

Literate Clojure

Brian Beckman

December 22, 2013

Contents

1	Introduction	1
1.1	How-To's	1
2	Tangle to Leiningen	2
2.1	Files in the Project Directory	4
2.1.1	.Gitignore	4
2.1.2	README.md	4
2.1.3	project.clj	5
2.2	The Documentation Subdirectory	5
2.3	Core Source File	5
2.3.1	The Namespace	6
2.3.2	Data Instances	6
2.3.3	A Test Excel Spreadsheet	7
2.4	Core Unit-Test File	7
3	A REPL-based Solution	9
4	Conclusion	10

1 Introduction

1.1 How-To's

This is a literate program.¹ That means that source code *and* documentation spring from the same, plain-text source files. That gives us a fighting chance of keeping knowledge and source coherent.

¹See http://en.wikipedia.org/wiki/Literate_programming.



Figure 1: This means “Fortune” and is pronounced “Fu”.

This file is named *ex2.org*. It’s an outline in plain text with obvious structure. Top-level headlines get a single star; second-level headlines get two stars, and so on; \LaTeX can be freely written anywhere; source-code examples abound to copy-and-paste, and text explaining how to build and run the source is nearby.

You can edit the file with any plain-text editor. Emacs offers some automation in generating the typeset output, *ed2.pdf*, and the source code of the application right out of the *org* file. To generate source code, issue the emacs command `org-babel-tangle`. To generate documentation, issue the emacs command `org-latex-export-to-pdf`.

We are working on a batch process via *make* so that you can just clone the repo, make whatever edits you like, type *make*, and have a complete PDF file and a complete directory full of source code.

2 Tangle to Leiningen

Let’s generate Leiningen projects.² Leiningen is the easiest way to use Clojure.³ Clojure is a 100% Java-compatible functional programming language; it is simple, straightforward, and arguably a great way to use Java post-2010. As with any Java-based language, there is significant “ceremony” in setting up code to run. Files must be in certain directories that correspond

²<http://leiningen.org>

³<http://clojure.org>

to namespaces and packages, and the ever-finicky *classpath* must be set up. This ceremony is often much more time-consuming than the code. Much of the value of the Clojure-Leiningen combination is that Leiningen automates almost all the ceremony.

After *tangling* this file, as directed in this section, you will have a Leiningen project. Go to the project directory (the one containing the file *project.clj*), and type `lein test` in a console or terminal window, running all the unit tests. Type `lein repl` to get an interactive session, in which you may run code from the project or any other experimental code. If you're using emacs, you can also run the repl directly in emacs, as described in section 3.

First, let's show the Leiningen project in detail. If you were to run the following command at a console prompt

```
$ lein new ex2
```

you would get the following source tree:

```
ex2
ex2/.gitignore
ex2/doc
ex2/doc/intro.md
ex2/project.clj
ex2/README.md
ex2/resources
ex2/src
ex2/src/ex2
ex2/src/ex2/core.clj
ex2/test
ex2/test/ex2
ex2/test/ex2/core_test.clj
```

We create the identical, base structure by typing

```
M-x org-babel-tangle
```

and no more, in our org-mode buffer in emacs (or, eventually, by typing *make* in the root directory, for non-users of emacs). Below, we tangle some more, application-specific code into that directory structure.

2.1 Files in the Project Directory

In our example, the top-level directory doesn't have a name – put our *org* file in that directory. The Leiningen project directory will have the same name as our *org* file. Our *org* file is named **ex2.org** and we want a directory tree rooted at **ex2** exactly as above.

Start with the contents of the project directory, **ex2**. Each org-mode babel source-code block will name a file path – including sub-directories – after a `:tangle` keyword on the `#+BEGIN_SRC` command of org-mode.

2.1.1 .Gitignore

First, we must create the `.gitignore` file that tells `git` not to check in the ephemeral *ejecta* of build processes like `maven` and `javac`. When we gain more confidence and adoption with `tangle` and `LATEX`, we will even ignore the PDF file and the generated source tree, saving *only* the *org* file in the repository.

```
/target
/lib
/classes
/checkouts
pom.xml
pom.xml.asc
*.jar
*.class
.lein-deps-sum
.lein-failures
.lein-plugins
.lein-repl-history
```

2.1.2 README.md

Next, we produce a `README.md` in markdown syntax for the entire project:

```
# ex2
A Clojure library designed to do SOMETHING.
## Usage
TODO
## License
Copyright © 2013 TODO
```

2.1.3 project.clj

Next is the `project.clj` file required by Leiningen for fetching dependencies, loading libraries, and other housekeeping. If you are running the Clojure REPL inside emacs, you must visit this file *after tangling it out of the org file*, and then run

```
M-x nrepl-jack-in
```

in that buffer (see more in section 3).

```
(defproject ex2 "0.1.0-SNAPSHOT"
  :description "DocJure's Excel Processor"
  :url "http://example.com/TODO"
  :license {:name "TODO"
            :url "TODO"}
  :dependencies [[org.clojure/clojure "1.5.1"]
                 [org.clojure/data.zip "0.1.1"]
                 [dk.ative/docjure "1.6.0"]
                ]
  :repl-options {:init-ns ex2.core})
```

2.2 The Documentation Subdirectory

Mimicking Leiningen's documentation subdirectory, it contains the single file `intro.md`, again in markdown syntax.

```
# Introduction to ex2
```

```
TODO: The project documentation is the .org file that produced
this output, but it still pays to read
http://jacobian.org/writing/great-documentation/what-to-write/
```

2.3 Core Source File

By convention, the core source files go in a subdirectory named `./ex2/src/ex2`. This convention allows the Clojure namespaces to map to Java packages.

The following is our core source file, explained in small pieces. The *org* file contains a spec for emitting the tangled source at this point. This spec is not visible in the generated PDF file, because we want to individually document the small pieces. The invisible spec simply gathers up the source of the small pieces from out of their explanations and then emits them into

the source directory tree, using another tool called *noweb*.⁴ This is not more complexity for you to learn, rather it is just a way for you to feel comfortable with literate-programming magic.

2.3.1 The Namespace

First, we must mention the libraries we're using. This is pure ceremony, and we get to the meat of the code immediately after. These library-mentions correspond to the `:dependencies` in the `project.clj` file above. Each `:use` or `:require` below must correspond to either an explicit dependency in the `project.clj` file or to one of several implicitly loaded libraries. Leiningen loads libraries by processing the `project.clj` file above. We bring symbols from those libraries into our namespace so we can use the libraries in our core routines.

To ingest and compile raw Excel spreadsheets, we use the built-in libraries `clojure.zip` for tree navigation and `clojure.xml` for XML parsing, plus the third-party libraries `clojure.data.zip.xml` and `dk.ative.docjure.spreadsheet`. The following brings these libraries into our namespace:

```
(ns ex2.core
  (:use [clojure.data.zip.xml :only (attr text xml->)]
        [dk.ative.docjure.spreadsheet] )
  (:require [clojure.xml :as xml]
            [clojure.zip :as zip]))
```

2.3.2 Data Instances

Next, we create a couple of data instances to manipulate later in our unit tests. The first one ingests a trivial XML file and the second one converts the in-memory data structure into a *zipper*,⁵ a very modern, functional tree-navigation facility. These instances will test our ability to freely navigate the raw XML form of Excel spreadsheets:

```
(def xml (xml/parse "myfile.xml"))
(def zippered (zip/xml-zip xml))
```

⁴<http://orgmode.org/manual/Noweb-reference-syntax.html>

⁵<http://richhickey.github.io/clojure/clojure.zip-api.html>

2.3.3 A Test Excel Spreadsheet

Finally, we use `docjure` to emit a test Excel spreadsheet, which we will read in our unit tests and verify some operations on it. This code creates a workbook with a single sheet in a rather obvious way, picks out the sheet and its header row, and sets some visual properties on the header row. We can open the resulting spreadsheet in Excel after running `lein test` and verify that the `docjure` library works as advertised.

```
(let [wb (create-workbook "Price List"
                          [["Name"      "Price"]
                           ["Foo Widget" 100]
                           ["Bar Widget" 200]])
      sheet (select-sheet "Price List" wb)
      header-row (first (row-seq sheet))]
  (do
    (set-row-style!
     header-row
     (create-cell-style! wb
                          {:background :yellow,
                           :font       {:bold true}}))
    (save-workbook! "spreadsheet.xlsx" wb)))
```

2.4 Core Unit-Test File

Unit-testing files go in a subdirectory named `./ex2/test/ex2`. Again, the directory-naming convention enables valuable shortcuts from Leiningen.

As with the core source files, we include the built-in and downloaded libraries, but also the `test framework` and the `core` namespace, itself, so we can test the functions in the core.

```
(ns ex2.core-test
  (:use [clojure.data.zip.xml :only (attr text xml->)]
        [dk.ative.docjure.spreadsheet])
  )
(:require [clojure.xml :as xml]
          [clojure.zip :as zip]
          [clojure.test :refer :all]
          [ex2.core :refer :all]))
```

We now test that the zippered XML file can be accessed by the *zipper* operators. The main operator of interest is `xml->`, which acts a lot like Clojure's *fluent-style*⁶ *threading* operator `->`.⁷ It takes its first argument, a zippered XML file in this case, and then a sequence of functions to apply. For instance, the following XML file, when subjected to the functions `:track`, `:name`, and `text`, should produce `'("Track one" "Track two")`

```
<songs>
  <track id="t1"><name>Track one</name></track>
  <ignore>pugh!</ignore>
  <track id="t2"><name>Track two</name></track>
</songs>
```

Likewise, we can dig into the attributes with natural accessor functions⁸ `#+name: docjure-test-namespace`

```
(deftest xml-zipper-test
  (testing "xml and zip on a trivial file."
    (are [a b] (= a b)
      (xml-> zippered :track :name text) '("Track one" "Track two")
      (xml-> zippered :track (attr :id)) '("t1" "t2"))))
```

Next, we ensure that we can faithfully read back the workbook we created *via docjure*. Here, we use Clojure's `thread-last` macro to achieve fluent style:

⁶http://en.wikipedia.org/wiki/Fluent_interface

⁷http://clojuredocs.org/clojure_core/clojure.core/-%3E

⁸Clojure treats colon-prefixed keywords as functions that fetch the corresponding values from hashmaps, rather like the dot operator in Java or JavaScript; Clojure also treats hashmaps as functions of their keywords: the result of the function call `({:a 1} :a)` is the same as the result of the function call `(:a {:a 1})`


```
(deftest docjure-test
  (testing "docjure read"
    (is (=

      (->> (load-workbook "spreadsheet.xlsx")
            (select-sheet "Price List")
            (select-columns {:A :name, :B :price}))

      [{:name "Name"          , :price "Price"}, ; don't forget header row
       {:name "Foo Widget", :price 100.0  },
       {:name "Bar Widget", :price 200.0  }])

    ))))
```

3 A REPL-based Solution

To run the REPL for interactive programming and testing in org-mode, take the following steps:

1. Set up emacs and nRepl (TODO: explain; automate)
2. Edit your init.el file as follows (TODO: details)

```
;;; To load org-babel in Emacs, add this code to initialization
(when (locate-file "ob" load-path load-suffixes)
  (require 'ob)
  (require 'ob-tangle)
  (require 'ob-clojure)
  (org-babel-do-load-languages
   'org-babel-load-languages
   '((emacs-lisp . t)
     (clojure    . t))))
;; Under nrepl.el + NREPL:
;; Patch ob-clojure to work with nrepl
(declare-function nrepl-send-string-sync "ext:nrepl" (code &optional ns))
(defun org-babel-execute:clojure (body params)
  "Execute a block of Clojure code with Babel."
  (require 'nrepl)
  (with-temp-buffer
    (insert (org-babel-expand-body:clojure body params))
```

```

((lambda (result)
  (let ((result-params (cdr (assoc :result-params params))))
    (if (or (member "scalar" result-params)
            (member "verbatim" result-params))
        result
        (condition-case nil (org-babel-script-escape result)
            (error result))))))
(plist-get (nrepl-send-string-sync
  (buffer-substring-no-properties (point-min) (point-max))
  (cdr (assoc :package params)))
:value)))

```

1. Start nRepl while visiting the actual `project-clj` file.
2. Run code in the org-mode buffer with `C-c C-c`; results of evaluation are placed right in the buffer for inspection; they are not copied out to the PDF file.

```

[(xml-> zippered :track :name text)      ; ("Track one" "Track two")
 (xml-> zippered :track (attr :id))]      ; ("t1" "t2")

(->> (load-workbook "spreadsheet.xlsx")
      (select-sheet "Price List")
      (select-columns {:A :name, :B :price}))

(run-all-tests)

```

4 Conclusion

Fu is Fortune.