DFB and Source Code Readme for the Mac OS/X Platform

Willie Rogers

November 9, 2011

1 Preliminaries:

This document is a reference on installing tools to support Data File Builder or Source Code compilation on the Mac OS/X Platform.

This document is relevant only if

1. You do not want to use one of the data models provided with MetaMap, but want instead to build a custom Metathesaurus, and
2. You plan to build your custom Metathesaurus.
3. You wish to modify and build the C and Prolog sources for MetaMap or its associated programs.

The Data File Builder module, which constructs a custom Metathesaurus, requires certain GNU utilities, which are included in Linux distributions, however, most but not all are included in Mac OS/X. If you plan to use the Data File Builder on Mac OS/X, you need to download and install the GNU utilities.

It is *essential* that the GNU utilities be available, because the Data File Builder scripts use the GNU versions of programs such as grep, cut, join, sort, etc., and which may not work properly if the BSD and SYS V versions of these programs are used instead.

The necessary GNU utilities may be freely downloaded and compiled using GCC or CLang compiler supplied with Xcode.

If the GNU utilities are already available on your system, you may skip to the Configuring the Shell Environment: (§ 4) section below.

2 Downloading Software:

2.1 Software necessary for the use of Data FIle Builder

GNU coreutils The GNU Core Utilities are the basic file, shell and text manipulation utilities of the GNU operating system.

GNU sed This is a stream editor. A stream editor is used to perform basic text transformations on an input stream (a file or input from a pipeline). The data file builder scripts are written to use GNU sed rather than the original BSD/UNIX sed.

The source code to these utilities are available at http://www.gnu.org/software/.

Pre-compiled (binary) versions of these utilities can be obtained at the following:

Fink Project http://www.finkproject.org/

Darwin Ports http://coreutils.darwinports.com/, http://darwinports.com/

Mac Ports http://www.macports.org/

2.2 Software necessary for Source Code Compilation

Apple Developer Tools These tools provide the compilers and linkers necessary to build applications on Mac OS/X. These tools are supplied with the computer but can be downloaded at the following url: http://developer.apple.com/technologies/tools/

SICStus Prolog 4.1.3 available from Swedish Institute of Computer Science (SICS), the SICStus Prolog website: http://www.sics.se/sicstus/

3 Installing GNU Packages:

Refer to the installation instructions supplied by the provider of coreutils package, source or binary (pre-compiled). Be sure to install the Apple Developer Tools if you're using the source package for coreutils.

4 Configuring the Shell Environment:

The GNU utilities were installed above in /usr/local/bin. Before running BuildDataFiles, ensure that the programs grep, sed, sort, cut, join, etc. are run from `/usr/local/bin` and not `/bin`, `/usr/bin`, or another directory (specified by the packager or yourself). The simplest way to do this is to run the following command (using `csh` or `tcsh`):

```
$ set path = ( /usr/local/bin $path )
```

or using `sh` or `bash`:

```
$ PATH=/usr/local/bin:$PATH
$ export PATH
```

The appropriate command should be added to your `.cshrc` or `.bashrc` file.

See the Data File Builder Documentation (http://metamap.nlm.nih.gov/datafilebuilder.pdf) for complete instructions on running BuildDataFiles.

5 Installing LVG on Mac OS/X

If you are using a Mac there are special instructions for installing LVG on Mac OS/X:

Download the "Lite" version of Lvg for the year you wish to use (the example uses LVG 2010.) (see Lexical Tools Download Page: http://lexsrv3.nlm.nih.gov/LexSysGroup/Projects/lvg/current/web/release/index.html)

extract it using tar and then for each of the files in the `lvg2010lite/bin` directory replace:

```
JAVA=/export/home/lu/Development/LVG/lvg2010/bin/jre1.6.0_14/bin/java
```

with:

```
JAVA=java
```

and:

```
LVG_DIR=/export/home/lu/Development/LVG/lvg2010
```

with:

```
LVG_DIR={where lvg is extracted}/lvg2010lite
```

Also modify the CLASSPATH variable: include both ${LVG_DIR} and lvg2011dist.jar.

Make sure all of the files in `lvg2010lite/bin` are executable:

```
chmod +x lvg2010lite/bin/*
```

Also modify the configuration file (${LVG_DIR}/data/config/lvg.properties.

The Java VM (java) is currently provided with Mac OS/X Leopard, Snow Leopard, and Lion (Note: MetaMap has not been tested on Mac OS/X Lion).

More detailed instructions on how to install Lvg manually are available at the Lexical Systems WebSite. See details at URL:

http://lexsrv3.nlm.nih.gov/LexSysGroup/Projects/lvg/current/docs/userDoc/install/installManual.html