

# How To Write A Minimal L<sup>A</sup>T<sub>E</sub>X XML Binding

Hang Yuan, Jinbo Zhang, Michael Kohlhase  
Computer Science, Jacobs University Bremen

March 10, 2015

## 1 Introduction

L<sup>A</sup>T<sub>E</sub>X has been widely used as a document processor among scholars, especially when one needs to use large quantities of mathematical representations. L<sup>A</sup>T<sub>E</sub>X is also a good choice for those who are meticulous about typographical quality of documents.

As a page formatting tool, the primary output format of the L<sup>A</sup>T<sub>E</sub>X formatter is PDF; which – with fixed page formats and limited interaction features – is only partially suited for usage in the modern web. The DLMF (Digital Library of Mathematical Functions) developed L<sup>A</sup>T<sub>E</sub>X XML, a flexible, semantics-preserving L<sup>A</sup>T<sub>E</sub>X to XML converter to fix this.

However, for every L<sup>A</sup>T<sub>E</sub>X class and package used in a document L<sup>A</sup>T<sub>E</sub>X XML needs a **L<sup>A</sup>T<sub>E</sub>X XML binding** – a configuration file that specifies the XML counterpart of the L<sup>A</sup>T<sub>E</sub>X command sequences provided by the respective class or package.

Even though the L<sup>A</sup>T<sub>E</sub>X XML distribution provides bindings for the most commonly used classes and packages, the availability of bindings is still the most severe bottleneck for L<sup>A</sup>T<sub>E</sub>X XML. The L<sup>A</sup>T<sub>E</sub>X XML documentation [**LaTeXML:manual**] is mostly written for developers and quite impenetrable for beginners.

To encourage binding development this how-to tutorial goes through the steps and pitfalls of creating a L<sup>A</sup>T<sub>E</sub>X XML class binding from scratch. This tutorial does not cover advanced topics related to L<sup>A</sup>T<sub>E</sub>X XML, for which we refer to the L<sup>A</sup>T<sub>E</sub>X XML manual [**LaTeXML:manual**].

We have developed a minimal document class `mockDoc` as an example for this how-to and will go through it step-by-step. All necessary files (and the development version of this tutorial) are available from [**mockDoc:git**], but are also included in the appendix of this document for reference.

This how-to tutorial is structured as follows: Section 2 briefly reviews L<sup>A</sup>T<sub>E</sub>X XML workflows and the files involved.<sup>1</sup> Section 6 concludes the tutorial.

EdN:1

## 2 Using LaTeXML

In this tutorial we assume a working installation of L<sup>A</sup>T<sub>E</sub>X XML – see [**LaTeXML:get**] for instructions – on a Unix-like system (Linux, Mac OS, etc.).

Given that, we use the command

```
latexmlc mockDoc.tex --format=XML --destination=mockDoc.xml --log=mockDoc.xml.log
latexmlpost --stylesheet=mockDoc.xsl -dest=mockDoc.html mockDoc.xml
```

for converting `mockDoc.tex` into `mockDoc.xml` and `mockDoc.xml` into `mockDoc.html` <sup>2</sup>

EdN:2

The conversion from L<sup>A</sup>T<sub>E</sub>X to XML is processed by L<sup>A</sup>T<sub>E</sub>X XML. Basically L<sup>A</sup>T<sub>E</sub>X XML maps the L<sup>A</sup>T<sub>E</sub>X markups to the XML markups, more specifically: macros, primitives and constructors. The post-processing mechanism such as conversion to HTML and XHTML is done by passing the documents through the post-processing filter modules.

<sup>1</sup>EdNOTE: MK: continue, when the struture is fixed

<sup>2</sup>EdNOTE: Actually, we should also have a HTML step, so that this shows the stylesheet as well. \*Already added a post processing chapter in the end.

## 2.1 Things We Need

**source** Here we use `mockDoc.tex` as a minimal example<sup>3</sup> see appendix ??<sup>4</sup>

EdN:3

(**L<sup>A</sup>T<sub>E</sub>X** class) we provide a L<sup>A</sup>T<sub>E</sub>X class `mockDoc.cls` for reference; sometimes it is useful to generate PDF for proofreading. Also, the normal situation in developing L<sup>A</sup>T<sub>E</sub>XML bindings is that the class/package pre-exists. This file won't be illustrated in this tutorial.

EdN:4

**L<sup>A</sup>T<sub>E</sub>XML binding** the core issue of this tutorial. We use `mockDoc.cls.ltxml` – Section ?? for a step-by-step explanation and appendix A.2 for the end result.

**RelaxNG schema** L<sup>A</sup>T<sub>E</sub>XML needs a RelaxNG schema to infer the output structure. We supply it in compact form (`mockDoc.rnc`; see appendix A.3, which can be converted to the XML form L<sup>A</sup>T<sub>E</sub>XML needs (`mockDoc.rng`) via `trang.jar`. The reason for writing `mockDoc.rnc` instead of `mockDoc.rng` is that, `mockDoc.rnc` is much shorter and easier to maintain.

After you have finished writing all the documents above, run the command mentioned before, and then you should be able to see the converted XML file for `mockDoc.tex`. In the following chapters we will explain how to construct `mockDoc.rnc` and `mockDoc.cls.ltxml`

These workflows can be automated via a Unix `makefile` (see appendix B), which re-generates everything when source files have changed. Then only need to issue the command:

```
make
```

## 3 The mockDoc Format

### 3.1 A minimal Document Format

Actually our `mockDocformat` is probably the smallest one in the world, it is only intended for this tutorial.

The L<sup>A</sup>T<sub>E</sub>X class only provides one environment: `document` and four macros: `\section`, `\subsection`, `\paragraph`, and `\newline`. A minimal example would be

Listing 1: A Minimal L<sup>A</sup>T<sub>E</sub>X Document

```
\documentclass{mockDoc}
\begin{document}
  \section{A brief introduction about Shelley}
    Percy Bysshe Shelley (4 August 1792 -- 8 July 1822) was one of the
    major English Romantic poets, and is regarded by some critics
    as amongst the finest lyric poets in the English language.
  \section{Ode to the West Wind (partial)}
    \subsection{I}
      \paragraph{1.}
        O wild West Wind, thou breath of Autumn's being, \
        \newline
        Thou, from whose unseen presence the leaves dead \
        \newline
        Are driven, like ghosts from an enchanter fleeing,
      \paragraph{2.}
        Yellow, and black, and pale, and hectic red, \
        \newline
        Pestilence-stricken multitudes: O thou, \newline
        Who chariotest to their dark wintry bed
      \paragraph{3.}
        The winged seeds, where they lie cold and low, \
        \newline
        Each like a corpse within its grave, until \newline
        Thine azure sister of the Spring shall blow
```

<sup>3</sup>EdNOTE: MK: make a minimal one, use that here

<sup>4</sup>EdNOTE: make other references

```

\paragraph{4.}
Her clarion o'er the dreaming earth, and fill \
  \newline (Driving sweet buds like flocks to feed
in air) \newline With living hues and odours
plain and hill:
\paragraph{5.}
Wild Spirit, which art moving everywhere; \newline
Destroyer and Preserver; hear, O hear!
\subsection{II}
\paragraph{1.}
Thou on whose stream, 'mid the steep sky's
commotion, \newline
Loose clouds like Earth's decaying leaves are shed,
\newline
Shook from the tangled boughs of Heaven and Ocean,
\paragraph{2.}
Angels of rain and lightning: there are spread \
\newline
On the blue surface of thine airy surge, \newline
Like the bright hair uplifted from the head
\paragraph{3.}
Of some fierce Maenad, even from the dim verge \
\newline
Of the horizon to the zenith's height, \newline
The locks of the approaching storm. Thou dirge
\paragraph{4.}
Of the dying year, to which this closing night \
\newline
Will be the dome of a vast sepulchre \newline
Vaulted with all thy congregated might
\paragraph{5.}
Of vapours, from whose solid atmosphere \newline
Black rain, and fire, and hail will burst: O hear!
\end{document}
%%% Local Variables:
%%% mode: latex
%%% TeX-master: t
%%% End:

```

We want to use this document class for generating XML documents, which use the five elements `document`, `\section`, `\subsection`, `\paragraph`, and `\newline`. The XML document corresponding to the  $\text{\LaTeX}$  document from Listing 1 is

Listing 2: A Minimal  $\text{\LaTeX}$  Document

```

<?xml version="1.0" encoding="UTF-8"?>
<?latexml searchpaths="/Users/yuancalvin/xmlTrans/howTo"?>
<?latexml class="mockDoc"?>
<?latexml RelaxNGSchema="mockDoc.rng"?>
<mock:document xmlns:mock="https://kwarc.info/projects/mockDoc">
  <mock:section>
    <mock:title>A brief introduction about Shelley</mock:title>
    <mock:p>Percy Bysshe Shelley (4 August 1792 8 July 1822) was one of the major
      English Romantic poets, and is regarded by some critics as amongst the
      finest lyric poets in the English language.
    </mock:p>
  </mock:section>
  <mock:section>
    <mock:title>Ode to the West Wind (partial)</mock:title>
    <mock:subsection>
      <mock:title>I</mock:title>
      <mock:paragraph>
        <mock:title>1.</mock:title>
        <mock:p>
O wild West Wind, thou breath of Autumns being, <mock:break/>Thou, from whose
unseen presence the leaves dead <mock:break/>Are driven, like ghosts from an
enchante fleeing,
        </mock:p>
      </mock:paragraph>
    </mock:subsection>
  </mock:section>
</mock:document>

```

```

    </mock:paragraph>
    <mock:paragraph>
      <mock:title>2.</mock:title>
      <mock:p>
Yellow, and black, and pale, and hectic red,<mock:break/>Pestilence-stricken
      multitudes: O thou, <mock:break/>Who chariotest to their dark wintry bed
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>3.</mock:title>
      <mock:p>
The winged seeds, where they lie cold and low,<mock:break/>Each like a corpse
      within its grave, until <mock:break/>Thine azure sister of the Spring shall
      blow
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>4.</mock:title>
      <mock:p>
Her clarion oer the dreaming earth, and fill <mock:break/>(Driving sweet buds like
      flocks to feed in air) <mock:break/>With living hues and odours plain and hill:
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>5.</mock:title>
      <mock:p>
Wild Spirit, which art moving everywhere; <mock:break/>Destroyer and Preserver;
      hear, O hear!
</mock:p>
    </mock:paragraph>
    </mock:subsection>
    <mock:subsection>
      <mock:title>II</mock:title>
      <mock:paragraph>
        <mock:title>1.</mock:title>
        <mock:p>
<!-- ***** mockDoc.tex Line 25 ****>-->Thou on whose stream, mid the steep skys
      commotion, <mock:break/>Loose clouds like Earths decaying leaves are shed, <
      mock:break/>Shook from the tangled boughs of Heaven and Ocean,
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>2.</mock:title>
      <mock:p>
Angels of rain and lightning: there are spread <mock:break/>On the blue surface of
      thine airy surge, <mock:break/>Like the bright hair uplifted from the head
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>3.</mock:title>
      <mock:p>
Of some fierce Maenad, even from the dim verge <mock:break/>Of the horizon to the
      zeniths height, <mock:break/>The locks of the approaching storm. Thou dirge
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>4.</mock:title>
      <mock:p>
Of the dying year, to which this closing night <mock:break/>Will be the dome of a
      vast sepulchre <mock:break/>Vaulted with all thy congregated might
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>5.</mock:title>
      <mock:p>
Of vapours, from whose solid atmosphere <mock:break/>Black rain, and fire, and hail
      will burst: O hear!

```

```

</mock:p>
  </mock:paragraph>
</mock:subsection>
</mock:section>
</mock:document>

```

5

EdN:5

Note the typical format-specific differences between the presentation-oriented L<sup>A</sup>T<sub>E</sub>X and more content-oriented XML formats. The sectioning is conveyed by macros in L<sup>A</sup>T<sub>E</sub>X – only giving the start cues (here the numbered section headings) – whereas the XML has start and end tags<sup>6</sup>

EdN:6

After you link `mockDoc.tex` and `mockDoc.cls.ltxml` by changing your document class in your `mockDoc.tex` into your L<sup>A</sup>T<sub>E</sub>XML binding name, in our case, “doc”. Put `mockDoc.cls.ltxml` and `mockDoc.tex` in the same folder, L<sup>A</sup>T<sub>E</sub>XML will load your binding file automatically, when it tries to do the conversion.

BOP:7

EOP:7

## 3.2 The RelaxNG Schema

Schema is a crucial document that decides how `mockDoc.xml` is constructed. When you are creating your own schema<sup>1</sup>, one good approach to test this is to create your expected `mockDoc\_sample.xml` by hand, according to your `mockDoc.tex`, then compare `mockDoc\_sample.xml` with the generated `mockDoc.xml`. You can easily accomplish this by using *emacs nxml mode*<sup>2</sup>, in which you have the freedom to write your expected `mockDoc.xml`, while validating your `mockDoc.xml` at the same time. If validation fails, you can see the error message instantly, such that you can debug your `mockDoc.xml` or schema accordingly.<sup>8</sup>

EdN:8

In our `mockDoc.rnc`:

```

document = element document {p, section*}
section = element section {title,(p |subsection)*}

```

you can easily see that, under a `document`, there can be either `p` or `section`, and under a `section` there can be a `title` followed by `p` or a `title` followed by a `subsection`. This is because in the first section in `mockDoc.tex`:

```

\section{A brief introduction about Shelley}
  Percy Bysshe Shelley (4 August 1792 -- 8 July 1822)...

```

there is no `subsection` but texts directly. But in the other `sections`, there are `subsections`. In your schema you need to consider all kinds of possible hierarchy of your elements.

## 4 How to Create a L<sup>A</sup>T<sub>E</sub>XML Binding

We now come to the central part of our tutorial: writing the L<sup>A</sup>T<sub>E</sub>XML binding itself. Generally, a L<sup>A</sup>T<sub>E</sub>XML binding file is a Perl module – and therefore underlies Perl syntax, but special high-level commands simplify expressing the L<sup>A</sup>T<sub>E</sub>X-to-XML relation.

### 4.1 Basic structure

Since L<sup>A</sup>T<sub>E</sub>X binding is a perl module, we need to initialize a binding file by adding the followings in the beginning of `mockDoc.cls.ltxml`:

<sup>5</sup>EdNOTE: MK: actually, we should make an idealized minimal XML example by remove all the XML-isms from the generated one and show it here.

<sup>6</sup>EdNOTE: MK: are there more? Here would be the place to discuss them.

<sup>7</sup>Old PART: MK: I do not understand this, what do you want to say here?

<sup>1</sup>Before you write your expected xml and schema, having a look at the links below can be beneficial: <http://relaxng.org/compact-tutorial-20030326.html>; <http://www.w3schools.com/xml/>.

<sup>2</sup>Here is a tutorial about Emacs nxml mode: <http://www.emacswiki.org/emacs/NxmlMode>

<sup>8</sup>EdNOTE: MK: convert all footnotes to citations!

```
package LaTeXML::Package::Pool;
use strict;
use LaTeXML::Package;
use warnings;
```

At the end of `mockDoc.cls.ltxml`, don't forget to include

```
1;
```

to make sure that perl works properly.

## 4.2 Configure namespace

With:

```
RegisterNamespace('mock'=>"https://kwarc.info/projects/mockDoc");
RelaxNGSchema("mockDoc.rng", 'mock'=>"https://kwarc.info/projects/mockDoc");
```

We declared the namespace associated the prefix `mock` with the namespace.<sup>9</sup>

EdN:9

## 4.3 Linebreaks

The next task is to teach  $\text{\LaTeX ML}$  new commands used in `mockDoc.tex`. Here is an example:

```
DefConstructor('\newline', "<mock:break/>");
```

This line defines how  $\text{\LaTeX ML}$  interprets `\newline`, as you see,  $\text{\LaTeX ML}$  will translate `\newline` to `<mock:break/>` in `mockDoc.xml`.

## 4.4 Sectioning

When dealing with section, things get a little tricky, with:

```
DefConstructor('\section{ }', "<mock:section><mock:title>#1</mock:title>");
```

we defined `\section`. But, think about the closing tags. In `mockDoc.tex`, we declared where the `\section` starts and where the next `\section` starts, nevertheless, we never wrote something like “Now close this section”. Here is why we need `mockDoc.rnc`. This schema file tells  $\text{\LaTeX ML}$  what the structure of our document, and with:

```
Tag('mock:section', autoClose=>1);
```

$\text{\LaTeX ML}$  will close the section tags (i.e, adding `</mock:section>`) whenever needed.

## 4.5 The Document Environment

You may think something like:

```
DefEnvironment('{document}', "<mock:document>#body</mock:document>");
```

is enough for defining document environment. You can try it, you will find that all spaces disappear. What we actually wrote in `mockDoc.cls.ltxml` is:

```
DefEnvironment('{document}', "<mock:document>#body</mock:document>", beforeDigest
=> sub { AssignValue(inPreamble => 0); });
```

This code can prevent the error mentioned before, however, the mechanism of the `beforeDigest` part is out of our discussion in this tutorial.

For an environment, we don't need care about auto-closing, since an environment is always like

```
\begin{*environment-name*}
content...
\end{*environment-name*}
```

where `\end{*environment-name*}` will indicate where to close the tags.

<sup>9</sup>EdNOTE: MK: we have to explain the why more, not just what to do; that is generally the case.

## 4.6 Auto-opening for Paragraphs

Since we also want to write some texts directly under `document`, without any `section`. At this circumstance, we need auto-open for `p`:

```
Tag('mock:p', autoOpen=>1);
```

which will surround such texts.

You now have a complete set of tiles to generate your XML file. Simply by using `makefile`, you should be able to see the generated `mockDoc.xml` in your current directory. It should be something similar to your expected `mockDoc\_sample.xml`.

## 5 Postprocessing for Web Workflow

After we obtain `mockDoc.xml`, we can further utilize the power of  $\text{\LaTeX}XML$  to convert it into some other useful formats such as HTML, HTML5 and XHTML.  $\text{\LaTeX}XML$  by default provides us with stylesheets for this conversion, however we are given freedom to customize this process by creating our own XSL and CSS stylesheets.

### 5.1 XSL Stylesheet

Similar to how we create RelaxNG schema, in `mockDoc.xsl`

```
<xsl:template match="/">
  <xsl:apply-templates/>
</xsl:template>
```

we associate the template with the root of XML and then within the main structure, we tells  $\text{\LaTeX}XML$  to use other template when needed.

```
<xsl:template match="mock:section">
  <section>
    <span style="font-size:200%">
      <xsl:apply-templates select="mock:title"/>
    </span>
    <xsl:apply-templates select="mock:p"/>
    <xsl:apply-templates select="mock:subsection"/>
  </section>
</xsl:template>
```

In case of the section template, we let  $\text{\LaTeX}XML$  generate a section in HTML and then apply the templates for the macros that appear in `mock:section` from our `mockDoc.xml` file. Our `mockDoc.xsl` should be consistent with our `mockDoc.rnc` as they both define the structure of our documents, as you can see from the section definition in our `mockDoc.rnc`, they both contain: `title`, `p` and `subsection` elements.

```
section = element section {title,(p |subsection)*}
```

## 6 Conclusion

<sup>10</sup> For a web workflow, where the ultimate goal is to generate HTML5, writing a document class from scratch may not be the most common workflow, since the majority of document classes in  $\text{\LaTeX}$  are derived in some way from `article.cls` and therefore the  $\text{\LaTeX}XML$  bindings can inherit the from `article.cls.ltxml`, but in some cases we want to use  $\text{\LaTeX}XML$  to generate other XML-based format. There we need the techniques in this tutorial. Examples are generating OMDoc from  $\text{sTeX}$ <sup>12</sup>

EdN:10  
BNP:11

EdN:12  
ENP:11

---

<sup>10</sup>EdNOTE: MK: say something here

<sup>11</sup>NEW PART: MK: do not forget to discuss this somewhere

<sup>12</sup>EdNOTE: MK: cite them from `kwarc.bib`, are there others?

# A Appendix

## A.1 The mockDoc Class

```
% File: doc.cls
% Author: Jinbo Zhang
% Date: 3 Feb, 2015

\NeedsTeXFormat{LaTeX2e}
\ProvidesClass{mockDoc}
\RequirePackage{ifthen}

\renewcommand\normalsize{\fontsize{10pt}{12pt}\selectfont}
\setlength{\textwidth}{6.5in}
\setlength{\textheight}{8in}

\newcommand\large{\@setfontsize\large\@xiipt{14}}
\newcommand\Large{\@setfontsize\Large\@xivpt{18}}

% define \paragraph
\newcommand{\paragraph}[1]{
  \newline\newline
  \bfseries #1
  \normalfont
}

% define \section
\newcounter{SectionCount}
\newcommand{\section}[1]{
  \ifthenelse{\value{SectionCount}=0}{\newline\newline\newline}
  \Large
  \stepcounter{SectionCount}
  \noindent\bfseries\arabic{SectionCount}\hspace{4mm} #1
  \normalfont
  \newline\newline
}

% define \subsection
\newcounter{SubCount}[SectionCount]
\newcommand{\subsection}[1]{
  \ifthenelse{\value{SubCount}=0}{\newline\newline}
  \large
  \stepcounter{SubCount}
  \bfseries\arabic{SectionCount}.\arabic{SubCount}\hspace{3mm} #1
  \normalfont
}

\endinput
```

## A.2 The mockDoc Class Binding

```
package LaTeXML::Package::Pool;
use strict;
use LaTeXML::Package;
use warnings;

#Document Structure
RegisterNamespace('mock'=>"https://kwarc.info/projects/mockDoc");
RelaxNGSchema("mockDoc.rng", 'mock'=>"https://kwarc.info/projects/mockDoc");

#-----
DefEnvironment('{document}', "<mock:document>#body</mock:document>", beforeDigest
=> sub { AssignValue(inPreamble => 0); });
DefConstructor('{\section{}}', "<mock:section><mock:title>#1</mock:title>");
DefConstructor('{\subsection{}}', "<mock:subsection><mock:title>#1</mock:title>");
```



```

DefConstructor('\paragraph{ }', "<mock:paragraph><mock:title>#1</mock:title><mock:p>
");
DefConstructor('\newline', "<mock:break/>");

#autoClose
Tag('mock:paragraph', autoClose=>1);
Tag('mock:section', autoClose=>1);
Tag('mock:subsection', autoClose=>1);
Tag('mock:p', autoClose=>1);
Tag('mock:p', autoOpen=>1);

#make sure Perl work
1;

```

### A.3 mockDoc RelaxNG schema

```

default namespace md = "https://kwarc.info/projects/mockDoc"

start = document
document = element document {p, section*}
section = element section {title,(p |subsection)*}
subsection = element subsection {title,paragraph*}
paragraph = element paragraph { title, p }
title = element title { text }
p = element p { (text|element break { empty })*}

```

### A.4 Generated XML

```

<?xml version="1.0" encoding="UTF-8"?>
<?latexml searchpaths="/Users/yuancalvin/xmlTrans/howTo"?>
<?latexml class="mockDoc"?>
<?latexml RelaxNGSchema="mockDoc.rng"?>
<mock:document xmlns:mock="https://kwarc.info/projects/mockDoc">
  <mock:section>
    <mock:title>A brief introduction about Shelley</mock:title>
    <mock:p>Percy Bysshe Shelley (4 August 1792 8 July 1822) was one of the major
      English Romantic poets, and is regarded by some critics as amongst the
      finest lyric poets in the English language.
    </mock:p>
  </mock:section>
  <mock:section>
    <mock:title>Ode to the West Wind (partial)</mock:title>
    <mock:subsection>
      <mock:title>I</mock:title>
      <mock:paragraph>
        <mock:title>1.</mock:title>
        <mock:p>
0 wild West Wind, thou breath of Autumns being, <mock:break/>Thou, from whose
      unseen presence the leaves dead <mock:break/>Are driven, like ghosts from an
      enchante fleeing,
        </mock:p>
      </mock:paragraph>
      <mock:paragraph>
        <mock:title>2.</mock:title>
        <mock:p>
Yellow, and black, and pale, and hectic red,<mock:break/>Pestilence-stricken
      multitudes: O thou, <mock:break/>Who chariotest to their dark wintry bed
        </mock:p>
      </mock:paragraph>
      <mock:paragraph>
        <mock:title>3.</mock:title>
        <mock:p>
The winged seeds, where they lie cold and low,<mock:break/>Each like a corpse
      within its grave, until <mock:break/>Thine azure sister of the Spring shall
      blow
        </mock:p>
      </mock:paragraph>
    </mock:subsection>
  </mock:section>
</mock:document>

```

```

    <mock:paragraph>
      <mock:title>4.</mock:title>
      <mock:p>
Her clarion o'er the dreaming earth, and fill <mock:break/>(Driving sweet buds like
flocks to feed in air) <mock:break/>With living hues and odours plain and hill:
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>5.</mock:title>
      <mock:p>
Wild Spirit, which art moving everywhere; <mock:break/>Destroyer and Preserver;
hear, O hear!
</mock:p>
    </mock:paragraph>
  </mock:subsection>
  <mock:subsection>
    <mock:title>II</mock:title>
    <mock:paragraph>
      <mock:title>1.</mock:title>
      <mock:p>
<!-- ***** mockDoc.tex Line 25 ***** -->Thou on whose stream, mid the steep skys
commotion, <mock:break/>Loose clouds like Earths decaying leaves are shed, <
mock:break/>Shook from the tangled boughs of Heaven and Ocean,
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>2.</mock:title>
      <mock:p>
Angels of rain and lightning: there are spread <mock:break/>On the blue surface of
thine airy surge, <mock:break/>Like the bright hair uplifted from the head
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>3.</mock:title>
      <mock:p>
Of some fierce Maenad, even from the dim verge <mock:break/>Of the horizon to the
zeniths height, <mock:break/>The locks of the approaching storm. Thou dirge
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>4.</mock:title>
      <mock:p>
Of the dying year, to which this closing night <mock:break/>Will be the dome of a
vast sepulchre <mock:break/>Vaulted with all thy congregated might
</mock:p>
    </mock:paragraph>
    <mock:paragraph>
      <mock:title>5.</mock:title>
      <mock:p>
Of vapours, from whose solid atmosphere <mock:break/>Black rain, and fire, and hail
will burst: O hear!
</mock:p>
    </mock:paragraph>
  </mock:subsection>
</mock:section>
</mock:document>

```

## A.5 XSL stylesheet

```

<?xml version="1.0" encoding="utf-8"?>
<!-- customized for mockDoc conversion -->

<!-- style sheet declration -->
<xsl:stylesheet version="1.0"
xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
xmlns:mock = "https://kwarc.info/projects/mockDoc"
exclude-result-prefixes="mock">

```

```

<!-- associate the template with the root of xml -->
<xsl:template match="/">
  <html>
    <head>
      <xsl:comment>automatically generated, handle with care</xsl:comment>
    </head>
    <xsl:apply-templates/>
  </html>
</xsl:template>

<xsl:template match="mock:document">
  <body>
    <xsl:apply-templates/>
  </body>
</xsl:template>

<xsl:template match="mock:section">
  <h1>
    <xsl:value-of select="mock:title" />
  </h1>
  <p>
    <xsl:value-of select="mock:p" />
  </p>
  <xsl:for-each select="mock:subsection">
    <h2>
      <xsl:value-of select="mock:title" />
    </h2>
    <xsl:for-each select="mock:paragraph">
      <h3>
        <xsl:value-of select="mock:title" />
      </h3>
      <p>
        <xsl:value-of select="mock:p" />
      </p>
    </xsl:for-each>
  </xsl:for-each>
</xsl:stylesheet>

</xsl:stylesheet>

```

## A.6 Generated HTML

```

<html><body>
<head><title>MockDoc HTML SAMPLE Output</title></head>
<section><span style="font-size:200%">
  Title:
  A brief introduction about Shelley<br></span><p>Percy Bysshe Shelley (4
    August 1792 &ndash; 8 July 1822) was one of the major English Romantic
    poets, and is regarded by some critics as amongst the finest lyric
    poets in the English language.
</p></section><section><span style="font-size:200%">
  Title:
  Ode to the West Wind (partial)<br></span><subsection><span style="font-size
    :150%">
  Title:
  I<br></span><paragraph><span style="font-size:120%">
  Title:
  1.<br></span><p>
0 wild West Wind, thou breath of Autumn&rsquo;s being, Thou, from whose unseen
  presence the leaves dead Are driven, like ghosts from an enchanter fleeing,
</p></paragraph><paragraph><span style="font-size:120%">
  Title:
  2.<br></span><p>
Yellow, and black, and pale, and hectic red, Pestilence-stricken multitudes: O thou,
  Who chariotest to their dark wintry bed

```

```

</p></paragraph><paragraph><span style="font-size:120%">
  Title:
  3.<br></span><p>
The winged seeds, where they lie cold and low,Each like a corpse within its grave,
  until Thine azure sister of the Spring shall blow
</p></paragraph><paragraph><span style="font-size:120%">
  Title:
  4.<br></span><p>
Her clarion o&rsquo;er the dreaming earth, and fill (Driving sweet buds like flocks
  to feed in air) With living hues and odours plain and hill:
</p></paragraph><paragraph><span style="font-size:120%">
  Title:
  5.<br></span><p>
Wild Spirit, which art moving everywhere; Destroyer and Preserver; hear, O hear!
</p></paragraph></subsection><subsection><span style="font-size:150%">
  Title:
  II<br></span><paragraph><span style="font-size:120%">
  Title:
  1.<br></span><p>Thou on whose stream, &rsquo;mid the steep sky&rsquo;s
    commotion, Loose clouds like Earth&rsquo;s decaying leaves are shed,
    Shook from the tangled boughs of Heaven and Ocean,
</p></paragraph><paragraph><span style="font-size:120%">
  Title:
  2.<br></span><p>
Angels of rain and lightning: there are spread On the blue surface of thine airy
  surge, Like the bright hair uplifted from the head
</p></paragraph><paragraph><span style="font-size:120%">
  Title:
  3.<br></span><p>
Of some fierce Maenad, even from the dim verge Of the horizon to the zenith&rsquo;s
  height, The locks of the approaching storm. Thou dirge
</p></paragraph><paragraph><span style="font-size:120%">
  Title:
  4.<br></span><p>
Of the dying year, to which this closing night Will be the dome of a vast sepulchre
  Vaulted with all thy congregated might
</p></paragraph><paragraph><span style="font-size:120%">
  Title:
  5.<br></span><p>
Of vapours, from whose solid atmosphere Black rain, and fire, and hail will burst:
  O hear!
</p></paragraph></subsection></section>
</body></html>

```

## B A Makefile for Automation

```

#makefile for using latexml and pdflatex to generate *.pdf and *.xml
#declaration of variables
#set .tex as source. In our case only mockDoc is available
#name .xml and .pdf based on .tex
SRC = $(shell ls *.tex)
XML = $(SRC:%.tex=%.xml)
PDF = $(SRC:%.tex=%.pdf)

all: $(XML) $(PDF)

mockDoc.rnc: mockDoc.rnc
    java -jar trang.jar -I rnc -O rng mockDoc.rnc mockDoc.rnc

#the codes below follow the usage of variables mentioned above
#$$: object filename. $<:source file name
$(XML): %.xml: %.tex mockDoc.rnc mockDoc.cls.ltxml
    latexmlc $< --format=XML --destination=$$ --log=$$.log

$(PDF): %.pdf: %.tex mockDoc.cls
    xelatex $<

```