Pandoc migration: how easy is it?

Daren Sawkey

January 31, 2017

Pandoc a universal document converter

About

Installing

Getting started

Demos ▾

Documentation ▾

Help

Extras

Releases

About pandoc

If you need to convert files from one markup format into another, pandoc is your swiss-army knife. Pandoc can convert documents in markdown, reStructuredText, textile, HTML, DocBook, LaTeX, MediaWiki markup, TWiki markup, OPML, Emacs Org-Mode, Txt2Tags, Microsoft Word docx, LibreOffice ODT, EPUB, or Haddock markup to

- HTML formats: XHTML, HTML5, and HTML slide shows using Slidy, reveal.js, Slideous, S5, or DZSlides.
- Word processor formats: Microsoft Word docx, OpenOffice/LibreOffice ODT, OpenDocument XML
- Ebooks: EPUB version 2 or 3, FictionBook2
- Documentation formats: DocBook, TEI Simple, GNU TexInfo, Groff man pages, Haddock markup
- Page layout formats: InDesign ICML
- Outline formats: OPML
- TeX formats: LaTeX, ConTeXt, LaTeX Beamer slides
- PDF via LaTeX
- Lightweight markup formats: Markdown (including CommonMark), reStructuredText, AsciiDoc, MediaWiki markup, DokuWiki markup, Emacs Org-Mode, Textile
- Custom formats: custom writers can be written in lua.

Pandoc understands a number of useful markdown syntax extensions, including document metadata (title, author, date); footnotes; tables; definition lists; superscript and subscript; strikeout; enhanced ordered lists (start number and numbering style are significant); running example lists; delimited code blocks with syntax highlighting; smart quotes, dashes, and ellipses; markdown inside HTML blocks; and inline LaTeX. If strict markdown compatibility is desired, all of these extensions can be turned off.

LaTeX math (and even macros) can be used in markdown documents. Several different

Pandoc migration: how easy is it?

1. Install

Pandoc is written in Haskell
need compiler, libraries

2. Run

```
$ pandoc classCategory.xml global.xml -f docbook -t rst -s -o fundamentals.rst
```

Not everything worked first time (figures, citations, math)

DocBook

```
<!-- ****************** Section (Level#1) ***************** -->
<sect1 id="sect.ClassCate">
<title>
Class Categories and Domains
</title>

<!-- ****************** Section (Level#2) ***************** -->
<sect2 id="sect.ClassCate.WhatIsClassCate">
<title>
What is a class category?
</title>

<para>
In the design of a large software system such as Geant4, it is
essential to partition it into smaller logical units. This makes
the design well organized and easier to develop. Once the logical
units are defined independent to each other as much as possible,
they can be developed in parallel without serious interference.
</para>

<para>
In object-oriented analysis and design methodology by Grady Booch
<citation>
<xref linkend="biblio.booch1994" endterm="biblio.booch1994.abbrev" />
</citation>
, class categories are used to
create logical units. They are defined as "clusters of classes that
```

rST

```
Class Categories and Domains
=============================

What is a class category?
--------------------------

In the design of a large software system such as Geant4, it is essential
to partition it into smaller logical units. This makes the design well
organized and easier to develop. Once the logical units are defined
independent to each other as much as possible, they can be developed in
parallel without serious interference.

In object-oriented analysis and design methodology by Grady Booch
[Booch]_, class categories are used to create
logical units. They are defined as "clusters of classes that are
themselves cohesive, but are loosely coupled relative to other
clusters." This means that a class category contains classes which have
a close relationship (for example, the "has-a" relation). However,
relationships between classes which belong to different class categories
are weak, i.e., only limitted classes of these have "uses" relations.
The class categories and their relations are presented by a class
category diagram. The class category diagram designed for Geant4 is
shown in the figure below. Each box in the figure represents a class
category, and a "uses" relation by a straight line. The circle at an end
of a straight line means the class category which has this circle uses
the other category.

.. image:: AllResources/Fundamentals/classCategory.png
```

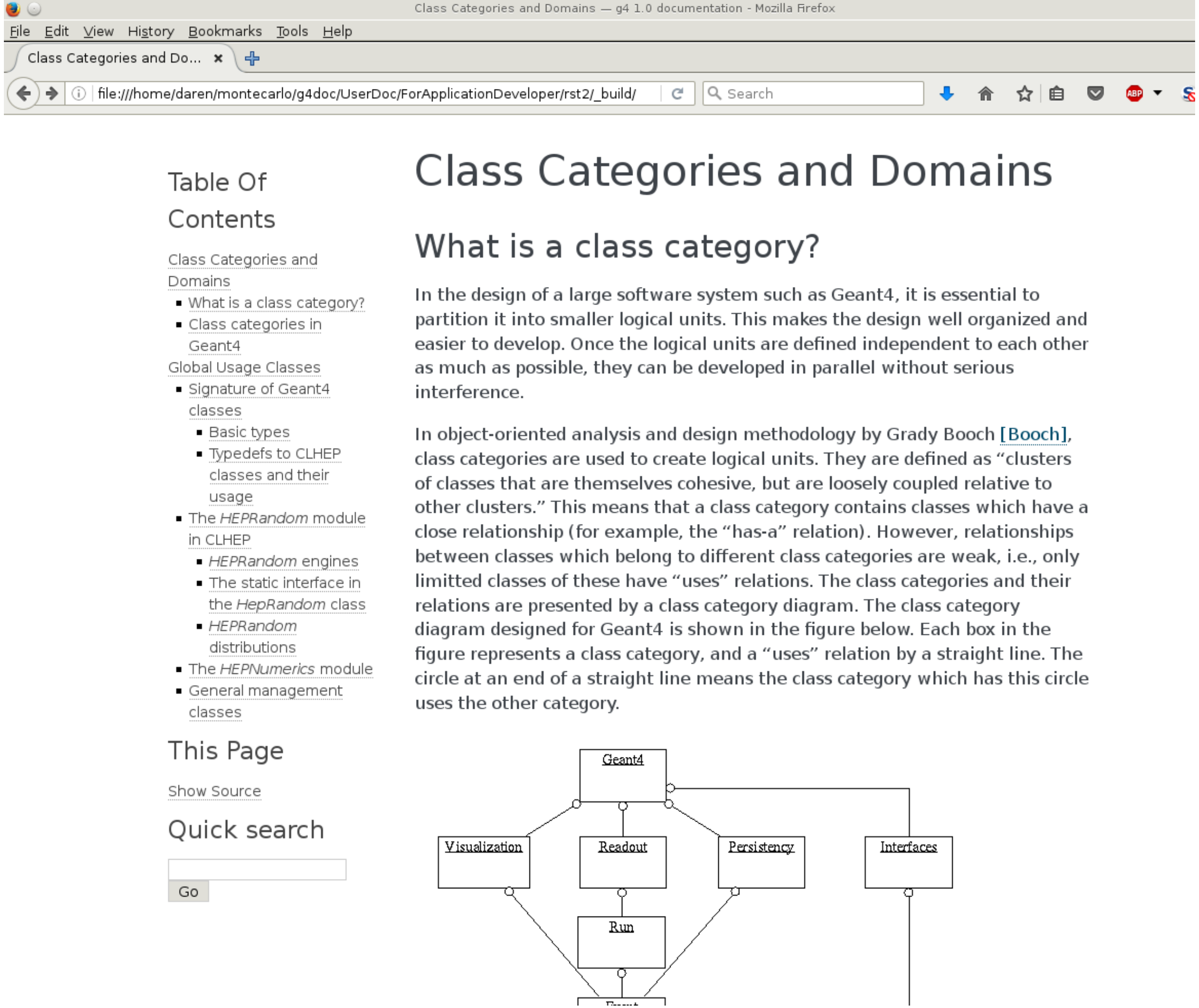

Class Categories and Domains — g4 1.0 documentation - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Class Categories and Do...
file:///home/daren/montecarlo/g4doc/UserDoc/ForApplicationDeveloper/rst2/_build/
Search

Table Of Contents

This Page

Show Source

Quick search

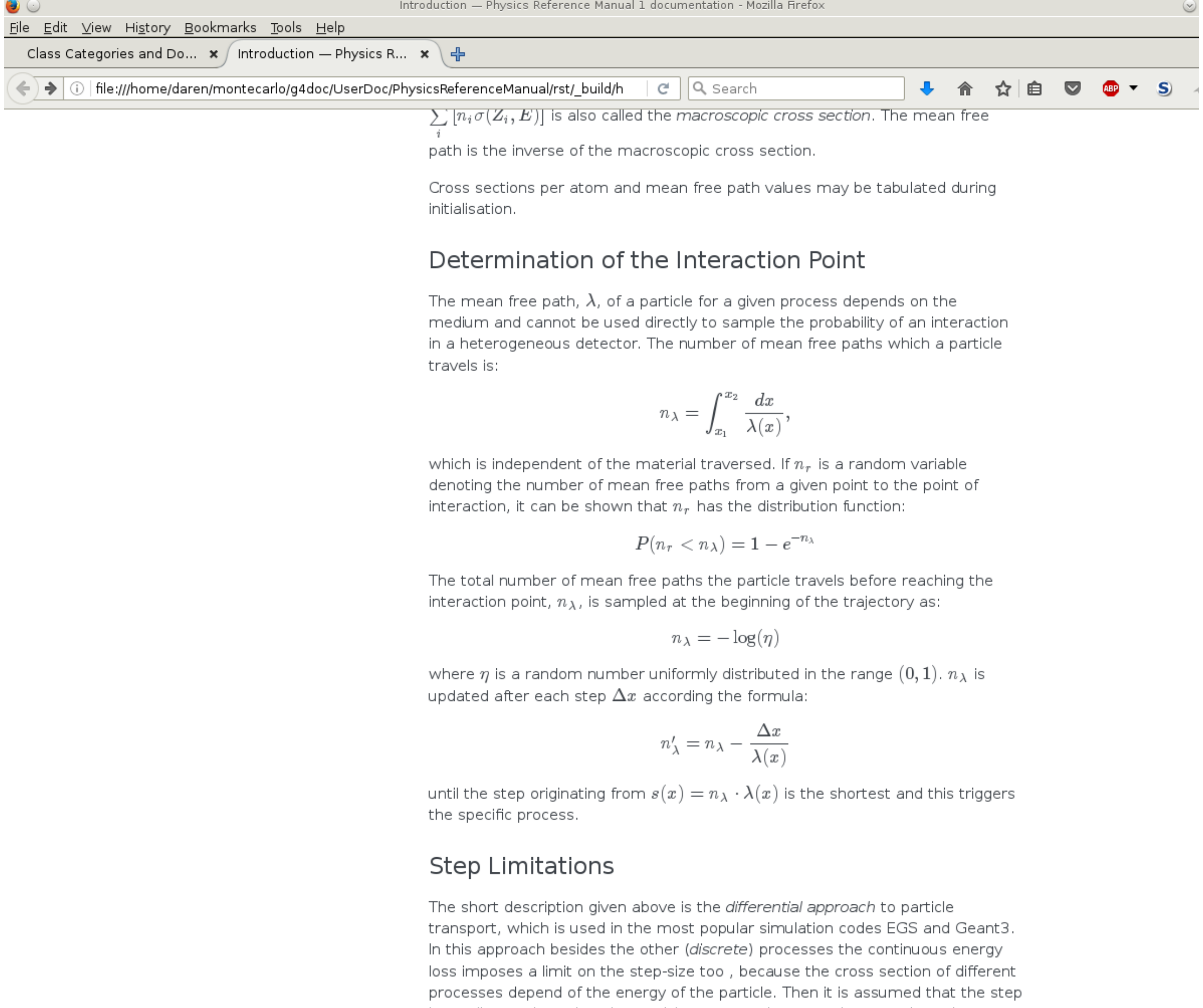
Go

Class Categories and Domains

What is a class category?

In the design of a large software system such as Geant4, it is essential to partition it into smaller logical units. This makes the design well organized and easier to develop. Once the logical units are defined independent to each other as much as possible, they can be developed in parallel without serious interference.

In object-oriented analysis and design methodology by Grady Booch [Booch], class categories are used to create logical units. They are defined as “clusters of classes that are themselves cohesive, but are loosely coupled relative to other clusters.” This means that a class category contains classes which have a close relationship (for example, the “has-a” relation). However, relationships between classes which belong to different class categories are weak, i.e., only limitted classes of these have “uses” relations. The class categories and their relations are presented by a class category diagram. The class category diagram designed for Geant4 is shown in the figure below. Each box in the figure represents a class category, and a “uses” relation by a straight line. The circle at an end of a straight line means the class category which has this circle uses the other category.


Introduction — Physics Reference Manual 1 documentation - Mozilla Firefox
File
Edit
View
History
Bookmarks
Tools
Help
Class Categories and Do...
Introduction — Physics R...
file:///home/daren/montecarlo/g4doc/UserDoc/PhysicsReferenceManual/rst/_build/h
Search

$\sum_i [n_i \sigma(Z_i, E)]$ is also called the *macroscopic cross section*. The mean free path is the inverse of the macroscopic cross section.

Cross sections per atom and mean free path values may be tabulated during initialisation.

Determination of the Interaction Point

The mean free path, λ, of a particle for a given process depends on the medium and cannot be used directly to sample the probability of an interaction in a heterogeneous detector. The number of mean free paths which a particle travels is:

$$n_\lambda = \int_{x_1}^{x_2} \frac{dx}{\lambda(x)},$$

which is independent of the material traversed. If n_r is a random variable denoting the number of mean free paths from a given point to the point of interaction, it can be shown that n_r has the distribution function:

$$P(n_r < n_\lambda) = 1 - e^{-n_\lambda}$$

The total number of mean free paths the particle travels before reaching the interaction point, n_λ, is sampled at the beginning of the trajectory as:

$$n_\lambda = -\log(\eta)$$

where η is a random number uniformly distributed in the range $(0, 1)$. n_λ is updated after each step Δx according the formula:

$$n'_\lambda = n_\lambda - \frac{\Delta x}{\lambda(x)}$$

until the step originating from $s(x) = n_\lambda \cdot \lambda(x)$ is the shortest and this triggers the specific process.

Step Limitations

The short description given above is the *differential approach* to particle transport, which is used in the most popular simulation codes EGS and Geant3. In this approach besides the other (*discrete*) processes the continuous energy loss imposes a limit on the step-size too , because the cross section of different processes depend of the energy of the particle. Then it is assumed that the step

Comments

- Only need to convert DocBook -> rST once
- => ease of use not a main factor
- Most time-consuming part will be proofreading/editing
    - E.g. Sphinx very picky about spaces in LaTeX mode
- Live demo: RadiotherapyResearchTools.com/geant4