

XML

and

***THE PARADOX OF
RUSSELL
(SET THEORY)***

By Fabio Alves

At this point in time, we can see the XML use has being mentioned as full-of-promises resource to widely offer information among internal organizations. It has its application but I do not see it as something that will bring only solutions. It can also bring some problems when it is applied. And that's why I am writing this article.

I have no doubts that XML has its benefits. Technology vendors are highly in favor of it because they do not need to open their industrial secrets among competitors.

Web Services and XML are bringing down the paradigm that browsers are the best place to host applications (It can be seen in IPODs).

And in any way I say XML invalidates its use or the use of Web Services. What I see is the need of be cautious regarding the security, performance and the instability of the applications to be built using XML as a component.

- **About Security**

Regarding security, it is clear that the XML message has information (Values) and the specification of the information (Tags). It's nothing bad for iPod's, but it is very severe when personal private information travels thru the internet. These data exposition can be considered a delicatessen for hackers.

Certificates are needed and I also see encryption as one. Encryption will be a performance issue. Certificates should not increase the complexity of applications. It delays the developing time.

- **About Performance**

Regarding performance, I can cite the increase of processor time use for packing and unpacking the messages.

It extrapolates the memory-processor binomial (Space-Time binomial), because more memory is needed for storing the information and more need of processor time to handle it.

The offset specification's elimination is not real. It is virtual. The use of XML makes the offset specification dynamic (Determined at run time) while the offset as we know is static (Determined at compiling time). The decrease in performance is inevitable.

The "well-formed XML" rules implies in two step message reading. One step is to check if the XML message is well-formed and another one to get the information.

In some tests I did for capturing information from XML strings and from HTML ones, I have noticed slow response time from the C# and java script code when handling big chunks of strings. It implies in innumerous substrings' comparisons.

At the mainframe side, in COBOL language, the STRING and UNSTRING commands are very versatile, although they highly consume processor time.

But for our hope, there is a positive example. SQL stores its information using a similar structure to XML. When a field is allowed to be "null" (It may not exist in the row) suggests that a "tag and un-tag set) may be missing. We do not need to reorganize SQL tables when we insert a new field in a row (Once it may be null). SQL keeps itself very efficient regarding it. What is the secret of SQL? I do not know and the vendors will barely open their algorithms (Industrial secret).

I see the optional use of Web Services without the XML as a need in case of high data volume messages.

- **About Instability**

The use of web services implies in encoding process. There are needs of encoding for URL parameters and also encoding for XML output message. The recursive use of encoding process might bring conversion errors when we do not how many times we should apply them.

The English language is less problematic about it because it has a smaller set of characters for being represented, but the problem is maximized with other languages.

Encoding implies also in a performance issue.

A XML is a represented by a variable-size string and the last one is defined by any chain (set) of characters ended in two binary zeros.

The XML data representation TAGS are hierarchically used for representing DATA and SETS OF DATA. The tags are built in the same way. That's where I see that problems can be caused using the Paradox of Russell as a reference.

So far I do not have a counter-example showing the instability of XML using the Paradox of Russell.

I am researching in the internet to find if anybody has raised the topic.

When I was in my graduation (1880-1990's), the "The Paradox of Russell" was showed to me in classes of Mathematics Fundamentals and Analysis. At that time, I did not give a relevant meaning to it, but it has always been in my mind as something pending to be

better understood and how it can negatively impact the applications of the beautiful and fundamental Set Theory.

To start, I would like to list the text bellow found in the internet, but related information can be found in books of Mathematics related to Algebra and Set Theory.

The discovery of this paradox over the Gottlob Frege's Set Theory, made him to establish restrictions about "Set of Set" constructions. A scientist that applies the Set Theory in different science fields must be cautious about it and also about the representation of "what is set" and "what is an element" among the objects the one is working with. The Frege's axioms needed some adjustments at this point because the discovery of the paradox. Frege was an innovator and by the time the paradox was discovered, the theory was already in large use. Nobody can deny the value and power of the Set Theory because it very naïve and understandable by most people.

1) From: http://www.gap-system.org/~history/HistTopics/Beginnings_of_set_theory.html

" ...

In 1899 [Cantor](#) discovered another paradox which arises from the "SET OF ALL SETS". What is the cardinal number of the set of all sets? Clearly it must be the greatest possible cardinal yet the cardinal of the set of all subsets of a set always has a greater cardinal than the set itself. It began to look as if the criticism of [Kronecker](#) might be at least partially right since extension of the set concept too far seemed to be producing the paradoxes. The 'ultimate' paradox was found by [Russell](#) in 1902 (and found independently by [Zermelo](#)). It simplify defined a set:

$A = \{ X \mid X \text{ is not a member of } X \}$.

[Russell](#) then asked :- *Is A an element of A?* Both the assumption that A is a member of A and A is not a member of A lead to a contradiction. The set construction itself appears to give a paradox.

[Russell](#) wrote to [Frege](#) telling him about the paradox. [Frege](#) had been near completion of his major treatise on the foundations of arithmetic. [Frege](#) added an acknowledgement to his treatise.

A scientist can hardly meet with anything more undesirable than to have the foundation give way just as the work is finished. In this position I was put by a letter from Mr. Bertrand Russell as the work was nearly through the press.

..."

2) From http://en.wikipedia.org/wiki/Russell's_paradox

“ ...

Informal presentation

Let us call a set "**abnormal**" if it is a member of itself, and "**normal**" otherwise. For example, take the set of all squares. That set is not itself a square, and therefore is not a member of the set of all squares. So it is "**normal**". On the other hand, if we take the complementary set that contains all non-squares, that set is itself not a square and so should be one of its own members. It is "**abnormal**".

Now we consider the set of all normal sets, R . Attempting to determine whether R is **normal** or **abnormal** is impossible: If R were a **normal** set, it would be contained in the set of normal sets (itself), and therefore be **abnormal**; and if it were **abnormal**, it would not be contained in the set of normal sets (itself), and therefore be **normal**. This leads to the conclusion that R is both normal and abnormal: Russell's paradox.

...”

The texts above give us Logics' references.

The use of XML in the commercial area is not a rocket science. But it surely is in the industrial area.

If I were an astronaut, I would never ride a rocket whose computers use XML to exchange messages with Earth computers.

■ ■ ■ ■ ■