Modeling Interactive Multimedia Presentation Systems Using Augmented Transition Networks

Shu-Ching Chen[1] Mei-Ling Shyu[2] Naphtali Rishe[1]

[1] High-Performance Database Research Center, School of Computer Science, Florida International University, Miami, FL 33199

[2] School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907

Abstract

In interactive multimedia presentations, users should have the flexibility to decide on various scenarios they want to see. This means that two-way communications should be captured by the conceptual model. An abstract semantic model, the augmented transition network (ATN), is proposed for modeling user interactions in multimedia presentations. In ATNs, each state node allows multiple outgoing arcs to model potential user interactions. At the decision point, the multimedia presentation system can use this information to display selection buttons so users can make their choices.

The superiority of modeling user interactions with ATN instead of the Timeline model or the Object Composition Petri Net model is discussed in this paper. Our results show that the ATN is effective for modeling user interactions in a multimedia presentation environment.

1 Introduction

Many conceptual models do not allow user interactions in multimedia presentations [1, 3, 8, 9, 11]. In these models, the presentation must follow the original design and there is no way for users to interact with the presentation. In other words, they do not allow the design and implementation of multiple choice user-directed presentations. Many applications such as multimedia document systems, digital libraries, and video games require this user interaction feature. Moreover, when document size increases, the complexity of the model increases. This complexity can make it very difficult for users to understand the representation of sequences when the model includes many media streams and a great deal of control information. As mentioned in [7], the most prevalent temporal model is the timeline [2], which aligns all events on a single time axis. Although this model provides a simple and graphical representation, it does not model user interaction features because it requires a total specification of all temporal relationships among media objects. In a video game, for example, users need to use the computer's mouse to select from various scenarios based on their decisions as the game proceeds. A single timeline does not work here because the start time of one selected scenario is known, but the end time can not be known before users make their choices. In order to solve this problem, different timelines are required to model different selection paths. Extra commands are also needed to explain the possible selection paths.

Little and Ghafoor [9] proposed an Object Composition Petri Net (OCPN) model, based on the logic of temporal intervals and Timed Petri Nets, to store, retrieve, and communicate between multimedia objects. OCPN is in the form of a *marked graph*, so each place in an OCPN has exactly one incoming arc and one outgoing arc. A *marked graph* can only model those systems whose control flow does not branch. This means that it can model parallel activities but not alternative activities [10] such as user interactions.

Since each state node in an augmented transition network (ATN) allows multiple outgoing arcs to model multiple selection paths, users can make selections based on their preference. Different choices lead to different presentation sequences. This allows user interactions to be modeled in ATNs [4, 5].

The rest of the paper is organized as follows. In Section 2, a Timeline model that allows user interactions

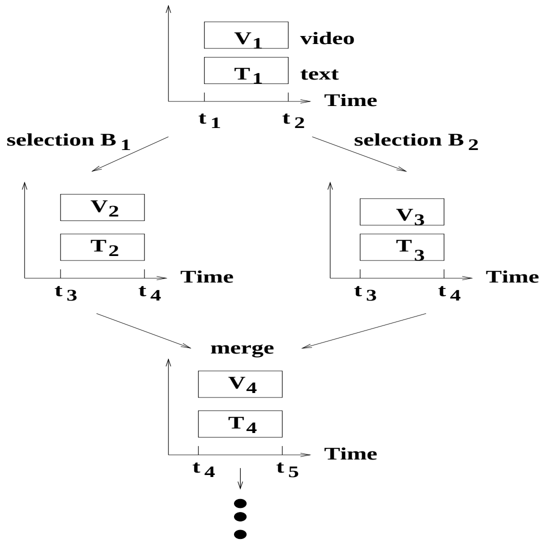

Figure 1: Timeline model with commands in user selections.

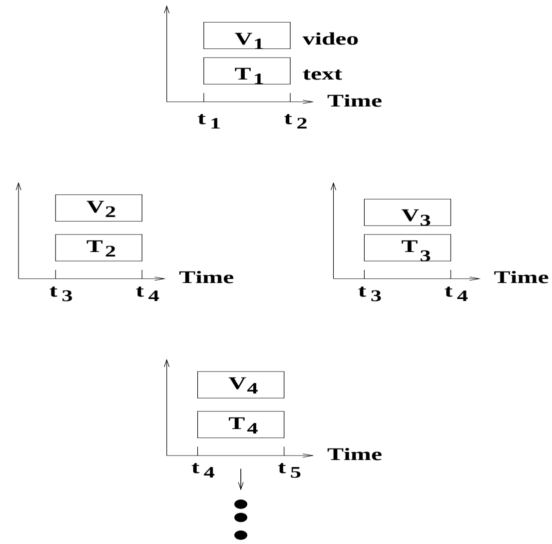

Figure 2: Timeline model with no commands in user selections.

to be modeled is introduced. The difficulties encountered in using the OCPN model to model user interactions is also presented. Section 3 discusses how user interactions are modeled by the proposed augmented transition network model. Conclusions are drawn in Section 4.

2 Related Work

In this section, the Timeline model and the Object Composition Petri Net (OCPN) model are discussed. We show how the Timeline model is used to model user interactions and why the OCPN model has difficulties with modeling user interactions.

2.1 Model User Interactions Using Timeline Models

The timeline model provides a simple and graphical representation. All media streams are aligned on a single time axis. Based on the representation of this model, users can easily see the temporal relations of each media stream in a presentation. The problems arise when we try to model user interactions with this model.

Figures 1 and 2 are two timelines that model user selections. In Figure 1, V_1 (Video 1) and T_1 (Text 1) are displayed beginning at time t_1 and ending at time t_2. The presentation then provides two selection buttons to let users make a selection. If B_1 is chosen, V_2 and T_2 are displayed from time t_3 through time t_4. If B_2 is chosen, V_3 and T_3, with the same duration as V_2 and T_2 are displayed. At time t_4, both selection paths merge and begin to display V_4 and T_4. This means that regardless of whether B_1 or B_2 is chosen, the presentation will share the same presentation sequence after time t_4. Figure 2 represents the same scenario as Figure 1 by using only a timeline with no commands. As Figure 2 shows, it is difficult to let users understand the selection sequences without additional commands and arcs.

2.2 Model User Interactions Using the OCPN Model

The Object Composition Petri Net (OCPN) model is based on the logic of temporal intervals and Timed Petri Nets. This model can be used to store, retrieve, and communicate between multimedia objects. OCPN is in the form of a *marked graph*, so that each place in an OCPN has exactly one incoming arc and one outgoing arc. Therefore, an OCPN can only model those systems whose control flow does not branch. Hence, it can model parallel activities but not alternatives such as user interactions. In applications such as computer-aided instruction (CAI), the situation depicted in Figure 1 can happen since CAI applications may involve extensive two-way communications and different selections may occasionally merge together. In these situations, the overhead inherent in the OCPN model makes it expensive for implementing a multimedia presentation system.

3 Model User Interactions Using Augmented Transition Network (ATN)

The augmented transition network (ATN), developed by Woods [12], has been used in natural language understanding systems and question answering systems for both text and speech. We use the ATN as a semantic model for multimedia presentations [4], multimedia database searching, the temporal, spatial, or spatio-temporal relations of various media streams and semantic objects [5], and multimedia browsing [6].

A multimedia presentation consists of media streams that are displayed together or separately across time. The arcs in an ATN represent the time flow from one state to another. An ATN can be represented diagrammatically by a labeled directed graph called a *transition graph*. The ATN grammar consists of a finite set of nodes (states) connected by labeled directed arcs. An arc represents an allowable transition from the state at its tail to the state at its head; the labeled arc represents a transition function. An input string is accepted by the grammar if there is a path of transitions that corresponds to the sequence of symbols in the string and that leads from a specified initial state to one of a set of specified final states. ATN differs from a finite state automata in that it permits recursion, so an ATN is a *recursive transition network*. Each nonterminal symbol consists of a subnetwork that can be used to model the temporal and spatial characteristics of semantic objects for images and video frames and of keywords for texts. In addition, a subnetwork can represent another presentation. Any change in one of the subnetworks will automatically change any presentation that includes these subnetworks. It is difficult to design a multimedia presentation from scratch using today's authoring environments. The subnetworks in ATN allow designers to use existing presentation sequences in their archives, which makes ATN a powerful model for creating a new presentation. This is similar to the *class* in the object-oriented paradigm.

In ATNs based on the selections of users, an input symbol is read and the corresponding conditions and actions are used to pass control to the desired state so the presentation can continue. Therefore, ATN allows nested selections and merges for the scenario in Figure 1.

As shown in Figure 3, arc symbol $V_1\&T_1$ in arc number 1 represents that V_1 and T_1 are displayed together. The & symbol as defined in [5] denotes concurrent display. After V_1 and T_1 are displayed, the control reaches a state node named $P_1/V_1\&T_1$. This state name means that at this stage presentation P_1 finishes the display of media streams V_1 and T_1. Then, two selection buttons B_1 and B_2 on arc number 2 are displayed. The control reaches a state node named $P_1/B_1\&B_2$. If B_1 or B_2 is selected, then arc number 3 or arc number 4 is traversed, respectively. The control is done by the condition/action table as shown in [4]. After B_1 is selected, arc number 5 is traversed and V_2 and T_2 are displayed. Otherwise, B_2 is selected and V_3 and T_3 are displayed. Then, arc number 7 or 8 is traversed. The arc symbols are the same for both arc numbers 7 and 8, which will display V_4 and T_4. Two selections then merge at a state node named $P_1/V_2\&T_2$ or $P_1/V_3\&T_3$, depending on the selection path that was previously selected, and arc number 5 is traversed.

Through this example, we clearly show that an ATN can model user interactions in a single framework, unlike timeline models that need to have several timelines to model the same situation.

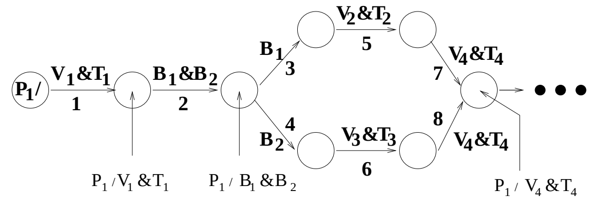

Figure 3: Augmented transition network with user interaction.

4 Conclusions

In this paper, we proposed an abstract semantic model, the augmented transition network (ATN), to model interactive multimedia presentation systems. By using the feature that allows each state node to have multiple outgoing arcs, user interactions in multimedia presentations can be modeled by the proposed ATN model. Examples showing the difficulties and overhead brought about by modeling user interactions with the Timeline model and the OCPN model are also presented in this paper. As we have shown, ATNs have advantages over the Timeline and OCPN models in modeling user interactions in multimedia presentation systems. Unlike the Timeline model, which needs to use different timelines to model user selections, and the OCPN model, which has difficulty in modeling user selections, the ATN model can be used to model user selections in a single framework. Therefore, the ATN model is good to model applications which require a lot of human interaction such as computer-aided instruction, multimedia document systems, digital libraries, and video games.

Acknowledgement

This research was supported in part by NASA (under grants NAGW-4080, NAG5-5095, NAS5-97222, and NAG5-6830), NSF (CDA-9711582, IRI-9409661, HRD-9707076, and ANI-9876409) and ARO (DAAH04-96-0049).

References

[1] Y.Y. Al-Salqan and C.K. Chang, "Temporal Relations and Synchronization Agents," *IEEE Multimedia*, pp. 30-39, Summer 1996.

[2] G. Blakowski, J. Huebel, and U. Langrehr, "Tools for Specifying and Executing Synchronized Multimedia Presentations," Proc. 2nd Int'l Workshop on Network and Operating System Support for Digital Audio and Video, pp. 271-279, 1991.

[3] H.J. Chang, T.Y. Hou, S.K. Chang, "The Management and Application of Teleaction Objects," *ACM Multimedia Systems Journal*, Vol. 3, pp. 228-237, November 1995.

[4] Shu-Ching Chen and R. L. Kashyap, "Temporal and Spatial Semantic Models for Multimedia Presentations," 1997 International Symposium on Multimedia Information Processing, pp. 441-446, Dec. 11-13, 1997.

[5] Shu-Ching Chen and R. L. Kashyap, "A Spatio-Temporal Semantic Model for Multimedia Presentations and Multimedia Database Systems," IEEE Trans. on Knowledge and Data Engineering, accepted for publication.

[6] Shu-Ching Chen, Srinivas Sista, Mei-Ling Shyu, and R. L. Kashyap, "Augmented Transition Networks as Video Browsing Models for Multimedia Databases and Multimedia Information Systems," 11th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'99), pp. 175-182, November 9-11, 1999, Chicago, IL, U.S.A.

[7] N. Hirzalla, B. Falchuk, and A. Karmouch, "A Temporal Model for Interactive Multimedia Scenarios," *IEEE Multimedia*, pp. 24-31, Fall 1995.

[8] C.C. Lin, J.X., S.K. Chang, "Transformation and Exchange of Multimedia Objects in Distributed Multimedia Systems," *ACM Multimedia Systems Journal*, Vol. 4, pp. 12-29, February 1996.

[9] T.D.C. Little and A. Ghafoor, "Synchronization and Storage Models for Multimedia Objects," *IEEE J. Selected Areas in Commun.*, Vol. 9, pp. 413-427, Apr. 1990.

[10] J.L. Peterson, "Petri nets," *ACM Comput. Surveys*, Vol. 9, pp. 223-252, Sept. 1977.

[11] Heiko Thimm and Wolfgang Klas, "δ-Sets for Optimized Relative Adaptive Playout Management in Distributed Multimedia Database Systems," in *IEEE 12th International Conference on Data Engineering*, New Orleans, Louisiana, pp. 584-592, 1996.

[12] W. Woods, "Transition Network Grammars for Natural Language Analysis," *Comm. of the ACM*, **13**, October 1970, pp. 591-602.