

Quassel IRC - Bug #640

iso-2022-jp encoding not honored anywhere at all.

03/21/2009 05:23 PM - salyavin

Status:	Feedback	Start date:	03/21/2009
Priority:	Immediate	Due date:	
Assignee:		% Done:	0%
Category:	Quassel Core	Estimated time:	0.00 hour
Target version:	Some future release	OS:	Any
Version:	0.9-pre		

Description

Trying out this program but I find it does not honor custom encoding. When setting up a network for example irc2.2ch.net I set custom encodings and set all three to ISO-2022-JP. When I enter the channels I see garbage characters. If I try /charset ISO-2022-JP that does not work either. This is 0.4.0. I have also tried this on other networks like say rizon then joining #japanese which is an ISO-2022-JP channel. Remember in networks I am setting custom encoding and setting all three to ISO-2022-JP. Of course konversation, xchat etc. work correctly. If I join a utf-8 Japanese channel quassel does display Japanese characters correctly but without iso-2022-jp the program is unusable for me. I am operating KDE 4.2 on 64 bit Gentoo Linux. This also applies in show channel list etc. like on irc2.2ch.net when all 3 custom encoding are set to iso-2022-jp many channel names are garbage characters (the channel names are encoded in iso-2022-jp and are Japanese characters).

History

#1 - 03/22/2009 02:25 AM - salyavin

I tried this with the gentoo ebuild (compile from source) x64 precompiled nightly binary listed on this site and gentoo ebuild from repository (did a checkout of source and compiled). I also tried connecting with a windows client on a work machine to the linux core server.

#2 - 03/22/2009 01:23 PM - EgS

Could you please tell us where we can test this?
You said something about a channel called #japanese, but on which network?

#3 - 03/22/2009 01:32 PM - EgS

Thanks to some mild poking in the eye from seezer, I managed to figure the network out ;)

#4 - 03/22/2009 03:04 PM - EgS

- Category changed from Quassel GUI (KDE) to Quassel Core
- Status changed from New to Assigned
- Assignee set to Sputnik

Reproduced the issue.

Assigning to Sput, as it seems to be an issue with the UTF-8 autodetection.
Perhaps we should disable the UTF-8 detection in some cases.

#5 - 03/22/2009 06:09 PM - salyavin

Thank you for looking at this. I am sure dealing with UTF-8 autodetection must be difficult.

BTW there are some iso-2022-jp channels on freenode like #wikipedia-ja-articles #wikinews-ja #wikipedia-ja-vandalism yet others are UTF-8 like #gentoo-ja or #nihongo irc2.2ch.net is a good test I think as it is a Japanese IRC server with some channel names, MOTD etc. in iso-2022-jp and I think all channels on that network are encoded iso-2022-jp. If it is down it is joined with irc.juggler.jp

Yes irc.rizon.net #japanese is in iso-2022-jp and has a decent amount of activity. It is joined with #japanese on irc2.2ch.net but sometimes too much English I think.

#6 - 03/26/2009 11:48 PM - salyavin

Just wanted to say I consider disabling UTF-8 autodetection as EgS suggests seems the easiest solution.

#7 - 10/18/2010 12:56 AM - Xzaws

Quassel 0.7.1; same problems. =/

#8 - 03/22/2011 01:43 AM - johu

Have a look at http://api.kde.org/4.x-api/kdelibs-apidocs/kdecore/html/classkhtml_1_1JapaneseCode.html and <https://projects.kde.org/projects/extragear/network/konversation/repository/revisions/master/show/src>

#9 - 10/07/2011 08:17 AM - khris

I agree with salyavin's opinion. It's a simple and easy solution that disabling UTF-8 auto-detection as an feature.

Guessing encoding is a over engineering for UTF-8 auto detection bugs. And multibyte encoding is not only japanese codecs but also korean's and chinese's too.

#10 - 03/26/2012 10:55 AM - rikai

Just confirming that this still exists in 0.8.0...

This really is a pretty big bug. It makes chatting in any channel using iso-2022-jp absolutely impossible.

Perhaps the priority of this bug should be bumped up, since it breaks the core function of quassel, chatting?

#11 - 01/29/2013 11:42 PM - Anonymous

- Assignee deleted (*Sputnick*)
- Priority changed from *Normal* to *Immediate*
- Version changed from *0.4.0* to *0.9-pre*
- OS changed from *Linux* to *Any*

#12 - 02/03/2013 03:00 PM - Anonymous

- Assignee set to *Anonymous*
- Target version set to *Some future release*

Fixed-ish in [PR 7](#) ; I added a blacklist of codecs which should bypass the detection. Testing showed that EUC-{JP,KR} and Shift-JIS worked fine, so I only added ISO-2022-JP to the blacklist. If others are wrongly detected, please nudge me.

#13 - 02/20/2013 07:46 PM - Anonymous

- Status changed from *Assigned* to *Resolved*
- Target version changed from *Some future release* to *0.9.0*

Merged in [4c80eeb2d07b5ca75fd399b51c939961fdff1670](#)

#14 - 06/20/2014 03:56 AM - Orbixx

- Status changed from *Resolved* to *Feedback*
- Target version changed from *0.9.0* to *Some future release*

This is still not resolved as of 0.10.0 on both client and core. Attempting to set iso-2022-jp via the network configuration UI results in it syncing with the core and then displaying default values again, disregarding the set values. Some other unusual behaviour is also exhibited with other random encoding selections, but not all of them.

Can anybody replicate this?