Pellet reasoner [6]. The reasoner is responsible for the classification mentioned in Section IIIA. For making the consistency check one introduces the OWL property "sameAs" between the IO and a candidate EO and feeds the reasoner with the ontology including the individuals. The reasoner will then discover if there is an inconsistency, and thus if the association is impossible. One similarly tries all report and existing object combinations.

Obviously an inconsistency might be due to an error in the report as there is usually some degree of uncertainty. This problem, although extremely important, is out of the scope of this paper. Here a report that is inconsistent with every known object is either handled as a new object, or it is left to an operator to make a decision.

C. Similarity

It will most likely turn out that the IO is consistent with many EOs. To choose the most likely EO, or at least to assist the operator in making the decision, we want to rank the possible associations. The more similar the IO is to an EO, the higher rank. One way of making the ranking is based on calculating a distance between two items in the ontology [7]. The distance is represented by the path from one item to the other via the closest common parent, and is defined so that adjacent objects high up in the hierarchy are considered more distant than those further down. This technique does however not utilize the information given by attributes and constraints. We will therefore instead introduce a technique that is highly inspired by Information Retrieval Theory applied to the Semantic Web. The main reference is [8]. The problem handled there is how to rank various web services with respect to the user needs. Two fundamental concepts in information retrieval are recall, *r*, and precision, *p*. Recall is the proportion of all relevant documents retrieved in answer to a search request. Precision is the proportion of retrieved documents that are actually relevant. For example, suppose that the user wants information about Italian (I) cars (C). One service may provide information about Italian sports (S) cars only; there the recall is low. A second service may provide information about all kinds of cars. The user will be flooded by irrelevant information; i.e., the precision is low. A third service provides information about sports cars. Here both precision and recall will be low. A perfect match between user request and service will give both the precision and the recall the value one.

In reality it is difficult to count the number of responses to compute the recall and precision. Instead one can try to estimate the quantities by looking at the number of overlapping features in the request and in the service announcement [8]. In this way,

$$r(F_R, F_S) \approx \frac{|F_R \cap F_S|}{|F_S|} \quad (1)$$

$$p(F_R, F_S) \approx \frac{|F_R \cap F_S|}{|F_R|} \quad (2)$$

where F_R is the set of features in the request, and F_S is the set of features in the service announcement. For example, in the case of cars we have $F_R\{I \quad C\}$, and the first service has $F_S = \{I \quad S \quad C\}$, giving $r = 2/3$, and $p = 1$.

To have a single measure for ranking purposes one can combine the recall and the precision according to one's needs. A standard way is to compute a weighted harmonic mean, also known as the F-measure:

$$q = \frac{1}{\frac{w_r}{r} + \frac{w_p}{p}} \quad (3)$$

where w_r and w_P are the weights given to precision and recall, $w_r + w_p = 1$.

Matters are complicated when these ideas are applied to an ontology, as it is not obvious how to count the number of features. A solution suggested in [8] is to consider the number of class ancestors (excluding "thing" at the top level but including the class in question) for the requestor and the service announcement, and thus, using (1) and (2), arrive at recall and precision based on class hierarchy, $r_A(C_R, C_S)$ and $p_A(C_R, C_S)$ respectively (*"A"* stands for "Ancestor"). The classes are further characterized by a number of properties, so similarly one computes recall and precision based on class properties $r_P(C_R, C_S)$ and $p_P(C_R, C_S)$. The issue is even more complex than that as the properties themselves form hierarchies and have restrictions, but that is out of the scope of this paper.

Now we would like to apply these ideas for measuring the similarity between the IO and EO in our data association problem. As we no longer deal with information retrieval, the original meaning of precision and recall is lost, but we will still keep the terminology. We thus simply compute $r_A(IO, EO)$, $p_A(IO, EO)$, $r_P(IO, EO)$ and $p_P(IO, EO)$, and combine them into a single measure that gives the desired ranking. In our tests it appears that precision and recall based on class hierarchy is less important than precision and recall based on properties. The reason for this is that the object classifications often land high up in the class hierarchy. We have therefore chosen the combined recall and precision values as

$$\begin{aligned} r(IO, EO) &= 0.35\, r_A(IO, EO) + 0.65\, r_P(IO, EO) \\ p(IO, EO) &= 0.35\, p_A(IO, EO) + 0.65\, p_P(IO, EO) \end{aligned} \quad (4)$$

The final ranking value *q* is computed from (3) and (4) with equal weights on precision and recall.

An example will clarify the calculations. Suppose the incoming object is of type ArmoredPersonnelCarrier with properties

-hasArmor LightArmor
-hasWeapon MachineGun
-hasWeight MediumWeight

and that the existing object is of type VehicleType and has the properties

-hasWeight MediumWeight
-hasAttribute Tracked

Referring to Figure 2 the IO has the three ancestors: ArmoredPersonnelCarrier, VehicleType and ResourceObject (Thing is not included), and the EO has the ancestors VehicleType and ResourceObject. In (1) F_R and F_S correspond to the ancestors of IO and EO respectively, and thus $r_A = 2/2 = 1$ and $p_A = 2/3$. Further, as there is only one common property, namely MediumWeight, $r_P = 1/2$ and $p_P = 1/3$. We combine the recall and precision according to (4) to $r = 0.68$ and $p = 0.45$, and then finally $q = 0.54$.

IV. Test Run

We will now illustrate the association process on the scenario of Section 2. The set of EOs is initially empty. For each IO the algorithm determines its classification according to the ontology, determines its consistency with respect to each EO, and computes the similarity values. If the ambiguity is too large, the decision is left to the operator.

We have not handled the information on the number of vehicles in a report; e.g. a report on a tank is regarded as equivalent to a report on a group of tanks. This is actually part of a much bigger issue, namely partial reports on units of arbitrary kind, and this is a topic for future studies.

First the UAV provides two IOs which are added to the ontology as new EOs with the same attributes.

ID	Type	Attributes
UAV1	VehicleType	hasWeapon: MachineGun
UAV2	VehicleType	hasWeapon: Cannon hasWeapon: MachineGun

ID	Type	Attributes
EO1	VehicleType	hasWeapon: MachineGun
EO2	VehicleType	hasWeapon: Cannon hasWeapon: MachineGun

Second, there is a report from the Ground Sensor Network (from now on we simplify the tables):

GSN1	VehicleType	LightWeight Wheeled

Following the ontology, this report is classified as an automobile. It is consistent with both EOs[2], so we compute the similarities. They turn out to be too low in both cases, so a third EO is added.

ID	p_A	r_A	p_P	r_P	q
EO2	0.67	1	0	0	0.28
EO1	0.67	1	0	0	0.28

EO3	Automobile	LightWeight Wheeled

Next report is from Reconnaissance Troop.

REC	ArmoredPersonalCarrier	MediumWeight MachineGun LightArmor

It is inconsistent with EO3 because of weights. The similarity values are

ID	p_A	r_A	p_P	r_P	q
EO1	0.67	1	0.33	1	0.62
EO2	0.67	0.67	0.33	0.33	0.45

This is deemed too ambiguous for automatic association. An operator decides that the report should go with EO1. Attributes are merged.

EO1	ArmoredPersonalCarrier	MachineGun LightArmor MediumWeight

Next report is only compatible with EO2.

[2] One might think that an automobile cannot carry weapons, and that therefore this report would be inconsistent with UAV1 and UAV2. However, nowhere in our ontology has it been stated that an automobile cannot carry weapons.

GSN2	VehicleType	HeavyWeight Tracked

However, the similarity q is only 0.35. It may be unwise to associate directly so we make an EO4, which later can be merged with EO2 when there is better evidence.

Finally,

TCR	Volvo-XC90	LightWeight Wheeled

is inconsistent with EO1 and EO4, because of weights. Similarities present strong evidence that this is EO3, and the association is carried out.

ID	p_A	r_A	p_P	r_P	q
EO3	0.75	1	1	1	0.95
EO2	0.5	1	0	0	0.23

Now we have the following EOs:

EO1	ArmoredPersonalCarrier	MachineGun LightArmor MediumWeight
EO2	VehicleType	Cannon MachineGun
EO3	Volvo XC-90	LightWeight Wheeled
EO4	VehicleType	HeavyWeight Tracked

where it is suspected that EO4 is identical to EO2 (truth is, it is).

V. Conclusion

In this paper we have shown a method for assisting the operator in associating intelligence reports with existing objects, where everything is expressed in an ontology, in this case using OWL. It is meant to be combined with association based on positional data, where however the latter can be so sparse that it is more or less useless. The method discussed here has the nice property that it is based entirely on the information contained in the classes and individuals, and there is no need to assign probabilities, beliefs, confidences or other similar measures, which could be quite difficult to do in a reliable way. Also the actual processing of such measures in connection with complex ontologies is a field open for research.

There are thus several challenges that we have not handled in this paper: The combination of ontological association with positional association; the uncertainties in the reports; the uncertainty in the association; the appropriate HMI for operators' support; and target hierarchies in the form of units or other aggregates. We still believe that, being entirely realistic for implementation, this approach could form a useful step in processing target reports in the ground domain.

VI. Acknowledgment

The authors would like to thank Mr. Johan Edlund for highly valuable discussions in the early phase of the project.

References

[1] J. Lagerlöf, "Situation assessment in the Ground Target Domain," Master's Thesis in Computing Science, University of Umeå, Sweden, 2008, available at http://www.cs.umu.se/education/examina/Rapporter/JennyLagerlof.pdf
[2] Multilateral Inteoperability Programme, http://www.mip-site.org
[3] World Wide Web Consortium, OWL information at http://www.w3.org/2004/OWL/
[4] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, P. F. Patel-Schneider (Eds.), *The Description Logic Handbook*, Cambridge University Press, Cambridge, UK, 2003
[5] M. M. Kokar, C. J. Matheus, J. A. Lethowski, K. Baclawski, and P. Kogut, "Association in Level 2 Fusion," *Proceedings of the SPIE Conference on Multisensor, Multisource Information Fusion*, Orlando, FL, April 2004.
[6] Clark & Parsia LLC, Pellet information at http://clarkparsia.com/pellet
[7] C. Bizer, R. Heese, M. Mochol, R, Oldakowski, R. Tolksdorf, R. Eckstein, "The Impact of Semantic Web Technologies on Job Recruitment Processes", *Proceedings of the 7th International Conference Wirtschaftsinformatik*, Bamberg, Germany, February 2005.
[8] D. Skoutas, A. Simitsis, and T. Sellis, "A Ranking Mechanism for Semantic Web Service Discovery," *Proceedings of the 4th International Workshop on Semantic Web for Services and Processes (SWSP ´07)*, in conjunction with the 2007 IEEE International Conference on Web Services (ICWS ´07), Salt Lake City, Utah, USA, July 9-13, 2007