Encoding and Rendering

May 6, 2009

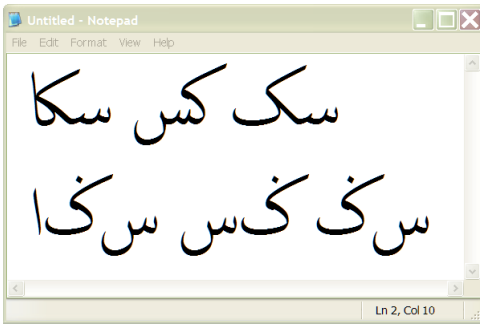
Encoding is how a character is saved and transmitted in computer. For instance, the UNICODE encoding for (alif) ا is 0627. There are many encoding exists, but today most of the computing world accept UNICODE as the standard for characters encoding.

Thus, for instance you wrote جاوي in `notepad.exe`, the computer will see this as a stream of codes: 0623 0627 0648 064*A*. It is the responsibility of the rendering engine and font technology to render it correctly as جاوي.

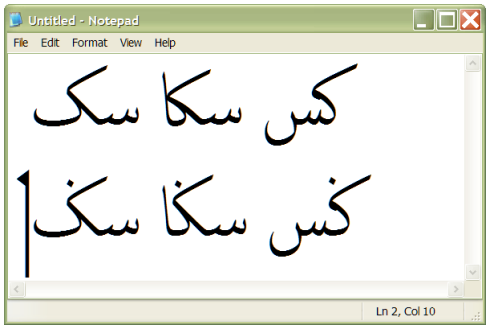
Rendering is how the characters is display on the computer. This involves 2 technology: font technology and rendering engine. On Window XP/Vista, the rendering engine is `usp10.dll`, available in `c:\Windows\system32`. The engine would render this Right-to-Left (RTL), and uses font technology to display the final form of the word. If you are having a problem in writing jawi, or the characters doesn't form correctly, you probably are having a rendering problem. Therefore, please check your `usp10.dll` version.

The current version for `usp10.dll` for Vista is 1.626.6001. The version for XP is 1.420.2600. Version 1.613.5291 and above support jawi fully. Version below 1.613.5291 may not render some characters correctly, like kaf ک and gaf ڬ.

The result from using usp10.dll version 1.1.420.2600.551 is shown below.



The same word using usp10.dll version 1.613.



In font, we have rules on how certain character changes shape due to it position in a word. For instance, the letter jim ج changes shape based on it position. So, the engine would use information from the font to render the correct shape. For instance, in the font rules for jim ج;

if jim in initial position, $0623 \rightarrow FE9F$

which mean, replace the shape ج with the shape جـ. The process is continue for the next characters until the end. Finally, the engine would render correctly جاوي.

You can change you `usp10.dll`.