

Smuxi - Feature # 1047: Per channel encoding support

Status:	New	Priority:	Normal
Author:	Vsyache Puz	Category:	Engine IRC
Created:	05/02/2015	Assigned to:	Mirco Bauer
Updated:	05/02/2015	Due date:	
Complexity:	Medium		
Subject:	Per channel encoding support		
Description:	<p>Encoding should be a configurable property of channel. It should derive it's value from similar property for IRC server. And server's encoding should derive it's default value from global application setting.</p> <p>User should be able to redefine encoding at each of three levels - 1) whole application 2) IRC server 3) particular channel</p> <p>User defined configuration should persist during program restart.</p> <p>There changes to the documentation should be added if necessary on</p> <ul style="list-style-type: none">- how to set these settings- how to backup settings <p>Example of usage:</p> <p>freenode server have UTF8 encoding, but it's #altlinux channel have KOI8-R encoding</p> <p>solarnet's servers have UTF8 encoding per whole server</p> <p>.Net have good internationalization features, and can wrap stream over stream (so one can change encoding without changing code much)</p>		

History

05/02/2015 02:44 PM - Mirco Bauer

Vsyache Puz wrote:

> .Net have good internationalization features, and can wrap stream over stream (so one can change encoding without changing code much)

That is an incorrect assumption. If the first stream reads the bytes from the IRC socket in UTF-8, then you can't put another stream on top that tries to re-encode from KOI8-R because the original bytes are already gone.

05/02/2015 02:46 PM - Mirco Bauer

It only works if the "middle-encoding" preserves the original byte values which most byte into character decoder or highlevel datatype implementations like String in C# doesn't do.

05/02/2015 02:48 PM - Mirco Bauer

So if one wants this feature to work in Smuxi he would need first to refactor the IRC library (SmartIrc4net) to read and expose IRC messages as byte[] instead of a String. Then Engine-IRC could decide in which character set it wants to decode based on the channel for example.

05/02/2015 02:52 PM - Vsyache Puz

it is not necessary to expose message as byte[]. One just need to pass encoding into library, and library should do conversion inside (if library have no way to set encoding, it should be reworked)

05/02/2015 03:02 PM - Mirco Bauer

Vsyache Puz wrote:

> it is not necessary to expose message as byte[]. One just need to pass encoding into library, and library should do conversion inside (if library have no way to set encoding, it should be reworked)

The IRC library that Smuxi uses (SmartIrc4net) has an encoding setting per connection and thus the message decoding is handled for all messages

and thus all channels the same way.

05/02/2015 03:02 PM - Vsyache Puz

One need to implement channel encoding support into SmartIrc4net.

Application will set the encoding, and library will return a properly decoded string.

05/02/2015 03:03 PM - Mirco Bauer

As I am also the author of SmartIrc4net I wouldn't implement such feature myself. Decoding messages differently depending on the channel. If someone else implements this I would happily accept the contribution.