Design of an Intelligent Distributed Multimedia Presentation System Using Temporal Algebra and SMIL

Anthony Y. Chang
Department of Information Technology
the Overseas Chinese Institute of Technology
Taichung, Taiwan, R.O.C.
email: achang@ocit.edu.tw

Abstract

We develop an intelligent multimedia presentation system that supports the presence of multiple distributed multimedia objects, as well as human-computer interaction. Several computation algorithms and operation tables which include a set of complete temporal logics are proposed. Using the authoring environment, authors could describe the temporal behavior of multimedia presentation, and specify the layout of the presentation on a screen and associate hyperlinks with media object over internet. The presentation system generates SMIL documents automatically after author designing. A SMIL player will be implemented for presenting another SMIL documents. The player could also analyzes the SMIL presentation contents, records grouping objects, time, linking, structure, and reasoning results.

Keywords: Multimedia Synchronization, SMIL, Multimedia Documents, Temporal Consistency, Temporal Algebra, Scheduling

This work is supported by the National Science Council of Taiwan, under Grant NSC-92-2213-E-240-003.

1. Introduction

To synchronize the various types of multimedia objects is the major challenge for multimedia systems. A three-level synchronization reference architecture consists of the media layer, the stream layer, and the object layer has been proposed by Steinmetz[3] . Little and Ghafoor[6] distinguish among a physical level for intrastream synchronization, a service level for interstream synchronization, and a human level for user specification in multimedia presentation.

There are some temporal frameworks to specify timing of a multimedia system. Little and Ghafoor[7] proposed the object composition Petri Net(OCPN). However it does not easily capture the distributed nature of multimedia application or react to the real world, just indicated in [13] The time-flow graph (TFG) is proposed in [14] but the graphs fail to represent interactive semantics and to reduce ambiguous situations. Time-line diagram is an general and intuitive model, such as QuickTime[15] and MAEStro[16] . In a time-line, it could compute the absolute and relative time precisely, but it fails to specify uncertain and indefinite temporal information.

This paper is based on our early research result about multimedia synchronization mechanisms [2] . We define a temporal algebra system for unifying and scheduling a multimedia presentation. The mechanism is efficiently to eliminate conflict specification and to generate synchronization scenarios. The authoring tools help uses designing a presentation more easily, also the system automatically encode a designed presentation as a consistent W3C standard Synchronized Multimedia Integration Language (SMIL) document.

SMIL is provided to describe multimedia presentation to be distributed over the web [10] [11] [12] It allows the sequencing of audio, video, text and graphic components to be described. Many video clips send SMIL documents before delivering video streams. We use temporal algebra to generate an intelligence presentation system and translate to a SMIL document for content exchanging in distributed environment.
2. Architecture

Multimedia application usually contains a number of multimedia resources to be presented sequentially or concurrently. These resources need to be arranged as layout. We design a visualized user interface by using temporal algebra [1] [2] and composing algorithms for controlling multimedia synchronization. The spatio-temporal derived engine constructs partial order relations about temporal and spatial relations as the core of the intelligence distributed presentation system compatible with SMIL modules. Give an overview of the system:

- Multimedia Resource Browser: before a user is about to design a presentation, multimedia resources are indicated from web allocation.
- Temporal Specification Editor: the presentation system allows authors to specify how components are related temporally to media objects or events in an interactive multimedia presentation.
- Spatial Specification Editor: the spatial relations are extended from temporal relations, to specify the spatial information by 2-tuple relations.
- Synchronization and Layout Engine: computes schedules and layouts between the resources, and translates algebra representation into SMIL documents or inverse.
- Presentation Database: a multimedia document database for grouping application or content exchange in distributed environment.

In following sections, we focus on temporal algebra system and the synchronization/layout engine that is the most important contribution in this paper.

3. Controlling Synchronization

Since the SMIL document is base on timeline model with a dynamic time graph, the inconsistency often occurs in both qualitative semantics and quantitative values. We propose a methodology based on point algebra to deal with qualitative and quantitative inconsistency. SMIL 2.0 defines 10 major functional grouping of attributes. The timing and synchronization module sets are the core of SMIL specification. A temporal scenario is a set of temporally events. There are three basic time containers in a SMIL presentation document:

- The <seq> container plays a sequence of children in which elements play one after the other.
- The <par> container plays a group of children in which multiple elements can share a common timebase and playback at the same time.
- The <excl> container plays one child at a time, but only one of the children can be active at the time.

Modeling a temporal scenario often requires synchronizing the distributed multimedia objects.

We propose a generic temporal constraint model to compose consistent scenarios without conflicts, to integrate event with point and interval temporal specification. This model can deal with both accurate and indefinite scenarios.

From the point view of any single process, events are ordered uniquely by times that shown on the local clock. Qualitative calculus is calculus of intervals instead of real numbers. To deal with qualitative representation, the real number in timeline is subdivided into three intervals: \([-\infty, 0]\), \([0, 0]\) and \([0, +\infty]\), and are denoted as \(<\), \(=\) and \(>\) (before, simultaneously, and after) for representing relations between two points. The notation R1 and R2 denotes the point relations over three points A, B, and C which A R1 B and B R2 C.

In order to express more precise relations without losing qualitative information, the temporal relations extended with qualitative mechanisms for handling quantitative information. To give a concrete form to the topic of temporal representation, consider the following variable and equations with quantitative and qualitative information.

Definition 3.1: A quantitative-qualitative variable is defined by a 2-tuple

\[ Q = (Q_r, V) \]

where \(Q_r\) is a qualitative relation, and \(V\) is a quantity which indicates degrees of \(Q_r\).

Definition 3.2: Formal Endpoint Relations
A formal endpoint relation \( Q_E = (E_R, V_E) \) is a quantitative-qualitative valuable.

where \( E_R \) is an endpoint relation based on the point space \( \{<, =, >\} \), and \( V_E \) is a quantity which expresses a quantitative value associated with \( E_R \) between two endpoints.

**Definition 3.3:** quantitative-qualitative functions

Addition, subtraction and unary minus are functions that take two quantitative variables and return a third:

\[
\begin{align*}
+ & : Q \times Q \rightarrow Q \\
- & : Q \times Q \rightarrow Q \\
[\cdot] & : Q \times Q \rightarrow Q
\end{align*}
\]

The meaning of some quantitative-qualitative calculus operators and equality are defined as follows.

**Example**

- Given \([d_A], [d_B] \) and \([\text{ABbb}]\)

  \[
  [\text{AB}] Q_E [\text{BE}] = [\text{AB}] Q_E [\text{BB}] + [d_A] \\
  [\text{AE}] Q_E [\text{BB}] = [-] [d_A] + [\text{AB}] Q_E [\text{BB}] \\
  [\text{AE}] Q_E [\text{BE}] = [-] [d_A] + [\text{AB}] Q_E [\text{BB}] + [d_A]
  \]

  where \([d_A], [d_B], [\text{AB}], [\text{BB}], [\text{AE}], \) and \([\text{BE}]\) are expressing duration of A, duration of B, begin of A, begin of B, end of A, and end of B, respectively.

The Synchronization and Layout Engine could translate the temporal algebra into SMIL or inverse. The following examples show how timing could be encoded with both.

<table>
<thead>
<tr>
<th>SMIL Specifications</th>
<th>Algebra Notations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(&lt;\text{seq}&gt;)</td>
<td>([\text{Ab}] = 3s)</td>
</tr>
<tr>
<td>(&lt;\text{img id=&quot;A&quot; src=&quot;a.txt&quot; begin=&quot;3s&quot;/}&gt;)</td>
<td>([\text{Ca}] = 10s)</td>
</tr>
<tr>
<td>(&lt;\text{par}&gt;)</td>
<td>([\text{Db}] = [\text{Cb}] + 25s = 7s)</td>
</tr>
</tbody>
</table>

The temporal algebra system is proved as an algebraic group, with associative and transitive relations. We could compute timing from serial specifications.

### 4. Spatial Layout

Since a 2-D relation is conjunction of two 1-D relations, we use the notation, \(Q_1 \times Q_2\), to denote a 2-D relation set where \(Q_i\) and \(Q_j\) are two 1-D relation sets. Thus \(f^1\) is a mapping from Cartesian product of two \(Q \times Q\) s to a \(Q \times Q\). Similarly \(f^2\) is obtained. The following are signatures of these functions:

\[
\begin{align*}
\forall i, j, i_1, i_2, j_1, j_2 & \in P (Q \times Q) \\
\forall i, j, i_1, i_2, j_1, j_2 & \in P (Q \times Q) \\
\forall i, j, i_1, i_2, j_1, j_2 & \in P (Q \times Q)
\end{align*}
\]

\[
\begin{align*}
f^1 &= Q \times Q \rightarrow 29\text{RelSet} \\
f^2 &= Q \times Q \times Q \times Q \rightarrow Q \times Q \\
f^3 &= Q \times Q \times Q \times Q \rightarrow Q \times Q \times Q \times Q
\end{align*}
\]

Functions \(f^1\), \(f^2\), and \(f^3\) are computed according to the following formulas:

\[
\begin{align*}
\forall i, j_1, i_2, j_2 & \in P (Q \times Q) \\
f^2 (i, j_1, i_2, j_2) &= \prod f^1 (i_1, i_2) \times f^1 (j_1, j_2) \\
\forall i, j_1, i_2, j_2 & \in P (Q \times Q) \\
f^3 (i, j_1, i_2, j_2) &= \prod f^1 (i_1, i_2) \times f^2 (j_1, j_2) \times f^3 (j_1, j_2)
\end{align*}
\]

Since multimedia presentations are dynamic, the definition of presentation layout is with respect to a time point. Each spatial relation is designed as well as its graphical user interface with ICON and values.
5. Parsing a SMIL Document

In order to identify SMIL documents and to play a SMIL presentation we parse the SMIL document into phrases based on SMIL tags. The SMIL has many functionality modules; we just focus on timing, Synchronization, media control and layout modules as an intelligent authoring tool for helping user design temporal synchronization and spatial layout more easily. After parsing a variable-length SMIL code, each temporal identifier is translate to consistent algebra notation and reasoning about complete temporal relations. A time-line scheduling is also provided for assistant of a presentation designing. The semantics of the SMIL document could be extracted from temporal relations of the beginning and ending point. It also be represented as a high level interval form.

With group manipulation, group objects of a presentation could be reused and be played independently. User could search presentation document, media attribute, and semantics from presentation database.

6. Conclusions

In this paper, we develop an intelligent methodology to diagnose the temporal consistency of SMIL document that supports the presence of multiple distributed multimedia objects, as well as human-computer interaction. This paper is based on our early research result about multimedia synchronization mechanisms. We develop a temporal algebra system as multimedia synchronization model to unify media presentation time and interaction event. The synchronization models are generalized by composing point temporal relations with qualitative and quantitative functions. We also construct an automatic reasoning framework in which to help user design presentation more easily. Temporal semantics can be extracted from the solutions. Authors could easily describe the temporal behavior of multimedia presentation associate hyperlinks with media object over internet.

The main contribution of this paper is to provide a temporal algebra system for consistent standard SMIL document with highly semantic. We hope this system could benefit other interesting researches and applications such as interactive multimedia course design presentation etc.

References