Modeling and Analysis of Spatiotemporal Behavior of Multimedia in SMIL

Stephen J.H. Yang[1], Kevin C.Y. Kuo[2], Norman W.Y. Shao[3], Ben J.D. Wu[4]

National Central University[1,4] *National Kaohsiung First University of Science and Technology[3]*

jhyang[1],wallace[4] @src.ncu.edu.tw *u9111805[2], u9115904[3]@ccms.nkfust.edu.tw*

Abstract*- Along with the broadband technology matured, more and more information can be found in the Internet. Using multimedia resources to express content will be an invertible trend. How to efficient and fast integrate multimedia has been a hot research topic.*

In design of multimedia content, there are three important considerations such as allocation of presentation window, scheduling, and flowchart planning. The other important issue is to verify the designed multimedia plan. Due to each object has its specific feature, the spatial and temporal behavior among these objects easily cause interference during content playing. This paper will focus on temporal and spatial behavior of multimedia in SMIL. We proposed the Reachability Tree analysis methods to detect spatiotemporal interference in the designed multimedia. In this way, teacher can be reminded to exclude the interference in multimedia content.

Keywords: SMIL, Petri nets, Reachability Tree, spatial and temporal behavior

1. Introduction

Currently using richly multimedia to express content is an invertible trend. There are many types of resource existed in the Internet, likes pictures, text, video chip, animation, audio, applet, and etc. Each type of resource has their dedicate features. World Wide Web Consortium (W3C) recommended Synchronized Multimedia Integration Language (SMIL) to enable simple authoring of interactive audiovisual presentations. The most benefit that SMIL contributed is the capability of integration among variety text, audio, video and images. Through SMIL player, a unified and outstanding multimedia content can be presented. SMIL syntax is similar with HyperText Markup Language (HTML) and easily for editing. Using the temporal syntax, SMIL can support ordering and synchronization relation among media. Some researcher has studied how to integrate different types of resources in a window with temporal sequence. This problem can be departed into representation of content and reduce interference.

When a teacher wants to design a scenario for a class, he can use an authoring tool to edit the appearing sequence among selected multimedia objects. Each object has its dedicated spatiotemporal features, likes duration and position in appearance. How to detect the temporal and spatial interference existed in multimedia plan? It is obviously that we need a modeling system, which can simulate the playing sequence and find out possible interference, bottleneck or deadlock.

We need to arrange playing sequence of objects without spatial and temporal interference. Petri Nets (PN) is a systematic analysis method that uses graph, mathematic, and module to simulate and represent the system. Since Carl Adam Petri proposed [1], there are many extended researches, likes Timed Petri Nets [2], Object Composition Petri Nets (OCPN) [3], and Extended Object Composition Petri Nets (XOCPN) [4] which extend from OCPN with spatial properties. The others like Coloured Petri Nets [5] and Multimedia Object Petri Nets [6] is derived from Petri nets. Considered the spatial or/and temporal relation, Chung proposed TPN [7] regards SMIL as a PN model. Hsu proposed Spatial-Temporal Relations Petri Nets (STRPN) [8] and Deng's "Presentation PN" [9]. Table 1 listed the above-mentioned model and described the comparison in advantage and disadvantage.

2. Temporal and Spatial Behavior of Multimedia in SMIL

In this section, we will discuss spatiotemporal relationship among objects.

2.1 Petri nets and SMIL

Petri nets provide a mathematical model to represent system. It only needs few symbols to depict process modeling about a system. The basic definition about Petri nets is shown in Table 2[10].

The basic concept of Petri nets is shown in Figure 1. The place is represented with a white circle in which "marking" can be stored; transition is represented with a box which indicating status shifts. The arc indicates the direction of flow. The arc is often labeled with weight, in general case with '1'. The black circle in place is the value of initial marking. It is used a number to represent the black circle more than three. For example, place may be the resource, likes video chip, text, image, or audio.

Table 1. Comparison among MOPN, TPN, STRPN, and Presentation PN

	MOPN	**TPN**	**STRPN**	**Presentation PN**
Objective	manifestation plan for universal access	transform SMIL temporal behavior into PN	use PN to model temporal and spatial relations between moving multimedia object	combine PN model and SMIL to reach the Multi-level Content Tree purpose
Extended PN	Classic PN	OCPN	OCPN	Classic PN
Advantage	access different manifestation plan in any device	optimize its delivery schedule and utilize the system resource	model temporal and spatial property of multimedia	emphasize workflow and content representation adaptation
Disadvantage	not to discuss the spatial arrangement	not to describe spatial layout	cannot integrate into SMIL	not to define spatial arrangement and spatiotemporal analysis

The marking can be treated as a time marking.

Table 2. Formal Definition of a Petri Net [10]

A Petri nets is a 5-tuple, $PN = (P, T, F, W, M_0)$ where:

$P = \{p_1, p_2, \cdots, p_m\}$ is a finite set of places,

$T = \{t_1, t_2, \cdots, t_n\}$ is a finite set of transitions,

$F \subseteq (P \times T) \cup (T \times P)$ is a set of arcs (flow relation)

$W : F \rightarrow \{1,2,3,\cdots\}$ is a weight function,

$M_0 : P \rightarrow \{0,1,2,3,\cdots\}$ is the initial marking,

$P \cap T = \varnothing$ 且 $P \cup T \neq \varnothing$

A Petri nets structure $N = (P, T, F, W)$ without any specific initial marking is denoted by ***N***.

A Petri nets with the given initial marking is denoted by (N, M_0).

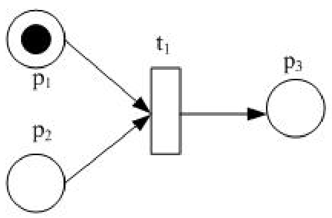

Figure 1. Petri nets Graph

SMIL is an XML-based extension markup language, which initiated from W3C. The currently version is 2.0. It provides a continuous multimedia presentation characterized with spatial and temporal synchronization of multiple integrated media which HTML limited.

In SMIL, every resource can be treated as object. Each object has its own attributes or properties and described in metadata format, likes identification, title, description, format, and etc. SMIL provide an integrated markup platform to organize objects. The user or teacher can use SMIL to arrange the appearing order and position among objects to express the learning objectives.

In summary, SMIL has five major features, as follows: [11]

- Media Content: used to integrate existing variety of multimedia.
- Layout: used to coordinate the scope of content displaying. In our study, layout is a key factor which effects the final representation of content.
- Timing: synchronization control among media is the most important contribution for SMIL. There are two basic elements which are "sequence" and "parallel" in temporal domain. The corresponding attributes are "begin", "end", and "duration".
- Linking: hyperlink function is similar with HTML.
- Adaptive: provide adaptive function to let people tailor content according to characteristics such as language, preference, and device capability.

An example of playing SMIL file is shown in Figure 2.

Figure 2. Example of playing SMIL file

During designing phase, user may not have a completely picture what he/her really wanted. User must wait for playing SMIL file then knows what exactly designed content is. Incorporated with Petri nets, the designed scenario can be visualized to help user design [12].

From Figure 1, it is found that there are four basic elements used in Petri nets such as place, transition, arc, and token. They are the key elements to establish the relationship between PN and SMIL. The mapping is shown in Table 3.

Table 3. Summary between Petri nets and SMIL

Petri nets	Place	Transition	Arc	Token
SMIL	Object property	Scene transition	Flow path	Play status

From Table 3, Place will be extended to define the attributes of multimedia objects, likes spatial and temporal information. Transition is responsible for the switching control among workflows. Arc represents the media workflow. Token will be the status of playing which represent as a black circle.

2.2 Temporal and Spatial Behavior Detecting

To detect conflict among objects, we use three major conditions, i.e. disjoint, meet, and overlap in temporal and spatial domain to illustrate. There are nine interference types existed, as follows: (1) same time with same position (2) same time with overlap position (3) same time with different position (4) overlap time with same position (5) overlap time with overlap position (6) overlap time with different position (7) disjoint time with same position (8) disjoint time with overlap position (9) disjoint time with different position. Audio objects do not need to occupy spatial resource, but it is sensitive with the other audio and sound of video object. This characteristic is very different from other type of object. Therefore, we separate audio object and sound of video from other type to discuss.

- Audio emphasized temporal property.
- Video emphasized both temporal and spatial property.
- Image emphasized spatial property.
- Text emphasized spatial property

Table 4 illustrates the detecting table of temporal conflict for audio and video media objects. Table 5 depicts the detecting table of spatial conflict for video, image, and text media objects.

Table 4. Detecting table of temporal conflict for audio and video media objects

	Same duration	Overlap duration	Disjoint duration
Same position	Yes	Yes	No
Overlap position	Yes	Yes	No
Different position	Yes	Yes	No

Table 5. Detecting table of spatial conflict for video, image, and text media objects

	Same time	Overlap time	Disjoint time
Same position	Yes	Yes	No
Overlap position	Yes	Yes	No
Different position	No	No	No

The comparison and feature between Table 4 and Table 5 is listed as shown in Table 6.

Table 6. Table 4 and Table 5 use method

Use Table	Table 4	Table 5
Purpose	check temporal conflicts between media objects	check spatial conflicts between media objects
Description	1. Row-spatial relation between two media objects. 2. Column-temporal relation between two media objects. 3. Media types: audio, video.	1. Row-spatial relation between two media objects. 2. Column-temporal relation between two media objects. 3. Media types: video, image and text.
Checking procedure	1. Check temporal conflict between media	1. Check spatial conflict between media type. If

	type first. If yes, give warning 2. Then check table Table5 (spatial conflict) between media type	yes, give warning 2. Then check is finish, it's no temporal and spatial conflict between media type

3. Formal Definition for Temporal and Spatial Conflict Analysis

To analyze temporal and spatial conflict, we proposed a modified PN Reachability Tree method to detect spatiotemporal conflict behavior, which produced via PN-based SMIL authoring tool [12]. For the need of PN Reachability Tree, we first explain the PN Reachability Tree, and then expand the definition of MOPN to proposed SMILPN.

3.1 Reachability Tree Logic

Reachability tree logic (RTL) [13] is a branching-time temporal logic in which the specification and verification is based on the reachability tree constructed from TPN. A branching-time temporal logic is built upon a linear-time temporal logic by associating two path quantifiers— for all ($\forall$) and there exist ($\exists$) to each firing sequence (path and firing sequence are used interchangeably in this paper). Reachability tree of RTL is different from the computation tree of CTL [14] in terms of the length of firing paths. The length of a firing sequence in a reachability tree is finite whereas the length of a path in a computation tree may be infinite. A path in a computation tree is as a full path because each path exhibits a full execution behavior of a computation. The primitive firing sequence is referred as the firing sequence from the initial marking (root) to any leaf node in a reachability tree

3.2 SMILPN Definition

The proposed SMIL based Petri Net (SMILPN) is used to model and analysis the SMIL-based multimedia plan. The formal definition of SMILPN is as follows.

$SMILPN = \{PN, TP, SP, ML\}$, where

1. $PN = (P, T, F, W, M_0)$ is a formal definition of PN,
2. Temporal Property:
   $TP = (DEL, TO, MOT, MPT)$,
   DEL is a time delay function of multimedia,
   $DEL : P \rightarrow \tau_{DEL}$, where $\tau_{DEL} = \{w \in \Re | w \geq 0\}$,
   TO is a time out function of multimedia,
   $TO : P \rightarrow \tau_{TO}$, where $\tau_{TO} = \{x \in \Re | x \geq 0\}$,
   MOT is original time function of multimedia,
   $MOT : P \rightarrow \tau_{MOT}$, where $\tau_{MOT} = \{y \in \Re | y \geq 0\}$,
   MPT is playtime function of multimedia,
   $MPT : P \rightarrow \tau_{MPT}$, where $\tau_{MPT} = \{z \in \Re | z \geq 0\}$,
   with limitation that $MOT \geqq MPT$
3. Spatial Property: $SP = (X, Y, W, H)$,
   X and *Y* are coordinate parameters of the upper-left corner of window for media object,
   $X : P \rightarrow \Psi_X$, where $\Psi_X = \{r \in \Re | r \geq 0\}$,
   $Y : P \rightarrow \Psi_Y$, where $\Psi_Y = \{s \in \Re | s \geq 0\}$,
   W is width function of media object,
   $W : P \rightarrow \Psi_W$, where $\Psi_W = \{u \in \Re | u \geq 0\}$,
   H is height function of media object,
   $H : P \rightarrow \Psi_H$, where $\Psi_H = \{v \in \Re | v \geq 0\}$,
4. $ML = \{ml_1, ml_2, \ldots, ml_n\}$ is a set of modalities, and MOD is a modality function such that
   $MOD : P \rightarrow ML$

The firing rules are explained as follows.

1. If the place exist a token and corresponded *MPT* of temporal property (TP) is played out, then we called it "token unlock."
2. For a transition, when corresponding input places have marked token and token is unlock, then these places can enable transition.

An enabled transition will be fired, and then the token of input places will be removed. At the same time, the token of output places will also be added.

4. Analysis of Temporal and Spatial Conflict

In this section, we will show how the Reachability Tree method could analyze temporal and spatial behavior. An example of Multimedia plan, which modeled by proposed SMILPN is shown in Figure 3. The corresponding relations of spatial property (SP) among multimedia objects are illustrated in Figure 4. To verify the result from Reachability analysis, firing sequence σ is used to search objects deadlock and find the playing place at the same time. The mentioned model is expressed via PN Reachability Tree. We use Reachability Tree tool [15] to analyze this model, the result is shown in Figure 5. In addition, the corresponding properties of related places are listed in Table 7. Deadlock is a logical error in processes or workflow. Detecting deadlock does not mean the multimedia plan can be executed without any interference, likes spatial overlapped. Due to the Reachability Tree only can detect the deadlock in multimedia plan; we develop a detecting algorithm based on proposed SMILPN definition. This algorithm will analyze the spatial and temporal relations among objects in the multimedia plan. All possible conflict will notify author to modify.

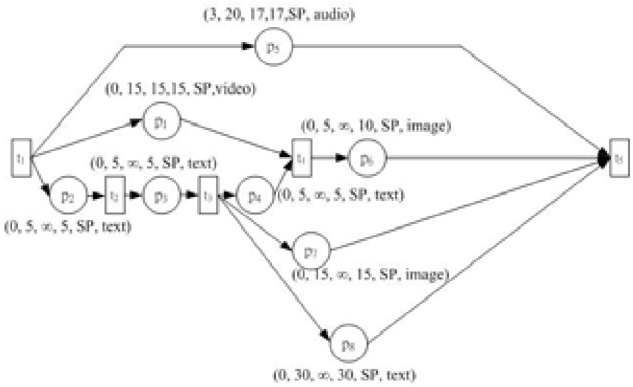

Figure 3. An example modeling SMIL-based multimedia plan with corresponding properties.

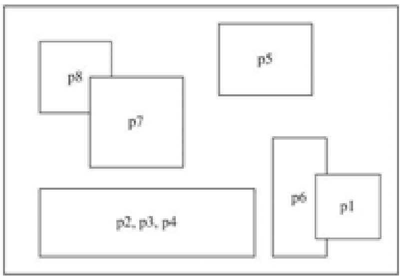

Figure 4. The spatial relation of example.

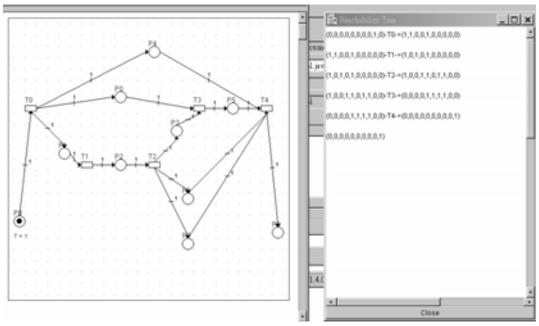

Figure 5. Reachability Tree of example.

Table 7. Property of Places in SMILPN model.

p1:(0, 15, 15, 15, SP, video)；
p2:(0, 5, ∞, 5, SP, text)；
p3:(0, 5, ∞, 5, SP, text)；
p4:(0, 5, ∞, 5, SP, text)；
p5:(3, 20, 17, 17, SP, audio)；
p6:(0, 5, ∞, 10, SP, image)；
p7:(0, 15, ∞, 15, SP, image)
p8:(0, 30, ∞, 30, SP, text)

The algorithm is as follows.

Algorithm:

From places of initial marking M0 in the Reachability Tree to analyze the spatiotemporal conflicts of multimedia objects in overlap duration (including spatial overlap between multimedia objects and two audio played at the same time)

for each $i, j \in N$, where T_0 is total played time before
 if $DEL(p_i) > DEL(p_j)$
 $(DEL(p_j)+\text{MIN}(MPT(p_i), MPT(p_j)))-(DEL(p_i))$ is overlap duration time and interval between $[T_0+DEL(p_i), T_0+(DEL(p_j)+\text{MIN}(MPT(p_i), MPT(p_j)))]$
 else if $DEL(p_i) < DEL(p_j)$
 $(DEL(p_i)+\text{MIN}(MPT(p_i), MPT(p_j)))-(DEL(p_j))$ is overlap duration time and interval between $[T_0+DEL(p_j), T_0+(DEL(p_i)+\text{MIN}(MPT(p_i), MPT(p_j)))]$
 else $DEL(p_i) = DEL(p_j)$
 $(DEL(p_i)+MPT(p_i))-(DEL(p_i))$ is overlap duration time and interval between $[T_0+DEL(p_i), T_0+(\text{MIN}(MPT(p_i), MPT(p_i))]$
 end if
end for

then analyze M1, M2, and M3 till state end

Reachability Tree analysis steps are as follows.

Step 1. Find out the spatiotemporal conflict and overlap duration of Mn states in Reachability Tree. For p1, p2, p5 of M1, the analysis as follows.

Step 2. Find out the overlap duration by above Algorithm.

(a) i=p1, j=p2 DEL(p1)=DEL(p2) → [T0, T0+5]
(b) i=p1, j=p5 DEL(p1)<DEL(p5) → [T0+3, T0+15]
(c) i=p2, j=p1 DEL(p2)=DEL(p1) → [T0, T0+5]
(d) i=p2, j=p5 DEL(p2)<DEL(p5) → [T0+3, T0+5]
(e) i=p5, j=p1 DEL(p5)>DEL(p1) → [T0+3, T0+15]
(f) i=p5, j=p2 DEL(p5)>DEL(p2) → [T0+3, T0+5]

Then remove the same interval, we could only select (a), (b), and (d). So we can show the overlap duration and interval as follows.

Part "(a)": p1, p2 overlap duration is (T0+5)-T0=5 and interval is from T0 to (T0+5),
Part "(b)": p1, p5 overlap duration is (T0+15)-(T0+3)=12 and interval is from (T0+3) to (T0+15),
Part "(d)": p2, p5 overlap duration is (T0+5)-(T0+3)=2 and interval is from (T0+3) to (T0+5).

Then we are able to continue next step.

Step 3. Find out the temporal and spatial conflict between multimedia objects in overlap duration by spatiotemporal conflict and type detecting table (Table 4 and Table 5).

We use spatiotemporal detecting table to check overlap duration, spatial property (SP) and compare to multimedia modalities. Then it will notify that user whether the plan has spatiotemporal conflict or not.

Part "(a)": in overlap duration, p1 and p2 are "Overlap time" and "Different position". By the detecting table, “video” and “text” have no temporal conflict and spatial conflict.

Part “(b)”: in overlap duration, p1 and p5 are “Overlap time” and “Different position”. By the detecting table, “video” and “audio” have temporal conflict but no spatial conflict.

Part “(d)”: in overlap duration, p2 and p5 are “Overlap time” and “Different position”. By the detecting table, “text” and “audio” have no temporal conflict and spatial conflict.

Step 4. Final result

Finally, we can find out p1 and p5 have temporal conflict and the overlap duration is 12 units.

Using the same ways, we can find out “(2) p1, p2, p5 of M2”, “(3) p1, p4, p5, p7, and p8 of M3”, and “(4) p5, p6, p7, p8 of M4”. The analysis result listed in Table 8.

Table 8. Analysis result from reachability tree method

Mn state	Conflicting places	Temporal conflict	Spatial conflict
(1) p1, p2, p5 of M1	p1, p5	Yes, 12 time unit	unknown
(2) p1, p4, p5 of M2	p1, p5	Yes, 12 time unit	unknown
(3) p1, p4, p5, p7, p8 of M3	p1, p5	Yes, 12 time unit	No
	p7, p8	No	Yes, 15 time unit
(4) p5, p6, p7, p8 of M4	p7, p8	No	Yes, 15 time unit

5. Summary

The proposed SMIL-based multimedia authoring environment, which combined SMIL and Petri net facilitate multimedia content generation. Although SMIL can provide excellent markup function on different type of multimedia objects, the spatiotemporal conflicts among objects are very difficult to detect and verify. All these temporal and spatial property of each object are modeled by proposed SMILPN. In this way, the designed content can be verified via Petri net analysis module. All the spatiotemporal conflicts could be effective detected, and then remind author to revise.

References

[1] C. A. Petri, “Kommunikation mit Automaten,” PhD thesis, 1962

[2] Merlin and D. Farber. "Recoverability of communication protocols – implications of a theoretical study,” IEEE Transactions on Communications, 1976.

[3] T. D. C. Little and A. Ghafoor, “Synchronization and Storage Models for Multimedia Objects,” IEEE Journal on Selected Areas in Communications, Vol. 8, No. 3, 1990, pp. 413-427.

[4] M. Woo et al., “A Synchronization Framework for Communication of Pre-Orchestrated Multimedia Information,” IEEE Network, pp.52-61, Jan./Feb. 1994

[5] K. Jensen. “Coloured Petri Nets: Basic Concepts, Analysis Methods and Practical Use. Vol. 1, Basic Concepts.” EATCS Monographs on Theoretical Computer Science, pp 1-234. Berlin: Springer-Verlag, 1992.

[6] Nabil R. Adam et al. “A Dynamic Manifestation Approach for Providing Universal Access to Digital Library Objects” IEEE Transactions On Knowledge And Data Engineering, Vol. 13, No. 4, pp.705-716 July/Aug. 2001

[7] S. M. Chung, A. L. Pereira, “Timed Petri Net Representation of the Synchronized Multimedia Integration Language (SMIL) of XML,” Proceeding of International Conference on Information Technology: Computers and Communications (ITCC’03), Apr., 2003.

[8] Ping-Yu Hsu et al. “STRPN: A Petri-Net Approach for Modeling Spatial-Temporal Relations between Moving Multimedia Object” IEEE Transactions on Software Engineering, Vol. 29, No.1, January 2003.

[9] Yu-Kuang Deng, Reuse and Adaptation Content Model for Multimedia Presentation, dissertation, 2002.

[10] Tadao Murata, "Petri nets: Properties, analysis and applications", Proceedings of the IEEE, Vol. 77, NO. 4, pp.541-580, Apr. 1989.

[11] L. Rutledge, "SMIL 2.0: XML for Web Multimedia" IEEE Internet Computing, pp.78-84, Sep.-Oct. 2001

[12] Stephen J.H. Yang, Norman W.Y. Shao, Kevin C.Y. Kuo, "A SMIL Editor and Rendering Tool for Multimedia Synchronization and Integration", National Computer Symposium 2003 ,Taiwan, page 56-62, Dec. 2003 .

[13] Yang, S.J.H. et al., “Specifying and Verifying Temporal Behavior of High Assurance Systems Using Reachability Tree Logic”, *IEEE 1998 High Assurance Systems Engineering Symposium*, pp. 150-156, Washington D.C., USA.

[14] M. Vazirgiannis et al., “Specifying and Authoring Multimedia Scenarios,” IEEE Multimedia, vol. 6, no. 3, pp.24-37, July-Sept. 1999.

[15]Richard Scott Brink, A Petri Net Design, Simulation, and Verification Tool, appeared in http://www.csh.rit.edu/~rick/thesis/thesis.html.