An Approach to Checking Consistency in Multimedia Synchronization Constraints

Huadong Ma\textsuperscript{aA} and Kang G. Shin\textsuperscript{b}

\textsuperscript{a} College of Computer Sci.& Tech., Beijing University of Posts & Telecom., Beijing 100876, China
\textsuperscript{b} RTCL, EECS Dept., The University of Michigan, Ann Arbor, MI 48109-2122, USA

Abstract—This paper proposes a new method for analyzing multimedia synchronization constraints based on the constraint graph and classification, which is essential in developing efficient support tools for constraint-based authoring systems. We specify the temporal relations between multimedia objects, and use a directed graph to represent the constraints among the objects in a multimedia scenario. We then analyze synchronization constraints based on graph theory, solving multiple problems in a unified theoretical framework: completeness and/or consistency checking, constraints relaxation and automatic temporal layout generation. We also discuss the effects of user interactive authoring. Compared to the other methods, it is simpler, more efficient, and easier to implement.

I. INTRODUCTION

One of the peculiar characteristics of multimedia data is that they are tightly-coupled through both spatial and temporal constraints. Thus, a key problem in multimedia systems is how to model spatial and temporal synchronization. The temporal relation is considered as the master relation in a multimedia scenario, and the spatial relation is a slave to the temporal relation. So we focus on the temporal synchronization.

The temporal relations can be specified by two basic models: time-line and constraint-based models. Constraint-based models have a definite advantage over the traditional time-line approach mainly to their flexibility. Although the constraint-based authoring systems have several distinct advantages, they must provide a mechanism for users to easily manipulate the underlying structures so as to implement completeness and consistency checking, constraints relaxation and layout generation.

The temporal synchronization was addressed in [3,4]. Temporal-synchronization models can be classified as a graph model [5,7], Petri-net based model [9], object-oriented model [8], or language-based model [2,6,10]. In a graph model, constraints are represented as edges in a graph, whereas a Petri-net model captures the temporal constraints by means of places and transitions. In an object-oriented model the constraints are modeled as object attributes, while a language-based model uses programming language constructs to express constraints.

An important issue, which has not yet been fully investigated, is consistency checking. Especially for multi-user collaborative authoring of a multimedia scenario, there usually exist inconsistencies in constraints specified by different users. In spite of its importance to the production of a correct presentation, only a few methods, such as Firefly [5], TCSP [2] and Elastic Time Model [7], explicitly provide such mechanisms. Another important issue is how to deal with the case when the satisfaction of synchronization constraints cannot be guaranteed, i.e., when an inconsistency is found. In such a case, it may be useful to provide a strategy to relax the constraints so as to obtain a feasible presentation. The simplest way is to explicitly require the insertion or the deletion of some constraints from the specification. Alternatively, constraints can be divided into mandatory and optional parts. Or one might provide some form of constraint relaxation by shrinking or stretching the duration of each object when a presentation is infeasible.

To our best knowledge, there has been no report in literature on checking for the incompleteness of synchronization constraints, hence no unique layout of presentation. We propose a new synchronization constraint analysis model, which differs from Firefly and Elastic Time models, for completeness and consistency checking, constraint relaxation and layout generation (either interactively or automatically). Main contributions of this paper are summarized as follows. We first propose the concept of constraint completeness and divide the constraints into different categories, then introduce the priority number to identify them so as to take different policies in removing inconsistencies, and use a directed graph to express the constraints among multimedia objects. Our algorithms improve efficiency in checking consistency with a simpler model. Moreover, we solve the above problems in a unified theoretical framework.

II. TEMPORAL SYNCHRONIZATION CONSTRAINTS

A. Temporal Relations

Assume that object A and B are scheduled in the interval $[p_A, t_A], [b_B, t_B]$, respectively. As shown in Table 1, there are 13 possible temporal relations between A and B, that is, equals, before, meets, starts, finishes, during, overlaps and their inverse relations [1].

<table>
<thead>
<tr>
<th>Relations</th>
<th>Symbols</th>
<th>Reverse</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A equals B</td>
<td>$e(A, B)$</td>
<td>$e(B, A)$</td>
<td>$b_A = b_B &lt; t_A = t_B$</td>
</tr>
<tr>
<td>A before B</td>
<td>$b(A, B)$</td>
<td>$b^{-1}(B, A)$</td>
<td>$b_A &lt; t_A &lt; b_B &lt; t_B$</td>
</tr>
<tr>
<td>A meets B</td>
<td>$m(A, B)$</td>
<td>$m^{-1}(B, A)$</td>
<td>$b_A &lt; t_A = t_B &lt; b_B$</td>
</tr>
<tr>
<td>A starts B</td>
<td>$s(A, B)$</td>
<td>$s^{-1}(B, A)$</td>
<td>$b_A = b_B &lt; t_A &lt; t_B$</td>
</tr>
<tr>
<td>A finishes B</td>
<td>$f(A, B)$</td>
<td>$f^{-1}(B, A)$</td>
<td>$b_B = b_A &lt; t_A &lt; t_B$</td>
</tr>
<tr>
<td>A during B</td>
<td>$d(A, B)$</td>
<td>$d^{-1}(B, A)$</td>
<td>$b_B &lt; b_A &lt; t_A &lt; t_B$</td>
</tr>
<tr>
<td>A overlaps B</td>
<td>$o(A, B)$</td>
<td>$o^{-1}(B, A)$</td>
<td>$b_A &lt; b_B &lt; t_A &lt; t_B$</td>
</tr>
</tbody>
</table>

In order to simplify the representation of temporal relations, we merge some of them, and introduce new notations as follows. $Cov(A, B, d_1, d_2) (d_1, d_2 \geq 0)$ means the presentation interval of B is covered by that of A. It is used to specify concurrent presentations of objects instead of $e, s, f, d$.

$Ovl(A, B, d_1, d_2) (d_1, d_2 > 0)$ means there is a partial overlap in presenting A and B. It is used to specify cross presentations of objects in the same way as $o$.

$Seq(A, B, d) (d \geq 0)$ means the presentation of B is after that of A with delay d. It is used to specify sequential presentations of objects instead of $m, b$. 

The work is supported by the National Natural Science Foundation of China, UN/US/IST Off-shore Project and US NSF under Grant EIA-9806280. The first author is a visiting scholar with RTCL, University of Michigan under the support of CSC. Correspondence email: {mahd,kgshin}@eecs.umich.edu.
Table 2 summarizes the definitions of three relations.

<table>
<thead>
<tr>
<th>Relations</th>
<th>Symbols</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequential</td>
<td>Seq(A,B,d)</td>
<td>( b_A &lt; t_A \leq b_B &lt; t_B ) if ( d_B = t_B - t_A )</td>
</tr>
<tr>
<td>Cover</td>
<td>Cov(A,B,d1,d2)</td>
<td>( b_A \leq b_B \leq t_A \leq t_B ) if ( d_1 = t_B - b_A, d_2 = t_A - t_B )</td>
</tr>
<tr>
<td>Overlap</td>
<td>Ovl(A,B,d1,d2)</td>
<td>( b_A &lt; b_B &lt; t_A &lt; t_B ) if ( d_1 = t_B - b_A, d_2 = t_B - t_A )</td>
</tr>
</tbody>
</table>

**Theorem 1.** The relation set consisting of Cov, Ovl and Seq is a complete set of temporal relations.

We can easily prove this theorem by representing all of 13 temporal relations using Cov, Ovl and Seq.

In our approach, they are used as temporal synchronization constraints illustrated in Figure 1.

![Figure 1. The diagram of temporal constraints](image)

**Example 1:** Consider a multimedia scenario which first presents a 10s logo consisting of Log-Ani (LA) and Logo-Music (LM) and then plays 20s Audio A1 with Video V1; Text T1 will be displayed while playing V1; at end, it plays 10s Exit-Music (EM) with Exit-Ani (EA), V1 may overlap with EM. Seven constraints are used to specify the scenario: \( R_1 = \text{Cov}(\text{LA,LM},0,0) \); \( R_2 = \text{Cov}(\text{A1,V1,4,0}) \); \( R_3 = \text{Cov}(\text{A1,V1,0,0}) \); \( R_4 = \text{Seq}(\text{A1, EM, 0}) \); \( R_5 = \text{Cov}(\text{EM,EA, 0, 2}) \); \( R_6 = \text{Cov}(\text{A1, T1, 1, 0, 4}) \); \( R_7 = \text{Ovl}(\text{V1, EA, 15, 5}) \). We assign the priority number 1 to \( R_1, R_3, R_5; 2 \) to \( R_2, R_4, R_6 \); and 3 to \( R_7 \).

**B. The constraint analysis model**

**Definition 1.** The Temporal Constraints Graph (TCG) is defined as a three tuple TCG=\((V,E,W)\), where \( V=\{O_1, \ldots , O_N\} \) is the set of vertices representing \( N \) media objects, \( E \subset V \times V \) is the set of edges specifying the temporal constraints between objects. There is a directed edge \( e(i,j) \) from \( O_i \) to \( O_j \) if there is a constraint \( C(O_i, O_j) \) between them, where \( C \) is one of three primitive constraints. \( W : E \rightarrow \{1, 2, 3\} \) is a mapping from \( E \) to integer for defining a priority number for each constraint.

In general, \(|E|\) is much smaller than \( N^2 \), so we assume that an input graph is represented in adjacency lists, in which \( e(i,j) \) is stored as an element of the list \( Adj[i] \). There may be more than one constraint between a pair of objects, but it obviously causes inconsistency. In this case, a best acceptable constraint is determined by comparing these constraints and reporting the inconsistency to the user. The best one should have the minimum priority number, or the one the user interactively chooses. The data structure for an edge in the list is as follows.

```c
  struct Edge {
    Integer idv; /* directed vertex number*/
    Constraint cot; /* constraint type: Cov, Ovl or Seq*/
    Integer pnum; /* priority number*/
    Edge *nextedge; /* pointer to next edge*/
  };
```

Thus the initial TCG will be structured by a preprocessing algorithm `PreTCG` on the above idea. There is at most one constraint for any pair of vertices in a TCG. Figure 2(a) shows the TCG of the scenario in Example 1.

![Figure 2. A TCG](image)

**III. CONSTRAINT CHECKING**

**A. Completeness Checking**

**Definition 2.** The constraints of a multimedia scenario are said to be complete if there is a given or implied constraint \( C(O_i, O_j) \) or \( C(O_j, O_i) \) for any pair of objects \( O_i, O_j \).

An undirected graph is connected if every pair of vertices are connected. So the concept of an undirected TCG can be used to map the completeness of temporal constraints.

**Definition 3.** If an undirected TCG is connected, it is called a Complete TCG (CTCG).

We need to determine if the TCG is a CTCG. If not, the constraints are not enough to fix the temporal relationships among the vertices in TCG. In general, an undirected graph has some connected components. An undirected graph is connected if it has exactly one connected component. We developed an algorithm `JudgeCTCG`, which finds the connected components of TCG. If a TCG has more than one connected component, the constraints are incomplete. Additional constraints should be added the sets of vertices of different components. Actually, the user just needs to add \( K-I \) constraints if there are \( K \) connected components. TCG will become a connected graph after appending constraints. Figure 2(b) is the CTCG of Example 1 by adding \( R_8 = \text{Seq}(\text{LA,V1,0}) \).

**B. Consistency Checking**

When checking consistency, a CTCG is considered as a connected, undirected graph. We can prove the following theorem by constructing an acyclic constraint graph (tree).

**Theorem 2.** The layout of presenting \( N \) media objects can be fixed by \( N-I \) constraints among them, where each object appears at least in one constraint.

Inconsistencies are usually due to the redundancy in constraints. If there are more than \( N-I \) edges in a CTCG representing constraints among \( N \) objects, there must exist redundant constraints. In order to find the best \( N-I \) constraints the user hopes to have, we consider the priority number of each constraint and then find an acyclic subset \( T \in E \) that connects all of vertices and whose total weight (priority number) \( w(T)=\sum_{c \in T} c.pnum \) is minimized. Since \( T \) is acyclic and connects all of vertices, it must be a spanning tree. So, the minimum spanning tree algorithm can be used to solve this problem.

Figure 3(a) gives the minimum priority number spanning tree (MPST) for the CTCG of Example 1. After finding the MPST \( T \)
of CTCG, we can obtain some constraints that are not selected by T if the number of CTCG constraints is greater than \( N - 1 \). These constraints are possibly inconsistent with those in T. In order to remove the inconsistencies, we need to determine if the remaining constraints are acceptable. We now describe the consistency checking method and the relaxation strategy.

**Table 3. Possible priority number combinations**

<table>
<thead>
<tr>
<th>Case</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.pnum</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>p</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

From the constraints in \( \alpha \) except \( e(i,j) \), we can get the temporal relation between \( O_i \) and \( O_j \) by using the definitions of constraints repeatedly in the order of cycle, and Table 4 shows the formula we used. Note that we should consider the direction of constraint (edge of directed TCG) to find the correct relation.

**Table 4. The constraint relation**

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Relations between objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq(A,B,d1)</td>
<td>( b_y = t_A + d_1, t_y = t_B + t_A + d_1 )</td>
</tr>
<tr>
<td>Cov(A,B,d1,d2)</td>
<td>( b_y = t_A + d_1, t_y = t_B - d_2 )</td>
</tr>
<tr>
<td>Ovl(A,B,d1,d2)</td>
<td>( b_y = b_A + d_1, t_y = t_B + d_2 )</td>
</tr>
</tbody>
</table>

Thus, the relaxation policy of inconsistent constraints can be decided by the user once they are found. We just give a default policy. Given \( dx = b_i - b_j, dy = t_i - t_j \), where \( b_i,t_i \) are the start time and end time of \( O_i \), respectively, there are 6 cases in total, which correspond to 6 different relations between \( O_i \) and \( O_j \), as shown in Table 5.

**Table 5. The possible relations**

<table>
<thead>
<tr>
<th>Case</th>
<th>Conditions</th>
<th>Impplied constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( dx &gt; 0, dy &gt; 0, b_i &lt; t_j )</td>
<td>( Cov(O_i, O_j, dx, dy) )</td>
</tr>
<tr>
<td>2</td>
<td>( dx &lt; 0, dy &gt; 0, b_i &lt; t_j )</td>
<td>( Seq(O_i, O_j, -dx, -dy) )</td>
</tr>
<tr>
<td>3</td>
<td>( dx &gt; 0, dy &lt; 0 )</td>
<td>( Cov(O_i, O_j, dx, -dy) )</td>
</tr>
<tr>
<td>4</td>
<td>( dx &lt; 0, dy &gt; 0 )</td>
<td>( Cov(O_i, O_j, -dx, dy) )</td>
</tr>
<tr>
<td>5</td>
<td>( b_i &gt; t_j )</td>
<td>( Seq(O_i, O_j, b_i - t_j) )</td>
</tr>
<tr>
<td>6</td>
<td>( b_i &gt; t_j )</td>
<td>( Seq(O_i, O_j, b_i - t_j) )</td>
</tr>
</tbody>
</table>

Now, we take \( C(O_i, O_j) \) to compare with the found relation \( F \), then decide how to use the strategies to relax the synchronization constraints in order to obtain a feasible specification.

1. If their constraint types and directions are same, they are compatible. They can be relaxed as follows:
   (a) If \( p \) is the same as C.pnum (i.e., C.pnum \( \neq (1,1), (2,2) \) or \( (3,3) \)), the parameters of C are relaxed as follows: \( C.d1 = (C.d1 + 2Ed1)/3 \), \( C.d2 = (C.d2 + 2Ed2)/3 \) (if necessary). \( C(O_i, O_j) \) is appended to the CTCG, and remove the constraint with priority number \( p \) which is the nearest to node \( i \) or \( j \) in \( \alpha \).
   (b) If C.pnum is \( (3,3) \), then keep the original constraints, and treat \( C(O_i, O_j) \) as an inconsistent constraint, so remove it from the constraint set.

2. If their types are same but their directions are different, only Cov constraints are compatible, i.e., \( C(O_i, O_j) = Cov(O_i, O_j, d1, d2) \). The constraints can be relaxed according to their priority numbers.

3. If they have different types, or have different directions and the same type except Cov, they are considered as incompatible. The relaxation cannot be done. If the user doesn't select one when their priority numbers are same, C is, by default, an inconsistent constraint, and remove it from the constraint set.

After relaxing all of redundant constraints, we can obtain the final TCG. The final TCG of Example 1 is Figure 3(a) with \( R6 = Cov(A1, T1, 2, 2) \) relaxed and \( R2 \) moved.

**IV. SUPPORT FOR INTERACTIVE SYSTEMS**

**A. Interactive Authoring**

Possible authoring operations for a user are as follows.

**Add an object AddObj(X):** addition of a new object also needs a constraint between X and an old object (vertex). If there are more than one such constraint, select the one with the minimum priority number first and add it to the TCG. At the same time, its addition to the old MPST forms the new MPST of modified graph. Then, consistency checking for additional constraints on X will be done by the algorithm in Section 3.2.

**Remove an object RmVObj(X):** deletion of a media object also causes inefficient constraints among X and other objects (vertices). If there is only one such constraint in the old MPST (X is a leaf node), just delete it to form a new MPST. Otherwise (X isn't a leaf node), we need to delete all constraints on X from the old TCG, and find a new CTCG and a new MPST. Moreover, consistency checking will be done again.

**Add a constraint AddCon(C):** addition of a new constraint possibly causes the change of MPST. A cycle will be formed after adding C to MPST. If \( C.pnum \) is smaller than another constraint, denoted as \( C' \), with the maximum priority number, then we can delete \( C' \) and add C in the old MPST to form the new MPST. Otherwise, consistency checking for C will be done.

**Remove a constraint RmVCon(C):** deletion of a constraint may cause incomplete constraints. Thus, we need to delete this constraint from the old TCG, and find a new CTCG and a new MPST. Moreover, consistency checking is also done again.

**Modify a constraint ModCon(C):** modification of a constraint can be divided into two cases: (1) \( C.pnum \) is unchanged, MPST will be unchanged and the other parameters of C are possibly modified if C is in MPST; (2) \( C.pnum \) is changed. We need to
reconstruct the MPST from CTCG with the modified C.

B. Generation of Temporal Layout

To generate temporal layout of a scenario, MPST is considered as a directed graph. We can use breadth-first traversal to find the unique paths from a given vertex $O_i$ (the source) to all other vertices. For example, the path from $O_1$ to $O_i$: $e(s, t), e(t, v), ..., e(v, i), (0 \leq p \leq N-2)$. 

Assume that the presentation interval of object $O_i$ is $[b_i, t_i]$. Once such a path is found, we can get the temporal relation between $O_i$ and $O_j$ according to the constraints in this path $b[i] = b_i + db[i], t[i] = t_i + dt[i]$, where $b[i], t[i]$ represent the start and end time of presenting $O_i$, respectively. Thus, we get a list of $2N$ timed events consisting of the start and end events, denoted as $P_i, P_j$.

$\Gamma : (P_0[s], b_0), (P_1[s], t_0), (P_1[1], b_1 + db[1]), (P_1[1], t_1 + dt[1]), ..., (P_N[N], b_N + db[N]), (P_N[N], t_N + dt[N])$

then replace $t_N$ by $b_N + t(O_2)$ and sort them in the temporal order. Note that $l(O_i)$ is the length of $O_i$,

$\Gamma : (a_1, b_1 + d[1]), (a_2, b_2 + d[2]), ..., (a_{2N}, b_{2N} + d[2N])$,

where $a_i$ is a start or end event of presenting an object $P_i$ or $P_j$. If the first event is scheduled at time 0, take $b_1 = -d[1]$, then $b_1$ is replaced in the above list by $-d[1]$, the final layout of presentation is as follows:

$\Gamma : (a_1, 0), (a_2, t[2]), ..., (a_{2N}, t[2N])$.

Figure 4 shows the generated temporal layout for Example 1.

![Figure 4. The layout of Example 1](image)

V. ANALYSIS AND COMPARISON

Suppose that there are $N$ media objects and $M$ constraints among them. Now we consider the efficiency of the algorithm.

Our approach consists of the following algorithms: PreCTCG preprocesses the temporal constraint graph taking $O(M \times (1 + \log(M/N))$ time. JudgeCTCG is used to judge if the CTCG is CTCG, and its total time is $O(N + M \log N)$. Finding MSPT takes $O(N \log M)$ time. Removing inconsistency takes $O(M + N + 1)N$ time.

Generating temporal layout includes the tree traversal and sorting timed events which take $O(N \log N)$ time. In general, $M > N-1$ and $M$ is about $N-1$. So, the total running time of our approach is $O(N \log M)$.

Many existing models do not provide consistency checking. Compared with several main consistency checking models, our method has the following advantages.

A1. The integrated modeling and verification of multimedia constraints. We proposed the concept of completeness of synchronization constraints, which had previously been overlooked. For convenience we divide the constraints into different categories and introduce the priority number to identify them so that we can take different policies in removing inconsistencies. By using a directed graph to express the constraints among objects, we solve the problems on completeness checking, consistency checking, constraint relaxation and layout generation.

A2. A simple method to develop the analysis tools to support a constraint-based system. Our method is based on graph theory which makes the problems in constraint-based systems some simple solved problems, i.e., the completeness of constraints is a mapping to the connectivity of CTCG, and finding best constraints is a mapping to the spanning tree of CTCG. OCPN [9] gives the semantic specification of temporal constraints, TCSP [2], Temporal Logic [10] and Time Automata [6] also efficiently specify the temporal synchronization. Although their proof may be used to check consistency, but it is difficult to develop automatic checking tools for practical multimedia authoring.

A3. More efficient solution. Our constraint analysis model is different from Firefly [5] and Elastic Time Model [7], although all of them are based on graphs. Firefly and Elastic Time Model specify the temporal relationship among the start and end events (in presenting objects) by using vertices to represent events and edge to label the relations between the events. In contrast, our model uses vertices to represent objects, and edges to label the constraints between objects. The complexity decreases due to only half of the number of vertices used. Moreover, by using our method, consistency checking will be finished before generating the feasible temporal layout.

A4. Flexible implementation. Our method supports the above problems in both interactive and batch authoring systems, especially for multi-user collaborative authoring mode. In an interactive authoring system, the user can interactively select the relaxing policy to facilitate consistency checking, constraint relaxation and layout generation. In a batch authoring system, the default processing will solve these problems automatically.

VI. CONCLUSIONS AND FUTURE WORK

This paper proposes a new method of analyzing multimedia synchronization constraints based on graphs and classification. This method provides a unified theoretical framework to solve the problems in constraint-based multimedia authoring systems, such as completeness and consistency checking, constraint relaxation and temporal layout generation. Compared to previous methods, the main advantages of our method are simple, efficient, and easy to implement and handle by nontechnical users. Because the consistency checking of spatial constraints is also important, our future work will consider consistency checking of both temporal and spatial constraints.

REFERENCES