

# A Temporal Reference Framework for Multimedia Synchronization<sup>1</sup>

M.J. Pérez-Luque<sup>2</sup> and T.D.C. Little

Multimedia Communications Laboratory  
Department of Electrical, Computer and Systems Engineering  
Boston University, Boston, Massachusetts 02215, USA  
(617) 353-9877, (617) 353-6440 fax  
*tdcl@bu.edu*

MCL Technical Report 04-15-1995

**Abstract**—The synchronization problem for audio/visual reproduction has consumed engineers since the advent of recorded audio and images and the first multimedia productions. As computers have evolved to support programmed reproduction of multimedia information, the complexity but not the character or interest have changed. In the digital domain, synchronization problems exist due to data distribution and communications, random events caused by human-computer interaction, and general computer and communications performance limitations. This rich research domain has led to numerous approaches to the modeling and execution of multimedia synchronization scenarios. Unfortunately, these approaches are difficult to compare and evaluate due to their varied theoretical bases and modeling techniques.

In this paper we develop a uniform, theoretical foundation for discussing multimedia synchronization and temporal specification. We propose a temporal reference framework and use it to compare existing temporal specification schemes and their relationships to multimedia synchronization. The ensuing comparison of existing specification and synchronization techniques demonstrates the utility of the framework.

---

<sup>1</sup>In *Proc. 1st Intl. Workshop on Multimedia Synchronization*, Tysons Corner, May 1995. A longer version appears in the *IEEE Journal on Selected Areas in Communications* (Special Issue: Synchronization Issues in Multimedia Communication), January 1996. This work is supported in part by the National Science Foundation under Grant No. IRI-9211165.

<sup>2</sup>Work performed while visiting from the Grupo de Tratamiento de Imágenes, E.T.S. Ingenieros Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain. Dr. Pérez-Luque's current address is the Laboratorio de Comunicación Multimedia, Facultad de Ciencias de la Información, Universidad de Navarra, 31080 Pamplona, Spain, *mjpl@mmlab.unav.es*.

**Keywords:** Models of time, temporal specification, temporal reference framework, multimedia synchronization.

## 1 Introduction

Multimedia synchronization is about providing coherent playout of orchestrated audio and visual information. It is a controversial topic in the multimedia computing community. On the one hand, providing “lip-sync” or “slide-show” capability requires no more than fast buses, buffering, and hardware decompression devices. On the other hand, complex multi-party human-computer interaction coupled with live and preorchestrated presentations are difficult to deal with in any formal or ad hoc manner. For these scenarios (e.g., the construction of virtual environments), the behavior of each participant must be understood and represented in the temporal domain. Once such a characterization is achieved, the nuances of the operating environment (communications, operating system, and playout devices) must be dealt with to achieve the desired multimedia presentation outcome.

For this reason, we divide the general multimedia synchronization problem into two parts. The first is one of modeling, representing, and specifying timing requirements of multimedia scenarios. The second problem is to achieve the temporal specification via synchronization. In this decomposition, synchronization reduces to the rendering of a temporal specification using some technique. In this paper we focus on the former problem; that of the specification of timing requirements. However, we also show the relationship between temporal specification and the rendering of a temporal specification (synchronization).

To appreciate the temporal specification problem (modeling and representing included) requires a basic understanding of formal specification techniques. These techniques isolate a *semantic domain* (i.e., the meaning of a specification and language) and a *syntactic domain* (i.e., the alphabet, and grammar of the specification language). Theory focuses on the semantics; however, the theory must be supported by models which require tools for visualization. These visualization tools exist as a syntax. Therefore, when we think we deal only with semantics, we are often crippled by a syntax as well. We have found this problem to arise quite frequently when considering models of time and their use in specifying *temporal scenarios*—sequences of events in time.

Consider the analogous problem of specifying or representing physical objects by computer. Computers can provide artificial representations of our real world (e.g., three di-

mensional rendering of objects on a computer screen). For obtaining these representations we need to find models to capture the nature of the real objects, mechanize the models, and find algorithms to transform (render) the mechanized models (representations) into the display space of the computer. Usually, more than one model can be found to represent a single object from the real world. For example, Fig. 1 shows simple models characterizing a table from the real world. The left-hand side describes the table. The middle illustrates three models and their corresponding representations. The right-hand side shows the object rendered on a computer display. The example illustrates different geometric modeling and representation approaches that can yield equivalent renderings of the synthesized object.

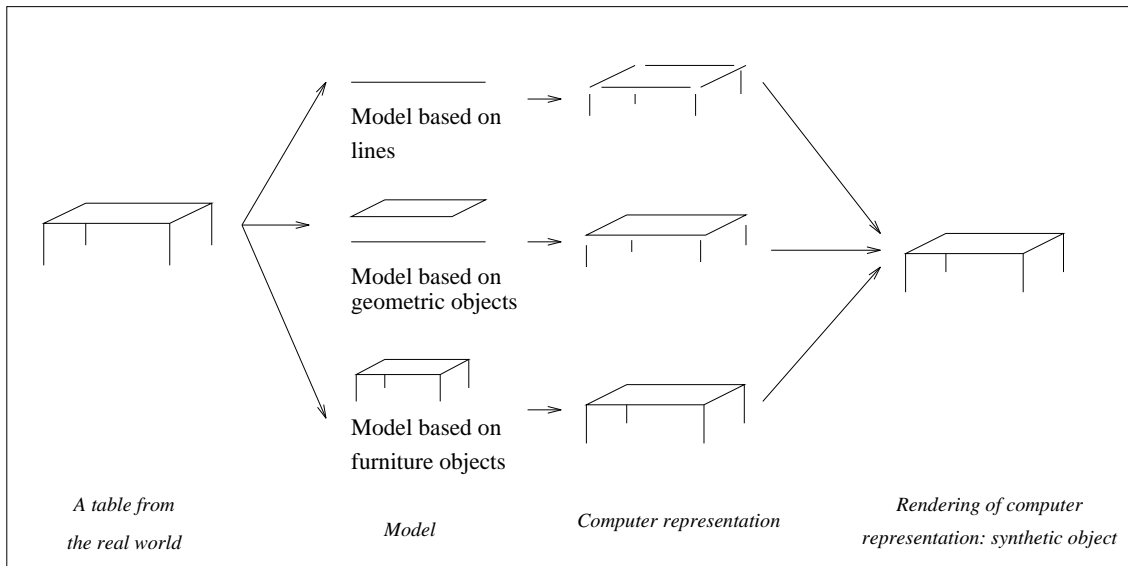


Figure 1: Geometric Representations of a Table

Time is another aspect of the real world that can be modeled. As in the case of the spatial example of Fig. 1, several models of time can be used to represent the same sequence of events. Consider Fig. 2 as an example of a short movie that has been recorded from the real world. We present four possible ways to describe the sequence of events in the time domain.

1. **An English description:** The car was initially stopped. Subsequently, the traffic light turned green and the car began moving. The car remained in motion until a gendarme signaled it to stop.
2. **A time of occurrence description:** At  $t = 0$ , the stopped car appears (beginning of the movie). At  $t = 5$ , the traffic light turns green. At  $t = 5$ , the car starts moving. At

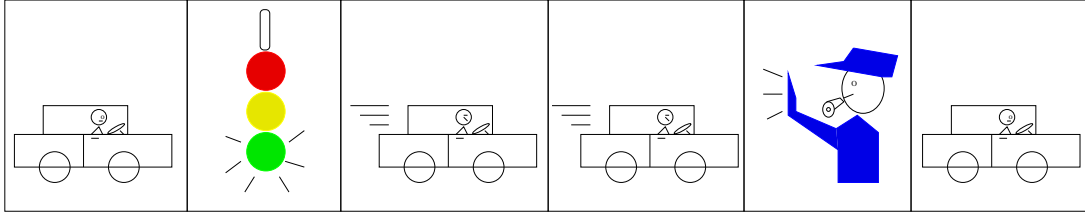


Figure 2: Frames Representing a Movie Recorded from the Real World

$t = 25$ , a gendarme signals the car to stop and the car stops immediately. At  $t = 29$ , the movie ends.

3. **An instant-based/temporal relationship description:** The car started moving at the same time the traffic light turned green. The car stopped moving at the same time the gendarme signaled the car to stop.
4. **An interval-based/temporal relationship description:** The car was initially stopped for 5 seconds; the traffic light was red for the same 5 seconds. After the first period, the car traveled for 20 seconds. Subsequently, the car stopped for 4 seconds (until the end of the movie) while the gendarme signaled the car to remain stopped.

Like the rendering of the table (Fig. 1), any of the descriptions above will lead to an appropriate temporal sequence for the movie, provided that we have suitable rendering mechanisms. This example illustrates how a single scenario can be modeled by a range of different approaches (models of time). The ensuing representations can be very different yet yield the same results. In the example, four different models capture the same scenario.

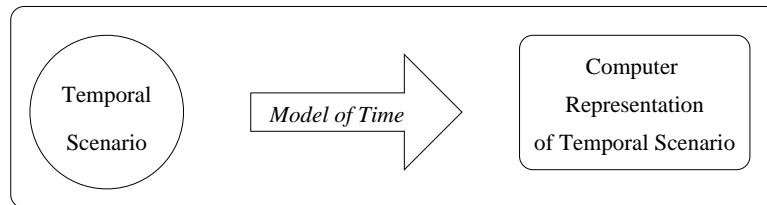


Figure 3: Modeling Temporal Information

The process of modeling temporal information is represented in Fig. 3. The left-hand side captures the *temporal scenario* that we seek to model. The right-hand side captures the representation of the scenario. The transformation from the scenario to the representation is

achieved by various temporal modeling techniques. Once a representation is achieved, mechanisms are required for final rendering in the time domain. For multimedia data presentation, this is equivalent to multimedia synchronization.

Our contribution is the development of a *temporal reference framework*: a uniform, theoretical basis for discussing multimedia synchronization and temporal specification. The proposed framework is most useful in the comparison and synthesis of temporal specification schemes. The remainder of the paper deals with the modeling and rendering process for temporal information in multimedia systems. In Section 2 we characterize temporal scenarios and present our terminology for models of time. Section 3 describes the proposed temporal reference framework. In Section 4 we show how the framework can be used to compare various existing specification approaches for multimedia object timing. Section 5 concludes the paper.

## 2 Modeling Temporal Information

To develop a unified theory for multimedia synchronization and representation, we require common terminology. Here we develop and define the basic terms that we use in the framework development, introduce and characterize the models of time used for multimedia specification, and present representations suitable for mechanization. First we consider a definition of the temporal activities we wish to model within the realm of computer-based multimedia applications.

### 2.1 Temporal Scenarios

A **Temporal Scenario** represents an instance of a set of activities that are in some way related in time. For example, consider the temporal scenario describing an instructor lecturing with slides. The time that each slide stays in the projector depends on the time expended in explanation. In this way, the activity called “slide in projector” is dependent on the activity called “instructor explains slide.” These two activities are *related* in time as well as *dependent* on time. This relation/dependency must be captured by a model of time.

There are two types of temporal scenarios: **determinate**, and **indeterminate**. Determinate temporal scenarios represent a set of real-world activities. As they are real-world activities, the temporal constraints are completely defined and determined by the physical

world. There is no indeterminacy in the temporal scenario. Fig. 2 is an example of such a scenario.

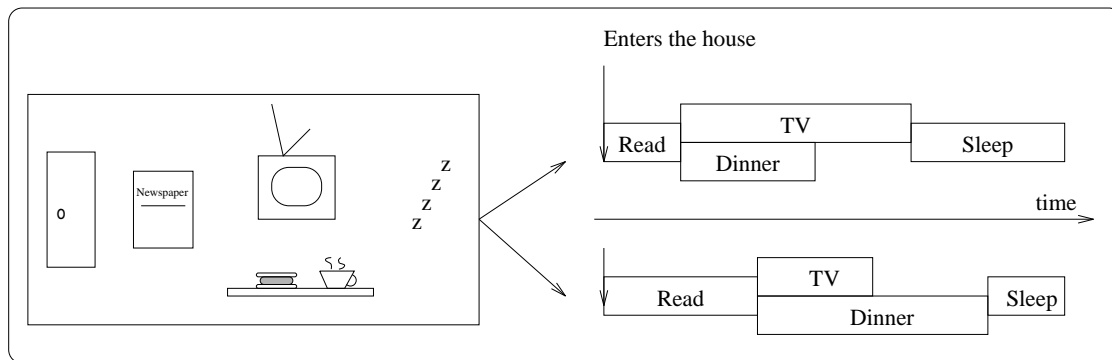


Figure 4: Example of an Indeterminate Temporal Scenario

For indeterminate temporal scenarios there is uncertainty or non-determinism. Fig. 4 shows an example of such a case. The left-hand side of the figure describes a symbolic description of an indeterminate temporal scenario. It consists of a common sequence of activities that an individual might do in the evening: opening the door, reading the newspaper, watching TV while eating dinner, and sleeping. This temporal scenario is the same every evening, but each day it has different realizations (e.g., arrival time, time spent reading the newspaper). Two of these realizations are illustrated on the right hand side of the figure. In these cases we need models of time that can deal with indeterminate cases.<sup>3</sup> When generating a model, we have sufficient information to describe the indeterminate temporal scenario, but the final realization is not specified or determined. Therefore, the description of the temporal scenario represents a superset of all possible final realizations. Ultimately, the final rendering or realization of the temporal scenario depends on external factors (e.g., human interaction). These cases of indeterminacy contrast scenarios that are determinate and therefore possess a unique temporal realization (e.g., the movie of Fig. 2).

Note that a temporal scenario does not imply a specific temporal model.<sup>4</sup> Temporal models should be able to capture the relationships and dependencies among activities for both determinate and indeterminate temporal scenarios. We now consider models of time for translation of temporal semantics to formal specification languages.

<sup>3</sup>Indeterminate temporal scenarios are studied because in multimedia systems there are two clear sources of indeterminacy: user interaction, in which the final realization is not known until the user interacts, and system failures, where the final realization is not known until run-time.

<sup>4</sup>*Model of time* and *temporal model* are used interchangeably throughout the paper.

## 2.2 Models of Time

*Activities* in the temporal scenario are called **events**. An **event** is an occurrence in time that can be instantaneous or can occur over some time period. The temporal scenario can be described either as a set of independent events (e.g., *I phoned you at 6:00*) or as a description that captures some temporal relationships among the events (e.g., *She whas cooking while I phoned you*).

In order to capture and describe the temporal scenario, we require the services of a model of time. The use of such model forces us to stray from a discussion of only semantics (i.e., we must choose some syntax).<sup>5</sup> A model of time can be viewed as the temporal semantics that are applied to yield a formal specification technique, language, and representation. Typically, these semantics are associated with a particular syntax for visualization purposes.

We characterize a **model of time** through three related components: the **basic time unit** of the model, the **contextual information** associated to the basic time units and, the **type of time representation Technique** expressing the basic units and their associated information. These three concepts completely describe a specific model of time and its power of expressivity.

**The basic time unit.** The basic time unit is the temporal unit used in characterizing a temporal scenario. There are two types of basic units: instants and intervals, defined as follows:

**Definition 1** *An instant is a zero-length moment in time [12].*

**Definition 2** *Let  $[S, \leq]$  be a partially ordered set, and let  $a, b$  be any two elements (time instances) of  $S$  such that  $a \leq b$ . The set  $\{x \mid a \leq x \leq b\}$  is called an **interval** of  $S$  denoted by  $[a, b]$ . Furthermore, any interval  $[a, b]$  has the following properties [22]:*

1.  $[a, b] = [c, d] \iff a = c \wedge b = d$
2. if  $c, d \in [a, b], e \in S$  and  $c \leq e \leq d$  then  $e \in [a, b]$
3.  $\#([a, b]) \geq 1$

---

<sup>5</sup>Recall from Section 1, formal language of specification technique is comprised of syntactical and semantic components.

The events of the temporal scenario are modeled through the basic time unit. time. The mapping from the temporal scenarios to the models of time is described by the following: (a) Models of time with intervals as basic time units can only model interval events; (b) interval events. Models of time with instants as basic time units can model interval, instants, and interval & instants together.<sup>6</sup>

**Contextual Information.** The contextual information comprising the model of time specifies the type of temporal information that can be associated with the basic time units of the model. The contextual information is a key principle in the selection of a model of time that is capable of expressing the semantics of a temporal scenario (relations/dependencies among events and indeterminacy included). The contextual information can be described as either quantitative or qualitative:

**Quantitative information.** Quantitative information is temporal information that can be expressed in time units (e.g.,  $t_1 = 6$  pm;  $\#[a, b] = 3$  hours). Quantitative information can refer to any temporal axis (*absolute* or *relative*) and can be expressed using a real or virtual measure of time (e.g., seconds, minutes, or bits in a constant-bit-rate stream).

**Qualitative information.** Qualitative information is the temporal information that is not quantifiable. For example, total and partial orderings of basic time units represent qualitative information. Other qualitative information relating basic time units are the following:

The **basic binary temporal relationships between instants** define the three ways in which two instants can be related [27]: *before*, *after*, and *at\_the\_same\_time*. Only one basic temporal relation can hold between two instants.

The **basic binary temporal relationships between intervals** define the thirteen ways in which two intervals can be related [12]. Only one basic temporal relation can hold between two intervals. These relationships are illustrated elsewhere [22]. Similarly, ***n*-ary temporal relationships among intervals** define the ways in which  $n$  intervals can be related in time [24].

**Indefinite temporal relationships** are those temporal relations between basic time units that are not explicitly or unambiguously given. Usually, these indefinite relations

---

<sup>6</sup>We could also think of modeling instants through intervals and consider the possibility of modeling interval & instants together through intervals. However, modeling instants through intervals is not very useful. Moreover, modeling instants and intervals through intervals is not always achievable. To the knowledge of the authors, there are no temporal techniques developed for this approach and we do not pursue the topic further.



are expressed as disjunctions of the basic temporal relationships. The number of indefinite relationships is calculated as the set of all possible disjunctions of the basic ones. For instance, there are three basic binary relations between instants, therefore, the number of indefinite relations is  $2^3 = 8$ . Two examples are the following {before or at the same time}, {at the same time or after}. For intervals, there are thirteen basic binary relations between intervals. Therefore, the number indefinite relations is  $2^{13} = 8192$ . Three examples are the following {starts or equals}, {finishes or overlaps}, {meets or after or overlaps}. An enumeration of these indefinite temporal relationships has been performed by Ladkin [19] and van Beek [26].

**Duration relationships** describe how the durations of two temporal intervals can be related [1] (e.g.,  $[a, b]$  is **shorter** than  $[c, d]$  iff  $\#[a, b] < \#[c, d]$ ). Duration relationships are usually specified in conjunction with temporal relationships among intervals.

**Time Representation Techniques** A time representation technique describes how time can be captured and mechanized in a computer environment. A particular representation occurs as a result of the application of a model of time. This is a concept that is intrinsically associated with the basic time units and contextual information of a model of time, i.e., a particular model of time will produce a specific time representation and will indicate the theory that will be required to interpret a resultant temporal specification.

To identify the type of time representation technique we follow Allen’s terminology and classification [2]. There are three types of time representation techniques: dating schemes, constraint propagation schemes, and duration schemes.

### 2.2.1 Classification of Models of Time

The three previous concept, the basic time unit, contextual information, and type of time representation, completely describe a *specific* model of time. There are five general types of models of time to which we have assigned names (Fig. 5). These are: (1) Quantitative Dates, (2) Qualitative Dates, (3) Qualitative Instants, (4) Qualitative Intervals, (5) Quantitative Intervals.

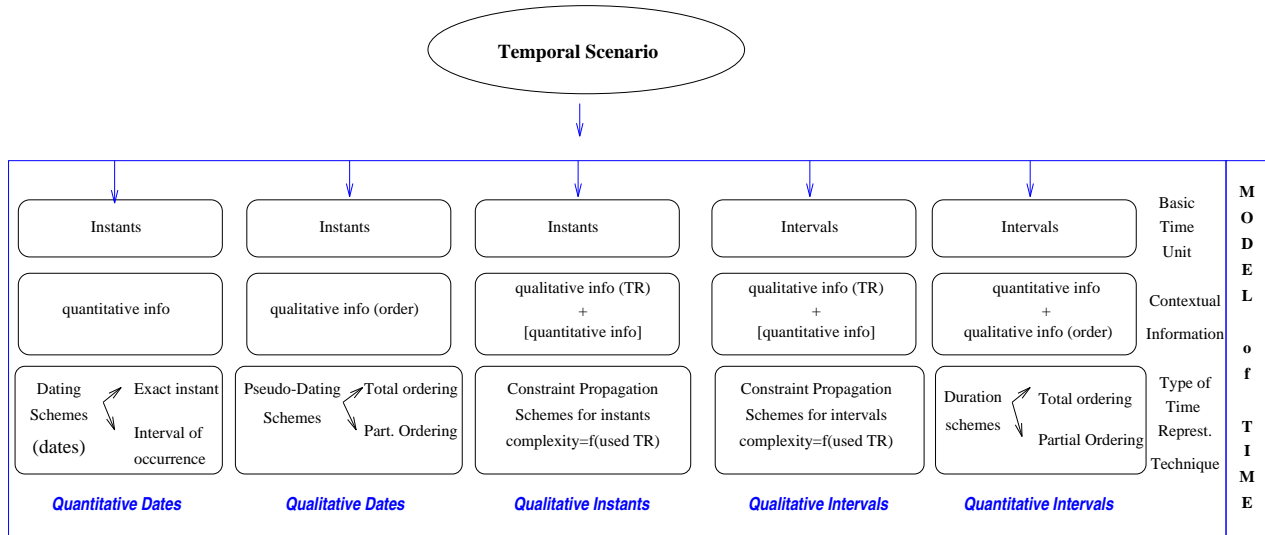


Figure 5: General Models of Time

### 2.2.2 Expressive Power of the Models of Time

The previous section described five general classes of models of time from which a temporal specification scheme can be derived. However, when one seeks to model a particular temporal scenario one could ask: how can I select a specific model of time for a given temporal scenario ? That question leads us to examine the expressive power of models of time and the equivalences between models of time (i.e., they have the same power of expressivity).

The expressive power describes the ability of a model of time to represent temporal scenarios. It can tell us about the expression of semantics for the temporal scenario and the type of events that can be modeled; the way relations/dependencies among events and the indeterminacy in the temporal scenario are captured by the model. The expressive power of the general models of time can be delineated, but each model must be supplemented by the specific contextual information.

Two or more models of time can be used to capture the same multimedia scenario, but that does not mean that they are completely equivalent. The equivalence of the temporal semantics must be studied mathematically by comparing the basic time unit and the contextual information for each model of time.

Some studies have been undertaken to determine equivalences among existing models of time. These demonstrate mathematically the complete or partial equivalence of some models.

For example, the equivalence between qualitative instants and qualitative intervals models in which the contextual information consists of temporal relations is extensively studied by Ladkin [19] and van Beek [26] (e.g., the indefinite temporal relations between instants is equivalent to 181 indefinite temporal relations between intervals). Wahl and Rothermel [28] have also studied these equivalences (Section 4.1).

## 2.3 Computer Representations

The goal of modeling temporal information is to find techniques for representing temporal scenarios for subsequent computer rendering and reproduction (Fig. 3). Following our terminology, we say that a mechanized representation corresponds to a *temporal specification*. More formally, a temporal specification is defined below.

**Definition 3** *A temporal specification is a well-formed sentence satisfying a particular temporal specification language.*

A specification language is the result of the definition of semantic and syntactic domains. Section 2.2 deals with the temporal semantics of the modeling process. Here we introduce the concept of syntactic domain and its influence in the resulting specification language.

The syntactic domain is defined in terms of a set of symbols (constants, variables, operators, etc.) and a set of grammatical rules. These symbols can either be textual or graphical in nature. The external appearance of a specification language corresponds to the combination of the syntactic symbols following the syntactic grammar. Many different specification languages can and have been defined for temporal representations because (1) different temporal semantics can be used for the same temporal scenario and (2) the same semantics can be represented in very different syntactic domains.

As long as the semantics of a specification language are correct, we can use any of the representations to describe a particular temporal scenario. Therefore, many different temporal specifications can be found for the same temporal scenario. In Section 4 we compare some of these representations.

In synopsis, the concept of a model of time has been introduced as influenced by three attributes: the basic time unit, contextual information, and the type of time representation technique. From these attributes, models of time are classified in five general categories. The ability of a model of time to describe a temporal scenario is determined by the expressivity of

the model. The general class to which the model belongs provides a rough characterization. However, its specific expressiveness can be identified. More than one model of time can be found for the description of a concrete temporal scenario. Complete equivalence of two models of time must be demonstrated mathematically. If two models of time are completely equivalent, a translation between their representation is achievable. Finally, the syntactic domain must be selected to achieve a specification language for a model of time, and this can be achieved with diverse syntax.

### 3 Temporal Reference Framework

Based on the terminology and concepts presented in Section 2, we now introduce our framework for the modeling and comparison of multimedia specification and synchronization techniques.

Multimedia synchronization techniques have been developed for the satisfaction of multimedia scenarios specified in some temporal language. In the design of these techniques, there are three processes that are clearly different (1) the definition of the specification language for a set of temporal scenarios, (2) the description of the multimedia scenario with the specification language, and (3) the creation of the synchronization mechanisms to render descriptions within the specification language. To clarify the relationship between a multimedia scenario and a temporal scenario we present the following definition:

**Definition 4** *A multimedia scenario is the spatial and temporal semantics of a multimedia presentation session, i.e., the orchestration of media in time and space that should be delivered to a user.*

A temporal specification language is used to provide a temporal specification for a multimedia scenario.<sup>7</sup> The synchronization mechanisms try to achieve the multimedia scenario that is described in the temporal specification. These terms and their relationships are illustrated in Fig. 6 and are defined as follows.

**Definition 5** *A temporal specification scheme is a specification language that is used to describe multimedia scenarios.*

---

<sup>7</sup>In the remainder of the paper we only consider the temporal aspect of multimedia scenarios.

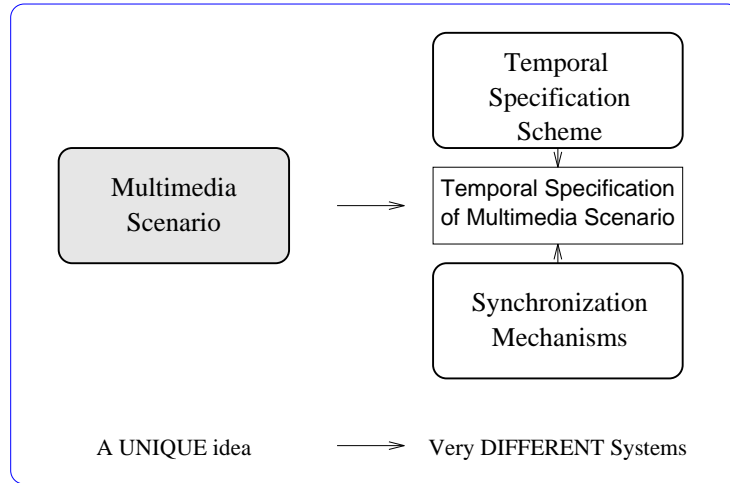


Figure 6: General View of Multimedia Synchronization Techniques

This specification language is defined by a semantic and a syntactic domain. The temporal specification scheme is equivalent to the time representation scheme. In multimedia definitions, the former usually is not restricted to temporal information, and we distinguish the two definitions. We also reiterate our distinction between specification and satisfaction of a specification in the definition of a synchronization mechanism.

**Definition 6** *A synchronization mechanism is a service or process that realizes a multimedia scenario.*

The synchronization mechanisms can be one that interprets general temporal specifications (e.g., from authoring tools) or one that is restricted to particular applications (e.g., lip-sync in audio and video). These mechanisms are varied and depend on the infrastructure, the architecture of the system, and the specific media types. Therefore, although they each share the common objective of satisfaction of a temporal specification (the multimedia scenario), must combine temporal information with other types of atemporal information in different ways.

Fig. 7 shows the integrated view of the temporal reference framework. On the left-hand side of the figure are the multimedia definitions; on the right-hand side are the corresponding temporal definitions.

In a distributed multimedia system (DMS), the synchronization is achieved by the combined effort of the different parts of the system (databases, network, operating systems,

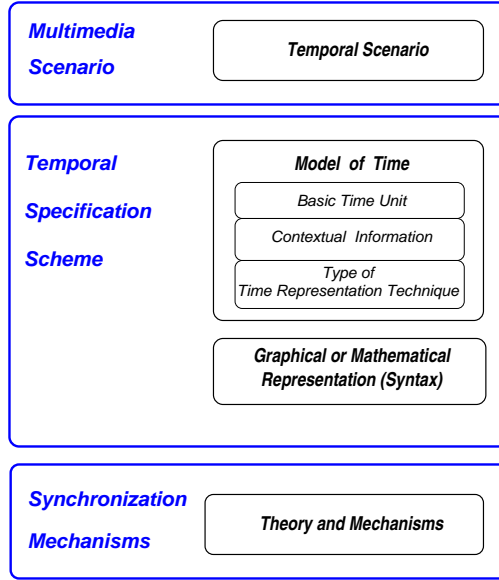


Figure 7: The Temporal Reference Framework

etc.). Our temporal reference framework is also useful in the analysis of the components of distributed multimedia systems. It can help to identify the temporal specification schemes of each component and the exchange of temporal information among the different parts of the system.

## 4 Use of the Framework for Analysis and Comparison of Approaches

In this section, we analyze a number of temporal specification and synchronization techniques that have appeared in recent literature. Our analysis, based on the temporal reference framework, focuses on models of time and begins with an investigation of temporal specification schemes and concludes with a study of the temporal aspects of several distributed multimedia systems.

### 4.1 Temporal Specification Schemes

Recent temporal specification schemes include Firefly [6], OCPN [22], etc. To compare these schemes we apply the proposed temporal reference framework as a guide. Each temporal specification scheme is compared in terms of structure of specification language (semantic

and syntactic domain) and equivalence of the models of time used (at the semantic level). The main differences between techniques are illustrated with figures.

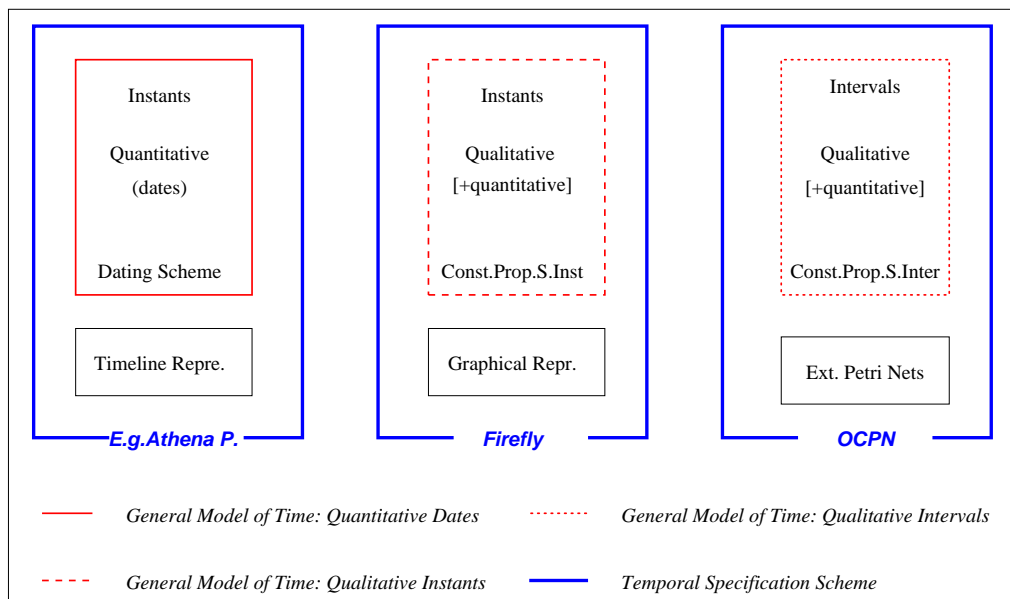


Figure 8: Comparison of Timeline, Firefly, and OCPN: Three Models of Time belonging to Different Classes of the General Classification

Fig. 8 shows a comparison of temporal specification schemes for (1) the timeline, (2) Firefly, and (3) the OCPN. As summarized in the figure, these three temporal specification schemes are distinct because of their different use of models of time. Therefore, they belong to different categories of our classification scheme. The timeline representation belongs to *quantitative dates*, Firefly belongs to *qualitative instants*, and OCPN belongs to *qualitative intervals*. As they use different models of time, they also use different syntax in their specification. Each is described in the following.

**The timeline approach** The timeline representation is the most basic method used for temporal specification. It consists of a dated timeline. A synchronization mechanism for this representation scheme can interpret the timeline and execute the appropriate actions at the indicated moments in time. Several synchronization systems use a timeline for the representation of the timing constraints (e.g., the Athena Muse Project [13] and Gibbs et al. [11]).

The model of time used for the timeline belongs to the general category of *quantitative dates*. The contextual quantitative information corresponds to the exact date (“at”) of the

basic time unit. The scheme can model homogeneous or heterogeneous temporal scenarios, but not ones with uncertainties.

**Firefly** Buchanan and Zellweger [6] propose a temporal specification scheme for the definition of general multimedia scenarios. They also develop algorithms for deriving the appropriate schedules for achieving the specified scenario.

The model of time used by Firefly belongs to the general category of *qualitative instants*. The contextual information corresponds to the basic binary temporal relationships between instants (i.e., only a single relationship between two instants). Associated quantitative information can also be used optionally. For modeling instants in the temporal scenario, no indeterminacy can be expressed in the temporal relations. For modeling intervals in the temporal scenario, indeterminacy in the temporal relationships can be expressed by a combination of the basic binary relationships between basic time units (instants). In any case, indeterminacy in the exact instant of occurrence can be described. A graphical representation (syntax) is used to capture the relations between the specified instants.

**OCPN** Little and Ghafoor [22] proposed the OCPN, a temporal specification scheme for the description of general multimedia scenarios. The authors also present algorithms for deriving a playout schedule from a specification of a temporal scenario [23, 10].

The model of time used for the OCPN belongs to the general category of *qualitative intervals*. The contextual information is both qualitative and quantitative. The temporal relationships considered are the thirteen basic binary temporal relationships. No indeterminacy can be expressed in the temporal scenario because the temporal relations are fixed (no indefinite temporal relations), and modified by quantitative information, which completely determines the scenario. The graphical representation (syntax) is based on an extended type of Petri net and retains most conventional Petri net semantics.

Fig. 9 shows a comparison of three related temporal specification schemes: Hoepner's path operators, the OCPN, and Wahl and Rothermel's operators [28]. In this comparison, the models of time of the temporal specification schemes belong to the same general category: *qualitative intervals*. The particular models of time used by each differ because of the contextual information that they consider. Furthermore, each technique uses a different representation. They are described in the following.



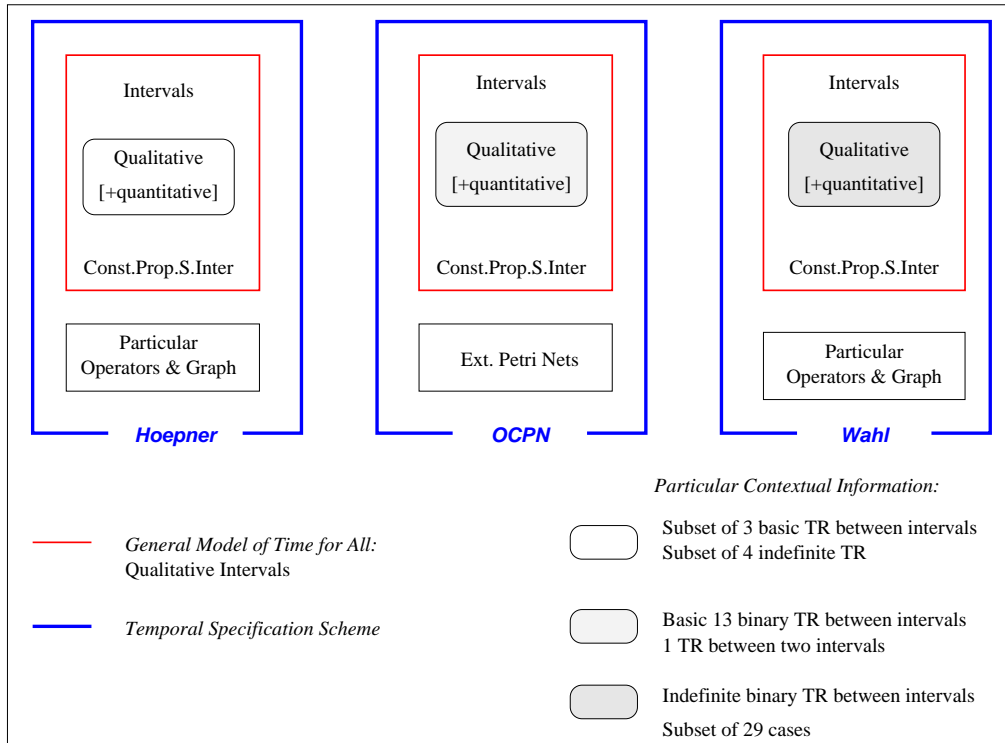


Figure 9: Comparison of Hoepner’s Path Operators, OCPN, and Wahl’s Specification: Same Class of General Models of Time, Different Contextual Information.

**Hoepner’s Path Operators** Hoepner [14, 15] defines a temporal specification scheme for the description of general multimedia scenarios. Her scheme consists of a set of path operators with an associated graphical representation. These operators are valid for any synchronization mechanism that can interpret them. One such mechanism is described in reference [15].

The model of time used belongs to the general category *qualitative intervals*. The specific contextual information for the qualitative information corresponds to a subset of the basic binary relationships (three) and to a subset of the indefinite temporal relationships (four) between intervals. In this temporal specification scheme, the syntax limits the extent of the possible temporal semantics, i.e., the path operators limit the number of binary temporal relationships that can be modeled. The expressive power is reduced to model only a subset of the possible temporal relations of the intervals in the temporal scenario. Indeterminacy in the temporal relationships (four indefinite relations) can be expressed.

**Wahl and Rothermel’s Temporal Specification Scheme** Wahl and Rothermel [28] have proposed a temporal specification scheme for multimedia scenarios. A common set of

operators describe the temporal relationships between intervals and the possible variations due to user interaction. A synchronization mechanism for realizing the specification has not been reported.

For this scheme, the model of time belongs to the general category of *qualitative intervals*. The contextual information is qualitative and can use quantitative information optionally. The temporal relationships considered are a set of 29 indefinite binary temporal relationships. Ten operators are used in conjunction with parameters to cover all possible cases and can be illustrated graphically. Indeterminacy of interval durations in the temporal scenario can be expressed in addition to indefinite temporal relationships between pairs of intervals.

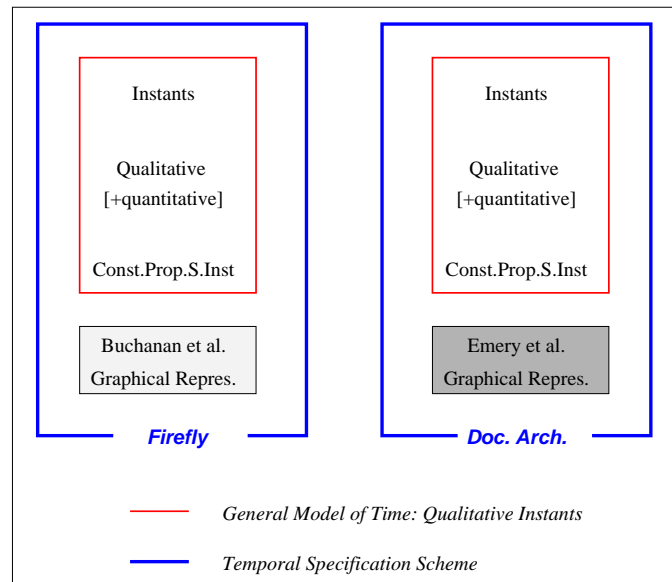


Figure 10: Comparison of Firefly and Emery’s Document Architecture: Same Model of Time, Different Graphical Representations.

Fig. 10 present a comparisons of temporal specification schemes with identical temporal semantics but different syntactic domains. In this case, the model of time used is *qualitative instants* with the three basic binary relationships between instants and quantitative information. Emery’s approach is explained as follows.

**Emery’s Temporal Specification Scheme** Emery [7] et al. propose a temporal specification scheme for the description of general multimedia scenarios. The temporal semantics are the same as Buchanan’s approach. (In this comparison we do not consider the concept of flexibility that is introduced by Buchanan; we are comparing this from the temporal specification perspective.) The graphical representation is different as are the synchronization

rendering mechanisms of the system. Buchanan’s approach has been developed in conjunction with a scheduling system, while Emery’s has been developed for a run-time system.

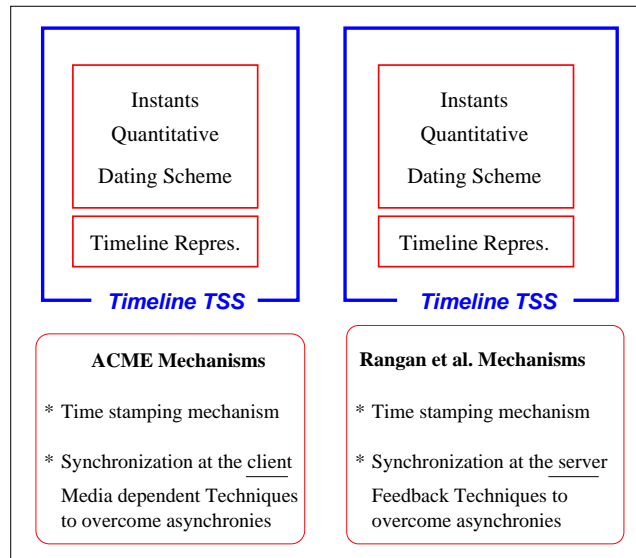


Figure 11: Comparison of the ACME System and the system of Rangan et al.: Same Temporal Specification Scheme, Different Mechanisms.

Fig. 11 illustrates two synchronization mechanisms that have been designed for the same temporal scenario: the synchronized delivery of audio and video. In both cases, the temporal specification is created in run-time with a time-stamping mechanism.

In this case, the re-synchronization mechanisms are different: the ACME [4] system tries to overcome asynchronies at the client side, while the system proposed by Rangan et al. [25] tries to solve the asynchronies at the server side. The different architectures of the systems also has an influence on the mechanisms. We illustrate these approaches to highlight that a temporal specification can be achieved by very different system architectures and mechanisms.

**Discussion** The aforementioned specification schemes can be examined for equivalence of temporal specification schemes. Fig. 12 shows a multimedia scenario and three possible temporal specification schemes for its representation. Using a natural language, the scenario can be described as a video of a person anchoring the news, a title of the broadcast company presented for the duration of the news broadcast, and two graphics that appear during the exposition. The Timeline, Firefly and OCPN schemes can be used for its specification.

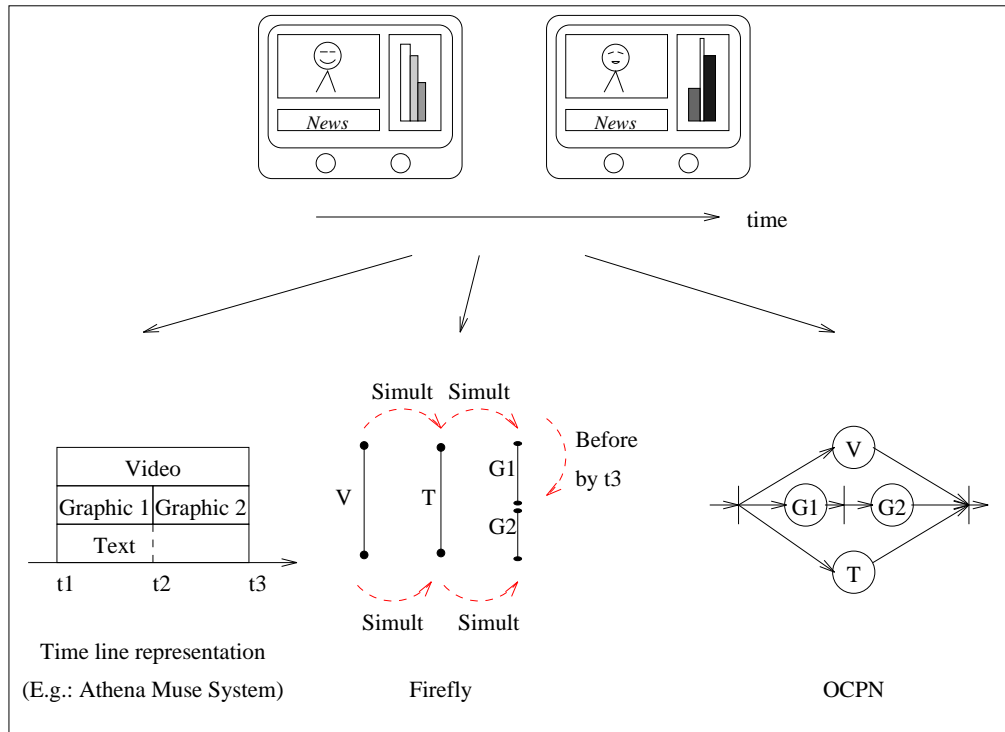


Figure 12: Several Temporal Specifications of a Particular Multimedia Scenario

The resultant temporal specifications are shown at the bottom of Fig. 12. The temporal specifications are different but equivalent for this particular multimedia scenario.

Few studies dealing with equivalent temporal specification schemes have been reported. Wahl and Rothermel [28] present a mathematical comparison of the equivalences among some temporal specification schemes: Timeline (e.g., [13]), OCPN [22], Firefly [6], MHEG [18], and Wahl and Rothermel’s operators [28]. The results of this study indicate that the listed approaches are equivalent for scenarios with no indeterminacy; however, only two of them (Firefly and Wahl and Rothermel’s technique) can be used for modeling indeterminate scenarios. Although this study effectively compares the equivalences at the semantic level, it does not examine translations among the schemes at the syntactic level.

In contrast, Erfle [8] studies possible translations among three temporal specification schemes. He proposes translations at the syntactic level, but does not use an adequate theoretical foundation for sufficient comparison at the semantic level. Compared schemes are OCPN [22], MODE [5], and Firefly [6], studied for the purpose of translation into HyTime [17]. Because HyTime uses a timeline model, the study results in a comparison of four tempo-

ral specification schemes. He demonstrates that HyTime is able to specify the same temporal constraints as the other three schemes, and details the equivalent syntactic translations for particular temporal scenarios. He admits that for user interaction (indeterminate temporal scenarios), there are some indeterminate semantics that cannot be translated directly to HyTime. A mathematical justification for these results (with the exception of MODE) can be found in reference [28].

Iino et al. [16] propose an object-oriented scheme for spatio-temporal synchronization of multimedia information. They also study the translation between the OCPN [22] and their proposed model. Their temporal specification scheme differs from the OCPN in two respects: (1) the contextual information associated with the OCPN consists of the basic binary temporal relationships, while the their model considers the  $n$ -ary basic temporal relationships [24]; and (2) the syntax of their model includes additional spatial and “processing” capabilities. We note that any scenario modeled with the OCPN can also be modeled with Iino’s scheme. The translation between the two temporal specification schemes is achievable with an increase in modeling power.

## 4.2 Distributed Multimedia Systems

It is particularly interesting to observe how the various system components follow different models of time in a complete distributed multimedia system. In this section we briefly show the utility of the temporal reference framework in the analysis of distributed multimedia systems. Again, our focus is on the identification of the different models of time used in each component.

Fig. 13 illustrates a system realization of the Firefly system (temporal specification plus scheduler). The figure shows the temporal specification scheme for a multimedia scenario that considers user interaction, i.e., indeterminacy in the final realization. With a language description it can be described as follows: “Object 2 is presented before object 3, and during the presentation of objects 2 and 3, object 1 is presented. Also, two related unpredictable events are also considered, event 4 triggers event 5 after 2 seconds, and they end together (the end is unpredictable).” For this system we see two models of time appearing: the Firefly temporal specification scheme, and the derived schedules. The former originates from the Firefly authoring tool in the specification of the multimedia scenario (*qualitative instants*), and the latter from the presentation playout schedules of the presentation system (*quantitative dates*). Because user interaction is considered, the derived presentation schedules are

provisional, and they are merged properly in real-time following the user's interaction.

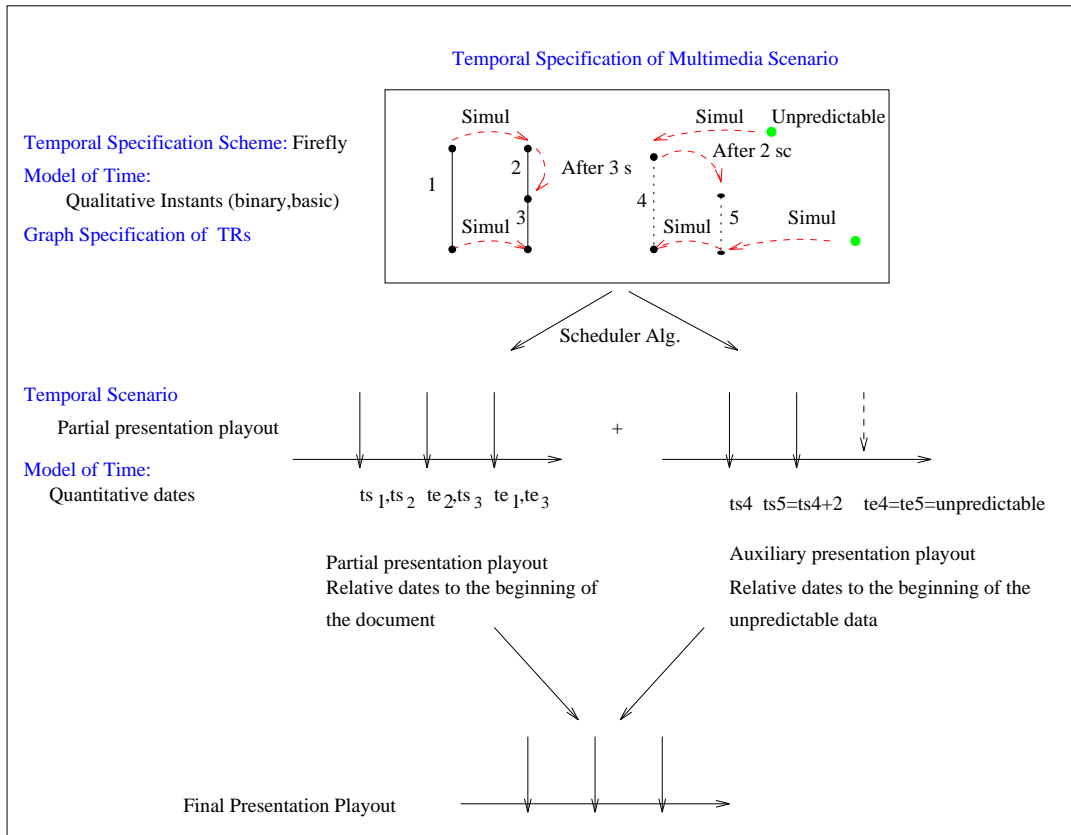


Figure 13: Firefly

Fig. 14 illustrates the operation of the distributed system proposed by Little and Ghafoor [23] for the same multimedia scenario of the previous example (unpredictable events are not considered here). The system uses a derivation of a playout schedule for presentation. As it assumes a distributed architecture, delay and jitter variations are anticipated during the presentation process. To overcome these problems, a secondary schedule is derived [10]. The system shown uses a scenario specification through the OCPN (*qualitative intervals*), and a playout specification (*quantitative dates*). The communication system also works according to the same model of time as the playout schedule. The quantitative information in the case of the playout schedule is different than the information used in the retrieval schedule (here, the dates are derived from the presentation playout schedule).

The identification of temporal structures in the overall multimedia system is very useful for comparison. Systems that are intended for different applications, and that appear to

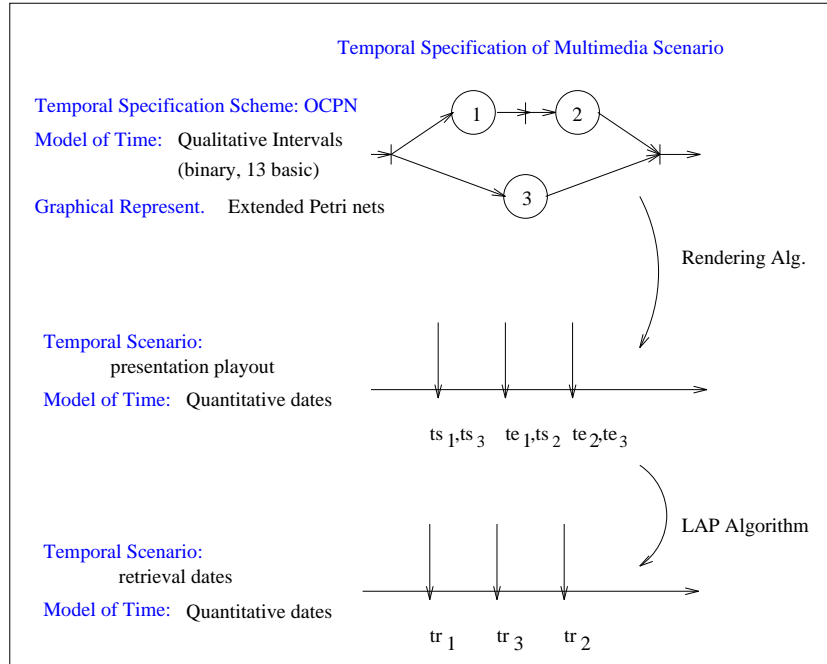


Figure 14: OCPN System

have little resemblance can have identical temporal structures. Fig. 15 compares temporal structures of two distributed multimedia systems: the OCPN system presented above, and a possible implementation following the temporal specification scheme proposed by Li Li et al. [20, 21]. In this case, the systems have the same temporal structure: (1) a temporal specification scheme for the multimedia scenario, on the left-hand side, OCPN, on the right-hand side, Temporal Graph Model (TGM); (2) the presentation system, both of them follow a presentation schedule that has been derived from the specification of the multimedia scenario (a *quantitative dates* model of time); (3) a communication system that follows a retrieval schedule that has been derived from the presentation schedule. Although the temporal specification schemes and the theory and mechanisms to derive the schedules and carry them out are different, the temporal structures are the same.

**Discussion** The identification of the temporal models used in each component of the system clarifies the way in which different approaches can be combined. For example, Amer et al. [3] have designed a communications protocol which takes advantage of partial orderings to improve transmission bandwidth utilization. One can envision the use of a specification

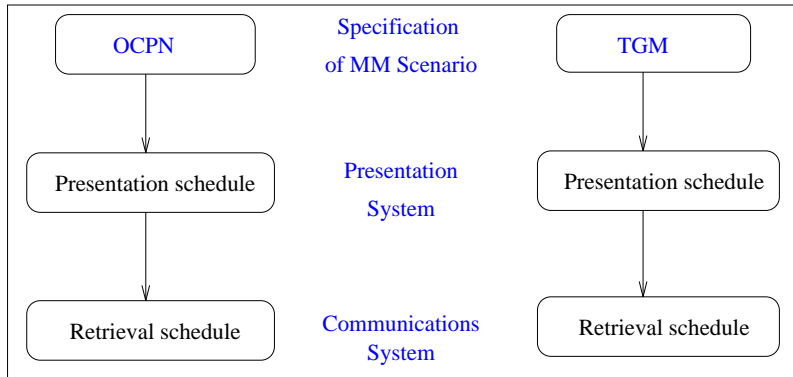


Figure 15: Comparison of Systems Derived from Li Li et al. and the OCPN

scheme such as the OCPN for the description of a multimedia scenario, the use of a presentation schedule for the playout mechanism, and the use of the aforementioned communications protocol for the transmission of the media units. To achieve this combination we need an algorithm that generates a partial ordering from the OCPN temporal specification.

Another protocol that has been proposed for the communication system is the Flow Synchronization Transport Protocol [9]. It has been designed for media units generated in real-time, and provides a constant virtual delay between sources and destinations. This protocol can also be used for the transmission of stored data by using a prespecified retrieval schedule (Fig. 16).

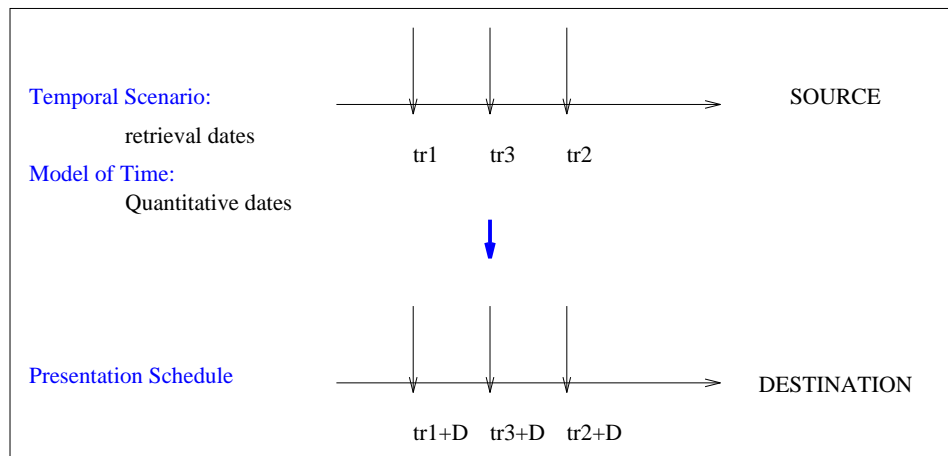


Figure 16: The Flow Synchronization Protocol



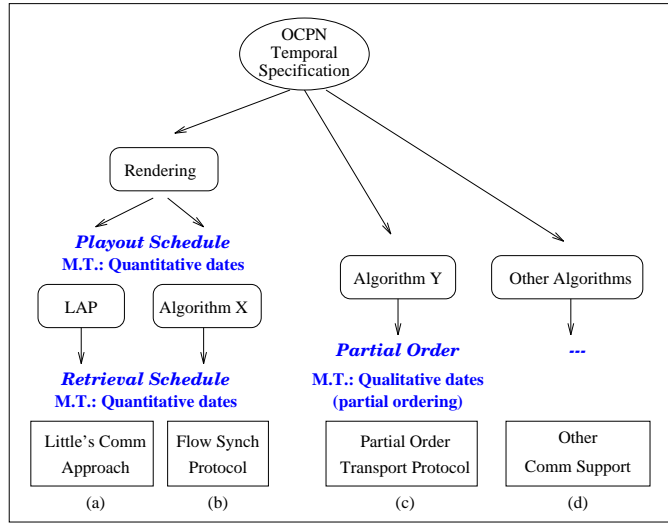


Figure 17: Using the OCPN with Different Communications Protocols

Fig. 17 illustrates several alternatives for the construction of distributed multimedia systems by combining a temporal specification scheme with different mechanisms for satisfying the communications component. The models of time for each part of the system are highlighted. In this figure, an OCPN specification is shown with a rendering algorithm that generates quantitative dates. This is linked to the LAP algorithm and might also be linked to Escobar’s flow synchronization protocol through some (undefined) algorithm  $X$ . The OCPN can also be applied to techniques such as the partial order protocol of Amer et al. through some (undefined) translation  $Y$  or to other communications approaches not considered here. In all cases the unifying basis is a model of time.

Fig. 18 shows another example of the combination of techniques to yield multimedia synchronization. Here we use the same communication system for different temporal specification schemes when uncertainties do not exist. From any of the temporal specification schemes we can derive a presentation schedule. With some (undefined) algorithm ( $X$ ) we can transform presentation schedules into retrieval schedules, and then use the flow transport protocol for communications.

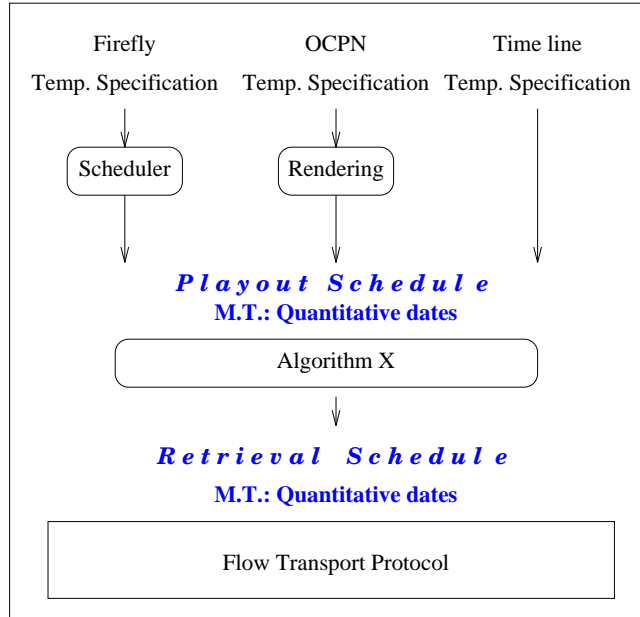


Figure 18: Using Flow the Synchronization Protocol for Several Temporal Specification Schemes

## 5 Conclusion

In this paper we have divided the general multimedia synchronization problem into two parts: one of modeling, representing, and specifying timing requirements of multimedia scenarios; and the other of achieving a temporal specification via synchronization methods. The former problem was the focus in our development of a temporal reference framework that can be used to evaluate and synthesize temporal specification schemes for the support of multimedia synchronization.

The temporal reference framework is based on existing temporal theory and modeling techniques and attempts to unify the terminology applied towards temporal specification for multimedia. The framework was applied to the comparison of existing approaches for multimedia synchronization to illustrate the differences of modeling power and to justify the development of the framework.

Our comparison of existing approaches indicates the utility of the framework. In addition, the analysis explains why there are so many synchronization frameworks, how a multimedia scenario can be represented with different temporal specification schemes, and why some specification schemes cannot model all scenarios.

## References

- [1] Allen, J.F. Maintaining Knowledge About Temporal Intervals. *Communications of the ACM*, 26(11):832–843, November 1983.
- [2] Allen, J.F. Time and Time Again: The Many Ways to Represent Time. *International Journal of Intelligent Systems*, 6:341–355, 1991.
- [3] Amer, P.D., Chassot, C., Connolly, T.J., Díaz, M., and Conrad, P. Partial Order Transport Service for Multimedia and Other Applications. *IEEE/ACM Trans. on Networking*, 2(5):440–456, October 1994.
- [4] Anderson, D. and Homsy, G. A Continuous Media I/O Server and Its Synchronization Mechanism. *Computer*, pages 51–57, October 1991.
- [5] Blakowski, G., Hübel, J., and Langrehr, U. Tools for Specifying and Executing Synchronized Multimedia. In R.G.Herrtwich, editor, *Proc. Network and Operating System Support for Digital Audio and Video*, pages 271–282. Springer-Verlag, 1991.
- [6] Buchanan, M.C. and Zellweger, P.T. Automatic Temporal Layout Mechanisms. In *Proc. ACM Multimedia 93*, pages 341–350, Anaheim, CA, August 1993.
- [7] Emery, J. and Karmouch, A. A Multimedia Document Architecture and Rendering Synchronization Scheme. In *Second International Conference Broadband Islands (Bridging the Services Gap)*, pages 59–69, June 1993.
- [8] Erfle, R. HyTime as the Multimedia Document Model of Choice. In *Proc. IEEE Intl. Conference on Multimedia Computing and Systems*, pages 445–454, Boston, MA, May 1994.
- [9] Escobar, J., Deutsch, D., and Partridge, C. Flow Synchronization Protocol. *IEEE/ACM Trans. on Networking*, 2(2):111–121, April 1994.
- [10] Gibbon, J. and Little, T.D.C. Real-Time Data Delivery for Multimedia Networks. In *Proc. of the 18th Annual Conference on Local Computer Networks*, pages 7–16, Minneapolis, MN, September 1993.
- [11] Gibbs, S. Composite Multimedia and Active Objects. In *Proc. Object-Oriented Programming Systems, Languages and Applications (OOPSLA'91)*, pages 97–112, 1991.
- [12] Hamblin, C.L. Instants and Intervals. In *Proc. of the 1st Conf. of the Intl. Society for the Study of Time*, New York 1972.

- [13] Hodges, M. and Sasnett, R. “Synchronization”. *Multimedia Computing: Case Studies from MIT Project Athena*, Addison-Wesley, 1993.
- [14] Hoepner, P. Presentation Scheduling of Multimedia Objects and its Impact on Network and Operating System Support. In R.G.Herrtwich, editor, *Proc. Network and Operating System Support for Digital Audio and Video*, pages 132–143. Springer-Verlag, 1991.
- [15] Hoepner, P. Synchronizing the Presentation of Multimedia Objects. *Computer Communications*, 15(9):557–564, November 1992.
- [16] Iino, M., Day, Y.F., and Ghafoor, A. An Object-Oriented Model for Spatio-Temporal Synchronization of Multimedia Information. In *Proc. IEEE Intl. Conference on Multimedia Computing and Systems*, pages 110–119, Boston, MA, May 1994.
- [17] ISO. Hypermedia/Timed Based Document Structuring Language (HyTime). *ISO/IEC IS 10744*, April, 1992.
- [18] ISO. Information Technology Coded Representation of Multimedia and Hypermedia Information Objects (MHEG). ISO/IEC JTC1/SC29/WG12. Standards document, February, 1993.
- [19] Ladkin, P. and Maddux, R. On Binary Constraint Networks. Technical Report Palo Alto, CA, Krestel Institute, 1988.
- [20] Li, L., Karmouch, A., and Georganas, N.D. Multimedia Teleorchestra with Independent Sources: Part 1 – Temporal Modeling of Collaborative Multimedia Scenarios. *ACM/Springer Multimedia Systems*, 1(4):143–153, 1994.
- [21] Li, L., Karmouch, A., and Georganas, N.D. Multimedia Teleorchestra with Independent Sources: Part 2 – Synchronization Algorithms. *ACM/Springer Multimedia Systems*, 1(4):154–165, 1994.
- [22] Little, T.D.C. and Ghafoor, A. Synchronization and Storage Models for Multimedia Objects. *IEEE Journal on Selected Areas in Communications*, 8(3):413–427, April 1990.
- [23] Little, T.D.C. and Ghafoor, A. Scheduling of Bandwidth-Constrained Multimedia Traffic. *Computer Communications*, 15(6):381–387, July/August 1992.
- [24] Little, T.D.C. and Ghafoor, A. Interval-Based Conceptual Models for Time-Dependent Multimedia Data. *IEEE Transactions on Knowledge and Data Engineering*, 5(4):551–563, August 1993.

- [25] Rangan, P.V, Vin, H.M., and Ramanathan, S. Designing an On-Demand Multimedia Service. *IEEE Communications Magazine*, pages 56–64, July 1992.
- [26] van Beek, P. and Cohen, R. Exact and Approximate Reasoning about Temporal Relations. *Computational Intelligence*, 6:132–144, 1990.
- [27] Vilain, M. and Kautz, H. Constraint Propagation Algorithms for Temporal Reasoning. In *Proc. AAAI-86, Artificial Intelligence*, pages 377–382, Philadelphia, August, 1986.
- [28] Wahl, T. and Rothermel, K. Representing Time in Multimedia Systems. In *Proc. IEEE Intl. Conference on Multimedia Computing and Systems*, pages 538–543, Boston, MA, May 1994.