Notes for a Collaboration: On the Design of a Wiki-type Educational Video Lecture Annotation System

Stefano Ferretti, Silvia Mirri, Marco Roccetti, Paola Salomoni
Department of Computer Science, University of Bologna,
Mura Anteo Zamboni 7, 40127 Bologna, Italy
{sferrett, mirri, roccetti, salomoni}@cs.unibo.it

Abstract

We describe a collaborative annotation system for the production of rich media e-learning contents. Our system exploits a wiki-like interface that allows cooperative users to enrich didactical multimedia lectures with additional information, such as captions, annotations and comments. Before delivering them to users, these additional contents are automatically adapted and integrated with video lectures based on user profiles and requests. This way, the additional contents collaboratively provided by users not only represent an effective alternative that students may follow to better understand the lecture, but also allow for a fine-grained customization of the didactic material. This is an important aspect which represents a step forward towards the Web(2.0)-ification of e-learning technologies. An experimental assessment shows the viability of our approach.

1. Introduction

The educational potential of Web 2.0-based technologies, thought as means for distributing and sharing e-learning material, is huge. Even if these new Web-based solutions were not explicitly designed for e-learning scenarios, it has been recognized that they can truly empower learners and create new exciting opportunities for cooperation and collaboration in distributed didactical contexts.

Nevertheless, various e-learning tools for collaboration and teamwork have been developed in the recent past and many other initiatives are actually under development [10]. Prominent examples are those which allow to produce Learning Objects (LOs) that comply to the SCORM standard [11]. Indeed, these tools include several features such as the possibility of using e-mailing lists and newsgroups, whiteboards, bulletin board systems, chats and online presentation tools. Yet, these e-learning systems are still far from truly meeting the requirements of the Web 2.0 philosophy, such as i) open participation, ii) data decentralization, iii) assembly from diverse sources, and iv) freedom to share and re-use [9].

Instead, today, new solutions for e-learning can be built which rely upon the concept to exploit user contribution and the strength of collective intelligence. Indeed, to fully let learners to proficiently enjoy e-learning contents produced in a collaborative way, the didactical material should be delivered to users according to a format that fully satisfies the learner's needs and expectations. User collaboration and content personalization thus become the two key issues to address in order to provide effective Web 2.0-based e-learning experiences.

With all this in view, we have designed a new collaborative annotation system for multimedia e-learning contents. Our system makes use of LOs encoded as (SCORM compliant) SMIL documents. Cooperation is allowed by enabling users to actively add textual information (such as captions and annotations) to SMIL video lectures, based on a wiki-like editing interface system.

Specifically, two main features are provided. First, upon the insertion of captions to be associated to a lecturer talk, the temporal starting point in time of each caption and its duration may be specified thanks to a wiki-type textual syntax. Second, annotations may be added to each slide, which represent alternative (or additional) descriptions to the images shown during the lecture. All these annotations are embedded in one or more separate HTML documents to be accompanied with the SMIL rich media content.

It turns out that the use of additional information, such as captions and annotations, not only provides further explanations that may be of help to the learner, but also improves the level of accessibility of multimedia contents [2, 6, 11, 12]. For example, blind people could gain a great advantage from having additional textual annotations that provide explanations to contents presented in the slides of the video lecture. Moreover, mobile users connected through a PDA with a low bandwidth (e.g., cellular) network would be happy to anticipate the textual contents of a given

0-7695-2997-6/07 $25.00 © 2007 IEEE
DOI 10.1109/ICSC.2007.18

lecture while revising an attended lesson during their return home by train, and delaying later in time the video and audio parts of that video lecture.

With this in view, our system is able to adapt these rich media contents based on the user profile and request [11]. For example, if the user is a hard of hearing one, the lecturer's speech (stored as an audio file in the rich media LO) is automatically replaced by textual captions (provided by some collaborative user and associated to the lecture).

Experimental results obtained from a preliminary evaluation assessment confirm the viability of our approach.

The remainder of this paper is organized as follows. Section 2 presents an overview of our system. Section 3 describes the annotation system we have devised for the cooperative editing of SMIL-based rich media contents. In Section 4, we report on an experimental assessment we conducted on our system. Finally, Section 5 provides some concluding remarks.

2. Our System at Work

Figure 1 shows a simplified view of our system. Lecturers and learners cooperate together as follows. A lecturer publishes a new SMIL-based, rich media lecture (publishing step p_1 in the figure) by uploading it on our system. Then, learners may enrich this didactical material by adding annotations to slides presented during the lecture, or even to add captions to the lecturer speech (editing steps e_1, e_2 in the figure). This is accomplished by exploiting a user interface which resembles a classic wiki.

Our system elaborates and manages such data edited by users to be properly inserted within the SMIL-based rich media content. This process is described in the next section.

Concurrently, e-learners are allowed to retrieve the didactical content by exploiting facilities offered by an adaptation system [11]. This is a distributed software component in charge of providing each user with an adapted, suitable version of the LO, based on her/his specific profile and request. Customization of contents is performed by considering: i) the personal characteristics and physical limitations that the user may have; ii) the technical characteristics of the device being used by the user. A service oriented architecture is devoted to manage all the transcoding steps (e.g., image conversion, audio conversion, video conversion, speech-to-text, conversion from SMIL to XHTML, text translation into different languages), needed to profitably adapt the LO to the specific user. The interested reader, may find a detailed discussion of such adaptation system in [11].

In practical cases, flexible systems for content adaptation strongly benefit from having collaborative users that cooperate to add annotations, comments and captions to the learning material. For example, the presence of captions associated to the lecturer speech could be of great help for those learners which are deaf, hard of hearing, or simply foreign students. The same holds for additional textual descriptions associated to the lecture slides. Moreover, captions are useful also for those students that exploit devices unsupplied with audio capabilities [1, 5, 11].

Thus, for instance, based on our system, a deaf user, which accesses the lecture through the use of a common PC, will receive a high quality SMIL video lecture with captions activated (delivery step d_1 in the figure). Instead, a mobile user joining the system through the use of a PDA, receives a specific version of the lecture that fits the technical characteristics of the exploited client device (delivery step d_2 in the figure). Finally, a blind user will receive a properly adapted XHTML document obtained through the transcoding of the SMIL video lecture (delivery step d_3 in the figure). This way, it is ensured that all the didactic material will be presented to the blind user as a linear sequence of contents (thus avoiding cognitive overload); then, textual contents can be converted at the client-side, for example by resorting to a screen reader or a Braille display.

Finally, our system is provided with appropriate packaging facilities so that the multimedia didactical material cooperatively produced by lecturers and students and properly customized by the adaptation system can be exported as SCORM-compliant LOs.

This paper specifically focuses on the annotation system that enables users to cooperatively contribute to enrich LOs. A detailed discussion is reported in the following section.

3. Designing a Wiki-like Video Annotation System

In this section, we present our solution for the cooperative production of rich media LOs. To this aim, we first outline the main issues concerned with the development of collaborative content management systems in the context of e-learning. Then, our approach is presented.

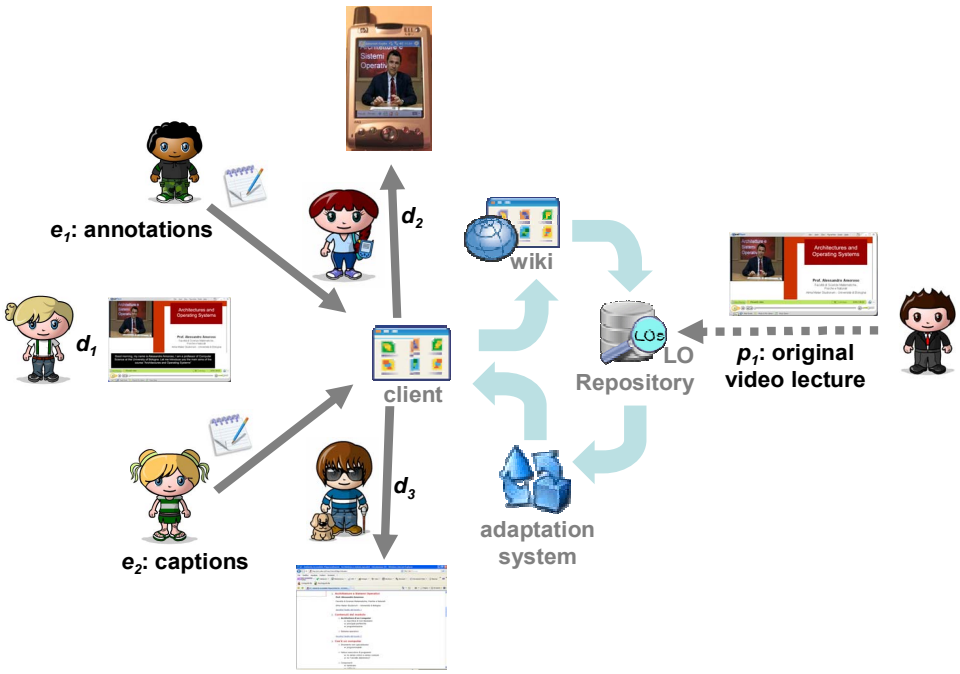

Figure 1. Overall System Architecture.

3.1. Collaborative Wikis: a Background

Today, wiki engines are certainly the most widely utilized content management systems and represent one of the key technologies for enabling user collaboration in the Web. These Web 2.0-based systems embrace an approach where contents are collaboratively edited by a multitude of users [9]. This particular philosophy of content editing now goes under the name of "*the wiki way*" [3].

In classic wiki systems, the typical source format to edit a new content is the *wikitext*, i.e., a simple kind of markup language that exploits plain text with a few simple conventions for linking, structuring and styling. The edited content is automatically managed to produce a final HTML document [3]. There is no commonly accepted standard for the wikitext language. The grammar, structure and keywords only depend on the particular wiki software exploited at the website. As an example, different wiki engines may have different syntax conventions to specify links.

Recently, new wiki engines have been developed which support complex editing methods, including ways to add text decorations, tables, images and formulas. For example, MediaWiki, the software exploited in Wikipedia, uses a markup language able to produce text decorations which can be equivalently obtained through common HTML tags. OpenWiki allows to write mathematical formulas based on MathML [8]. In TWiki, users can include Latex commands [13]. The SnipSnap wiki engine enables to write organigrams and diagrams [7].

However, while these systems generally result as effective and easy to use, a clear limitation is that they only enable the editing of spare HTML documents. There is no opportunity for users to collaboratively operate on rich media contents, or even to add to these contents some additional information. We claim that to really exploit the power of user contribution, the strength of the collective intelligence and to maximize the cooperation among users, new solutions must be devised for the production, the management and the presentation of rich media contents.

Following the wiki way, we are looking for an easy-to-use system interface that allows to add to video lectures new multimedia contents and enrich the didactical material. Thus, each student would be enabled to add and synchronize his own provided media resources (e.g., audio, video, text, images) with other ones already available in the system. To reach this goal, a practicable choice is that of structuring multimedia LOs as SMIL-based video lectures.

In this work, we focus on an approach to provide users with the possibility of adding textual information such as captions, or annotations encoded as alternative textual representation of media contents (the possibility of adding and synchronizing other media types to the video lecture in an ongoing work).

3.2. Blending SMIL and Wiki for Flexible Collaboration

Our system exploits didactical contents which are encoded as SMIL documents and packed as SCORM-compliant LOs. To enable users to enrich the didactical material with annotations, comments and captions, SMIL video lectures must have a predefined structure. In particular, at least two specific regions should be available in the SMIL document: i) a region devoted to show the captions, and ii) a region where image slides are shown (annotations and comments are associated to these slides). This way, when converting a new textual information inserted by some users, based on its type (i.e., caption, annotation) our scheme associates that data to the related region of the SMIL LO.

As to captions, these must be synchronized to the lecturer speech. Thus, for each phrase, the starting point in time and the duration of the related caption must be specified (i.e., the caption should remain visible while the lecturer is pronouncing the sentence).

In SMIL, there are three different attributes for specifying the timing properties of a given SMIL element: i) `begin`, which identifies the time for the explicit begin of an element; ii) `end`, which identifies the explicit end of an element; iii) `dur`, which identifies the explicit duration of an element. With this in view, we added to the wikitext a specific syntax for each of these three attributes (see Table 1).

Table 1. Syntax to Edit Captions: Main Issues.

SMIL attribute	`<text>`	`<textstream>`
`begin`	$	$$
`end`	__	__$
`dur`	::	

In particular, in our wiki syntax a `$` symbol corresponds to the `begin` SMIL attribute, the `__` symbol corresponds to the `end` attribute and, finally, the `::` symbol corresponds to the `dur` SMIL attribute. Summing up, to add a caption, a user should exploit a syntax which is as follows

```
$'start'__'end' 'sentence'
```

or

```
$'start'::'dur' 'sentence'
```

where `'start'` is the starting point, specified in seconds, `'end'` represents when the caption should disappear (time specified in seconds), `'dur'` represents the duration of the element, and `'sentence'` is the specific caption to be shown.

Figure 2. SMIL Video Lecture, courtesy of Prof. Amoroso (University of Bologna).

As an example, consider the SMIL video lecture shown in Figure 2. Such lecture is composed of three regions: i) a region associated to the video showing the lecturer ii) a region showing the image slides, and iii) a region, initially hidden, devoted to show captions. Moreover, an audio flow representing the lecturer speech is included in the SMIL document. If a cooperative user edits the following lines of code

```
$3::7 Good morning, my name is
      Alessandro Amoroso, I am
      professor of Computer Science
      at the University of Bologna
```

Figure 3. SMIL Video Lecture with Captions, courtesy of Prof. Amoroso (University of Bologna).

(or, alternatively, `$3__10 Good morning, ...`) then the specified caption is associated to the lecture, starting from the third second of the lecture, for 7 seconds. Thus, this caption becomes available to be visualized, once the SMIL lecture is played out, depending on the preferences and the profile of the student enjoying the lecture. Figure 3 shows the resulting SMIL video lecture with the region, which is associated to captions, activated.

To obtain this result, once the user adds a caption using our wiki-like interface, the system automatically

adds to the document the SMIL markup code needed to display that caption for the specific time interval, i.e., it adds a `<text>` element in the proper position of the SMIL document and specifies its duration using the related timing attributes.

For instance, w.r.t. the previous example, the following SMIL code fragment is inserted into the document (some syntactic trivia has been removed, for the sake of better comprehension)

```
<text begin="3s" dur="7s"
      region="region_caption">
      Good morning, my name is ...
</text>
```

where the `begin` and `dur` attributes specify the timing properties of the caption, while the `region` attribute specifies that the caption is visualized in the region named `region_caption`.

To add several captions to be shown at different times during the lecture, the user is simply asked to write them in sequence in the (wiki-like) form, using the syntax described above. All these statements will be automatically put in sequence in the SMIL document, by including them in a `<seq>` … `</seq>` SMIL structure.

Apart from the classic use of direct captions in SMIL documents, it is well known that Real currently provides the best support for text captions in SMIL documents, using its XML-based RealText format. The idea is that it is possible to create a `.rt` file, using this proprietary textual format, where all captions with their related timing properties are included. Then, such file is associated to a given region of the SMIL document.

In our system, we included the possibility of creating captions using such a technique. The user must only use a different syntax, whose meaning is very similar to the previous one (see Table 1). Basically, to specify the starting point of a caption and storing it in a `rt` file, the `$$` symbol must be used. The ending point of the caption is indicated with the `__$` symbol. Hence, w.r.t. the previous example, if the user adds the following lines of code (instead of those shown above)

```
$$3__$10 Good morning, my name is ...
```

then, the system adds to the SMIL document the following code fragment

```
<textstream src="captions.rt"
      region="region_caption" …/>
```

which associates to the specific region devoted to show the captions (i.e., `region_caption`), a specific text file (i.e., `captions.rt`) where captions are stored, encoded based on the RealText format. Instead, the caption is added to this last file using the following RealText syntax

```
<Time begin="0:00:03.0"
      end="0:00:10.0" /><clear/>
      Good morning, my name is ...
```

As to the additional annotations that can be associated to the slides shown during the video lecture (e.g., step e_1 in Figure 1), the wiki-like interface provides the user with a list of the slides related to the lecture. For each slide, a button is associated. Once a particular button is clicked, annotations can be added (using a classic wikitext syntax) for the related slide.

Such annotations are then translated and stored as a HTML document to be associated as an alternative media for the image representing the slide. This is accomplished by adding to the SMIL tag related to the visualization of the slide, a `longdesc` attribute which specifies the location of the HTML file containing the alternative annotations. Translation for wikitext to HTML is performed as in all common wiki systems.

The following example shows a SMIL code fragment where, for a given slide (encoded as an image stored in the `1.gif` file), the alternative data contained in the `1.html` document is associated (using the `longdesc` SMIL attribute):

```
<img region="region_slide"
     src="1.gif" dur="15s"
     alt="…" longdesc="1.html" />
```

4. An Experimental Assessment

In this section, we report on a performance evaluation we conducted to assess the efficacy of our system. Basically, our SMIL-based LOs were made available for annotation, distribution and customization. Users were allowed to enrich multimedia contents with textual annotations, comments and captions, that were then employed to customize the LO to be delivered to other students.

Needless to say, when distributing this collaboratively produced e-learning material, the fundamental aspect to validate is concerned with the overhead due to the request and adaptation of these contents, based on the additional data provided by users. In particular, scalability is the main factor to assess, due to the fact that potentially a high amount of students could ask for some particular LO (which should be specifically customized for each different learner).

With this in view, we carried out a specific experimentation. Several concurrent requests were randomly generated for dozens of rich media LOs, collaboratively produced by resorting to our annotation system. For each request, a specific user profile was

randomly selected, thus imposing a specific resource customization [11].

For the sake of scalability, the adaptation unit was equipped with a caching system, so as to avoid subsequent identical adaptations for same resources [11]. The cache size was set to 5-8 GB at each node; caches were refreshed every 20 minutes (while each run lasted one hour).

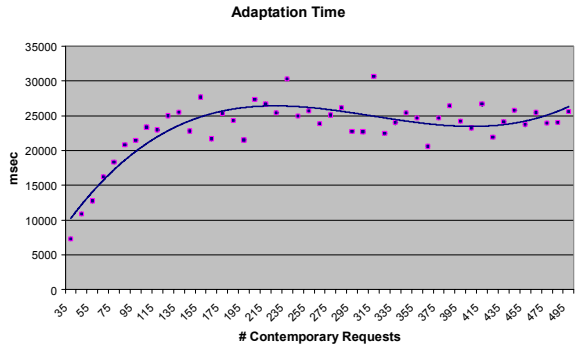

Figure 4. Average Adaptation Time.

The number of contemporary random client requests was set to vary from 0 up to 500. Figure 4 reports the average time required to receive, analyze the request, adapt contents and deliver them back to the client, depending on the number of contemporary client requests.

As shown in the figure, the system results quite scalable, since response times do not increase linearly with the number of contemporary requests. This is probably due to the effective use of the caching system, which strongly reduces the number of performed transcoding activities.

5. Conclusions and Future Work

In this paper, we have described a new annotation tool for the production of rich media SMIL LOs in collaborative e-learning environments. Users are allowed to enrich LOs by adding information such as annotations, comments and captions to the available SMIL video lectures. This way, rich media contents can be properly customized before being delivered to specific users.

Encouraging users to cooperate and produce highly detailed rich media contents has a very positive effect for the provision of accessible and flexible educational experiences to users. We claim that this is the way for an effective Web(2.0)-ification of e-learning.

We are currently investigating the possibility of extending our system so as to allow students to add multimedia contents (i.e., not only textual data) to the available LOs. This way, it will be possible to have an open, cooperative system for the production of SMIL based multimedia presentations.

6. Acknowledgments

This work is supported by the Italian M.I.U.R. under the MOMA and DAMASCO initiatives.

7. References

[1] F. Azouaou, C. Desmoulins, W. Chen, "Semantic Annotation Tools for Learning Material", *Proc. of Workshop on Applications of Semantic Web Technologies for Adaptive Educational Hypermedia*, August 2004, The Netherlands, 22-26.

[2] J.A. Barron, L. Fleetwood, A E. Barron, "E-Learning for Everyone: Addressing Accessibility", *Journal of Interactive Instruction Delivery*, 16(4):3-10, 2004.

[3] A. Désilets, L. Gonzalez, S. Paquet, M. Stojanovic, "Translation the Wiki way", *Proc. of the ACM International Symposium on Wikis,* Denmark, August 2006, 19-32.

[4] S. Ducasse, L. Renggli, R. Wuyts, "SmallWiki: a meta-described collaborative content management system", *Proc. of the ACM International Symposium on Wikis*, San Diego, California, October 2005, 75-82.

[5] D. Gasevic, J. Jovanovic, V. Devedzic, "Enhancing Learning Object Content on the Semantic Web", *Proc. of the IEEE International Conference on Advanced Learning Technologies (ICALT'04)*, 2004.

[6] M. L. Guenaga, D. Burger, J. Oliver, "Accessibility for e-Learning Environments", *Computers Helping People with Special Needs*, LNCS, Springer, 2004, 157-163.

[7] Jugel M.L., Schmidt S.J., "SnipSnap: the easy weblog and wiki software", http://www.snipsnap.org/.

[8] OpenWiki: http://www.openwiki.com.

[9] T. O'Reilly, "Web 2.0 Compact Definition: Trying Again", 2006, http://radar.oreilly.com/archives/2006/12/web_20_compact.html.

[10] M.F. Paulsen, "An Analysis of Online Education and Learning Management Systems in the Nordic Countries". *Online Journal of Distance Learning Administration*, 5(3), 2002.

[11] P. Salomoni, S. Mirri, S. Ferretti, M. Roccetti, "Profiling Learners with Special Needs for Custom E-Learning Experiences, a Closed Case?", *Proc. of the 4th International Cross-Disciplinary Conference on Web Accessibility (W4A 2007)*, ACM Press, Banff (Canada), May 2007.

[12] A. Savidis, C. Stephanidis C, "Developing inclusive e-learning and e-entertainment to effectively accommodate learning difficulties" *ACM SIGACCESS Accessibility and Computing*, ACM Press, 83:42-54.

[13] P. Thoeny, "TWiki: Enterprise Collaboration Platform", http://twiki.org.