

Using MeCab from MATLAB® (in 64bit windows 7 + MATLAB R2014b)

This document describes a procedure to build MeCab 0.966 in 64bit windows 7 environment using Visual C++ 2010 compiler. The built DLL file can be used directly from MATLAB using loadlibrary command.

Copyright MathWorks, 2015

According to the author, MeCab is a Japanese morphological analyzer, and stands as "Yet Another Part-of-Speech and Morphological Analyzer".

MeCab Home Page

<http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

Software Used in this document:

1. 64bit Windows 7 (Have not tested in 64bit Windows 8 , but probably OK also)
2. Microsoft Visual Studio 2010 (Visual C++ 2010 is used)
3. MATLAB version R2014b or R2015a (for 64bit windows)

Software Download

1. Download & Install MeCab 0.966 binary package (includes compiled dictionary file)
Install the package in default C:\Program Files (x86)\MeCab directory.
<http://code.google.com/p/mecab/downloads/detail?name=mecab-0.996.exe>
2. Download and extract MeCab 0.966 source zip file
MeCab 0.966 download link (This document assumes source code from 0.966 version only)
<http://code.google.com/p/mecab/downloads/detail?name=mecab-0.996.tar.gz>
3. Unpack mecab-0.966.tar.gz to your local directory (winzip worked nicely here).
Open "src" directory of the package in windows explorer.

Preparing Makefile.msvc

1. Locate Makefile.msvc.in in "src" directory
2. Copy & Rename Makefile.msvc.in to Makefile.msvc
3. Open Makefile.msvc in text editor (or double click the file in MATLAB to open in MATLAB editor)
4. Locate and fix line 6, 8, and 9, with text highlighted in red ("**X64**", "**0.996**", "**102**")

```
5: CFLAGS = /EHsc /O2 /GL /GA /Ob2 /nologo /W3 /MT /Zi /wd4800 /wd4305 /wd4244
6: LDFLAGS = /nologo /OPT:REF /OPT:ICF /LTCG /NXCOMPAT /DYNAMICBASE /MACHINE:X64 ADVAPI32.LIB
7: DEFS = -D_CRT_SECURE_NO_DEPRECATED -DMECAB_USE_THREAD \
8:       -DDLL_EXPORT -DHAVE_GETENV -DHAVE_WINDOWS_H -DDIC_VERSION=102 \
9:       -DVERSION="0.996" -DPACKAGE="\"mecab\" \" \
10:      -DUNICODE -D_UNICODE \
11:      -DMECAB_DEFAULT_RC="\"c:\\Program Files\\mecab\\etc\\mecabrc\""
```

Modifying common.h

Add "#include <iterator>" to common.h file under line 16

```
14: #include <string>
15: #include <iostream>
16: #include <sstream>
17: #include <iterator>
```

Modifying feature_index.cpp

Open feature_index.cpp, and go to line 356. Change line 356 cast definition to below ("size_t" to "unsigned int")

```
353:         if (!r) goto NEXT;
354:         os_ << r;
355:     } break;
356:     case 't': os_ << (unsigned int)path->rnode->char_type;     break;
357:     case 'u': os_ << ufeature; break;
358:     case 'w':
359:         if (path->rnode->stat == MECAB_NOR_NODE) {
360:             os_.write(path->rnode->surface, path->rnode->length);
```

Modifying writer.cpp

Open writer.cpp, and go to line 260. Add cast of (unsigned int) in front of lattice->size()

```
257:         // input sentence
258:         case 'S': os->write(lattice->sentence(), lattice->size()); break;
259:         // sentence length
260:         case 'L': *os << (unsigned int)lattice->size(); break;
261:         // morph
262:         case 'm': os->write(node->surface, node->length); break;
```

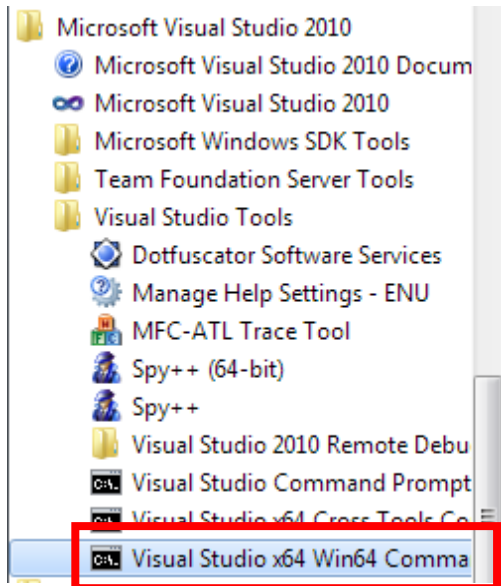
Modifying mecab.h

Open mecab.h, and go to line 1125, and line 1414. Fix SIWG to SWIG

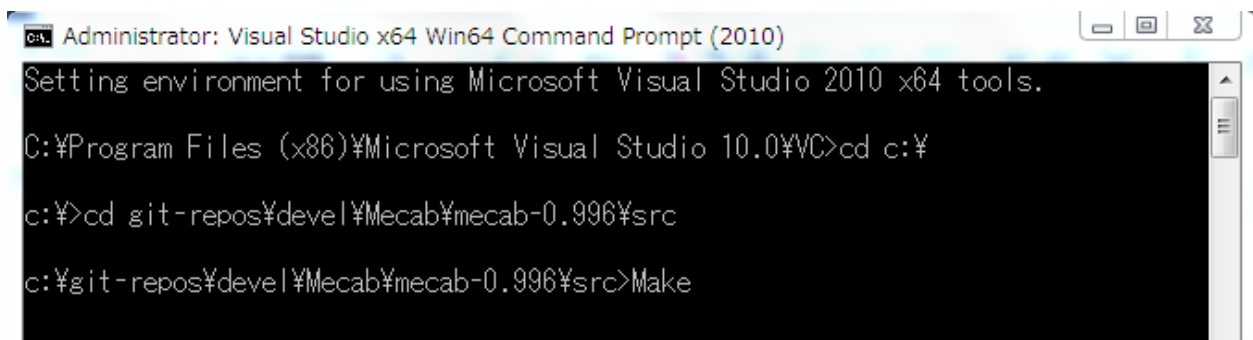
```
1123:     virtual ~Model() {}
1124:
1125:     #ifndef SWIG
1126:     /**
1127:     .
1128:     .
1129:     .
1412:     virtual ~Tagger() {}
1413:
1414:     #ifndef SWIG
1415:     /**
```

Building libmecab.dll

After above fixes are in place, locate "Visual Studio x64 Win64 Command Prompt" from Start → "Microsoft Visual Studio 2010" menu.



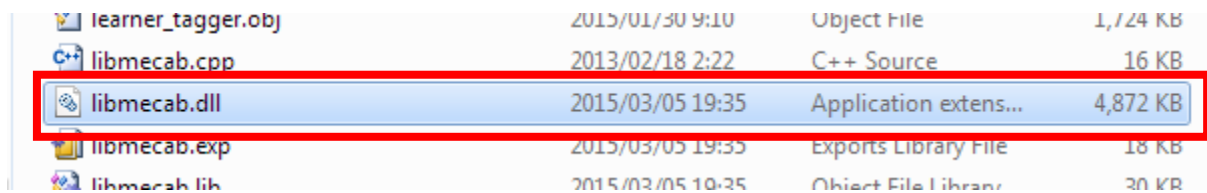
1. From the visual studio command prompt, change directory to where the mecab source files are located (which you have just performed series of source code changes)
2. Type "Make" at the src directory (Make.bat is the build batch file provided to compile the source code)



```
Administrator: Visual Studio x64 Win64 Command Prompt (2010)
Setting environment for using Microsoft Visual Studio 2010 x64 tools.
C:\Program Files (x86)\Microsoft Visual Studio 10.0\VC>cd c:\
c:\>cd git-repos\devel\Mecab\mecab-0.996\src
c:\git-repos\devel\Mecab\mecab-0.996\src>Make
```

The directory above is an example. Please change directory to where you have unpacked the mecab zip files.

After series of messages, you should see libmecab.dll generated in the src directory



learner_tagger.obj	2015/01/30 9:10	Object File	1,724 KB
libmecab.cpp	2013/02/18 2:22	C++ Source	16 KB
libmecab.dll	2015/03/05 19:35	Application extens...	4,872 KB
libmecab.exp	2015/03/05 19:35	Exports Library File	16 KB
libmecab.lib	2015/03/05 19:35	Object File Library	30 KB

Using libmecab.dll from MATLAB

It is now time to use generated x64 libmecab.dll from MATLAB

Copy libmecab.dll, and mecab.h into the same directory as a MATLAB script to call the library. Upon executing the following code, you may find warnings being output in MATLAB command window, upon loadlibrary command.

Sample code to call the dll

```
%% Loading DLL
fname_lib = 'libmecab.dll';
fname_header = 'mecab.h';

[notfound, warnings] = loadlibrary(fname_lib, fname_header);

%% (optional) view available functions of the DLL
% libfunctionsview('libmecab')

%% Calling MeCab morphological analyzer
argv = libpointer('stringPtrPtr', {'MeCab'});
argc = 1;

mecab = calllib('libmecab', 'mecab_new', argc, argv);

%% test MeCab functionality
input = '太郎は次郎が持っている本を花子に渡した。';
result = calllib('libmecab', 'mecab_sparse_tostr', mecab,
input);

%% Release DLL from the memory
clear
unloadlibrary('libmecab')
```
