SMIL: Synchronized Multimedia Integration Language 2.0

Nabil LAYAÏDA
INRIA Rhône-Alpes – SYMM WG/W3C, Monbonnot
Nabil.Layaida@inrialpes.fr

Introduction

- Web evolution
  - Diversity of formats and platforms (no common exchange format, …)
  - Multimedia and the web were developed in distinct parallel worlds!
- Multimedia on the web: an integration problem at two levels
  1. Between media objects (mp3, video, text, ..)
  2. With the Web (web technologies)
- Context of this work : W3C
  - SYMM Working Group (SMIL 1.0 et 2.0)

Plan

- Introduction
- Goal and design principles
- Organization of a SMIL document (.smi)
  - Spatial et temporal aspects
  - Animations and transitions
  - Hypermedia time-based Links
  - Scalability framework
- Conclusions and future work

Goal of the project : version 2

- Document Text Format for the integration of media items (the html of multimedia).
- Use Web technologies where applicable for multimedia : XML, Namespaces, Schemas, ...
  DOM.
- Promotion of the notion of time-based and Synchronized documents at the scale of the web.
- Remain neutral to access protocols and to media formats RTP, RTSP, Mpeg,…
- Bring together major players in multimedia around an open format (the impossible challenge !).
Involved Companies and Organizations

- Application developers
  - Oratrix, Real Networks, Microsoft, IBM, Macromedia, Intel, Philips, Panasonic, Nokia

- Products
  - Public Institutions: INRIA, CWI, NIST, WGBH …Experimental Syst.

Strengths of SMIL
- Version 1.0 is a success … given the no of ??
- Very simple to learn and to use.
- Growing integration with other web standards.

SMIL 2.0: Design principles
Meta-language which allows the description of multimedia documents ranging from the simplest to the very complex.

SMIL 2.0: Functional spaces
The functional aspects covered in SMIL 2.0 are:

- Layout -- positioning on the screen and audio channels
- Content Control -- content selection, adaptation and optimization
- Structure -- the glue between the different modules
- Metainformation -- metadata
- Timing and Synchronization -- the heart of the language
- Linking -- hypermedia time-based navigation
- Media object -- basic media objects
- Time manipulations -- accelerate / decelerate time
- Transition effects -- fade-ins, visual effects …

SMIL 2.0: Languages space
A profile:
- A language corresponds to a particular application (DTD, Schema)
- A composition of the functional space (modules)
- Integration with extra-SMIL modules (Animation SVG)

SMIL 2.0 Language Profile (SMIL Profile):
- Successor of SMIL 1.0 (backward compatible)
- XML Language syntax + a semantics
- Composition of most of the SMIL 2.0 modules

SMIL 2.0 Basic Language Profile:
- Language for 3G phones and PDAs ...
- A scalability framework to handle heterogeneous devices

XHTML + SMIL
- Basic medias are XHTML elements
- Fusion (thanks to namespaces) of two XML languages

SMIL DOM 1-2
SVG
Animation
Transition
...
Typical SMIL Documents

• A set of “components” accessible via urls, the content is not included in a SMIL file
• These components may have different media types: audio, video, text, image, etc.
• Synchronization: intra- inter-objects et lip-sync
• User Interactions: TAC (Global and VCR like) and spatial and temporal links, dynamic changes to the course of a presentation (events)

Plan

• Introduction
• Goal and design principles
• Organization of a SMIL document (.smi)
  – Spatial et temporal aspects
  – Animations and transitions
  – Hypermedia time-based Links
  – Scalability framework
• Conclusions and future work

Organization of a SMIL document

Two parts:
• Head: contains information of document level
• Body: contains the temporal scenario, animations, transitions and the media objects

Structure of a document

toto.smi

head

Layout

Region

Audio Channel

Media

Transition

Animation

body

par.seq

switch

Media

Transition
Document Head

- META element: description of the document properties and metadata (RDF)
  - Title, author, expiration date, keywords, summary, …
  - … the MPEG 7 of SMIL!

Spatial Aspect

- Layout element
  - Hierarchical Model with *hotspots* (regPoints)
  - Layers instead of text flow (text flow for text!)
  - Simple Positioning close to the CSS Model 2+1/2 D (x, y, z-index, fit)
Spatial Aspect

Document Body: Synchronization

Contains the temporal scenario of the document
- A scenario is defined recursively: Schedule elements
  - Schedule = Parallel | Seq | Excl
  - Media object
    - anchors (starting/arriving)
    - Switch
    - priorityClass
    - Prefetch

Basic Media Objects

Media Objects marked-up with:
- Audio, Video, Text, Img, Textstream, Animation, Ref, Param, et ... Prefetch

Attributes:
- Src: identifies the basic media object file (URL)
  rtsp://rtsp.example.org/video.mpg
- Type: mime type (eg. video/mpeg)
- Region: identifies the drawing surface
- Dur: duration of the media object

Synchronization Attributes

The Dur (duration) attribute:
- “intrinsic”: the duration corresponds to the duration of the external file.
- “explicit”: the duration is specified in the document (dur= “15 s”)

The repeat attribute:
RepeatCount=“3” repeats the simple duration of the media.
RepeatDur=“12 s” :
Synchronization Attributes

The begin, end attributes:
- Value (begin= "13 s") : offset relative to the parent element.
- Reference to another clock : (begin= "e2.end + 5 s ")
- Reference to the absolute time reference: (begin= “wallclock(2001-01-01Z)"
- Reference to an asynchronous event (interactivity): (begin= “button.click”)

Media Clipping

- Spatial Clipping using regions and sub-regions
- Temporal Clipping using clip-begin et clip-end attributes (media objects are external files)

The sequential element : seq

- Semantics : play in sequence a set of media objects
- Attributes
  - Fill : used to make the object « persist » on the screen
  - Remove : removes the object at end time
  - Freeze : keeps the last frame at end time

```
<seq>
  <image id="a" regionName="x" src="wait.gif" fill="freeze"/>
  <video id="b" regionName="x" src="video.au" dur="20 s" />
</seq>
```

The parallel element : par (1)

- Semantics :
  - Play in parallel a set of media objects
  - End time : maximum duration of child objects
- Attributes :
  - endSync : Last (Rendez-vous)
  - Dur : reference clock of the par(Wall clock)
  - Begin/End : Synchronization Arc
Parallel element: \( \text{par} (2) \)

Synchronization Arcs and Events

Allows the description of graph structures:

```
<par>
  <audio id="a" src="audio.au" begin="id(b)" />
  <video id="b" src="video.au" end="id(c)" />
  <text id="c" src="text" begin="id(d)" end="id(a)" />
  <image id="d" src="image.gif" begin="id(a)" />
</par>
```

Triggering of objects on events:

```
<par>
  <img id="a" src="image" />
  <video id="b" src="video.mpg" begin="a.activateEvent" end="a.activateEvent" />
  <text id="c" end="b.focusInEvent" />
</par>
```

syncBehavior and syncTolerance

- syncBehavior
  - canSlip: the synchro is loose, child elements can slip from the parent clock
  - locked: the Synchronization is hard (lipsync), amount of tolerated slipping (syncTolerance).
  - Independent: synchro completely independent

- syncTolerance = "amount of jitter"
- syncMaster = "true" clock ticker of the par element

excl and priorityClass elements

- Semantics:
  - Play a set of media objects one at a time
  - End time: same as par/seq with the addition of

```
<excl dur="indefinite" endSync="all">
  <priorityClass id="ads" peers="defer">
    <video id="pub1" />
    <video id="pub2" />
  </priorityClass>
  <priorityClass id="program" peers="stop" higher="pause">
    <video id="program1" />
    <video id="program2" />
    <video id="program3" />
    <video id="program4" />
  </priorityClass>
</excl>
```
Switch element

- An element to choose from a set of content equivalent objects
- Choice is based on attribute values
  - language, screen size, depth, bitrate, systemRequired
  - … and user preferences

Plan

- Introduction
- Goal and design principles
- Organization of a SMIL document (.smi)
  - Spatial et temporal aspects
  - Animations and transitions
  - Hypermedia time-based Links
  - Scalability framework
- Conclusions and future work

Animations

Definition:
- A set of attributes are target of the animation
- A function (calc mode) makes these attributes evolve
- A control on the instants where the changes are applied

Syntax
- `animateMotion`: graphical movements of elements
- `animate`: generic animation applied to element attributes `from/to/by/calcMode`
- `set`: discrete change of an attribute value at a given instant
- `animateColor`: animation in the color space

Animations

Calc Mode: discrete, list of values with linear, log interpolation
Transitions

Element: transition
- **Type** and **Subtype** (transition repository + variant)
- **transIn** and **transOut** attributes

```xml
<transition id="wipe1" type="zigZagWipWipe" subtype="leftToRight" dur="1s"/>
<transition id="wipe2" type="veeWipe" subtype="leftToRight" dur="1s"/>
</seq>
```

Hypermedia time-based links

- Compatible with (Xlink/Xpointer)
- Extension to the semantics of URLs
  - `http://foo.com/path.smil#ancre(begin(id(anchor)))`
  - Two types (a: whole object, area: part of it)
  - Jump over time and space !!
- Attribute show
  - Replace (default value)
  - New (fork)
  - Pause (procedure call)

Plan

- Introduction
- Goal and design principles
- Organization of a SMIL document (.smi)
  - Spatial et temporal aspects
  - Animations and transitions
  - Hypermedia time-based Links
  - Scalability framework
- Conclusions and future work
Combination of both ...

With the animation of coords

A word on scalable profiles

Based on CC/PP and static content negotiation: user agent and server
Correspondence between the namespace prefix and module names (URIs)
Uses the `systemRequired` attribute and the `switch` element

Scalability

An example of capabilities description

The user agent must support SMIL 2.0 entirely

The user agent must support `time+contain+media` modules

Prefetching strategies

Goal: optimize the QoS by reducing the download delays: explicit method
An example

Prefetch an image so it is made available for display immediately after the video:

```xml
<smil xmlns="http://www.w3.org/2001/SMIL20/CR/Language">
    <body>
        <seq>
            <par>
                <prefetch id="endimage" src="http://www.example.org/logo.gif"/>
                <text id="interlude" src="http://www.example.org/pleasewait.html" fill="freeze"/>
            </par>
            <video id="main-event" src="rtsp://www.example.org/video.mpg"/>
            <img src="http://www.example.org/logo.gif" dur="5s"/>
        </seq>
    </body>
</smil>
```

The prefetch element

- The prefetch element gives the authors the control to enhance network transfers.
- SMIL documents must be playable even if prefetch elements are ignored.
- If a prefetch element is ignored, its Synchronization must be enforced, e.g. if a prefetch element has a `dur="5s"`, depending elements must behave accordingly.

The prefetch element

The prefetch element supports the following attributes:

- `mediaSize` values: `bytes-value | percent-value`
  - Defines how much of the resource to fetch as a function of the file size of the resource. To fetch the entire resource without knowing its size, specify 100%. The default is 100%.

- `mediaTime` values: `clock-value | percent-value`
  - Defines how much of the resource to fetch as a function of the duration of the resource. To fetch the entire resource without knowing its duration, specify 100%. The default is 100%.
  - For discrete media (non-time based media like text/html or image/png) using this attribute causes the entire resource to be fetched.

- `bandwidth` values: `bitrate-value | percent-value`
  - Defines how much network bandwidth the user agent should use when prefetching. To use all that is available, specify 100%. The default is 100%.

Plan

- Introduction
- Goal and design principles
- Organization of a SMIL document (.smi)
  - Spatial et temporal aspects
  - Animations and transitions
  - Hypermedia time-based Links
  - Scalability framework
- Conclusions and future work
Conclusion

• More visible impact on industry: HTML browsers (IE), ++ browsers (RealOne), ++ authoring tools, ++ smil servers.
• Declarative Markup and specification very appreciated.
• 3GPP adopted SMIL Basic for MMS.
• XMT – Part of Mpeg 4 uses SMIL syntax.
• SVG+Animation Profile (Adobe, …).

Perspectives

• Finer Control on the text media: timed-text (RealText), audio, …
• “streamable” SMIL for real time transmissions.
• SMIL 2.0 DOM: API for the scripting of multimedia presentations (atomic updates, affects the timing model, …).
Web Site: http://www.w3.org/AudioVideo/