GROMACS - Task #1947

Consider fixing corner case in TinyXML2

04/19/2016 01:39 AM - Mark Abraham

Status: New
Priority: Normal
Assignee: Mark Abraham
Category: 
Target version: 
Difficulty: uncategorized

Description

In https://gerrit.gromacs.org/#/c/5653/ we identified that TinyXML2 can't read an XML element that is non-empty and only whitespace. It's possible we might want to use that some time, though for now we don't (and can't do so inadvertently).

For a rainy day, see if TinyXML2 can be patched. See discussion at https://github.com/leethomason/tinyxml2/issues/432 (and others of their issues).

Associated revisions

Revision 2f00ac17 - 04/19/2016 04:54 PM - Mark Abraham

Replace libxml2 with tinyxml2 for use in test code

Building libxml2 on complex machines has been problematic for users and developers, and there is no pressing reason to need the full capabilities of libxml2 for reading and writing data for test cases. Support for detecting libxml2 is retained, because other code in Gerrit might plan to use it, but it currently is not needed for anything.

This is version 3.0.0 of tinyxml2, modified only so that XMLPrinter::PrintSpace will write XML files with nesting indentation of 2 spaces rather than 4, to match existing XML files in the GROMACS repo. Also, a cppcheck suppression was added, and fix proposed upstream that helps avoid it.

The resulting XML is identical in format to the current form in the repo, except that empty text fields are rendered with '<string></string>' rather than using the '<string/>' shorthand. Those have been regenerated.

One known limitation is that "\r" is not handled correctly in text nodes (string or text block), as it is converted to "\n". libxml2 did not handle it correctly in text block either, and we don't need this behaviour to work, so it is no longer tested for either case.

Another limitation is that a non-CDATA text element that is not empty and contains only whitespace is written correctly by TinyXML2, but is parsed to an empty element upon reading. So GROMACS cannot rely on refdata that wants to write a whitespace-only non-empty String (a TextBlock is OK), which we might want to do some time. Attempting to do so throws an exception, so we can't inadvertently run into trouble.

Some of the helper code that reports how strings in refdata do not match has been slightly improved, so that the string that doesn't match is wrapped in single quotes, so that the manner in which whitespace is not matched can be more easily seen on the terminal.

Refs #1947

Change-Id: I6153136b67b41e7141fc3f1fc8ee4005a72c90f1

History

#1 - 04/19/2016 01:40 AM - Gerrit Code Review Bot

Gerrit received a related patchset '17' for Issue #1947.
Uploader: Mark Abraham (mark.j.abraham@gmail.com)