

Half-steps toward LMNL

Wendell Piez
Mulberry Technologies

Abstract

Overlap in markup occurs where some markup structures do not nest, such as where the sentence and phrase boundaries of a poem and the metrical line structure describe different hierarchies. LMNL (Layered Markup and Annotation Language) is a model for representing textual data, designed to recognize and account for layer separation and markup overlap. LMNL is specified as a data model, not as a syntax — but without a syntax and an API it's very difficult to experiment with the model. I demonstrate a subset of LMNL using an XML syntax and some severe restrictions on LMNL (thus “half-LMNL”). Using an attribute structure for milestone marking and correspondence allows the input to be processed as XML and parsed into a tree. If this tree is flattened to reduce all XML markup to empty XML elements demarcating fragments of text, it can be transformed again to produce a modified “reified LMNL” model including overlapping ranges. This XML representation of a LMNL model takes the form, in effect, of standoff markup, although the technique preserves tag ordering (as full LMNL would not).

Half-steps toward LMNL

Table of Contents

Background and update.....	1
Building XML models of LMNL.....	2
Demonstrations.....	5
Enjambment in Milton.....	6
Annotations in a scholarly edition.....	8
Current conclusions.....	10
Appendixes: code samples.....	12
A fragment of <i>Paradise Lost</i>	12
A half-LMNL model of the text.....	12
Enjambment Analysis XSLT.....	15
Scholarly annotations.....	16
TEI source, with CLIX.....	16
Processing results: the extracted notes.....	20
Notes extraction XSLT.....	21
Footnotes.....	25
Bibliography.....	27
The Author.....	27

Half-steps toward LMNL

Wendell Piez

§ Background and update

LMNL is an approach to the modeling of textual data first presented at this conference in 2002. “Layered Markup and Annotation Language” provides for representing textual data in terms of “ranges” (which may overlap) and “layers” over those ranges. LMNL is thus designed explicitly to recognize and account for overlapping phenomena and (eventually) concurrent hierarchies, such as can be observed in extant literary and linguistic artifacts (wherever else the model may be applicable). Identified with an abstract data model, not a syntax, LMNL is only, as of this date, specified in draft. Its authors consider it to be a project ongoing. [LMNL Extreme 2002], [LMNL 2002].¹

Paradoxically, although LMNL avoids specifying itself as a markup syntax, since it's fundamentally a model of how text markup might be applied and construed (over and above its suitability as a way of modeling the results of operations over text), it proves very difficult to work with absent any method of text markup altogether. LMNL cannot get completely away from its conception as a “markup and annotation” language. This results in a bootstrapping problem: the design problem of the LMNL object model must take the form, at some point (at least if it is to be computationally useful), of specifying an API of some kind, but there is no way of fully clarifying the requirements to be addressed by an API without understanding functional requirements of actual applications, no way of listing functional requirements without use cases and an understanding of what problems are best addressed through this model of text (intuitions not being concrete enough), no way of getting close to those problems without working with actual texts — and no actual texts without a syntax — and a parser implementing that syntax — to complement the data object model that is not yet fully conceptualized.

This conundrum is reflected in the principal design problems now facing LMNL, any of which might be decided absent considerations of a syntax, of use cases and of the operations we want to perform, but which might not best be. In particular, LMNL faces two distinct issues, which yet may not be so distinct:

- Are “tags” phenomena that have existence in their own right? Or alternatively, are range-starts and -ends ordered with respect to one another? LMNL as drafted supposes they are not (their order is a function only of their offsets, so two ranges that begin at the same offset effectively begin together, irrespective of the order of tags used to represent them) but then runs into difficulties representing certain structures easy to represent with marked-up text, such as ordered point phenomena (think of an HTML anchor surrounding an inline reference to an image, as in 1). In part this may be because LMNL is conceived in a way to make it possible to implement as a set of operations over a text layer, or over ranges, and tag ordering questions are harder to deal with when starting with no tags at all.
- What exactly are layers and how are they best represented (if at all) in a LMNL syntax? (In particular, it could be useful to specify layers to contain a LMNL equivalent of XML mixed content, or layers that stipulate content models. But these ideas go beyond layers as presently specified.)

It is not readily apparent that these issues are not interrelated. If point phenomena or events represented by tags (as opposed to ranges identifiable only by the character offsets where they begin and end) have order in the data model, what effect does this have on the possible relations between ranges, and hence on layering and on the way operations are best defined over ranges and layers? (2 illustrates this question.) Are layers a solution to the relative range-ordering problem, or merely a complication? LMNL's basic conception posits layering as a basic feature of the language; but it has not yet been seen exactly what layers are best conceived and what they will be useful for. Again, a hermeneutic circle prevents us from simply announcing a solution: we do not yet know because we haven't tried it, and we haven't tried it yet because we don't know.

This paradox cannot stop progress for long, however. Nor has it. By exploring the edges of “LMNL-space” it can become clearer what the best solutions are to outstanding questions. While this work has not been well coordinated or publicized (LMNL's loose organization and lack of formal support make such coordination impractical), it has nonetheless continued (if sometimes below the surface) since the core concept of LMNL was introduced at this conference two years ago.² Even stop-and-go progress gets us some distance eventually; the staying power of these ideas is also suggested by the parallel efforts of

Figure 1

```
[a [href]graphic.jpg{1}][img [src]graphic.jpg{}]{a}
```

```
[a [href]graphic.jpg{}][img [src]graphic.jpg{}]
```

The first bit of code appearing here is an HTML-style hyperlink (expressed in LMNL bracket-brace syntax): an anchor “wraps” an image. But since the `img` range is empty (has no string length), both it and the (likewise empty) `a` range occur in the same position (at the same offset) with respect to the document’s string content; thus, in LMNL, the first example is indistinguishable from the second. While one can imagine learning to live with this, and while the design has considerable virtues in some respects, markup languages (if not all kinds of text processing applications) have traditionally taken tag ordering to be significant, even critically so.

Figure 2

```
[seg]And chiefly thou,{seg}[seg] O Spirit,{seg}
```

```
[seg]And chiefly thou,[seg]{seg} O Spirit,{seg}
```

If LMNL does not distinguish between these cases, it will have no way of saying the second one is “wrong”. This may or may not prove to be a problem: at least when using inline markup, it seems counter-intuitive to argue that the second case shows no overlap.

other projects formally unrelated to LMNL (as evidenced at this conference in papers by Durusau, DeRose and Witt, at least). Moreover, LMNL itself appears to be “sticky”, at least judging by not only the continued level of interest and confidence expressed by members of the markup community, but also the surprising and welcome demonstration of LMNL this summer by an independent researcher³ at this year’s joint conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities (Gothenburg, Sweden). — a very encouraging development.

§ Building XML models of LMNL

Progress is occurring because if defining LMNL as a data model faces us with a development Catch-22, it simultaneously offers a way out. Jeni Tennison demonstrated a year ago that LMNL can be usefully, if awkwardly, represented in XML, by proposing a **reified LMNL** as a solution to certain problems in rendering LMNL from an instance syntax [LMNL 2002]. The XML format serves as an intermediate format between LMNL expressed syntactically, and LMNL proper (an abstract information set accessible via an API). In the paper just cited, Alexander Czmiel has paralleled this effort in using standard XML (JDOM) processing over an XML “reification” of LMNL (XfOS) that he designed for his own system [Czmiel 2004]. Simultaneously, Steve DeRose, with others involved in the OSIS project, has worked out a neat method for representing ranges in XML syntax [DeRose 2004]. The present paper essentially builds directly on all this work by processing XML marked with CLIX (DeRose’s proposal) into yet another reified form, and processing this, again (following Tennison’s lead), with XML tools (in this case XSLT).

This method has variously been referred to as “Trojan milestones” (not, here, simply a classical reference, but in honor of the developer who first suggested it) or as “CLIX”. CLIX follows the convention of representing ranges with milestones (an approach sanctioned by TEI, for example [TEI P4]); but it manages to evade many of the modeling questions that are raised by this approach by loosening up the rules regarding which element type(s) might be expected to serve as such milestones. In CLIX notation, it is the presence of a dedicated attribute that indicates that a particular XML element serves here a CLIX marker (and implicitly that we can elsewhere find a corresponding CLIX marker) rather than an ordinary instance of the element:

```
<line n="Bk1.1" l:sID="L1"/><s><seg>Of Man's first disobedience,</seg>
<seg>and the fruit<line n="Bk1.2" l:eID="L1"/>...</seg>...</s>
```

(The opening line of Milton’s *Paradise Lost*⁴, with ranges tagged with CLIX milestones.) The virtue of this approach is in how rough-and-ready it is: it has the advantage of allowing the notation of overlap to

occur in a very ad-hoc way, so the marking up of instances can proceed without a wholesale redevelopment of document types.⁵ Deploying CLIX gets us past having to parse any provisional LMNL syntax, into a realm where the ranges themselves can be recognized, related to one another, and processed.

We accept two limitations on the generic LMNL model in order to work like this. First and most importantly, we leave aside the LMNL construct of “layers”, which is to say ranges of ranges. Although layers provide what is potentially one of the most powerful features of the language, the details of how they should work are still being explored. There is so much to be gotten from “flat LMNL” that we can proceed safely without layers, leaving them aside until they are needed. Secondly, for the sake of simplicity we will restrict names to those that play well in XML. 3 shows a brief sample instance.

4 diagrams how this document is processed; the XML serialization of the result appears in 5. The sequence of stylesheets operates on the CLIX-enhanced XML as follows:

- A “flattener” reduces all markup, whether regularly nested XML, or milestones, to a CLIX (milestone) representation. Content of a CLIX marker is taken as an anonymous annotation on the range marked by that marker. The transform preserves the original “containment” structure of the original XML only implicitly: what started as XML elements appear only ranges among other ranges. This transform also provides a place to layer document-type-specific processing such as templates to recognize TEI milestones, etc.
- An “extractor” renders the sequences of text nodes with markers created by the first pass into a model representing “half-LMNL”, which is defined as a text layer marked with range markers, which can have annotations (which are similarly text layers with markers).
- The extracted model is enhanced in a subsequent stage, deriving information, such as range lengths, useful for later processing. Strictly speaking this is not adding information that was not there (since if we are willing to pay a performance cost we could always derive these later); it just makes properties more accessible that are important to the data model.
- Canonical Reified LMNL can be straightforwardly derived from half-LMNL, if wanted. But half-LMNL presents enough information about ranges to be useful for range-based processing, while preserving information about tag ordering should we need it.

This input can be flattened to make a document that reduces all XML markup to empty XML elements, demarcating fragments of text. Notice this step entails deliberately throwing information away: all range-start and -end markers will be represented by empty elements in the XML, with no nesting. The result of this transformation makes no distinction between markup that specified the original source tree, and markup that took the form only of CLIX markers set therein.

We can then process a liminally marked-up document with another generic transform to produce a modified reified LMNL model including overlapping ranges.⁶ Again, this model takes the form of an XML document, and can hence be processed with standard tools to produce an in-memory tree addressable with XPath. Accordingly, these three transforms in series⁷ provide an effective end-run around a true parse of a “native” LMNL syntax, using CLIX notation in XML as a means to the end.

Figure 3

```

<?xml version="1.0" encoding="UTF-8"?>
<snip xmlns:clx="http://lmnl.org/namespace/clix">
<p><q clx:sID="q1"/>The thing is<q clx:eID="q1"/>, argue
researchers, <q clx:sID="q2"/>structures in text can
overlap.</p>

<p><comment clx:sID="1"/>There can be quotes that align,
or don't, with the arrangement of the text into
paragraphs. There can be interfilings of notes and
marginalia. There can be metaleptic narrative structures,
stories folding into stories or presenting topological
puzzles in their narrative makeup (devices like this are
as old as epic poetry). There can be reference structures
or indexes laid over the top, such as the canonical
referencing schemes to scripture or to the works of Plato
and Aristotle. There can be glosses, footnotes, parallel
commentaries.<comment clx:eID="1">XML might do any of
these beautifully, but generally not more than one at a
time; and in XML one has to consider the design up front
and must design to particular functional requirements or a
particular <called>vision</called> of a text; were markup
to allow overlap it would be easier to play around
with.</comment><q clx:eID="q2"/></p>
</snip>

```

Even a simple example like this one can demonstrate the use of in-line markers to denote ranges that fall across XML element boundaries. Here, too, one of the CLIX markers is “opened up”: its content can be attributed to an anonymous metarange (annotation) in the LMNL model.

Figure 4



This diagram depicts the chain of transformations that turn a CLIX-annotated XML instance into a reified “half-LMNL” instance. The process is one of (a) explicating tags as elements in a flattened model, and then (b) rearranging and enhancing the information presented by the model so it is amenable for downstream processing (text nodes are merged and range offsets and lengths are calculated). We could

easily take this to full reified LMNL with another simple transform; because we want to preserve the possibility of asserting significance in tag-ordering, however, these demonstrations work over the half-LMNL.

Figure 5

```
<?xml version="1.0" encoding="utf-8"?>
<l-:document xmlns:l-="http://lmnl.org/namespace/halfLMNL/reified">
  <l-:content>
    The thing is, argue researchers, structures in text can overlap.

    There can be quotes that align, or don't, with the arrangement of
    the text into paragraphs. There can be interfilings of notes and
    marginalia. There can be metaleptic narrative structures, stories
    folding into stories or presenting topological puzzles in their
    narrative makeup (devices like this are as old as epic poetry).
    There can be reference structures or indexes laid over the top,
    such as the canonical referencing schemes to scripture or to the
    works of Plato and Aristotle. There can be glosses, footnotes,
    parallel commentaries.
  </l-:content>
  <l-:range-start name="snip" id="snip-1" where="0" length="608"/>
  <l-:range-start name="p" id="p-1" where="1" length="64"/>
  <l-:range-start name="q" id="q1" where="1" length="12"/>
  <l-:range-end name="q" id="q1" where="13"/>
  <l-:range-start name="q" id="q2" where="34" length="573"/>
  <l-:range-end name="p" id="p-1" where="65"/>
  <l-:range-start name="p" id="p-2" where="68" length="539"/>
  <l-:range-start name="comment" id="1" where="68" length="539"/>
  <l-:range-end name="comment" id="1" where="607">
    <l-:annotation>
      <l-:content>XML might do any of these beautifully, but
      generally not more than one at a time; and in XML one has
      to consider the design up front and must design to
      particular functional requirements or a particular vision
      of a text; were markup to allow overlap it would be easier
      to play around with.</l-:content>
      <l-:range-start name="called" id="called-1" where="203"
      length="6"/>
      <l-:range-end name="called" id="called-1" where="209"/>
    </l-:annotation>
  </l-:range-end>
  <l-:range-end name="q" id="q2" where="607"/>
  <l-:range-end name="p" id="p-2" where="607"/>
  <l-:range-end name="snip" id="snip-1" where="608"/>
</l-:document>
```

“Half-LMNL” is essentially a plain-text rendition of the source document with standoff markup to identify ranges defined over the text. (In this code listing, cosmetic line breaks have been added.) Where annotations occur, their content (which in LMNL can include ranges of their own) can be handled the same way as the base text.

It is worth noting that the final XML representation we choose to work with, “half-LMNL”, takes the form, effectively, of a representation of the input using standoff markup. The resemblance to standoff markup is coincidental, but significant and consequential: here we are benefiting from the particular strength of standoff markup, its straightforward representation of arbitrary overlapping structures without ambiguity, without having to contend with its primary challenge, namely the difficulty of its maintenance (since this model can be generated on the fly from conventional inline markup as needed).

Then we can process the reified LMNL result to report fragments of our data, construct operations over it and so forth.

§ Demonstrations

Two running examples demonstrate what can be done with this model using available technologies such as XSLT. The first implements a process that, while fairly rudimentary, takes advantage of LMNL’s capabilities in representing multiple concurrent hierarchies in a single instance. The second, somewhat more ambitiously, works over a text with several concurrent structures, one of which (arbitrarily annotated segments) includes ranges that overlap with each other. This routine demonstrates not only the extraction of arbitrary ranges, but how they can be represented in a LMNL syntax (or again, in CLIX) for further processing.

Enjambment in Milton

An XSLT stylesheet can be applied to the half-LMNL version of the opening lines of *Paradise Lost* to analyze where in the passage line ends correspond with phrase boundaries in the grammar of the passage. This is of interest because the placement of **enjambment** and its converse, **end-stopping**, is a significant feature in Milton's prosody. End-stopping occurs in verse when the boundaries of verse lines correspond with the boundaries of grammatical units (sentences or phrases, indicated with pauses when speaking). Enjambment, its opposite, occurs when lines "run on" from one to the next without pause. This feature of verse, particularly in Milton, is interesting enough for us to postulate the usefulness of an automated routine to identify where it occurs, in a text marked up with both verse and grammatical hierarchies.

Figure 6

```
<?xml version="1.0"?>
<quote xmlns:l="http://lxml.org/namespace/clix">
<line n="Bk1.1" l:sID="L1"/><s><seg>Of Man's first disobedience,</seg>
<seg>and the fruit<line l:eID="L1"/>
<line n="Bk1.2" l:sID="L2"/>Of that forbidden tree whose mortal taste
<line l:eID="L2"/>
<line n="Bk1.3" l:sID="L3"/>Brought death into the World,</seg> <seg>
and all our woe,</seg><line l:eID="L3"/>
<line n="Bk1.4" l:sID="L4"/><seg>With loss of Eden,</seg> <seg>till
one greater Man <line l:eID="L4"/>
<line n="Bk1.5" l:sID="L5"/>Restore us,</seg> <seg>and regain the
blissful seat,</seg><line l:eID="L5"/>
<line n="Bk1.6" l:sID="L6"/><seg>Sing,</seg> <seg>Heavenly Muse,</seg>
<seg>that,</seg> <seg>on the secret top <line l:eID="L6"/>
<line n="Bk1.7" l:sID="L7"/>Of Oreb,</seg> <seg>or of Sinai,</seg>
<seg>didst inspire <line l:eID="L7"/>
<line n="Bk1.8" l:sID="L8"/>That Shepherd who first taught the chosen
seed <line l:eID="L8"/>
<line n="Bk1.9" l:sID="L9"/>In the beginning how the heavens and earth
<line l:eID="L9"/>
<line n="Bk1.10" l:sID="L10"/>Rose out of Chaos:</seg> <seg>or,</seg>
<seg>if Sion hill <line l:eID="L10"/>
<line n="Bk1.11" l:sID="L11"/>Delight thee more,</seg> <seg>and
Siloa's brook that flowed <line l:eID="L11"/>
<line n="Bk1.12" l:sID="L12"/>Fast by the oracle of God,</seg> <seg>I
thence <line l:eID="L12"/>
<line n="Bk1.13" l:sID="L13"/>Invoke thy aid to my adventurous
song,</seg><line l:eID="L13"/>
<line n="Bk1.14" l:sID="L14"/><seg>That with no middle flight intends
to soar <line l:eID="L14"/>
<line n="Bk1.15" l:sID="L15"/>Above th' Aonian mount,</seg> <seg>while
it pursues <line l:eID="L15"/>
<line n="Bk1.16" l:sID="L16"/>Things unattempted yet in prose or
rhyme.</seg></s><line l:eID="L16"/>
<line n="Bk1.17" l:sID="L17"/><s><seg>And chiefly thou,</seg> <seg>O
Spirit,</seg> <seg>that dost prefer <line l:eID="L17"/>
<line n="Bk1.18" l:sID="L18"/>Before all temples th' upright heart and
pure,</seg><line l:eID="L18"/>
<line n="Bk1.19" l:sID="L19"/><seg>Instruct me,</seg> <seg>for Thou
know'st;</seg> <seg>Thou from the first <line l:eID="L19"/>
<line n="Bk1.20" l:sID="L20"/>Wast present,</seg> <seg>and,</seg>
<seg>with mighty wings outspread, </seg><line l:eID="L20"/>
<line n="Bk1.21" l:sID="L21"/><seg>Dove-like sat'st brooding on the
vast Abyss, </seg><line l:eID="L21"/>
<line n="Bk1.22" l:sID="L22"/><seg>And mad'st it pregnant:</seg> <seg>
what in me is dark <line l:eID="L22"/>
<line n="Bk1.23" l:sID="L23"/>Illumine,</seg> <seg>what is low raise
and support;</seg><line l:eID="L23"/>
<line n="Bk1.24" l:sID="L24"/><seg>That,</seg> <seg>to the height of
this great argument,</seg><line l:eID="L24"/>
<line n="Bk1.25" l:sID="L25"/><seg>I may assert Eternal
Providence,</seg><line l:eID="L25"/>
<line n="Bk1.26" l:sID="L26"/><seg>And justify the ways of God to
men.</seg></s></quote>
</line l:eID="L26"/>
```

This XML version of the opening lines of *Paradise Lost* provides the source code for the first running example. The XML structure indicates sentence and phrase boundaries (marked with `s` and `seg` elements), while leaving lines indicated only by milestones. (We note in passing that this is not the most usual way of marking up this text, whose typographic requirements suggest it be broken out first into

lines; but it is easy to imagine a scenario where phrase boundaries are of primary importance and only milestones, in the way of HTML **br** elements, mark where lines break.)

While here, both line starts and ends are indicated, this is only a convenience (although perhaps a consequential one). With modest extra effort, specialized milestone-marking protocols can be supported.

Having processed the CLIX-marked source code of a passage in Milton (see 6) into a half-LMNL representation, we can run an enjambment-analysis stylesheet on this result, to produce a plain-text version that shows where enjambment appears (see 7).

Figure 7

```

001  [~  Of Man's first disobedience, and the fruit
002  ~~  Of that forbidden tree whose mortal taste
003  ~]  Brought death into the World, and all our woe,
004  [~  With loss of Eden, till one greater Man
005  ~]  Restore us, and regain the blissful seat,
006  [~  Sing, Heavenly Muse, that, on the secret top
007  ~~  Of Oreb, or of Sinai, didst inspire
008  ~~  That Shepherd who first taught the chosen seed
009  ~~  In the beginning how the heavens and earth
010  ~~  Rose out of Chaos: or, if Sion hill
011  ~~  Delight thee more, and Siloa's brook that flowed
012  ~~  Fast by the oracle of God, I thence
013  ~]  Invoke thy aid to my adventurous song,
014  [~  That with no middle flight intends to soar
015  ~~  Above th' Aonian mount, while it pursues
016  ~]  Things unattempted yet in prose or rhyme.
017  [~  And chiefly thou, O Spirit, that dost prefer
018  ~]  Before all temples th' upright heart and pure,
019  [~  Instruct me, for Thou know'st; Thou from the first
020  ~]  Wast present, and, with mighty wings outspread,
021  []  Dove-like sat'st brooding on the vast Abyss,
022  [~  And mad'st it pregnant: what in me is dark
023  ~]  Illumine, what is low raise and support;
024  []  That, to the height of this great argument,
025  []  I may assert Eternal Providence,
026  []  And justify the ways of God to men.
```

Here, end-stopping is indicated by brackets in the margin, whereas enjambment is indicated by tilde signs. For example, ~] indicates that the line is enjambed at the start, but end-stopped at the end. Even these 26 lines demonstrate vividly how Milton uses enjambment, or conversely end-stopping, to rhetorical effect.⁸

The stylesheet that performs this processing, an XSLT 1.0 stylesheet, works straightforwardly: it simply iterates over the starts of **line** ranges, writing out each line in turn, but checking as it does so for correspondences with the starts and ends of **seg** ranges. The stylesheet is not long; in 50 lines including cosmetic blank lines, it uses three key declarations and three templates. (See “**Enjambment Analysis XSLT**”.)

Now admittedly, given sufficient dedication and persistence, this particular output could be obtained using an XSLT transform operating directly over the raw XML, as marked with milestones to indicate the **line** ranges. Yet to conduct this analysis by examining ranges directly is attractive not just for its notional purism: it is also more robust in the face of greater complexity and variation in the source, such as if inline emphasis were also present in some of the lines. The LMNL-based solution (here it is half-LMNL but since this particular stylesheet does not make use of tag-ordering, it could equally well be run over LMNL in its full form) actually defines the problem of observing enjambment in terms of the **co-termination** of ranges that are allowed to overlap: it can recognize the line as an actual line.

Consequently the entire methodology is more secure and straightforward, as well as more general, than would be a straight hack in XSLT. (One indication of this is that the same set of stylesheets would work perfectly well if our original input markup had had the sentence/phrase structures, not the lines, marked by milestones.) Accordingly this demonstration also points the way to operations that could be conducted over a text originating in a syntax more straightforwardly representative of LMNL's native structures, such as the proposed “dragon-tooth” (brackets-and-braces) syntax.

Annotations in a scholarly edition

Enjambment has the virtue for these demonstrations of being a phenomenon that is defined in terms of overlap; yet it is not of very wide interest. Half-LMNL also allows more general-purpose kinds of processing over ranges that may happen to overlap. A common requirement for overlap arises when we provide texts with glosses, marginalia or any kind of annotation, when we wish to associate the annotations not simply with points in the text, but with spans or ranges of text.⁹ In this case, where a **note** appears in this TEI source text, an **sID** or **eID** attribute, in a namespace reserved for CLIX, is taken to identify a LMNL range marker; its content is deemed to be an anonymous metarange (or **annotation** in LMNL parlance) attached to this range.

This strategy allows for the annotation of arbitrary ranges of text in the source, irrespective of element boundaries or overlaps between ranges so annotated (or any other ranges). Incidentally, because we now have the element content of our LMNL marker to work within, we can leverage XML and even more CLIX markup to provide any annotation with LMNL ranges of its own. Since LMNL (unlike XML) allows its annotations (XML's analog being attributes) to have structure of their own, there is no problem here.

The demonstration text is a rendition of an historical document translated and edited by Prof JRR Tolkien, *The Red Book of Westmarch*, which relates an account of ancient history, particularly as it relates to early efforts in the development of technology standards. Of particular interest here is that the English-language rendition of this text provided by Prof Tolkien¹⁰ contains dialogue passages that also overlap structural boundaries between paragraphs.

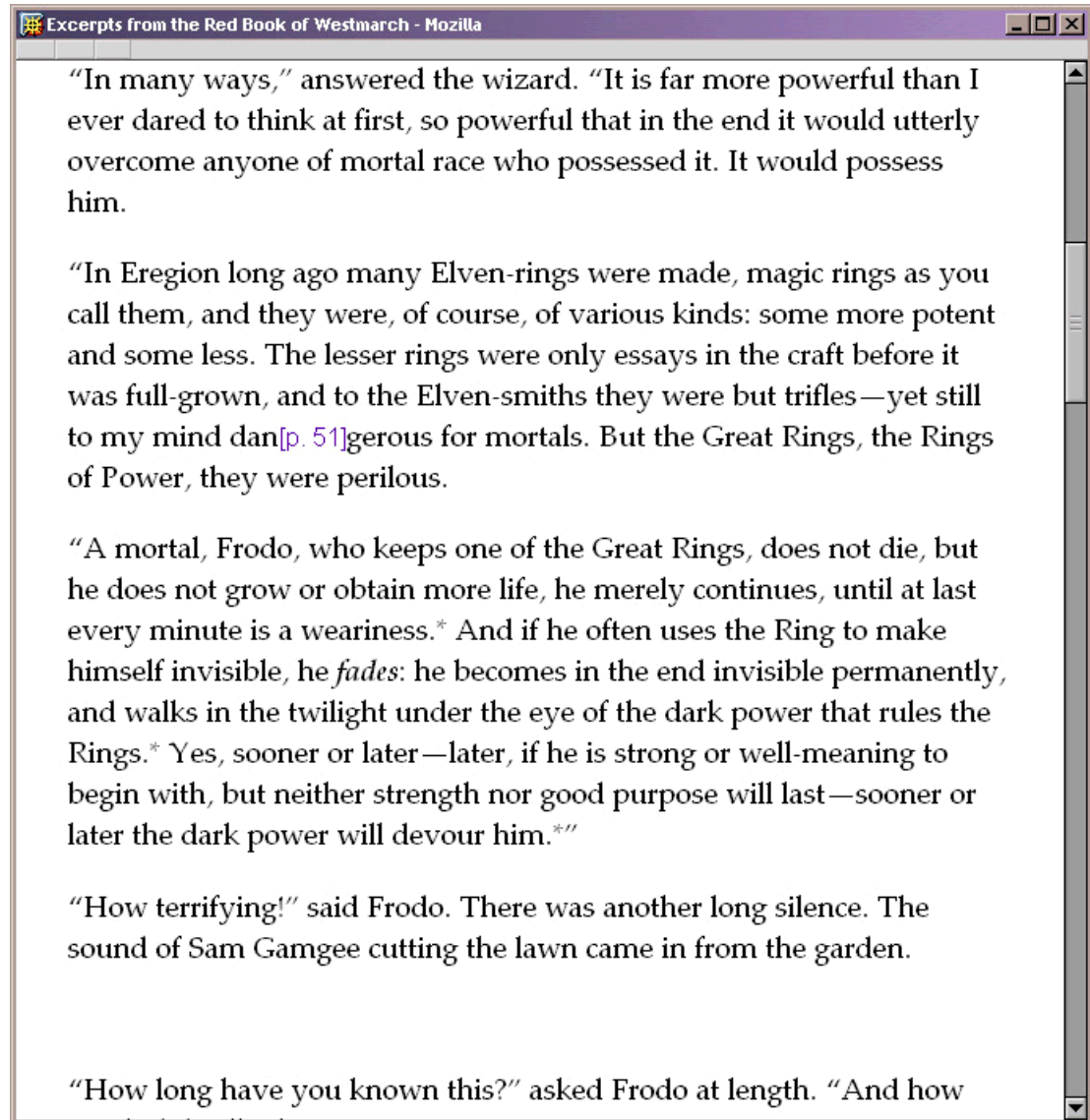
There are at least four concurrent hierarchies here:

- page divisions (of this one single edition);
- expository divisions such as chapters, chapter sections (marked in the original only by white space), paragraphs and verse segments;
- narrative and dialogical divisions (since lines of dialogue can spill across paragraphs and start and stop mid-paragraph, to say nothing of pages; and
- spans of arbitrary interest, which may be annotated.

This demonstration is interested primarily in the last of these, though the techniques it demonstrates could be applied to take account of any of them.

A reading rendition of the TEI document¹¹ appears in 8, with an excerpt from the result of processing a “notes extractor” stylesheet.

Figure 8



```
<p><note 1:sID="note1"/>A mortal, Frodo, who keeps one of the Great
Rings, does not die, but he does not grow or obtain more life, <note
1:sID="note7"/>he merely continues, until at last every minute is a
weariness.<note 1:eID="note1">The ring has the power of suspending
the rules of mortality and prolonging life. That is, it <emph>
preserves</emph> (though not in the same way as the Elf-rings whose
science it borrows). In this way, like any technology, it promises
control over time and space.</note> <note 1:sID="note2"/>And if he
often uses the Ring to make himself invisible, he <emph>fades</emph>:
he becomes in the end invisible permanently, and walks in the
twilight under the eye of the dark power that rules the Rings.<note
1:eID="note2">Another power the Ring has is to make its wearer
invisible [...].</note> Yes, sooner or later—later, if he is strong
or well-meaning to begin with, but neither strength nor good purpose
will last—sooner or later the dark power will devour him.<note
1:eID="note7">From the very first, accounts of what happens with the
Ring are completely negative: the temptation it represents is
described more than it is demonstrated. The Ring is beautiful as an
aesthetic object, although simple; but its treacherous nature is
spelled out early.</note><q 1:eID="q4" who="Gandalf"/></p>
```

An excerpt from the CLIX-enhanced TEI source of a passage from the Red Book: appearing here is a reading version in a browser (a client-side transformation of the TEI source) with an excerpt of the code behind part of it. In the code listing, a considerable part of one of the annotations is elided for clarity; also emphasis has been added: CLIX markers appear in bold, while a single **note** range (`note7`), with the annotation that appears on its end-marker, has been set off in italics.

Note that, in the reading rendition, the annotations can be represented only at points (and as simple asterisks, they even disappear into the text), and are not properly associated with the spans to which they belong. This means that some functionalities, such as highlighting the spans with a mouseover, or extracting the spans (what we will do in this example), are difficult, even prohibitively so in the general case.

We can use the same transformation pipeline over this source to create a half-LMNL model of the text-with-annotated ranges. Over this model, we can again conduct operations using an available XML toolkit. While again XSLT 1.0 (though not exactly operating in classic mode) proves up to the task, an XSLT 2.0 version can take advantage of XPath 2.0 operators for examining relative positions of nodes;¹² this is very helpful when the operation needs to determine which ranges begin and end within which other ranges — something the processor needs to be able to determine to write its output correctly.

The output of this stylesheet simply lists the ranges of text in the source that have annotations attached, with their annotations. (One can say “simply” since the operation is, after all, fairly simple: what is easy to miss is how this simplicity contrasts with what we would encounter trying to achieve the same output without a LMNL, or half-LMNL, layer.) Since the half-LMNL model preserves information about all the ranges spanning over both text and annotations, the output can also be enhanced by routines (named templates in the XSLT implementation) that can “write tags” where ranges are demarcated in the original. Operating straightforwardly to do this, of course, we cannot guarantee well-formedness in the XML sense; so instead we write LMNL “dragon teeth” (calling the syntax in honor of Cadmus, the founder of letters). Where a range of text has been identified as a **note** range, we write out the range, starting it with markers to show what other ranges have been started before or along with this one. As we write this range we can also write out tags to show where other ranges occur over this text. Finally, we write out close tags for those ranges that have yet to close where the glossed text stops (achieving tag balancing while not precluding overlap, as is demonstrated by the overlapping **q** and **p** ranges). We can also customize the tag writing so that when tags are written out to show a quote (a **q** range) has started, we can annotate the range with a metarange showing whose quote it is (information that was stored in an attribute in the incoming XML).

§ Current conclusions

Albeit in a fairly primitive way, these stylesheets already demonstrate the potential of more generalized applications, even in XSLT. Creating XML output (as opposed to the LMNL output marked with bracket-brace syntax) could be accomplished by a range of different approaches, starting (in cases where we knew in advance which sets of ranges in the input could be rendered as XML in the output) with brute force: for structures known to be nestable, write out XML tags; otherwise write out CLIX milestones.¹³ More attractive, perhaps, is the notion of using XSLT (or any technology) to interrogate a LMNL structure in a heuristic mode, to determine and report, either *a priori*, or by reference to an XML schema, where XML could be straightforwardly extracted. A LMNL-to-XML transform that could be specified by means of a generic XML schema is a particularly intriguing possibility.

In order to support this kind of “XML induction”, it would be very useful to have a set of operations defined to support comparison of ranges (where and how do they overlap?) and the retrieval of sets of related ranges (given a range, retrieve those ranges that start within it, and so forth). These operations are already specified, in large part, as part of the LOM Draft Specification (which provides an API for full LMNL); a mapping that describes a similar (though not identical) set of operations, and which accounts for both LMNL and half-LMNL (which must have different rules regarding the relative start and end positions of ranges) is also being worked on (see the LMNL web site at <http://lmnl.net>). In defining these operations, the distinction between half-LMNL and full LMNL (which tells not about range-starts and -ends but about ranges themselves), becomes critical to determining the boundaries of any sets of XML elements we would infer. Half-LMNL, in this context, is useful precisely because it helps us distinguish between cases where (a) we do not know, or do not have warrant to infer, or do not care about, the relative ordering of range-starts and -ends, and correspondingly (b) when this knowledge is important to us (perhaps because it apparently expresses the intent of the creator in some respect) and needs to be left alone and not regularized.

Figure 9

```

---
[body][text]{TEI.2}[q [who]Gandalf{}][p]A mortal, Frodo, who keeps one
of the Great Rings, does not die, but he does not grow or obtain more
life, he merely continues, until at last every minute is a weariness.
{q}[p]{body}{text}{TEI.2}

The ring has the power of suspending the rules of mortality and
prolonging life. That is, it [emph]preserves[emph] (though not in the
same way as the Elf-rings whose science it borrows). In this way, like
any technology, it promises control over time and space.

---
[body][text]{TEI.2}[q [who]Gandalf{}][p>he merely continues, until at
last every minute is a weariness. And if he often uses the Ring to
make himself invisible, he [emph]fades[emph]: he becomes in the end
invisible permanently, and walks in the twilight under the eye of the
dark power that rules the Rings. Yes, sooner or later—later, if he is
strong or well-meaning to begin with, but neither strength nor good
purpose will last—sooner or later the dark power will devour him. {q}
{p}{body}{text}{TEI.2}

From the very first, accounts of what happens with the Ring are
completely negative: the temptation it represents is described more
than it is demonstrated. The Ring is beautiful as an aesthetic object,
although simple; but its treacherous nature is spelled out early.

---
[body][text]{TEI.2}[q [who]Gandalf{}][p>And if he often uses the Ring
to make himself invisible, he [emph]fades[emph]: he becomes in the end
invisible permanently, and walks in the twilight under the eye of the
dark power that rules the Rings. {q}[p]{body}{text}{TEI.2}

Another power the Ring has is to make its wearer invisible. [...]

```

Part of the output of the notes-extractor stylesheet operating on the half-LMNL rendition of the *Red Book* sample text. (See “Notes extraction XSLT” to see the stylesheet and complete output.) Again, italics are added for emphasis: here they indicate where particular spans of text are annotated by more than a single **note** range. Also worth observing is how **p** and **q** ranges, as reported by the tag-writing layer that shows what ranges a particular span of text appears in, overlap, whenever a character starts speaking in one paragraph and completes his line in a later one.

One of the most instructive aspects of this work is the revelation that XML, even with only XSLT 1.0, can be perfectly well hacked to do this kind of work, at least at small scales, using techniques that must surely have been developed and applied any number of times — reified LMNL in its proposed form is not quite standoff markup, but half-LMNL might as well be.

This in turn suggests that the “problems of overlap” are really not so intractable, if we approach XML not as a “way” (a philosophy as much as a technology of markup) but as a tool that can be applied creatively to the job. That is, we should not lament XML's shortcomings and try to “fix” it, so much as understand it for what it is, use it for the jobs it is good at, and not try to force it to be something it's not. Although not ideal for work at the layer where LMNL needs to operate (these range-retrieval and range-comparison methods might surely be implemented more neatly in an object-oriented framework, in XQuery or in a database), XML proves to be surprisingly up to the job.

Getting even this far, also, has the considerable virtue of clarifying what needs to happen next with LMNL. Fleshing out operations and experimenting further with actual applications of multiple concurrent hierarchies and related phenomena will further clarify the extent to which LMNL as a markup application (or class of applications) benefits from, or can disregard, a concept (tag ordering) close to the core conception of markup languages up to this point. To test this further, the best approach would seem to be to compare full LMNL directly, in operation, to its reified markup-derived forms. This we are starting to do.

Interestingly, it may also demonstrate that a particular syntax should not have to be the stickiest point (and the main virtue of bracket-and-brace dragons' teeth may be that they are *not* XML), and that we should proceed to work on refining the LMNL model and operations over it without worrying about the details of syntax. (It took some years of working with SGML before anything like a common data model emerged — in the form of ESIS — eventually to be superseded by the present plethora of XML data models; so there's no hurry.)

Likewise it may be a less urgent concern how rapidly LMNL's development progresses in general. Experiments like those presented here should make the case for the practicability of LMNL's core concepts; the lower-level operations these stylesheets perform (such as the retrieval of all ranges that start or end within a given range) can be formalized further and perhaps implemented on other platforms; and likewise, given success at the level of application, it may be easier to motivate the development of a parser for a more suitable syntax for LMNL than XML+CLIX.

§ Appendixes: code samples

The complete suite of stylesheets for the production of half-LMNL will be made available for download (check the conference *Proceedings*); appearing here is the code of the two custom stylesheets, one for each demonstration.

A fragment of Paradise Lost

A half-LMNL model of the text

Serialized as XML, the half-LMNL result of processing the CLIX-marked instance in 6 looks like this:

```
<?xml version="1.0" encoding="utf-8"?>
<l-:document xmlns:l-="http://lmdl.org/namespace/halfLMNL/reified">
  <l-:content>
Of Man's first disobedience, and the fruit
Of that forbidden tree whose mortal taste
Brought death into the World, and all our woe,
With loss of Eden, till one greater Man
Restore us, and regain the blissful seat,
Sing, Heavenly Muse, that, on the secret top
Of Oreb, or of Sinai, didst inspire
That Shepherd who first taught the chosen seed
In the beginning how the heavens and earth
Rose out of Chaos: or, if Sion hill
Delight thee more, and Siloa's brook that flowed
Fast by the oracle of God, I thence
Invoke thy aid to my adventurous song,
That with no middle flight intends to soar
Above th' Aonian mount, while it pursues
Things unattempted yet in prose or rhyme.
And chiefly thou, O Spirit, that dost prefer
Before all temples th' upright heart and pure,
Instruct me, for Thou know'st; Thou from the first
Wast present, and, with mighty wings outspread,
Dove-like sat'st brooding on the vast Abyss,
And mad'st it pregnant: what in me is dark
Illumine, what is low raise and support;
That, to the height of this great argument,
I may assert Eternal Providence,
And justify the ways of God to men.
  </l-:content>
  <l-:range-start name="quote" id="quote-1" where="0"
length="1121"/>
  <l-:range-start name="line" id="L1" where="1" length="42">
    <l-:annotation name="n">
      <l-:content>Bk1.1</l-:content>
    </l-:annotation>
  </l-:range-start>
  <l-:range-start name="s" id="s-1" where="1" length="681"/>
  <l-:range-start name="seg" id="seg-1" where="1" length="28"/>
  <l-:range-end name="seg" id="seg-1" where="29"/>
  <l-:range-start name="seg" id="seg-2" where="30" length="86"/>
  <l-:range-end name="line" id="L1" where="43"/>
  <l-:range-start name="line" id="L2" where="44" length="42">
    <l-:annotation name="n">
      <l-:content>Bk1.2</l-:content>
    </l-:annotation>
  </l-:range-start>
  <l-:range-end name="line" id="L2" where="86"/>
  <l-:range-start name="line" id="L3" where="87" length="46">
    <l-:annotation name="n">
      <l-:content>Bk1.3</l-:content>
    </l-:annotation>
  </l-:range-start>
  <l-:range-end name="seg" id="seg-2" where="116"/>
  <l-:range-start name="seg" id="seg-3" where="117" length="16"/>
  <l-:range-end name="seg" id="seg-3" where="133"/>
  <l-:range-end name="line" id="L3" where="133"/>
  <l-:range-start name="line" id="L4" where="134" length="40">
```

```

<l-:annotation name="n">
  <l-:content>Bk1.4</l-:content>
</l-:annotation>
</l-:range-start>
<l-:range-start name="seg" id="seg-4" where="134" length="18"/>
<l-:range-end name="seg" id="seg-4" where="152"/>
<l-:range-start name="seg" id="seg-5" where="153" length="33"/>
<l-:range-end name="line" id="L4" where="174"/>
<l-:range-start name="line" id="L5" where="175" length="41">
  <l-:annotation name="n">
    <l-:content>Bk1.5</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-5" where="186"/>
<l-:range-start name="seg" id="seg-6" where="187" length="29"/>
<l-:range-end name="seg" id="seg-6" where="216"/>
<l-:range-end name="line" id="L5" where="216"/>
<l-:range-start name="line" id="L6" where="217" length="45">
  <l-:annotation name="n">
    <l-:content>Bk1.6</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-start name="seg" id="seg-7" where="217" length="5"/>
<l-:range-end name="seg" id="seg-7" where="222"/>
<l-:range-start name="seg" id="seg-8" where="223" length="14"/>
<l-:range-end name="seg" id="seg-8" where="237"/>
<l-:range-start name="seg" id="seg-9" where="238" length="5"/>
<l-:range-end name="seg" id="seg-9" where="243"/>
<l-:range-start name="seg" id="seg-10" where="244" length="27"/>
<l-:range-end name="line" id="L6" where="262"/>
<l-:range-start name="line" id="L7" where="263" length="36">
  <l-:annotation name="n">
    <l-:content>Bk1.7</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-10" where="271"/>
<l-:range-start name="seg" id="seg-11" where="272" length="12"/>
<l-:range-end name="seg" id="seg-11" where="284"/>
<l-:range-start name="seg" id="seg-12" where="285"
length="125"/>
<l-:range-end name="line" id="L7" where="299"/>
<l-:range-start name="line" id="L8" where="300" length="47">
  <l-:annotation name="n">
    <l-:content>Bk1.8</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-end name="line" id="L8" where="347"/>
<l-:range-start name="line" id="L9" where="348" length="43">
  <l-:annotation name="n">
    <l-:content>Bk1.9</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-end name="line" id="L9" where="391"/>
<l-:range-start name="line" id="L10" where="392" length="36">
  <l-:annotation name="n">
    <l-:content>Bk1.10</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-12" where="410"/>
<l-:range-start name="seg" id="seg-13" where="411" length="3"/>
<l-:range-end name="seg" id="seg-13" where="414"/>
<l-:range-start name="seg" id="seg-14" where="415" length="32"/>
<l-:range-end name="line" id="L10" where="428"/>
<l-:range-start name="line" id="L11" where="429" length="49">
  <l-:annotation name="n">
    <l-:content>Bk1.11</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-14" where="447"/>
<l-:range-start name="seg" id="seg-15" where="448" length="57"/>
<l-:range-end name="line" id="L11" where="478"/>
<l-:range-start name="line" id="L12" where="479" length="36">
  <l-:annotation name="n">
    <l-:content>Bk1.12</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-15" where="505"/>
<l-:range-start name="seg" id="seg-16" where="506" length="48"/>
<l-:range-end name="line" id="L12" where="515"/>
<l-:range-start name="line" id="L13" where="516" length="38">
  <l-:annotation name="n">

```

```

        <l-:content>Bk1.13</l-:content>
    </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-16" where="554"/>
<l-:range-end name="line" id="L13" where="554"/>
<l-:range-start name="line" id="L14" where="555" length="43">
    <l-:annotation name="n">
        <l-:content>Bk1.14</l-:content>
    </l-:annotation>
</l-:range-start>
<l-:range-start name="seg" id="seg-17" where="555" length="67"/>
<l-:range-end name="line" id="L14" where="598"/>
<l-:range-start name="line" id="L15" where="599" length="41">
    <l-:annotation name="n">
        <l-:content>Bk1.15</l-:content>
    </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-17" where="622"/>
<l-:range-start name="seg" id="seg-18" where="623" length="59"/>
<l-:range-end name="line" id="L15" where="640"/>
<l-:range-start name="line" id="L16" where="641" length="41">
    <l-:annotation name="n">
        <l-:content>Bk1.16</l-:content>
    </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-18" where="682"/>
<l-:range-end name="s" id="s-1" where="682"/>
<l-:range-end name="line" id="L16" where="682"/>
<l-:range-start name="line" id="L17" where="683" length="45">
    <l-:annotation name="n">
        <l-:content>Bk1.17</l-:content>
    </l-:annotation>
</l-:range-start>
<l-:range-start name="s" id="s-2" where="683" length="437"/>
<l-:range-start name="seg" id="seg-19" where="683" length="17"/>
<l-:range-end name="seg" id="seg-19" where="700"/>
<l-:range-start name="seg" id="seg-20" where="701" length="9"/>
<l-:range-end name="seg" id="seg-20" where="710"/>
<l-:range-start name="seg" id="seg-21" where="711" length="64"/>
<l-:range-end name="line" id="L17" where="728"/>
<l-:range-start name="line" id="L18" where="729" length="46">
    <l-:annotation name="n">
        <l-:content>Bk1.18</l-:content>
    </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-21" where="775"/>
<l-:range-end name="line" id="L18" where="775"/>
<l-:range-start name="line" id="L19" where="776" length="51">
    <l-:annotation name="n">
        <l-:content>Bk1.19</l-:content>
    </l-:annotation>
</l-:range-start>
<l-:range-start name="seg" id="seg-22" where="776" length="12"/>
<l-:range-end name="seg" id="seg-22" where="788"/>
<l-:range-start name="seg" id="seg-23" where="789" length="17"/>
<l-:range-end name="seg" id="seg-23" where="806"/>
<l-:range-start name="seg" id="seg-24" where="807" length="34"/>
<l-:range-end name="line" id="L19" where="827"/>
<l-:range-start name="line" id="L20" where="828" length="48">
    <l-:annotation name="n">
        <l-:content>Bk1.20</l-:content>
    </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-24" where="841"/>
<l-:range-start name="seg" id="seg-25" where="842" length="4"/>
<l-:range-end name="seg" id="seg-25" where="846"/>
<l-:range-start name="seg" id="seg-26" where="847" length="29"/>
<l-:range-end name="seg" id="seg-26" where="876"/>
<l-:range-end name="line" id="L20" where="876"/>
<l-:range-start name="line" id="L21" where="877" length="45">
    <l-:annotation name="n">
        <l-:content>Bk1.21</l-:content>
    </l-:annotation>
</l-:range-start>
<l-:range-start name="seg" id="seg-27" where="877" length="45"/>
<l-:range-end name="seg" id="seg-27" where="922"/>
<l-:range-end name="line" id="L21" where="922"/>
<l-:range-start name="line" id="L22" where="923" length="43">
    <l-:annotation name="n">
        <l-:content>Bk1.22</l-:content>
    </l-:annotation>
</l-:range-start>

```



```

</l-:range-start>
<l-:range-start name="seg" id="seg-28" where="923" length="23"/>
<l-:range-end name="seg" id="seg-28" where="946"/>
<l-:range-start name="seg" id="seg-29" where="947" length="29"/>
<l-:range-end name="line" id="L22" where="966"/>
<l-:range-start name="line" id="L23" where="967" length="40">
  <l-:annotation name="n">
    <l-:content>Bk1.23</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-end name="seg" id="seg-29" where="976"/>
<l-:range-start name="seg" id="seg-30" where="977" length="30"/>
<l-:range-end name="seg" id="seg-30" where="1007"/>
<l-:range-end name="line" id="L23" where="1007"/>
<l-:range-start name="line" id="L24" where="1008" length="43">
  <l-:annotation name="n">
    <l-:content>Bk1.24</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-start name="seg" id="seg-31" where="1008" length="5"/>
<l-:range-end name="seg" id="seg-31" where="1013"/>
<l-:range-start name="seg" id="seg-32" where="1014"
length="37"/>
<l-:range-end name="seg" id="seg-32" where="1051"/>
<l-:range-end name="line" id="L24" where="1051"/>
<l-:range-start name="line" id="L25" where="1052" length="32">
  <l-:annotation name="n">
    <l-:content>Bk1.25</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-start name="seg" id="seg-33" where="1052"
length="32"/>
<l-:range-end name="seg" id="seg-33" where="1084"/>
<l-:range-end name="line" id="L25" where="1084"/>
<l-:range-start name="line" id="L26" where="1085" length="35">
  <l-:annotation name="n">
    <l-:content>Bk1.26</l-:content>
  </l-:annotation>
</l-:range-start>
<l-:range-start name="seg" id="seg-34" where="1085"
length="35"/>
<l-:range-end name="seg" id="seg-34" where="1120"/>
<l-:range-end name="s" id="s-2" where="1120"/>
<l-:range-end name="line" id="L26" where="1120"/>
<l-:range-end name="quote" id="quote-1" where="1121"/>
</l-:document>

```

It may be of particular interest to innovators of approaches to markup that treat multiple concurrent hierarchies, or anything related to the “overlap” problem, whether and how this model (or a reified true LMNL model derived from it) maps to their own constructs.

Enjambment Analysis XSLT

This stylesheet effectively uses LMNL, although it is written to run over half-LMNL to save a final step. But note that there is no reliance on tag-ordering to disambiguate semantics here; on the contrary, its key (which retrieves **seg** elements by their offset) is used only to find where **seg** and **line** structures co-occur — precisely not which one comes first.

```

<?xml version="1.0" encoding="UTF-8" ?>
<xsl:stylesheet version="1.0"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:l="http://lmnl.org/namespace/halfLMNL/reified"
  >

  <!-- query document

  input: reified LMNL with line and seg ranges
  output: lines that break across segs (enjambment)

  -->

  <xsl:output method="text"/>

  <xsl:variable name="text-layer"
    select="/l-:document/l-:content"/>

  <xsl:key name="range-end-by-id" match="l-:range-end" use="@id"/>

```

```

<xsl:key name="seg-by-start" match="l-:range-start[@name='seg']"
  use="@where"/>

<xsl:key name="seg-by-end" match="l-:range-end[@name='seg']"
  use="@where"/>

<xsl:template match="/">
  <xsl:apply-templates
    select="l-:document/l-:range-start[@name='line']"/>
</xsl:template>

<xsl:template match="l-:range-start[@name='line']">
  <xsl:variable name="line-end"
    select="key('range-end-by-id',@id)"/>
  <xsl:text>&#xA;</xsl:text>
  <xsl:number count="l-:range-start[@name='line']" format="000"/>
  <xsl:text> </xsl:text>
  <xsl:choose>
    <!-- no enjambment if a seg that begins with the line -->
    <xsl:when test="key('seg-by-start',@where)"></xsl:when>
    <xsl:otherwise>~</xsl:otherwise>
  </xsl:choose>
  <xsl:choose>
    <!-- no enjambment (end-stopped) if a seg ends where the line
    ends -->
    <xsl:when test="key('seg-by-end', $line-end/@where)">]
    </xsl:when>
    <xsl:otherwise>~</xsl:otherwise>
  </xsl:choose>
  <xsl:text> </xsl:text>
  <xsl:call-template name="range-text"/>
</xsl:template>

<xsl:template name="range-text">
  <xsl:param name="range" select="."/>
  <xsl:value-of select="substring($text-layer, ($range/@where+1),
    ($range/@length))"/>
</xsl:template>

</xsl:stylesheet>

```

Scholarly annotations

TEI source, with CLIX

An edition marked up in TEI, with CLIX markup for notes and quotes as well as implicit “range markers” in the form of **pb** milestones, all to be rendered into LMNL as overlapping phenomena, appears as follows¹⁴:

```

<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="teitext.xsl"?>
<TEI.2 xmlns:l="http://lmnl.org/namespace/clix">
  <teiHeader type="text" status="new">
    <fileDesc>
      <titleStmt>
        <title type="main">Excerpts from the Red Book of
          Westmarch</title>
        <title type="sub">with LMNL annotations</title>
      </titleStmt>
      <editionStmt>
        <edition/>
      </editionStmt>
      <publicationStmt>
        <availability status="restricted">
          <p>The original text is copyright 1954, 1965,
            1966 by JRR Tolkien; 1965 ed. renewed 1993,
            1994 by Christopher R Tolkien, John FR Tolkien
            and Priscilla MAR Tolkien. These excerpts
            appear in the spirit of fair use, as a
            demonstration of electronic encoding
            technologies, and are not for sale. Notes are
            in the public domain.</p>
        </availability>
        <date>2003-05-27</date>
      </publicationStmt>
      <sourceDesc default="NO"><p>Original text scanned from
        Ballantine edition; notes provided by the encoder.</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>

```

```

</fileDesc>
<encodingDesc>
  <editorialDecl default="NO">
    <p>Basic TEI tagging of paragraphs, inline
    phenomena, page breaks, certain quotes (including
    dialogue).</p>
    <p>CLIX tagging of passages with annotations
    (appearing as the content of notes elements),
    speeches (which overlap with paragraphs)</p>
  </editorialDecl>
</encodingDesc>
</teiHeader>
<text>
  <body>

<pb n="50"/>

<p>Next morning after a late breakfast, the wizard was sitting with
Frodo by the open window of the study. A bright fire was on the
hearth, but the sun was warm, and the wind was in the South.
Everything looked fresh, and the new green of Spring was shimmering
in the fields and on the tips of the trees' fingers.</p>

<p>Gandalf was thinking of a spring, nearly eighty years before,
when Bilbo had run out of Bag End without a handkerchief. His hair
was perhaps whiter than it had been then, and his beard and
eyebrows were perhaps longer, and his face more lined with care and
wisdom; but his eyes were as bright as ever, and he smoked and blew
smoke-rings with the same vigour and delight.</p>

<p>He was smoking now in silence, for Frodo was sitting still, deep
in thought. Even in the light of morning he felt the dark shadow of
the tidings that Gandalf had brought. At last he broke the
silence.</p>

<p><q l:sID="q1" who="Frodo"/>Last night you began to tell me
strange things about my ring, Gandalf,<q l:eID="q1"/> he said. <q
l:sID="q2" who="Frodo"/>And then you stopped, because you said that
such matters were best left until daylight. Don't you think you had
better finish now? You say the ring is dangerous, far more
dangerous than I guess. In what way?<q l:eID="q2"/></p>

<p><q l:sID="q3" who="Gandalf"/>In many ways,<q l:eID="q3"/>
answered the wizard. <q l:sID="q4" who="Gandalf"/>It is far more
powerful than I ever dared to think at first, so powerful that in
the end it would utterly overcome anyone of mortal race who
possessed it. It would possess him.</p>

<p>In Eregion long ago many Elven-rings were made, magic rings as
you call them, and they were, of course, of various kinds: some
more potent and some less. The lesser rings were only essays in the
craft before it was full-grown, and to the Elven-smiths they were
but trifles—yet still to my mind dan<pb n="51"/>gerous for mortals.
But the Great Rings, the Rings of Power, they were perilous.</p>

<p><note l:sID="note1"/>A mortal, Frodo, who keeps one of the Great
Rings, does not die, but he does not grow or obtain more life,
<note l:sID="note7"/>he merely continues, until at last every
minute is a weariness.<note l:eID="note1">The ring has the power of
suspending the rules of mortality and prolonging life. That is, it
<emph>preserves</emph> (though not in the same way as the Elf-rings
whose science it borrows). In this way, like any technology, it
promises control over time and space.</note> <note l:sID="note2"/>
And if he often uses the Ring to make himself invisible, he <emph>
fades</emph>: he becomes in the end invisible permanently, and
walks in the twilight under the eye of the dark power that rules
the Rings.<note l:eID="note2">Another power the Ring has is to make
its wearer invisible. Interestingly, many or most of the Ring's
powers only come into play when it is being worn (that is, being
claimed). Invisibility amounts to the power to act without taking
responsibility—without being seen or recognized. But the Ring's
power is double-edged: while it allows its wearer to go unseen by
those from whom he wishes to hide, it exposes him to the
observation of the will or purpose identified with the Ring, that
is the Dark Lord. (Note that the Dark Lord himself exists as little
more than a disembodied principle of malice, and never takes much
more tangible form in the novel, acting for the most part through
his agents.) Thus the struggle between Frodo and the Ring is deeply
personal, a struggle between Frodo and himself: the Ring is
precisely a means by which Frodo can indulge his wish to hide,
though at the price of being known by the malevolence expressed in

```

that hiding, whether that be identified as Frodo's own more selfish motives, or a motive <soCalled>built in</soCalled> to the Ring by its creator. To wear the Ring is thus to take on a burden of guilt or knowledge of one's weakness and fallibility, as well as exposure and subservience to the greater power that made it and whose purpose is expressed in it: to claim the Ring is to embrace, to one's own ultimate destruction, a fallen condition.</note> Yes, sooner or later—later, if he is strong or well-meaning to begin with, but neither strength nor good purpose will last—sooner or later the dark power will devour him.<note l:eID="note7">From the very first, accounts of what happens with the Ring are completely negative: the temptation it represents is described more than it is demonstrated. The Ring is beautiful as an aesthetic object, although simple; but its treacherous nature is spelled out early.</note><q l:eID="q4" who="Gandalf"/></p>

<p><q l:sID="q5" who="Frodo"/>How terrifying!<q l:eID="q5"/> said Frodo. There was another long silence. The sound of Sam Gamgee cutting the lawn came in from the garden.</p>

<milestone unit="div"/>

<p><q l:sID="q6" who="Frodo"/>How long have you known this?<q l:eID="q6"/> asked Frodo at length. <q l:sID="q7" who="Frodo"/>And how much did Bilbo know?<q l:eID="q7"/></p>

<p><q l:sID="q8" who="Gandalf"/>Bilbo knew no more than he told you, I am sure,<q l:eID="q8"/> said Gandalf. <q l:sID="q9" who="Gandalf"/>He would certainly never have passed on to you anything that he thought would be a danger, even though I promised to look after you. <note l:sID="note3"/>He thought the ring was very beautiful, and very useful at need; and if anything was wrong or queer, it was himself. He said that it was <q>growing on his mind</q>, and he was always worrying about it; but he did not suspect that the ring itself was to blame. Though he had found out that the thing needed looking after; it did not seem always of the same size or weight; it shrank or expanded in an odd way, and might suddenly slip off a finger where it had been tight.<note l:eID="note3">Bilbo evidencing the effect of mixed subjectivity. Interestingly, Gandalf draws attention to the Ring's independence and willfulness by noting how it grows and shrinks (a physical effect), leaving implicit its moral effects both of causing Bilbo to fret and of confusing Bilbo as to the source of his own feelings. Thus Gandalf is willing to attribute responsibility and self-interest to the Ring despite (and by way of observing) Bilbo's being deceived on that very point (about what's Bilbo's fault and what isn't).</note><q l:eID="q9"/></p>

<p><q l:sID="q10" who="Frodo"/>Yes, he warned me of that in his last letter,<q l:eID="q10"/> said Frodo, <q l:sID="q11" who="Frodo"/>so I have always kept it on its chain.<q l:eID="q11"/></p>

<p><q l:sID="q12" who="Gandalf"/>Very wise,<q l:eID="q12"/> said Gandalf. <q l:sID="q13" who="Gandalf"/>But as for his long life, Bilbo never connected it with the ring at all. He took all the credit for that to himself, and he was very proud of it. Though he was getting restless and uneasy. <emph>Thin and stretched</emph> he said. A sign that the ring was getting control.<q l:eID="q13"/></p>

<p><q l:sID="q14" who="Frodo"/>How long have you known all this?<q l:eID="q14"/> asked Frodo again. <q l:sID="q15" who="Gandalf"/>Known?<q l:eID="q15"/> said Gandalf. <q l:sID="q16" who="Gandalf"/>I have known much that only the Wise know, Frodo. But if you mean "known about <emph>this</emph> ring", well, I still do not <emph>know</emph>, one might say.<note l:sID="note6"/>There is a last test to make. But I no longer doubt my guess.</p>

<pb n="52"/>

<p>When did I first begin to guess?<q l:eID="q16"/> he mused, searching back in memory.<note l:eID="note6">Gandalf's recollection leads to a narrative-within-the-narrative, which itself relates to earlier stories.</note> <q l:sID="q17" who="Gandalf"/>Let me see—it was in the year that the White Council drove the dark power from Mirkwood, just before the Battle of Five Armies, that Bilbo found his ring. A shadow fell on my heart then, though I did not know yet what I feared. I wondered often how Gollum came by a Great Ring, as plainly it was—that at least was clear from the first. Then I heard Bilbo's strange story of how he had "won" it, and I could not believe it. <note l:sID="note4"/>When I at last got the truth out

of him, I saw at once that he had been trying to put his claim to the ring beyond doubt.<note l:eID="note4">Notice both Bilbo's proprietary attitude towards the ring, and his wish to keep it hidden.</note> Much like Gollum with his "birthday present". The lies were too much alike for my comfort. Clearly the ring had an unwholesome power that set to work on its keeper at once. That was the first real warning I had that all was not well. I told Bilbo often that such rings were better left unused; but he resented it, and soon got angry. There was little else that I could do. I could not take it from him without doing greater harm; and I had no right to do so anyway. I could only watch and wait. I might perhaps have consulted Saruman the White, but something always held me back.<q l:eID="q17"/></p>

<p><q l:sID="q18" who="Frodo"/>Who is he?<q l:eID="q18"/> asked Frodo. <q l:sID="q19" who="Frodo"/>I have never heard of him before.<q l:eID="q19"/> <q l:sID="q20" who="Gandalf"/>Maybe not,<q l:eID="q20"/> answered Gandalf. <q l:sID="q21" who="Gandalf"/>Hobbits are, or were, no concern of his. Yet he is great among the Wise. He is the chief of my order and the head of the Council. His knowledge is deep, but his pride has grown with it, and he takes ill any meddling. The lore of the Elven-rings, great and small, is his province. He has long studied it, seeking the lost secrets of their making; but when the Rings were debated in the Council, all that he would reveal to us of his ring-lore told against my fears. So my doubt slept—but uneasily. Still I watched and I waited.</p>

<p>And all seemed well with Bilbo. And the years passed. Yes, they passed, and they seemed not to touch him. He showed no signs of age. The shadow fell on me again. But I said to myself: <q>After all he comes of a long-lived family on his mother's side. There is time yet. Wait!</q></p>

<p>And I waited. Until that night when he left this house. He said and did things then that filled me with a fear that no words of Saruman could allay. I knew at last that something dark and deadly was at work. And I have spent most of the years since then in finding out the truth of it.<q l:eID="q21"/></p>

<p><q l:sID="q22" who="Frodo"/>There wasn't any permanent harm done, was there?<q l:eID="q22"/> asked <pb n="53"/>Frodo anxiously. <q l:sID="q23" who="Frodo"/>He would get all right in time, wouldn't he? Be able to rest in peace, I mean?<q l:eID="q23"/></p>

<p><q l:sID="q24" who="Gandalf"/>He felt better at once,<q l:eID="q24"/> said Gandalf. <q l:sID="q25" who="Gandalf"/>But there is only one Power in this world that knows all about the Rings and their effects; and as far as I know there is no Power in the world that knows all about hobbits. Among the Wise I am the only one that goes in for hobbit-lore: an obscure branch of knowledge, but full of surprises. Soft as butter they can be, and yet sometimes as tough as old tree-roots. I think it likely that some would resist the Rings far longer than most of the Wise would believe. I don't think you need worry about Bilbo.</p>

<p>'Of course, he possessed the ring for many years, and used it, so it might take a long while for the influence to wear off— before it was safe for him to see it again, for instance. Otherwise, he might live on for years, quite happily: just stop as he was when he parted with it. <note l:sID="note5"/>For he gave it up in the end of his own accord: an important point.<note l:eID="note5"/>Bilbo and Sam are the only characters in the novel to give up the Ring willingly; Gandalf and Galadriel (and somewhat less momentarily, Aragorn and Faramir) also refuse it when offered, or when they have an opportunity to take it (also not insignificant).</note> No, I was not troubled about dear Bilbo any more, once he had let the thing go. It is for <emph>you</emph> that I feel responsible.</p>

<p>Ever since Bilbo left I have been deeply concerned about you, and about all these charming, absurd, helpless hobbits. It would be a grievous blow to the world, if the Dark Power overcame the Shire; if all your kind, jolly, stupid Bolgers, Hornblowers, Boffins, Bracegirdles, and the rest, not to mention the ridiculous Bagginses, became enslaved.<q l:eID="q25"/></p>

<p>Frodo shuddered. <q l:sID="q26" who="Frodo"/>But why should we be?<q l:eID="q26"/> he asked. <q l:sID="q27" who="Frodo"/>And why should he want such slaves?<q l:eID="q27"/></p>

<p><q l:sID="q28" who="Gandalf"/>To tell you the truth,<q l:eID="q28"/> replied Gandalf, <q l:sID="q29" who="Gandalf"/>I

```

believe that hitherto— hitherto, mark you—he has
entirely overlooked the existence of hobbits. You should be
thankful. But your safety has passed. He does not need you—he has
many more useful servants—but he won't forget you again. And
hobbits as miserable slaves would please him far more than hobbits
happy and free. There is such a thing as malice and revenge.<q
l:eID="q29"/></p>

<p><q l:sID="q30" who="Gandalf"/>Revenge?<q l:eID="q30"/> said
Frodo. <q l:sID="q31" who="Frodo"/>Revenge for what? I still don't
understand what all this has to do with Bilbo and myself, and our
ring.<q l:eID="q31"/></p>

<p><q l:sID="q32" who="Gandalf"/>It has everything to do with it,<q
l:eID="q32"/> said Gandalf. <q l:sID="q33" who="Gandalf"/>You do
not know the real peril yet; but you shall. I was not sure of it
myself when I was last here; but the time has come to speak. Give
me the ring for a moment.<q l:eID="q33"/></p>

<pb n="54"/>

</body></text></TEI.2>

```

Processing results: the extracted notes

The full result of the notes-extractor stylesheet running over the half-LMNL derived from the preceding:

```

---
[TEI.2][text][body][q [who]Gandalf{}][page [n]51{}][p]A mortal,
Frodo, who keeps one of the Great Rings, does not die, but he does
not grow or obtain more life, he merely continues, until at last
every minute is a weariness. {q}{p}{page}{body}{text}{TEI.2}

The ring has the power of suspending the rules of mortality and
prolonging life. That is, it [emph]preserves{emph} (though not in
the same way as the Elf-rings whose science it borrows). In this
way, like any technology, it promises control over time and space.

---
[TEI.2][text][body][q [who]Gandalf{}][page [n]51{}][p]he merely
continues, until at last every minute is a weariness. And if he
often uses the Ring to make himself invisible, he [emph]
fades{emph}: he becomes in the end invisible permanently, and walks
in the twilight under the eye of the dark power that rules the
Rings. Yes, sooner or later—later, if he is strong or well-meaning
to begin with, but neither strength nor good purpose will
last—sooner or later the dark power will devour him. {q}{p}{page}
{body}{text}{TEI.2}

From the very first, accounts of what happens with the Ring are
completely negative: the temptation it represents is described more
than it is demonstrated. The Ring is beautiful as an aesthetic
object, although simple; but its treacherous nature is spelled out
early.

---
[TEI.2][text][body][q [who]Gandalf{}][page [n]51{}][p]And if he
often uses the Ring to make himself invisible, he [emph]
fades{emph}: he becomes in the end invisible permanently, and walks
in the twilight under the eye of the dark power that rules the
Rings. {q}{p}{page}{body}{text}{TEI.2}

Another power the Ring has is to make its wearer invisible.
Interestingly, many or most of the Ring's powers only come into
play when it is being worn (that is, being claimed). Invisibility
amounts to the power to act without taking responsibility—without
being seen or recognized. But the Ring's power is double-edged:
while it allows its wearer to go unseen by those from whom he
wishes to hide, it exposes him to the observation of the will or
purpose identified with the Ring, that is the Dark Lord. (Note that
the Dark Lord himself exists as little more than a disembodied
principle of malice, and never takes much more tangible form in the
novel, acting for the most part through his agents.) Thus the
struggle between Frodo and the Ring is deeply personal, a struggle
between Frodo and himself: the Ring is precisely a means by which
Frodo can indulge his wish to hide, though at the price of being
known by the malevolence expressed in that hiding, whether that be
identified as Frodo's own more selfish motives, or a motive "built
in" to the Ring by its creator. To wear the Ring is thus to take on
a burden of guilt or knowledge of one's weakness and fallibility,

```

as well as exposure and subservience to the greater power that made it and whose purpose is expressed in it: to claim the Ring is to embrace, to one's own ultimate destruction, a fallen condition.

```

---
[TEI.2][text][body][page [n]51{}][p][q [who]Gandalf{}]He thought
the ring was very beautiful, and very useful at need; and if
anything was wrong or queer, it was himself. He said that it was
[q]growing on his mind[q], and he was always worrying about it; but
he did not suspect that the ring itself was to blame. Though he had
found out that the thing needed looking after; it did not seem
always of the same size or weight; it shrank or expanded in an odd
way, and might suddenly slip off a finger where it had been tight.
[q]{p}[page]{body}{text}[TEI.2]

```

Bilbo evidencing the effect of mixed subjectivity. Interestingly, Gandalf draws attention to the Ring's independence and willfulness by noting how it grows and shrinks (a physical effect), leaving implicit its moral effects both of causing Bilbo to fret and of confusing Bilbo as to the source of his own feelings. Thus Gandalf is willing to attribute responsibility and self-interest to the Ring despite (and by way of observing) Bilbo's being deceived on that very point (about what's Bilbo's fault and what isn't).

```

---
[TEI.2][text][body][page [n]51{}][p][q [who]Gandalf{}]There is a
last test to make. But I no longer doubt my guess.{p}[page][page
[n]52{}][p]When did I first begin to guess?[q] he mused, searching
back in memory. {p}[page]{body}{text}[TEI.2]

```

Gandalf's recollection leads to a narrative-within-the-narrative, which itself relates to earlier stories.

```

---
[TEI.2][text][body][page [n]52{}][p][q [who]Gandalf{}]When I at
last got the truth out of him, I saw at once that he had been
trying to put his claim to the ring beyond doubt. {q}[p][page]
{body}{text}[TEI.2]

```

Notice both Bilbo's proprietary attitude towards the ring, and his wish to keep it hidden.

```

---
[TEI.2][text][body][page [n]53{}][q [who]Gandalf{}][p]For he gave
it up in the end of his own accord: an important point. {p}[q]
{page}{body}{text}[TEI.2]

```

Bilbo and Sam are the only characters in the novel to give up the Ring willingly; Gandalf and Galadriel (and somewhat less momentarily, Aragorn and Faramir) also refuse it when offered, or when they have an opportunity to take it (also not insignificant).

Notes extraction XSLT

Most of this stylesheet is a prototype library of templates used for “tag writing” the output ... using bracket-and-brace syntax. Without taking any trouble to assure that ranges are nested according to LMNL rules, they simply “blurt out” whatever range-markers (begin- or end-range events) they may happen to run into. Were they provided with enough logic (either hard-wired by range type, or by some other means such as an external spec or schema) to assure that the output would be well-formed, these routines could write out XML instead. The main logic of the stylesheet does nothing but retrieve all **note** range-start markers, processing each noted range in turn. First the noted range itself is written out; then its annotation (the note itself, which appears on the end-range-marker) is written out, only without the contextual tagging provided for the extracted range.

Again, there is not special reliance here on half-LMNL logic; this stylesheet could also be written at the “full LMNL” level. It also points the way to generalizing tag-writing as a useful technique for LMNL processing in general.

The code listing for the entire stylesheet appears, plus a copy of the entire output resulting from running it on the sample text.

```

<?xml version="1.0" encoding="UTF-8" ?>
<xsl:stylesheet version="2.0"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:l-="http://lmnl.org/namespace/halfLMNL/reified">

```

```

<!-- query document
input: reified LMNL with note ranges
output: notes with the ranges they annotate
-->

<xsl:output method="text"/>

<xsl:key name="range-end-by-id" match="l-:range-end" use="@id"/>

<xsl:key name="range-start-by-id" match="l-:range-start"
use="@id"/>

<xsl:template match="/">
  <xsl:apply-templates
    select="l-:document/l-:range-start[@name='note']"/>
</xsl:template>

<xsl:template match="l-:range-start[@name='note']">
  <xsl:text>---&#xA;</xsl:text>
  <xsl:apply-templates select="." mode="range-marked"/>
  <xsl:text>&#xA;</xsl:text>
  <xsl:text>&#xA;</xsl:text>

  <!-- we assume the annotation itself is on the end-range
marker -->
  <xsl:apply-templates
    select="key('range-end-by-id',@id)/l-:annotation[not(@name)]"
    mode="write-annotation"/>
  <xsl:text>&#xA;</xsl:text>
  <xsl:text>&#xA;</xsl:text>
</xsl:template>

<xsl:template match="l-:range-start" mode="range-marked">
  <xsl:variable name="this-start" select="."/>
  <xsl:variable name="this-end" select="key('range-end-by-id',
  @id)"/>
  <xsl:variable name="already-started"
    select="$this-start/preceding-sibling::l-:range-start[
    key('range-end-by-id', @id) >> $this-start]"/>
  <!-- $already-started are start-range markers before ours
whose proper end-ranges do not appear before this one -->
  <xsl:variable name="not-yet-ended"
    select="$this-end/following-sibling::l-:range-end[
    key('range-start-by-id', @id) &lt;&lt; $this-end ]"/>
  <!-- $not-yet-ended are end-range nodes after ours whose
start-ranges appear before this end-range
(incidentally these will belong to "unclosed"
already-started and starting-inside ranges) -->
  <xsl:variable name="starting-inside"
    select="$this-start/following-sibling::l-:range-start[.
    &lt;&lt; $this-end]"/>
  <!-- -->
  <xsl:variable name="ending-inside"
    select="$this-end/preceding-sibling::l-:range-end[
    . >> $this-start]"/>
  <!-- and a way to get end-ranges for end-overlapping
and containing ranges (ranges started but not yet ended)
-->
  <xsl:apply-templates select="$already-started" mode=
"wrap-range-marker">
    <xsl:sort select="count(preceding-sibling::l-:range-start)"
order="ascending"/>
    <!-- in XSLT 2.0 we sort to get a forward sequence -->
  </xsl:apply-templates>
  <xsl:call-template name="write-marked-range">
    <xsl:with-param name="start" select="$this-start"/>
    <xsl:with-param name="end" select="$this-end"/>
    <xsl:with-param name="range-events"
      select="$starting-inside | $ending-inside"/>
  </xsl:call-template>
  <xsl:text> </xsl:text>
  <xsl:apply-templates select="$not-yet-ended" mode=
"wrap-range-marker"/>
</xsl:template>

<xsl:template name="range-text">
  <xsl:param name="range-start" select="."/>
  <xsl:variable name="text-layer" select=
"preceding-sibling::l-:content"/>
  <xsl:value-of select="substring($text-layer,
($range-start/@where+1), ($range-start/@length))"/>

```



```

</xsl:template>

<xsl:template name="write-marked-range">
  <xsl:param name="start" select="."/ >
  <xsl:param name="end" select="key('range-end-by-id', @id)"/ >
  <xsl:param name="range-events" select="/.."/ >
  <xsl:variable name="text-layer" select="
preceding-sibling::l-:content"/ >
  <xsl:choose>
    <xsl:when test="$range-events">
      <xsl:variable name="next-event" select="$range-events[1]"/ >
      <xsl:value-of
        select="substring($text-layer, ($start/@where + 1),
          ($next-event/@where - $start/@where))"/ >
      <xsl:apply-templates select="$next-event" mode="
write-range-marker"/ >
      <xsl:call-template name="write-marked-range">
        <xsl:with-param name="start" select="$next-event"/ >
        <xsl:with-param name="range-events" select="
$range-events[position() &gt; 1]"/ >
      </xsl:call-template>
    </xsl:when>
    <xsl:otherwise>
      <xsl:value-of
        select="substring($text-layer, ($start/@where + 1),
          (($end/@where - $start/@where))"/ >
    </xsl:otherwise>
  </xsl:choose>
</xsl:template>

<xsl:template name="write-annotation" match="l-:annotation"
mode="write-annotation">
  <xsl:param name="range-events"
    select="l-:range-start | l-:range-end"/ >
  <xsl:param name="start-offset" select="0"/ >
  <xsl:param name="text-layer" select="l-:content"/ >
  <xsl:variable name="length"
    select="string-length($text-layer)"/ >
  <xsl:choose>
    <xsl:when test="$range-events">
      <xsl:variable name="next-event" select="$range-events[1]"/ >
      <xsl:value-of
        select="substring($text-layer, ($start-offset + 1),
          ($next-event/@where - $start-offset))"/ >
      <xsl:apply-templates select="$next-event" mode="
write-range-marker"/ >
      <xsl:call-template name="write-annotation">
        <xsl:with-param name="text-layer" select="$text-layer"/ >
        <xsl:with-param name="start-offset"
          select="$next-event/@where"/ >
        <xsl:with-param name="range-events" select="
$range-events[position() &gt; 1]"/ >
      </xsl:call-template>
    </xsl:when>
    <xsl:otherwise>
      <xsl:value-of
        select="substring($text-layer, ($start-offset + 1),
          ($length - $start-offset))"/ >
    </xsl:otherwise>
  </xsl:choose>
</xsl:template>

<xsl:template match="l-:range-start" mode="wrap-range-marker">
  <xsl:text></xsl:text>
  <xsl:value-of select="@name"/ >
  <xsl:apply-templates select="self::*[@name='q' or
    @name='page']/l-:annotation" mode="annotation-marker"/ >
  <xsl:text></xsl:text>
</xsl:template>

<xsl:template match="l-:annotation" mode="annotation-marker"/ >

<xsl:template match="l-:annotation[@name='who' or @name='n']"
mode="annotation-marker">
  <xsl:text> [</xsl:text>
  <xsl:value-of select="@name"/ >
  <xsl:text>]</xsl:text>
  <xsl:call-template name="write-annotation"/ >
  <xsl:text>{]</xsl:text>
</xsl:template>

```

```

<xsl:template priority="10" mode="wrap-range-marker"
  match="l-:range-start[@name='note'] |
  l-:range-end[@name='note']"
/>

<xsl:template match="l-:range-end" mode="wrap-range-marker">
  <xsl:text></xsl:text>
  <xsl:value-of select="@name"/>
  <xsl:text></xsl:text>
</xsl:template>

<xsl:template match="l-:range-start" mode="write-range-marker">
  <xsl:text></xsl:text>
  <xsl:value-of select="@name"/>
  <xsl:text></xsl:text>
</xsl:template>

<xsl:template match="l-:range-start[@name='page']" mode=
"write-range-marker">
  <xsl:text></xsl:text>
  <xsl:value-of select="@name"/>
  <xsl:apply-templates
    select="l-:annotation[@name='n']" mode="annotation-marker"/>
  <xsl:text></xsl:text>
</xsl:template>

<xsl:template match="l-:range-end" mode="write-range-marker">
  <xsl:text></xsl:text>
  <xsl:value-of select="@name"/>
  <xsl:text></xsl:text>
</xsl:template>

<xsl:template match="l-:range-start[@name='soCalled'] |
  l-:range-end[@name='soCalled']" mode="write-range-marker">
  <xsl:text></xsl:text>
</xsl:template>

<xsl:template match="l-:range-start[@name='note'] |
  l-:range-end[@name='note']" mode="write-range-marker"/>

</xsl:stylesheet>

```

```

---
[TEI.2][text][body][q [who]Gandalf{}][page [n]51{}][p]A mortal,
Frodo, who keeps one of the Great Rings, does not die, but he
does not grow or obtain more life, he merely continues, until at
last every minute is a weariness. {q}[p][page][body]{text}[TEI.2]

The ring has the power of suspending the rules of mortality and
prolonging life. That is, it [emph]preserves{emph} (though not in
the same way as the Elf-rings whose science it borrows). In this
way, like any technology, it promises control over time and
space.

---
[TEI.2][text][body][q [who]Gandalf{}][page [n]51{}][p]he merely
continues, until at last every minute is a weariness. And if he
often uses the Ring to make himself invisible, he [emph]
fades{emph}: he becomes in the end invisible permanently, and
walks in the twilight under the eye of the dark power that rules
the Rings. Yes, sooner or later—later, if he is strong or well-
meaning to begin with, but neither strength nor good purpose will
last—sooner or later the dark power will devour him. {q}[p][page]
{body}{text}[TEI.2]

From the very first, accounts of what happens with the Ring are
completely negative: the temptation it represents is described
more than it is demonstrated. The Ring is beautiful as an
aesthetic object, although simple; but its treacherous nature is
spelled out early.

---
[TEI.2][text][body][q [who]Gandalf{}][page [n]51{}][p]And if he
often uses the Ring to make himself invisible, he [emph]
fades{emph}: he becomes in the end invisible permanently, and
walks in the twilight under the eye of the dark power that rules
the Rings. {q}[p][page][body]{text}[TEI.2]

Another power the Ring has is to make its wearer invisible.

```

Interestingly, many or most of the Ring's powers only come into play when it is being worn (that is, being claimed). Invisibility amounts to the power to act without taking responsibility—without being seen or recognized. But the Ring's power is double-edged: while it allows its wearer to go unseen by those from whom he wishes to hide, it exposes him to the observation of the will or purpose identified with the Ring, that is the Dark Lord. (Note that the Dark Lord himself exists as little more than a disembodied principle of malice, and never takes much more tangible form in the novel, acting for the most part through his agents.) Thus the struggle between Frodo and the Ring is deeply personal, a struggle between Frodo and himself: the Ring is precisely a means by which Frodo can indulge his wish to hide, though at the price of being known by the malevolence expressed in that hiding, whether that be identified as Frodo's own more selfish motives, or a motive "built in" to the Ring by its creator. To wear the Ring is thus to take on a burden of guilt or knowledge of one's weakness and fallibility, as well as exposure and subservience to the greater power that made it and whose purpose is expressed in it: to claim the Ring is to embrace, to one's own ultimate destruction, a fallen condition.

```

---
[TEI.2][text][body][page [n]51{}][p][q [who]Gandalf{}]He thought
the ring was very beautiful, and very useful at need; and if
anything was wrong or queer, it was himself. He said that it was
[q]growing on his mind[q], and he was always worrying about it;
but he did not suspect that the ring itself was to blame. Though
he had found out that the thing needed looking after; it did not
seem always of the same size or weight; it shrank or expanded in
an odd way, and might suddenly slip off a finger where it had
been tight. {q}[p]{page}[body]{text}{TEI.2}

```

Bilbo evidencing the effect of mixed subjectivity. Interestingly, Gandalf draws attention to the Ring's independence and willfulness by noting how it grows and shrinks (a physical effect), leaving implicit its moral effects both of causing Bilbo to fret and of confusing Bilbo as to the source of his own feelings. Thus Gandalf is willing to attribute responsibility and self-interest to the Ring despite (and by way of observing) Bilbo's being deceived on that very point (about what's Bilbo's fault and what isn't).

```

---
[TEI.2][text][body][page [n]51{}][p][q [who]Gandalf{}]There is a
last test to make. But I no longer doubt my guess.{p}[page][page
[n]52{}][p]When did I first begin to guess?{q} he mused,
searching back in memory. {p}[page][body]{text}{TEI.2}

```

Gandalf's recollection leads to a narrative-within-the-narrative, which itself relates to earlier stories.

```

---
[TEI.2][text][body][page [n]52{}][p][q [who]Gandalf{}]When I at
last got the truth out of him, I saw at once that he had been
trying to put his claim to the ring beyond doubt. {q}[p][page]
[body]{text}{TEI.2}

```

Notice both Bilbo's proprietary attitude towards the ring, and his wish to keep it hidden.

```

---
[TEI.2][text][body][page [n]53{}][q [who]Gandalf{}][p]For he gave
it up in the end of his own accord: an important point. {p}[q]
[page][body]{text}{TEI.2}

```

Bilbo and Sam are the only characters in the novel to give up the Ring willingly; Gandalf and Galadriel (and somewhat less momentarily, Aragorn and Faramir) also refuse it when offered, or when they have an opportunity to take it (also not insignificant).

Notes

1. This is a single one of several research projects and initiatives devoted to overlapping hierarchies, which must be one of the healthiest "problems" facing markup today. See [Durusau/O'Donnell 2002-], [Cover 2003] for starters. Significant contributions to this discussion have been made by Patrick Durusau, C.M. Sperberg-McQueen, Claus Huitfeldt,

David Durand and others; several other papers on the topic also appear in the present conference proceedings.

2. As of October 2002, LMNL had a web site, which has disappeared and reappeared in the time since. (When <http://www.lmnl.org> does not work, <http://www.lmnl.net> might.) Although at the time of writing it is still due for maintenance, it is nonetheless worth checking for current status.
3. Alexander Czmil of the University of Cologne has represented LMNL in XML in the back end of a prototype annotations system for transcribing, translating and commenting on Ancient Egyptian papyrus fragments. See [Czmil 2004].
4. Milton's "Sing, Heavenly Muse" must be the "Hello World" of overlapping markup projects.
5. Another approach would certainly be to put the milestones themselves in a separate namespace. The tradeoffs here mostly have to do with validation requirements and tools support.
6. Straight-up LMNL does not provide for a simple way to order the starts or ends of coterminous ranges. So for example if two ranges start at the same point with respect to character data, LMNL does not assert whether one starts before the other. Until user requirements are better understood, however, a model closer to the marked-up text, "half-LMNL" preserves tag ordering just in case it may be called for. In other respects, however, this is essentially reified LMNL as proposed by Jeni Tennison [LMNL 2002].
7. It is worth noting that while much of this processing could be achieved with a single transformation, both maintenance and performance suggest we implement the conversion as a series of simple transformations. A particular advantage of breaking out the steps is that it makes it possible both to generalize this processing over the broadest range of inputs, and to modularize it so it can be easily enhanced for special cases.
8. Rhetorical impressions made by enjambment and end-stopping in this passage: first, a series of enjambed lines (6-13) serves to emphasize the Hebraic (as opposed to pagan and classical) identification of Milton's Muse; second, a series of end-stopped lines at the end of the passage (24-25) bring it to a climactic conclusion. Particular examples of enjambment (for example, "Soar / Above th' Aonian mount" also have their own local significance. Interestingly, either enjambment or end-stopping can be used to add (different kinds of) emphasis to line ends.
9. A note at a point, such as the footnotes provided for in the Extreme document model in which this paper is submitted, is easy to accommodate in XML; but an annotation of a span or range is confounded by XML's single-hierarchy world view: arbitrary ranges may overlap either or both element boundaries, and other ranges associated with other annotations. While a range can be marked up with milestone elements, processing it on the far end is difficult or impossible using current tools. Reduced to half-LMNL, it is all one.
10. Under the title *The Lord of the Rings*, this translation of the Red Book is copyrighted by Prof Tolkien and his descendants (1954, 1965, renewed 1995).
11. This is plain simple TEI except that the **note** and **q** elements are used in CLIX mode, not as normally.
12. The << and >> operators both take two nodes as arguments and return Boolean "true" or "false", depending on which argument node follows which in document order. This avoids a tortuous XPath 1.0 idiom for node comparison and allows a processor to optimize the operation.
13. A stylesheet that performs this operation over flattened CLIX input has now been prototyped by the author, and will be included in the stylesheet distribution (to be made available in the online conference proceedings).
14. Note that this example, as well as other code examples, has been cleaned up slightly for presentation; so if transcribed into a LMNL processor (even one constructed of an XML pipeline as demonstrated here) they will not, quite, make it "round-trip" (their whitespace should be normalized).

Bibliography

- [Cover 2003] Cover, Robin. *Markup Languages and (Non-) Hierarchies*. At <http://xml.coverpages.org/hierarchies.html>. Page last modified: July 15, 2003.
- [Czmiel 2004] Czmiel, Alexander. *XML for Overlapping Structures (XfOS) using a non XML Data Model*. ALLC/ACH 2004. Gothenburg, Sweden.
- [DeRose 2004] DeRose, Steven J. *Markup Overlap: A Review and a Horse*. Extreme 2004. Montreal, Canada.
- [Durusau/O'Donnell2002-] Durusau, Patrick, and Matthew Brook O'Donnell. *Overlapping Hierarchies / Concurrent Markup*. Society for Biblical Literature web site: <http://www.sbl-site2.org/Overlap/>. As of April 25, 2004.
- [LMNL 2002] *Layered Markup and Annotation Language*. Web site at www.lmnl.net.
- [LMNL Extreme 2002] Tennison, Jeni, and Wendell Piez. *LMNL: the Layered Markup and Annotation Language*. Extreme Markup Languages Conference. Montreal, 2002.
- [TEI P4] Sperberg-McQueen, C.M., and Lou Burnard, eds. *TEI Guidelines for Electronic Text Encoding and Interchange (P4)*. 2001. On line at <http://etext.virginia.edu/teip4/>.
-

The Author

Wendell Piez

Mulberry Technologies

Wendell Piez is an XML Consultant at Mulberry Technologies, Inc. His work with markup languages stems from broader interests in rhetoric and media theory. When not writing XSLT stylesheets, he may be practicing Tai Chi “push hands” or taking digital photographs of historic Shepherdstown, West Virginia.

Extreme Markup Languages 2004®

Montréal, Québec, August 2-6, 2004

*This paper was formatted from XML source via XSL
by Mulberry Technologies, Inc.*