Concurrent Markup for XML Documents

Patrick Durusau (Society of Biblical Literature)
  pdurusau@emory.edu

Matthew Brook O’Donnell (OpenText.org)
  matt@opentext.org

XML Europe
May 22, 2002
• **Background**: biblical texts  
  complex textual transmission  
  multi-level analyses

• **Previous Insights** *(Durusau & O’Donnell 2001)*:  
  > Markup is *metadata about PCDATA*  
  > Typically recorded *in-line* utilizing a *tree* notation  
  > Metadata can be recorded *on* each PCDATA item instead of *around* it (BUVH)
(1) Concurrent Markup Revisited: A Partial Typology for Overlapping Hierarchies

(2) Testing the BUVH Model
   Milton’s *Paradise Lost*

(3) ‘Down from the trees’: Future Investigations
Case 1: No-overlap

Some sample PCDATA
Overlapping Hierarchies

Case 13: No-overlap

```xml
<a>
  Some sample PCDATA
</a>
```
Overlapping Hierarchies

Case 2: Shared start/end point

Some sample PCDATA
Overlapping Hierarchies

Case 12: Shared end/start point

Some sample PCDATA
Overlapping Hierarchies

Case 3: ‘Classic’ overlap

Some sample PCDATA
Overlapping Hierarchies

Case 11: ‘Classic’ overlap

```
<text>
<a>
  <b>Some sample</b>
</a>
</text>
PCDATA
```

```
<text>
  <b>Some sample</b>
</text>
PCDATA
```

```
<text>
  <b>Some sample PCDATA</b>
</text>
```
Overlapping Hierarchies

Case 4: Shared end point

Some sample PCDATA
Overlapping Hierarchies

Case 10: Shared end point

Some sample PCDATA
Overlapping Hierarchies

Case 5: One element contains other

Some sample PCDATA
Overlapping Hierarchies

Case 9: One element contains other

Some sample PCDATA
Overlapping Hierarchies

Case 6: Shared start point

```html
<a>
  <b>Some sample PCDATA</b>
</a>
```
Overlapping Hierarchies

Case 8: Shared start point

Some sample PCDATA
Case 7: Shared start and end points

Some sample PCDATA
Overlapping Hierarchies

Cases 1, 5, 9 & 13 handled by traditional markup

Tree model imposes parent-child relationship here
Overlapping Hierarchies

‘Classic’ overlap requires alternative methods

Common solutions:

- Milestones
- Stand-off markup
- Non-SGML/XML approaches e.g. MECS/TexMECS (Sperberg-McQueen and Huitfeldt)

Limitations
Overlapping Hierarchies

Hierarchical assumptions of tree model

Some sample PCDATA

Some

PCDATA

PCDATA

PCDATA

PCDATA
Bottom-Up Virtual Hierarchies

Overview