NSM INPUT FILES

July 17, 2009

1 Overview

An NSM input file is a text file, with some special tags which show to NSM-DALIA what to do.

When NSM-DALIA parses such a file, it works in normal file-reading mode (simply outputting each character found), until a tag is encountered at the beginning of a new line. Tags begin with a "@" character, followed by a letter. Tags are case-sensitive, so @p and @P are different commands!

If you write an NSM input file with a word processor (instead of a text editor), you will have to save your file in pure text format (".txt"), ANSI-encoded (NOT unicode or UTF8. Support for such encodings will be hopefully made available in later versions).

Here is a quick tour of the main tags implemented so far. Other will be made available as the program is developed.

@p Following text is split into sentences (fullstop-separated), parsed and transformed into NSM-PROLOG notation, until an @e tag is found. Then, normal file-reading mode is resumed. You will find many example of how this works in the "sentences.txt" file, in the demo directory. See what happens when you run a pf("demo/sentences.txt") command in NSM-DALIA: the sentences between a @p ... @e block are replaced (in the output) by their NSM-PROLOG analysis, while all the text outside such blocks is output as it is. Please note that the text in a @p .. @e block is NOT in NSM format, but sentences must be fullstop-separated, and indentation plays no role.

@P Text in a @P ... @e block is considered to be in NSM- standard format. See examples of this in the "texts.txt" file, in the "demo" directory. The NSM standard format uses newline as sentence separator, no punctuation, and indentation as a means of showing quotation. If you want to split a long line, but you want the parser to consider the new line as a continuation of the previous sentence, put a "/" or a "\" at the end of the first part, and indentation and newline will be not considered. Ex. in the following text, all indentations and newlines are significative:

```
this person thinks like this
    these people are not people like me
    these people are bad people
```

But you can also write it like this:

```
this person thinks like this
    these people \
```

1

```
            are not people like me
    these people \
            are bad people
```

@t Text in a @t ... @e block is translated from current_language NSM into curernt_l2 NSM. If current_language and current_l2 are the same, text is not first parsed and then re-generated in the same language, but simply output as it is. Text is in normal format, with fullstops showing sentence boundaries.

@T As before, but this time, text must be in NSM standard format.

@g Text in a "@g ... end." block is read as NSM-PROLOG formulas, and replaced by its equivalent in the target language. Note that this block is ended by an "end." tag (with fullstop!), and not by the usual "@e" tag. If generation fails for some reason, the formula is output unchanged. You can see how this and the following tag work by making NSM-DALIA parse the file "text-gen.txt", in the "demo" directory. Text is in normal format.

@G As before, but with text in NSM standard format. The end of the block is given by and "end." command (with fullstop), instead of the usual "@e" tag.

@@ This lets you write a "@" character at the beginning of a line (writing a simple "@" in that position would start a tag-recognition process).

2   License

This file is part of NSM-DALIA, an extensible parser and generator for NSM grammars.

Copyright (C) 2009 Francesco Zamblera.

This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program. If not, see `http://www.gnu.org/licenses/`.