StoryML: An XML Extension for Woven Stories

Petri Gerdt[1], Piet Kommers[2], Jarkko Suhonen[1], and Erkki Sutinen[1]

[1] Dept. of Computer Science, University of Joensuu, Finland
{pgerdt,suhonen,sutinen}@cs.joensuu.fi
[2] Dept. of Educational Technology, Univ. Twente, The Netherlands
P.A.M.Kommers@edte.utwente.nl

Abstract. StoryML is an XML-based representation of metadata elements connected to collaboratively written stories. The StoryML specification gives means to interpret and specify the core characteristics of stories. Hence, StoryML supports the functionality of information retrieval, filtering and adaptive representation of stories. These intelligence properties make collaboratively written stories a significant platform for truly activating, open learning environments.

1 Introduction

In this paper we elaborate the concept of *woven stories*, first presented in [3], into the direction of an intelligent learning environment. A woven story is the result of several authors' efforts in a shared writing space, where the authors may write story sections and link them together. A woven story is thus a hyperdocument or alternatively hyperspace, which consists of an arbitrary set of story sections and links between them. The woven stories idea differs from collaborative writing in that the participating authors may not have a common goal to write a document jointly. Instead, they explore their ideas and express their views, for example different authors might want to change the ending of a story by writing an alternative ending in a new story section. A woven story is visualized as a graph, where the story sections are represented by nodes and the relations between them by links. We have explored the concept of woven stories in a experimental Internet-based application labeled Woven Stories 2, or WS2 for short [2].

The opportunity to adapt the hyperspace, consisting of stories, according to the specific characteristics and needs of the users is interesting. This adaptation means for example that the stories would be represented to different users in different ways. Also it would be useful to the user of woven stories to retrieve subsets of collaboratively written stories. To achieve adaptation, filtering and retrieval in woven stories, we need a way to define an extension for collaboratively written stories.

We have developed a XML-based specification for woven stories which is referred to in this paper as *StoryML*. The StoryML specification specifies a data structure for capturing the characteristics of a woven story: the story sections, associations and the authors who have participated in the authoring effort. We present a description of the core element set of the StoryML in Section 3.

S.A. Cerri, G. Gouardères, and F. Paraguaçu (Eds.): ITS 2002, LNCS 2363, pp. 893–902, 2002.

One of the key concepts in the development of StoryML is metadata. With metadata elements one can store all the necessary information related to a specific domain. The core StoryML elements have been inspired by the Dublin Core metadata initiative [1]. Furthermore, StoryML draws elements from issues related to collaborative writing and collaborative concept mapping.

The woven stories concept supports open problem solving in a way that it does not force the user to act in a predefined manner. The intelligence of the woven stories empowered by StoryML comes with the adaptation of stories. This adaptation can be utilized for example during the learning process to illustrate individualized learning paths.

Traditionally, intelligent tutoring systems have supported the learner in his or her learning process to follow a more or less optimal root from a given starting point into the desired learning goal or objective. The tutoring system has tried to help the learner according to his or her preferences as given by the user model. In the scheme of woven stories, however, we emphasize the active role of the learner. The intelligence, or adaptivity, of a learning environment should be guaranteed by tools which help the learner to actively compose a meaningful environment for the learning process. To achieve this functionality an effective metadata scheme, in this case StoryML, is needed.

2 Woven Stories

Our view on systems realizing the woven stories idea is a family of web-based applications that allow users to compose their stories, and link appropriate story sections with pre-existing sections authored by someone else. As a co-authoring environment, woven stories support the user not only as his or her individual cognitive tool, but also as a shared platform to reflect ideas and thought processes of other users with related interests. Thus, a group of users can apply woven stories to tasks such as creative, open problem solving [2].

2.1 Uses of Woven Stories

The combination of the fact that human narrative thoughts may build upon existing recent expressions and at the same time benefit from alternative thought directions from the parallel co-authors is a potential source of progress in the coming future. The key issue is what mechanisms we develop as viable and fruitful ones for this intertwining process. In these efforts to make meaningful linking and reuse of knowledge sources on the web our woven stories project is not unique. A relevant discipline in this respect is the one initiated by research groups under the IEEE like the group in standard-upper-ontology [9]; its main goal is to find a solid basis for classifying and interrelating conceptual entities automatically. The lesson drawn from their work so far is that linguistic expressions can be utilized to convey the thinking process rather than to capture and represent the solid meaning underneath. This is the reason why woven stories do not attempt to rely on formalisms for the interlinking of narrative elements; it attempts to facilitate the human participants to anticipate to thought-provoking ideational junctions, without giving away one's own complete message.

The utilisation of woven stories is in its early days; a number of application domains have been identified so far:

1. The educational use is in composition and structural awareness in formulating thoughts and accepting lateral connotations and even conflicting perspectives.
2. Communication and mutual understanding in multicultural settings. Explaining technological and scientific constructs soon promotes a paternalistic and colonialist relation between teacher and audience. Woven stories allow us to anticipate to the mental process of successive integration and conceptual reconciliation; the new topic should become interwoven with familiar, preferably ideographic elements from the local culture, myths and episodic repertoire.
3. In the construction of new knowledge also the idea of a matrix of cross-fertilizing ideas seems much more fruitful than the mechanism of comparing conflicting ideas pair wise and deciding 'which is the true one' all the time.
4. The notion that concepts can better be understood in competing contexts is a well-received in rhetoric and theorem-proving: The attack by hybrid non-native seems to articulate the identity of a concept rather than threatening it. A good example is the ontology browser [7].
5. Finally, woven stories help us to redefine the most intrusive element we allowed in cognitive psychology, learning theory and theories about social representations: Concepts themselves. Based upon our pre-occupation with how to emulate human intelligence, we have adopted "object-oriented" ramifications about how to imagine "concepts". This is not a trivial entity as a factor "X" in order to reason about anonymous phenomena. As long as we confront students and designers with the name for a mental expression, we should have at least attempt to clarify its nature. It will be clear that in epistemic analyses we do not successfully characterize the phenomenological nature of concepts. Woven stories and concept mapping are the two more prominent methods to make students and designers aware of the perceptual and reflective nature of concepts.

2.2 The Dependency between Woven Stories and Concept Mapping

It is not so long ago that we started to think again about explicit representations of our thoughts. Cognitivism gave an impulse to accept the human mind as a world than can effectively be explored. Simon and Newell's work opened a new tradition of defining human problem solving as a rule-driven production process. The succeeding era of building expert systems showed exactly to what point this formalism works: Just before we ask ourselves how the human mind succeeds in controlling the direction and the span of the search.

It seems that both the procedural and the declarative representation are necessary. On top of that, the human problem solving capacity needs configurational awareness. This is exactly enabled by schematic representations. The types of knowledge can be classified along two dimensions: Episodic vs. semantic and declarative vs. procedural. It has become clearer lately how episodic memory plays a crucial role in the transition from declarative into procedural knowledge. Students with a good short-term memory capacity tend to rely less on episodic knowledge; these students seem to have the natural gift to transform new information into operational knowledge. The students with a weaker short-term memory need to reconstruct earlier given information via situational cues and via semantic elaborations like analogue reasoning; we call them students with a holistic style, as they need to recruit a large repertoire of prior knowledge before the factual date can be reconstructed.

The serialistic students proved to be much quicker learners. However in the mid and long term the serialists show a weaker integrated knowledge structure; the speed of memory prevents them from making connections with what they already know. After few years of study the holistic-style students show an easier and more flexible problem-solving style [5,6].

The study further showed that concept mapping fits best to the natural study style of the holists. It proved to be difficult to make serialistic students benefit from the concept mapping approach. The typical reason why holists benefited more from the concept map was that it releases the short-term memory load. The typical correlation between holistic style and the problems to perceive and report information in the original sequence seems to do with the lack of episodic encoding. Here we expect to find woven stories as a beneficial compensation for the problems holistic students have in chronologic reports. In summary: woven stories and concept mapping have the natural tendency to be mutually supportive. In global terms we may expect that the serialistic students will benefit more from concept mapping, while the holistic students get a typical compensation from the woven stories approach.

The clear distinction between stories and concept maps is the episodic nature of stories. A node in a concept map is often related only to its neighbouring concepts. One might say that a concept map represents only the local information of a particular concept. On the contrary stories typically represent episodes that consists of a several sub-sequent story sections connected by links. The analysis of episodes could be useful to make assumptions about the students' reasoning during the learning process. Moreover, information on different *episode patterns* could be stored by the StoryML specification, and actual story lines corresponding to a given pattern could be retrieved on demand. An example of episode patterns is a sequence A-B-A which represents a three step dialog, between students A and B.

3 Definition of StoryML

In a general sense, metadata can be called as a "data about data". The term "metadata" itself refers to background information about something[4]. Metadata can be associated to data which stores descriptive information of learning objects existing in a particular learning space. In many cases, metadata is dependent on the usage of the learning object. For example in the case of adaptive learning materials there can be metadata elements which store the data used in the adaptation or personalization process.

The reasons for creating and using metadata in general situations is to improve the possibilities of retrieval as well as to support control and management of the described resources. To make the creation process of metadata elements easier and to achieve interoperability between systems from diverse origins, several metadata standards have been developed. To describe the essential characteristics of a particular woven story we need a method to store the metadata for the attributes of such data entities. In the context of StoryML the metadata elements describe the characteristics and background information of stories.

StoryML draws elements from issues related to collaborative writing and collaborative concept mapping. One of the key ideas in woven stories and StoryML is the combination of collaboratively writing and linking a set of story sections. Collaboratively written text needs a special storing format, as any part of it may include contributions from many authors, and these must be separable in order to indicate different contributions. Collaborative linking of a set of story sections needs a visualization that shows clearly the structure of relations between them. In woven stories this is done by representing the collection of sections as a graph that is stored in the StoryML. The visual objects in the graph are linked to the story sections; in fact, the graph visualization represents the entire woven story. Furthermore the graph representation needs a special storing format to indicate contributing authors. The StoryML core element set is a stand-alone representation of a woven story at a given time. We present a summary of the core element set in Subsection 3.1.

3.1 The StoryML Core Element Set

The StoryML core element set is divided into five main categories represented by separate elements, which are summarized in Table 1.

Table 1. The main categories of the StoryML core element set.

Element	Description
General	General information about the woven story.
Content	The content (story sections) of the woven story written by an arbitrary number of authors.
Association	Contains the associations or links included in the woven story.
Visualization	Contains the representation of the woven story as a graph.
Author	Data about the authors who have contributed to the woven story.

The *General element* and its subelements store general information about a particular woven story, such as the name and intention of it. A woven story has a name that is used to refer to it as a whole. As a woven story can be used to serve different goals, an intention statement is included in the General element to let the authors decide and document a common goal. In addition to the intention the General element includes a description of the woven story which summarizes the content of the woven story in a free form. The description can for example be an invitation or an announcement to new writers, or simply an abstract of the content. The General element has a subelement, Creator, which stores information about the software tool used to implement the woven story. The General element includes a unique identifier, by which possible extensions are linked to the woven story.

The *Content element* stores the data of the set of story sections belonging to a woven story. The sections may include text, links and images and they may be written by an arbitrary number of authors. Thus the story sections are represented by a flexible structure which can be reformed to reflect the potentially complex authorship of a single section.

The *Association element* describes the associations between the collection of story sections in a woven story. These associations are restricted in the core element set of the StoryML to include only the links among the story sections and links to external content (not for example version history, which can be viewed as an association between versions too). Each of the associations may be described by a free form description.

The structure of a woven story is visualized by a graph. The *Visualization element* stores information that describes the visual representation that the authors have decided to use for their woven story. The Visualization element includes subelements for nodes and arcs, which together form the basis of the graph representation. The nodes represent separate story sections in the woven story, every one of the nodes representing a specific section. The only exception to this are external links, in which case a node represents external content located somewhere in the web. The arcs which connect nodes represent links (associations) between the story sections.

The *Author element* stores information about the authors who have participated in the writing of the woven story. This data includes personal data submitted by the authors themselves, such as their names, affiliations and interests. Two important information fragments stored within this element are the login name and color of each user. Both of these are required to indicate authorship in the woven story. The story sections may include a list of authors, who have participated in the writing of it, and their contributions in the actual text may be colored according to their allotted colors.

3.2 Data Management and StoryML

It is unrealistic to think that the user of the woven stories environment stores or even has to store all the necessary data to the StoryML elements. There are three different ways to collect the data in the environment. First, the data can be explicitly stored by the user or group of users. Secondly, the environment using StoryML can automatically or semi-automatically create some of the elements dynamically according to the content and relationships between story sections. Thirdly, data can be retrieved from the available user model. The elements that can be generated automatically by the application are, for example, the author of the node, the time of the contribution or the ID-number of the story node.

The user model of the authors can be used to get the data determining the visual appearance of a particular story node. In this way the system can adapt the representation of the story according to the specific preferences of the user. One of the most intuitive visual attributes would be the color of the nodes and arcs used to highlight a part of a woven story. The user model can be also dynamic, so that it can change according to the behavior of the user.

From the computational point of view, the methods of storing the metadata of a particular story section is quite easy to implement. The difficult part emerges when we are dealing with complex relations between the story nodes. Of course the students can add the relational information by themselves, but it is difficult to imagine that all the students can relate their ideas to the writings of other students or even to their own writings. Besides, it is basically impossible for an author to figure out complex relations between his own text and the contributions of the others.

A rich, ontologically sound relational information about the story nodes can help us to improve the functionality of a woven story environment. We can apply the relation information to make the environment adaptive to the needs of different users. Also the filtering of the woven story can be achieved using the relational information of the story.

A solution to the problem of creating relational or content dependent metadata elements is a content analysis of the story nodes. There are several heuristics to analyse the content of the text. For example the <!metaMarker> program is designed to provide an "information context" in the form of a rich set of metadata elements for a variety of time and resource intensive tasks [8]. <!metaMarker> has been originally developed to automatically organize customer service requests or incoming email streams according to their subject contents. It also automatically identifies such things as the emotional "tone" of the message and the intention or goal of the author of the message. The <!metaMarker> program uses Natural Language Processing (NLP) and Machine Learning (ML) for automatic metadata generation.

To adapt <!metaMarker> to extract metadata elements specific to educational material, the initial target elements from GEM and Dublin Core were categorized into three groups depending on how they will be extracted. Some elements such as "author" or "publisher" were directly extracted from the texts by applying educational material specific sublanguage grammar. Other elements such as "quality" or "relation", which are implicit in the texts, were derived through the discourse model analysis of the educational materials. This gives us an opportunity to apply such applications as <!metaMarker> also in the woven stories environment to create the relevant metadata elements of story nodes.

Of course there are still problems in the automatic text analysis. It is not entirely certain that we will get the needed information. Furthermore it is highly possible that in some cases we will get a false result from the automatic analyser. One solution to this is to make the creation of the metadata elements semi-automatic. In a semi-automated system a program can automatically create essential information or hints to the person responsible for creating the metadata of the text or in our case the metadata of a story section. The person can complement the automatic analysis according to the information received from the system. In this way we can combine the speed of automatic text processing with the opportunity of having a kind of backup check of the information received from the system.

In the context of the woven stories, the metadata creator can be the author of a story node, peer users in the system or the teacher/creator of the writing situation. In educational context the teacher could be able to add relative information of story nodes according to the information received from an automatic content analyser. This will help the student, because complex relations between story nodes are not trivial. An automatic system can help the teacher to help the student to relate his writing to other writings in many different ways.

4 Applications of the StoryML

The idea of woven stories is very useful in *distance education* settings. The students can use the system to collaboratively write stories or various types of documents. They can see the development of the story and the relations between different nodes in the story line. On the other hand the woven stories concept can be seen as a way to represent the learning process of the student or students. During the learning process of a particular subject there might be several interactions between different students or nodes in the story line. The learning process (path) can be analysed by examining various sections of the story.

An example of the educational use of woven stories is the case of a distributed learning task. Let us consider a situation where a group of students receive a learning goal that they must achieve collaboratively. In this situation the woven story application can support the students by presenting the individual efforts. Hence, the students can figure out if they have reached the given goal. For example, the parallel learning paths of the group members might be represented as alternative paths between two nodes A and B.

StoryML can be constructed to support the *analysis of the nature of discursion in stories.* For example during the assessment of collaborative work it would be useful to get information about students who have encouraged or opposed other students during the collaborative work. Those students who have been active and really contributing to the ongoing work could receive better marks for the assignment. This would give a teacher an opportunity to assess the students' participation and contribution to the learning activity processed by the woven stories system. In addition, the student could use the above information to self-evaluate his own contribution to the collaborative learning task. Maybe the information received from the application could help the student to make new discoveries of his own thinking.

The *time dimension* of stories is also important. It gives the students means to self-reflect their own or other students' paths during the learning process. In this way the students are able to make conclusion about the relations of story sections or the development of ideas.

The StoryML specification gives us *tools to define different relations between story sections*. The system can automatically or semi-automatically inform a student or a teacher about different factors in a particular woven story. For example a woven stories application using StoryML can notify the student about similar topics or the same kind of story sections in the hyperspace that are related directly to what he is writing or doing. On the other hand it might be useful to inform a student about totally different contents or opinions, so that the student might get some novel ideas from other students. Also, if the story nodes represent the learning process or path of particular student, the application could advice the student to take the next step in the learning path. For example, in the case where the student is stuck in his studies and he is making the same mistakes (circling in the learning space), the application could show him the right direction.

The StoryML specification supports the *implementation of adaptive functionality*. The user model maintained by an woven stories application makes it possible to present the hyperspace according to specific characteristics of the user. One example is to present the hyperspace differently to a heterogeneous population of users. Thus, an woven stories application can present different presentations of the same hyperspace by showing different types of story lines and relations according to several attributes present in the StoryML. The StoryML specification can act as a basis of the adaptation function.

The educational application of StoryML specification supports the means to adaptively offer particular learning materials to the learner. With the woven stories application the learner can choose from a repository of learning materials (e.g. from particular course or several courses) such objects that are relevant for his own learning goal or motivation. Hence, the StoryML specification gives us a tool to represent the individual *learning stories* of a particular learner. In this way the different learning styles and learning goals are represented in a form of various learning paths or stories with in the same learning space. The learner can for example browse through different kinds of learning paths and select the one that fits to his own learning style.

To develop this idea further, we can imagine that the learner makes his own contributions and associations during the learning process. With the woven stories application the learner can modify the existing learning materials (e.g. content and relationships) and hence enhance the repository with the modified, graph-structured representation of the contents. In this way the repository of learning material comes richer every time there is a novel way of understanding the knowledge and information presented through the learning material. Furthermore, the learning material becomes more adaptive and personalized.

5 Conclusion

The StoryML specification introduced in this paper serves as a starting point to the formalization that can be used to represent the hyperspace of woven stories in an educationally meaningful way. The continuing development of both the concept of woven stories and the metadata attached to a particular woven story is underway. The StoryML core set will develop further as we get more insight into the nature of the data stored in StoryML, and its actual utilization used by a student. Furthermore, we are enhancing our existing woven stories application, the WS2, to support the StoryML specification in order to exploit the possibilities offered by it. Knowledge and experience gathered during the development process of the application as well as feedback from its users will be essential in the further development of StoryML.

From the educational point of view the concept of woven stories empowered by StoryML supports flexible and adaptive ways of learning. This modification of woven stories supports the learner individually, not by forcing or guiding her to behave in a predefined way.

References

1. Dublin Core Metadata Initiative. Dublin Core metadata element set, version 1.1: reference description, 1999. Available at: http://dublincore.org/documents/1999/07/02/dces/
2. P. Gerdt, P. Kommers, C. Looi, E. Sutinen: Woven Stories as a Cognitive Tool. *Cognitive Technology 2001*, LNAI 2117, Springer-Verlag, 2001, 233 – 247.
3. Harviainen, T., Hassinen, M., Kommers, P., Sutinen, E.: Co-authoring Stories over the Internet. In *Advanced Research in Computers and Communications in Education* (Eds. Cumming, G., Okamoto, T., Gomez, L.), IOS Press, Amsterdam, 1999.
4. Heery, R.: Review of metadata formats. *Program*, 30(4):345-373, 1996.
5. Huai, H.; *Cognitive Style and Memory Capacity: Effects of Concept Mapping as a Learning Method*. Doctoral Thesis, University of Twente, 2000.
6. Huai, H., Kommers, P.: Concept Mapping as a Learning Strategy for Autonomous Students with a Serialistic Cognitive Style. In *International Journal of Continuing Engineering Education and Life-Long Learning*, 11 (1/2), 135 – 151, 2001.
7. The Ontology Browser. Internet WWW-site, url: http://ontology.teknowledge.com:8080/rsigma/SKB.jsp (December 12, 2001).
8. Paik, W. Automatic Generation of Eduactional Metadata Elements to Enable Digital Libraries. *Proceedings of the 9th Internaltional Conference on Computers in Education* (ICCE/SchoolNet2001), Seoul, Korea, November 12-15, 2001, 1406 – 1413.
9. Standard Upper Ontology (SUO) Working Group, IEEE P1600.1, Internet WWW-site, url: http://suo.ieee.org/, 2002 (January 31, 2002).