



HAL
open science

Ontology-based Semantic Relatedness for Detecting the Relevance of Learning Resources

Amel Yessad, Catherine Faron, Rose Dieng-Kuntz, Mohamed Tayeb Laskri

► To cite this version:

Amel Yessad, Catherine Faron, Rose Dieng-Kuntz, Mohamed Tayeb Laskri. Ontology-based Semantic Relatedness for Detecting the Relevance of Learning Resources. *Interactive Learning Environments*, 2011, 19 (1), pp.63-80. 10.1080/10494820.2011.528888 . hal-01174614

HAL Id: hal-01174614

<https://hal.science/hal-01174614>

Submitted on 27 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ontology-based Semantic Relatedness for Detecting the Relevance of Learning Resources

A. Yessad¹, C. Faron-Zucker², R. Dieng-Kuntz¹ & M.T. Laskri³

¹*Edelweiss, INRIA Sophia Antipolis, 2004 route des Lucioles, BP 93, 06902 Sophia Antipolis, France {Amel.Yessad, Rose.Dieng}@inria.fr;* ²*Kewi, I3S, Université de Nice Sophia, CNRS, CNRS, 930 route des Colles, BP 145, 06930 Sophia Antipolis cedex, France, faron@polytech.unice.fr;* ³*GRIA-LRI, Université Badji Mokhtar, BP 12, 23000 Annaba, Algérie, laskri@univ-annaba.orgAnnaba,*

Ontology-based Semantic Relatedness for Detecting the Relevance of Learning Resources

Semantic relatedness measures have proved to be useful for a number of applications, such as querying personalized web resources, word sense disambiguation or real-word spelling error correction. Most semantic relatedness measures between concepts are based on concept hierarchy of the domain ontology. In this paper we propose a semantic relevance measure that expresses the semantic relatedness between a learning resource and the learning context of a learner. In our case, both the learning resource and the learning context are described by graphs using the learning concepts of the domain model. Our semantic relevance measure aims to detect the relevance of a learning resource for a particular learning context of a learner. In our work, the semantic relevance measure is based on the assignment of relative weights to the learning concepts describing the learning resource according to their relationships with the current concept interesting the learner. The proposed measure achieves better results than similar measures and yet is much simpler than most of them. It is shown to achieve a correlation from 0.627 to 0.945 with expert ratings. This measure is implemented and used in a learning organizer, a system which generates adaptive hypermedia courses and reuses learning resources from distant web repositories, called « Organisateur de Parcours Adaptatifs de Formation » (OrPAF) (Yessad, Faron-Zucker, Dieng-Kuntz, & Laskri, 2008a, 2008b). In OrPAF, learning resources are annotated in order to be queried for a particular learning context which is represented by a map of annotated learning concepts, called Adaptive Cognitive Map (ACM). The proposed semantic relevance measure is used in order to automatically detect the learning resource relevance.

Keywords: Semantic web; Educational hypermedia; Learning resource; Semantic annotation; Semantic relatedness;

Introduction

The Web is becoming a de-facto standard platform for providing various kinds of learning resources, to support teaching in a university or a technical training company. One of the greatest benefits of the Web is that learning resources created to support a specific course no longer remains the only resources that the learners can use during the course. Many researches propose personalized learning supports in order to reuse and share learning resources from distributed repositories (Nejdl, Wolf, Qu, Decker, Sintek, Naeve, Nilsson, Palmer, & Risch, 2002; Miklos, Neumann, Zdun, & Sintek, 2003; Dolog, Henze, Nejdl, & Sintek, 2004). The personalization of existing learning resources can be a solution to the problem of developing online courses. However, a personalized learning that uses distributed metadata of learning resources is still an unsolved problem in the e-Learning research area.

Considering the increasing re-use of learning resources from the Web it becomes almost impossible for learners, experts and formation supervisor to get an overview of all the available information relevant to their current needs, tasks, roles and goals. And even if they find some materials which seem suitable, they are not able to assess completely whether the found content is entirely appropriate for their goals, their current knowledge and cognitive state. For that reason, learning resources searched from web repositories must be first subject to a pedagogy engineering work in order to make them reusable in the context of a particular formation for particular learners. This engineering work is time and effort consuming in the design step of an e-learning system.

In this area, researches attempt to specify interoperable metadata standards to capture and formalize the semantic meaning of learning resources in order to use, adapt and share these resources in different e-learning context. The IEEE/IMS LOM specification constitutes the best example of these standards. The IEEE/IMS LOM specification describes nine major categories of metadata which are used to describe a learning resource and where the implementation of metadata is language independent. XML is the most common language binding used presently, but other language bindings are becoming available such as the RDF-LOM binding [Nilsson, Palmer, & Brase, 2003]. Despite the positive steps taken with the development of metadata standards, they suffer of some complaints. The most important complaint is that value spaces of LOM metadata are not necessary shared and common to all the resource repository systems. This approach offers very little or no semantic metadata such as the actual metadata standards that are meant to help reuse in fact do the opposite [Mohan & Greer, 2003]. There is a growing concern though towards the need of extending the IEEE/IMS LOM standard, in the context of the Semantic Web so as to allow improved semantic annotation of learning resources. Many researches have been associating ontology with various aspects of learning resource in order to improve the quality of metadata and encourage reuse. The Semantic Web for E-Learning (SW-EL) field has shown the greatest activity in this trend with several interesting and recurring practices (Aroyo & Dicheva, 2004; Yessad & Laskri, 2006).

We propose a learning organizer which generates adaptive hypermedia courses and reuses learning resources from distant web repositories, called « Organisateur de Parcours Adaptatifs de Formation » (OrPAF) (Yessad et al., 2008a, 2008b). The learning organizer aims to improve the learning process efficiency (1) by providing the learner with adaptive course structure according to his/her level of knowledge, learning goal and time constraints; and (2) by reusing learning resources of different web repositories as adaptive course content:

On the one hand, the adaptive course structure is a map of relevant learning concepts. We call this map Adaptive Cognitive Map (ACM) which is a sub-graph of the graph of learning concepts which describes the domain model (e.g. Algebra domain, the algorithmic and programming languages domain) and adapted to the learner model (e.g. beginner).

On the other hand, the adaptive content is made of learning resources that are queried from distributed web repositories (e.g. ARIADNE) and recommended to the learner. These resources are already annotated with the LOM metadata but are still very difficult to be automatically reused because of the semantic lack of the LOM

metadata. In order to improve the automatic reuse of learning web resources, we propose to enrich their LOM metadata by conceptual annotations as an additional layer of description. For instance, a conceptual annotation concerns resource topics, resource author or resource role related to the learning resource. The annotation task is semi-automatic, assisted by domain experts. Once the annotation task is performed, the problem is to automatically detect the semantic relevance of a learning resource for the learning context of a learner. In this paper, we propose a novel approach based on the learning context to discover the semantic relevance of annotated learning resources. The semantic relevance measure of a learning resource is based on relative weights of the learning concepts related to the learning resource. The semantic relevance measure expresses the relatedness between the learning resource and the learning context which is represented by the ACM and the current concept interesting the learner. Whereas the most relatedness measures are based on the matching between terms, the relevance measure is based on the semantic representation of the domain and the matching between concepts. The originality of our approach is that the semantic relevance measure considers the global learning context and not only the current concept interesting the learner.

This paper is organized as follows: section 2 presents how we represent the learning knowledge by learning ontology and learning models and we show how we annotate a learning resource by using these learning ontology and learning models. We explain, in section 3, the proposed approach to detect the semantic relevance of a learning resource. Section 4 presents a related work and section 5 presents the evaluation approaches and results of the proposed semantic relevance measure.

Learning Knowledge

Because of the increasing complexity and heterogeneity of knowledge in e-learning systems (e.g. domain knowledge, learner knowledge, pedagogical knowledge), we require an efficient and modular knowledge organization. We represent our learning organizer knowledge in two levels: the meta-model level and the model level.

Learning Ontology

The learning ontology we developed stands for the meta-model and the backbone of the learning process. It describes classes and properties that are instantiated in order to specify the domain of interest (e.g. Algebra), profiles of the learners (e.g. beginner), pedagogical strategies (e.g. deductive strategy) and annotations of learning resources (e.g. learning resource role).

Classes and properties are seen as general objects. So, the learning ontology (meta-model) is instantiated to construct three learning models: a domain model, a learner model and a pedagogical model. Contrary to the learning ontology, these models describe specific objects. For instance, in the learning ontology, we describe types of learning concepts of a domain (e.g. a medium concept) and relationships between these types (e.g. prerequisite

relationship) whereas in the domain model we describe concrete learning concepts and their relationships. For instance, the concept *EquivalenceClass* and the concept *EquivalenceRelation* are instances of the class *LearningConcept* defined in the meta-model (cf. figure 1); and in the domain model, the concept *EquivalenceClass* is related to the concept *EquivalenceRelation* by the relationship *hasPrerequisite* (cf. figure 2). The classes of the learning ontology are hierarchically organized with one general and abstract class called *Class*. The class *Class* has three direct subclasses: *LearningConcept*, *LearningPerson* and *LearningPedagogy* which are used to construct respectively the domain model, the learner model and the pedagogical model. For example, all learning concepts of the domain model are instances of subclasses of the class *LearningConcept*. Similarly, properties of the learning ontology are hierarchically organized with one

general and abstract class called *property*. The class *property* is specialized in three subclasses: *learningConceptProperty*, *learnerProperty* and *pedagogicalProperty*.

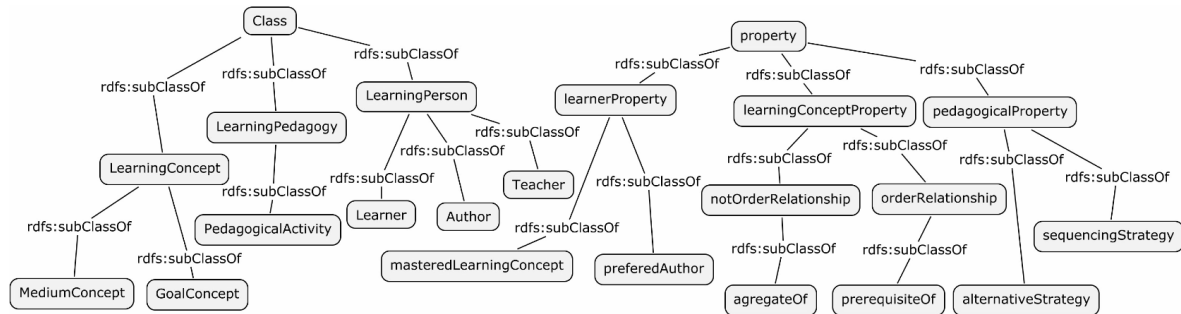
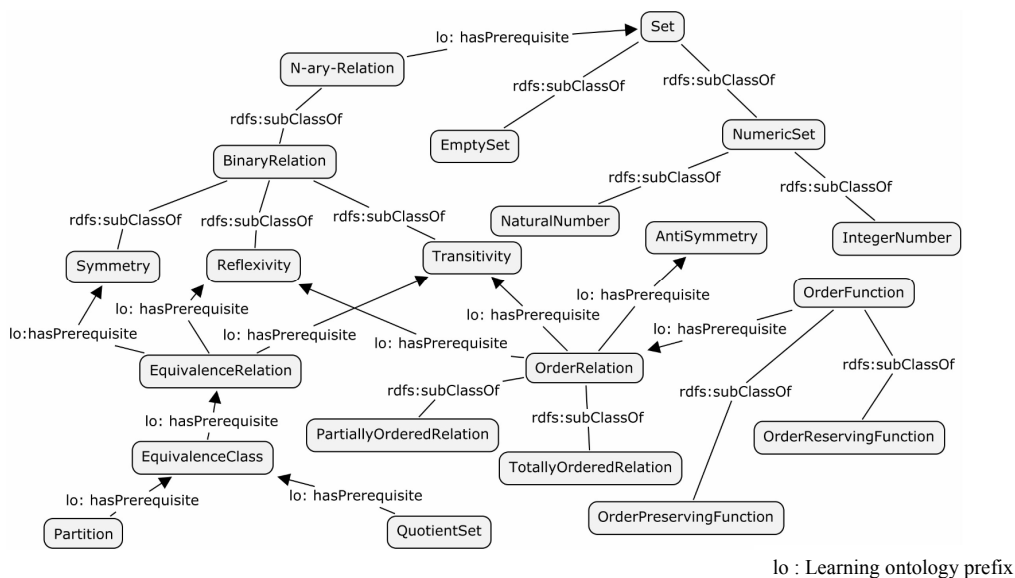


Figure 1. An excerpt of our learning ontology

Learning Models

The learning ontology is instantiated into three learning models: the domain model, the learner model and the pedagogical model.

Domain model. The domain model represents the domain of interest where the learner evolves. A specific domain of interest (e.g. algebra, algorithmic and programming languages) is described by learning concepts and their relationships. In figure 2, we show a fragment of the domain knowledge covering learning concepts of « Algebra » domain, including the *subClassOf* and the *hasPrerequisite* relationships between learning concepts. Learning concepts related with an order relationship (e.g. *hasPrerequisite*) must be taught following a certain order.



lo : Learning ontology prefix

Figure 2. An excerpt of the domain model « algebra »

Learner model. The learner model captures knowledge and preferences of the learner. It represents what the system knows about the learner. In our learning organizer, we represent the knowledge of the learner by the

so-called overlay model (Galeev, Tararina, & Kolosov, 2004). For instance, as shown in figure 3, if the learning concept *EquivalenceRelation* is mastered by the learner Maria, the knowledge $\langle Maria, masteredLearningConcept, EquivalenceRelation \rangle$ occurs in the learner model, else the concept *EquivalenceRelation* is unknown by the learner Maria. The learner model changes during the learning process when the learner passes tests. In this way, our learning organizer provides mechanisms for self regulated learning (Pintrich & Schunk, 2002; Perry, Phillips, & Hutchinson, 2004).

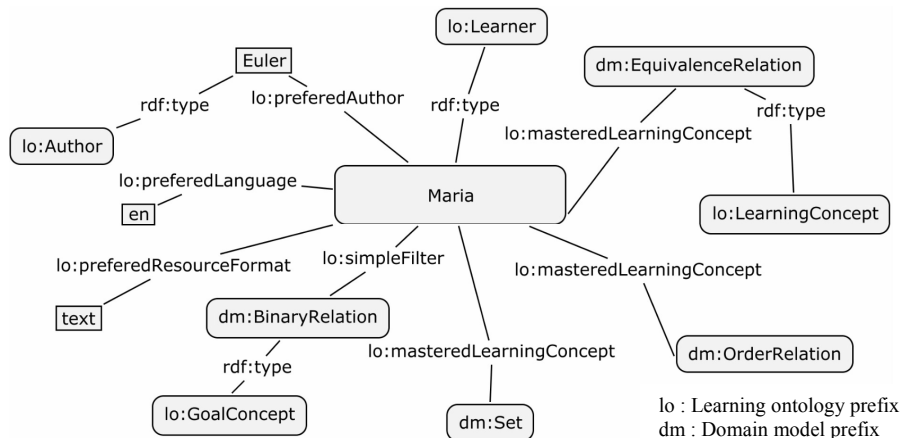


Figure 3 An excerpt of a learner model

Pedagogical model. The structure of the domain model alone is not sufficient to decide how to present the selected learning concept to the learner, i.e. what pedagogical type of learning resources to select, or how to sequence several learning resources to teach a given learning concept. For this purpose, we define a pedagogical model which describes pedagogical strategies to teach learning concepts. It describes different pedagogical activities (e.g. exercise, lecture) and their relationships. For instance, as shown in figure 4, the *Definition* activity (instance of the class *PedagogicalActivity* in the meta-model) must precede the *Exercise* activity; both activities are related by the *sequencingStrategy* property which is defined in the meta-model. The *alternativeStrategy* relationship between pedagogical activities means that learning resources related to these pedagogical activities can be accessed by a learner in any order whereas the *sequencingStrategy* relationship requires an order in the presentation of learning resources to the learner.

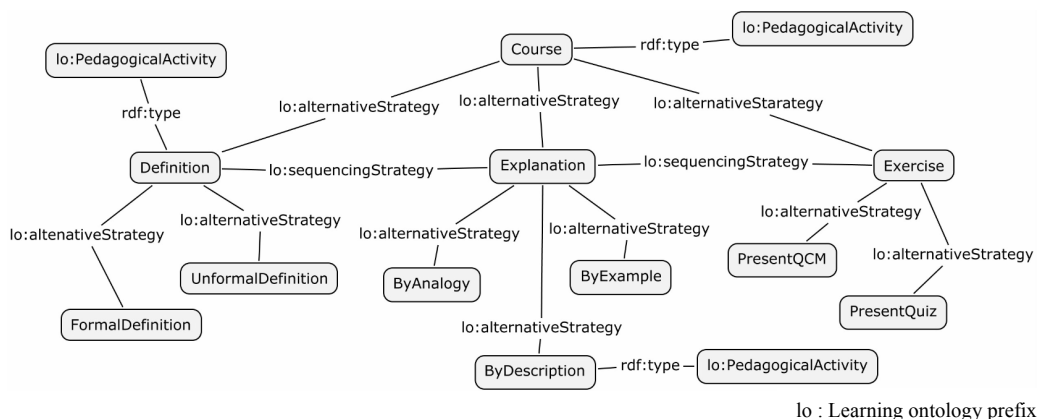


Figure 4. An excerpt of the pedagogical model

Adaptive Conceptual Map

An ACM is a course structure generated and displayed by the learning organizer in order to help the learner to construct a «correct» mental representation of the learning domain. It is a sub-graph of the domain model and contains learning concepts that the learner must learn to achieve his goal in required time. An ACM enables the learner to navigate between the learning concepts of the course. We distinguish between three cognitive maps constructed from the goal concept set G of the domain model (DM): simple $M_s(G)$, hierarchical $M_h(G)$ and relational $M_{r,m}(G)$. Each of these maps results from the application of a specific filter on the domain model and only the concepts which pass the filter are displayed to the learner.

- Simple $M_s(G)$ is the smallest map. It is composed of the goal concept and all concepts related to it directly or by transitive closure of order relationships.
- Hierarchical $M_h(G)$ is the cognitive map that extends the simple cognitive map to descendants and ascendants of the goal concept.
- Relational $M_{r,m}(G)$ is the cognitive map that extends the simple cognitive map to all concepts related to the goal concept by a path of relationships, the length of this path being less than m .

OrPAF implements each of these three possible filters. For the same concept goal, the filter depends on learner temporal constraints: the simple filter for learners with hard temporal constraints, the hierarchical filter for learners with medium temporal constraints and the extended filter for learners with flexible or no temporal constraints. These approaches can be compared to the micro, the meso and the macro learning approaches (Hug, 2005).

Before displaying the ACM to a learner, an additional adaptation layer is applied on it. It consists of applying rules in order to annotate each learning concept, similarly to the link annotation technique in adaptive Hypermedia systems (De Bra, Brusilovsky, & Houben, 1999). Different icons represent different states of a learning concept (cf. figure 5). A learning concept is accessible to the learner if all its prerequisites are mastered by the learner. Elsewhere, the learning concept is not accessible to the learner and no learning resources are attached to it. A concept without prerequisites is always accessible. In our organizer, we distinguish also between mastered concepts and not yet mastered concepts. Graphical icons are used to represent the difference between these three learning concept states. For instance, figure 5 presents a hierarchical ACM to learn the algorithmic and programming languages domain, the goal concept is *Procedure*, the learning concept *Operator* is mastered, the learning concept *Procedure* is not mastered and the learning concept *SemanticLanguage* is not accessible. The state of a learning concept changes from inaccessible state to accessible state if its prerequisites become mastered. Also, it changes from non mastered state to mastered state. These alterations in the ACM result from updating the learner model.

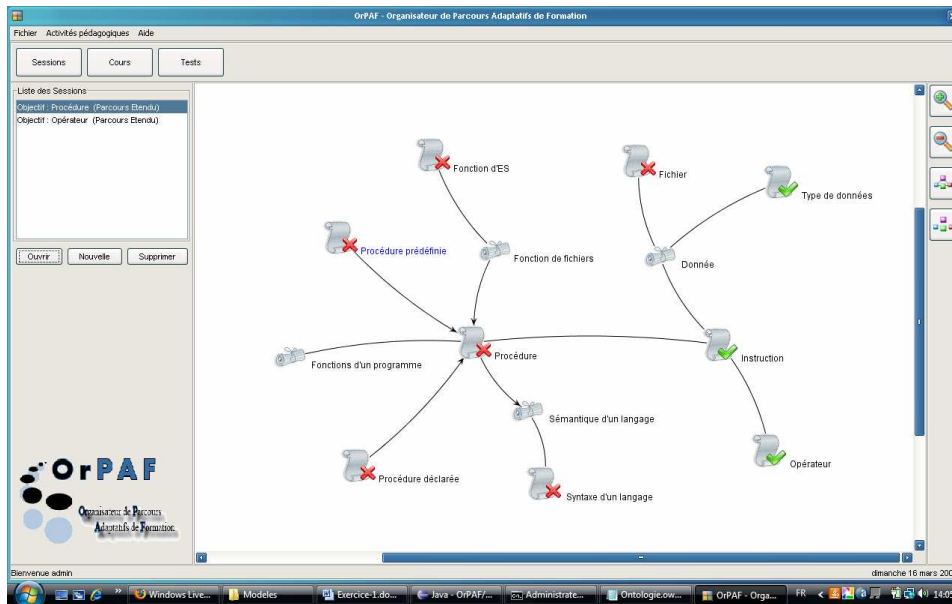


Figure 5. A screenshot of a course structure

Conceptual Annotation of Learning Resources

Once a distant learning resource is downloaded, it is submitted to a semi-automatic annotation process assisted by a teacher/expert; and finally the learning resource is stored in a local repository. Conceptual annotation of the learning resource are constructed by instantiating some classes and properties from both the domain model and the pedagogical model. For instance, as shown in figure 6, *Course1* is a learning resource. The role of *Course1* is *FormalDefinition* which is defined as *PedagogicalActivity* in the pedagogical model. The topics containing in the learning resource *Course1* are *reflexivity*, *Symmetry* and *Transitivity*, which are defined as *LearningConcept* in the domain model.

Characteristics (e.g. Resource topics, resource role, resource author, resource interactivity type) of the learning resource are manually identified by teacher/human expert and related to the learning resource in order to construct a conceptual annotation. This conceptual annotation is automatically generated by the system according to the learning models. The so-built conceptual annotation is then added to the learning resource metadata (in our case, the RDF-LOM binding metadata). This conceptual annotation is used during the learning process to detect the matching between the learning resource and the learning context of a particular learner.

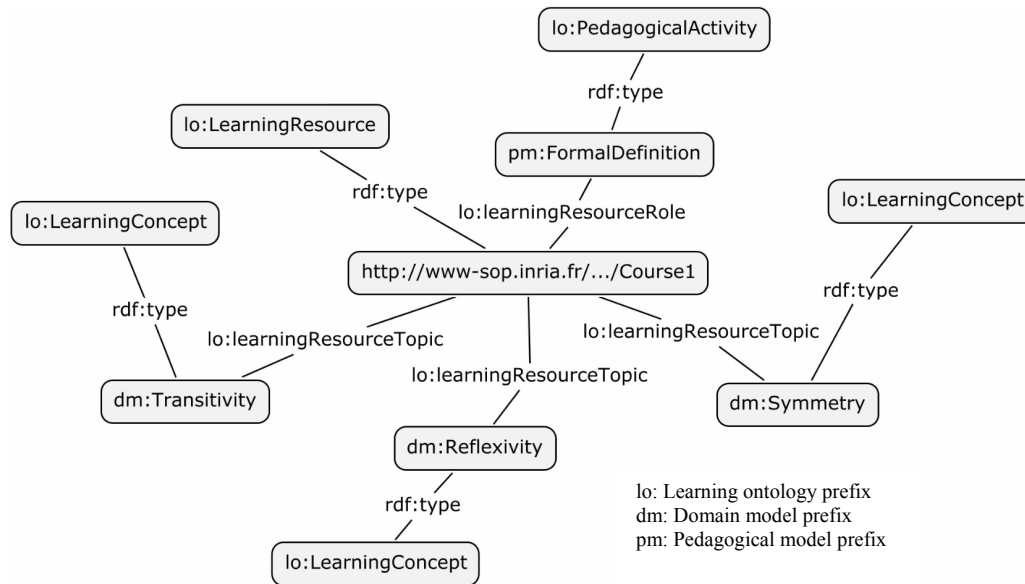


Figure 6. An excerpt of the conceptual annotation of the learning resource <http://www-sop.inria.fr/.../Course1>

Detection of the Semantic Relevance of Learning Resources

The problem of automatically recommending an adaptive content to a learner is still not resolved. We propose an approach which is based on previous semantic models and consists of searching relevant learning resources adapted to a particular learning context in order to recommend them to the learner. In OrPAF, a query component is developed in order to query learning resources (Yessad et al., 2008b). The learning ontology is described in OWL language and both the learning models and the resource annotation are described in RDF language. We extract knowledge from the learning ontology, the learning models and the resource annotation by submitting SPARQL query to Corese, an ontology-based search engine for the semantic web (Corby, Dieng-Kuntz, Faron-Zucker, & Gandon, 2006). In order to detect the resource relevance for a learner we extract the relationships that connect the concepts of the resource annotation, to the current concept in the ACM. Based on these relationships, we compute and assign a relative weight to each learning concept of the resource annotation. These weights are then used to compute the semantic relevance of the learning resource. Our approach consists of two phases: a semantic analysis and a customization analysis. The semantic analysis is based on the concepts of the ACM, their states and their relationships. In the semantic analysis, only the learning concepts of the resource conceptual annotation are used in the calculus of the semantic relevance. The other information related to the learning resource (e.g. learning resource role, learning resource author) are used in the customization analysis.

The result of the semantic analysis is learning resources which are relevant for the learning context but not necessarily adaptive for the learner preferences. Thus, we require a customization analysis for recommending only the learning resources adapted to the preferences of the learner.

Semantic Analysis

The learner can consult resources associated to one accessible learning concept in his/her ACM by a simple click on it. The query component searches for relevant learning resources from the local repository. It computes the relevance of a learning resource by matching its conceptual annotation with the learning

context: a given ACM and the learning concept currently selected in it. The relevant resources are recommended in decreasing order of relevance.

In order to compute the semantic relevance of a learning resource to the learning context, the learning concepts related to the learning resource by the *learningResourceTopic* property (cf. figure 6) are extracted from the resource annotation. For each extracted learning concept a relative weight is calculated. The weight of a concept depends on the path of the relationships that connect this concept to the current concept. The semantic relevance is then computed by using these relative weights.

Let $P \langle c_1, c_2, c_3, \dots, c_n \rangle$ be a learning path of length n and composed of concepts c_i . Let W_{c_i/c_1} ($i > 1$) be the weight of concept c_i relative to current concept c_1 in the ACM. Let W_{c_1} be the weight of the current concept c_1 . We define the relative weight by the following recursive relation:

$$\begin{aligned} W_{c_2/c_1} &= 1/a \\ W_{c_i/c_1} &= (1/a) W_{c_{i-1}/c_1}, i > 2 \end{aligned}$$

where a is a variable whose value is as follows:

$a=2$ if the relationship between c_{i-1} and c_i is *subClassOf* or its inverse relationship

$a=3$ if the relationship between c_{i-1} and c_i is *hasPrerequisite* or its inverse relationship

$a=5$ if the relationship between c_{i-1} and c_i is *aggregationOf* or its inverse relationship

The weight W_{c_1} of the current concept c_1 is a big number N .

When there are several relative weights for one learning concept (due to graph cycles) we consider the smallest value. Once the relative weight of each learning concept related to the learning resource by the *learningResourceTopic* property is computed, the semantic relevance SR of the learning resource can be measured as follows:

Let E be the set of concepts both present in the learning resource annotation and the accessible concepts of the ACM, let F be the set of concepts present in the learning resource annotation and not present in the accessible concepts of the ACM, and let c be the current concept of the ACM.

$$SR = \frac{\sum_{x \in E} w_{x/c}}{1 + \sum_{y \in F} w_{y/c}}$$

The definition of SR reflects the fact that the weight of a concept depends on the current concept and the state of the ACM and therefore on the learning context. A resource is relevant if the learning concepts related to it by the *learningResourceTopic* property have important relative weights and are largely similar to the accessible learning concepts of the ACM. Otherwise, the resource is less or not relevant.

For instance, as shown in figure 7, the weight of the learning concept *BinaryRelation* relatively to the current concept *OrderRelation* is $1/6$ because the path between these both concepts is composed of one *hasPrerequisite* relationship ($a=3$) and one *subClassOf* relationship ($a=2$): $W_{BinaryRelation/OrderRelation} = 1/2 * W_{Transitivity/OrderRelation}$ and $W_{Transitivity/OrderRelation} = 1/3$.

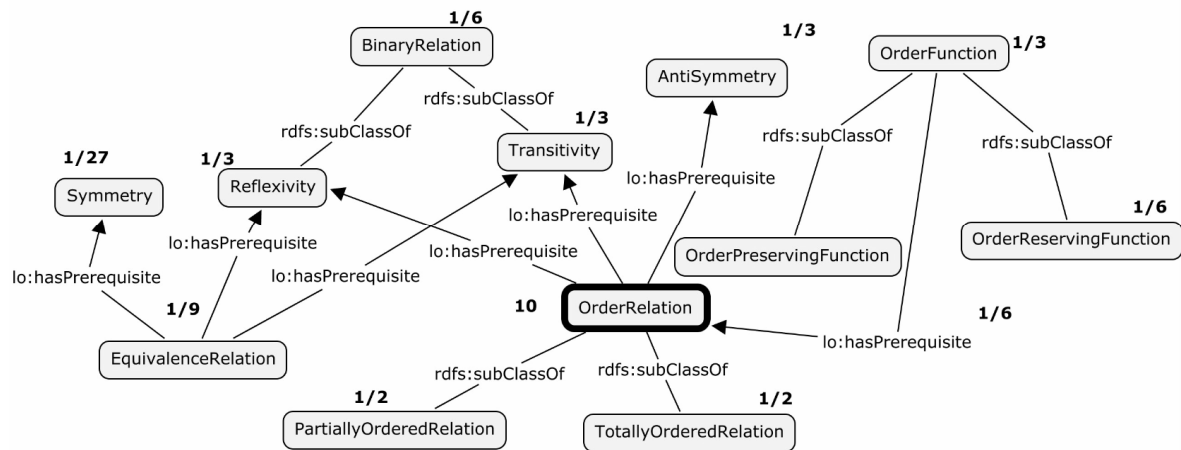


Figure 7. Assignment of relative weights to the concepts related to the learning concept *OrderRelation* in algebra domain

Customization Analysis

Each learning resource is described with a conceptual annotation. This conceptual annotation describes learning concepts related to the learning resource by the *learningResourceTopic* property (cf. figure 6.). In addition to the learning concepts, a conceptual annotation describes other LOM metadata which are used to decide if the resource matches with the learner preferences (e.g. *preferredAuthor*, *preferredResourceFormat*, etc.) or not. Once a set of candidate learning resources are selected as a result of the semantic analysis, we must verify that these resources match with the learner preferences. For instance, we verify that the learning resource has a particular interactivity type (e.g. video) or that the language used in the resource is the english language adapted to the learner preferences

Related Work

In our work, the proposed semantic relevance measure detects the degree of similarity between a learning resource and a learning context represented by an ACM and a current learning concept. It may be considered as the inverse of the semantic distance between the learning resource and the learning context. In the semantic analysis, the calculus of the semantic relevance concerns only learning concepts related to the learning resource. Other characteristics like the resource pedagogical role or the preferred resource format are used in the customization analysis. Thus, the computing of the semantic relevance is reduced to the computing of the semantic similarity between networks of learning concepts. The idea of computing the semantic proximity on semantic networks goes back up to researches of [Quillian, 1968; Collins & Loftus, 1975] on semantic human memory.

The structure of concepts supporting semantic similarity in many researchers is the hierarchy. In concept hierarchy, the subsumption relationship is used to compute semantic similarity between concepts. We can consider that there are two big families of approaches to compute the semantic similarity [Corby, Dieng-Kuntz, & Faron-Zucker, 2004; Rada, Mili, Bicknell, & Blettner, 1989; Wu & Palmer, 1994]. On the one hand, approaches that use only the subsumption relationship to compute the semantic similarity [Corby, Dieng-Kuntz, & Faron-Zucker, 2004; Rada, Mili, Bicknell, & Blettner, 1989; Wu & Palmer, 1994]. On the other hand, approaches which include besides the subsumption relationship other information. This information can be statistics about the use of concepts in corpus [Resnik, 1995; Jiang & Conrath, 1997]. Particularly, the

researches of [Jiang et al., 1997; Lin, 1998] introduce the notion of conditional probability of encountering an instance of a child-concept given an instance of a parent-concept. More recently, [Zhong, Zhu, Li, & Yu, 2002] define a similarity measure between conceptual graphs: a query graph and a resource graph. While most previous researches base the calculus of the semantic proximity on the hierarchical structure of concepts, our approach takes into account all the semantic relationships between the concepts (e.g. *hasPrerequisite* relationship) in the domain model. In addition, the semantic similarity proposed for our learning organizer depends on the learning context by using relative weights whereas the semantic similarity in the above researches depends of the depth of the concepts in the hierarchy which is a constant value and learning context-independent.

Evaluation of the Semantic Relevance Measure

In order to evaluate our semantic relevance measure, we adopt two approaches :

The first evaluation approach is the comparison between the SR results and the expert ratings. These expert annotators had been asked to judge the relevance of resources for different learning contexts in the algorithmic and programming languages domain. The average relevance scores of experts are correlated with the semantic relevance values generated by our measure. In this experimentation context, we also studied results provided by a measure of the semantic similarity between conceptual graphs SoG proposed in (Zhong et al., 2002). We aimed to compare results obtained by our semantic relevance measure SR, the semantic similarity SoG and expert ratings. The comparison between SR and SoG is possible for two reasons: Similarly to the SoG measure, the SR measure computes the matching between two conceptual graphs: an ACM which represents a query graph and a learning resource annotation which represents a resource graph. In addition, similarly to the entry concept used in the SoG measure, we consider the current concept as the entry of the ACM (the query graph) and the most important learning concept related to the learning resource as the entry of the learning resource annotation (the resource graph). However, contrary to the SoG measure, the SR measure doesn't depend on the depth of the learning concepts in the domain model hierarchy but depends on the learning context. In addition, the weights used in the SoG are specified by users within the UI or defined as default values whereas the weights used in the SR are automatically calculated and depend on the learning context.

The second evaluation approach consists to measure the performance of the semantic relevance measure SR in detecting relevant resources and avoiding to recommend irrelevant resources to the learner.

Comparison with Expert Ratings of Relevance

Data. We asked three experts to note on a scale from 0 to 10 the relevance of twenty learning resources for six different learning contexts in the algorithmic and programming languages domain.

First, we explained to experts our aim of developing OrPAF and asked them to annotate twenty learning resources from Algorithmic and programming languages domain. They annotated resources with learning concepts from the domain model. After this annotation task, we presented to the experts six different ACM (three simple ACMs and three relational ACMs). In each ACM, the current concept was explicitly showed. The experts gave notes on scale from 0 to 10 which represent the relevance of each resource for each learning context. Finally, we studied both the correlation between our semantic measure of relevance SR and the expert ratings and the correlation between the semantic measure of similarity SoG (Zhong et al., 2002) and the expert ratings.

Results and discussion

Table 1. Correlation coefficients between expert ratings of similarity and both SR and SoG measures

Similarity measures	Simple ACM			Relational ACM		
	ACM ₁	ACM ₂	ACM ₃	ACM ₄	ACM ₅	ACM ₆
SR measure	0.883	0.945	0.964	0.627	0.690	0.691
SoG measure	-0.124	0.008	-0.678	-0.495	0.003	-0.097

Table 1 shows correlation scores for both the SR measure and the SoG measure with expert ratings of relevance. The correlation coefficients of the SR measure vary from 0.627 to 0.964. The semantic relevance scores converge to expert ratings in simple ACMs. Given that the SR measure is the same formula for all ACMs, these scores may be explained that the increasing of the size of an ACM disorient the expert who must make more effort to determine the relevance of a learning resource for a learning context. The more the size of the ACM is small, the better the expert estimates the relevance of a learning resource.

While the correlation coefficients of the SR measure are acceptable and even good for small ACMs, the correlation coefficients of the SoG measure are bad and diverge with expert ratings. This can be explained by the fact that the SoG is more adapted and gives better results for linguistic ontologies with a large specter of semantic relationships between concepts.

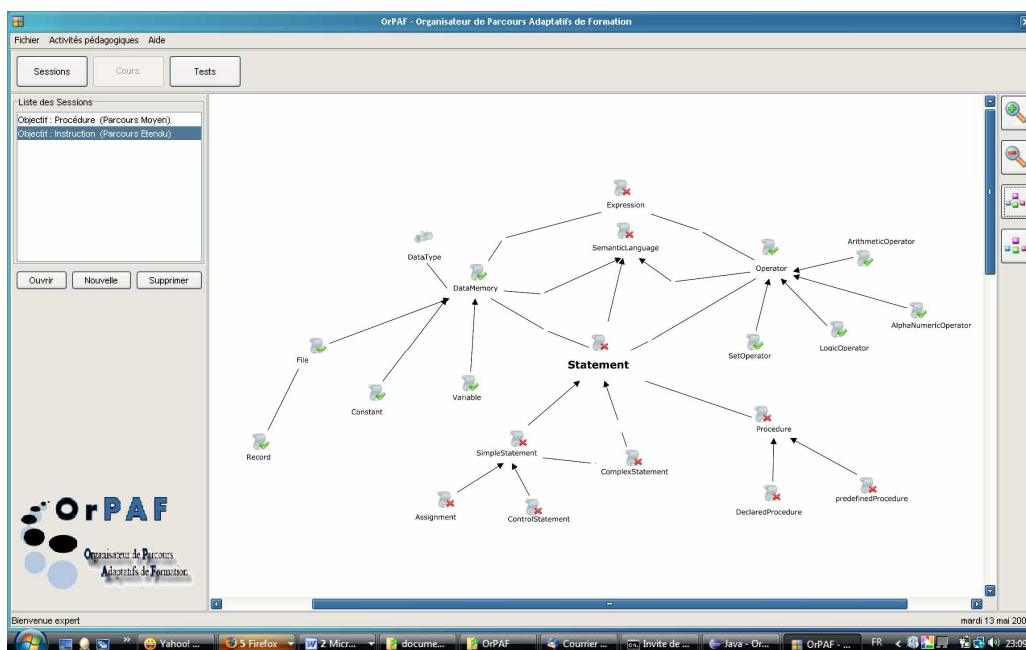


Figure 8. An annotated course structure or learning context

Performance of Semantic Relevance Measure

Data. We took a corpus of thirty six annotated learning resources relative to concepts in Algorithmic and programming language domain. Fifteen of these learning resources were judged relevant by experts for a learning context (cf. figure 8).

The detection of the relevance of the learning resources can be viewed as a retrieval task and evaluated in terms of precision (P), recall (R), and F-measure (F). In addition, we defined the scope of a query as the maximum authorized number of retrieved resources for any query. The scope of a query is a priori defined by the domain expert and so defines the size of the learning resource GUI container. We varied the scope of the query in order to study how the number of detecting relevant resources evolves with the number of retrieved learning resources. We define P, R and F as follows:

$$P = \frac{|\{\text{relevant resources}\} \cap \{\text{retrieved resources}\}|}{|\{\text{retrieved resources}\}|}$$

$$R = \frac{|\{\text{relevant resources}\} \cap \{\text{retrieved resources}\}|}{|\{\text{relevant resources}\}|}$$

$$F = \frac{2 * (P * R)}{(P + R)}$$

Results and discussion. In table 2, learning resources are sorted by decreasing order of SR. The rows in bold type represent the relevant learning resources (fifteen resources) whereas the rows in italic type are the irrelevant learning resources for the learning context of figure 8. We observe that most the relevant learning resources have the best relevance scores.

In figure 9, the values of precision range from 66.6% to 84.6% and stay stabilized even when we increase the scope of the query. These values express that the relevance of the learning resources retrieved are acceptable and the error varies from 15.4% to 33.4%. Another result is that the recall or the detection of relevant learning resources increases with the scope of the query and the ascent of irrelevant learning resources decreases with the increasing of the scope. The value thirteen of the scope appeared as a good value in order to obtain a good precision and a good recall. Results show that the proposed semantic measure SR is a reliable filter permeable to irrelevant resources.

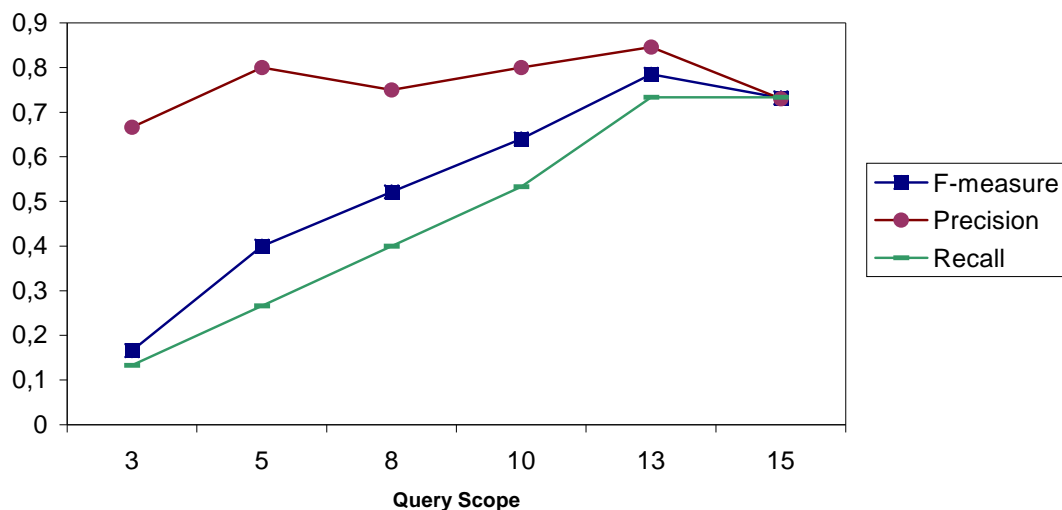
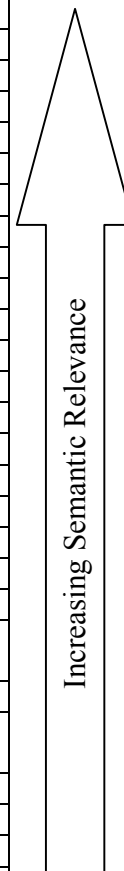


Figure 9. Relevance detection F-measure, Recall and Precision by query scope

Table 2. SR measure values of the learning resources for the learning context of figure 8

Learning resource	Conceptual annotation (learning concept list)	SR
Learning resource 7	Operator, arithmetic operator, Alphanumeric operator, logic operator	0.833
Learning resource 28	Data type, data memory	0.444
<i>Learning resource 15</i>	<i>Data, date type, data declaration,</i>	<i>0.410</i>
Learning resource 5	Variable, constant	0.333
Learning resource 32	Operator	0.333
Learning resource 34	Data type, set operator	0.277
Learning resource 23	Record, variable	0.222
<i>Learning resource 19</i>	<i>File, record, file function</i>	<i>0.205</i>
Learning resource 16	Logic operator	0.166
Learning resource 21	Set operator	0.166
Learning resource 24	Arithmetic operator	0.166
Learning resource 33	Variable	0.166
Learning resource 35	Constant	0.166
<i>Learning resource 11</i>	<i>Choice statement, logic operator</i>	<i>0.148</i>
<i>Learning resource 27</i>	<i>Control statement, logic operator</i>	<i>0.133</i>
Learning resource 9	Data type	0.111
Learning resource 4	Data type, array	0.108
<i>Learning resource 20</i>	<i>Data type, simple data, structured data, array</i>	<i>0.097</i>
Learning resource 31	Record	0.055
Learning resource 36	Record	0.055
<i>Learning resource 1</i>	<i>Algorithm</i>	<i>0</i>
<i>Learning resource 2</i>	<i>Machine code</i>	<i>0</i>
<i>Learning resource 3</i>	<i>Iterative statement</i>	<i>0</i>
<i>Learning resource 6</i>	<i>IO statement</i>	<i>0</i>
<i>Learning resource 8</i>	<i>Assignment statement</i>	<i>0</i>
<i>Learning resource 10</i>	<i>Functional algorithm</i>	<i>0</i>
<i>Learning resource 12</i>	<i>procedure, parameter, local parameter, global parameter</i>	<i>0</i>
<i>Learning resource 13</i>	<i>Algorithm, program, programming language</i>	<i>0</i>
<i>Learning resource 14</i>	<i>Structured program, non structured program, recursive program</i>	<i>0</i>
<i>Learning resource 17</i>	<i>Array sorting</i>	<i>0</i>
<i>Learning resource 18</i>	<i>Function, predefined function</i>	<i>0</i>
<i>Learning resource 22</i>	<i>Statement</i>	<i>0</i>
<i>Learning resource 25</i>	<i>Expression</i>	<i>0</i>
<i>Learning resource 26</i>	<i>Control statement, choice statement</i>	<i>0</i>
<i>Learning resource 29</i>	<i>Semantic language</i>	<i>0</i>
<i>Learning resource 30</i>	<i>Syntax language</i>	<i>0</i>



Conclusion

Reusing and sharing learning resources in adaptive manner is a challenge for reducing the time and the effort of E-Learning system development. The semantic web technologies like the ontologies, the semantic annotations and the LOM standard are used in an attempt to improve the learning resource metadata and so make the learning resource more meaningful for programs.

In this paper, we proposed an approach that consists in annotating a learning resource by a conceptual annotation which is constructed by instantiating some classes and properties from our models of knowledge. These conceptual annotations are an additional layer of description that extends the LOM metadata of a learning resource. We propose to detect the relevance of an annotated learning resource for a learner via two analyses: the semantic analysis and the customization analysis. In the semantic analysis, a relative weight is assigned to each learning concept related to the learning resource. These weights depend on the learning context, i.e. the ACM and the current concept interesting the learner and are used to compute a semantic relevance of the learning resource. The learning resources selected in the semantic analysis are submitted to the customization analysis. The customization analysis consists of verifying that these learning resources are adapted to the learner preferences.

The proposed semantic relevance measure was evaluated by two approaches. On the one hand, we calculated the correlation coefficients between the SR results and the expert ratings on a corpus of twenty learning resources. In addition, we compared these results with the correlation coefficient obtained between the SoG results and the expert rating on the same corpus. The SR results were significantly closer to the expert ratings than the SoG results.

On the other hand, we used information retrieval measures in order to evaluate the performance of the semantic relevance SR in detecting relevant resources. This approach gave very satisfactory results.

References

- Aroyo, L., & Dicheva, D. (2004). The new challenges for e-learning: The educational semantic web. *Educational Technology & Society*, 7 (4), 59-69.
- Collins, A., & Loftus, E. (1975). A Spreading Activation Theory of Semantic Processing. *Psychological Review*, (vol. 82, pp. 407-428).
- Corby, O., Dieng-Kuntz, R., & Faron-Zucker, C. (2004, August 22-27). Querying the semantic web with the Corese search engine. In *Proceedings of European Conference on Artificial Intelligence* (pp. 705-709). Valencia, Spain: IOS Press.
- Corby, O., Dieng-Kuntz, R., Faron-Zucker, C., & Gandon, F. (2006). Searching the semantic web: Approximate query processing based on ontologies. *IEEE Intelligent Systems Journal*, 21(1), 20-27.
- De Bra, P., Brusilovsky, P., & Houben, G.-J. (1999). Adaptive hypermedia: From systems to framework. *ACM Computing Surveys*, 31(4).
- Dolog, P., Henze, N., Nejd, W., & Sintek, M. (2004, May 17-22). Personalization in distributed e-learning environments. In *Proceedings of Thirteenth International World Wide Web Conference* (pp. 170-179). New York, USA.
- Galeev, I., Tararina, L., & Kolosov, O. (2004, August 30-September 1). Adaptation on the basis of the skills overlay model. In *Proceedings of IEEE International Conference on Advanced Learning Technologies* (pp. 648-650). Joensuu, Finland: IEEE Computer Society.
- Hug, T. (2005, May 6-8). Micro learning and narration. In *Proceedings of fourth Media in Transition Conference*, MIT, Cambridge (MA), USA.
- Jiang, J., Conrath, D. (1997). Semantic Similarity based on Corpus Statistics and Lexical Taxonomy. In *Proceedings of International Conference on Research in Computational Linguistics*. Taiwan.
- Lin, D. (1998, July 24-27). An information-theoretic definition of similarity. In *Proceedings of 15th International Conference on Machine Learning* (pp. 296-304). Madison, Wisconsin USA.
- Miklos, Z., Neumann, G., Zdun, U., & Sintek, M. (2003, October 20-23). Querying semantic web resources using TRIPLE views. In *Proceedings of the 2nd International Semantic Web Conference (ISWC)*, Sanibel Island, Florida, USA.

- Mohan, P., & Greer, J. (2003, June 23-28). Learning objects: Current status and future directions. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, Honolulu, Hawaii.
- Nejdl, W., Wolf, B., Qu, C., Decker, S., Sintek, M., Naeve, A., Nilsson, M., Palmer, M., & Risch, T. (2002, May 7-11). EDUTELLA: a P2P networking infrastructure based on RDF. In *Proceedings of 11th World Wide Web Conference*, Hawaii, USA.
- Nilsson, M., Palmer, M., & Brase, J. (2003, November 20-21). The LOM RDF binding-Principles and implementation. In *Proceedings of Annual Ariadne Conference*, Leuven, Belgium.
- Perry, N.E., Phillips, L., & Hutchinson, L.R. (2004). Preparing student teachers to support for self-regulated learning. *Elementary School Journal*, 106(3), 237-254.
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications* (2nd ed.). Upper Saddle River, NJ: Merrill, Prentice Hall.
- Quillian, M.R. (1968). Semantic Memory. *Semantic Information Processing*, M. Minsky, the MIT Press, Cambridge, Mass.
- Rada, R., Mili, H., Bicknell, E., & Blettner, M. (1989). Development and application of a metric on semantic nets. *IEEE Transaction on Systems, Man, and Cybernetics*, 19(1), 17-30.
- Resnik, P. (1995). Semantic similarity in a taxonomy: An information-based measure and its applications to problems of ambiguity in natural language. *Journal of Artificial Intelligence Research*, 11, 95-130.
- Wu, Z., & Palmer, M. (1994, June 27-30). Verb semantics and lexical selection. In *thirty second Annual Meeting of the Associations for Computational Linguistics* (pp. 133-138). New Mexico State University, Las Cruces, New Mexico, USA.
- Yessad, A., & Laskri, M.T. (2006, December 7-9). Ontology-driven dynamic course generation for web-based education. In *Proceedings of Maghrebian Conference on Information Technologies (MCSEAI)*, Agadir, Morocco.
- Yessad, A., Faron,- Dieng-Kuntz, R., Laskri, M.T (2008a, June 30-July 4). Ontology-Driven Adaptive Course Generation for Web-based Education. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, Vienna, Austria,
- Yessad, A., Faron-Zucker, C., Dieng-Kuntz, R., & Laskri, M.T. (2008b). Adaptive Learning Organizer for Web-based Education. To appear in *International Journal of Web-based Learning and Teaching Technologies*.
- Zhong, J., Zhu, H., Li, J., & Yu, Y. (2002, July 15-19). Conceptual graph matching for semantic search. In *Proceedings of 10th International Conference on Conceptual Structures, ICCS* (pp. 92-106). LNCS 2393, Springer Verlag, Borovets, Bulgaria.