Subject: [Proposal] IsDoubleWidth() implementation for double width unicode chars.
Posted by Oblivion on Mon, 18 Nov 2019 11:40:32 GMT
View Forum Message <> Reply to Message

Hello,

I propose adding the below code to charsets functions.
It simplifies Eastern-Asian/CJK character handling, and seems to cover more than IsCJKIdeograph().
Code is a rip-off from Markus Kuhns wcwidth() implementation. It is widely used, and doesn't seem to have a problematic license.
Other parts of the original code can be added as separate functions too, but while rendering double width characters, this function can be safely used.

```
bool IsDoubleWidth(int c)
{
 // This function is taken from Markus Kuhn's wcwidth implementation.
 // For license and implementation details, see:
 // https://www.cl.cam.ac.uk/~mgk25/ucs/wcwidth.c

 return c >= 0x1100
   && (c <= 0x115F        // Hangul Jamo init. consonants
    || c == 0x2329
    || c == 0x232A
    || (c >= 0x2E80 && c <= 0xA4CF && c != 0x303F)                // CJK ... Yi
    || (c >= 0xAC00 && c <= 0xD7A3)     // Hangul Syllables
    || (c >= 0xF900 && c <= 0xFAFF)     // CJK Compatibility Ideographs
    || (c >= 0xFE10 && c <= 0xFE19)     // Vertical forms
    || (c >= 0xFE30 && c <= 0xFE6F)     // CJK Compatibility Forms
    || (c >= 0xFF00 && c <= 0xFF60)     // Fullwidth Forms
    || (c >= 0xFFE0 && c <= 0xFFE6)
    || (c >= 0x20000 && c <= 0x2FFFD)
    || (c >= 0x30000 && c <= 0x3FFFD)
    );
}
```

Best regards,
Oblivion