Samp*L*e: Towards a Framework for System-supported Multimedia Authoring

Kateryna Falkovych, Frank Nack, Jacco van Ossenbruggen, Lloyd Rutledge
CWI, Amsterdam
Firstname.Lastname@cwi.nl

Abstract

Much current research on hypermedia generation accepts user input only at the start of an otherwise fully-automated process. However, since multimedia presentation creation is often a complex and creative process, it has multiple phases which would each benefit from human intervention. This paper presents a hypermedia generation model that lets the user influence all phases of this computer-assisted human-guided process. The main focus is on providing extra support for helping the user find relevant media items and combine them meaningfully into a rich and coherent multimedia presentation. Like fully-automated systems, our approach uses explicit knowledge about the presentation's topic domain, narrative structures, hypermedia presentation and distinctions between media modalities. This paper presents a motivating scenario that is used to derive a number of system requirements and to discusses the pros and cons of the presented approach.

1. Introduction

The production of multimedia presentations is a complex, resource demanding, distributed, and creative process. The aim of the process is to provide engaging and relevant information by composing a multi-dimensional network of relationships between different kinds of audio-visual information units.

We divided this process into five stages: theme identification, specification of presentation structure, collection of material, arrangement of material, and presentation creation. At the first stage, a topic of the presentation is chosen within a domain of interest. The scope of the presentation is outlined at this stage by defining the genre of the presentation as well as main and secondary characters. This information allows specification of the logical structure of the presentation at the second stage. At the third stage appropriate material is selected and placed within each part of the logical structure. The particular ordering of the selected material is done at the fourth stage. At the fifth stage the final presentation is created by determining spatial-temporal relationships between selected items and defining stylistic aspects.

Even though we presented the stages above in a sequential manner, we find that the various stages are typically iterative and mutually interrelating. While this increases the complexity of the process it also leads to semantically rich and engaging presentations.

Many multimedia presentations, such as those made by pupils, students, researchers or managers, are manually crafted presentations. The support in manual production comes mainly in form of production environments such as Director, Premiere, Photoshop, Flash, FrontPage, PowerPoint and others. These tools are, however, not equipped to support the complex processes of content and design development, as they assume that the user has a sufficient level of expertise [4].

In this paper, we present a framework that supports the five steps of the presentation generation process. Our approach combines the creative strength of humans with the analytical and procedural power of machines. It allows a way of presentation generation in which the user has full control over the presentation creation process, but at the same time is facilitated with ontology-based

and context-oriented information at those stages where she lacks knowledge or skills.

Our system, called Samp*L*e (Semi-Automatic Multimedia Presentation generation Environment), is connected to a large multimedia database. In order to create a presentation, the material from the database is used.

Since different users have different levels of expertise in the domain as well as various experiences with presentation building process, the system should be able to support any type of a user by providing support at each stage of presentation creation process. Depending on the user's level of expertise, created presentations can be included into the system repository, enriching it with new points of view on the material and new interrelation structures between media items. In such a way, stored presentations might help novice users (such as students having an assignment to build a presentation) get an insight into the domain and get an idea about possible topics and designs for their own presentations. Additionally, the system is also in the position to offer assistance on the presentational level by providing various ways of structuring a presentation.

The paper is structured as follows. Section 2 discusses related work in which we show connections between our approach and existing presentation generation techniques. Further discussion follows the five stage process described in the Introduction. Section 3 presents the theme identification stage. Section 4 discusses presentation structure specification. Section 5 describes the elaboration on the created structure by facilitating content selection. Arrangement of selected media items is presented in Section 6. Section 7 discusses the last stage of the process where the presentation is created according to the proper style and duration. Finally Section 8 outlines conclusions and future work.

2. Related work

The development of the Samp*L*e framework took account of existing work in two related research domains, namely authoring environments for the fully manual generation of presentations and techniques from the field of automatic presentation generation.

The usual computer support for manually crafted presentations is manifested in form of production environments, such as image editing tools (Photoshop, Illustrator, GIMP, or Maya), new media authoring tools, such as Director/Shockwave, Flash, Dreamweaver, Frontpage, PowerPoint, GRiNS, Authorware, HyperCard, or WWW presentation technology, such as HTML and SMIL. Although these tools provide the user with much freedom during the development process, they are not equipped to support the complex process of content development, because they assume that the user has a sufficient level of expertise.

Systems that facilitate automated presentation generation, on the other hand, are typically applied for dynamic, interactive environments that do not allow the intervention of a human during presentation generation at any time after the initial request, such as web sites of museums [8], real time instruction generation [1], discourse driven hypermedia presentations [9] or user tailored biographies of artists [6]. These systems provide ontology-based description layers for content and presentation aspects that permit reliable control mechanisms to establish presentations that are flexible enough to respond to changing individual user needs and user groups. The complexity of logical structure developments and the complex choice of stylistic aspects regarding conveying certain information typically result, however, in quite simple structures and styles and hence unexciting presentations.

The blend of these two methodologies for multimedia presentation generation can lead to an improvement of mainly human generated presentations.

First attempts to make use of description structures and mechanisms taken from automatic processes as the basis for support in manual authoring environments are described on the content level, such as for the authoring of motion pictures [2], or on a task-oriented level, such as in supporting the early exploration of design ideas [3]. The advantage of this system-guided approach is that it facilitates the creation of attractive presentations by

utilizing user abilities to create rich and diverse discourse structures and presentation styles by offering support on both content and task levels at any time if required.

Our approach, which we describe in the rest of the paper, uses similar strategies. The distinction from established work is that we address all stages of the production process where the cited work concentrates on particular phases or tasks only.

In the next sections we discuss each stage of the presentation production process outlined in the introduction and consider the user actions and system support involved. The stages are described based on a scenario in which a student needs to make a presentation about the Dutch art movement 'De Stijl'. The scenario is suited to exemplify the requirements needed for system-guided presentation generation.

3. Theme identification stage

The student is interested in the Dutch art movement called 'De Stijl' , and wants to make a presentation that conveys the interdisciplinary structure of this art movement in an engaging way. After starting the system, the student defines ' De Stijl' as the main topic of the presentation and chooses essay *as the genre for the presentation. Next, the user defines ' De Stijl' also to play the role of the main character within the narrative associated with the chosen genre. Since she is interested in the relationships between 'De Stijl' and Cubism, she defines Cubism to be the secondary character.*

In the scenario above, genre is understood as a kind of narrative which identifies the emphasis of the work. For example, the biography genre denotes a narrative that describes a person' s life (main character) including relationships with other people (related characters). Essay designates an analytic composition dealing with its subject from a personal point of view.

Requirements The system should thus support different genres and be able to provide template storylines and roles that are relevant for these genres. The system could, for example, support genres such as biography, monograph, essay, analysis.

In the next section we describe how the choice of genre and roles influences the logical structure of the presentation proposed by the system and how the user can personalize and adapt this structure.

4. Specification of presentation structure

Based on the selection made in the previous stage, the system selects an appropriate genre-specific presentation template defining the default presentation structure. This template is used to generate a first draft of the presentation structure. Since the user defined 'De Stijl' and Cubism as roles in the narrative, the system will try to introduce these concepts in the first part of the presentation. It will generate a relatively detailed introduction about "De Stijl" using the overview *narrative structure, and only a short* definition *of Cubism since it plays a secondary role. For the middle part, it will look for relations between these two concepts. The system finds that Cubism had influences on 'De Stijl' . The influences of Cubism on ' De Stijl' will be used as a basis for the relationship between the two roles in the narrative. It concludes with a summary of later art movements that were influenced by 'De Stijl' . The user wants to alter this default presentation structure. She decides that she prefers two art movements to be equally presented in the introduction. For that she has to let the system know that Cubism should also be introduced using the* overview *narrative structure. After the alteration for the whole presentation structure is complete the user proceeds with the next step.*

The system repository contains template genre-specific presentation structures. A presentation structure consists of conceptual parts arranged within a top-level structure, such as Prologue, Elaboration, and Epilog. The common way of developing the narrative of essay is to start with introducing the main character and related secondary characters (Prologue), then to elaborate on the main character' s major achievements including the roles of the secondary characters (Elaboration), and conclude with outlining the significance of the main character' s achievements and their influences on future developments (Epilog). The content of conceptual parts

within each top-level division depends on the user choice of characters.

The adaptation of presentation structure consists either of the extension or reduction of the number of conceptual parts within each element of the top-level structure, or in changing descriptions of the parts and thus their conceptual meaning. The conceptual meaning is considered here consisting of two parts: information about the content of a media item and information about the narrative structure of it. The narrative structure is understood as the type of narrative a media item belongs to, such as description, overview, summary. For instance, the description "Overview of ' De Stijl' " denotes that this conceptual part of the presentation should contain information about ' De Stijl' (content) which is expressed as narrative structure overview.

Changes in description of sections result, as we will see in Section 5, in a different content selection behaviour of the system. Important is that this mechanism also allows the enrichment of structures for a genre by storing the new template genre structure in the database.

Requirements Domain knowledge is required to adjust the frame presentation structure of genres according to the chosen characters and also to derive relationships such as that Cubism started before and influenced ' De Stijl' .

Next, the system should be able to reflect changes the user makes to the proposed presentation structure (as changing the section description). Thus, all concepts that the user operated upon during the specification of the presentation structure should be related to the domain ontology (to let the system know the content of the media items) and to the narrative structure ontology (to deduce the preferred type of narrative the user wants to have for the certain content as overview or definition).

In the next section we will outline how the established structure facilitates the ability of the system to guide the content collection process.

5. Collection of the material

In the next phase, the user wants to fill in the logical structure of her presentation with actual media items by, for example, using the interface shown in Figure 1. She uses the check-boxes on the top of the screen to select preferred media types. She clicks on a description of a specific part of the presentation, and the system will present her with relevant media items collected from the database, with a top ranking for those items that will best fit in the particular narrative context of that part of the presentation structure.

The system represents the search results as icons, distinguishable by media type, where the content of each icon can be viewed by clicking on it. Icons are connected with links, representing conceptual relations between corresponding media items. In this way, our user sees that the textual information object containing information about principles of 'De Stijl' and the image representing a drawing where those principles are reflected are conceptually both about 'De Stijl' principles.

Based on these suggestions, our user fills in all the parts of the presentation with the media items she likes best. She drags and drops selected media items to the right bottom panel where they will appear together with their relations to other objects. This process repeats until all the sections of the presentation structure are filled with selected material. Every time she clicks on the section description, chosen information objects will be visible on the right bottom panel. If the user is not satisfied with the selection set, a new search can be initiated for each section and a new selection can be made.

The section description, which basically specifies content requirements, and the position of the section in the overall narrative structure (e.g. the De Stijl overview is inside the Prologue section), which puts some restrictions on the narrative structure of media items, both form the query for the retrieval process. The advantage of this approach is that restrictions of the search include not only searching for media items

according to their content but also implicit elimination of them with regard to their narrative structure (for example, due to the fact that Prologue usually contains abridged narrative structures, the choice should be for media items which contain summarized or short description about the subject of interest). Thus, smaller and more relevant sets of retrieved objects are shown to the user. The pre-selection of information objects based on their narrative structure can be cancelled by switching off the option "Context search" on the top panel. Then all the items with

and also introduces the need for different kinds of annotations of media items. In the previous section we have already mentioned the roles of the domain and narrative structure ontology for the development of presentation structures. We also described the importance of concepts in the description of each part of the presentation for defining the potential content and as a consequence the retrieval of it. Now we outline a complete framework for meta-data.

Each media item has to be annotated with the

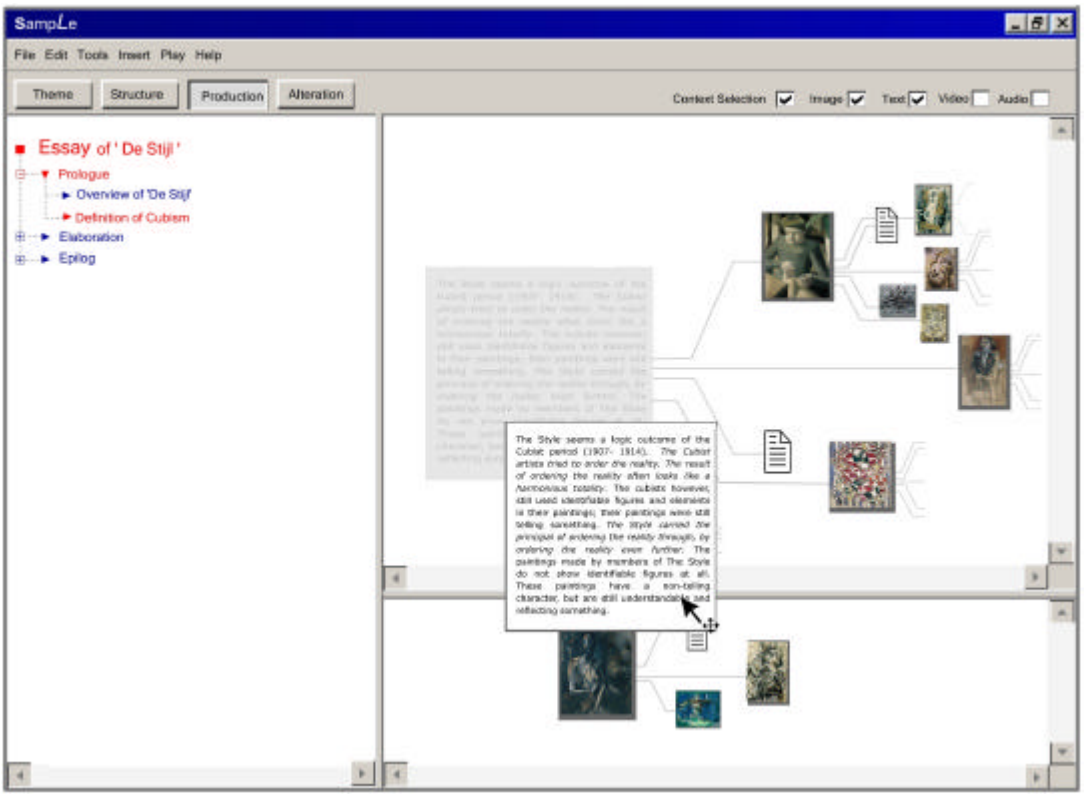

Figure 1. Selection of the material

relevant content will be retrieved without any specific preferences on their narrative structures.

At this stage selected media items are not ordered with the respect to the final presentation layout, but rather located within the presentation structure (they are in the right place but not necessarily in the right order).

Requirements The retrieval process described above determines new requirements for ontological structures

concepts from the domain ontology allowing the system to identify its content. Moreover, a media item should be annotated with concepts from the narrative structure ontology to provide the information for Samp*L*e to retrieve for the particular context of the current presentation step. In addition, since the system provides the option of choosing between preferred media types of the objects to be retrieved, we require items to be

annotated with concepts from the media ontology, which specifies external (e.g. painting) and internal (e.g. image) representations of objects and their format (e.g. jpg) The structure of media ontology was inspired by the IMAT project [5].

On the base of this framework for annotations relationships between objects can be explained. In order to introduce relationships (links) between media items in a visual way, the system must be able to derive them. As mentioned above, annotations contain descriptions about content, narrative structures and media types of items. All the annotations are written and stored using RDF [7] format. An example annotation of the textual item discussing principles of ' De Stijl' could look as presented below:

```
<rdf:Description about="d:item132">
    <c:principle name="P1"/>
    <c:art_movement name="De Stijl" />
    <s:structure>summary</s:structure>
    <m:media_type>
        <m:internal>text</m:internal>
        <m:external>text</m:external>
        <m:format>txt</m:format>
    </m:media_type>
</rdf:Description>
```

where ' c' stands for the namespace of the domain ontology, 's' for the narrative structure ontology, and ' m' for the media ontology.

The annotations for the image reflecting ' De Stijl' principles contain:

```
<rdf:Description about="d:item122">
    <c:principle name="P1"/>
    <c:painting style="De Stijl"/>
    <s:structure>summary</s:structure>
    <m:media_type>
        <m:internal>image</m:internal>
        <m:external>drawing</m:external>
        <m:format>jpg</m:format>
    </m:media_type>
</rdf:Description>
```

The fact that the attribute *name* of the class *Principle* has the same value (*P1*) for both media items determines the relationship between these objects and forms the basis for the visualization on the screen.

Having explained the selection of media items for the intended content of the presentation, the next section discusses arrangement of media items within each section to form a coherent presentation.

6. Arrangement of the material

In this step our user orders the material selected into a coherent presentation, and again, she is supported by the system. For example, our user selected 3 texts for the Prologue. *One object contains a general overview about 'De Stijl' (according to the content annotations), another one presents the principles of this movement, and finally the third one gives the overview of Cubism. The system suggests putting the Cubism text at the end of the Prologue, because of its secondary role in the narrative. Additionally, it suggest putting the overview before the text about the movement's principles because in general, it is better to put general information before specific information. Our user is content with the system's suggestion for the order of the* Prologue *and other parts and decides to make no further adaptations.*

In order to be able to arrange media items into a meaningful structure the system should be equipped with rules making the arrangement possible. These rules need to take into account not only the content and narrative structure of the media items but also the presentation structure the media items need to fit in. For example, the rule that puts generic before specific may only apply in the Prologue, but not in the other sections.

Requirements The system needs rules that can suggest a coherent ordering of material based on the media items' meta-data and the presentation structure.

In the next section the created presentation has to be conveyed in the right form with regard to the presentation type, style, and duration.

7. Presentation creation

Once the theme, structure, and general arrangement have been determined, this final phase generates the presentation itself based on the output of these previous phases. The user is confronted with the choice of different

presentation types (slide-show, interactive or non-interactive presentation), among which she chooses non-interactive presentation. The user defines the style of the presentation by specifying font types, colour schema, scenes layout, and animation effects. The final presentation is evaluated with respect to the required duration. If the duration exceeds the desired one, the user might reconsider the selection of items for different sections of the presentation in order to reduce their amount.

To realize this process in our system, we build on techniques from automatic presentation generation engines. The Cuypers system is one example of how presentations can be automatically generated from the abstract presentation settings together with user top-level requirements about the intended content of the presentation [8]. In the context of Samp*L*e, the content of the presentation is explicitly specified during previous four stages of the cooperative human-computer process and at this stage is convey according to the user preferences.

Requirements A main principle for all of Samp*L*e's stages is the ability of the user to directly modify the output from one stage before it is input for transformation to the next. For the presentation creation stage, this means the user is given the presentation as a rough draft that she can polish up. The user can use the tailor-made editor integrated in the system for editing that has to be appropriately designed.

8. Conclusions

In this paper we focused on solutions for the problem of an authoring environment that supports five stages of the presentation generation process. Our approach supports a way of presentation generation where the user has complete control over the presentation creation process, but at the same time is facilitated with ontology-based and context-oriented information. We described the underlying framework of our approach and discussed the interrelationships between different types of meta-data.

Future work will concentrate on the realization of internal processes for the specified framework. The proposed meta-data requirements will be verified according to their completeness.

The possibility to enrich the system repository by storing successful presentations has to be integrated into the system. For that the way should be found to manage with new presentation structures in the relation with old ones.

Building the presentation at the last stage of the process includes combining various media items into scenes. By doing that new relations between items can be discovered (e.g. a subsection of the selected textual item describes a part of the image). The challenge of creating new annotations will be addressed. Annotations used by the system fall into complicated infrastructure a user cannot be faced with. Thus, a semi-automatic way of creating new annotations has to be found, where user intentions are understood by the system, and complete relation infrastructure is filled in without the user intervention.

Acknowledgements

Part of the research described here was funded by the Dutch national ToKeN2000/CHIME project.

References

[1] Andre, E., Muller, J., and Rist, T. (1996). WIP/PPP: Knowledge-Based Methods for Fully Automated Multimedia Authoring. London, UK, 1996.

[2] Baecker, R., Rosenthal, A.J., Friedlander, N., Smith, E. and Cohen, A. (1996). A Multimedia System for Authoring Motion Pictures. In Proceedings of the 4th ACM International Conference on Multimedia, pp. 31- 42, Boston, USA, Nov. 18 - 22, 1996.

[3] Bailey, B. P., Konstan, J.A., and Carlis, J.V. (2001). DEMAINS: Designing Multimedia Applications with Interactive Storyboards. In Proceedings of the 9th ACM International Conference on Multimedia, pp. 241 -250, Ottawa, Canada, Sept. 30 - Oct. 5, 2001.

[4] Candy, L., and Hori, K. (2003) The Digital Muses: HCI In Support of Creativity. In: Interactions: New Visions of

Human-Computer Interaction 10(4), pp. 44-54, July/August 2003.

[5] Kabel, S.C., Wielinga, B.J., de Hoog, R. (1999) Ontologies for Indexing Technical Manuals for Instruction. AI-ED (Artificial Intelligence in Education) conference 1999, Le Mans, France, Proceedings workshop on ontologies for intelligent educational systems, pp. 44-53.

[6] Kim, S., Alani, H., Hall, W., Lewis, P.H., Millard, D.E., Shadbolt, N.R., and Weal, M.J. (2002), Artequakt: Generating Tailored Biographies with Automatically Annotated Fragments from the Web . Available at http://eprints.aktors.org/archive/00000068/

[7] Lassila, O., and Swick, R.R. (eds.) Resource Description Framework (RDF) Model and Syntax Specification. World Wide Web Consortium (W3C) Recommendation, February 22nd, 1999.

[8] van Ossenbruggen, J., Geurts, J., Hardman, L., and Rutledge, L. (2003). Towards a Formatting Vocabulary for Time-based Hypermedia. In: The Twelfth International World Wide Web Conference, Budapest, Hungary pp. 384-393, ACM Press, May 20-24, 2003

[9] Rutledge, L., Alberink, M., Brussee, R., Pokraev, S., van Dieten, W., and Veenstra, M. (2003). Finding the Story --- Broader Applicability of Semantics and Discourse for Hypermedia Generation. In: Proceedings of the 14th ACM conference on Hypertext and Hypermedia, Nottingham, UK August 26-30, 2003.