

---

Subject: Re: string filtering bug

Posted by [cbpporter](#) on Tue, 21 Jul 2009 19:44:45 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

This is not a bug, rather a design choice. String is not Utf8, it is 8 bit. You can store Utf8 in it, but U++ doesn't handle code points if you don't convert to WString, and even then only 16 bit, not true Unicode and will fail if you do real internationalization.

Generally you can store Utf8 without problems, but if you need to iterate over an Utf8 string you are left to your own devices. Utf8 (and even Utf32) are not indexable. There is no way to implement a fast []. There is a way to implement an amortized cost [], but the best way is to use an iterator which has excellent performance with the limitation that you can only go ahead or backwards in a linear fashion. Fortunately, in most cases this is what you need.

---