Erudite

Mariano Montone

April 9, 2015

Contents

1	Introduction			
2	Invocation			
3	3 Implementation			
4	Backends			
	4.1	LaTeX	11	
	4.2	Sphinx	11	
5	Commands			
	5.1	Commands definition	13	
	5.2	Commands list	13	
		5.2.1 Input type	13	
		5.2.2 Output type	14	
		5.2.3 Title	14	
		5.2.4 Author	14	
		5.2.5 Chunks	14	
		5.2.6 Extraction	14	
		5.2.7 Ignore	15	
		5.2.8 Include	16	
6	Eru	Erudite syntax		
6.1 Syntax definition		Syntax definition	17	
	6.2 Commands list		17	
		6.2.1 Section	17	

iv CONTENTS

	6.2.2	Subsection	18
	6.2.3	Subsubsection	18
	6.2.4	Verbatim	18
	6.2.5	Code	18
	6.2.6	Lists	18
	6.2.7	Emphasis	19
	6.2.8	Bold	19
	6.2.9	Italics	19
	6.2.10	Inline verbatim	19
	6.2.11	Reference	20
6.3	Syntax	formatting	20
	6.3.1	Latex output	20

Introduction

Erudite is a system for Literate Programming in Common Lisp.

Some of its salient features are:

- Documentation is written in Common Lisp comments. This is very useful because you can work with your program as if it were not a literate program: you can load it, work from SLIME, etc, directly.
- Multiple outputs. Like Latex, Sphinx, etc.
- Automatic indexing and cross-references.
- A command line interface.
- It is portable. You can compile and use in several CL systems.

Invocation

Erudite is invoked calling erudite function.

```
(defun call-with-destination (destination function)
    ((null destination)
     (with-output-to-string (output)
       (funcall function output)))
    ((pathnamep destination)
     (with-open-file (f destination : direction : output
                     :if-exists :supersede
                     :if-does-not-exist :create)
       (funcall function f)))
    ((streamp destination)
     (funcall function destination))
    (t (error "Invalid destination: ~A" destination))))
(defmacro with-destination ((var destination) &body body)
  \hbox{`(call-with-destination , destination}\\
                          (lambda (,var) ,@body)))
(defun erudite (destination file-or-files
                &rest args &key (output-type *output-type*)
                             (syntax *syntax*)
                             &allow-other-keys)
 "Processes literate lisp files and creates a document.
  Args: - destination: If NIL, output is written to a string. If a pathname, then a file is
      created. Otherwise, a stream is expected.
         - files: Literate lisp files to compile
         - args: All sort of options passed to the generation functions
         - output-type: The kind of document to generate.
                        One of :latex, :sphinx
                        Default: :latex
         - syntax: The kind of syntax used in the literate source files.
                       One of: :erudite, :latex, :sphinx.
                       Default: :erudite"
 (with-destination (output destination)
    (let ((*output-type* output-type)
          (*syntax* syntax))
      (apply #'gen-doc output-type
             output
             (if (listp file-or-files)
                 file-or-files
                 (list file-or-files))
             args))))
```

Implementation

First, files with literate code are parsed into *fragments*. Fragments can be of type *documentation* or type *code*. *documentation* is the text that appears in Common Lisp comments. *code* fragments are the rest.

```
(defmethod process-file-to-string ((pathname pathname))
  (let ((*current-path* (fad:pathname-directory-pathname pathname)))
    (with-open-file (f pathname)
      (post-process-output
       (with-output-to-string (s)
         (process-fragments
          (split-file-source
           (extract-chunks f))
          s))))))
(defmethod process-file-to-string ((files cons))
  (post-process-output
   (with-output-to-string (s)
     (process-fragments
      (loop
         :for file :in files
         :appending (let ((*current-path* (fad:pathname-directory-pathname file)))
                      (with-open-file (f file)
                        (split-file-source
                         (extract-chunks f)))))
      s))))
(defmethod process-file-to-string :before (pathname)
 (setf *chunks* nil
        *extracts* nil))
(defmethod process-file-to-string :after (pathname)
 (setf *chunks* nil
        *extracts* nil))
(defun process-string (string)
 (let ((*chunks* nil)
        (*extracts* nil))
    (post-process-output
     (with-input-from-string (f string)
       (with-output-to-string (s)
         (process-fragments
          (split-file-source
           (extract-chunks f))
          s))))))
(defun post-process-output (str)
  "Resolve chunk inserts and extract inserts after processing"
```

```
(with-output-to-string (output)
    (with-input-from-string (s str)
      (loop
         :for line := (read-line s nil)
         :while line
         :do
         (cond
           ((scan "^__INSERT_CHUNK__(.*)$" line)
            (register-groups-bind (chunk-name)
                 ("^__INSERT_CHUNK__(.*)$" line)
Insert the chunk
               (let ((chunk (find-chunk chunk-name)))
                 (write-chunk chunk-name
                              (get-output-stream-string (cdr chunk))
                              output))))
           ((scan "^__INSERT_EXTRACT__(.*)$" line)
            (register-groups-bind (extract-name)
                 ("^__INSERT_EXTRACT__(.*)$" line)
Insert the extract
```

The parser works like a custom look-ahead parser, with a whole file line being the slice looked ahead. And is implemented in Continuation Passing Style.

```
(defun extract-chunks (stream)
  "Splits a file source in docs and code"
  (with-output-to-string (output)
      :with current-chunk := nil
      :for line := (read-line stream nil)
      :while line
      :do
      (cond
         ((scan "@chunk\s+(.+)" line)
          (register-groups-bind (chunk-name) ("@chunk\\s+(.+)" line)
            (setf current-chunk (list :name chunk-name
                                       :output (make-string-output-stream)))
            (write-chunk-name chunk-name output)
            (terpri output)))
          (push (cons (getf current-chunk :name)
                      (getf current-chunk :output))
                *chunks*)
          (setf current-chunk nil))
         (current-chunk
          (let ((chunk-output (getf current-chunk :output)))
            (write-string line chunk-output)
            (terpri chunk-output)))
         (t
          (write-string line output)
          (terpri output)))))
(defun split-file-source (str)
 "Splits a file source in docs and code"
  (with-input-from-string (stream str)
    (append-source-fragments
```

```
:for line := (read-line stream nil)
        :while line
        :collect
        (parse-line line stream)))))
(defun parse-line (line stream)
   (parse-long-comment line stream)
   (parse-short-comment line stream)
   (parse-code line stream)))
(defun parse-long-comment (line stream)
  "Parse a comment between #| and |#"
TODO: this does not work for long comments in one line
  (when (equalp (search "#|" (string-left-trim (list #\ #\tab) line))
We've found a long comment Extract the comment source
    (let ((comment
           (with-output-to-string (s)
First, add the first comment line
              (register-groups-bind (comment-line) ("\\\#\\|\\s*(.+)" line)
                (write-string comment-line s))
While there are lines without |#, add them to the comment source
              (loop
                :for line := (read-line stream nil)
                :while (and line (not (search "|#" line)))
                :do
                (terpri s)
                 (write-string line s)
                :finally
Finally, extract the last comment line
                     (register-groups-bind (comment-line) ("\\s*(.+)\\|\\#" line)
                       (when comment-line
                         (write-string comment-line s)))
                     (error "EOF: Could not complete comment parsing"))))))
      (list :doc comment))))
(defun parse-short-comment (line stream)
  (when (equalp
         (search *short-comments-prefix*
                  (string-left-trim (list #\ #\tab)
                                    line))
         0)
A short comment was found
    (let* ((comment-regex (format nil "~A\\s*(.+)" *short-comments-prefix*))
            (with-output-to-string (s)
               (register-groups-bind (comment-line) (comment-regex line)
                 (write-string
                  (string-left-trim (list #\; #\ )
```

```
comment-line)
                 s)))))
      (list :doc comment))))
(defun parse-code (line stream)
  (list :code line))
(defun append-to-end (thing list)
    ((null list)
     (list thing))
     (setf (cdr (last list))
           (list thing))
     list)))
(defun append-source-fragments (fragments)
  "Append docs and code fragments"
  (let ((appended-fragments nil)
        (current-fragment (first fragments)))
    (loop
       :for fragment :in (cdr fragments)
       (if (equalp (first fragment) (first current-fragment))
The fragments are of the same type. Append them
           (setf (second current-fragment)
                 (with-output-to-string (s)
                   (write-string (second current-fragment) s)
                   (terpri s)
                   (write-string (second fragment) s)))
else, there's a new kind of fragment
           (progn
             (setf appended-fragments (append-to-end current-fragment appended-fragments))
             (setf current-fragment fragment))))
    (setf appended-fragments (append-to-end current-fragment appended-fragments))
    appended-fragments))
(defun process-fragments (fragments output)
  (when fragments
    (let ((first-fragment (first fragments)))
      (process-fragment (first first-fragment) first-fragment
                        output
                        (lambda (&key (output output))
                           (process-fragments (rest fragments) output))))))
(defgeneric process-fragment (fragment-type fragment output cont))
(defmethod process-fragment ((type (eql :code)) fragment output cont)
  (write-code (second fragment) output *output-type*)
  (funcall cont))
(defmethod process-fragment ((type (eql :doc)) fragment output cont)
  (with-input-from-string (input (second fragment))
    (labels ((%process-fragment (&key (input input) (output output))
               (flet ((process-cont (&key (input input) (output output))
                         (%process-fragment :input input :output output)))
                 (let ((line (read-line input nil)))
                   (if line
                       (maybe-process-command line input output #'process-cont)
                       (funcall cont :output output)))))
      (%process-fragment))))
```

```
(defun find-matching-command (line)
 (loop
     :for command :in *commands*
     :when (match-command command line)
     :return command))
(defmethod maybe-process-command (line input output cont)
  "Process a top-level command"
  (let ((command (find-matching-command line)))
    (if command
        (process-command command line input output cont)
        (process-doc *syntax* *output-type* line output cont))))
(defmethod process-doc ((syntax (eql :latex)) output-type line stream cont)
 (write-string line stream)
 (terpri stream)
 (funcall cont))
(defmethod process-doc ((syntax (eql :sphinx)) output-type line stream cont)
 (write-string line stream)
  (terpri stream)
  (funcall cont))
(defmethod process-doc ((syntax (eql :erudite)) output-type line stream cont)
  (let ((formatted-line line))
    (loop
      :for syntax :in *erudite-syntax*
       :while formatted-line
      :when (match-syntax syntax formatted-line)
      :do
       (setf formatted-line (process-syntax syntax formatted-line stream output-type))
       :finally (when formatted-line
                  (write-string formatted-line stream)))
    (terpri stream)
    (funcall cont)))
(defmethod write-code (code stream (output-type (eql :latex)))
 (write-string "\\begin{code}" stream)
 (terpri stream)
 (write-string code stream)
 (terpri stream)
 (write-string "\\end{code}" stream)
  (terpri stream))
(defmethod write-chunk-name (chunk-name stream)
 (write-string "<<<" stream)
  (write-string chunk-name stream)
  (write-string ">>>" stream))
(defmethod write-chunk (chunk-name chunk stream)
  (write-code (format nil "<<~A>>=~%~A" chunk-name chunk)
              stream *output-type*))
```

Code blocks in Sphinx are indented. The indent-code function takes care of that:

```
lines))))
(defmethod write-code (code stream (output-type (eql :sphinx)))
  (write-string ".. code-block:: common-lisp" stream)
  (terpri stream)
  (write-string (indent-code code) stream)
  (terpri stream))
```

Backends

Erudite supports LaTeX and Sphinx generation at the moment.

4.1 LaTeX

```
(defgeneric gen-doc (output-type output files &rest args))
(defmethod gen-doc ((output-type (eql :latex)) output files
                    &key (title *title*)
                      (author *author*)
                      template-pathname
                      (syntax *syntax*)
                      (document-class *latex-document-class*)
                      &allow-other-keys)
  "Generates a LaTeX document.
  Args: - output: The output stream.
         - files: The list of .lisp files to compile
         - title: Title of the document
        - author: Author of the document
         - template-pathname: A custom LaTeX template file. If none is specified, a default
            template is used."
  (let ((*latex-document-class* document-class))
    (let ((template (cl-template:compile-template
                     (file-to-string (or template-pathname
                                          (asdf:system-relative-pathname
                                           :erudite
                                           "latex/template.tex")))))
          (body (process-file-to-string files)))
      (write-string
      (funcall template (list :title (or title
                                           *title*
                                           (error "No document title specified"))
                               :author (or author
                                            (error "No document author specified"))
                               :body body))
      output))
   t))
```

4.2 Sphinx

Sphinx is the other kind of output apart from LaTeX.

```
(defmethod gen-doc ((output-type (eql :sphinx)) output files &key prelude postlude syntax &
   allow-other-keys)
 "Generates Sphinx document.
  Args: - output: The output stream.
         - files: .lisp files to compile.
        - prelude: String (or pathname) to append before the Sphinx document.
        - postlude: String (or pathname) to append after the Sphinx document."
 (when prelude
    (write-string
     (if (pathnamep prelude)
         (file-to-string prelude)
        prelude)
    output))
 (write-string (process-file-to-string files) output)
 (when postlude
    (write-string (if (pathnamep postlude)
                      (file-to-string postlude)
                      postlude)
                  output)))
```

Commands

```
Commands are held in *commands* list

(defvar *commands* nil)

(defun find-command (name &optional (error-p t))

(let ((command (gethash name *commands*)))

(when (and error-p (not command))

(error "Invalid command: ~A" command))

command))
```

5.1 Commands definition

5.2 Commands list

5.2.1 Input type

```
(setf *syntax* (intern (string-upcase syntax) :keyword)))
(funcall cont)))
```

5.2.2 Output type

5.2.3 Title

5.2.4 Author

5.2.5 Chunks

5.2.6 Extraction

5.2. COMMANDS LIST

Build and register the extracted piece for later processing Redirect the output to the "extract output"

Restore the output

5.2.7 Ignore

```
(defvar *ignore* nil)
(define-command ignore
  (:match (line)
    (scan "@ignore" line))
  (:process (line input output cont)
            (setf *ignore* t)
            (funcall cont)))
(define-command end-ignore
  (:match (line)
    (scan "@end ignore" line))
  (:process (line input output cont)
            (setf *ignore* nil)
            (funcall cont)))
(defmethod process-doc :around (syntax output-type line stream cont)
 (if *ignore*
      (funcall cont)
      (call-next-method)))
(defmethod process-fragment :around ((type (eql :code)) fragment output cont)
 (if *ignore*
      (funcall cont)
      (call-next-method)))
(defmethod maybe-process-command :around (line input output cont)
```

```
(if (and *ignore* (not (match-command 'end-ignore line)))
     (funcall cont)
     (call-next-method)))
```

5.2.8 Include

```
(defvar *include-path* nil)
(define-command include-path
  (:match (line)
    (scan "@include-path\\s+(.+)" line))
  (:process (line input output cont)
            (register-groups-bind (path) ("@include-path\\s+(.+)" line)
              (setf *include-path* (pathname path))
              (funcall cont))))
(define-command include
  (:match (line)
    (scan "@include\\s+(.+)" line))
  (:process (line input output cont)
            (register-groups-bind (filename-or-path) ("@include\\s+(.+)" line)
              (let ((pathname (cond
                                 ((fad:pathname-absolute-p
                                  (pathname filename-or-path))
                                 filename-or-path)
                                 (*include-path*
                                 (merge-pathnames filename-or-path
                                                   *include-path*))
                                (t (merge-pathnames filename-or-path
                                                     *current-path*)))))
```

Process and output the included file

```
(write-string (process-file-to-string pathname) output)
(terpri output)
(funcall cont))))
```

Erudite syntax

Erudite formatting operations are held in *erudite-syntax* list

```
(defvar *erudite-syntax* nil)

(defun find-syntax (name &optional (error-p t))
  (let ((command (gethash name *erudite-syntax*)))
     (when (and error-p (not command))
        (error "Invalid syntax: ~A" command))
     command))
```

6.1 Syntax definition

6.2 Commands list

6.2.1 Section

```
(format-syntax output (list :section title)))
nil))
```

6.2.2 Subsection

6.2.3 Subsubsection

6.2.4 Verbatim

6.2.5 Code

6.2.6 Lists

```
(define-erudite-syntax begin-list
  (:match (line)
```

6.2. COMMANDS LIST

```
(scan "@list" line))
  (:process (line output output-type)
            (format-syntax output (list :begin-list))
(define-erudite-syntax end-list
  (:match (line)
    (scan "@end list" line))
  (:process (line output output-type)
            (format-syntax output (list :end-list))
            nil))
(define-erudite-syntax list-item
  (:match (line)
    (scan "@item" line))
  (:process (line output output-type)
            (regex-replace "@item" line
                           (lambda (match)
                              (format-syntax nil (list :list-item)))
                           :simple-calls t)))
```

6.2.7 Emphasis

6.2.8 Bold

6.2.9 Italics

6.2.10 Inline verbatim

```
(format-syntax nil (list :inline-verbatim text)))
:simple-calls t)))
```

6.2.11 Reference

6.3 Syntax formatting

6.3.1 Latex output

```
(defvar *latex-document-class* :article)
(defun format-syntax (destination syntax)
 (if (null destination)
      (with-output-to-string (stream)
        (%format-syntax *output-type* (first syntax) stream syntax))
      (%format-syntax *output-type* (first syntax) destination syntax)))
(defmethod %format-syntax ((output-type (eql :latex))
                           (selector (eql :section))
                           stream
                           syntax)
  (ecase *latex-document-class*
    (:article (format stream "\\section{~A}" (second syntax)))
    (:book (format stream "\\chapter{~A}" (second syntax)))))
(defmethod %format-syntax ((output-type (eql :latex))
                           (selector (eql :subsection))
                           syntax)
  (ecase *latex-document-class*
    (:article (format stream "\\subsection{~A}" (second syntax)))
    (:book (format stream "\\section{~A}" (second syntax)))))
(defmethod %format-syntax ((output-type (eql :latex))
                           (selector (eql :subsubsection))
                           stream
                           syntax)
  (ecase *latex-document-class*
    (:article (format stream "\\subsubsection{~A}" (second syntax)))
    (:book (format stream "\\subsection{~A}" (second syntax)))))
(defmethod %format-syntax ((output-type (eql :latex))
                           (selector (eql :begin-verbatim))
                           stream
                           syntax)
 (format stream "\\begin{verbatim}"))
(defmethod %format-syntax ((output-type (eql :latex))
                           (selector (eql :end-verbatim))
                           stream
                           syntax)
```

```
(format stream "\\end{verbatim}"))
(defmethod %format-syntax ((output-type (eql :latex))
                           (selector (eql :inline-verbatim))
                           stream
                           syntax)
 (format stream "\\verb|~A|" (second syntax)))
(defmethod %format-syntax ((output-type (eql :latex))
                            (selector (eql :begin-code))
                           stream
                            syntax)
 (format stream "\begin{code}"))
(defmethod %format-syntax ((output-type (eql :latex))
                           (selector (eql :end-code))
                           stream
                           syntax)
 (format stream "\\end{code}"))
(defmethod %format-syntax ((output-type (eql :latex))
                           (selector (eql :begin-list))
                           stream
                           syntax)
 (format stream "\\begin{itemize}"))
(defmethod %format-syntax ((output-type (eql :latex))
                            (selector (eql :end-list))
                           stream
                           syntax)
 (format stream "\\end{itemize}"))
(defmethod %format-syntax ((output-type (eql :latex))
                            (selector (eql :list-item))
                           stream
                           syntax)
 (format stream "\\item" (second syntax)))
(defmethod %format-syntax ((output-type (eql :latex))
                           (selector (eql :emph))
                           stream
                           syntax)
 (format stream "\\emph{~A}\" (second syntax)))
(defmethod %format-syntax ((output-type (eql :latex))
                            (selector (eql :bold))
                           stream
                            syntax)
 (format stream "\\textbf{~A}" (second syntax)))
(defmethod %format-syntax ((output-type (eql :latex))
                            (selector (eql :italics))
                           stream
                            syntax)
 (format stream "\\textit{~A}" (second syntax)))
(defmethod %format-syntax ((output-type (eql :latex))
                            (selector (eql :ref))
                           stream
                            syntax)
  (format stream "\\verb#~A#" (second syntax)))
```