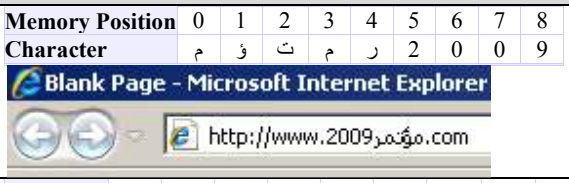
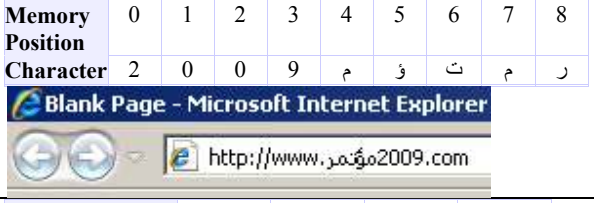
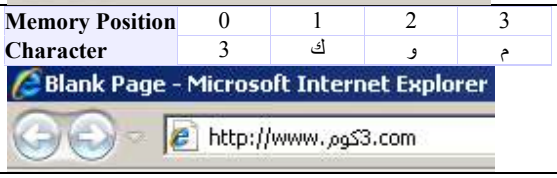
Some Concerns on the new IDNA & IDNA-Bidi Rules

Raed Al-Fayez, Abdulaziz Al-Zoman, SaudiNIC, 19 July 2009

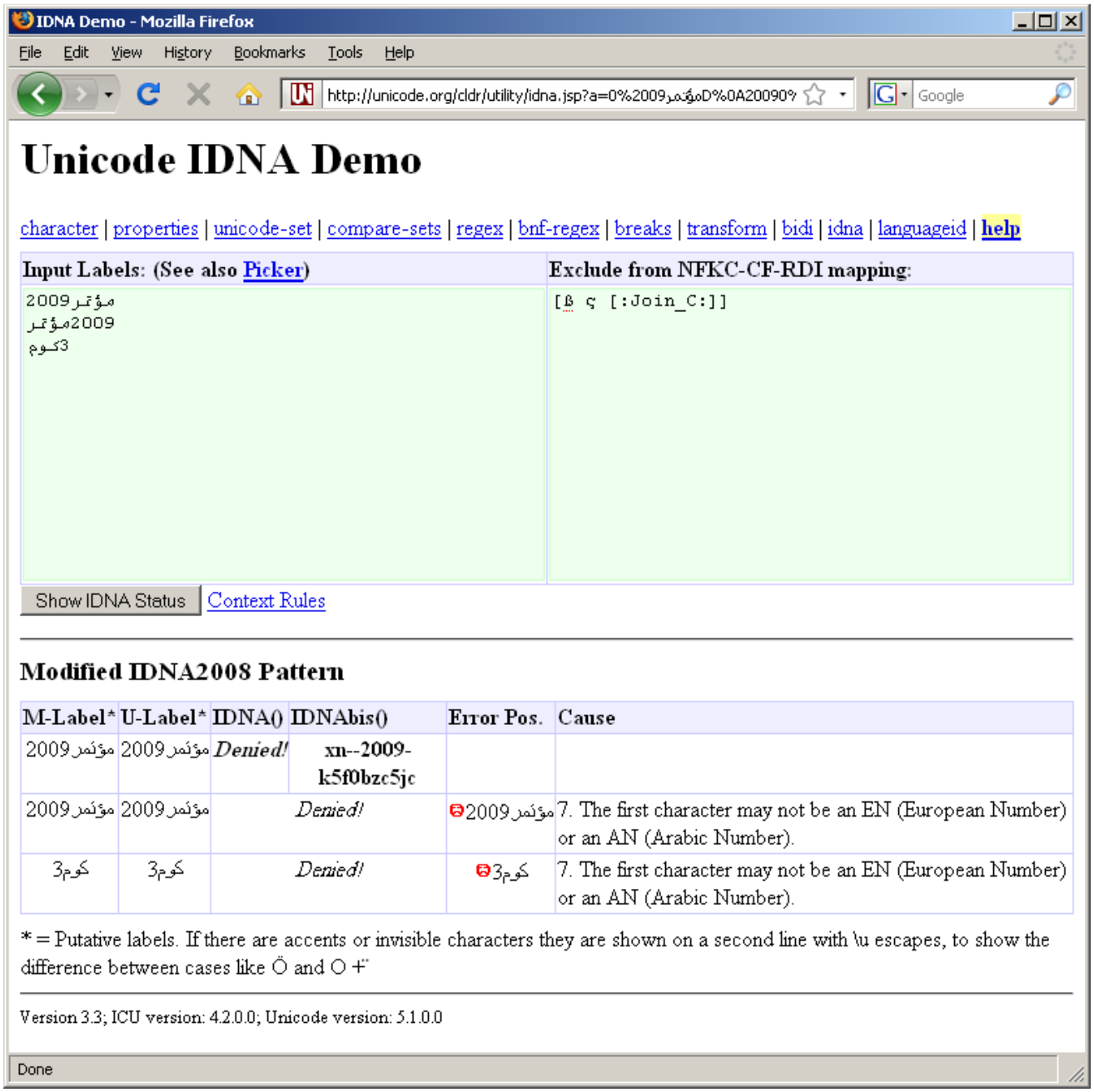
Problem description:

After testing the current IDNA implementation on this link: http://unicode.org/cldr/utility/idna.jsp, we have noticed that some of the domain names that start with a Number or ends with a number (which of course always needed in domain labels) may be blocked from protocol level because of IDNA and/or IDNA-Bidi rules and the reason given was "***The first character may not be an EN (European Number) or an AN (Arabic Number)***"!

Examples:

#	Input string (Arabic language using Arabic script)	Display	English translation
1	مؤتمر2009		Confrances2009
2	2009مؤتمر		2009Confrances
3	3كوم		3Com

The next figure shows the results of the test for the above "input strings". As it can be seen, the last two input strings (#2 and #3) are denied by the protocol.



To run the same test, please click on the following link:

http://unicode.org/cldr/utility/idna.jsp?a=%D9%85%D8%A4%D8%AA%D9%85%D8%B12009%0D%0A2009%D9%85%D8%A4%D8%AA%D9%85%D8%B1%0D%0A3%D9%83%D9%88%D9%85&f=[%C3%9F+%CF%82+[%3AJoin_C%3A]]

Recommendations:

The new IDNAbis() should be revised to allow the use of digits at the beginning of a label.