

A fuzzy contextual approach towards intelligent educational content adaptation

Phivos Mylonas

*National Technical University of Athens
Image, Video and Multimedia Laboratory
Zographou Campus, PC 15773, Athens, Greece
fmylonas@image.ntua.gr*

Abstract

Fuzzy contextual information is utilized within this paper in order to assist the process of educational multimedia content adaptation. Towards this scope, we extend on previous work in order to address the problem of educational personalization at a semantic level. In different previous works we have developed algorithms to address computationally efficient handling of fuzzy relations, and theory to address knowledge representation, thematic categorization and user profiling. In this work we take two further steps: first, we integrate ontologies in the original knowledge modeling approach and then we explain how these diverse algorithms and methodologies can be combined in order to approach a greater goal, that of educational content personalization. Prior work on web-based educational content adaptation is now extended with Semantic Web functionalities and features. The basis of this work is the role of context, as it improves the quality of the produced results, by utilizing fuzzy algebra, fuzzy sets and relations, towards efficient manipulation of the utilized user profiles. The novelty of the presented work is the ontological representation of contextual knowledge and its utilization within the personalized e-learning process.

1 Introduction

One of the most interesting problems in the field of e-learning applications is the adaptation of content to the end-users. It is also a common fact that most modern e-learning frameworks are lacking with respect to educational content representation and user personalization. In our approach, automated representation and extraction of user profiles, to be used in an e-learning content offering system, forms the proposed solution to the problem. In this approach we present such a profile-based system, by which educational content is matched to its context, so that it can be adapted to its user's needs and capabilities. Current effort extends

previous work on profile knowledge representation. This framework has been designed to enable learners (i.e. teachers and students) to gracefully increase their Information and Communication Technologies (ICT) knowledge [9] and provide them with credible information and feedback.

Nowadays, with the continuous growing of multimodal educational content and its adaptation needs, learning and multimedia communities are both realizing that the amount of accessible, processable and searchable data currently available brings two more parameters in play; those of personalization and semantics. Several types of personalization exist when dealing with interactive services, applications and content delivery to the end-user [2], where different content can be generated for different individual users or even classes of users. With multimedia retrieval building upon and borrowing from text retrieval in many aspects, most current e-learning content adaptation problems may generally be classified under the generic umbrella of semantic multimedia personalized adaptation [12], [13].

In this process, ontologies [8] do come into play and represent a formal way to model knowledge and provide a solid basis upon which semantic knowledge representation and operations could be developed. On the other hand, their size can make their handling problematic, let alone their practical utilization. As a result, caution must be taken when they are introduced in the education process.

In this work we integrate, on the one hand, an ontological representation for knowledge together with various independent previous works and exploit context on the other, in order to provide a framework upon which a contextualized multimedia personalization can be based. We utilize fuzzy algebra and fuzzy relations, based on the fuzzy properties a knowledge representation scheme should have in order to be informative enough. The ultimate goal of this research effort may be summarized in setting the basic principles to establish an efficient e-learning framework, based on educational content adaptation and personalization.

The outline of this paper is as follows. In section 2 we integrate the proposed approach with modern ontological

modeling. Section 3 briefly explains how the above knowledge can be used in order to extract the user profile structures described in detail in [5]. Finally, section 4 lists our concluding remarks that support the proposed approach.

2 Ontological knowledge representation

Among all possible knowledge representation formalisms, ontologies [8] present a number of advantages. In the context of this work, ontologies are suitable for expressing multimedia content semantics in a formal machine-processable representation that allows manual or automatic analysis and further processing of the extracted semantic descriptions. As an ontology is a formal specification of a shared understanding of a domain, this formal specification is usually carried out using a subclass hierarchy with relationships among the classes, where one can define complex class descriptions. In general, we may decompose an ontology O into two parts, i.e.

1. the set S of all its concepts and
2. the set $SR(s_i, s_j) : S \times S \rightarrow \{0, 1\}$ of all semantic relations amongst any two given concepts s_i, s_j

More formally:

$$O = \{S, SR(s_i, s_j)\}, \quad i, j = 1 \dots n \quad (1)$$

The proposed knowledge model is based on a set of concepts and relations between them, which form the basic elements towards semantic interpretation of the present research effort. Building on ontology-based semantic structures and semantic metadata, a personalization system builds and utilizes an explicit awareness of information about the end user, either directly provided by the user, or implicitly evidenced along the history of his/her actions [10]. Although almost any type of relation may be included to construct such knowledge representation, the two categories commonly used are taxonomic (i.e. ordering) and compatibility (i.e. symmetric) relations. However, compatibility relations fail to assist in the determination of the context and the use of ordering relations is necessary for such tasks. Thus, a first main challenge is the meaningful exploitation of information contained in taxonomic relations for the task of context exploitation within semantic image segmentation and object labeling. In addition, for a knowledge model to be highly descriptive, it must contain a large number of distinct and diverse relations among concepts. A major side effect of this approach is the fact that available information will then be scattered among them, making each one of them inadequate to describe a context in a meaningful way. Consequently, relations need to be combined to provide a view of the knowledge that suffices for context definition and estimation. In this work we utilize ten

relation types, whose semantics are defined in the MPEG-7 standard [1] and are presented in the following Table 1.

Table 1. Fuzzy semantic relations utilized.

Name	Inverse	Symbol	Meaning	Example	
				x	y
Example	ExampleOf	$Ex(x, y)$	y is an example of x	teacher	Dr. Ruth
Specialization	Generalization	$Sp(x, y)$	y is a specialization in the meaning of x	student	first-grader
Part	PartOf	$P(x, y)$	y is a part of x	school	yard
Similar	Similar	$Sim(x, y)$	x is similar to y	car	automobile
Instrument	InstrumentOf	$Ins(x, y)$	y is an instrument of or is employed by x	cut	knife
inContext	inContextOf	$Ct(x, y)$	y is on context of x	school	blackboard
Location	LocationOf	$Loc(x, y)$	y is the location of x	school concert	stage
Substance	SubstanceOf	$Sub(x, y)$	y is a substance or constituent of x	martini	gin
Property	PropertyOf	$Pr(x, y)$	y is a property of x	banana	ripeness
Patient	PatientOf	$Pat(x, y)$	y is affected by or undergoes the action of x	give	book

A last point to consider when designing such a knowledge model is the fact that real-life data often differ from research data. Real-life information is in principal governed by uncertainty and fuzziness, thus its modeling is based on fuzzy relations. For the problem at hand, the above commonly encountered crisp relations can be modeled as fuzzy ordering relations and can be combined for the generation of a meaningful fuzzy taxonomic relation. Consequently, to tackle such complex types of relations we propose a “fuzzification” of the previous ontology definition, as follows. A domain-specific, “fuzzified” version of a concept ontology may be described by O_F (eq. 2), where S represents the set of all possible concepts it describes, $F(SR(s_i, s_j)) = sr(s_i, s_j) : S \times S \rightarrow [0, 1]$ denotes a fuzzy ontological relation amongst two concepts s_i, s_j and $SR(s_i, s_j)$ denotes the non-fuzzy semantic relation amongst the two concepts. All above MPEG-7 originated relations are represented as an RDF graph and constitute the contextual knowledge to be used during the next phase.

$$O_F = \{S, sr(s_i, s_j)\}, \quad i, j = 1 \dots n \quad (2)$$

In the aforementioned relations, fuzziness does not enlarge the crisp ontology size and has the following meaning: High values of $Sp(x, y)$, for instance, imply that the meaning of y approaches the meaning of x , in the sense that when an image is semantically related to y , then it is most probably related to x as well. On the other hand, as $Sp(x, y)$ decreases, the meaning of y becomes “narrower” than the meaning of x , in the sense that an images relation to y will not imply a relation to x as well with a high probability, or to a high degree. Summarizing, the value of $Sp(x, y)$ indicates the degree to which the stored knowledge shows that an occurrence of y in an image implies relation to x . Likewise, the degrees of the other two relations can also be interpreted as conditional probabilities or degrees of implied relevance. The latter imply that, for example, $x \neq y \implies Sp(x, y) < 1$ since, if $x \neq y$, then we cannot be sure that both x and y are related to a given content, without first examining the contents context; at this point it is important to remind the reader that x and y are not terms but concepts, which means

that $x \neq y$ indicates/ensures a difference in a conceptual level.

$P(x, y)$ means that y is a part of x , e.g. x could be a *boat* and y could be a *sail*. So it is expected that the role of $P(x, y)$ is the opposite of that of $Sp(x, y)$, i.e. $Sp(x, y)$ means that the meaning of x “includes” the meaning of y . Finally, $Pr(x, y)$ means that y is a property of x . Based on these roles and semantic interpretations of all semantic relations, it is easy to see that a meaningful combination of them, utilizing inverse functionality where it is semantically appropriate (i.e., where the meaning of one relation is semantically contradictory to the meaning of the rest on the same set of concepts), could be utilized for the sake of educational content adaptation.

The representation of our concept-centric contextual knowledge model follows the Resource Description Framework (RDF) standard proposed in the context of the Semantic Web [7]. RDF is the framework in which Semantic Web metadata statements are expressed and usually represented as graphs. The combination of relations used may thus be visualized as a graph, in which every node represents a concept and each edge between two nodes constitutes a contextual relation between the respective concepts. Additionally each edge has an associated degree of confidence, which represents the fuzziness within the context model. Representing the graph in RDF is a straight forward task, since RDF structure itself is based on a similar graph model (Figure 1).

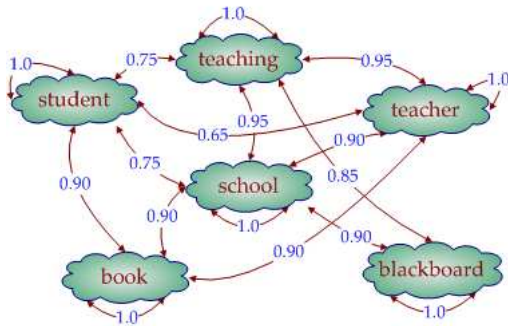


Figure 1. A small sample of the contextualized ontology; it represents relation *inContext* (*Ct*).

The reification technique [6] is used further, in order to achieve the desired expressiveness and obtain the enhanced functionality introduced by fuzziness. Representing the degree of confidence associated with each relation is carried out using reification, i.e. making a statement about the statement, which contains the degree information. Representing fuzziness with reified statements is a novel but acceptable way, since the reified statement should not be asserted auto-

matically. For instance, having a statement such as: “*Blackboard PartOf Classroom*” and a degree of confidence of “0.85” for this statement, does obviously not entail, that a *blackboard* is always a part of a *classroom*, as it may belong to a school laboratory or other facility. A small clarifying example is provided in the following Figure 2 for an instance of the specialization relation $Sp(x, y)$. In the example, the RDF subject *chalk* has *specializationOf* as an RDF predicate and *chemistrylab* forms the RDF object. Additionally, the proposed reification process introduces a statement about the former statement on the *specializationOf* resource, by stating that 0.65 is the membership degree to this relation.

3 Profiling

The user profile clustering approach described in detail in [5] forms the basic procedure, with the aid of which each learner is automatically categorized to a specific profile class that characterizes his behavior and his future interests and choices within the system. The general structure of the proposed algorithm combines traditional hierarchical clustering [3] principles with one a posteriori classification step. The latter is analytically presented in [4], whereas the clustering steps are summarized in the following:

1. Turn each input element into a singleton, i.e. into a cluster of a single element
2. For each pair of clusters c_1, c_2 calculate their Euclidean distance $d(c_1, c_2)$
3. Merge the pair of clusters that have the smallest Euclidean distance
4. Continue at step 2, until a meaningful termination criterion is satisfied, such as the thresholding of the value of the Euclidean distance

According to the produced cluster to which each user belongs, educational content, appropriately selected and associated a priori to each potential cluster by the systems experts, is offered to him. Because of flexibility and protection of crucial personal data reasons, the step of user characterization is only provided as an added value characteristic to the users that are willing to use it. Suitable verification procedures ensure that content offering filtering features are only enabled according to each end users will.

In order to extract user profiles from educational content with the highest possible confidence, the above agglomerative clustering approach may be used. Clearly, the definition of inter-cluster distances is the fine issue that determines the success of the approach. Our proposal is to use the intensity of the context, as the latter is defined by the ontological context approach presented in Section 2. This indicates

```

<rdf:Description rdf:about="#statement1">
<rdf:subject rdf:resource="#&dom;chalk"/>
<rdf:predicate rdf:resource="#&dom;specializationOf"/>
<rdf:object>rdf:resource="#&dom;chemistrylab"</rdf:object>
<rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement"/>
<context:specializationOfrdf:datatype="http://www.w3.org/2001/XMLSchema#float">
0.65</context:specializationOf>
</rdf:Description>

```

Figure 2. Fuzzy relation representation: RDF reification.

the degree to which the ontological knowledge indicates that the concepts contained in two clusters are indeed semantically related. The agglomerative clustering approach helps us overcome most of the problems faced in the process. For example, by ignoring clusters of low cardinality we remove concepts that are misleading and should not be considered in the estimation of user profiles. In this manner, our approach outperforms previous conventional personalization methodologies [12] that usually locate terms in queried content without taking into account its actual context. E-learning personalization frameworks also retrieve multimedia educational content and documents that aren't necessarily related to each other by semantic continuity. An efficient content browsing system like the one introduced briefly herein, based on semantic content analysis, user profiling, adaptation and context could provide users/learners with more effective educational content-browsing capabilities and increase their ICT knowledge.

Furthermore, its evaluation process extracts statistical analysis results that provide useful information about the end users' opinions regarding new technologies. The latter are then evaluated, in order to specify specific, underlying trends, needs and interests. The automatic profiling process, which takes place within the proposed approach, provides extremely useful and fully personalized information. According to this information, educational resources are linked to each profile category, whereas a group of domain experts is responsible to provide a set of multimedia enhanced e-courses and material to the end-users.

4 Conclusions and future work

This paper comes after a list of previous works on various theoretical and practical topics dealing with user profiles and educational content adaptation and before another similar list. Its purpose is to identify independent previous works that may be combined in order to come closer to the ultimate goal of e-learning personalization, indicate the way these works can be combined and identify the main directions for future and further research on context representation. Going further from our previous works, we have

focused on using an ontological modeling for contextual knowledge, thus allowing our approach to utilize knowledge bases developed by other researchers and making it easier for it to be integrated with other modern approaches. We have also provided a computationally efficient solution to the handling of the considered ontological relations and outlined the ways in which educational content adaptation and user modeling could be approached.

Utilization of fuzzy algebra and fuzzy sets and operations in the personalization process increases the flexibility and adaptability of the proposed approach and provides an innovative view on the entire research field. Bringing this one step further, introduction of contextual information optimizes personalization results and increases their level of accuracy and functionality, especially in comparison to non-contextual approaches. In other words, the described use of contextualized educational content adaptation ensures that the obtained personalized outcome will be of great interest to the real-life end-users of any e-learning content offering system based on top of this approach.

As a main field for further research and improvement we certainly identify the task of semiautomated generation of the required ontological knowledge. Where other approaches focus on crisp ontological relations our approach is based on a fuzzified version of this kind of knowledge. Current manual specification of these fuzzy degrees is both time consuming and subjective and as a result an automated or semi-automated methodology to perform this step is certainly welcomed. Other than that, much more is to be done in the field of experimental evaluation of similar methodologies, for all or us working in the fields of context and personalization, as ground truth is practically impossible to generate. The overall proposed contextual knowledge framework could be easily adapted to other e-learning schemes, mainly due to the robustness and clarity of its main entities.

References

- [1] A. B. Benitez, D. Zhong, S.-F. Chang, and J. R. Smith, "MPEG-7 MDS Content Description Tools and Appli-

- cations”, *Lecture Notes in Computer Science*, vol. 2124, pp. 41-52, Springer-Verlag, Berlin, 2001.
- [2] N. Correia, and M. Boavida, “Towards an Integrated Personalization Framework: A Taxonomy and Work Proposals”, *In Proceedings of WPTEPW*, Malaga, Spain, May 2002.
- [3] S. Miyamoto, *Fuzzy Sets in Information Retrieval and Cluster Analysis*, MA: Kluwer, Boston, 1990.
- [4] Ph. Mylonas, M. Wallace, and S. Kollias, “Using k-nearest neighbor and feature selection as an improvement to hierarchical clustering”, *In Proceedings of 3rd Hellenic Conference on Artificial Intelligence*, Samos, Greece, May 2004.
- [5] Ph. Mylonas, P. Tzouveli, and S. Kollias, “Intelligent content adaptation in the framework of an integrated e-learning system”, *In Proceedings of 16th ACM Conference on Hypertext & Hypermedia - Workshop on Combining Intelligent & Adaptive Hypermedia Methods/Techniques in Web-Based Education Systems*, Salzburg, Austria, September 2005.
- [6] RDF Reification, W3C Recommendation, February 2004, http://www.w3.org/TR/rdf-schema/#ch_reificationvocab
- [7] RDF, W3C Recommendation, January 2004, <http://www.w3.org/RDF/>
- [8] S. Staab, and R. Studer, *Handbook on ontologies, international handbooks on information systems*, Springer-Verlag, Heidelberg, 2004.
- [9] P. Tzouveli, N. Tsapatsoulis, and S. Kollias, “On the Development of User Adapted E-learning Schemes”, *In Proceedings of NHIE'03*, Santorini, Greece, 2003.
- [10] D. Vallet, P. Castells, M. Fernandez, Ph. Mylonas, and Y. Avrithis, “Personalized Content Retrieval in Context Using Ontological Knowledge”, *IEEE Transactions on Circuits and Systems for Video Technology*, Volume 17, Issue 3, pp. 336-346, March 2007.
- [11] N. Voisine, S. Dasiopoulou, V. Mezaris, E. Spyrou, T. Athanasiadis, I. Kompatsiaris, Y. Avrithis, and M. G. Strintzis, “Knowledge-assisted video analysis using a genetic algorithm”, *In Proceedings of 6th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2005)*, Montreux, Switzerland, April 2005.
- [12] M. Wallace, K. Karpouzis, G. Stamou, G. Moschovitis, S. Kollias, and C. Schizas, “The Electronic Road: Personalised Content Browsing”, *IEEE Multimedia*, 10(4), pp. 49-59, 2003.
- [13] M. Wallace, K. Karpouzis, M. Stefanou, I. Maglogianis, and S. Kollias, “Electronic Roads in Historical Documents: a Student Oriented Approach”, *Education and Information Technologies*, 9(3), pp. 271-289, 2004.