INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG2000/M5728
20 – 24 March 2000

Title: **Report of the AHG on MPEG-7 Semantic Information Representation**

Source: **Ad Hoc Group on MPEG-7 Semantic Information Representation**

Source: **R. Leonardi (Univ. of Brescia) on behalf of AHG co-chairs**
M. Tekalp (Univ. of Rochester), R. Mehrotra (Kodak), H.K. Rising III (Sony)

Status: **AHG Report**

---

1 Semantics DS AHG Mandates (from N3138 [1])

1. Provide a clear definition and role of each DS which is part of the revised *Semantic DS*, possibly providing examples for explanation.
2. Debate on the exact definition of the *Concept DS*, so that a unique interpretation can be provided.
3. Debate on the composition of the *State DS* and whether it provides a necessary functionality.
4. Clarify how semantic information can be embedded in the *Segment DS*, in particular whether attaching a *SemanticDescription DS* is the most appropriate way.
5. Get feedback from participants involved in the Validation Experiment on *Semantic DS*

2 Summary of AHG Activity

This AHG has been created at the last MPEG meeting in Maui to work in parallel with the CE experiment on the *Semantic DS*, so as to continue the refinement, both in terms of significance, usage and syntax of the DS's that have been proposed during the Maui meeting [2] . Following the discussions on the email reflector, the results of a meeting of the US delegation in February and of the discussion during the AHG meeting Mar. 19th, 2000, in Noordwijkerhout, some clarifications were made, though a total convergence has not yet been reached. During the US delagate meeting, an alternative syntax had also been proposed for consideration of the continuation of the CE; it is likely that this will lead to the formulation of competitive solutions to select the best syntax and elementary components of the *Semantic DS* during the CE process to follow after the 51st MPEG meeting. Below, some listings of the discussions that took place, in reference to the individual mandates of this AHG.

3 Results of Individual Mandates

The results of the work on the individual mandates are as follows:

3.1 Mandate 1

- **Provide a clear definition and role of each DS which is part of the revised *Semantic DS*, possibly providing examples for explanation.**

Result: There was a strong opinion shared on the reflector that parts of the *Semantic DS* could not be understood from the semantic point of view from output document N3123 of the Maui meeting, which limited the work that could be carried out for the CE. Participants who participated to the editing of document N3123, which took part of the break-out on Semantics in Maui, posted a definition of most items on the reflector (see in particular email of H.K. Rising, III, on xx.xx.2000), which is summarized hereafter:

- *Semantic DS*: container box of all semantically related components.
- *SemanticDescription DS*: In order to accommodate for multiple semantic descriptions of a single piece of content, one needs a DS that can occur multiply inside the *Semantic DS*. The *SemanticDescription DS* is meant to allow many different descriptions, even orthogonal in terms of semantics, with a same content, depending on the perspective of the description creator. That entity is not an object, since a semantic description is not an object, necessarily, likewise it is not an event, or a single concept, necessarily. Hence the *SemanticDescription DS* is a container element, of all the important components of the *Semantic DS* with the exception of holding a possible annotation (*Identifier DS* has a *Type D*, i.e. a single string, an unique *id*, and possibly an *annotation D* which could contain any free text), and contains a single attribute, the *UsageLabel D*. This binary label is either descriptive or accessing, depending on the function of the description (see Maui input document M5611 [3]).
- *Concept DS*: See mandate M2.

- *Event DS*: It represents a particular semantic situation (e.g., an interaction between objects, a goal in a soccer game, …) which can be well referenced in time and/or space. It may be composed of other events. This defines an event-event relationship of the type "is composed of". Events may be related to other events; in such case, the event-event relationship is of the type "any type" (though it may be precisely defined thanks to the *annotation D* which may be attached to the *Semantic Link DS*). When an event represent an interaction between objects, an event-object relationship of the type "involves" is created to link the different participating objects. To identify the reference in time of the *Event DS*, one or more temporal references (each representing a duration or a date) may be attached to it, the *STime DS*. To identify the reference in space of the *Event DS*, one or more spatial references (each representing a geographic location, defined by its coordinates or a name) may be attached to it. An *Event DS* may contain several *Medialocator DS* to link directly to the audio-visual content, which allow to identify the segments when the event is actually taking place. D's can be attached to an *Event DS* through the inclusion of a *Segment DS*. Finally, an *Event DS* may have a weight attached to it like for any other component of a *SemanticDescription DS*, the semantics of which may enable ordering, confidence, perceptual significance of the event itself with respect to other events forming part of the description.
- *Object DS*: It represents a particular piece of content (in time and space) to which a physical significance can be attached. It may be composed of other objects. This defines an object-object relationship of the type "is composed of". If an object interacts with different objects (i.e. other separate physical entities), the interaction constitutes an event, and the two or more objects interacting are connected to the event through an object-event relationship of the type "is present during", while the converse event-object relationship is of the type "involves". In some cases, when an object has a certain attribute related to an action it is performing which cannot be well referenced in time (in the sense that it is a permanent situation of the object), such action is called a "state", and defined through the *State DS*. An *Object DS* may contain several *Medialocator DS* to link directly to the audio-visual content. D's can be attached to an *Object DS* through the inclusion of a *Segment DS*. Finally, an *Object DS* may have a weight attached to it like for any other component of a *SemanticDescription DS*, the semantics of which may enable ordering, confidence, perceptual significance of the object itself with respect to other objects forming part of the description.
- *Semantic Link DS*: This DS allows to represent all types of relationships defined so far between objects, events, and their mutual interactions. Beyond being able to specify more precisely the type of existing relationship, through the use of an associated *annotation D*, there are 4 basic types of relationships:
  1. Is composed of: for object-object, event-event hierarchies.
  2. Is present during: for object-event interactions and event-event interactions.
  3. Involves: for event-object interactions and event-event interactions.
  4. Unspecified: for other types of event-event relationships (in this last case, it is expected that the *annotation D* be used for specifying the relationship).

  Note that object-object interactions (i.e., non hierarchies) need to be represented through the definition of an event, and two or more object-event relationships. Note also that events and objects are allowed implicitly to recurse through the usage of object-object relationships and event-event relationships of the type "is composed of". Finally, as in

the case of objects, events, and concepts, a weight DS may be attached to a particular relationship for comparison purposed with respect to other types of relationships.

- *STime DS*: This DS allows to represent time information associated with an event. It includes a *time DS*, or a *duration* or a *reference* to an already defined *STime DS*. In this last case, it is assumed that the same *STime DS* is defined elsewhere, and that it is simply referred to, here. A free text description may be attached to better explain the listed temporal references.
- *SLocation DS*: This DS allows to represent a location information associated with an event. It includes mandatorily a *name* or a *reference* to an already defined *SLocation DS*. In this last case, it is assumed that the same *SLocation DS* is defined elsewhere, and that it is simply referred to here. It is also possible to have in the *SLocation DS* a list of absolute spatial coordinates defining the locations, in terms of latitude, longitude and altitude. (Note: These may or may not be associated to the list of location names, according to the current syntax. This means that it is left to the application scenario to determine the usage of the absolute coordinate specification).
- *SWeight DS*: This DS has a *value* associated with it, a *target type* specifying its semantics, and an *id*. It should allow to rank order different components of the *Semantic DS*.
- *State DS*: See mandate M3.

Recommendation: Produce during this meeting a precise definition of each individual component for the *Semantic DS*, and verify its DDL syntax, and UML representation. In case the *Semantic DS* participants cannot converge on a unique syntax solution, all necessary clarifications should be provided for the individual components of this alternative solution as well to enable the success of a CE to follow.

3.2 Mandate 2

- **Debate on the exact definition of the *Concept DS*, so that a unique interpretation can be provided.**

Results:

This DS was introduced on the belief that a description might involve one or more concepts,

which would need further description, or could be taken from previously described work. It represents a template for a certain semantic item (modeled in abstract terms) which would enable a recipient of the description to have a priori knowledge of the model used by whoever created the description, even though only parts of the template are actually instantiated in reference to the audio-visual document being described. As such, a *Concept DS* enables a portable, non-specific description of an object, event, or combination of concepts, objects and events. A concept is inherently recursive; it is not a specific object or event. Portability means that a concept is the appropriate way to create a template for a description or part of a description that will be used often, as part or parcel of several semantic descriptions.
A concept allows the description of a combination of objects, events, and other concepts to be used to generate and complex or abstract concepts. The *Concept DS* can also cover the

creation of generic or abstract events/objects (needed as per documents M4754 Vancouver, M5190 Melbourne, M5593 Maui). A concept may include objects or events to create an instance of the concept through such objects or events (i.e., describe the concept by an example).
As most components of the *Semantic DS*, the *Concept DS* contains an *Identifier DS*, meaning that it has a single string (the *Type D*) assigning a semantic to it, an optional free text description, since the *Annotation D* is in the *Identifier DS*, and a unique *id* for referencing. It may have a weight attached to it like for any particular component of a *SemanticDescription DS*, the semantics of which may enable ordering, confidence, perceptual significance of the concept itself with respect to other concepts forming part of the description. In addition, it contains, objects, events, and other concepts. An item that clearly appeared to be missing from the document describing the CE (N3123) as part of the Semantic DS, is represented by the links and/or relationships between events and objects which are used to define the concept itself, and possibly links and/or relationships that may exist between various types of concepts which are not part of a same recursion.

Recommendation:
As the benefit for a *Concept DS* is not yet well accepted by all participants in the AHG, it is necessary to identify examples that demonstrate its benefit of its usage in the *Semantic DS*.
If examples are difficult to derive, a detailed document should be provided to understand the functionality of the *Concept DS*.

3.3 Mandate 3:

- **Debate on the composition of the *State DS* and whether it provides a necessary functionality.**

Result: *State DS* is a description scheme for referring to the state of an object, particularly when the object is in this state with continuity, without a specific time information that can be associated with it. It contains an *identifier*, hence an *annotation* and an *id*. States of objects are different from events, though some participants to the AHG believe that the state of an object can be represented thanks to a specific type of event and an adequate event-object relationship. The *State DS* appears useful for creating semantic states. In fact, there is a lot of justification for defining some events to represent a change in state of one or more objects. A *State* appears to represent a local or remote list of attributes describing the state of an object. The transition between states defines a certain type of events. In this respect, it should be noted that a subset of the list of attributes in a state is also a description of a state. This allows the implementation of *contexts:* multiple descriptions of the same object using different subsets of a list depending on which is appropriate to the object in the context in which it is being described in each description

Recommendation: Continue to address the meaning of such DS, and find concrete examples demonstrating the advantage of having this type of representation with respect to a simple event based description of object continuous state of affair.

3.4 Mandate 4

- **Clarify how semantic information can be embedded in segment DS, in particular whether attaching a *SemanticDescription DS* is the most appropriate way.**

Result: This activity has been delayed, due to the lack of convergence of ideas.

Recommendation: Get back to it once a core *Semantic DS* has been established.

3.5 Mandate 5

- **Get feedback from participants involved in the Validation Experiment on *Semantic DS*.**

Result: Apart from the discussion that took place regarding the incomplete status of document N3123, it appears clearly that a natural language description can be included in any part of the Semantic DS, considering the "free text" or "annotation" DS which can be attached at any level of a semantic description. It is thus necessary to define a series of criteria for the Core Experiment that ensure that the Semantic DS structure better represent the natural language description without (or with as little as possible) natural language description attached to the "annotation" portion of each DS. Another aspect which came from some of the CE participants is that the "conceptual modeling" process using entity-relationship is inadequate to represent concepts for the *Semantic DS*, and that an "object-oriented design" approach seems more appropriate.

Recommendation: Revise the CE methodology to enable a good performance evaluation of the proposed *Semantic DS* solutions. If possible, try to evaluate the different components independently so that a core DS structure can be used to represent semantic information. In order to reach satisfactory results, the following steps are considered as mandatory:

1. Select one or more syntax to represent semantic information.
2. Define precisely each component DS of the *Semantic DS*, providing possibly examples for clarification, and the functionality of such a component in the representation of semantic information.
3. Select a common corpus for the validation.
4. Define a methodology for the CE, trying to assess a "measurable" benefit of the proposed solutions, and demonstrating that much simpler one (such as "free text annotation") cannot reach the same performance.

4 References

[1] Ad Hoc Group on MPEG-7 Semantic Information Representation, *ISO/IEC JTC1/SC29/WG11 MPEG99/N3138*, Maui, HI, December 1999.

[2] Core Experiment on the Semantic DS, *ISO/IEC JTC1/SC29/WG11 MPEG99/N3123*, Maui, HI, December 1999.

[3] A Useful Generalization of the Syntactic/Semantic Division and its Application to Collections. *H.K. Rising and A. Tabatabai, Sony Corporation, contribution M5611* to MPEG99, Maui, HI, December 1999.