

Support for Arbitrary Regions in XSL-FO

A proposal for extending XSL-FO semantics and processing model

Ana Cristina B. da Silva
Joao B. S. de Oliveira
-
PUCRS/FACIN
{benso,oliveira}
@inf.pucrs.br

Fernando T. M. Mano
Thiago B. Silva
Leonardo L. Meirelles
PUCRS/CPSE
{fernando,bueno,
meirelles}
@cpts.pucrs.br

Felipe R. Meneguzzi
Fabio Giannetti
-
Hewlett-Packard
{felipe.meneguzzi,
fabio.giannetti}
@hp.com

ABSTRACT

This paper proposes an extension of the XSL-FO standard which allows the specification of an unlimited number of arbitrarily shaped page regions. These extensions are built on top of XSL-FO 1.1 to enable flow content to be laid out into arbitrary shapes and allowing for page layouts currently available only to desktop publishing software. Such a proposal is expected to leverage XSL-FO towards usage as an enabling technology in the generation of content intended for personalized printing.

Categories and Subject Descriptors

I.7 [Computing Methodologies]: Document and Text Processing; I.7.2 [Document and Text Processing]: Document Preparation—*Desktop publishing, Format and notation, Markup languages, Photocomposition/typesetting*

Keywords

XML, XSL-FO, L^AT_EX, SVG, Digital Printing

1. INTRODUCTION

The XSL-FO standard describes XML documents separating content and layout information. This has led it to be considered an interesting alternative for publishing workflows. One of the main advantages of using XSL-FO in publishing lies in it being an open standard based on XML. Moreover, the processing of a document based on this standard can be logically broken down into multiple stages that can be distributed among specialized service providers.

From a formatting point of view, the XSL-FO format provides constructs for specifying page layouts in which content flows can be positioned automatically. Such capability is important as it eases the process of paginating complex content. Nevertheless the current version of the standard has limitations regarding the type of layout over which flow content can be placed. In particular, XSL-FO version 1.0 defines five possible regions within a page, only one of which can be used for the disposition of flow content. Such limitation is overcome in version 1.1 through

a construct called **fo:flow-map**, which allows multiple content flows to be mapped into specific regions within a page. Despite such improvements in version 1.1, the current page model allows only rectangular regions in which content is laid out. If XSL-FO is to be used as the base format for a Digital Printing workflow, more flexibility in the definition of content holding regions is required. Therefore, we propose to define regions using arbitrary shapes as a means to allow for more flexible content layout within XSL-FO.

This article describes an extension of XSL-FO 1.1 that allows the definition of an unlimited number of page regions, each one having arbitrary geometric shape. Such extension empowers the XSL-FO standard with document-typesetting capabilities with complex graphical compositions.

2. ARBITRARILY-SHAPED FORMATTING OBJECTS

The new page model for the proposed extensions allows the specification of any number of content-bearing page regions as well as the five regions specified in XSL-FO 1.1. These additional regions may possess an arbitrarily complex geometric specification, limited only by the language chosen for its description. Furthermore, our departure from the original XSL-FO page-region organization implies that overlapping of arbitrary portions of multiple adjoining regions is possible (Figure 2).

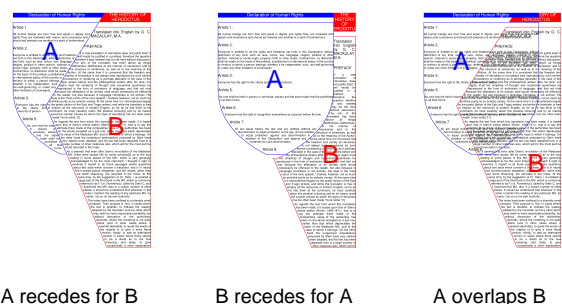


Figure 1: Simple arbitrary shapes.

In order to cope with the new possibilities, the proposed representation allows regions to have depth values so as to

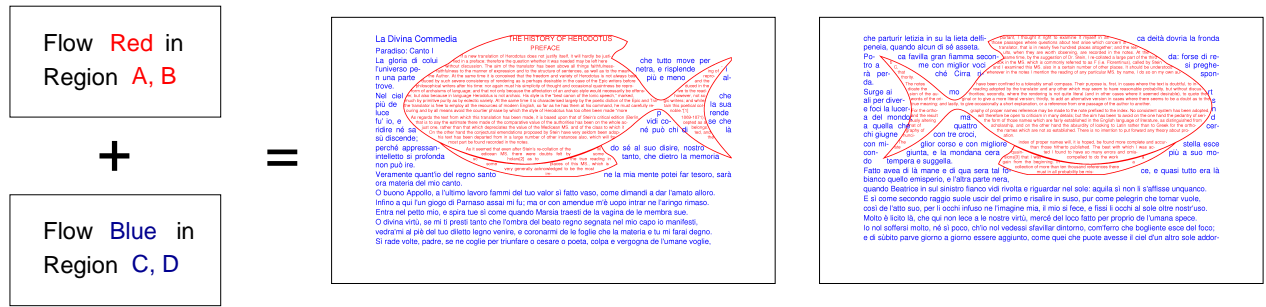


Figure 2: Possible page layout in extended XSL-FO.

allow region areas to be prioritized. A region can also have a specific behavior associated to it which is considered when overlapping with other regions occur, thus allowing a user to control the interaction among overlapping regions.

The description of arbitrary page regions in the extended XSL-FO specification is accomplished adopting two mechanisms similar to the currently supported pagination model. As previously seen there are two ways of mapping a flow inside a page region, a direct one using the region-name as referencing attribute and an indirect one using the flow mapping. The proposed approach accomplishes a similar effect regarding the association of shapes to regions enabling the direct embedding of a shape description within a new region-arbitrary element as well as a region mapping.

2.1 Direct and Indirect Mapping

In the direct mapping approach, an **fo:region-arbitrary** element, declared as a child of the **fo:simple-page-master** element is introduced. Unlike its XSL-FO 1.1 siblings, any number of different **fo:region-arbitrary** elements can be declared within a given **fo:simple-page-master** as long as they are uniquely named. The geometric outline of an arbitrary region is specified using an external vector format. Within our proposal the chosen format is SVG due to its XML nature and consequent easier adaptation of an XSL-FO parser to cope with it. For example, an SVG-specified arbitrary region named Region1 in Figure 3.

```
<fo:region-arbitrary region-name="Region1" ...>
  <svg ... > ...
</fo:region-arbitrary>
```

Figure 3: Arbitrary Region SVG specification.

The support for indirect mapping adds flexibility to the shape reuse across multiple pages and page sequences. Each shape will be defined inside an **fo:shape** element as part of the layout master set definition (Figure 4).

This mapping can be applied over legacy regions to support non-rectangular shapes in the "border" layout, already available in XSL-FO 1.1, or in combination with the **fo:region-arbitrary** to achieve a "free-form" layout.

2.2 Free-from layout versus Border Layout

In the original XSL-FO model only four regions can be

```
<fo:shape-map shape-map-name="document-regions">
  <fo:shape-assignment>
    <fo:shape-source shape-name-reference="poly1"/>
    <fo:region-target region-name-reference="left"/>
  </fo:shape-assignment>
  <fo:shape-assignment>
    <fo:shape-source shape-name-reference="poly2"/>
    <fo:region-target region-name-reference="body"/>
  </fo:shape-assignment>
</fo:shape-map>
```

Figure 4: Shape Map specification.

described (roughly corresponding to the header, footer and both margins of a page) and all that is not included in one of these regions comprises the body of the page. A similar effect can be immediately achieved through the use of four shapes. Figure 1 shows three page models that could be obtained through the use of arbitrary shapes, demonstrating how layering options can achieve multiple layouts using the same set of shapes.

3. CONCLUDING REMARKS

This paper describes a proposal for the extension of the XSL-FO standard to be able to layout content into multiple arbitrarily shaped page regions. The extension is intended to leverage the XSL standard for usage in the typesetting of complex document layouts usually found only in expensive desktop publishing applications. An implementation of such standard enables the generation of personalized documents in layouts only encountered in one-of-a-kind graphic presentations.

The proposed extensions reuse many concepts from the SVG standard, as well as concepts currently being proposed for its upcoming version. Such reuse is not incidental, as ensuing implementations of the arbitrary layout algorithms and shape encoding schemes will be able to share many common components. At some level this sharing of components is currently being done in the implementation of FOP, as well as in our extended prototype.

The next step in the development of the arbitrary shapes extensions of XSL-FO is its analysis by the community. Such an analysis should uncover potential flaws in the specification or point out improvements, and potentially lead to its incorporation into a future version of the XSL-FO standard.