Authoring Environment for Interactive Multimedia Documents

Muriel Jourdan, Nabil Layaïda and Loay Sabry-Ismail

OPERA project, INRIA Rhône-Alpes.
655 avenue de l'Europe, 38330 Montbonnot, France.
e-mail: {Muriel.Jourdan, Nabil.Layaida and Loay.Sabry}@inrialpes.fr
tel : +33 4 76 61 52 00, fax : +33 4 76 61 52 52
http://opera.inrialpes.fr/OPERA/

ABSTRACT

In this paper, we make an analysis of the main requirements needed for building an authoring environement for interactive multimedia documents. These requirements are of different nature as they are related to a considerable number of computer science fields. Then, we present our contribution to meet these requirements. The particularity of our project is to consider the problem of multimedia authoring as a whole and to develop a real running application to validate our theoretical results.

1 Introduction

The recent advances in multimedia systems, together with the advent of high speed networks, have paved the way to a new generation of applications. In particular, authoring environments have found in multimedia the means of increasing the richness of information contained in electronic documents. Multimedia documents compose in time and space different types of elements like video, audio, still-picture, text, synthesized image, etc. Interactive multimedia documents aim at transforming the reader from passive : (he cannot interact with the document presentation) to an active one. For example hyperlinks are used to navigate freely inside the same document and/or between different documents.

As compared to classical documents, multimedia documents are characterized by their inherent temporal dimension. Basic media objects, like video, have intrinsic durations and they can be temporally organized by the author. This temporal organization is called the temporal scenario of the document.

This temporal dimension raises numerous difficulties. For instance, the famous WYSIWYG paradigms cannot be applied. Edition and presentation phases are not the same, so we must distinguish between the **specification** phase (or editing phase) of the temporal scenario and the **presentation** phase.

The purpose of this paper is to present the different issues involved in the design of a multimedia authoring environment and how these issues are related to each other. Such a general presentation (like in.[7]) of this research area is necessary since it is still a new domain of interest (the first papers in the domain appeared at the beginning of the nineties) and we think that the problematic is still not clear. Defining what are the requirements of a multimedia authoring environment is difficult and still an open issue, since experiences in this area are limited. Let us point out that the first experiments for an electronic book, merging video, text and user interactions, took place in 1983 (D. Backer).

We consider that the four following questions must be taken into account when designing a multimedia document authoring environment:

- which kind of scenario can be expressed ?
- How easy is it for an author to create his document ?
- Is it possible to specify incorrect scenario (it is not possible to execute the specification) and in the positive case, is it possible to prevent these errors statically ?
- Can the presentation layer respect the exact meaning of the specification without being disturbed by the system load and by the distribution of media objects all over the network;

The paper is organized in two parts. In the first one, we are discussing about each question above-mentioned, outlining difficulties raised by each individual case. We are also interested in the interactions between these different issues of the authoring problematic. One of our claims is that these issues cannot be addressed in an independent way. This first part begins with the description of a working example. In the second part of the paper, we present our approach which is the only one at the present time to take the authoring problematic as a whole and does not make hypothesis in order to simplify the problem. Two important choices have been made : firstly, to provide the author with a declarative and hierarchic specification language of multimedia documents based on an extension of Allen's interval operators; secondly, to use an extension of temporal constraint networks to manage the temporal scenario of a document. Moreover, we define a general architecture of the Madeus application to demonstrate the validity of the two above-mentioned choices.

2 Example

We would like to build a multimedia document to present a family of four members. The idea is that each person can be presented by a sequence composed of a still picture of his when he was born, a short video about his life (its duration is about one minute) and a recent photo of him. Each of these three elements is associated with a textual part which gives a relevant commentary on it. We suppose that the reader of the document could interrupt the video when he wants in order to see the last still picture faster. The sequence composed of the three elements, the three associated textual parts and the button of interaction will be called S in the rest of the paper. In addition, each family member can add a personal audio message (its duration is about 1 minute and 30 sec.). This audio message must start with the sequence S and must end before it. The figure Fig. 1 presents a possible execution of the scenario of one member of the family. In this execution, the reader (by clicking on the button) interrupts the video 40 seconds after its beginning.

The first screen of the document is composed of a still picture of all the members of the family. Once the presentation of this person ends, the first screen reappears.

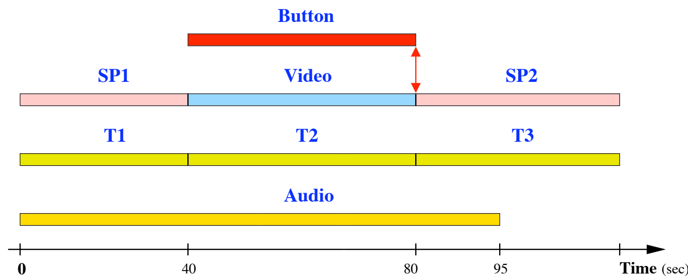


Figure 1: Possible execution for the presentation of one person

3 Multimedia authoring issues

3.1 A large set of possible temporal scenarios

3.1.1 Requirements

The relevant question here is : What kind of temporal scenario would an author like to write? Answering to this question is not trivial since the needs in this area are strongly related to authoring practical experiments which is not so widespread as authoring classical documents. We do not want to give an exhaustive list of expressivity requirements, but only to present the four points which seem necessary to build a non-trivial scenario and which raise numerous problems.

- A wide variety of basic objects: text, video, audio, still pictures, virtual animations, JAVA applets, ...
- Global and local interactions: The idea is to provide the reader with two kinds of interactions : global ones to interrupt all the current activities in the document in order to start other activities in another part of either the same document or another document. In our example, the first screen is composed of a set of global interactions. At the opposite, the button which interrupts the video in the sequence is a local interaction since it does not interrupt the audio message which is active at the same time of the video.
- Causal relations between events: Some behaviors of a scenario are expressed like causal relation between the temporal events (start or end of objects). In our example, this is the case with the button associated to the video : if the reader clicks on it the video is interrupted.
- Relative positioning of events (synchronization): they are used to express the equality or the precedence between temporal events. In our example, we used such synchronization many times: textual commentary related to components of the sequence must start (resp. end) when the component related to it starts (resp. ends). This is also the case with the audio message which must start at the same time as the sequence S and finish before it.

3.1.2 Difficulties

As far as the temporal expressivity is concerned, difficulties are mainly introduced by relative positioning of events, since it introduces the notion of temporal consistency of the specification.

It is important to understand well the distinction between causal relations between events and equality of events. Let's take the following example : A and B are respectively a video and an audio. Consider the two following specifications :

- (1) Start points of A and B as well as their end points are linked by an equality relation.
- (2) Start points of A and B are synchronized but the end of the shorter object causes the end of the longer one.

In the two cases, the two objects begin and end at the same instants, but in the first one the duration of the two objects must be the same (otherwise it is not possible to meet the specification) and in the second one whatever the values of durations are, the specification can be met. Thus, the strong difference between the two scenarios is that in the second case, one of the objects does not deliver all its contain to the reader of the document.

Some propositions in the domain of authoring of multimedia documents like CMIFed[3] controllable, only take into account causal relations avoiding the difficult task to check the temporal consistency of the documents. Others, like FireFly[2] or ISIS,[13] only consider relative positioning. These approaches take the problems of scenario specification as a set of linear inequalities to solve. Duration of objects are considered as : the system could choose a value between a minimal and a maximal bound. They statically compute a time-map which satisfies all the constraint. The two major drawbacks are that:

- they cannot manage local user interactions but only global ones. A global interaction is a relation between two different sets of inequalities which represent different non connected sub-parts of the document. The activation of a global interaction in their framework only implies to follow another statically pre-build time-map. On the contrary, local interactions interfere directly in one set of inequalities. Statically computing all the solution is not further possible.
- they are obliged to suppose that objects like video, audio, JAVA applets have a statically known duration. Indeed, objects duration are statically computed for each objects before the presentation phase. In order to satisfy these durations at the presentation time they interrupt the object if it is too long and wait in the opposite case. From our point of view this is not the good way to manage such objects.

In fact, we are convinced that the real difficulties of the multimedia documents authoring problematic are :

- to merge the two types of specification : causal and relative positioning. This is what we have done in our example : the sequence S contains a causality relation since a user-interaction could interrupt the video, and S must start at the same time as an audio message and must end after it. Causal relations introduce the distinction for each objects of two kinds of ending : the normal one and the interruption. This distinction must be taken into account while considering the semantics of relative positioning operators : which kind of objects ending do they consider ?
- to consider that some objects have a statically unpredictable duration. They introduce the distinction between two kinds of objects : the **controllable** ones : the system can

choose their values between a range of possible ones and the **uncontrollable** one : they take a value at the presentation time. In some cases, the uncertainty on this value could be restricted as far as quality requirements are considered. For instance, a video duration can have a range of possible values depending on its acceptable playing rates. We know the lower and upper bound values, but we don't know exactly which value it will take during the presentation.

These two points are related due to local interactions since they are a typical example of objects with statically unpredictable duration and they are always used with causal operators.

3.2 Static Analysis of the document specifications

When an author specifies a multimedia document, the document must be checked before passing to the presentation phase, in order to be sure that it is error free. We only consider here errors due to the temporal dimension of the multimedia document. It is obvious that the presence and the nature of temporal errors are strongly related to (1) the way in which the author specifies his document and (2) the expressiveness power of the specification language itself. For instance, with some approach (like SRT[8]) it is possible to express temporal deadlock : an object A is waiting for the object B to end and vice versa.

Approaches which are based only on causal operators (like CMIFed[3]) have not any problems of temporal consistency. At the opposite, relative positioning of events is a great root of errors. The trivial case is to specify that two objects have the same duration, although their interval of possible durations have an empty intersection. Another example, which is not so trivial, is to specify that the end of a video is synchronized with the end of an audio. If you consider that video and audio are uncontrollable objects, their durations can not be statically controlled and are only known when they end, it is not possible to guarantee that this specification is respected. In all these case you have to warn the author of the existence and place of the inconsistency.

It is not too difficult to express the problem of consistency if you only consider relative positioning of events and controllable objects. In this case, it is a classical problem of constraint satisfaction : a specification is consistent if and only if there exist a value for each start and end points which satisfy all the constraints. Difficulties arise when you consider uncontrollable objects. The notion of consistency becomes more complex[10] since you have to take into account all the possible durations for the uncontrollable object and the ordering between events.

Let's take two simple specifications (the video A is an uncontrollable object):

- (1) Present a still picture after playing a video A. During the presentation of this sequence, a text must move from a point X of the screen to a point Y.
- (2) The same as (1) but reverse the order of the still picture and the video.

Intuitively, the specification (2) is temporally consistent whereas the (1) is not if you consider that the duration of the video is only known when it ends. In the two examples, you could choose either the duration of the still picture and the speed of the text motion. The problem is to know if there exists some values for this two variables such that whatever the video duration is the end of the sequence is synchronized with the time point where the text reaches the Y point. It is true with the specification (2) since you could always adapt the duration of the still picture when knowing the real duration of the video. As a consequence, the duration of the sequence composed by the video and the still-picture could be statically known and you can choose at the beginning of the sequence the right speed for the motion text. This is not the case in the specification (1).

The same situation arises in our working example. In order to be sure that the sequence S will finish after the audio message, the duration of the second still picture of the sequence depends on the real duration of the video, which could be interrupt by a user-interaction.

The problem of uncontrollable variables in a constraint satisfaction problem is studied in another application domain: planning in robotics by[15]

3.3 User-friendly interface

We totally agree with the following citation : "Easy-to-use tools for creating, manipulating, and presenting multimedia content, will be an important factor in the creation of multimedia materials by nonspecialists. The most powerful contemporary authoring tools use a programming paradigm to allow complex interactions and synchronization relationships to be defined. This approach will be supplanted by paradigms that are more visual in nature and which allow the author to focus on the conceptual level of the presentation".[7] Today, the most used commercial authoring tool is Director[5] which used a proprietary language of script.

Here the relevant question is how long does it take an author to design or to modify his scenario ? Modification facility is an important functionality to take into account when designing an authoring environment, since building an interactive multimedia document is a "specify, test and modify" process . We identify three important aspects :

- Spatial and temporal positioning of document components must be expressed in a relative way, thus the system must support temporal and spatial formatting functionalities. In other words, it should be the system's task (and not the author's) to fix the exact duration and timing (resp. the exact coordinates) of each object. This is the only way to provide the user with easy to modify specifications. Obviously, the difficulty lies in the temporal dimension. The main difference from the problem of spatial formatting is the uncertainty of the values of the durations of some objects, local user-interactions and causal operators. The quality of the diagnosis in case of inconsistency is another point to take into account.

- Edition and presentation phases must be really integrated. It must be easy and fast to toggle between the two phases, so that the author can interact during the presentation phase to modify his document. This is the only way to be as close as possible to the WYSIWYG paradigm successfully used for static documents, but unusable for multimedia documents due to the time dimension.

- Hierarchic decomposition of the document must be supported. This is necessary to be able to manage large documents. However, this hierarchy must not be too constraining for the author. Some approaches (SRT, CMIFed) are based on the idea that the document is a tree structure with a temporal operator associated to each node. This solution has two major drawbacks : some temporal compositions can not be specified and the modification of the scenario is not easy.

3.4 An efficient presentation layer

We can identify several objectives for the presentation layer[11]: (1) providing the reader with an accurate rendition of what the author has specified, (2) supporting Temporal Access Control operations like: stop, resume, fast-forward, fast-rewind, as well as hyperlinks to allow the user to manipulate the multimedia document's presentation, (3) and supporting the distribution of media objects across the network. This support for the distribution of media objects is becoming more and more necessary due to the increase in the number of the widely spread servers (WWW servers and others), holding media objects all over the network.

Difficulties arise due to the fact that multimedia applications are, in fact, soft real time systems and CPU time consuming, so the applications have to fight against limitations of resources (such as CPU availability, disks bandwidth, networks bandwidth, network accessing delays, availability of media channels, data buffers sizes, etc.). For example, media objects like video and still pictures need a considerable CPU processing time for image decompression, while media objects like audio and video, are really sensitive to the system load as they have deadline times that *must* be respected. So, Quality of Service (QoS) issues must be taken into consideration to guarantee, at least, minimum requirements of performance. Examples of situations that must be avoided are:

- The author specifies that two videos play in sequence. During the presentation, the first finishes but a considerable delay between the two videos is encountered (due to the time needed by decompression or the network accessing delay), in a way that it affects the quality of presentation.
- The author specifies that two videos play in parallel together with an audio and the CPU is overloaded by the decompression of the videos in a way that the quality of the heard audio is deteriorating.

The resource limitations, previously pointed out, are unpredictable in nature. So, the task of the presentation system can be simplified, if delay values (of network and disk accesses as well as the buffer's queue waiting time) and CPU availability can be approximately anticipated.

Thus, it is important to have an efficient presentation layer that can make a tradeoff between the quality of presentation and the availability of resources, in such a way that it gives the reader an acceptable rendering of the multimedia document.

3.5 Synthesis

In order to synthesize this first part of the paper, our point of view is that the five following points are the central issues of the authoring multimedia documents problematic :

- Relative positioning of objects are necessary due to the user-interface requirements but this kind of operators are the principle reason of temporal inconsistencies and require the presence of a temporal formatter.
- Causality operators are necessary to provide the reader with local interactions. They introduce for each object, the distinction between two kinds of ending: normal and interrupted. This distinction must be taken into account while considering the semantics of relative positioning operators.

- Some objects have no statically controllable durations and we have to preserve their meaning (not to cut them). They make the notion of temporal inconsistencies more complex.
- Hierarchic decomposition is necessary to represent the logical structure but must not be used to represent the temporal organization of the document.
- The presentation phase of a document is not so simple as it seems to be at a first glance : it must be really integrated with the editing phase and it requires a technically high system of management.

Nowadays, no environment proposes an appropriate global answer for the authoring interactive multimedia documents problematic. Some of them (Firefly,[2] ISIS,[13] Madeus) take the difficult problem of temporal formating as the central issue, while others (CMIFed,[3] HTPSN based environment[17]) only consider causal operators to specify the temporal scenario of the document.

Madeus is the only system who study how mixing both temporal formating and causal operators and temporal formating with uncontrollable durations. We have not yet solved all the problems raised by these two points but some promising directions are delineated. Moreover, we have a sound basis application architecture that provides us with a fast way to validate our ideas.

4 Madeus approach

Two important choices have been made in Madeus : firstly, to provide the author with a declarative and hierarchic specification language of multimedia documents based on an extension of Allen's interval operators; secondly, to use an extension of temporal constraint networks to manage the temporal scenario of a document during its static analysis, and also to support the presentation phase. We begin to present the specification language that we proposed, and then the internal structure that we used to manage it both at the edition end presentation phases.

4.1 A declarative and hierarchic language of multimedia documents specification

4.1.1 Basic objects

Madeus basic objects are video, audio, still picture, text, temporal delay and button. Buttons are used in the scenario to integrate local user-interactions. All objects are abstracted as intervals. The interval associated to a button starts at the instant when the button can be activated and ends when it is really activated. Intervals are typed as controllable (text, still picture, temporal delay) or uncontrollable (video, audio, button). A minimal and maximal duration is associated to each duration. Maximal duration could be infinite, this is the case of a button. In case of a controllable interval, the temporal formater could choose any value within these limits. In case of uncontrollable interval, the temporal formater is informed that the real duration of the object will be within these limits.

4.1.2 Allen's interval algebra and our extensions

Allen's interval algebra[1] defines seven ways (without considering the symmetric relations) in which two intervals can be related in the time dimension. Briefly, the idea is that an interval can be strictly before (resp. after) another one, they can start (resp finish) at the same time, and so on. One of the most interesting operators is the one which states that two intervals are equal : they start and finish at the same time.

Allen's algebra expresses relative positioning of intervals, we extend it with causal operators to express causal relations between ending events:

- A par_min B: the two intervals start at the same time and the shorter causes the termination of the other one ;
- A par_master(A) B : the two intervals start at the same time and the end of A (which is the master) causes the termination of B, if and only if B is still playing.

We have, also, extended the algebra with the possibility to quantify free delays between intervals like in the following example:`I1 before(30,40) I2`. This means that I1 is before I2 and the delay between them is greater than 30 units of time and lower than 40.

4.1.3 Hierarchic decomposition

A document is composed of a set of objects (composed or basic ones) which are related in time. A hierarchic decomposition gives the logical structure of the document, where the siblings at each hierarchical level can be temporally related. Figure Fig. 2 gives some ideas about the hierarchic decomposition and the symbolic representation of the working example. The principal document has a root object called Menu which has four children (Option1, ..., Option4) of type button. These children are temporally related by the par_min operator, and each one of them is associated with a hyperlink to another corresponding document.

It is important to see that the hierarchic decomposition only fixes boundaries for the temporal relations.

Hyperlinks are managed like in HTML by associating a special attribute to an object. This attribute defines a component of the document to reach if the user clicks on the object.

4.1.4 Semantic and consistency issues

The semantics of this new language of specification which merges both causal and relative operators, and controllable and uncontrollable objects is not easy to define. First, we have to know what is the relevant behavior of some specification example. Second, we have to formalize the semantics in order to fix the execution that can be associated with a given specification. Finally, we have to decide whether a specification is temporally consistent. Let us take an example to illustrate the first point. Consider the following specification :
A meets B ; A par_min I
what is its meaning :

- Always inconsistent due to the fact that A can be interrupted ?

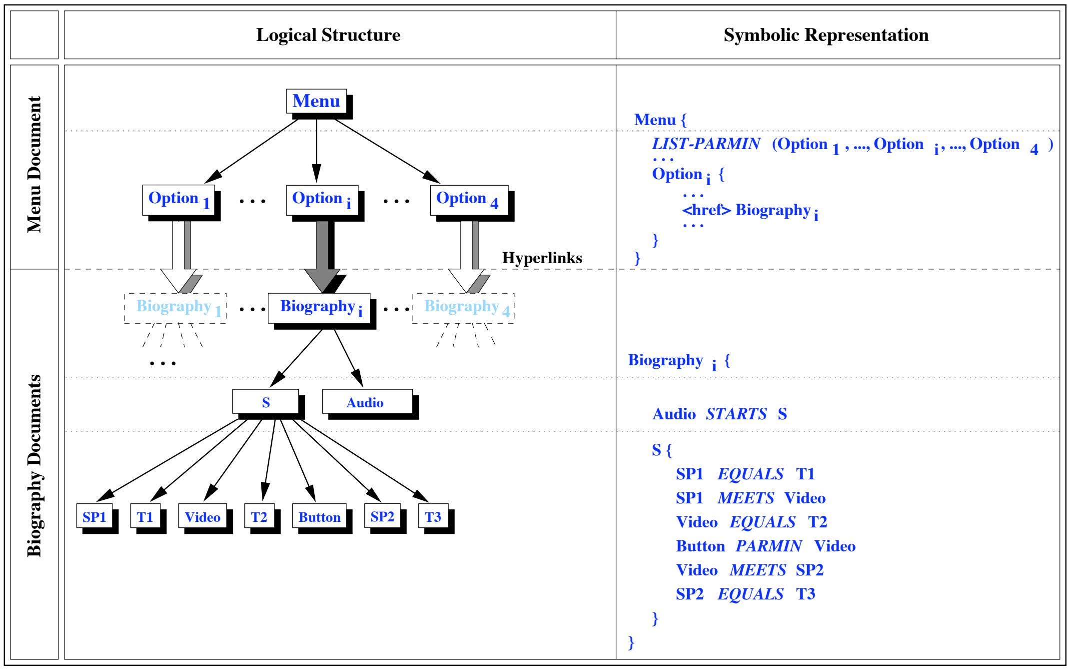


Figure 2: Symbolic and hierarchic specification of the working example

- Consistent if and only if A is shorter than I ?
- Always consistent: If A is not interrupted, B will start when A ends, otherwise B will not start.
- Always consistent: B starts when A either normally ends or is interrupted.

It seems to us that the last case is the most relevant but we have to confront this choice to the practice.

We are currently working on the formalization of such a semantics but it is not a simple task. The definition of a consistent specification is not easy to achieve too. The intuitive idea is clear: a scenario is consistent if and only if it is always possible to adapt the durations of some controllable objects to compensate the durations of uncontrollable ones. However, in order to formalize this idea we have to mix universally and existentially quantifiers and to build a partial order on variables from the specification.

4.2 An extension of constraint networks to manage temporal scenarios

In order to manage the temporal specification of the document, our approach consists in incrementally (temporal relations are analyzed one by one) translating the specification into a hierarchy of extended temporal constraints networks. Each time a relation is analyzed, the consistency of the scenario is checked. We first present the constraint temporal networks formalism as defined in[4] and then present the extended form we used. Moreover we give in the second part, ideas about how we plan to check the correctness of a scenario.

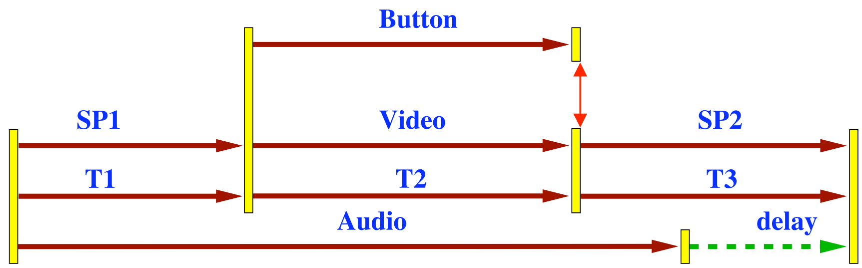


Figure 3: An example of Madeus constraint network.

4.2.1 Temporal constraint networks formalism

The formalism of temporal constraint networks is used for different kinds of applications (such as planning and scheduling of robotics), which has been defined to capture time information. A temporal constraint network is a Directed Acyclic Graph (DAG) where nodes represent time points and an edge (i, [min,max], j) represents a temporal interval with a duration range from min to max between the two time points i and j. Different kinds of algorithms exist to answer questions like : does a solution exist ? Can a particular set of values be considered as a solution for the problem ?

In our context of multimedia document a temporal constraint network describes the temporal ordering of media objects by capturing an information about their logical and quantitative temporal dependencies. The range of possible durations of a dynamic object is modeled by an edge labeled by its appropriate interval. Free or quantified delays are used to measure the temporal metric distance between objects or to enforce precedence relationships between them.

Before we go on, let us introduce the following notion used in the sequel : A temporal chain [i,j, ...s] is a sequence of contiguous edges where each edge s only related to one successor and one predecessor.

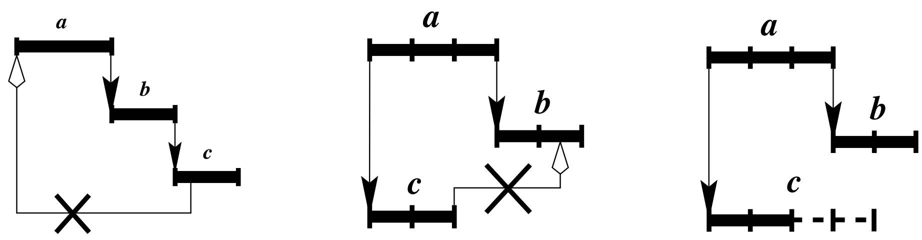
4.2.2 Madeus temporal constraint networks

In MADEUS, the previous definition of a constraint network has been extended to capture two types of information : the controllability of duration and the causality relations between events.

We need to distinguish between controllable objects whose duration values can be chosen by the system from a set of possible values, and uncontrollable objects which take random duration values at presentation time. We add an additional labelling to intervals which indicates whether the temporal distance is controllable or not (respectively $[\text{min,max}]^c$ and $[\text{min,max}]^u$).

In order to express causality in temporal scenarios, we add causal relations between temporal instants, i.e. nodes of the network.

For instance, figure Fig. 3 shown the extended networks associated to the presentation of one person in our working example. Nodes of the constraint network appear in yellow. The causal operator par_min is represented by a two ways arrow. A non null delay is introduced since the audio must finish before the sequence S.



(a) Qualitative *(b) Quantitative* *(c) Indeterministic*

Figure 4: Three kinds of inconsistency

4.2.3 Static Analysis

Given one temporal network of this form, the relevant question is to check the existence of at least one solution for the presentation of the specification. Three kinds of inconsistencies can occur in the specification of a scenario (see figure Fig. 1). For example, given three objects a, b and c, specifying that a meets b and b meets c makes the specification c overlaps a inconsistent. This is called a qualitative inconsistency because it results from the nature of the relations independently of the durations of the objects they link. However, qualitative consistent specifications may be inconsistent regarding the explicit durations of the related objects. For example specifying that a meets b and a starts c makes the qualitatively consistent specification c overlaps b inconsistent, if duration (a)=3 seconds and duration (c)=2 seconds. Finally, in the example we show another type of inconsistencies due to the presence of indeterministic intervals. Given three objects a, b and c where duration (a)=3 seconds, duration (b)=2 seconds and c is an indeterministic object whose duration is bounded in the interval [2, 4], specifying that a meets b, a starts c and c overlaps b leads to inconsistency since the end of c may occur at any time, therefore this specification may possibly fail.

Only the first one could be manage with classical algorithms of the domain, for the others we have to develop new one

- **qualitative:** a topological sort of the constraint network is maintained in order to avoid this kind of inconsistency by detection of cycles.
- **quantitative**: by using constraints propagation techniques,[16] each time the network is modified, we check that all chains parallel to each other still have non-empty intersection of their allowable duration ranges. Algorithms used for the constraint propagation are polynomial ones. Moreover, they support incremental modification as far as a new relation is inserted or a the duration of an object is changed. As a consequence, this kind of modification are quickly managed by the system. The problem is more difficult (this is well known in the area of constraint propagation techniques) as far as a relation (or an object) is suppressed.
- **indeterministic**: the idea here is to perform worst case analysis to obtain dates of the form (inf, sup) on each node of the networks. Then, we have to take into account the fact that some duration are controllable or not to decide the correctness of the specification. The idea here is to see if it is possible to recover dynamically, at the presentation stage, the indeterministic behavior of uncontrollable intervals by using the flexibility of some controllable objects which are played in the future.[10] We are first looking for a solution

based on recovery in the same chain of the network but plan to do better in exploiting controllable intervals in concurrent chains.

4.2.4 Presentation phase

Once the temporal validity of the scenario is checked, the difficulty is to compute the solution dynamically at the presentation stage. It cannot be done statically due to the uncontrollable durations. Starting from this time, the temporal constraint network is managed by the presentation layer like an execution support. The advantage of this solution (having the same structure as editing and presentation phase support is obvious) is the rapidity by which we can toggle between these two phases.

The presentation progresses as follows: Firing a node consists in starting all the objects (text, video, delay, audio, ...) associated with its outgoing edges. A node can be fired if and only if all the objects associated with its incoming edges have finished. Causal relations on each node are used to propagate end information. This internal structure which is the support of the presentation phase is rich enough to anticipate the executional state of the whole document at different future instants in the presentation life time. This ability of anticipation will allow us to manage better the media objects distribution on the World Wide Web.

5 Conclusion

Our approach is based on the following remarks :

- Using Allen's algebra to specify a scenario is a good idea since it is rather intuitive, easy to modify and efficient algorithms to detect inconsistencies exist.
- Unfortunately, Allen's algebra is not a complete answer to the problem of scenario authoring of multimedia documents: what about causal operators, uncontrollable objects presentation system ?

We are working on integrating these lacking features to the original framework of Allen. It raises non trivial theoretical problems. We are, also, working with great care in the development of a real running application.[6] The interface of our application is shown figure. Currently, the editing phase is mainly based on an integrated textual editor, although it is possible to modify the temporal scenario by selecting objects during the presentation phase and then setting the temporal relation holding between them with a palette of operators. Having an interactive authoring graphical user interface, is an important goal for us.

Several works are currently progressing on: spatial formatting and visualization of the temporal scenario.

Concerning the spatial formatting, we would like to introduce functionalities such as : alignments, centering, etc. As the case in the temporal dimension, the author is allowed to put spatial relations between objects of the same hierarchical level. Moreover, we do not want to use the restricting group/ungroup method as in most graphical interfaces. We plan to manage these spatial constraints by incremental dataflow solvers.[12] This is also for us an opportunity to clearly understand two different techniques to solve constraints: constraint propagation solvers and incremental dataflow solvers.

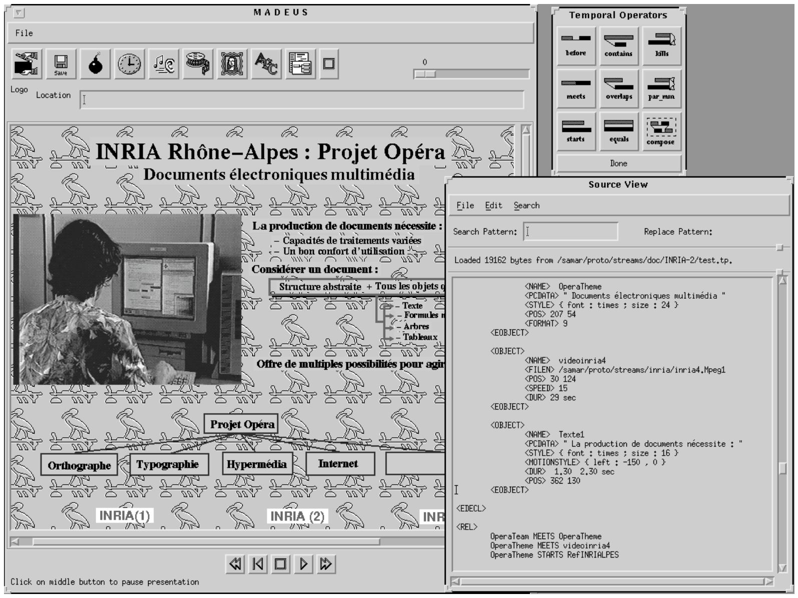


Figure 5: Interface of the Madeus prototype

The aim of this paper is twofold: first, we would like to convince the multimedia community of the interest and the diversity of the multimedia authoring documents issue. This is a big research topic that is strongly interrelated with other research areas (planning in robotics, formal verification, Human-Computer Interface, etc.). Currently, only a few research projects choose to work in this direction (ISIS,[13] Firefly,[2] CMifed,[3] environment based on a temporal extension of petri-nets[17]). The second aim of this paper is to show how the Opera project, which is interested in editing electronic document environments (Web), proposes with the Madeus prototype a global approach to the problem. We have presented and argued our three key contributions : using a declarative and hierarchic specification language of multimedia documents based on an extension of Allen's interval operators and managing the temporal part by an extended form of temporal constraint networks.

It is worth noting that the Madeus approach to represent the temporal dimension of multimedia documents, has interested the members of WWW Consortium (W3C) to extend the current HTML format in order to represent interactive multimedia documents on the Web.[9]

6 REFERENCES

[1] ALLEN J.F., "Maintaining Knowledge about Temporal Intervals", *Communications of the ACM,* vol. Vol. 26, num. No. 11, pp. 832-843, november 1983.

[2] BUCHANAN M.C., ZELLWEGER P.T., "Automatic Temporal Layout Mechanisms", *Proceedings of the First ACM International Conference on Multimedia,* pp. 341-350, Anaheim, California, august 1993.

[3] BULTERMAN D.C.A., VAN ROSSUM G., VAN LIERE R., "A Structure for Transportable, Dynamic Multimedia Documents", *USENIX Conference,* pp. 137-1559, Nashville, TN, june 1991.

[4] DECHTER R., MEIRI I., PEARL J., "Temporal Constraint networks", *Artificial Intelligence,* vol. 49, , pp. 61-95, 1991.

[5] DIRECTOR 4.0, "User's Guide", Macromedia.

[6] JOURDAN M., LAYAÏDA N., SABRY-ISMAIL L., "Time Representation and Management in Madeus : an authoring environment for Multimedia documents ", *in Proceedings of Multimedia Computing and Networking 1997 ,* SPIE 3020, San-Jose, February 1997.

[7] KOEGEL BUFORD J. F. , *Multimedia Systems,* vol. ACM Press, Addison Wesley, , 1994.

[8] KIM W., KENCHAMMANA-HOSEKOTE D. et SRIVASTA J. ., "Synchronization Relation Tree : A model for Temporal Synchronization in Multimedia Presentation", *Technical Report TR92-42, Dept. of Computer Science, Univ. of Minnesota, 1992 .*

[9] LAYAÏDA N, "Issues on Temporal Representation of Multimedia Documents", *International Workshop on Real-Time Multimedia and the World-Wide-Web (RTMW'96),* Sophia-Antipolis, October 1996.

[10] LAYAÏDA N., SABRY-ISMAIL L., "Maintaining Temporal Consistency of Multimedia Documents Using Constraint Networks", *Multimedia Computing and Networking 1996, M. Freeman, P. Jardetzky, H. M. Vin, ed., pp. 124-135, SPIE 2667,* San-José, USA, february 1996.

[11] LIAO W. and LI V., "Synchronization of distributed multimedia systems with user interactions", *International Conference on Multimedia Modeling ,* Toulouse 1996.

[12] SANNELLA M., MALONEY J., FREEMAN-BENSON B. and BORNING A., "Multi-way versus one way constraints in user interfaces: Experiences with the Deltablue algorithm", *Softwares Practices and Experiences,* vol. 23, num. 5, 1993.

[13] SONG J., DOGANATA Y, KIM M. and TANTAWI A., "Modeling Timed User-Interactions in Multimedia Documents", *Proceedings of the IEEE International Conference on Multimedia Computing and Systems,* 1996.

[14] VAN ROSSUM G., JANSEN J. , MULLENDER K. and BULTERMAN D., "CMIFed : a presentation Environment for Portable Hypermedia Documents", *Proceedings of the ACM Multimedia Conference,* California (USA) 1993.

[15] VIDAL T and H.FARGIER, "Contingent duration in temporal CSP : from consistency to controllabilities", *4th Int. Workshop on Temporal Representation and Reasoning (TIME97),* Daytona Beach, Florida, USA, May 10-11, 97.

[16] VILAIN M., KAUTZ H. A., "Constraint Propagation Algorithms for Temporal Reasoning", *Proceedings of AAAI,* pp. 377-382, Philadelphia, august 1986.

[17] WILLRICH R., de SAQUI-SANNES P., SENAC P. and DIAZ M., "Hypermedia Document Design Using the HTSPN Model", *International Conference on Multimedia Modeling,* Toulouse, 1996.