tccpage.py: Dynamic generation of TCC-style web pages

John W. Shipman
2009-10-10 14:33

Table of Contents

1. Introduction ............................................................................................................................ 1
   1.1. How to get this publication ............................................................................................. 2
2. Overview ................................................................................................................................ 2
3. The XHTML to be generated ..................................................................................................... 2
4. Code prologue ......................................................................................................................... 3
5. Manifest constants ................................................................................................................... 4
   5.1. CSS_URL: Stylesheet location ......................................................................................... 4
   5.2. TOP_NAV_SEP: Separator string for the top navigational bar .................................. 4
   5.3. TCC_MAIN_URL: TCC mainpage URL ........................................................ .................. 4
   5.4. TCC_LOGO_URL: TCC logo URL .................................................................................. 4
   5.5. BOT_NAV_SEP: Separator for page-bottom links ...................................................... 4
6. class TCCPage: The page object interface ......................................................................... 5
   6.1. TCCPage.__init__(): Constructor ................................................................................ 6
   6.2. TCCPage.__createHead(): Set up page heading .......................................................... 7
   6.3. TCCPage.__createBody(): Set up page body ............................................................... 7
   6.4. TCCPage.__topNav(): Set up top nav bar ................................................................... 8
   6.5. TCCPage.__topNavItem(): Add one top navigational item ........................................ 10
   6.6. TCCPage.__titleBlock(): Set up page title block ....................................................... 10
   6.7. TCCPage.__botNav(): Set up page-bottom navigational links ..................................... 11
   6.8. TCCPage.__botNavItem(): Generate page-bottom navigational link ........................ 12
   6.9. TCCPage.__colophon(): Add colophon section ......................................................... 13
   6.10. TCCPage.write(): Output the finished page ............................................................... 14
7. class NavLink: Describes one navigational feature .......................................................... 15

1. Introduction

The PyStyler application is used to maintain the bulk of the pages in the Tech Computer Center¹ page structure and the TCC Help System². This application is rather antiquated, the design substantially stable since 1996; see Building informational webs with PyStyler³ (PDF format) for user-level documentation.

One major limitation of this system is that it treats its web pages as static structure, to be maintained with an ordinary text editor.

¹ http://www.nmt.edu/tcc
² http://www.nmt.edu/tcc/help/
³ http://www.nmt.edu/tcc/projects/pystyler/
A number of TCC scripts generate Web pages outside the PyStyler system. Some scripts generate one or more static pages; other scripts are CGI scripts that generate pages dynamically in response to user requests. Ideally, PyStyler should be rewritten so that its content template can be applied easily to dynamically generate pages as well as to static pages inside a PyStyler document tree. This is a fairly sizeable project, however.

Here, we describe a Python module named \texttt{tccpage.py} that is intended for use by Python scripts that generate web pages. Its purpose is to mimic the appearance of pages generated by PyStyler so that they blend visually with static pages.

\begin{warning}
Here's why this is a short-term solution. The PyStyler system uses a template file named Template to describe the overall layout of a TCC page. The templates used in the TCC web and TCC help web haven't changed much lately, so they're the basis of the style imitated by the \texttt{tccpage.py} module. But if in the future we want to modify these templates, we'll have to modify the logic in the \texttt{tccpage.py} module in parallel. The best long-term solution is to do a whole new version of PyStyler that has a single source for template information that can drive both static and dynamic page generation.
\end{warning}

1.1. How to get this publication

This publication is available in Web form\footnote{http://www.nmt.edu/tcc/projects/tccpage/} and also as a PDF document\footnote{http://www.nmt.edu/tcc/projects/tccpage/tccpage.pdf}. Please forward any comments to \texttt{tcc-doc@nmt.edu}.

2. Overview

This module uses the techniques of XML generation described in \textit{Python and the XML Document Object Model (DOM)}\footnote{http://www.nmt.edu/tcc/help/pyxml/}.

Here's the general procedure for generating a TCC-like web page:

1. Import the \texttt{tccpage.py} module as:

\begin{verbatim}
import tccpage as tp
\end{verbatim}

2. Instantiate a \texttt{TCCPage} object and supply it with the desired navigational links. For the constructor and other interfaces to the \texttt{tccpage.py} class, see Section 6, \textit{class TCCPage: The page object interface} (p. 5).

3. If the title is not a simple string—for example, if it contains markup such as \texttt{filename} or \texttt{userinput} tags—the content can be added to the \texttt{.headTitle} and/or \texttt{.bodyTitle} elements.

4. Add the page's content to the \texttt{.content} attribute.

5. Use the \texttt{.write()} method to send the finished page to its destination.

3. The XHTML to be generated

Here is a general outline of the page to be generated.

\footnote{http://www.nmt.edu/tcc/projects/tccpage/} \footnote{http://www.nmt.edu/tcc/projects/tccpage/tccpage.pdf} \footnote{http://www.nmt.edu/tcc/help/pyxml/}
1. A **DOCTYPE** identifier for XHTML 1.0. Although the existing *PyStyler* web uses HTML, we want to move toward XHTML, and there’s no reason not to use it for generated pages.

2. A **head** element containing:
   a. The page’s **title** element.
   b. A link to the standard TCC stylesheet at `/tcc/style.css` on the server.

3. The **top navigational bar**, a set of **top navigational links**. This must fit on a single line, with the links separated by the string " / ".
   Each link has a **short name** consisting of a generic word or phrase such as “Next” or “Site map”. The word or phrase always appears, even if it is not a link. Thus the layout of the top nav bar is static over the whole site.

4. The formal page header, formatted as a one-row, two-column table with the page title on the left and the TCC logo graphic on the right.

5. The page body, followed by a horizontal rule (**hr** element).

6. A set of **bottom navigational links**. Each appears on a separate line, and the entire line is omitted if that link points nowhere.
   Each bottom nav link consists of three parts:
   - A boldfaced generic link type, which will be the same as that link in the top nav link, e.g., “Next”, followed by a colon and space.
   - One or more destination links, each of which uses the page title (or a close approximation) as its link text. If there are multiple links, they are separated by semicolons.

7. Another horizontal rule, followed by the **colophon**:
   - The “author credit”, which for dynamic pages will credit not a human author, but the script that wrote the page.
   - The date and time when the page was generated.
   - The URL of the page, if known.

### 4. Code prologue

The implementation of `tccpage.py` is included here in “literate programming” style. For information on the tool used for code generation, see *A source extractor for lightweight literate programming*.

The actual code starts with the conventional Python documentation string:

```python
"""tccpage.py: For dynamic generation of web pages in standard TCC style.
For documentation, see:
http://www.nmt.edu/tcc/projects/tccpage/
"""
```

Next, some imports. We need the `sys` module so we can use `sys.stdout` as the default output file. We need the `time` module for timestamping the page.

---

# Imports
import sys
import time

Also, we’ll need to import the xmlcreate.py module discussed in *Python and the XML Document Object Model (DOM)*.

import xmlcreate as xc

## 5. Manifest constants

Next we set up some manifest constants, grouped at the top of the script for easy maintenance.

### 5.1. CSS_URL: Stylesheet location

Gives the URL of the stylesheet for TCC web pages.

```python
CSS_URL = "http://www.nmt.edu/tcc/style.css"
```

### 5.2. TOP_NAV_SEP: Separator string for the top navigational bar

This is the string that separates elements of the page-top navigational bar.

```python
TOP_NAV_SEP = " / "
```

### 5.3. TCC_MAIN_URL: TCC mainpage URL

This is the URL that people reach if they click on the TCC logo.

```python
TCC_MAIN_URL = "http://www.nmt.edu/tcc/"
```

### 5.4. TCC_LOGO_URL: TCC logo URL

This is the URL of the official TCC logo.

```python
TCC_LOGO_URL = "http://www.nmt.edu/tcc/images/logo.png"
```

### 5.5. BOT_NAV_SEP: Separator for page-bottom links

When a particular kind of page-bottom link has multiple destinations (e.g., a “See also” link), this string is placed between the links.

---

8 [http://www.nmt.edu/tcc/help/pyxml/]
6. class TCCPage: The page object interface

Here’s the interface to this class:

```python
BOT_NAV_SEP = "; "

```# Functions and classes
#--------------------------------------------------------
# - - - - - c l a s s   T C C P a g e   - - - - -

class TCCPage:
    """Represents one page in the TCC style.

Exports:
    TCCPage ( title=None, navList=None, author=None, url=None ):
        [ (title is the page's title as a string, default empty) and
         (navList describes the page's standard set of
         navigational links as NavLink objects, default empty) and
         (author is the page's author credit string, default empty) and
         (url is the page's URL as a string, default empty) ->
         return a new, empty TCCPage object representing
         those values ]
    .headTitle:
        [ self's head's title as a DOM title element,
          initially the title argument, read/write ]
    .bodyTitle:
        [ self's body title as a DOM h1 element, initially the
          title argument, read/write ]
    .content:
        [ a DOM div element in self's body between the top
          and bottom nav links ]
    .address:
        [ self's address as a DOM address element, initially
          the author argument, read/write ]
    .url:
        [ as passed to constructor, read-only ]
    .write ( outFile=None ):
        [ outFile is a writeable file, defaulting to sys.stdout ->
          outFile +=: an XHTML representation of self ]

State/Invariants:
    .doc:  [ self's Document object ]
    .head:  [ self's head element ]
    .body:  [ self's body element ]
```

The `navList` argument describes a sequence of standard navigational links that appear on the bottom of each page, and optionally on the top as well. This argument is a list of zero or more NavLink objects,
each of which describes one navigational link and the place or places where that link goes. See Section 7, “class NavLink: Describes one navigational feature” (p. 15).

6.1. TCCPage.__init__(): Constructor

The constructor for this class creates the XHTML page as a DOM Document tree, fills in the top and bottom content, and creates self.content as an empty div element.

```python
# - - - T C C P a g e . __ i n i t __ - - -

def __init__ ( self, title=None, navList=None, author=None, url=None ):
    """Constructor for a TCCPage."""

We divide the construction of the page into four parts: creation of a Doctype object for XHTML; creation of the Document object; creation of its head element; and creation of its body element.

```
6.2. TCCPage.__createHead(): Set up page heading

This method sets up the page’s head element and all its content.

```python
# - - - T C C P a g e . _ _ c r e a t e H e a d - - -
def __createHead ( self):
    """Set up the page head element.
    [ self.doc := self.doc with a new head element added
      containing self's head content
    self.head := that element ]
    """
```

First we create the actual head element.

```python
#-- 1 --
# [ self.doc := self.doc with a new head element added
#    self.head := that element ]
self.head = xc.Element ( self.doc.root, "head" )
```

Next we set up the title element, if any, and set attribute self.headTitle to point to it. Unless the value of self.title is an empty string, we then add the title text.

```python
#-- 2 --
# [ self.head := self.head with a new title element
#    added containing self.title (if nonempty)
#    self.headTitle := that element ]
self.headTitle = xc.Element ( self.head, "title" )
if self.title:
    xc.Text ( self.headTitle, self.title )
```

The stylesheet link is pro forma, and links to the URL given by CSS_URL.

```python
#-- 3 --
# [ self.head := self.head with a stylesheet link added ]
xc.Element ( self.head, "link", rel="stylesheet",
            href=CSS_URL )
```

6.3. TCCPage.__createBody(): Set up page body

This method sets up the overall page body structure.

```python
# - - - T C C P a g e . _ _ c r e a t e B o d y - - -
def __createBody ( self):
    """Set up the page body.
    [ self.doc := self.doc with a new body element added
      containing self's body content
    self.body := that element
    self.bodyTitle := the h1 element containing self.title
    ]
    """
```
The principal divisions of the body content, in order, are:

1. Top navigational bar.
2. Title block with TCC logo.
3. The `div` element we export as attribute `self.content`. Here, the caller adds the actual page content.
4. Horizontal rule and bottom navigational links.
5. Another horizontal rule and the colophon section.

6.4. `TCCPage.__topNav()`: Set up top nav bar

The top navigational bar consists of a series of strings separated by " / ". The content comes from the elements of the `self.navList` argument, each of which is a `NavLink` object. Each string may be a link or not, depending on whether the corresponding object's `destList` has any members or not.
"""Build the row of navigational links across the page top.

[ (self.navList as invariant) and
  (self.body as invariant) ->
  self.body := self.body with the top nav bar added ]
"""

First we create a `div` element to hold the pieces of the nav bar.

```python
#-- 1 --
# [ self.body := self.body with a new div element added
# navBar := that element ]
navBar = xc.Element ( self.body, "div" )
navBar["class"] = "top-nav"
```

Now all that remains is to add one link for each element of `self.navList`. The only tricky part is getting the `TOP_NAV_SEP` separator between elements, but not initially or finally. I am indebted to Dr. Allan M. Stavely for the best (and easiest to verify) way to do this. We set a variable called `prefix` to the empty string initially. Each time through the loop, we add a copy of `prefix`, then we add the new content, and then we set `prefix` to `TOP_NAV_SEP`.

```python
#-- 2 --
prefix = ""

#-- 3 --
# [ navBar += (prefix) + (navBar with top elements of
#   self.navList added, separated by TOP_NAV_SEP
# prefix := TOP_NAV_SEP ]
for navItem in self.navList:
  #-- 3 body --
  # [ navItem is a NavLink object ->
  #   if navItem.noTop ->
  #     I
  #   else ->
  #     navBar += (prefix) + (navItem as a top nav link) ]
  if not navItem.noTop:
    #-- 3.1 --
    # [ if prefix is "" ->
    #   I
    # else ->
    #   navBar += prefix ]
    if prefix:
      xc.Text ( navBar, prefix )
    #-- 3.2 --
    # [ navBar += navItem as a top nav link ]
    self.__topNavItem ( navBar, navItem )
  #-- 3.3 --
  prefix = TOP_NAV_SEP
```

New Mexico Tech Computer Center  tccpage.py  9
6.5. TCCPage.__topNavItem(): Add one top navigational item

This method translates one NavLink object into a page-top navigational link. If the given navItem.destList has a URL in it, the link has link text navItem.shortName and points at that URL. (This is intended only for links such as “Next” that have a single URL. If there are multiple URLs, we use the first one.) If there are no URLs, the navItem.shortName attribute is inserted as plain text (not a link).

```python
# - - - T C C P a g e . _ _ t o p N a v I t e m - - -

def __topNavItem ( self, navBar, navItem ):
    """Add one item to the top navigational bar.

    [ (navBar is an xc.Element) and
      (navItem is a NavLink) ->
      navBar += navItem as a top nav link ]
    """

if navItem.destList:
    #-- 1.1 --
    # [ navBar += a link to navItem.destList[0] with text
    #   navItem.shortName ]
    title, url = navItem.destList[0]
    a = xc.Element ( navBar, "a", href=url )
    xc.Text ( a, navItem.shortName )
else:
    # [ navBar += navItem.shortName as text ]
    xc.Text ( navBar, navItem.shortName )
```

The test “if navItem.destList:” succeeds if that list is nonempty. It fails if navItem.destList is either None or an empty list.

6.6. TCCPage.__titleBlock(): Set up page title block

The purpose of this method is to output the small table (one row, two columns) that positions the main page title h1 element to the left of the TCC logo graphic.

Here's the XHTML we're generating:

```html
<table width="100\%">
<tr valign="top">
    <td align="left"><h1>title text</h1>
    <td align="right">
        <a href="http://www.nmt.edu/"> <img src="/tcc/images/logo.png" alt="Tech Computer Center logo"></a>
    </td>
</tr>
</table>
```

Here's the code.

```python
# - - - T C C P a g e . _ _ t i t l e B l o c k - - -
```
def __titleBlock ( self ):
    """Set up the main page title block.
    
    [ self.title as invariant ->
        self.body := self.body with a title block added
        self.bodyTitle := the h1 within that title block ]
    """

    #-- 1 --
    # [ self.body := self.body with a new table element added
    # table := that table element ]
    table = xc.Element ( self.body, "table", width="100%%" )

    #-- 2 --
    # [ table := table with a new tr element added
    # row := that tr element ]
    row = xc.Element ( table, "tr", valign="top" )

    #-- 3 --
    # [ row := row with a new cell added containing self.title
    #   inside an h1 element
    # self.bodyTitle := that h1 element ]
    td = xc.Element ( row, "td", align="left" )
    self.bodyTitle = xc.Element ( td, "h1" )
    if self.title:
        xc.Text ( self.bodyTitle, self.title )

    #-- 4 --
    # [ row := row with a new cell added containing the
    #   TCC logo as a link to TCC_MAIN_URL ]
    td = xc.Element ( row, "td", align="right" )
    a = xc.Element ( td, "a", href=TCC_MAIN_URL )
    img = xc.Element ( a, "img", alt="Tech Computer Center logo",
                      src=TCC_LOGO_URL )

6.7. TCCPage.__botNav(): Set up page-bottom navigational links

This method outputs the hr element below the page body, followed by page-bottom navigational links made from self.navList.

def __botNav ( self ):
    """Add the page-bottom navigational links section.
    
    [ (self.body as invariant) and (self.navList as invariant) ->
        self.body := self.body with bottom nav links added
        from self.navList ]
    """

First we add the separator rule.
Then we add the links from `self.navList`.

```python
#-- 1 --
# [ self.body += an hr element ]
exc.Element ( self.body, "hr" )
```

For each `NavItem` object in `self.navList`, there are three cases:

- If the `NavItem` object’s `.destList` is empty, no content is generated.
- For some links, such as “Site map”, the short name is the same as the title. It would be silly to generate the text “Site map: Site map” with the second repetition being the link.
  
  In that case, the `.destList` attribute has only one (title, URL) tuple, and the title is empty. We generate a link containing the boldfaced `.shortName` attribute, pointing at the given URL.

- In the general case, we start by generating the boldfaced `.shortName` attribute, but not as a link. This is followed by all links to each element of `.destList`, using the title of the (title, URL) tuple as the link text and the URL as the destination.

```python
#-- 2 --
# [ self.body += bottom nav links made from self.navList ]
for navItem in self.navList:

#-- 2 body --
# [ navItem is a NavLink ->
#   self.body += a bottom nav link made from navItem ]
self.__botNavItem ( navItem )
```

### 6.8. TCCPage.__botNavItem(): Generate page-bottom navigational link

For each `NavItem` object in `self.navList`, there are three cases:

- If the `NavItem` object’s `.destList` is empty, no content is generated.
- For some links, such as “Site map”, the short name is the same as the title. It would be silly to generate the text “Site map: Site map” with the second repetition being the link.
  
  In that case, the `.destList` attribute has only one (title, URL) tuple, and the title is empty. We generate a link containing the boldfaced `.shortName` attribute, pointing at the given URL.

- In the general case, we start by generating the boldfaced `.shortName` attribute, but not as a link. This is followed by all links to each element of `.destList`, using the title of the (title, URL) tuple as the link text and the URL as the destination.

```python
# - - - T C C P a g e . _ _ b o t N a v I t e m - - -
def __botNavItem ( self, navItem ):
    """Generate one page-bottom navigational link.
    
    [ navItem is a NavLink object ->
      self.body += a bottom nav link made from navItem ]
    ""
```

First we eliminate the case where nothing is generated.

```python
#-- 1 --
# [ if navItem.destList is empty or None ->
#   return
# else ->
#   self.body += a new, empty div element
#   div := that div element ]
if not navItem.destList:
    return
else:
    div = xc.Element ( self.body, "div" )
```
Next we check for the special case where the short name is the same as the title.

```python
#-- 2 --
# [ if navItem.destList[0] has an empty title ->
#   div += a link to the url of navItem.destList[0]
#   whose link text is navItem.shortName
#   return
# else -> I ]
title, url = navItem.destList[0]
if not title:
  #-- 2.1 --
  # [ self.body += a link to the url of
  #   navItem.destList[0] whose link text is
  #   navItem.shortName
  #   return ]
a = xc.Element ( div, "a", href=url )
b = xc.Element ( a, "b" )
xc.Text ( b, navItem.shortName )
return
```

In the general case, we first generate the boldfaced short name, then add all the elements of `navItem.destList` separated by `BOT_NAV_SEP`.

```python
#-- 3 --
# [ div += (navItem.shortName, boldfaced) + ": "
#   prefix := "" ]
b = xc.Element ( div, "b" )
xc.Text ( b, navItem.shortName )
xc.Text ( b, ": " )
prefix = ""
#-- 4 --
# [ div += (prefix) + (elements of navItem.destList,
#   made into links, separated by BOT_NAV_SEP)
#   prefix := BOT_NAV_SEP ]
for title, url in navItem.destList:
  #-- 4 body --
  # [ div += (prefix) + (a link to url with link
  #   text=title)
  #   prefix := BOT_NAV_SEP ]
  if prefix:
    xc.Text ( div, prefix )
a = xc.Element ( div, "a", href=url )
xc.Text ( a, title )
prefix = BOT_NAV_SEP
```

### 6.9. TCCPage.__colophon(): Add colophon section

The last part of the page to be generated contains a horizontal rule, the `address` element, the time of last update, and the URL (if known).
def __colophon ( self ):
    """Add the colophon section to the page.
    [ (self.author as invariant) and (self.url as invariant) ->
        self.body := self.body with the colophon added
        self.address := the address element within that
        colophon ]
    """

First, the horizontal rule:

```python
#-- 1 --
# [ self.body += an hr element ]
xc.Element ( self.body, "hr" )
```

Next, the address element. We insert self.author if it is nonempty.

```python
#-- 2 --
# [ self.body += a new address element containing
#     self.author
#     self.address := that element ]
self.address = xc.Element ( self.body, "address" )
if self.author:
    xc.Text ( self.address, self.author )
```

We package the timestamp in a div element to make it small, since this is not content most people care about.

```python
#-- 3 --
# [ self.body += current time ]
div = xc.Element ( self.body, "div" )
xc.Text ( div, "Last updated: %s" %
    time.strftime ( "%Y-%m-%d %H:%M %Z" ) )
```

Finally, we add another small heading with the URL, if it is known.

```python
#-- 4 --
# [ if self.url ->
#     self.body += self.url ]
if self.url:
    div = xc.Element ( self.body, "div" )
    xc.Text ( div, "URL: " )
    tt = xc.Element ( div, "tt" )
    xc.Text ( tt, self.url )
```

6.10. TCCPage.write(): Output the finished page

This method sends the XHTML to the output.
7. class NavLink: Describes one navigational feature

Most navigational links have three attributes:

• A short name such as “Previous” or “Help”.
• The link text to be used for this link in the bottom nav links. This is generally the page title or a close approximation; see the discussion of “navigational shock” in the PyStyler documentation.
• The destination URL of the link.

There are two principal design complications for navigational links:

• Some nav links have multiple destinations. For example, in the TCC help system, some pages list one or more “See also” links: the first is always the parent page in the structure, but additional related pages may be named.

• Some links do not appear in the top nav bar. For example, in TCC help system pages, the “See also” links appear only in the bottom nav link section.

So, here’s the interface to the NavLink object that represents one navigational link.