SMIL:
Synchronized Multimedia Integration Language

Lloyd Rutledge and Lynda Hardman

Multimedia and Human-Computer Interaction Group
CWI: Centrum voor Wiskunde en Informatica
Amsterdam, The Netherlands
W3C SYMM working group

The Problem: Multimedia

Lots of Bits

- Images, audio and video are beyond Internet design specs
- Results in space/time constraints at:
  - the server
  - the network(s)
  - the client

Not All Bits are Equally Important

- Time between samples often more important than bits in sample, for example lip synchronization (but not always...)

Content may be Distributed Across Network

- Need to synchronize presentation

SMIL

Synchronized   Multimedia   Integration   Language

SMIL is about timing...

SMIL

Synchronized　**Multimedia**　**Integration**　**Language**

SMIL is about timing...

not just graphics...

SMIL

Synchronized   Multimedia   Integration   Language

SMIL is about timing...

not just graphics...

combining Web resources...

SMIL

Synchronized   **Multimedia**   **Integration**   **Language**

SMIL is about timing...

not just graphics...

combining Web resources...

in an XML syntax

Synchronized Multimedia Integration Language (SMIL)

Main Points

- Pronounced *smile*
- Multimedia for the Web — for multimedia what HTML is for hypertext
- Integration format for presentable mono-medium formats

Structure

- *SMIL 2.0* is a "meta-language"
- *SMIL Profile*, *SMIL Basic* and *XHTML+SMIL* set as among possible subsets

Status

- SMIL 1.0 became W3C Recommendation on 15th June 1998
- SMIL 2.0 is now in Last Call and will become a W3C Recommendation soon
  - includes SMIL Profile and SMIL Basic
- XHTML+SMIL comes after SMIL 2.0

Main themes

- Powerful timing and synchronization
- Adaptive to users and systems
- Models a flexible but consistent presentation and user interface

SMIL Applications
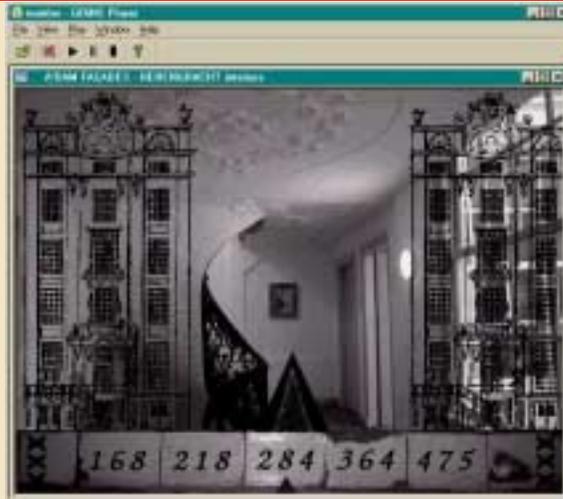
Infotainment

Accessibility

Conceptual Art

SMIL 2.0 extension over SMIL 1.0

Much Much More

- SMIL 1.0 spec is 30 pages, SMIL 2.0 spec is 300 pages

Animation

- Values of SMIL constructs change over time
- Enables more vibrant presentation
- Incorporation with SVG

Timing Integration

- Use of SMIL constructs in other document sets
- Enables, for example, HTML+SMIL in Internet Explorer
- Raises issues of semantic significance of hierarchy

Broadcasting/streaming

- No preload or full download
- Use of non-predictive events in timing
- Need to maintain hard synchronization
- Large potential use of SMIL

SMIL 2.0 Modules

SMIL is broken up into separate modules

- Thus not all of SMIL 2.0 needs to be used in one instance

The SMIL 2.0 Sections of Modules are:

- Animation
- Content Control — *selection, adaptation and optimization*
- Layout
- Linking — *navigation*
- Media Object — *media content that is integrated into presentation*
- Metainformation — *machine-processible data about the presentation*
- Structure — *base elements for high-level SMIL structure*
- Timing and Synchronization — *~100 pages!!*
- Time Manipulations — speed of integrated media
- Transition Effects — *fades and wipes*

SMIL 2.0 Profiles

What is a profile?

- A language for which a browser can be built
- A combination of modules from the SMIL 2.0 "meta-language"
- Possibly non-SMIL constructs with SMIL constructs

SMIL 2.0 Language Profile (SMIL Profile)

- What is typically thought of as SMIL 2.0
- Most of SMIL 2.0 features in one profile

SMIL 2.0 Basic Language Profile (SMIL Basic)

- Intented for mobile devices
- Assumes restricted processing ability

XHTML+SMIL

- Applies timing to text-based display
- XHMTL-based layout

SMIL 1.0

- Backwards-compatable — can be played on SMIL Profile browsers

SMIL Implementors

RealNetworks

- RealPlayer 8 — time-focussed media types
  - Current SMIL 1.0 support, anticipated SMIL Profile support
- 3rd party creation tools
- Clear Leader for SMIL Players

ORATRiX

- GRiNS authoring environment and free player
  - current SMIL 1.0 and anticipated SMIL Profile authoring
  - current free player for SMIL Profile

Microsoft

- Internet Explorer 5.5 supports XHTML+SMIL prototype
  - support for XHTML+SMIL standard anticipated with its release

Apple

- Quicktime 4.1 supports SMIL 1.0

A Sample Presentation

The Network News

On demand on your screen

Formatted text, video and audio

Local anchor setup

Top Story:

Growth of the
World-Wide Web
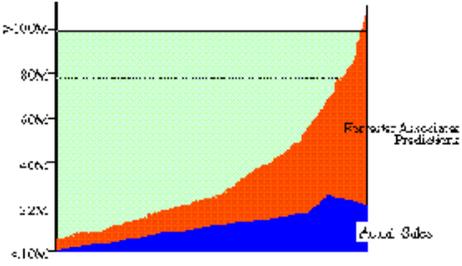
Amsterdam

Graph appears during spoken commentary

Remote Correspondent

Top Story:

Growth of the *World-Wide Web*

Amsterdam

Los Angeles

First video finishes, second video plays

Following a Link

Top Story:

Growth of the
World-Wide Web

Amste...

Los Angeles

File   Edit   View   Go   Bookmarks   Options   Directory   Window   Help

Dick Bulterman   **CWI**

HOME
SWEET
HOME

Document: Done.
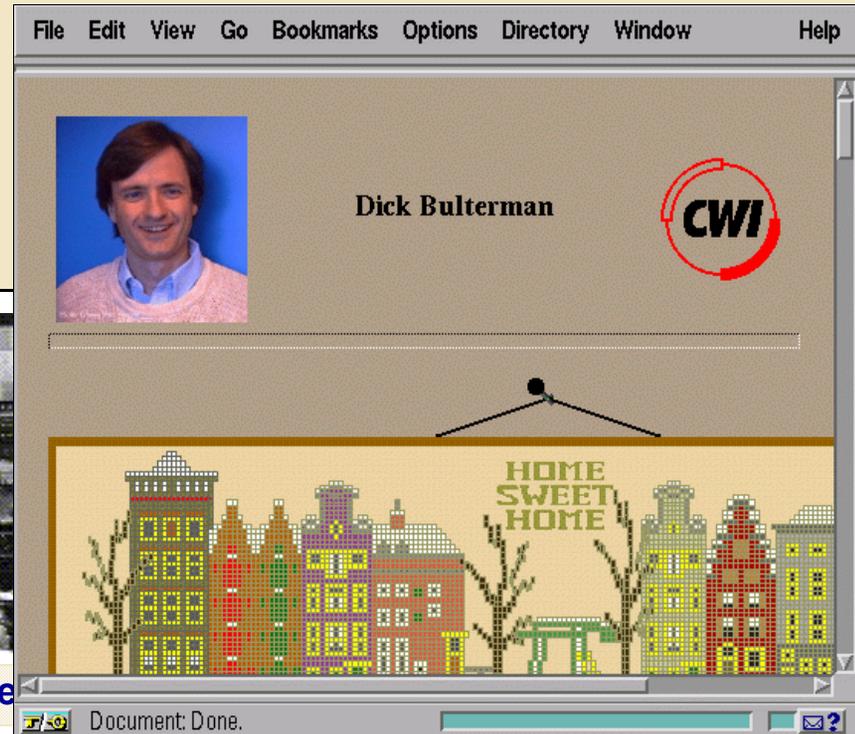
At any point during the video the viewer can request extra information.

CWI spin-off Oratrix

GRiNS market leader in SMIL authoring systems

Distribution agreement with Real Networks

Co-founder Oratrix

- prize-winning business plan for McKinsey's New Venture 1998

So what do we need to specify?

Content

(part of) media item

Spatial layout

Alternative content

bandwidth
task
user characteristics
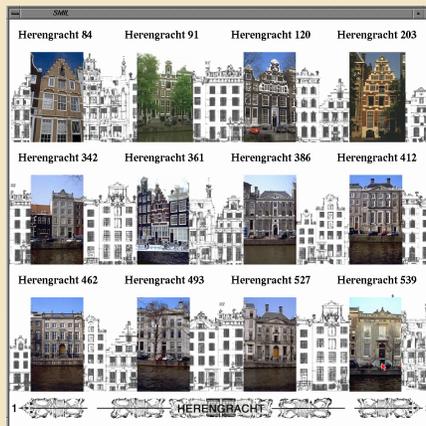
Semantic annotations

meta-data

Temporal layout

Links

source and destination
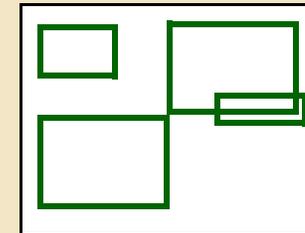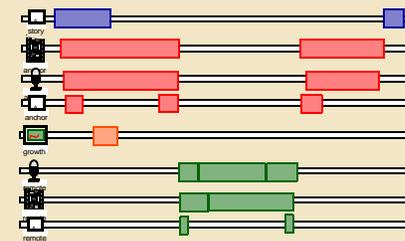
SMIL as XML Markup

Integration language

- Media elements referred to, not included

SMIL is XML

- Defined with XML DTD
- Can be hand-authored
- Declarative language
  - attribute/value pairs
- Integrable with XML environments

Relationship with Other W3C Recommendations

- Again, SMIL is XML
- Basic layout isomorphic and replacable with CSS
- Shares constructs with (X)HTML
- SMIL 2.0 "Family" languages enable new SMIL-based XML formats

XML

Foundation Syntax for all Documents

xxxxx

xxxx=xxxx
xxxx=xxxx

file

xxxxx

xxxx=xxxx
xxxx=

xxxxx

xxxx=xxxx
xxxx=

text content
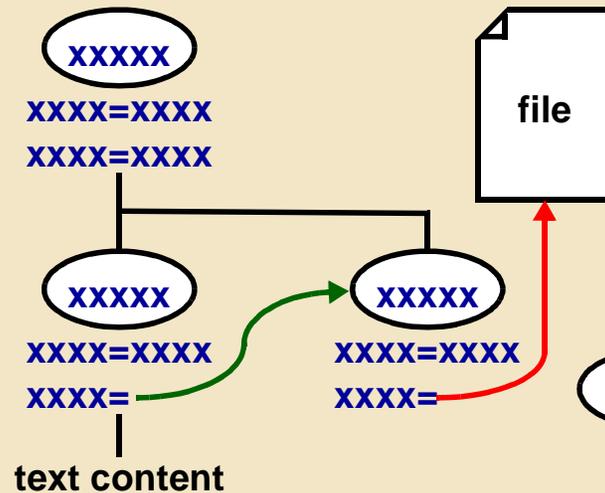
Document Type Definitions (DTDs)

smil

media file

head

body

layout

par

seq

ref

region
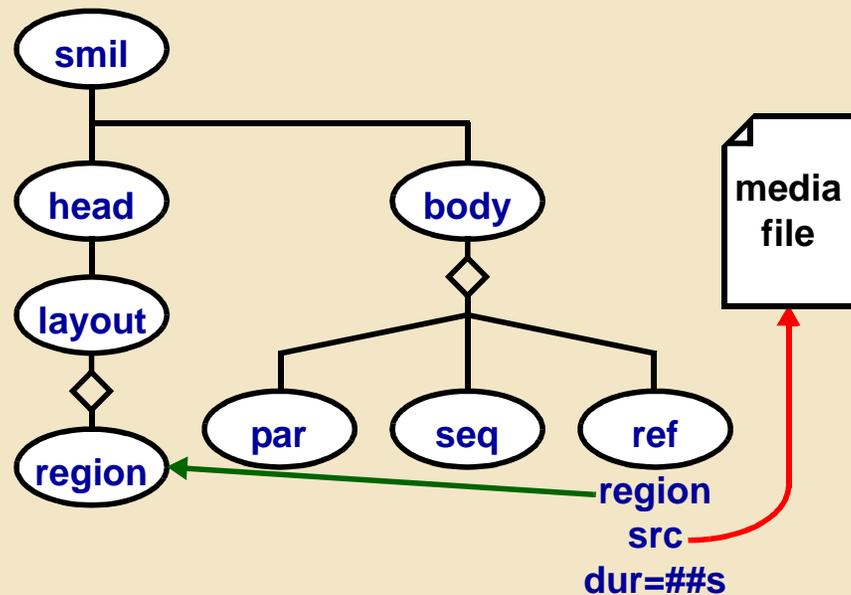
region
src
dur=##s

An XML (SMIL) Document

```xml
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE smil PUBLIC "-//W3C//DTD SMIL 1.0//EN"
                      "http://www.w3.org/AudioVideo/Group/SMIL10.dtd">
<smil>
  <head>
    <meta name="sync" content="soft"/>
    <layout>
      <root-layout id="SMIL-" width="492" height="810"/>
      <region id="address-region" width="50%" height="8%"/>
      <region id="image-region" top="8%" height="91%"/>
    </layout>
  </head>
  <body>
    <seq>
      <par>
        <text type="text/plain" region="address-region"
              src="Herengracht284.txt" dur="2s"/>
        <img region="image-region"
              src="http://www.amsterdam.nl/bmz/adam/pics/h284.jpg"/>
      </par>
      <par>
        <text type="text/plain" region="address-region"
            src="Herengracht539.txt"/>
        <img region="image-region"
            src="http://www.amsterdam.nl/bmz/adam/pics/h539.jpg" dur="2s"/>
      </par>
    </seq>
  /body
</smil>
```

Content — Instance of Media Item
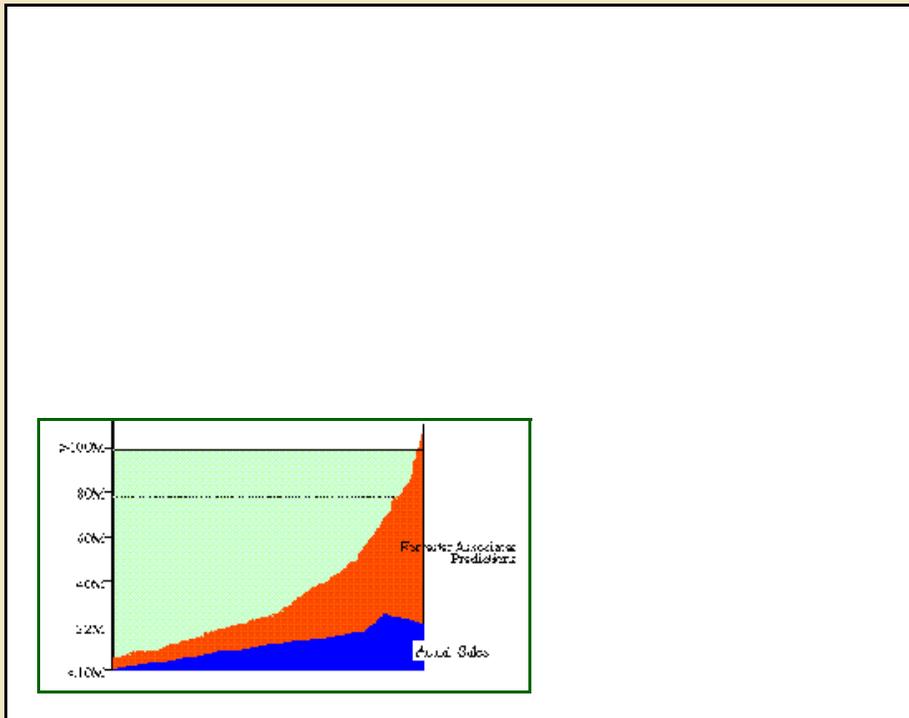
I   Media item, or part

II   Extent, position and z-index

III Duration

IV Alternate Content

V   Link end-points

VI Associated semantics

We will return to all these points at the end.

Media Object Elements

ref, **text**, **textstream**, **img**, **audio**, **video** **and** **animation**

```
<ref src="anything.???" ... />
<text src="caption.html" ... />
<textstream src="stockticker.rtx" ... />
<img src="graph.jpg" ... />
<audio src="http://www.w3c.org/SYMM/joe-audio.wav" ... />
<video src="rtsp://www.cwi.nl/SMIL/video.rm" ... />
<animation src="cute.anim" ... />
```

|			
video	audio	text	images
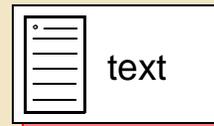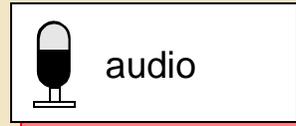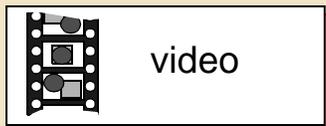
The `src` attribute is a URI, locating the data.

Names are for readability and are not used for determining data type.

# Specification of part of media item
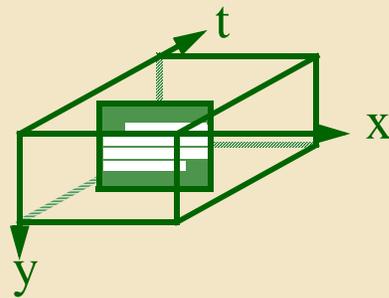

text — string
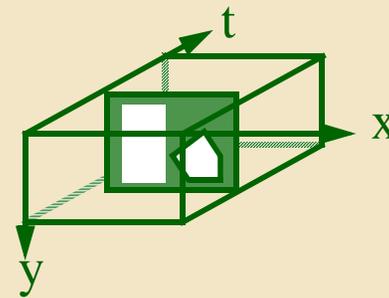
image — area
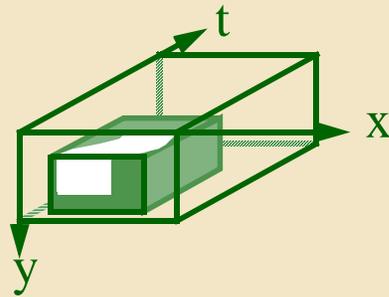
video — (moving) area
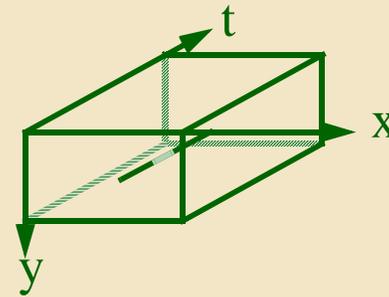
audio — phrase

Clips in time

Time and space treated independently.

- Spatial clipping done via region mechanism, discussed later

- Time restricted to a single extent
  - a contiguous section of a continuous media object can be specified

t_1 t_2

duration
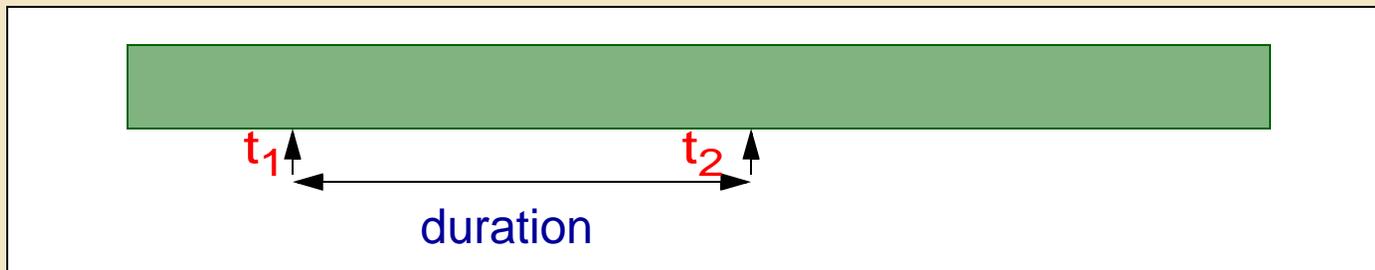
The `clipBegin` and `clipEnd` attributes

```
<video src="the.news/mpeg/zoomin.mpv"
    clip-begin="smpte=00:01:19:20"
    clip-end="smpte=00:01:38:40" ... />
```

- syntax of values is ...

New Media Constructs in 2.0

Parameter Control

- Application of media-specific parameters to media playback
- Handling of repeat instrinsic to media
- What to do when media ends

Media Clips Markers

- Use of media clips defined internally in media

Brush Element

- Paints a solid color on the screen

Spatial layout

**Top Story:**

Growth of the
World-Wide Web
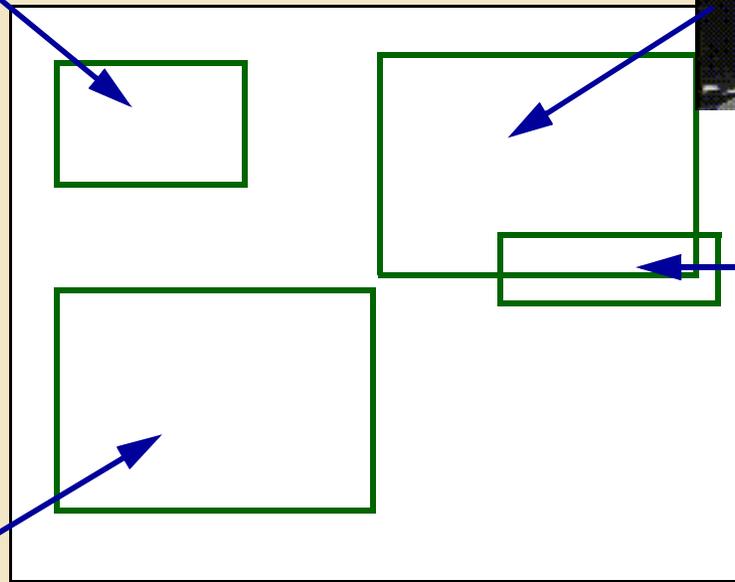
Amsterdam
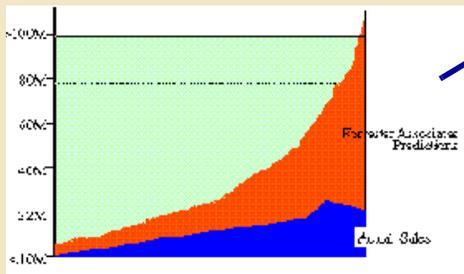
Possible ways to specify layout
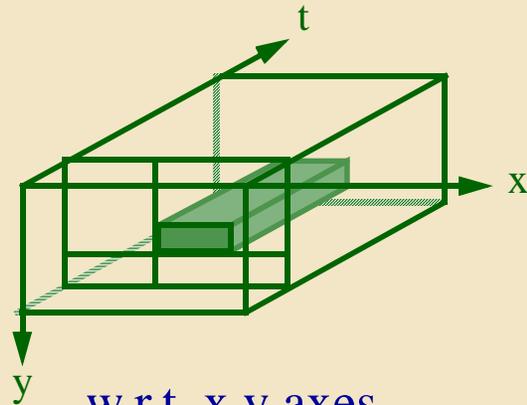
w.r.t. x,y axes
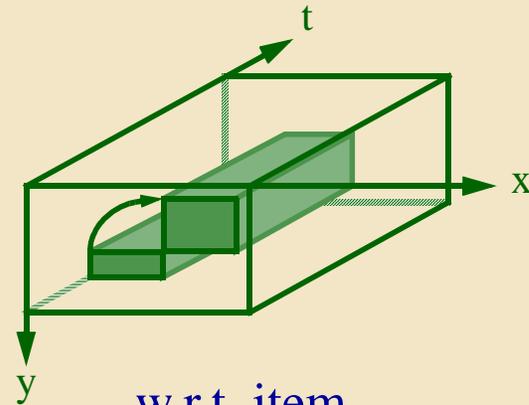
w.r.t. item
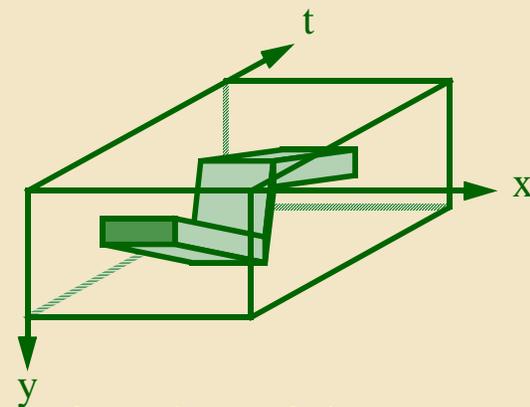
function of time
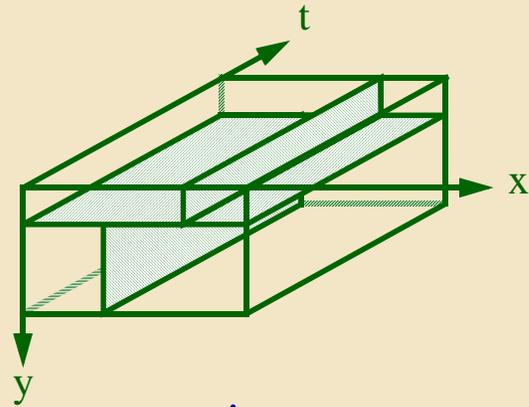
regions

Region

Each media object instance contains a region reference:

- allows author to know where object will be played

  `<video src="anchor.mpg" region="V-main" />`

The `region` is defined by:

x,y ● W H and Z
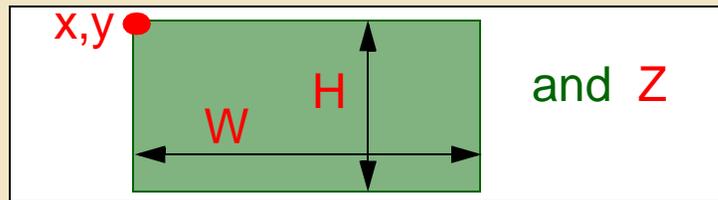
```
<region id="V-main" top="5%" left="50%" height="100%"
    width="100%" z-index="3" />
<region id="V-remote" top="10" left="100" height="200"
    width="200" z-index="3" />
```

- An "id" for each region is required. Its value is an XML identifier.
- Length values are percentage values or in pixels. The unit "px" may be omitted.
- The z-index gives the stacking order (highest integer stacks on top).

The WebNews Layout

T-title

root-layout

V-remote

Top Story:

Growth of the
World-Wide Web

Amsterdam

Los Angeles

V-main

T-anchor
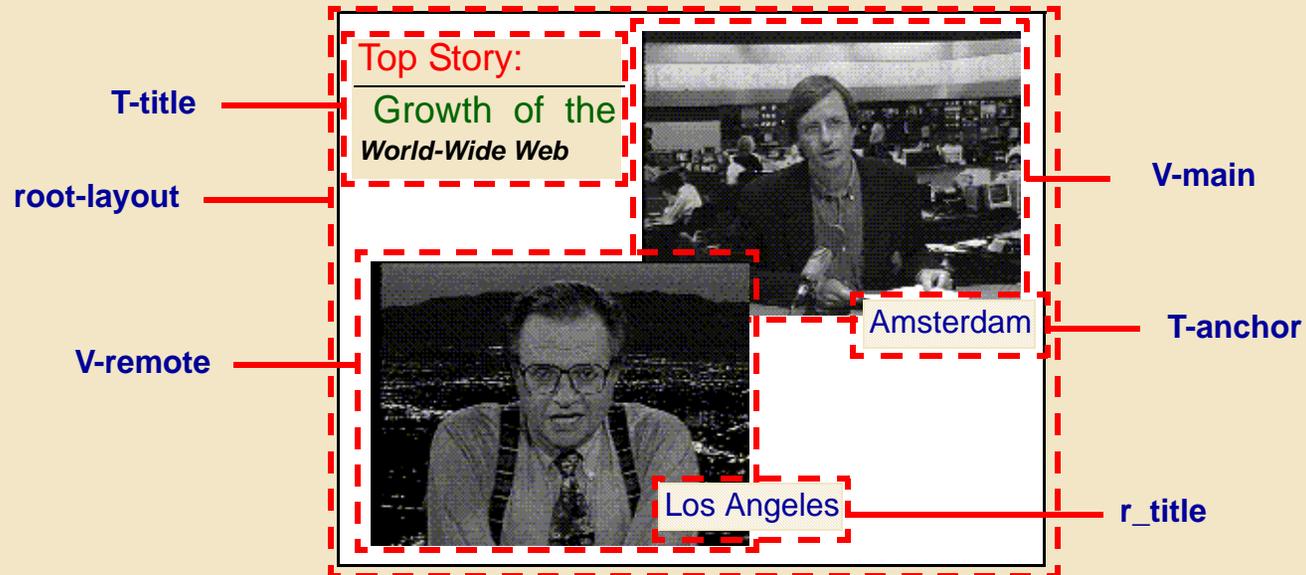
r_title

```
<layout>
 <root-layout width="721" height="587" id="matise" />
 <region id="T_title" left="2%" top="5%" width="40%"
        height="24%" z-index=2 />
 <region id="V-remote" left="3%" top="44%" width="54%"
        height="40%" z-index=3 />
 ...
</layout>
```

Clips in space

The `fit` attribute

hidden (default)

**media item
not scaled**

hidden (default)

**media item
not scaled**

scroll

**media item
not scaled**

meet

**aspect ratio
preserved**

slice

**aspect ratio
preserved**

fill

**aspect ratio
not preserved**

Layout Adaptation in SMIL

SMIL documents can adapt to devices with different screen sizes
- layout relative to the dimensions of the player's viewport
- alternative layout strategies

Switch on `layout` and `region`
- Allow assigning test attributes to SMIL `layout` and `region` elements
- Examples
  - make room for subtitles
  - rearrange for varying screen size

New Layout Constructs in 2.0

Audio Control

- Adjustment of volume of integrated audio media

Multiple Windows

- Regions placed in one of many windows

Hierarchical Layout

- Regions placed within regions
- Introduced relative placement of regions

Extended Adaptivity

- Adaptivity of layout components rather than choosing between layouts

(Potential) Use of CSS with SMIL

CSS in Current HTML

- Attribute transform
- Mapping to text-based presentation model
- Next step: applies to (text-based) XML

CSS in Current SMIL

- Alternative to SMIL Basic Layout
  - Not currently implemented
- Layout
  - isomorphic with CSS2
  - CSS2 as alternative layout
- With embedded XML/HTML for text

SMIL Basic Layout Positions Entire Media Object

- If media object is XML, CSS2 can specify object's appearance in its position
- CSS2 specifies aspects of object

Example: CSS with HTML

```
A:link {
  color: blue;
  text-decoration:underline;
}
A:visited {
  color: purple;
  text-decoration: underline;
}
P.Body {
  text-align: left;
  text-indent: 0.000000pt;
  margin-right: 18.000000pt;
  margin-left: 18.000000pt;
  font-size: 11.000000pt;
  font-weight: medium;
  font-style: Regular;
  color: #000000;
  vertical-align: baseline;
  text-transform: none;
  font-family: Minion;
}
```

III Temporal Layout
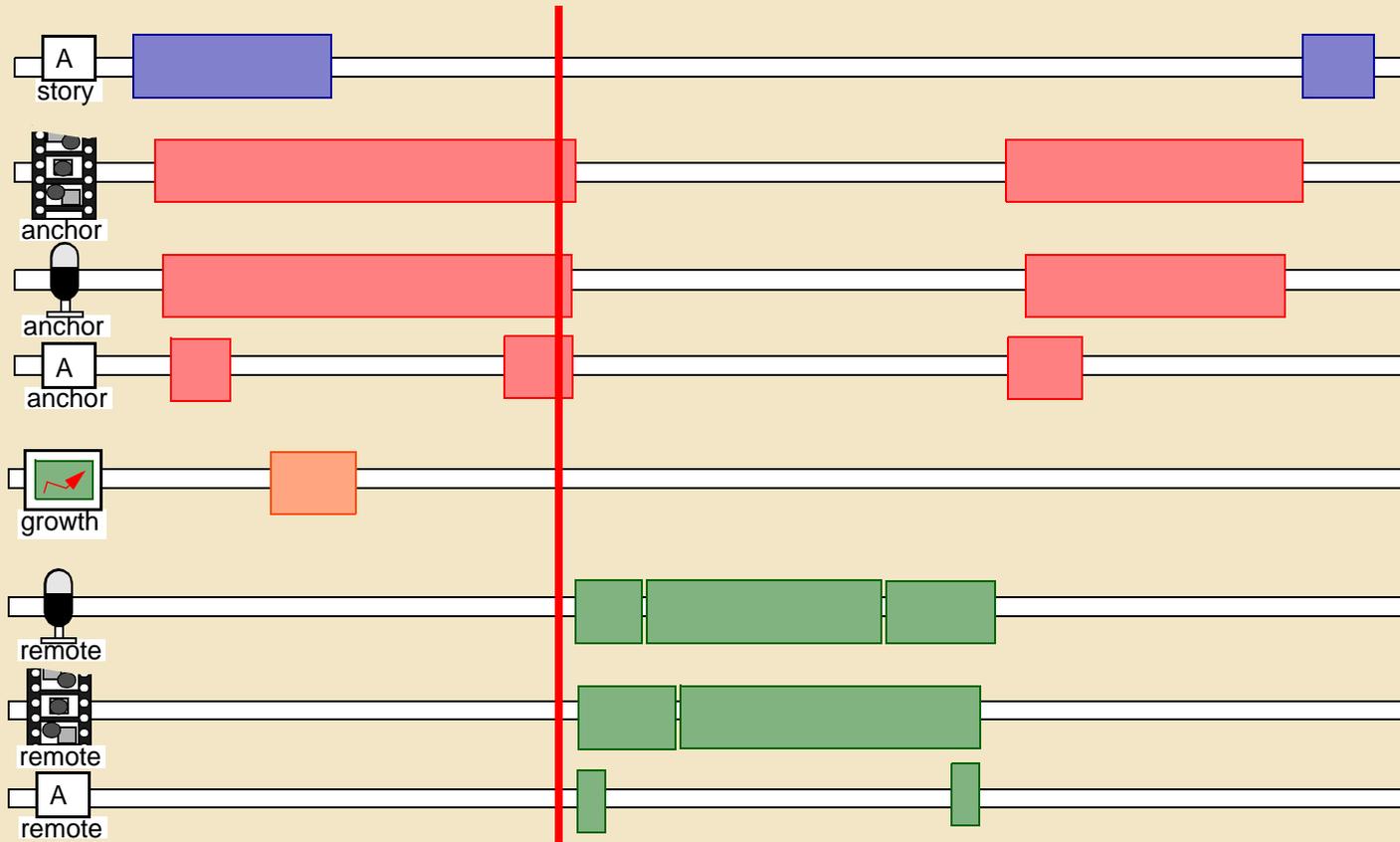
Which time?

Types of time:

- media item time axis
  - video divided in frames, audio sampled at 44kHz

- document time
  - image starts at certain time and ends at a later time

start time   end time

- run-time presentation
  - video data bits get caught up in network, so end time is delayed

start time   end time

Duration of a media object element

Intrinsic

- derived from content of media item

    audio (or video) lasts 5.3 seconds

- intrinsic duration of discrete media, such as text or image, is zero.

Explicit

- an explicit duration can be given
  The **dur** attribute, value is a clock-value or "indefinite".

  `<video src="zoomin.mpv" region="V-main" dur="4s" />`

    media object stops after 4 seconds

  `<video src="zoomin.mpv" region="V-main" dur="6.5s" />`

    media object stops after 6.5 seconds

  - in this case, the audio track just stops and the last frame of the video remains

Duration of a Media Object Element ctd.

An object can have its duration extended by repeating the content.

The `repeat` attribute

```
<video src="zoomin.mpv" region="V-main" repeat="3" />
```
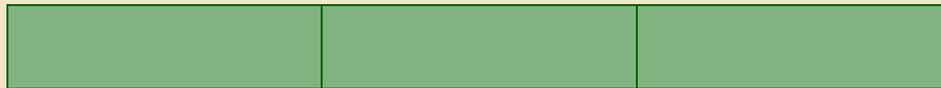
media object stops
after 15.9 seconds

```
<video src="zoomin.mpv" region="V-main" repeat="3" dur="11s" />
```

media object stops
after 11 seconds

```
<video src="zoomin.mpv" region="V-main" repeat="indefinite" />
```
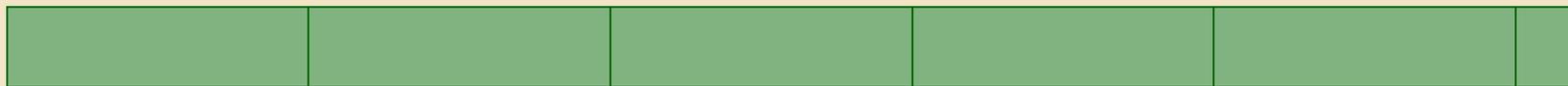
media object stops
when parent stops

Attribute value of repeat is an integer or "indefinite".

Start time of elements—par

The `par` element groups elements which are played in parallel

- Children of a par element are started at the same time

```
<par>
 <text src="leader_title.html" region="m_title" dur="5s" />
 <video src="cnn.mpv" region="V-Main" />
 <audio src="cnn.aiff" region="music" />
</par>
```
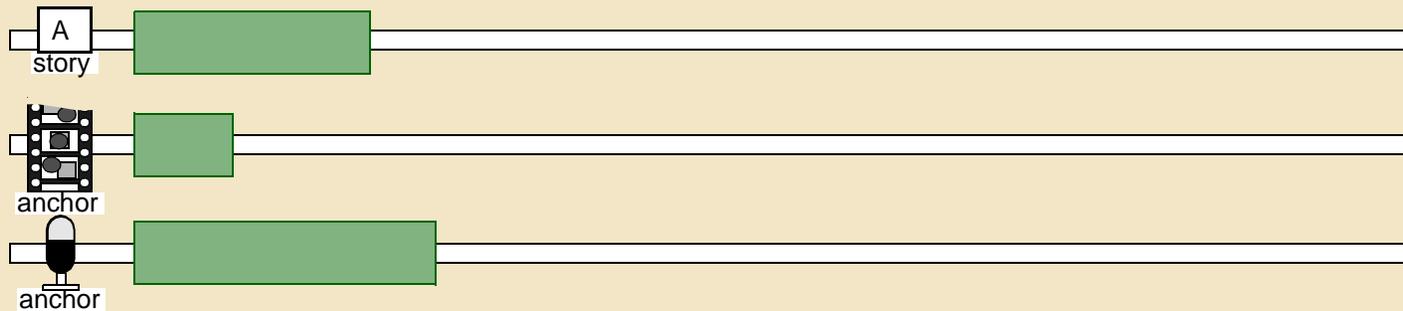

- The start time of a child of a **`par`** element is equal to the start time of the **`par`** element itself.

Start time of elements—seq

The `seq` element groups elements which are played sequentially

- Children are played one after the other, based on the textual order

```
<seq>
 <video src="logo.mpv" region="V-main" />
 <video src="anchor.mpv" region="V-main" />
</seq>
```
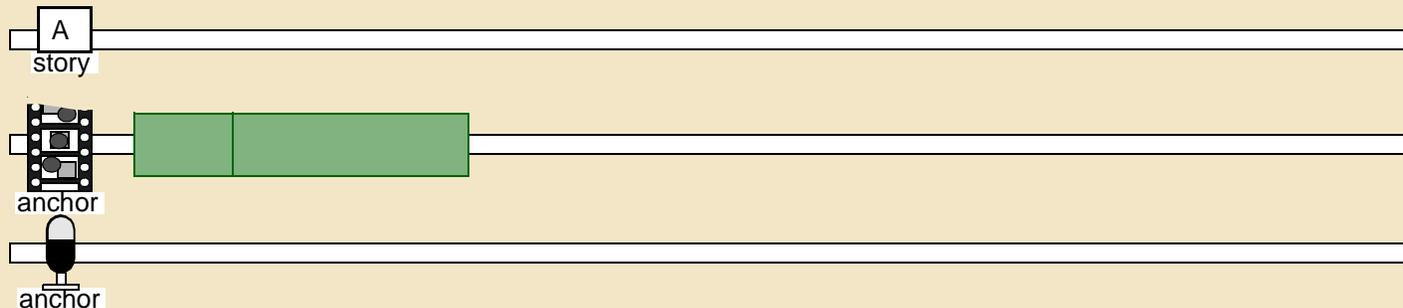
- The start time of the first child of a seq element is
  the start time of the seq element itself.
- The start time of the next child is
  the end time of the previous child.

Par's and seq's can be nested

```
<seq>
 <par>
      <text src="leader_title.html" region="m_title" dur="5s"/>
      <video src="cnn.mpv" region="V-Main" />
      <audio src="cnn.aiff" region="music" />
 </par>
 <par>
      <text src="story_title.html" region="m_title" dur="2s" />
      <video src="anchor.mpv" region="V-Main" />
      <audio src="anchor.aiff" region="music" />
 </par>
<seq>
```
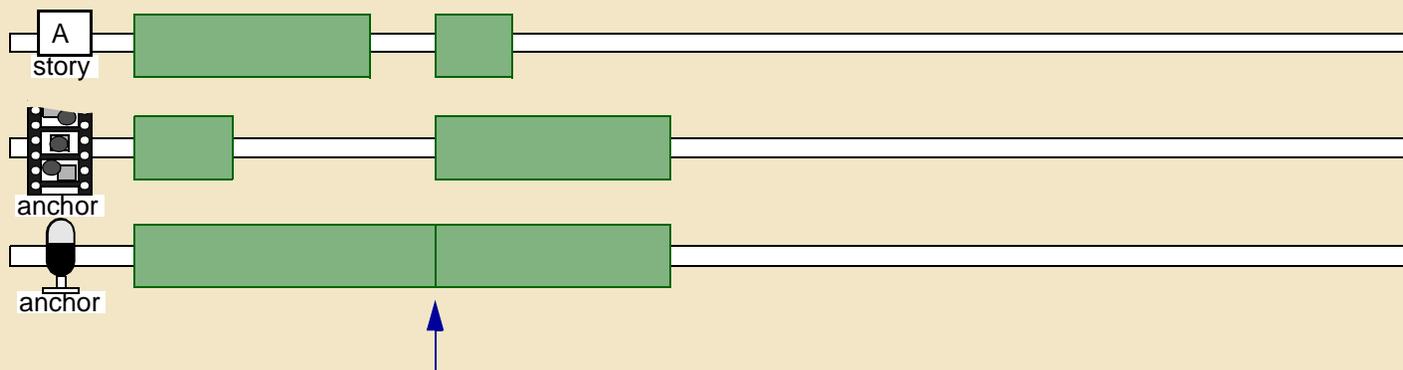
Explicit start time in a par element

The `begin` attribute, delay-value

```
<par>
 <text src="leader_title.html" region="m_title" dur="5s" />
 <video src="cnn.mpv" region="V-Main" begin="1.4s" />
 <audio src="cnn.aiff" region="music" />
</par>
```


- Video is delayed until 1.4s after the start of the **par** element.

Start time relative to another element

The `begin` attribute, event-value

```
<par>
 <text src="leader_title.html" region="m_title" dur="5s" />
 <video id="v1" src="cnn.mpv" region="V-Main" begin="1.4s" />
 <audio src="cnn.aiff" region="music" begin="id(v1)(0.5s)" />
</par>
```

- Audio is delayed until 0.5s after the start of video element "v1".

End time of media object element

A media object element with an implicit or explicit duration and a start time has an end = begin + duration.

```
<video src="cnn.mpv" region="V-Main" begin="4s" />
```
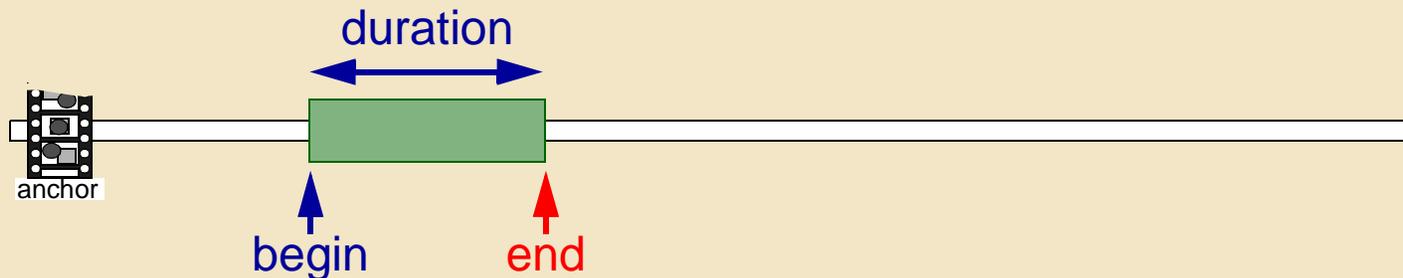
duration

begin end

The end attribute.   Syntax same as begin attribute.
A media object element with an explicit start time and an explicit end has a duration = end - begin.

```
<text src="title.html" region="m_title" begin="4s" end="8s" />
```

duration

begin end
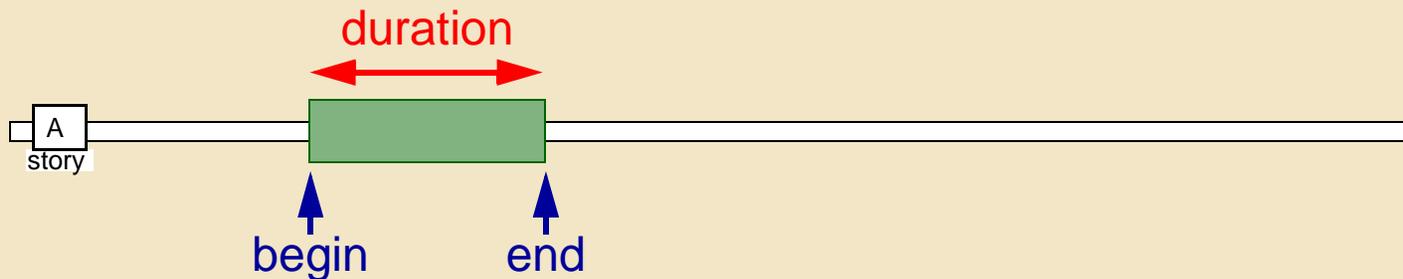
End time of par element—first

endsync

- **par** can end when the **first** element to finish ends

```
<par endsync="first">
 <text src="leader_title.html" region="m_title" dur="5s" />
 <video id="v1" src="cnn.mpv" region="V-Main" begin="1.4s" />
 <audio src="cnn.aiff" region="music" begin="id(v1)(0.5s)" />
</par>
```


end time
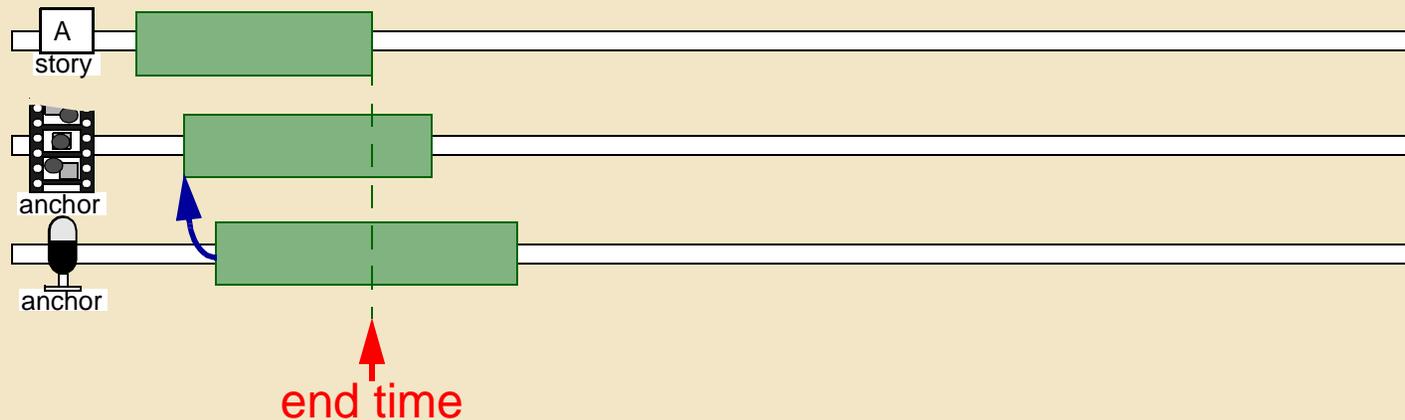
- **par** can end when the referenced element ends: **id(Id-value)**
- **par** can end when the **last** element to finish ends (default)

New Timing Constructs in 2.0

Animation

- Changing of numeric constructs over time — such as region placement
- Applied to SVG

Transitions

- Standard list of types, with timing

Manipulation

- Changing of media playback speed

Events

- List of DOM events that can trigger SMIL timing

SMIL 1.0 restrictions lifted

- Long synchronization
- Negative begin times

Much, much more ...

Alternate content

... explosive growth of the WWW ...

audio

text
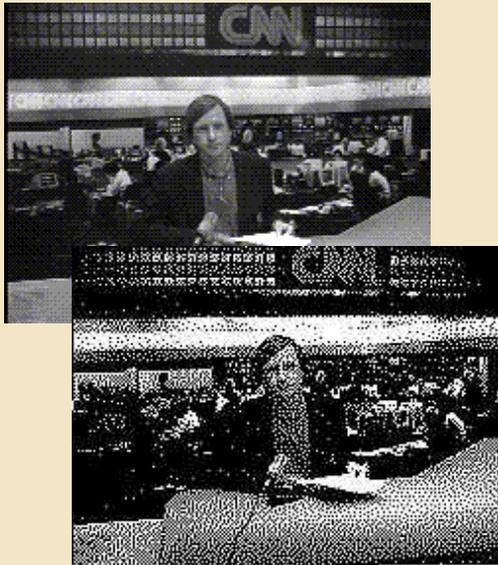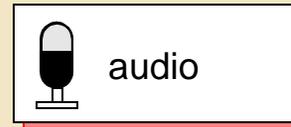
... explosive growth of the WWW ...
... explosieve groei van het WWW ...
... eksplozivni rast WWW ...
... crescita esplosiva della WWW ...

Adaptation Issues

Adaptation for User

- Disabilities
- Language
- Previous knowledge

Adaptation for Environment

- Delays: bandwidth, available CPU time
- Available processing: media peripherals, browser additional features

Adaptation for Document Purpose

- Selection of appropriate content
- Media items have different meanings in different focus
- Progression of presentation to meet purpose

W3C Web Accessibility Initiative (WAI)

- Guidelines for accessible (text-based) Web documents
- Meaningful values for attributes like `alt`, `title`, `abstract` and `longdesc`
- Meaningful content of link triggers (`a` element)
- How to apply these to a fixed timeline?

Specifying Adaptation in SMIL

Temporal Adaptation

- Handling delays of download and processing
- Explicit and implicit time
- Temporal hierarchy of parallel and sequential composites
  - sets points in presentation progression for stronger sychronization

`switch` Element

- At most one of the children of a switch element is played.
- The first acceptable element is chosen, so ordering should be best first.
- Works on anything the browser wants
- Test attributes can be combined

`skip-content` Attribute

- How to adapt for SMIL "dialects"
- Ignore unknown elements within sub-tree or ignore whole sub-tree

SMIL Test Attributes

Selecting Content Alternatives

- `system-bitrate` — bandwidth of Web connection
  - can switch media: video -> image -> text
- `type` — mime type of media object
- `system-required` — select if certain processing available

Selecting for User

- `system-language` — what language the user understands best
- `system-captions` — show content if user want closed captioning (subtitles)
  - usually single content of switch (on or off)
- `system-overdub-or-caption` — choice between audio or text

Adaptive Visual Complexity

- `system-screen-size`, `system-screen-depth`
- Switch on structure, not content

Extension Attributes for Particular Domain

- Won't be recognized by all browsers
- Potential examples — knowledge level, audience profile, length of time

Specifying alternative behavior

switch

- At most one of the children of a switch element is played.
- The first acceptable element is chosen, so ordering should be best first.

```
<switch>
 <audio system-bitrate="44000" src="hi-res.aiff" />
 <audio system-bitrate="16000" src="low-res.aiff" />
</switch>
```

- Test attributes can be combined.

New Content Control in 2.0

Prefetch

- Control, timing, and adaptation of pre-loading media before its presentation
- Helps whole presentations progress with fewer hitches

Custom Test Attributes

- Anyone can define adaptive test attributes for use in SMIL

Linking

Top Story:

**Growth of the
World-Wide Web**

Amsterdam

Los Angeles

File   Edit   View   Go   Bookmarks   Options   Directory   Window   Help

Dick Bulterman   CWI

HOME
SWEET
HOME

Document: Done.

XVII
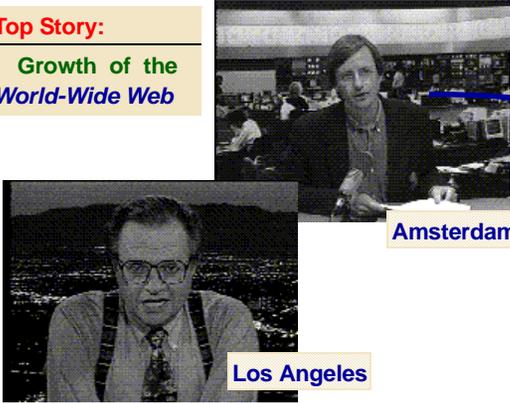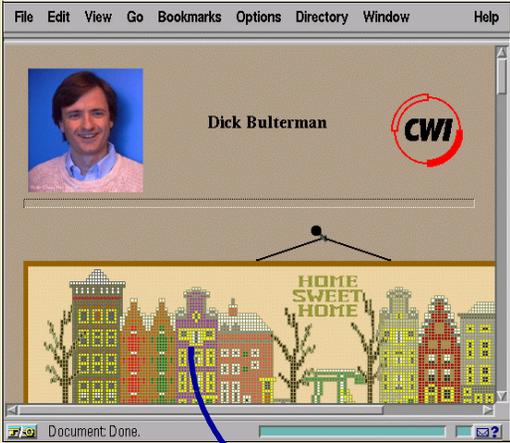
Link from element to presentation

The `<a>` element — similar to HTML `<a>` element.

- Source is unaffected and destination, `href`, is shown in `new` window.

```
<a show="new" href="archives-dcab.smi">
 <video src="zoomin.mpv" region="V-Main" />
</a>
```
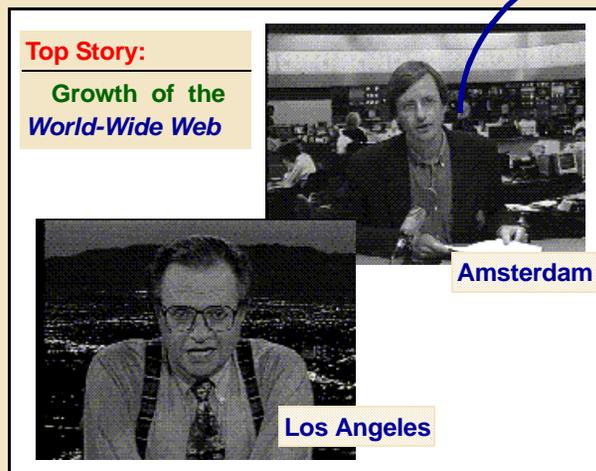

- Source may also `pause` while destination is shown,
- or destination may `replace` the source (default).

Link from element to element

Linking to SMIL fragments

- Destination element within another SMIL document uses **#** connector.

```
<a show="new" href="time-time.smil#XVII">
 <text src="archives-dcab.html" region="I-Main"
         dur="indefinite" />
</a>
```
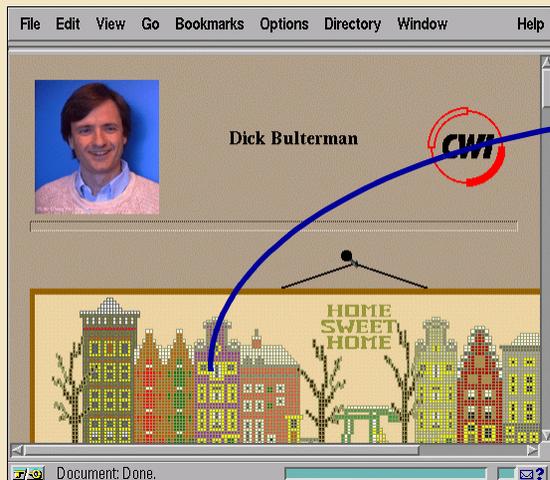

- Destination presentation starts as if the presentation had been fast-forwarded to the beginning of the element designated by the fragment.

Link from element to subpart of media object

The `area` element
The `area` element allows the specification of temporal and spatial subparts of a media object element.

- Spatial subparts use the **`coords`** attribute (similar to HTML image maps).

```
<video src="zoomin.mpv" region="V-Main" >
 <area id="mic" coords="40%, 70%, 55%, 100%" />
</video>
```


Defined w.r.t. media object, not w.r.t. region

`fit="slice"`

- Order of **`coords`** is `left-x`, `top-y`, `right-x`, `bottom-y`.

- Temporal subparts use the **`begin`** and **`end`** attributes.

```
<video src="zoomin.mpv" region="V-Main" >
 <area id="graph-ref" begin="4.3s" end="6.8s" />
</video>
```

Areas as source and destination of a link

- **href** needed if used as source, **id** needed if used as destination

```
Source document (image in SMIL, area and link defined in SMIL):
 <img src="home-sweet.gif" region="I-Main">
        <area href="time-time.smil#gable-3" show="new"
            coords="35%, 5%, 40%, 95%" />
 </img>
Destination document "time-time.smil" (image in SMIL):
 <img src="XVII.tiff" region="house-right" >
        <area id="gable-3" coords="30%, 0%, 70%, 100%" />
 </img>
```
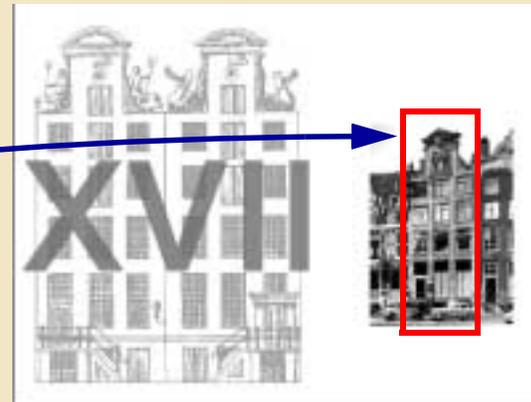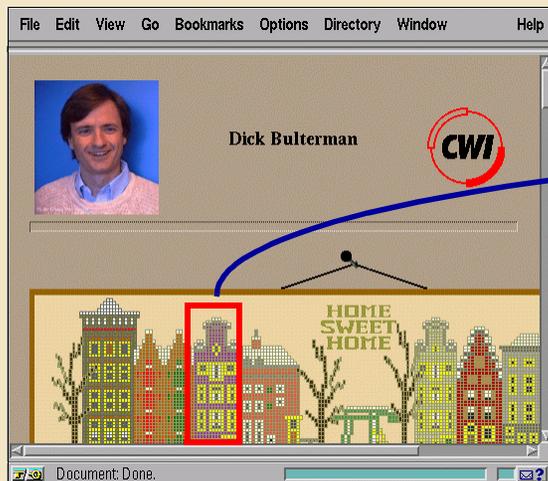
Semantic annotations

SMIL 3.4. `meta` element defines properties of a document

- The `name` attribute is the property and the `content` attribute gives the value.

  ```
  <meta name="title" content="Web News, 15th June 1998" />
  <meta name="base" content="http://www.cwi.nl/SMIL/webnews/" />
  ```

- The list of properties (values of `name` attribute) are open-ended.

Attributes on `par`, `seq` and media object elements
`abstract`, `author`, `copyright`, `title` (recommended)

Attributes on media object elements
`alt` (contains alternative text, recommended),
`longdesc` (supplement to `alt`, but longer and should include descriptions of areas)

Attributes on `region` elements
`title` (recommended)

High-Level Structure of Document

Partitioning in Sections

```
<smil>
 <head>
        <meta>
            ... information about the document ...
        </meta>
        <layout>
            ... layout definition ...
        </layout>
 </head>
 <body>

            ... objects and temporal relations ...
            ... including links and area objects ...

 </body>
</smil>
```

What's next?

SMIL 2.0 becomes a recommendation early in 2001
- Draws attention to the standard
- Stabilized to enable wide-spread implementation and adoption
- First players scheduled for release with recommendation

SMIL 2.0 becomes more implemented
- More browsers introduced
- More existing Web browsers add SMIL to languages shown
- SMIL browsers show more and more media
  - SVG?
  - All show XHTML?

SMIL 2.0 becomes more used

New Profiles Introduced from Outside W3C?

SMIL 2.5 and 3.0?

SMIL's Relationship with Other W3C Recommendations

SMIL Documents are XML Documents

- SMIL syntax is defined by an XML DTD

Private Extensions must use Namespaces

- `Skip-content` attribute allows content of non-SMIL elements to be played
- `System-required` attribute states the subtree requires the named implementation

SMIL Layout and CSS-2

- SMIL basic layout is consistent with the visual rendering module in CSS-2
  - it introduces the "fit" attribute
  - it is otherwise a subset.
- SMIL basic layout applies only to media object elements.
- SMIL media object elements refer to a region
  - CSS-2 "region" elements refer to the media object elements.

Summary

Media object element revisited

```
<video id="vid1" region="R_video"
    src="rtsp://www.w3.org/CoolStuff.rm"
    clip-begin="smpte=00:01:19:20"
    clip-end="smpte=00:01:38:40"
    begin="3s"
    dur="22s"
    end="21s"
    alt="Video of Joe chatting to Tim"
    longdesc="Joe and Tim are in a meeting room.  Joe is on
    the
        left and Tim is on the right"
    title="Joe greets Tim"
    system-bitrate="28800">
 <area id="joe" begin="0s" end="5s" coords="0%,0%,50%,50%"
        href="http://www.w3.org/" />
 <area id="tim" begin="5s" end="10s" coords="50%,50%,50%,50%"
        href="http://www.w3.org/Tim" />
</video>
```

Links

More info:

- http://www.cwi.nl/SMIL{/Tutorial}
- http://www.w3.org/TR/REC-smil
- http://www.w3.org/TR/smil20
- http://www.oratrix.com/
- http://www.real.com/