ON THE USAGE OF XML FILE FORMAT IN GEODYNAMICS

V.Ya. CHOLIY

Kyiv Shevchenko University
01601, Ukraine, Kyiv, Glushkova str. 2
charlie@univ.kiev.ua

ABSTRACT. We recommend the extended usage of XML data format for the representation of geodynamic observations and processing results. A short introduction to the technology and a simple example for a Consolidated Laser Ranging Data Format (CRD) data file are presented.

TECHNOLOGY

XML = eXtensible Markup Language is the result of a very successful attempt to create a language for standardized text representation of the structured information in data files. XML is a metalanguage: it has no operators, it does not define the algorithms, it does not make calculations. The only purpose of XML is to explain the structure of the files, normally called XML-documents.

Two main features make XML attractive. At first, the document becomes self-explanatory. It itself explains its structure by allowing putting tags into the file. Secondly, the data stay separated from the explanations. Multiple explanations are allowed in the same file. There are two different approaches to make the file self-explanatory: Document Type Declaration (DTD) or XML-scheme. Both of them explain dependencies between data elements of the file and their attributes, value ranges, numbers of tags, etc.

To recognize the internal structure of an XML file and its semantics, specially-designed parsers are used. These are SAX and DOM parsers. To use SAX (Simple Application Programmer Interface or API for XML), the programmer should explain how to process elements of the document. When the parser loads some item, it uses that explanation to do the preprocessing. So, the parser just reacts on loading events. SAX-parsers don't create the internal document structure in memory and don't consume the memory resources. In contrary to SAX, the DOM (Document Object Model) parsers build the document tree. The document components occupy the tree leaves and branchings. The DOM-parser guarantees the access to any of the document elements with DOM-API.

It is possible to process the XML files with specially designed XSLT (eXtensible Stylesheet Language for Transformations). XSLT (written in XML) is a query and transformation language, based upon XML syntax. It is widely used for making transformation of XML files, for example to html, txt or other representations, even to binary ones. It may be used to create reports or just preprocess data, stored in an XML file.

XML is ideal for data storage. There is a tiny difference with respect to the ordinal read/write process. We parse the file during 'read' and serialize it during 'write'. With the standard XML instrumentation, there is no need to write something else. One should just prepare the correct file structure descriprion and put it into the file.

Shortly summarizing:

- the user should concentrate on the data explanation, not on file processing algorithms;
- there is no limitation on the file structure if it is representable as a tree;
- the data may be commented;
- we are able to store text, binary and classified data together in the same file;
- the existing APIs allow effective access;
- there is the possibility to add pre- or post-processing of the data in parsing or serialization.

EXAMPLE

Below, there is a portion of the example CRD file (from Ricklefs R., 2000) followed by its representation in XML syntax.

```
H1 CRD 1 2007 3 20 14
H2 MLRS 7080 24 19 4
H3 LAGEOS2 9207002 5986 22195 0 1
H4 1 2006 11 13 15 25 4 2006 11 13 15 44 40 0 0 0 0 1 0 2 0
C0 0 532.000 std1
60 std1 5 2
11 55504.9728030 0.047379676080 std1 2 120 18 94.0 -1.000 -1.000 -1.0 0.0 0
20 55504.9728030 801.80 282.10 39 1
40 55504.9728030 0 std1 -1 -1 0.000 -913.0 0.0 56.0 -1.000 -1.000 -1.0 3 3 0
11 55988.9809589 0.044893190432 std1 2 120 19 83.0 -1.000 -1.000 -1.0 0.0 0


<?xml version="1.0" encoding="utf-8"?>
<!-- <!DOCTYPE CrdFile SYSTEM "crd.dtd"> -->
<CrdFile Version="1" Year="2007" Month="3" Day="20" Hour="14">
<Station Name="MLRS" CDPIdentifier="7080" CDPTwoDigot="24" CDPOccupancy="19" TimeScale="4">
<Target Name="LAGEOS2" COSPARId="9207002" SIC="5986" NORAD="22195" TimeScale="0" Type="1">
<Session Type="1" DataReleaseFlag="0" TropoApplied="false" CenterOfMassApplied="false"
    ReceiveAmplApplied="false" StationSysDelayApplied="true"
    SpacecraftSysDelayApplied="false" RangeType="2" DataQuality="0">
<Start Year="2006" Month="11" Day="13" Hour="15" Minute="25" Second="4"/>
<Finish Year="2006" Month="11" Day="13" Hour="15" Minute="44" Second="40"/>
<SystemConfig Type="0" Wavelength="532.000" Id="std1"/>
<Compatibility Id="std1" SCH="5" SCI="2"/>
<Range Moment="55504.9728030" FlightTime="0.047379676080" Id="std1" Event="2"
    NPWindow="120" RawNumber="18" RMS="94.0"/>
<Meteo Moment="55504.9728030" Pressure="801.80" Temperature="282.10" Humidity="39"
    Origin="1"/>
<Calibration Moment="55504.9728030" Type="0" Id="std1" SystemDelay="-913.0"
    DelayShift="0.0" DelayRMS="56.0" CalibrType="3" CalibrShiftType="3" Channel="0"/>
<Range Moment="55988.9809589" FlightTime="0.044893190432" Id="std1" Event="2"
    NPWindow="120" RawNumber="19" RMS="83.0"/>
<Meteo Moment="55988.9809589" Pressure="801.50" Temperature="282.80" Humidity="39"
    Origin="1"/>
```

There is a reference to the "crd.dtd" file in the commented, second line of the XML file. It is a definition of the CRD file with XML syntax. Our recommendations for RINEX, SINEX, CRD, MERIT and some other data file formats will be collected soon on the author's web page (http://space.univ.kiev.ua/choliy/).

REFERENCES

Ricklefs R., Moore C., 2009, "Consolidated Laser Ranging Data Format CRD v.1.01".