Moss, L., Sleeman, D. and Sim, M. (2010) *Reasoning by analogy in the generation of domain acceptable ontology refinements.* Lecture Notes in Computer Science, 6317 . pp. 534-543. ISSN 0302-9743

Reasoning by Analogy in the Generation of Domain Acceptable Ontology Refinements

Laura Moss[1,2,3], Derek Sleeman[1,2], and Malcolm Sim[2]

[1] Department of Computing Science, University of Aberdeen, Aberdeen, AB24 3FX
[2] University Section of Anaesthesia, Pain, & Critical Care Medicine, Glasgow Royal Infirmary, Glasgow, G4 0SF.
[3] Department of Clinical Physics, University of Glasgow, Glasgow, G12 8QQ
Email: lauramoss@nhs.net

Abstract. Refinements generated for a knowledge base often involve the learning of new knowledge to be added to or replace existing parts of a knowledge base. However, the justifiability of the refinement in the context of the domain (domain acceptability) is often overlooked. The work reported in this paper describes an approach to the generation of domain acceptable refinements for incomplete and incorrect ontology individuals through reasoning by analogy using existing domain knowledge. To illustrate this approach, individuals for refinement are identified during the application of a knowledge-based system, EIRA; when EIRA fails in its task, areas of its domain ontology are identified as requiring refinement. Refinements are subsequently generated by identifying and reasoning with similar individuals from the domain ontology. To evaluate this approach EIRA has been applied to the Intensive Care Unit (ICU) domain. An evaluation (by a domain expert) of the refinements generated by EIRA has indicated that this approach successfully produces domain acceptable refinements.

Key words: Ontology Refinement, Analogical Reasoning, Medicine

1 Introduction & Related Work

Bundy [3] suggests that ontologies (like any defined knowledge base) evolve over time; a model can only capture a finite description of the world, decisions made in the modelling of the domain can be *"overturned by experimental evidence or changes in specification"*. Previous approaches to (semi-automatic) ontology construction (also considered as ontology learning [16]) generally rely on a large text corpus to automatically extract concepts and relationships, for example, Text2Onto [5]. The validation and verification of ontologies is also a widely covered research topic (for a summary see [14]). The competency of an ontology is often evaluated using several distinct approaches: the first by testing the ontology against a list of competency questions which it should be able to answer, and the second by checking aspects of structural consistency using tools such as ODEClean [8]. For an ontology, incompleteness and incorrectness can exist at both the structural level (TBox) and the instance (individual) level (ABox). The field of ontology evolution can provide some support for ontology refinement (e.g. [12]) and attempts have recently been made at automating the process [15]. However, the majority of previous work (ontology learning, validation, and evolution) has focused on the refinement (or creation) of the taxonomic structure of an ontology, and removing inconsistencies in ontologies (e.g. [9],[10]); the validity of values associated with individuals in an ontology has been largely neglected.

The approach presented in this paper generates refinements for *individuals* in an ontology (rather than the TBox) and differs from previous time consuming and knowledge intensive approaches by using *existing* domain knowledge contained in analogous individuals. Gentner [7] describes reasoning by analogy as *"a kind of reasoning that applies between specific exemplars or cases, in which what is known about one exemplar is used to infer new information about another exemplar"*. The use of reasoning by analogy in ontology engineering is mainly confined to reuse processes such as mapping, and merging of ontologies, and has largely not been explored in ontology refinement; one notable exception is the LEARNER system [4] which generates refinements for incomplete ontologies. The approach described in this work extends previous work by generating refinements (automatically) for *both* incomplete and incorrect ontology individuals.

The rest of this paper is organised as follows; section 2 discusses the approach implemented in EIRA to generate ontology refinements and its application to the Intensive Care Unit (ICU) domain; section 3 details an evaluation of the refinements generated; section 4 discusses conclusions and future work.

2 Reasoning by Analogy to Generate Refinements

2.1 EIRA

EIRA, an existing knowledge-based system has previously been applied in the ICU domain to produce explanations for anomalous patient responses to treatment[4]. Figure 1 provides an overview of EIRA (the following numbers in brackets correspond to numbers in the figure). EIRA's explanation generation process starts when an ICU clinician enters an anomaly into EIRA (1). EIRA can also detect (if possible) additional anomalies at the same time as the clinician detected anomaly (2). The explanations produced by EIRA are generated by the application of strategies with (medical) domain knowledge represented in several ontologies (3)(4). A number of explanations for the anomaly are then presented to the ICU clinician (5). To identify the strategies (algorithms) implemented in EIRA, interviews were held with ICU clinicians during which they were asked to provide explanations for a number of pre-identified anomalous patient responses to treatment. These interviews and explanations were subsequently analysed and high level strategies used by the clinicians were extracted and implemented in EIRA.

[4] For further details see [11]

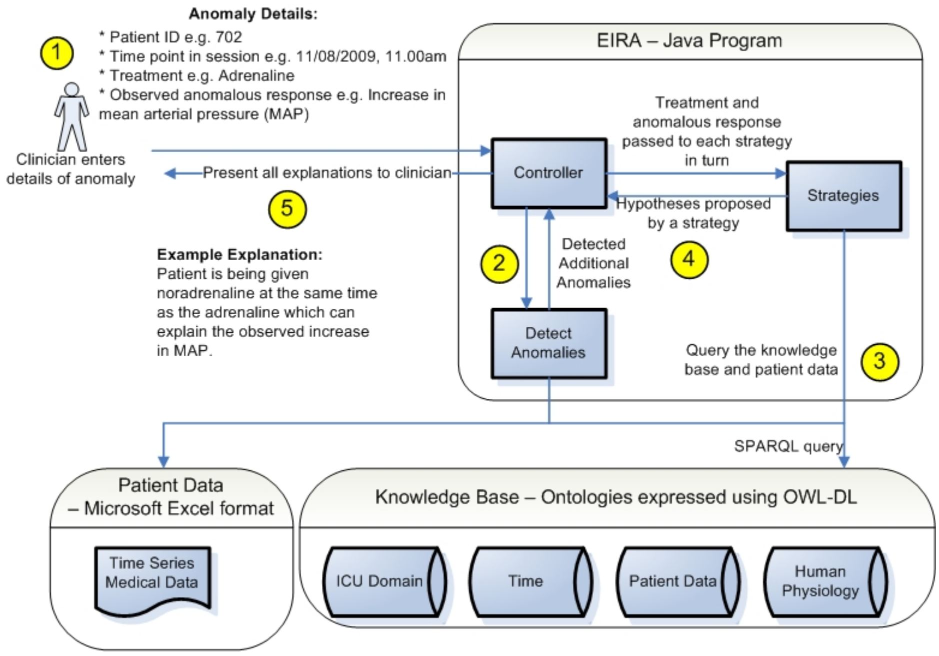

Fig. 1. Overview of EIRA

2.2 Generating Ontology Refinements

EIRA has been recently extended to reason by analogy to generate ontology refinements. For each explanation produced by EIRA, the clinician using the tool can respond by dismissing it, fully agreeing with it, or specifying that the explanation requires further (clinical) investigation. If the clinician dismisses the explanation, EIRA analyses the reasoning processes that generated the explanation[5]. An incorrect explanation indicates that either the strategies used by EIRA to produce the explanation were inappropriate or parts of EIRA's knowledge base are incorrect. The work described focuses on the refinement of EIRA's knowledge base; approaches to refine incorrect or insufficient rules (as strategies can be considered to be rules) have previously been explored in the theory revision literature (e.g.[6]). Of course, the knowledge base in EIRA may not only contain incorrect knowledge, it may also be incomplete, which prevents explanations from being generated. Incomplete knowledge may be identified when a query of the knowledge base fails to return any results. To refine the knowledge base (for both incorrect and incomplete individuals) it is proposed that knowledge about concepts and individuals which are *similar* to the erroneous concept(s) is used to suggest refinements. The following definition of similarity

[5] It is assumed that the ICU clinician provides a gold standard to commence the *generation* of a refinement, however, a consensus between multiple clinicians will be required before the refinement is actually *implemented* in the knowledge base

has been applied: two individuals can be considered similar if they are both instances of the same class (or immediate super-class).

The process of generating a refinement starts when either EIRA has detected missing information in the knowledge base or a clinician has stated that an explanation is incorrect. In both cases the SPARQL [2] queries used by EIRA are examined to identify the areas of the knowledge base potentially requiring refinement[6]. For each SPARQL query, the following high-level steps are followed:

1. Determine the type (class) of the individual containing a missing or incorrect property value(s) (the subject of the SPARQL triple).
2. Examine other individuals of the same (or similar) type.
3. Identify from these, individuals which have at least one value for the property (the predicate) used in the SPARQL triple.
4. Suggest refinements to the knowledge base which involve the property values of the other individuals.

To allow the domain expert/s to consider the impact of a proposed refinement to the ontology, EIRA identifies (using the Explanation Ontology, described later) previously correct explanations generated from the parts of the domain ontology for which a refinement has now been generated. A domain expert is asked to consider the impact of the refinement which may result in the following responses: 1) decline the current refinement as the associated change to a previous explanation is not acceptable, 2) accept the refinement because it does not change previous explanations, and 3) re-analyse a (previously correct) explanation as it is now considered to be incorrect in light of the proposed refinement.

Generating Refinements for Incomplete Individuals To identify incomplete parts of EIRA's knowledge base, the results from SPARQL queries used in EIRA are examined; when a SPARQL query fails to return any results, the parts of the ontology queried require further investigation.

A SPARQL query can fail to return any results when either the information in the ontology does not match the components of the query, or information is missing from the knowledge base. When a SPARQL query fails, it is examined further by dividing the query into (triple) patterns, investigating each pattern in turn. The following types of triple patterns are used in EIRA's algorithms (the examples are taken from the ICU domain)[7]:

– **A** - RDF-Term, RDF-Term, Variable, e.g.
  <http://www.owl-ontologies.com/unnamed.owl#Hypoadrenal Crisis>
  <http://www.owl-ontologies.com/unnamed.owl#additionalSymptoms>
  ?additionalSymptom.

[6] SPARQL is a W3C recommended query language for RDF (Resource Description Framework [1].), and is used to query ontologies.

[7] It is acknowledged that other types of triple patterns (e.g. Variable, Variable, Variable) could theoretically occur in a SPARQL query, however, they are not currently used in the implementation of EIRA

- **B** - Variable, RDF-Term, RDF-Term, e.g.
  ?symptom
  <http://www.owl-ontologies.com/unnamed.owl#clinicalFeatures>
  <http://www.owl-ontologies.com/unnamed.owl#DecreaseHeartRate>.

- **C** - Variable, RDF-Term, Variable, e.g.
  ?condition,
  <http://www.owl-ontologies.com/unnamed.owl#additionalSymptoms>
  ?symptom.

Triple pattern of type A is the simplest to investigate. In the example given, if the triple pattern did not return any results, it can be determined that 'Hypoadrenal Crisis' has no associated additional symptoms in the ontology. In this example, additional symptoms for 'Hypoadrenal Crisis', can be suggested by using individuals which are similar to 'Hypoadrenal Crisis'. To enable this the class (or super-class) of the subject in the triple (T) being examined is determined and other individuals of this class are identified. For each individual identified, if an object (property value) is associated via the same predicate (property) as in T, then the object is suggested as a refinement for the original triple. The algorithm only examines the class and super-class of the individual as moving further up the class hierarchy reduces the likelihood that suggestions acceptable to a domain expert will be found. Following the above example, the ontology is queried to determine the type (class) of 'Hypoadrenal Crisis', which in this case is 'Adrenal Disorder'. Other individuals with type 'Adrenal Disorder' are then retrieved from the ontology (e.g. 'Phaeocromocytoma'), if none are found then the super-class of the type ('Adrenal Disorder') is examined (e.g. 'Metabolic Disorder'). When another, similar, individual is identified it is examined to determine if the 'additionalSymptoms' property relates it to another individual. For example, it can be determined that 'Phaeocromocytoma' is associated with 'HighHeartRate' and 'Anxiety' via the 'additionalSymptoms' property. A refinement for the 'Hypoadrenal Crisis' class may then be suggested in which 'HighHeartRate' and 'Anxiety' are associated with 'Hypoadrenal Crisis' via the 'additionalSymptoms' property.

For triple pattern types B and C, if no results are returned by the query these triple patterns are not investigated further unless the triple patterns exist as part of a sequence of triple patterns. For example, if triple pattern B in the example given does not return any results, then it can be inferred that no individuals exist in the ontology with a value for the 'clinicalFeatures' property of 'DecreaseHeartRate' and hence it is not an appropriate scenario for refinement. Similarly, if triples of pattern type C fail, then it is not possible to make a suggestion as no other individuals exist in the ontology to compare against. The only situation where it may be possible to make suggestions for values using type B or C triples is when they occur as part of a sequence of triple patterns within the same query (i.e. the SPARQL query examines several parts of the knowledge base), in these scenarios information can be used from the other triple patterns in the query. For example, the following SPARQL query contains two triple patterns (the first of type B and the second of type A):

SPARQL query: SELECT ?symptom WHERE ?symptom
<http://www.owl-ontologies.com/unnamed.owl#clinicalFeatures>
<http://www.owl-ontologies.com/unnamed.owl#IncreaseHeartRate>.
<http://www.owl-ontologies.com/unnamed.owl#Phaeocromocytoma>
<http://www.owl-ontologies.com/unnamed.owl#additionalSymptoms> ?symptom

The first triple pattern (TP1) would not be examined as its subject is an unknown concept (i.e. a variable, '?symptom'). The algorithm would then proceed to examine the second triple pattern (TP2); if it does return results and further, if the variables in both triple patterns have the same variable name (i.e. '?symptom'), then TP2 can be used to enable a further examination of TP1. This is achieved by substituting '?symptom' in TP1 with each value for '?symptom' from the results of TP2. For example, once it has been determined that 'Phaeocromocytoma' has the 'additionalSymptom', 'HighHeartRate', it is possible to examine TP1 by replacing '?symptom' with 'HighHeartRate'.

Generating Refinements for Incorrect Individuals For each explanation generated by EIRA, which the domain expert classifies as incorrect, the associated SPARQL queries are identified from an Explanation Ontology. The Explanation Ontology models (domain-independently) the explanations generated by EIRA. Figure 2 provides a visualisation of the Explanation Ontology.

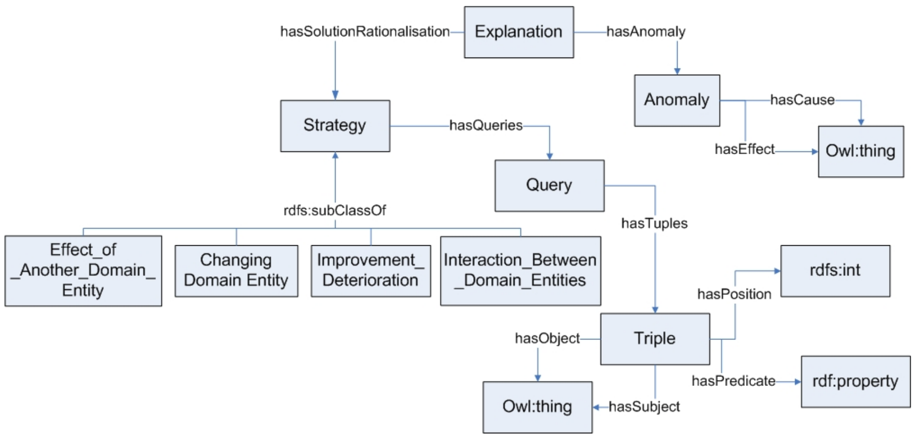

Fig. 2. Explanation Ontology Schema

The identified SPARQL queries are then segmented into triple patterns and examined further (as discussed in the previous section). The following types of refinements can be suggested by EIRA: replace the property value, remove the property value associated with an individual, and change the property associating the individual and property value. EIRA systematically determines if refinements can be generated for each type of refinement. For example, if the ontology (incorrectly) contains the ‘expectedEffect’ property relating the individual, ‘Verapamil’, with the property value, ‘IncreaseHR’ (i.e. verapamil is expected to increase a patient’s heart rate), then the following refinements may be suggested by EIRA:

1. **Replace property value**, e.g. replace ‘IncreaseHR’ with another value, for example, ‘DecreaseHR’.
2. **Remove the property value associated with an individual**, e.g. remove the property value, ‘IncreaseHR’, associated with ‘Verapamil’ via the ‘expectedEffect’ property.
3. **Change the property associating the individual (subject) and property value (object)**, e.g. relate ‘Verapamil’ and ‘IncreaseHR’ via another property, for example, the ‘conditionalDrugEffect’ property

The first type of refinement involves the examination of similar individuals to suggest a replacement value (object) which the property (predicate) should relate the (triple’s) subject to. To enable this, EIRA determines if the class (or super-class) of the subject in the triple (T) contains other individuals. For each individual (I) identified, the value (V) related to I by the same property as T is noted. If V is associated with at least 25%[8] of the individuals examined then V is recommended as a refinement for T. Following the above example, if ‘Verapamil’ (the subject) and ‘Propranolol’ are individuals of the same super-class and it is observed that the property ‘expectedEffect’ relates ‘Propranolol’ to ‘DecreaseHR’, then the refinement suggested is to set ‘DecreaseHR’ as the value of the ‘expectedEffect’ property for ‘Verapamil’; that is, expectedEffect(Verapamil, IncreaseHR) becomes, expectedEffect(Verapamil, DecreaseHR).

The second type of refinement removes the erroneous property value, and so the subject is no longer related to the object via the property. To achieve this EIRA examines each triple (T) in the SPARQL query and retrieves other individuals of the same class (or super class if no instances exist) as the subject in T. If greater than 50% of the individuals do not have the same property as T then it is suggested that the property associated with T is removed.

The third type of refinement replaces the property which associates the subject and object. To suggest a replacement property, individuals of the same class (or super-class) as the subject in the triple (T) are examined by EIRA. For each individual retrieved, the property which associates the individual with the object in T are noted (if the association occurs). If an identified property occurs in more than 50% of individuals examined then the property is suggested as a refinement for T. Following the previous example, if ‘Verapamil’ (the subject) and ‘Propranolol’ are individuals of the same super-class and it is observed that ‘Propranolol’ is related to the ‘IncreaseHR’ individual via the ‘conditionalDrugEffect’ property, then the algorithm suggests that the ‘expectedEffect’ property relating ‘Verapamil’ and ‘IncreaseHR’ is effectively replaced with the ‘conditionalDrugEffect’ property. In this case, the refinement of the knowledge

[8] The threshold levels applied in the three types of refinements have been agreed with an ICU clinician as sensible levels for this domain.

base would involve: expectedEffect(Verapamil, IncreaseHR) being replaced by conditionalDrugEffect(Verapamil, IncreaseHR).

3 Evaluation

To evaluate whether the use of reasoning by analogy in EIRA (to suggest refinements to incomplete and incorrect ontology individuals) produced acceptable results, refinements generated by EIRA were presented to an ICU clinician for evaluation. To generate the refinements, test cases were created, each consisting of a medical treatment and anomalous response (as identified by an ICU clinician [11]), this information was entered into EIRA and on each occasion EIRA failed to produce an explanation, or produced an incorrect explanation, the subsequent ontology refinements generated by EIRA were noted.

3.1 Evaluating Refinements Generated for Incomplete Individuals

The test cases did not query all parts of the knowledge base (as the queries are based on the anomalies entered) and hence refinements were not generated for all the properties in the domain ontology. To produce a complete set of refinements, examples of missing property values (not previously identified by the test cases) were manually identified from the knowledge base and details entered into EIRA.

EIRA produced (manually and automatically as described above) 46 refinements which were presented to the clinician. The clinician was asked to state whether each refinement was clinically acceptable. For 7 out of the 46 possible refinements a mistake had been made and refinements had been generated for properties which should contain property values unique to the individual (e.g the name of a drug) and these were removed[9]. A further 2 refinements were not commented on by the clinician as he felt he lacked the required clinical knowledge.

For the remaining 37 refinements which the clinician did comment on, a total of 23 refinements (62.2%) were accepted by the ICU clinician and 14 refinements (37.8%) were classified as not acceptable; for these 14 refinements, the clinician stated that the property values suggested were not acceptable. In addition, for 7 cases, the clinician did not agree that the two compared individuals were similar (the majority of these cases had generated refinements judged as unacceptable). For these later cases it is suggested that the taxonomy (TBox) of the ICU domain ontology requires further refinement.

3.2 Evaluating Refinements Generated for Incorrect Individuals

Incorrect explanations generated by EIRA (as identified previously by an ICU clinician in [11]) were selected from the Explanation Ontology and refinements

[9] These refinements are not considered as incorrect in the evaluation as the clinician agreed that it is completely meaningless to generate refinements for these properties. EIRA has subsequently been updated.

subsequently generated. A total of 72 explanations (instances) were contained in the Explanation Ontology. 11 of these explanations had previously been evaluated by an ICU clinician as incorrect (the remaining 61 were evaluated as correct or required further clinical investigation). For 7 of the 11 incorrect explanations, the ICU clinician had previously suggested why the explanation was incorrect (i.e. suggested a refinement to the knowledge base) in interviews described in [11].

The evaluation of the refinements generated for an incorrect knowledge base consisted of two stages: the first identifies if the refinements made by the ICU clinician (for 7 of the incorrect explanations) could be reproduced by EIRA, and the second determines if refinements generated by EIRA for the remaining 4 'unexplained' incorrect explanations were acceptable to a domain expert. In the first stage of the evaluation, EIRA reproduced 4 out of the 7 refinements suggested by the ICU clinician; the refinements EIRA produced for the remaining 3 explanations (which did not match the previous clinician's refinements) were subsequently evaluated by an ICU clinician. In the second stage of the evaluation, refinements were generated by EIRA for the remaining 4 (out of 11) incorrect explanations for which the ICU clinician *did not* previously suggest any refinements. A total of 19 refinements were generated by EIRA and viewed by an ICU clinician. For two (out of the 19 refinements) generated by EIRA, the clinician again felt he lacked the required clinical knowledge to evaluate the refinements. Out of the remaining 17 refinements, the clinician agreed that 10 out of the 17 (58.8%) refinements were acceptable. If the refinements from both sets of evaluation are considered, the overall acceptance of the refinements generated by EIRA for an incorrect knowledge base is 61.9%.

4 Conclusions & Future Work

The results indicate that (ABox) ontology refinements, acceptable to a domain expert for both an incomplete and incorrect ontology, can be generated by reasoning by analogy. Further, this approach has the following advantages: *existing domain knowledge* contained in ontology individuals is used, thereby avoiding the requirement for additional domain datasets (which are often unavailable) and/or time consuming to acquire; this work indicates that *domain acceptable* refinements can be generated when the TBox of an ontology is organised from the same perspective as the properties for which the property values are being transferred; and finally the use of SPARQL queries and domain expert feedback allows for relatively easy *identification* of ABox elements requiring refinement.

Plans for future work include: the implementation of more complex definitions of similarity in EIRA as the ICU clinician disagreed at several points that the identified individuals were similar, Ricklefs et al [13] provide a comparison of such ontology similarity metrics; an extended evaluation, firstly by asking further ICU clinicians to evaluate EIRA's refinements, both individually and as a group to form a consensus evaluation, and secondly, an evaluation to determine whether the refinements accepted by the domain expert improve the number of satisfactory explanations generated by EIRA; finally, an exploration of the use of reasoning by analogy in EIRA to generate *new* domain knowledge.

Acknowledgments. Kathryn Henderson and Jennifer McCallum (CareVue Project) & the staff and patients of the ICU Unit, Glasgow Royal Infirmary. This work was an extension of the routine audit process in Glasgow Royal Infirmary's ICU; requirements for further Ethical Committee Approval have been waved. This work was supported under the EPSRC's grant number GR/N15764.

References

1. RDF - Accessed October 2009. http://www.w3.org/RDF/.
2. SPARQL - Accessed October 2009. http://www.w3.org/TR/rdf-sparql-query/.
3. Bundy, A. Why Ontology Evolution Is Essential in Modelling Scientific Discovery. In *Proceedings of the 2008 AAAI Fall Symposium on Automated Scientific Discovery*, 2008.
4. Chklovski, T. Learner: A System for Acquiring Commonsense Knowledge by Analogy. In *Proceedings of 2nd International Conference on Knowledge Capture*, 2003.
5. Cimiano, P., Völker, J. Text2Onto. In *Natural Language Processing and Information Systems*, pages 227–238, 2005.
6. Craw, S., Sleeman, D. Automating the Refinement of Knowledge-Based Systems. In *Proceedings of the Ninth European Conference on Artificial Intelligence*, 1990.
7. Gentner, D. Analogical Reasoning, Psychology of. In *Encyclopedia of Cognitive Science*. Nature Publishing Group, 2002.
8. Guarino, N., Welty, C.A. An overview of OntoClean. In *Handbook on Ontologies in Information Systems*, pages 151–172, 2004.
9. Horridge, M., Parsia, B., Sattler, U. Laconic and Precise Justifications in OWL. In *The Semantic Web - ISWC 2008*, 2008.
10. Lam, J., Pan, J., Sleeman, D., Vasconcelos, W. A Fine-Grained Approach to Resolving Unsatisfiable Ontologies. In *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI-2006)*, 2006.
11. Moss, L., Sleeman, D., Sim, M., Booth, M., Daniel, M., Donaldson, L., Gilhooly, C., Hughes, M., Kinsella, J. Ontology-Driven Hypothesis Generation to Explain Anomalous Patient Responses to Treatment. *Knowledge-Based Systems*, 23:309–315, 2010.
12. Noy, N., Chugh, W., Liu, M., Musen, M. A Framework for Ontology Evolution in Collaborative Environments. In *Proceedings of the Fifth International Conference on the Semantic Web (ISWC-2006)*, 2006.
13. Ricklefs, M., Blomqvist, E. Ontology-Based Relevance Assessment: An Evaluation of Different Semantic Similarity Measures. In *On the Move to Meaningful Internet Systems: OTM 2008*, pages 1235–1252, 2008.
14. Sure, Y., Gómez-pérez, A., Daelemans, W., Reinberger, M.L., Guarino, N., Noy, N.F. Why Evaluate Ontology Technologies? Because They Work! *IEEE Intelligent Systems*, 19:74–81, 2004.
15. Zablith, F., Sabou M., d'Aquin, M., Motta, E. Using Background Knowledge for Ontology Evolution. In *Proceedings of International Workshop on Ontology Dynamics (IWOD) at ISWC*, 2008.
16. Zhou, L. Ontology Learning: State of the Art and Open Issues. *Information Technology and Management*, 8:241–252, 2007.