
Subject: Re: LoadFile problem with accented chars
Posted by [koldo](#) on Wed, 11 Feb 2009 14:05:52 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello luzr

Everything solved. I inclose you some of the proposed functions:

```
String LoadFileBOM(const char *path)
{
    String s = LoadFile(path);
    if (((s[0]&0xFF) == 0xFF) && ((s[1]&0xFF) == 0xFE)) { // UTF16 Little Endian
        StringBuffer ws = s.Mid(2);
        s = ToUtf8((wchar *)ws.Begin(), ws.GetCount()*sizeof(char)/sizeof(wchar));
    } else if (((s[0]&0xFF) == 0xEF) && ((s[1]&0xFF) == 0xBB) && ((s[2]&0xFF) == 0xBF)) // UTF8
        s = s.Mid(3);
    else // May be ISO8859-1
        s = ToUtf8(ToUnicode(s, CHARSET_ISO8859_1));
    return s;
}

bool SaveBOMUtf8(Stream& out, const String& data) {
    if(!out.IsOpen() || out.IsError())
        return false;
    unsigned char bom[] = {0xEF, 0xBB, 0xBF};
    out.Put(bom, 3);
    out.Put((const char *)data, data.GetLength());
    out.Close();
    return out.IsOK();
}

bool SaveFileBOMUtf8(const char *path, const String& data)
{
    FileOut out(path);
    return SaveBOMUtf8(out, data);
}
```

When loading it checks the BOM if it is UTF-16 little endian or UTF-8. If there is no BOM it is considered to be ISO8859-1. It always return a UTF-8 String.

When saving it always save to UTF-8.

If they are right I will do the rest of functions.

There was no problem with EditString. My error was because it handles UTF-8 but not ISO8859-1 chars.

Best regards
Koldo
