Interval Expressions - a Functional Model for Interactive Dynamic Multimedia Presentations

Chérif Keramane, Andrzej Duda
IMAG-LSR
BP 53, 38041 Grenoble Cedex 9, France
{Cherif.Keramane, Andrzej Duda}@imag.fr

Abstract

We propose a new model for structured temporal composition of interactive dynamic multimedia presentations. It extends the notion of basic media segments to include executable code, live feeds, and links. The model is based on Interval Expressions *that involve media segments of unknown duration.*

1. Introduction

Multimedia representation and computing have made significant progress during last several years. However, the advances have not led to satisfactory proposals for rich dynamic adaptive story environments [3]. Existing models of multimedia presentations allow static composition of different media having inherent temporal behavior such as video and audio. Temporal composition defines synchronization among media segments according to some static temporal scenario. Usually, a presentation is played back linearly with little (simple Temporal Access Control functions) or no user control over the flow of presentation. In opposition to the traditional multimedia presentations, we are interested in a new class of multimedia presentations that include rich dynamic interactive scenarios. Such scenarios integrate dynamic user control over the flow of presentation, advanced processing of media content, and diverse sources of media streams such as live feeds or teleconferencing streams.

Existing temporal models for multimedia can be divided into two classes: *point-based* and *interval-based*. In point-based models, the elementary units are points in a time space. Each event in the model has its associated time point. An example of the point-based approach is *timeline*, in which media objects are placed on several time axes called *tracks*, one per each media type. The timeline model is applied in HyTime [5] and in other propositions. The model is well suited for temporal composition of media segments of known durations, however it falls short for unknown durations. Some authors have proposed to use relations between interval end points for temporal composition of multimedia. The temporal point nets [2] can deal with intervals of unknown duration, but their use is difficult and results in complicated, unstructured graphs. In addition to that, their use may lead to an inconsistent specification in which contradictory conditions are specified for intervals. In this case, a verification algorithm must check for temporal inconsistency.

Interval-based models consider elementary media entities as time intervals ordered according to some relations. Existing models are mainly based on the relations for expressing the knowledge about time [4, 1]. Giving any two time intervals, they can be arranged according to seven relations : *before, meets, overlaps, finishes, during, starts, equals*. A relation expresses the inequalities between the end points of intervals. If the relations are used for multimedia composition, then the duration of intervals may vary only within the limits defined by the inequalities of a given relation. This drawback makes the relations not suitable for specifying composition of intervals with unknown duration. Consider for example an existing relation *before* between intervals a and b (see Figure 1). When we increase the duration of interval a, the relation changes from *before* to *during* passing through intermediate relations *meets*, *overlaps*, and *finishes*.

Another problem with the relations is their descriptive character—they allow expression of an existing, *a posteriori* arrangement of intervals, but they do not express any causal or functional relation between intervals. For example, relation *meets* is ambiguous—it only states that the end of the first interval coincides with the end of the second one, but it does not say whether the first interval *starts* the second one, whether the second interval *stops* the first one or whether it is a pure coincidence. So, the relations can be useful for characterizing an existing, *instantiated* presentation (a presentation for which all start and termina-

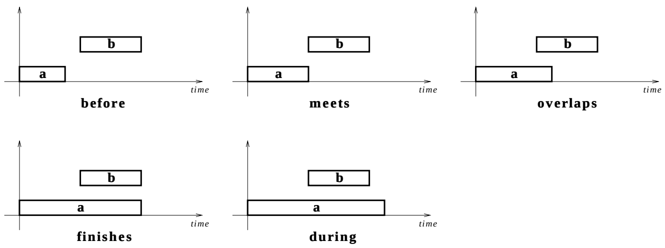

Figure 1. Relations change when the interval duration is modified.

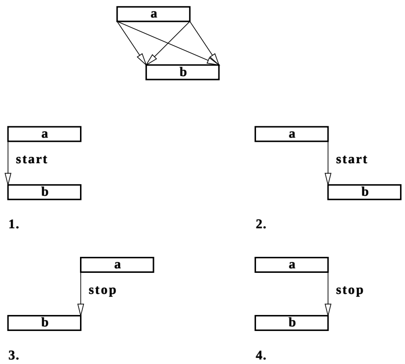

Figure 2. Four possible relations between intervals.

tion instants of media segments are known). The third problem with the relations is related to inconsistent specifications that can be introduced to a multimedia presentation. Detecting inconsistent specification requires algorithms of complexity $[O(N^2)]$, where N is the number of intervals [1, 6].

Many existing representations are conceptually equivalent to the Allen relations, for example OCPN [7]. Several other models have been proposed: LMDM [9], FLIPS [8], and TIEMPO [10]. Unlike previous models, *Interval Expressions* offer a different paradigm that allow specification in a structured manner of functional relations between intervals of unknown duration. We extend the notion of basic media segments to include executable code, live feeds, and links. For temporal composition, we explore the same approach as *Algebraic Video* [11] that used a set of basic operations on video segments. Temporal composition based on algebraic construction of nested, encapsulated expressions supports sharing and reusing of media content.

2. Interval Expressions Model

The elementary entities of our model are *media segments* that have inherent temporal behavior. They may be traditional multimedia objects containing video, audio, animations, images or text as well as new special objects such as live media streams, executable code or links to other presentations. For temporal com-

Figure 3. Two relations for equality.

position, media segments are viewed as *time intervals* independently of their content or other characteristics. Time interval a is defined by its end points ($\underline{a} \leq \bar{a}$) as $a = \{t \mid \underline{a} \leq t \leq \bar{a}\}$. We define the following types of media segments:

- Set M of traditional multimedia objects such as video, audio, animations, images, text, or pure delay. Segments of type M are time intervals with duration that is either *implied*, i.e. determined by the rate of their capture or *synthetic*, i.e. a delay is associated with a given segment.
- Set S of live media streams such as live feeds, teleconferencing streams, or even complex presentations that arrives on a network connection.
- Set P of programs. When activated, a program begins its execution. A program can take into account user interactions.
- Set L of links. A link refers to another media segment that will be activated when a link is activated.
- Set $\emptyset$ of empty segments.

Given these different types, we define the basic set I_{base} of media segments as follows:

$$I_{base} = M \cup S \cup P \cup L \cup \emptyset \qquad (1)$$

It represents all media segments that can be used as arguments in interval operators described below.

Let us consider relationships between any two media segments considered as time intervals (see Figure 2). Each end point of an interval may perform an action on any of the end points of another interval: the action aimed at the beginning of an interval corresponds to *start* or activation and the action aimed at the end of an interval corresponds to *stop* or termination. The actions represent *causal* relations that have real system interpretation.

There are four possible relations with a single action involving a and b (shown in Figure 2) and four inverse relations involving b and a. In addition, we can consider relations with more than one action. Many of such relations are trivial, for example relations $\underline{a} \longrightarrow \underline{b}$ and $\underline{a} \longrightarrow \bar{b}$ make the duration of b null. There are two non trivial relations concerned with equality of interval durations (see Figure 3). In the first relation, a *starts* b

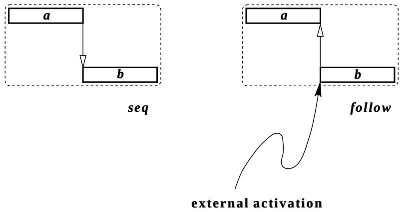

Figure 4. Sequential temporal composition operators

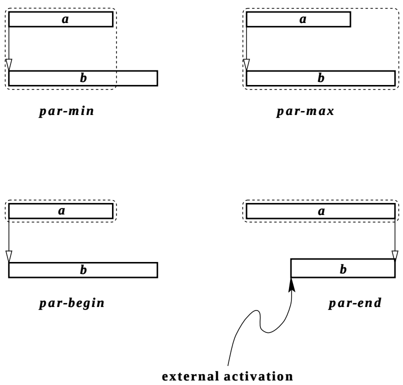

Figure 5. Parallel temporal composition operators

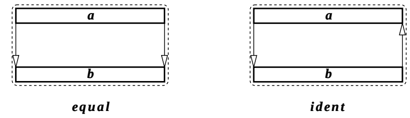

Figure 6. Equality temporal composition operators

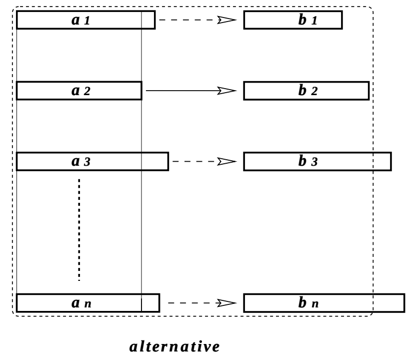

Figure 7. Alternative temporal composition operator

and a *stops* b. In the second one, a *starts* b and b *stops* a.

Given these possible relations between intervals, we define a high-level model that provides a means for encapsulation and structuring. It is based on the following functional operators defined below (for each operator, we give its semantics and the definition of the result interval):

`seq` $(a,\ b) \mapsto (\underline{a}, \bar{b})$ defines composition in which the end of interval a *starts* interval b.

`follow` $(a,\ b) \mapsto (\underline{a}, \bar{b})$ defines composition in which the beginning of interval b *stops* interval a; interval b is activated externally.

`par-begin` $(a,\ b) \mapsto (\underline{a}, \bar{a})$ defines composition in which the beginning of interval a *starts* interval b.

`par-end` $(a,\ b) \mapsto (\underline{a}, \bar{a})$ defines composition in which the end of interval a *stops* interval b; interval b is activated externally.

`par-min` $(a,\ b) \mapsto (\underline{a}, \min\ (\bar{a}, \bar{b}))$ defines composition in which the beginning of interval a *starts* interval b; the result interval is stopped when the first of the two interval terminates.

`par-max` $(a,\ b) \mapsto (\underline{a}, \max\ (\bar{a}, \bar{b}))$ composition in which the beginning of interval a *starts* interval b; the result interval is stopped when the last of the two interval terminates.

`equal` $(a,\ b) \mapsto (\underline{a}, \bar{a})$ defines composition in which interval a *starts* and *stops* interval b.

`ident` $(a,\ b) \mapsto (\underline{a}, \bar{b})$ defines composition in which the beginning of interval a *starts* b and the end of interval b *stops* a.

`alternative` $((a_1, b_1), \ldots, (a_n, b_n)) \mapsto (\underline{a}_1, \bar{c})$ defines the composition in which intervals $a_i,\ i = 1, \ldots, n$ are *started* in parallel; the end of the first of them *stops* all parallel intervals and *starts* its associated interval b_i: $\bar{c} = \bar{b}_j$ where $\bar{a}_j = \min_{i=1,\ldots,n}(\bar{a}_i)$.

`loop` $a \mapsto (\underline{a}, \infty)$ defines composition in which the interval a is repeated for ever and can only be stopped by other intervals.

The operators are represented graphically in Figures 4, 5, 6, 7. Dashed boxes show the result interval for each operator (encapsulation boundary).

An operator takes one, two, or more intervals as arguments and returns an interval as a result. `alternative` operator defines a correspondence between two sets of intervals: each interval of the first set has its corresponding interval in the second set. When `alternative` begins, all intervals of the first set behave as `par-min` with n arguments: they are activated in parallel and the shortest interval terminates all parallel intervals. Then, it selects its corresponding interval from the second set. The operator allows composition of conditional presentations in which the user can choose the flow of presentation.

Complex temporal multimedia compositions can be built by nesting operators. The set I of interval expressions that define a temporal multimedia composition is build from the set of media segments I_{base} by induction according to the following rule:

$$I_0 = I_{base} \quad (2)$$

$$I_{n+1} = \bigcup_{f \in \mathcal{F}} f(I_n^{ar(f)}) \cup alternative(\mathcal{P}(I_n \times I_n)) \quad (3)$$

$$I = \bigcup_{n=0}^{\infty} I_n \quad (4)$$

where $\mathcal{P}(X)$ is the set of all subsets of set X, $\mathcal{F}$ is the set of the operators except **alternative**, and $ar(f)$ the arity of operator f.

The operators can be used to form nested *Interval Expressions* allowing specification of temporal compositions by means of encapsulation in a well-structured way. Nesting allows building more complex hierarchical well-structured presentations and encourages sharing and reuse of parts of presentations.

The structure defined by an interval expression is invariant with respect to the duration of component intervals. This means that an expression defines the temporal structure of a presentation. Its playback may result in many different temporal schedules depending on the duration of media segments. However, relations between segments are always maintained according to the temporal structure.

By construction, our model guarantees the absence of temporal inconsistencies. As each result interval is temporally consistent with the arguments of an operator, nesting preserves temporal consistency.

Forming interval expressions provides encapsulation boundaries. In this way, the result interval forms an autonomous entity separated from outside. For example, when a *starts* b (Figure 2), the result interval is defined by the end points $(\underline{a}, \bar{b})$. The activation of the result interval implies activation of a that in turn causes activation of b.

Operators `follow` and `par-end` require external activation. For example, relations 3 and 4 (Figure 2) in which a *stops* b can only be encapsulated, if we assume that interval b is of type S (or is a complex expression that begins with a media segment of type S), because for this type of media segments the instant of its activation is determined by the source of a media stream independently of the initial activation of interval a. So, the result interval for operators `follow` and `par-end` provides encapsulation boundary, however the operators require the second argument to be a live media stream or a complex expression beginning with a live media stream.

3. Conclusions

We have presented a new model for structured temporal composition of interactive dynamic multimedia presentations. Our work differs from earlier works in several major aspects. We have extended the notion of basic media segments to include executable code, live feeds, and links. In this way, we can take into account user interactions, content-sensitivity, new interesting sources of multimedia data, and provide support for sharing and reuse. These new media segment types are integrated in a seamless way within our temporal composition model. The model is based on *Interval Expressions* that involve media segments of unknown duration. We have defined a set of operators that express causal relations between intervals. *Interval Expressions* provide a means of encapsulation and structuring: compound encapsulated intervals can be specified in terms of elementary media objects as building blocks. The model addresses the temporal consistency problem—*Interval Expressions* guarantees the absence of temporal inconsistencies by construction.

References

[1] J. F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11), Nov. 1983.

[2] M. C. Buchanan and P. T. Zellweger. Automatic temporal layout mechanisms. In *Proc. First ACM International Conference on Multimedia.*, pages 341–350, Anaheim, CA, Aug. 1993.

[3] G. Davenport. Still seeking : Signposts of things to come. *IEEE Multimedia*, 2(3), 1995.

[4] C. Hamblin. Instants and intervals. In *Proc. of the 1st Conf. of the Intl. Society for the Study of Time*, pages 324–331, New York, 1972.

[5] ISO. Information technology hypermedia/time-based structuring language (HyTime). *ISO International Standard*, (ISO/IEC IS 10744), Aug. 1992.

[6] N. Layaida and C. Keramane. Maintaining temporal consistency of multimedia documents. In *Effective Abstractions in Multimedia Layout, Presentation and Interaction ACM 95 Workshop*, San Francisco, CA, 1995.

[7] T. Little and A. Ghafoor. Interval-based temporal models for time-dependent multimedia data. In *Proc. IEEE Conference on Data and Knowledge Engineering*, Lake Buena Vista, FL, Aug. 1993.

[8] J. Schepf, J. D. Konstan, and D. Du. Doing flips : Flexible interactive presentation synchronization. In *Proc. 1995 International Conference on Multimedia Computing and System*, Washington, DC, May 1995.

[9] G. A. Schloss and M. J. Wynblatt. Building temporal structures in a layered multimedia data model. In *Proc. Second ACM International Conference on Multimedia.*, pages 271–278, San Francisco, CA, 1994.

[10] T. Wahl, S. Wirag, and K. Rothermel. Tiempo : Temporal modeling and authoring of interactive multimedia. In *Proc. 1995 International Conference on Multimedia Computing and System*, Washington, DC, May 1995.

[11] R. Weiss, A. Duda, and D. Gifford. Composition and search with a Video Algebra. *IEEE Multimedia*, 2(1), 1995.