Synchronization in Multimedia Documents*

Peter King¹, Helen Cameron¹, Howard Bowman², and Simon Thompson²

¹ Department of Computer Science, University of Manitoba, Winnipeg, Manitoba, R3T 2N2, Canada
² Computing Laboratory, University of Kent at Canterbury, Canterbury, Kent, CT2 7NF, United Kingdom

Abstract. This paper presents a taxonomy of possible synchronization relationships between pairs of items in multimedia documents. Several existing approaches to the synchronization of entire items are reviewed. We then discuss classes of synchronization based upon dynamic events or conditions occurring within media items and their internal structure. We present a taxonomy of seventy-two possible such relations, which are illustrated by numerous examples and which are formalized in the authors' temporal logic notation, Mexitl. The ideas are then applied to provide a description of the lip-synchronization problem.

1 Introduction

Multimedia documents, their description and means for their authoring, are the subject of a considerable volume of current research and development work. The term "multimedia" refers to a document containing continuous or time dependent components, which are termed media items; see Erfle [Erf93]. A major part of the task facing the author of such a document is the specification of the temporal relationships among the media items in such a document. The present paper discusses our approach to this question.

In this paper we are primarily interested in temporal constraints between pairs of media items. Many authors have made use of the well-known taxonomy described by Allen [All83], who presents a complete set of thirteen such binary relationships or constraints between media items, as a starting point for defining a set of temporal relations. Shih et al [SHT96] have pointed out that the Allen relations are not sufficient from the viewpoint of authoring since they do not contain any precise timing information, which is typically a basis for item synchronization in multimedia documents. To this drawback, we add that the Allen relations treat each of the two media items as continuous and indivisible. There is no (direct) provision for synchronizing one object upon events or conditions occurring dynamically in the second object. The main objective of this paper is to address this latter point.

* The authors acknowledge the assistance of the British Council in providing travel grants, and the Natural Sciences and Engineering Council of Canada for research grants.
We will present what we consider to be a complete set of binary media item synchronization relationships. We will do this by presenting a set of relationships as the Cartesian product of various classes of primitive synchronization conditions which may occur in the two items which are to be related.

Each such synchronization condition will be discussed from three viewpoints. First, we will present an informal description of the synchronization relation. Second, we will present examples from hypothetical multimedia documents, illustrating the synchronization relation. Third, we will give a formal description of the synchronization relation. While most of our discussion will be limited to synchronization relations between a pair of media items, the ideas certainly extend to several items.

The formal descriptions of the synchronization relations will be given in the Mexitl notation. Mexitl [BCKT97a,BCKT97b] is a formal notation developed by the authors for use in specifying multimedia document. Mexitl is based on an interval temporal logic. It is a central component of the long-term goals of the authors: to make use of formal methods in the development of an authoring tool, which would thereby provide a means to address issues such as consistency verification, modeling, prototyping, and specification refinement. Although Mexitl itself is not the primary subject matter of this paper, it is nonetheless a convenient vehicle for the formal expressions we wish to provide, and equally, we seek to support the claim that Mexitl is a complete formalism in the pragmatic sense mentioned above. Accordingly, a brief description of Mexitl will be included.

The remainder of the paper will be organized as follows. Section 2 reviews the Allen set, and also makes use of it to introduce the Mexitl language. We will also elaborate on some of the shortcomings of the Allen set in this application area. Section 3 presents an informal introduction to our view of item synchronization, and presents a number of illustrations. Section 4 defines synchronization events and synchronization items, and uses these notions to provide a new taxonomy of possibilities for media item synchronization. Section 5 uses these notions to develop a more complex example, that of lip-synching, and also includes a Mexitl version of this problem. Section 6 concludes.

2 Synchronization of Continuous Items: Allen and Mexitl

In this section, we are concerned with synchronizing two continuous items, where the synchronization is defined in a manner which is independent of any events or conditions occurring in the two items. The items, therefore, are to be regarded as indivisible. The taxonomy of binary relations between two temporal intervals introduced in Allen [All83] is well known, and has been used by a number of authors [JLSI97,KS95,Kin96,MHM96,SHT96]. We re-introduce it here for completeness and as a means of introducing some of the components of the Mexitl notation.