Semantics-based Dynamic Hypermedia Adaptation using the Hidden Markov Model

Jayan C Kurian, Payam M. Barnaghi, Michael Ian Hartley
School of Computer Science and Information Technology, The University of Nottingham (Malaysia Campus),
Jalan Broga, 43500 Semenyih, Selangor Darul Ehsan, Malaysia.
eyx5jkc@nottingham.edu.my

Abstract - Information collection, selection, structuring, and presentation design are the core considerations for general hypermedia presentation generation systems. The content collection process can be enhanced by retrieving semantically related information objects, relevant to the topic selected by an author. Once relevant information objects are available, the content selection process suggests semantically related resources for the author's selection based on data usage history. The information objects are represented by media assets and descriptive documents. The semantic web technology that allows resource interoperability can be used for content description and interpretation of these information objects. By utilizing a semi-automatic approach, authors can be assisted at different stages of the presentation generation process. In this research, adaptation constraints are established independent of the author's proficiency (i.e. novice, intermediate or expert) by applying the Hidden Markov Model methodology. Semantically related media objects are suggested to authors for selection based on their interactive behaviour and the strength of semantic relations. The paper describes an application of the Hidden Markov Model in the initial authoring phases of semi-automatic hypermedia presentation systems.

Keywords: Hypermedia Presentation Generation, Semantic Web, Semi-Automatic Authoring, Hidden Markov Model, Adaptation.

1. Introduction

Multimedia authoring provides an effective way to communicate the goals of a presentation coherently. Adaptation is one of the features in user modeling that will support customization of semantic search strategies based on the author requirements. The approach towards semi-automatic authoring in a semantic environment will facilitate different types of authors in getting through the presentation generation process starting from initial exploration of a domain to the final presentation. For semantic-based authoring, semantic web document representation standards [1],[2] are used to represent media assets in a machine accessible form. In this context, relations between domain conceptual structures are explicitly defined by an ontology that makes contents of multimedia objects accessible through a rich metadata model.

The functionality of hypermedia systems can be enhanced by making it personalized or adaptable [3] to the author's requirement. In our proposed system, we describe adaptation (i.e. content suggestion) in the content selection context, once the author identifies her/his topic of interest. For adaptation, we employ the Hidden Markov Model [4], a statistical model that can determine hidden states from observable parameters. It is our belief that the proposed adaptation model supports authors' for generating adaptable hypermedia presentations.

The paper is organized as follows. The next section describes multimedia systems for presentation generation and section 3 describes knowledge representation and the data model. The system architecture is described in section 4. Section 5 introduces the Hidden Markov Model for semantic content suggestion and section 6 concludes the paper and discusses future work.

2. Multimedia Presentation Generation

We investigate and describe systems that employ adaptation, temporal constraints, and automated presentation design in multimedia presentations. Then, we briefly introduce several authoring systems that use semantic web technology as means for generating presentation contents with emphasis on the semi-automatic authoring process.

The various interaction styles used for hypermedia navigation and an adaptive web interface that generates semantically related multimodal output are described by Taib et al. [5]. The multimodalities (i.e. written text, graphics, and speech) are classified by output modality classification methodologies. The users are classified into predefined profiles (e.g. text profile or multimedia profile) depending on their interactive behaviour while progressing through the authoring process. Predefined presentation templates are used for output generation that adapts progressively according to the interaction styles of users. Thus, the paper describes an approach to classify users into interaction style profiles and highlights the need for adaptation in hypermedia systems.

The work of Dalal et al. [6] introduces a knowledge-based system that generates customized temporal multimedia presentations. The order and duration of information objects are represented using temporal constraints and is achieved using negotiation process at run-time. The adaptation process employed here is tailored for different caregivers in a medical domain. An instance hierarchy creates the knowledge representation structure for each patient using domain and concept hierarchies. The presentation structure represented as a directed acyclic graph exchanges the information among various system components and expresses the communicative goals to be adapted to different caregivers. The media coordinator makes a consistent and synchronized presentation by allowing media-specific components to access and update the presentation plan. The inconsistencies in presentation plan are resolved using a constraint solver. The paper introduces user modeling concepts and signifies the importance of temporal and spatial constraints in the context of synchronized multimedia presentation systems.

Andre et al. [7] describes an approach towards fully automated presentation design in the context of personalized multimedia presentation generation. The paper discusses prototype systems [e.g. WIP and PPP] that produce presentations based on a given set of presentation parameters and by considering temporal coordination of different media items. The structure of coherent media items are described using the generalization of speech act theory, and the rhetorical structure theory, for communication between multiple media parts of heterogeneous media objects. The presentation structure is generated by utilizing the knowledge-base components. The presentation strategies select the relevant content, and structure it for delivering through an appropriate medium for target consumers. The qualitative and quantitative constraints are taken into consideration for building up a temporal constraint network for presentation acts and the temporal coordination that facilitates presentation generation. From the paper, we realize the need for considering the presentation parameters, and the temporal constraints of media objects for an efficient adaptive system.

Topia [8] uses RDF multimedia repository of Rijksmuseum collection [9] and creates hypermedia presentations as a result of a query. The SemInf system [10] semi-automatically generates multimedia presentations by combining semantic inferencing with multimedia presentation tools. In this context, Dublin Core (DC) [11] metadata and SMIL [12] presentation formats are used in generating multimedia presentations. The Artequakt [13] project generates artist's biographies by applying semantic associations between different entities that represent the artist's personal and professional life. The aim of the DISC [14] system is to build a multimedia presentation about a certain topic by traversing a semantic graph that consists of the domain ontology of classes, instances, and relations between them, together with the media resources related to instances.

A hypermedia presentation generation system in a multi-facet environment is described in SampLe [15]. The system uses semantic web technologies to support authors during presentation generation process. The process is divided into four phases: topic identification, discourse structure building, media material collection, and production of final-form presentation. SampLe supports authors during every phase of the process, independent of a particular workflow. This is achieved using ontology-based and context oriented information, as well as semantic interrelationships between different types of meta-data.

3. Knowledge Representation and Data Model

For knowledge representation, we employ an ontology that defines a common vocabulary for machine-interpretable definitions of common concepts in a domain and relations among them. The domain ontology gives information related to domain concepts. The media ontology gives media specific information and the discourse ontology narratively structures presentation contents. Protégé [16], an ontology editor tool is used to develop the knowledge-base for the domain ontology. This is represented in RDF(S) [17]. A simplified version of the Neural Network domain ontology created for the proposed system is illustrated in Fig. 1.

A data model represents the basic guidelines for annotating various media items. The data model adopted from [18] has the components content schema, semantic schema, and media schema that describes multimedia objects proficiently. The content schema is represented by Dublin Core attributes (e.g. title, identifier), semantic schema is represented by Learning Objective Metadata (LOM) [19] attributes (e.g. language, level), and media schema is represented by MPEG-7 [20] attributes (e.g. media type, media URI).

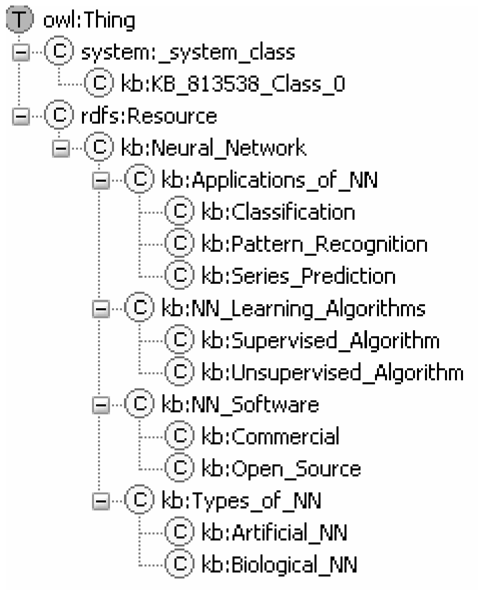

Fig.1. Neural Network domain ontology

The described data model has been chosen since the representation describes media contents, semantics and the media attributes effectively. The metadata attributes for media resources are generated manually. The data model allows the representation of external media objects already annotated with Dublin Core, Learning Objective Metadata, and MPEG-7 attributes. The ontology and the data model make the knowledge representation structure independent from data representation structure. In the data model, the metadata attribute “language” specifies the adaptation component for the presentations to be customized according to language specifications. The metadata attribute “level” specifies the content proficiency to provide customizable presentations. To describe the continuous media objects, MPEG-7 standard is chosen that represents the spatial and temporal aspects of multimedia objects. MPEG-7 Multimedia Description Schemes can effectively describe multimedia entities. MPEG-7 Visual Description Tools describes the visual features (e.g. color, motion). MPEG-7 Audio provides the standard for describing the audio contents (e.g. sound recognition). The data model enables us to represent content specific information (e.g. who, what), media specific information (e.g. size, height), and semantic specific information (e.g. when, how) of multimedia objects. Thus, the data model represents media dependent features of multimedia resources efficiently for enhanced data selection.

4. Adaptable Authoring System Architecture

In the proposed system, domain ontology is represented in RDF/XML, and the media objects are stored in a database or are coming from heterogeneous resources. Jena [21], an open source Java based RDF repository and reasoning engine, is used to query RDF/XML knowledge-base.

We adopt the extensive architecture proposed by Bunt et al. [22] that describes multimodal interaction and coordination for information presentation. The architecture identifies the functional and technical requirements for intelligent multimodal systems. The main components of the architecture are multimodal input, multimodal integration, and multimodal output. The multimodal input component caters to a mixture of input modalities (e.g. text, audio, and video). The multimodal integration component interacts with various modeling components for adaptation. The content management component interacts with the multimodal integration component to provide appropriate presentation content. The presentation is delivered by the multimodal output component that interacts with the application interface. The logical diagram of the presentation generation process is illustrated in Fig.2.

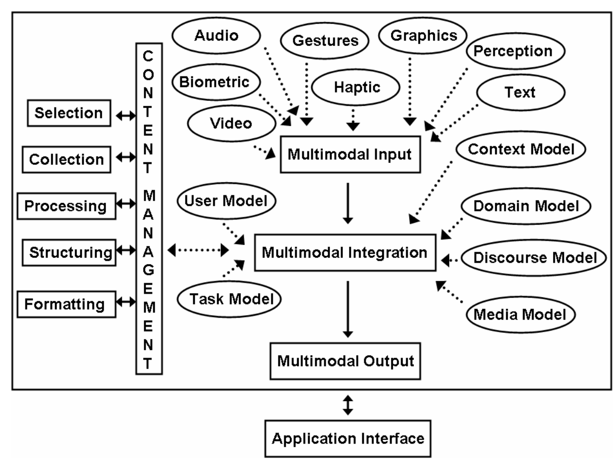

Fig.2. Logical diagram of the presentation generation process

The content management component has been designed with reference to the standard reference model [23] that describes an implementation-independent view of the processes required for the generation of intelligent multimedia presentations. The conceptual design of standard reference architecture modularizes multimedia presentation generation process into five layers: control layer, content layer, design layer, realization layer, and the presentation display layer.

In our methodology, the author selects a theme for the presentation (e.g. Lecture Notes) followed by a title (e.g. Introduction to Neural Network Architecture) that is supported by the discourse ontology. This is represented by the control layer of the standard reference architecture. The author selects relevant concepts by browsing through the domain ontology. Depending on author's selection, related concepts are suggested by the system based on data usage history. Once the contents are selected, corresponding media items are added and ordered. This is represented by the content layer of reference architecture. The design layer conveys the presentation layout structure for the media objects. The realization layer integrates displayable media objects with the layout information. The presentation display layer converts the realization layer representation into a hypermedia presentation that will be generated in the form of SMIL. The system architecture is illustrated [24] in Fig. 3.

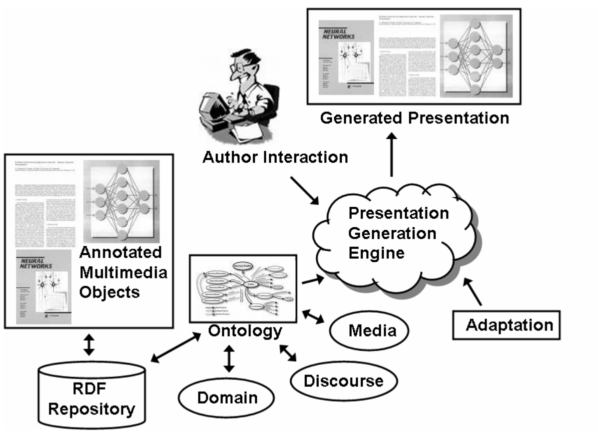

Fig.3. Logical diagram of the system

5. Semantic Content Suggestion using the Hidden Markov Model

A Markov process [4] is defined as a process which moves from state to state depending on the previous state of the process. Our objective is to suggest semantically related information objects to authors based on data usage history, without explicitly determining the author types. The Hidden Markov Model, which is an extension to the Markov Model, can make probabilistic assumptions of the hidden states based on observable states. Here the hidden states refer to the author types and the observable states refer to the author data usage history. Thus, the Hidden Markov Model can be used to examine and predict semantically related information objects.

In the proposed system, media items are annotated with the domain ontology concepts: classification (CS), pattern recognition (PR), and series prediction (SP). The Markov model takes into account the number of authors browsing from classification to pattern recognition or classification to series prediction or vice-versa. The closest probabilities between the concepts indicate that the annotated media items are strongly related. The farthest probabilities between the concepts indicate that the media items are vaguely related. Thus, the strong and weak semantic relations between media assets can be predicted for suggesting to potential authors. The author's proficiency (i.e. novice, intermediate, or expert) is the hidden state in this model and is not determined explicitly. The author's proficiency is not determined since average proficiency may vary depending on the cohort of authors and annotating media resources based on the author's proficiency limits the knowledge space for potential authors. In this context, the Markov Model holds its significance since the model predicts the probability based on the previous state and the probability for future events can be determined by extending the model to "n" previous states.

The browsing pattern of authors is given in Table 1 and the calculated transition probabilities are given in Table 2. Based on these values, the authors' browsing behaviour of concepts can be predicted by the Markov Model. If the calculated probability of authors' accessing the classification and series prediction concepts are near, it can be predicted that the classification and series prediction concepts are semantically related.

Table 1
Browsing pattern of the authors

Concepts	Authors	Browsing From			Browsing To			Authors
		CS	PR	SP	CS	PR	SP	
CS	200	0	35	25	0	20	20	220
PR	500	20	0	20	35	0	15	490
SP	300	20	15	0	25	20	0	290

Table 2
Transition probabilities

	CS	PR	SP
CS	160/200 = 0.80	20/200=0.10	20/200=0.10
PR	35/500=0.07	450/500=0.90	15/500=0.03
SP	25/300=0.083	20/300=0.067	255/300= 0.85

Using the Markov Model, authors' browsing behaviour of concepts is predicted as 0.234 for classification, 0.483 for pattern recognition, and 0.283 for series prediction. The concept pattern recognition has the highest probability of selection followed by the concepts series prediction, and classification. The nearest probability between the concepts classification and series prediction implies that they have a strong semantic relation. Moreover, the concepts classification and pattern recognition implies a weak semantic relation since their probabilities are far-off. The subsequent predictions based on previous probability results give the probability values as 0.245 for classification, 0.477 for pattern recognition, and 0.278 for series prediction. In this case, it is evident that the probability difference between the concepts classification and series prediction has reduced. This signifies that the strength of semantic relation has increased. Thus, the authoring system suggests semantic contents related with the author's preferred topic. Moreover, based on predicted semantic relations, additional media items annotated with related concepts can be supplied to the multimedia repository for generating resourceful hypermedia presentations. To make the system effective at the initialization phase, semantic search strategies have to be incorporated for supporting the authors' in content selection since the system suggests concepts based on data usage history.

6. Conclusion

This paper describes an ongoing research to develop a semi-automatic hypermedia authoring system. The research concentrates on the content collection and content selection phases of authoring, and suggests semantically related information objects to potential authors. The presentation generation process describes the adaptation component in the context of user modeling for hypermedia presentations. The Hidden Markov Model is employed for semantic content suggestion based on data usage history. The architecture of the proposed authoring system complies with the multimodal interactive information presentation architecture and the standard reference architecture, which describes extensively the fundamental authoring stages of intelligent multimedia presentations.

Future work concentrates on a prototype implementation to evaluate the Hidden Markov Methodology based on user trails. Once the adaptation component is realized, an integrated system can generate hypermedia presentations that would adapt dynamically to authors' proficiencies.

References

[1] The Resource Description Framework (RDF), available at: <http://www.w3.org/RDF>

[2] The Web Ontology Language (OWL), available at: http://www.w3.org/2004/OWL/

[3] P. Brusilovsky, "Methods and techniques of adaptive hypermedia", User Modeling and User-Adapted Interaction, 1996.

[4] K. Seymore, A. McCallum, and R. Rosenfeld, "Learning Hidden Markov Model Structure for Information Extraction", AAAI 99 Workshop on Machine Learning for Information Extraction, 1999.

[5] R. Taib, N. Ruiz, "Multimodal interaction styles for hypermedia adaptation", Proceedings of the 11th international conference on Intelligent user interfaces, 2006.

[6] M. Dalal, S. Feiner, K. McKeown, S. Pan, M. Zhou, T. Hollerer, J. Shaw, Y. Feng, and J. Fromer, "Negotiation for automated generation of temporal multimedia presentations", ACM Multimedia, 1996.

[7] E. Andre, J. Muller, and T. Rist, "WIP/PPP: Automatic generation of personalized multimedia presentations", In Proc. of the 4th ACM Int. Multimedia Conference (Multimedia'96), 1996.

[8] L. Rutledge, M. Alberink, R. Brussee, S. Pokraev, W. Van Dieten, and M. Veenstra, "Finding the Story — Broader Applicability of Semantics and Discourse for Hypermedia Generation", In Proceedings of the 14th ACM Conference on Hypertext and Hypermedia, pp. 67–76, 2003.

[9] Rijksmuseum Amsterdam, a museum of Dutch art and history, available at: <http://www.rijksmuseu m.nl>

[10] S. Little, J. Geurts, and J. Hunter, "Dynamic Generation of Intelligent Multimedia Presentations through Semantic Inferencing", In 6th European Conference on Research and Advanced Technology for Digital Libraries, pp. 158–189, 2002.

[11] Dublin Core Metadata Element Set Version1.1 (DC), Reference Description, Dublin Core Metadata Initiative, July 1999, available at: <http://dublincore.org/docum ents/1999/07/02/dces>

[12] Synchronized Multimedia Integration Language (SMIL), available at: <http://www.w3.org/TR/REC-smil>

[13] S. Kim, H. Alani, W. Hall, P.H. Lewis, D.E. Millard, N. Shadbolt, and M.J. Weal, "Artequakt: Generating Tailored Biographies with Automatically Annotated Fragments from the Web", Workshop on Semantic Authoring, Annotation & Knowledge Markup, 15 European Conf. on Artificial Intelligence (ECAI), pp.1-6, 2002.

[14] J. Geurts, S. Bocconi, J. van Ossenbruggen, and L. Hardman, "Towards Ontology-driven Discourse: From Semantic Graphs to Multimedia Presentations", In Second International Semantic Web Conference (ISWC2003), pp. 597–612, 2003.

[15] K. Falkovych, and S. Bocconi, "Creating a Semantic-based Discourse Model for Hypermedia Presentations:

(Un)discovered Problems", In Workshop on Narrative, Musical, Cinematic and Gaming Hyperstructure, 2005.
[16] Protégé, an ontology editor and knowledge-base framework, available at: <http://protege.stanford.edu>
[17] Resource Description Framework (RDF) Schema, available at: <http://www.w3.org/TR/rdf-schema>
[18] P.M. Barnaghi, and S.A. Kareem, "Ontology-Based Multimedia Presentation Generation", IEEE TENCON 2005 Conference, 2005.
[19] Learning Object Metadata (LOM), available at: <http://ltsc.ieee.org/wg12/20020612-Final-LOM-Draft.html>
[20] B.S. Manjunath, P. Salembier, and T. Sikora, "Introduction to MPEG-7: Multimedia Content Description Interface", John Wiley, 2002.
[21] Jena, a semantic web framework for Java, available at: <http://jena.sourceforge.net>
[22] H. Bunt, M. Kipp, M. Maybury and W. Wahlster, "Fusion and Coordination for Multimodal Interactive Information Presentation", In: O. Stock and M. Zancanaro (eds), Multimodal Intelligent Information Presentation, Springer, 2005.
[23] M.Bordegoni, G.Faconti, M.T.Maybury, T.Rist, S.Ruggieri, P.Trahanias, and M.Wilson, "A Standard Reference Model for Intelligent Multimedia Presentation Systems", Proceedings of IJCAI-97 Workshop on Intelligent Multimodal Systems, International Joint Conferences on Artificial Intelligence Inc., 1997.
[24] Web resources, available at: www.w3.org, gbn. glenbrook.k12.il.us, www.cdc.gov, www.inns.org, www.uni-duisburg.de.