Subject: How to display Traditional Chinese (Big-5)?
Posted by yoco on Tue, 20 Jun 2006 20:04:16 GMT
View Forum Message <> Reply to Message

Thanks at first.

I am learning U++ in recently days.
I found that, When I try to read a string from a text file with Traditional Chinese(Big-5),
And display it as a static text.
It does not display Chinese correctly, but only some strange symbols..

I had read the "Internationalization and translation files",
but still don't know how to display Chinese.
How can I done this work?

Thanks again.
Best wishes

Subject: Re: How to display Traditional Chinese (Big-5)?
Posted by mirek on Wed, 21 Jun 2006 04:59:06 GMT
View Forum Message <> Reply to Message

yoco wrote on Tue, 20 June 2006 16:04Thanks at first.

I am learning U++ in recently days.
I found that, When I try to read a string from a text file with Traditional Chinese(Big-5),
And display it as a static text.
It does not display Chinese correctly, but only some strange symbols..

I had read the "Internationalization and translation files",
but still don't know how to display Chinese.
How can I done this work?

Thanks again.
Best wishes

Frankly, perhaps by starting explaining us what is "big-5" and how things are supposed to work...

To put it straight, U++ is still under development and far east issues are one of areas that are not quite finished due to lack of information....

Speaking about it, do you see the right symbols in TheIDE?

Mirek

Subject: Re: How to display Traditional Chinese (Big-5)?
Posted by yoco on Fri, 23 Jun 2006 01:04:34 GMT
View Forum Message <> Reply to Message

Thank you for your reply

After read lots of documents and traced lots of source code,
Finally I display Chinese successfully.

The "Big5" is the name of a Charset Encoding,
just like "WIN1250" is a name of a charset encoding. In Taiwan,
"Big5" is the encoding that most people using.

http://en.wikipedia.org/wiki/Big5

After read the cource code,
I know Upp havn't support Big5 yet,
So I saved my text file as UTF-8,
But I still can't display Chinese correct.

After tracing into the source codes of Upp, I found that,
in the function Draw::DrawText() in Draw.cpp,
it send a CHARSET_DEFAULT to function TextUnicode() as a charset parameter,
but the charset in my file is UTF-8.
So the function TextUnicode() can't convert the UTF-8 to wchar in the right way.
And ExtTextOutW() does not get the right wchar string.

So I modify the definition of CHARSET_DEFAULT from 0 to 255 in Charset.h, and it WORKS!
The Chinese display correctly.

I don't know is that the right way to do?
or I just missed some instruction in the manual which I should known at first?
Is there are any "official" way to change the default charset?
--
Sorry about my poor English O_O

Subject: Re: How to display Traditional Chinese (Big-5)?
Posted by mirek on Fri, 23 Jun 2006 04:15:22 GMT
View Forum Message <> Reply to Message

Have you tried putting "SetDefaultCharset(CHARSET_UTF8)" at the beggining of
GUI_APP_MAIN? (I mean, without patching CHARSET_DEFAULT to 255).

Mirek

Subject: Re: How to display Traditional Chinese (Big-5)?
Posted by yoco on Fri, 23 Jun 2006 17:35:17 GMT
View Forum Message <> Reply to Message

No, I didn't.
I didn't know that I can set default charset by this function.
Thank you for tell me that
Is it on the manual already? (I mean user can set default charset through this function.)

I have another problem.
Since the upp does not support Big5,
so I decide to use UNICODE in my application.

I set default charset to CHARSET_UNICODE in the beginning of the program.
And save my text file in UNICODE,
But it display my UNICODE text fall.

==========

My code..

```
class test : public WithtestLayout<TopWindow>
{
public:
typedef test CLASSNAME;
String s ;
test()
{
FileIn fin ( "test.txt" ) ; // In unicode format
s = fin.GetLine() ;
}
virtual void Paint(Draw& w)
{
w.DrawRect(GetSize(), SWhite);
w.DrawText( 0, 0, s, Arial(16), Black);
}
};
```

==========

I found that the definition of CHARSET_UNICODE and CHARSET_UFT8 are both 255,
so in the function

```
WString ToUnicode(const char *src, int l, byte charset){
charset = ResolveCharset(charset);
if(charset == CHARSET_UTF8)
return FromUtf8(src, l);
WStringBuffer result(l);
ToUnicode(result, src, l, charset);
```

return result;
}

it always pass the string to the function FromUtf8(),
but the original string read from the file is UNICODE already.

Do I must save my text file in UTF-8?

Thanks.

Subject: Re: How to display Traditional Chinese (Big-5)?
Posted by mirek on Fri, 23 Jun 2006 17:57:49 GMT
View Forum Message <> Reply to Message

yoco wrote on Fri, 23 June 2006 13:35No, I didn't.
I didn't know that I can set default charset by this function.
Thank you for tell me that
Is it on the manual already? (I mean user can set default charset through this function.)

I have another problem.
Since the upp does not support Big5,
so I decide to use UNICODE in my application.

I set default charset to CHARSET_UNICODE in the beginning of the program.
And save my text file in UNICODE,
But it display my UNICODE text fall.

==========

My code..

```
class test : public WithtestLayout<TopWindow>
{
public:
typedef test CLASSNAME;
String s ;
test()
{
FileIn fin ( "test.txt" ) ; // In unicode format
s = fin.GetLine() ;
}
virtual void Paint(Draw& w)
{
w.DrawRect(GetSize(), SWhite);
w.DrawText( 0, 0, s, Arial(16), Black);
}
};
```

==========

I found that the definition of CHARSET_UNICODE and CHARSET_UFT8 are both 255,
so in the function

```
WString ToUnicode(const char *src, int l, byte charset){
charset = ResolveCharset(charset);
if(charset == CHARSET_UTF8)
return FromUtf8(src, l);
WStringBuffer result(l);
ToUnicode(result, src, l, charset);
return result;
}
```

it always pass the string to the function FromUtf8(),
but the original string read from the file is UNICODE already.

Do I must save my text file in UTF-8?

Thanks.

Files in U++ are considered to be the stream of bytes.

To read 16-bit unicode file, you should read individual words. Use Get16le (or Get16be for big-endian files) to read individual characters.

Of course, UTF-8 is possible and good alternative. However, while UTF-8 is great for latin alphabets, it is less ideal for chinesse - in latin languages, UTF-8 can reduce the size of file (because most characters are from basic ASCII set and therefore represented by single), while for chinesse you end with 3-byte combos.

Mirek

Subject: Re: How to display Traditional Chinese (Big-5)?
Posted by yoco on Fri, 23 Jun 2006 21:24:04 GMT
View Forum Message <> Reply to Message

Mission complete

I am so happy >_<~

I love Ultimalte++ !!

Thank you all~
