分散システム特論レポート

SMIL(Synchronized Multimedia Integration Language), Version1.0

情報工学専攻 修士 1 年 廣川研究室所属
西野　謙吾

平成 12 年 12 月 23 日

1 SMIL の概要

SMIL は、W3C(World Wide Web Conference) の同期マルチメディアワーキンググループによって開発され、1998 年 W3C によって勧告された。SMIL は、WWW 上で 動画、静止画、音声、文字など様々な形式のデータの再生を制御して同期させる言語である。どの位置に、どのタイミングで、どのくらいの時間表示するのかといったことが制御できる。また、インターネット上で文書やデータを交換したり配布したりするときの標準となる可能性の高いマークアップ言語である XML(eXtensible Markup Language) で記述されている。SMIL を使って、製作者は次のようなことができる。

- 表現の時間的な挙動の記述
  プレゼンテーションの動作を記述できる。何秒後にどの映像を何秒間放映するなどと時間軸に沿って自分の思い通りにタイミングを指定できる。例えば、「画像を 5 秒間表示した後に、音声が流れだす。」とか「オーディオを鳴らしながら、2.7 秒目から画像を表示する。」といった処理を記述できる。
- 表現のスクリーン上のレイアウト
  スクリーン上のプレゼンテーションのレイアウトを記述する。
- パイパーリンクとメディアオブジェクトの連結
  メディアオブジェクトと呼ばれる画像、動画、音声等のコンポーネントとハイパーリンクを結びつけることにより、例えば関連するホームページにリンクを張って、視聴者に詳しい情報を提供するといったことが出来る。つまり、クリッカブルマップを実現できる。

2 SMIL を構成する要素

SMIL は XML で記述されており、それぞれの役割を持つタグの組み合わせで表現される。その組み合わせ規則は DTD(Document Type Definition) で定義されている。DTD とは XML の構造を記述したものであり、要素名、要素間の可能な組み合わせ、各要素タイプで使用可能な属性などのメタ情報を記述するところである。

SMIL は DTD において、smil タグをルートタグとした木構造として表現される。smil タグは子供要素として文書のヘッダである head タグ、文書の本体である body タグを持つ。文書のヘッダにおいては、表現の時間的挙動と関係しない情報が記述される。文書の本体においては、逆に、時間的な挙動およびリンクの挙動に関連する情報を記述する。

2.1 文書のヘッダ

文書のヘッダは head タグ以下のタグ群によって表現さる。直接の子供要素として layout、meta、switch タグを持つ。

- **layout**
文書の本体の要素を文書の視覚的、あるいは聴覚的抽象レンダリング画面にどのように位置付けるかを決定する。文書が layout 要素を記述しないとき、body 要素の位置付けは実装依存となる。

- **meta**
文書の管理情報、つまりそのファイルのタイトル、編集者、ファイルの有効期日などを定義している。そして、必要に応じてアプリケーションにその情報を渡す。

- **switch**
制作者が，受け付け可能な要素をそこから 1 つだけ選択するべき代替的要素の組を指定することを可能にする。要素は次のようにして選択される。 実装のアプリケーションは switch 要素内で発生する順番に要素を評価し、switch 内部の他の要素すべてを排斥して，受け付け可能な最初の要素が選択される。そこで，制作者は，もっとも好ましいものから最も好ましくないものへと選択肢を並べるべきである。

2.2 文書の本体

文書の本体では、メディアオブジェクトと呼ばれるメディアをどのように同期させ表示するかを表現する。メディアオブジェクトとは、静止画、動画、音声、アニメーション、テキスト等のメディアデータを表し、メディアオブジェクトエレメントと呼ばれる要素によって、それらを SMIL プレゼンテーションへ取り込むことを可能にする。メディアオブジェクトエレメントとそれに対応するメディアの組み合わせは次の表のようになる。メディアを 7 つに分類しているが、実際どのようなファイルを扱えるかは実装依存である。

メディアオブジェクトエレメント	メディア
animation	アニメーション
audio	音声
img	静止画像
ref	animation,audio,img 以外の属性
text	プレーンテキスト
textstream	ストリーミングテキスト
video	動画

次に、上に述べたメディアオブジェクトの時間的同期をとる同期要素として、par、seq タグがある。

2.2.1 並列同期要素 — par

par タグは、子供要素としてメディアオブジェクトや同期要素自身をとり、それらを並行処理する。図 1 は、以下の par 記述による実際の動作を表したものである。まず、image1 を t=0 から、10 秒間表示する。それに平行して t=3 から audio1 を 5 秒間再生する。

```
<par>
<img src=”image1” begin=”0s” dur=”10s”/>
<audio src=”audio1” begin=”3s” dur=”5s”/>
</par>
```

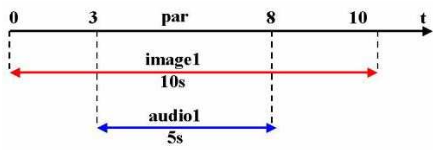

図 1: par タグによる同期

2.2.2 逐次同期要素 — seq

seq タグは、子供要素としてメディアオブジェクトや同期要素自身をとり、それらを逐次処理する。図 2 は、以下の seq 記述による実際の動作を表したものである。まず、image1 を 5 秒間表示し、それを表示し終わった段階で image2 を 3 秒間表示する。

```
<seq>
<img src=”image1” dur=”5s”/>
<img src=”image2” dur=”3s”/>
</seq>
```

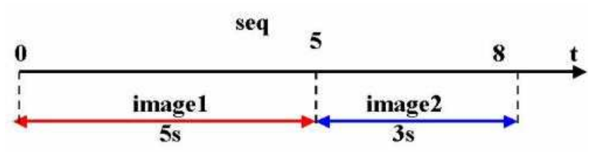

図 2: seq タグによる同期

3 SMIL の実装例 — Real Player

SMIL を実装するアプリケーションとして Real Player が挙げられる。SMIL 文書は、図 3 に示すように XML パーサーによって構文解析され、文書中の各要素をノードとした DOM ツリ-とよばれる木構造に変換される。そして、その DOM ツリ-にアプリケーションがアクセスし、プレゼンテーションに必要な情報を取り出す。この場合、XML パーサーの機能は Real Player 自身が持つ。

4 SMIL Boston

ここでは、SMIL1.0 と SMIL Boston(SMIL2.0 のドラフト) との相違点として顕著なものを述べる。

- モジュール化
  SMIL の機能を分類し、意味的に関係をもつ機能群からなる「モジュール」に整理したこと

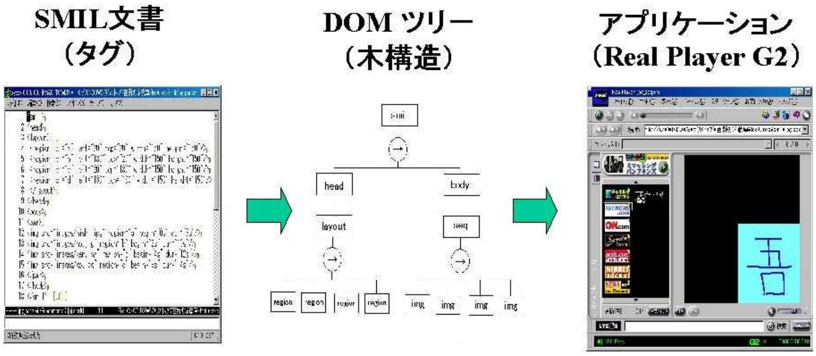

図 3: RealPlayer での SMIL 文書の取り扱い

である。これは単にグループ分けされたという以上のいみを持つ。モジュール化することで、他の XML アプリケーションへの統合を容易に行なうことが可能となるからである。

- イベントによる時間モデル
  イベントによる時間記述を行なうことで、ユーザがキーボードやマウスを操作することで発生するユーザイベントなどによってプレゼンテーションを制御し、よりインタラクティブなプレゼンテーションを作成することが出来る。

- アニメーション機能の提供
  ここでいうアニメーションとは、プレゼンテーション中では、画像、グラフィックス、テキストなどの「オブジェクト」に相当する XML 要素に対して与えられた座標、色、サイズ、角度などの属性の値を時間の経過に従って変化させることである。その結果が反映され、例えば、オブジェクトが画面上を移動することになるアニメーション効果が得られる。

5 まとめ

SMIL は、WWW 上で 動画、静止画、音声、文字など様々な形式のデータの再生を制御して同期させる言語で、XML で記述されている。どの位置に、どのタイミングで、どのくらいの時間表示するのかといったことが制御できる。SMIL の製作者は次のようなことができる。

- 表現の時間的な挙動の記述
- 表現のスクリーン上のレイアウト
- パイパーリンクとマルチメディアオブジェクトの連結

そして、Real Player での SMIL 文書の取り扱いについて述べた。また、SMIL2.0 のドラフトである SMIL Boston の仕様について SMIL1.0 と大幅に違うところを述べた。