Contextualizing Multimedia Semantics Towards Personalised eLearning

Emmanuel Eze, Tanko Ishaya, Dawn Wood
Centre for Computing
The University of HULL, Scaborough Campus, UK
{ E.Eze, T.Ishaya, D.M.Wood}@dcs.hull.ac.uk

ABSTRACT: Recent developments of the Internet and the World Wide Web (WWW) have resulted in the proliferation of multi-media resources and learning with existing multimedia web resources on the Web is becoming more prevalent and important. While recent standardization efforts in eLearning such as LOM, SCORM, and IMS Learning Design work towards learning content description, packaging, and delivery; existing eLearning solutions still lack the ability to adequately use multimedia resources to provide a learner with personalised learning resources. Effective use of multimedia for web-based learning provides quality interactive learning experience, but current techniques do not adequately provide a semantic approach for organising multimedia resources. With evolving trend in learning through the use of web technology, eLearning systems are expected to provide personalised learning resources for effective learning. We have accordingly proposed a way of organising existing multimedia resources based on contexts towards providing a truly personalised eLearning experience. This paper presents a framework for achieving context-based multimedia semantic annotation and organisation towards personalised learning.

Categories and Subject Descriptors
H. 5.1[Multimedia Information Systems]: Audio Input/Output:
K.3 [Computers and Education]

General Terms
Multimedia Semantics, eLearning

Keywords: Web based learning, WWW, Multimedia technology

Received 5 March 2006; Revised and accepted 20 April 2006

1. Introduction

Developments in the Internet and the World Wide Web (WWW) technologies have led to an evolving trend in web-based learning. Electronic learning (eLearning) is now one of the most fast growing trends in computing and higher education (Gerhard and Mayr, 2002; Ishaya and Wood, 2005; Wood and Ishaya, 2005) and certainly becoming a dominant way of learning in workplace settings across other organizations as presented by Mungania (2003). Despite intensive developments in the area of eLearning technology and the wide variety of learning environments from many different vendors (e.g. WebCT and Blackboard), there is increasing evidence of dissatisfaction felt by both instructors and learners. One of the causes of this dissatisfaction is that these environments have not been designed to include the disposition of individual learners. There is evidence therefore, that the future growth of eLearning may well be constrained due to learner dissatisfaction. While recent standardization efforts in eLearning such as LOM, SCORM, and IMS work towards learning content description, packaging, and delivery; existing eLearning solutions still lack the ability to provide a learner with personalised learning resources.

Since we all learn and perceive things differently, it is therefore fundamental that learners are presented with learning content that are most interesting and appealing to them. The pedagogy of education has long established the importance of using various means of communication with learners (at different levels) for effective learning to take place. It is also a well known fact from commercial media research\(^1\) that different TV shows appeal to different demographics and sex. Research by Al-Khalifa (2005) investigates the relationship between cognitive styles and preference to visual abilities. The proliferation of multimedia resources on the Web presents the eLearning community with opportunities to provide quality learning experience, but current approaches do not adequately provide a semantic approach for organising multimedia resources to provide a learner with tailored learning. This paper accordingly presents a way of organising existing multimedia resources based on contexts towards providing a truly personalised eLearning experience. Wood and Ishaya, 2005 described a model of personalised learning through the study of human personality traits based on an earlier model of personality (Loehlin, 1968), which identifies the elements that can be captured for an eLearning profile. Thus, this paper presents an eLearning framework for achieving multimedia context-aware personalised learning by extending the mapping that exists between personality traits and individual preference for inherent multimedia features as it affect learning.

In the next section, a background on current use of multimedia for learning and personality traits for personalised learning has been reviewed. Section 3 presents a framework for personalised multimedia learning process, highlighting the three main components of the framework. Section 4 presents a context-based multimedia annotation approach with detailed illustration through prototype experimentation for the multimedia metadata component of the framework and section 5 concludes the paper with further research development.

2. Background

This research is based on two related contexts – the current use of multimedia for learning and personality traits for personalised learning.

The integration of technology into the realm of teaching and learning has long since evolved learning from the traditional classroom instructor-led teaching into a flexible learner-centric knowledge-oriented solution. Traditional multimedia eLearning is usually a passive, linear, boring experience merely targeted at delivering information to learners without really getting them actively involved in the learning process.

There are a number of existing models, which aim at providing standards that enable the sharing, reuse, importing, and exporting of learning objects among eLearning systems. IEEE LTSC Learning Object Model (LOM)\(^2\) and Shareable Courseware Object Model (SCORM)\(^3\) are among the most influential of these models. LOM and SCORM specification have not taken into consideration all pedagogical issues in relation to adaptive multimedia eLearning. There are a lot of efforts from content-based multimedia retrieval but the

\(^1\) http://www.nelsonmedia.com
\(^2\) http://ltsc.ieee.org/wg12/
\(^3\) http://adlnet.org/
effectiveness of proposed approaches are still constrained by the semantic gap between low-level multi media representation and high-level human perception of multimedia content (Aigrain et al, 1996 and Lefevre et al, 2003). One of the greatest challenges for learners in a multimedia e-learning environment is the lack of support for intelligent multimedia semantics processing and pedagogies in learning practices. Commercially available eLearning systems like BlackBoard and WebCT tend to apply basic multimedia representation techniques, which merely facilitates knowledge access to learners at a very basic level (Al-Khalifa, 2005 and Ohene-djan and Naqvi, 2005). There is a clear gap in terms of technology development and applications to support learner-centric eLearning. Recent research efforts in multimedia semantic and knowledge organisation has shown that context can be used to add a new pragmatic layer on top of the existing feature and low-level oriented multimedia representation to facilitate high-level semantic integration and interoperation (Huang et al, 2006).

To improve the efficiency of learning, eLearning systems can be designed in a way to cope with different learning styles and goals of students. eLearning systems can offer several features to compensate the differences in learning outcomes caused by various learning styles and individuality of learners. The differences between good learners and bad learners usually lie in their use of learning strategies (Ishaya and Wood, 2005). The key challenge therefore, is how technology can be leveraged to ensure a comprehensive, flexible, and natural learning process.

2.2 Personality traits for personalised learning

One of the driving forces behind eLearning developments is the potential to deliver tailored, one-to-one learning experience. Research has generally focused on adapting content, usually in relation to their learning styles (Abdullan and Davis, 2005 and Stash, 2004), or adapting the learning path as experienced by Power et al (2005), in relation to what the learner knows. Both these methods capture a limited model of the student that can only be used for a small range of adaptations. While this is certainly a step in the right direction, a much more flexible student model is required. Looking at everyday interactions one element that is used frequently to assess student potential is their personality.

Psychologists have been studying personality and its relation to human behaviour for over a hundred years. Recently one particular theory has come to the forefront, that of the five factor model (often referred to as the Big Five), by McCrea and Costa (2003). This model describes personality as bipolar traits, or high level factors, that all individuals have to one degree or another. These traits are most commonly known as Neuroticism, Extraversion, Openness to experience, Agreeableness, and Conscientiousness. These are assessed via survey instruments such as the sixty item NEO Five + Factor Inventory (NEO-FFI) and the more comprehensive, 180 item, NEO Personality Inventory (NEO-PI) (McCrea and Costa, 2003). Both these instruments have been rigorously tested over the last ten years, across sex, age, and cultures. These instruments have been recently used in a variety of educational settings to identify the correlation between certain learning behaviours and attitudes, and the learners’ personality (Chamorro-Premuzic et al, 2005 and Komarrajlu and Karau, 2005).

From an extensive literary review of personality psychology two models were developed for use within eLearning software system, as described in Wood and Ishaya (2005). The first, models the high-level components of the behavioural process, such as personality, motivation and cognition. While the second focuses on the elements of personality which are divided into common static aspects and individual dynamic aspects. This captures the trait aspects described above as the static part of personality, the initial population of the learners profile via survey instruments. The dynamic aspects are captured and updated during interaction with the eLearning environment. These include, but are not limited to, such elements as personal values, learning goals, current abilities, beliefs, preferences and habits.

Given an understanding of the behavioural process, eLearning systems can use personality to predict learner behaviour and therefore adapt all aspects accordingly. The main research challenge is how to establish a clear mapping between individual personality traits and multimedia learning resources for true personalised learning experience. Thus, the next section presents a comprehensive framework towards addressing the stated challenge.

3. Framework for Personalised Multimedia Learning Resources

Since existing research such as in Ishaya and Wood (2005), have shown a relationship between personality traits and preferred learning styles, the underlying hypothesis driving this research is that there is a relationship between personality traits and multimedia learning resources. Section 3.1 will briefly present a list of personalised multimedia resource requirements and the proposed framework in section 3.2.

3.1 Requirements for Personalised Multimedia Resources

Personalization requires at least some basic user profile to specify various multimedia content semantics that are particularly useful for a learner within a given context. Firstly, the system must be capable of generating profiles with minimal interruption for user learning preferences (Kim et al, 2002 and Ohene-Djan and Naqvi 2005). The model of personality provides the ability to look at profiling structures behind individual preferences and at a much greater granularity and therefore should have the flexibility to deal with both specialised and general predictions, without leaning towards one or the other (Ishaya and Wood, 2005).

The main focus here is towards personalized support and removal or replacement of transactional model of learning. It should cope with changes in varied delivery modes in relation to appropriate learning styles, strategies, and prior knowledge of individuals. Secondly, as shown in Huang et al (2006), there must be multimedia semantic and knowledge organisation component to facilitate high-level semantic integration and interoperability in order to determine an appropriate mechanism for selection of suitable multimedia content to suit a particular learning context. Lastly, the need to define a set of computable parameters that provide consistent mapping between both the static and dynamic aspects of personality profiles and annotated multimedia metadata.

3.2 A proposed development framework

This section presents the framework, linking the processes described in section 3.1. While, different aspects of the framework are being implemented, focus of the implementation presented in this paper is on the multimedia metadata generation component.

\[1\] http://ltsc.ieee.org/wg12/

\[2\] http://adlnet.org/
In order to demonstrate the potential benefit of the proposed component of the framework. Next section presents detailed description and demonstration and multimedia contents vary. 

The most challenging problem in multimedia related management is the semantics capturing of media resources. In the framework shown in figure 1, we have taken a context-based approach to solve the problem of multimedia semantic extraction and indexing – the multimedia metadata manager. Multimedia learning resources are annotated based on various contextual dimensions that could be identified for the resources. The annotated metadata is stored in a machine-readable (RDF/XML) format ready for efficient context-based retrieval. RDF has a specific declarative semantics, which is specified independently of any RDF processor. This made it more suitable for modelling context. XML is used for basic resource content description, since its meaning is only determined by the actions that programs undertake on it. The RDF/XML combination helps to facilitate interoperability of the model.

The profile matching ontology is defined to provide a mapping for automatic generation of contextual metadata semantics of multimedia based a particular learner profile for personalised learning resources from the multimedia metadata manager (repository). Different factor analysis is performed at this stage to enable the extraction of predominant directions on which different personality traits and multimedia contents vary. 

Next section presents detailed description and demonstration of an implementation of multimedia metadata manager component of the framework.

4. Experimentation

In order to demonstrate the potential benefit of the proposed framework, two models have been defined and demonstrated. The personality component presented in Wood and Ishaya (2005), established the need for a link between individuals and learning styles. Huang and Eze (2005) defined a multimedia semantics integration model based on the context mediation with implementation presented in the following sections.

4.1. Multimedia Context formalization and representation

Context has been identified from the literature to play a crucial role in human knowledge representation, reasoning, and perception (Huang and Eze, 2005). Therefore, multimedia information retrieval systems need the ability to represent, utilise and reason about context to help improve semantic representation and management of multimedia resources. The framework relies on the identification and use of contextual information about the multimedia resources to enhance semantic understanding of such multimedia content for personalised eLearning resources.

Context is used to add a new pragmatic layer on top of the existing feature and low-level oriented multimedia representation to facilitate high-level semantic integration and interoperation. Typical multimedia contextual sources include:

- literal statements, such as free semantic annotation of multimedia resources like image, audio, video, etc.
- knowledge sources, such as information from web scraping, transcription of the audio part of video files
- entity’s properties and general descriptive metadata, such as author, title, date of publication, etc.
- inference and deductions based on a well-defined ontology

The context data is stored in such a way that higher level facts can be inferred from the individual pieces of context data. An ontology knowledge base of explicit concepts is maintained to facilitate mapping and inference deduction from the recognised contexts. The contextual information is represented as metadata using XML and RDF. RDF has a specific declarative semantics, which is specified independently of any RDF processor. This made it more suitable for modelling context. XML is used for basic resource content description, since its meaning is only determined by the actions that programs undertake on it. The RDF/XML combination helped to facilitate interoperability of the model. A generic context mediation model is presented to facilitate the integration, understanding, and discovery of multimedia semantics.

The multimedia semantic context descriptive (CON) model has been formalized using Feature Notation (Scheurer, 1994), which consists of the various members defined below: Given a set of multimedia object, X; let s represent semantically isolated segments in the object x ∣ x ∈ X (i.e. shot/track in video/audio, parts in image or text). EP is the set of all possible Extracted Properties. Various extractable media information like media type, media format, number of frames, file size, etc are all members of this set. Only two conditions are imposed on EP: i) this set may not be empty. This constraint was designed to rule out the possibility of not having some basic media information; ii) this set is a disjoint subset of all resource descriptors (RD).

\[
\begin{align*}
EP &\subseteq \mathcal{A} \\
\text{IsDisjoint}(EP, RD) &\quad EP \perp_{d} RD
\end{align*}
\]
ARD is the set of all possible annotator’s resource description. Annotator here can be a human user or a software agent. Members of this set include basic descriptive metadata such as author, title, URI, date of publication, etc. This set is also constrained to be nonempty to avoid the possibility of not having at least a description for the multimedia object. This set like the EP set, is also a disjoint subset of RD.

\[
\text{ARD: Set of annotator’s resource description} \\
\text{ARD} \subseteq \exists \_a \_E \\
\text{IsDisjoint(ARD, RD)} \\
\text{ARD} \_a \_D \_R \_D
\]

KS is the set of all possible contextual or knowledge sources. Members of this set include information from webscraping, audio transcription for video resource, ontology knowledge base, etc. This set is constrained to be nonempty and also a disjoint subset of semantic description scheme (DS), as it adds uniqueness to the model and enforces clearer automated semantic discovery and organisation.

\[
\text{KS: Set of all contextual or knowledge sources} \\
\text{KS} \subseteq \exists \_a \_E \\
\text{IsDisjoint(KS, DS)} \\
\text{KS} \_a \_D \_D
\]

ASD is the set of all possible annotator’s semantic description. This set comprises free textual annotation of multimedia objects. ASD is a disjoint subset of DS and can be empty.

\[
\text{ASD: Set of annotator’s semantic description} \\
\text{IsDisjoint(ASD, DS)} \\
\text{ASD} \_a \_D \_D
\]

DS is a strict union of the disjoint sets KS and ASD. Therefore the set inherits all constraints imposed on KS and ASD. DS is clearly nonempty since one of its disjoint subsets (KS) is nonempty.

\[
\text{DS: Set of all semantic description scheme} \\
\text{KS} \_a \_E \_a \_D \_S \_D \\
\text{and ASD}
\]

RD is a strict union of EP and ARD, and thus inherits all constraints imposed on both sets. Hence this set is also nonempty.

\[
\text{RD: Set of all resource descriptors} \\
\text{EP} \_a \_E \_a \_D \_R \_D \\
\text{and ARD}
\]

Based on the above definitions, the following relations can be defined:

\[
y \text{ReRD} x \iff x \in X \land y \text{ is a resource descriptor for } x \text{ based on } s | s \in RD \\
y \text{ReDS} x \iff x \in X \land y \text{ is a semantic description scheme for } x \text{ based on } s | s \in DS
\]

Finally, the semantics of multimedia object, x in context represented as CON(x) can be formally defined as:

(1)
modified on the Resource Description panel. Temporal audio/video resources can also be browsed in Semantic Segments on the preview panel. With references to ontologies and schemas, integrated semantic descriptions of a knowledge objects could be saved in the centralised knowledge base, in the options of XML or RDF.

5. Conclusion and further research

With evolving trend in learning through the use of web technology, eLearning systems are expected to provide personalised learning resources for effective learning, but current approaches does not adequately provide a semantic approach for organising multimedia resources toward facilitating personalised eLearning experience. This requires the provision of both personalised learning and a broader range of learning material, such as video, pictures and sound.

This paper has highlighted the three elements that are needed to bring personalised multimedia to the learner. A flexible, cross-domain profiling method is provided by the personality represented component. The framework describes the semantic description of multimedia that identifies the specific context of learning materials, and the ontological matcher that maps various elements from the profile to the media context.

The experimental prototype presented implements one of the key components of our framework – the multimedia data manager. It demonstrates a new context-based approach for annotating the semantics of multimedia resources semi-automatically for personalised delivery of learning resources. We are currently working towards extending this prototype to support personality capture and representation. Our future work will include further research and experimental prototype that implements the profile ontology matcher for an effective personalised learning experience.

References


meta-search agent." Electronic Commerce Research and Applications 1 (2) 150-173.


Emmanuel Eze received his BSc Honours Degree in Computer Science from The University of Nigeria in 1999. He is presently concluding his PhD in Computer Science at the Centre for Internet Computing, The University of Hull, UK, which commenced in 2003. He has also been providing IT consultancy service to a variety of companies in the area of Internet-based software technologies. His research interests include Multimedia Semantics, Multimedia Ontology, Contextual Knowledge Engineering, Internet Technologies, and the building of Software Architectures. Emmanuel is a member of BCS, ACM, and IEEE Computer Society and has published papers in leading conferences and journals.

Tanko Ishaya is a lecturer in computing at the University of Hull, Scarborough Campus. He obtained his MSc and PhD in Computation from the University of Manchester in 1997 and 2001 respectively. His research focuses on approaches and technologies to enhance the provision of personalised eLearning, eBusiness, and Information engineering. He is currently investigating on how context and human personality traits could help in providing adaptable multimedia eLearning. He serves as a programme committee member on many International conferences and a member of the editorial board for Interdisciplinary Journal of Knowledge and Learning Objects, Informing Science Press, since 2005.

Dawn Wood is a postgraduate doctoral student at the Centre for Internet Computing, University of Hull. She holds a first class degree in Computer Science from Hull University. Her current publication list includes several conference and journal papers on the issue of personalisation within eLearning. Miss Wood’s interests include the personalisation of the Semantic Web, Multi-agent systems for eLearning applications, ontology development, and metadata for learner support systems.