

Base Language - Bug #6268

Characters encoding issues with UTF-8 for emoji and confusable symbols.

04/08/2022 06:29 PM - Eugenie Lyzenko

Status:	New	Start date:	
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:		case_num:	
billable:	No		
vendor_id:	GCD		
Description			

History

#1 - 04/08/2022 06:32 PM - Eugenie Lyzenko

- File `gcd_fwd-3821c-13749__refs_6193.v1.bzr.patch` added

The patch to unlock negative ASC code conversion.

#2 - 04/08/2022 06:34 PM - Eugenie Lyzenko

- File `emojis_confusables_working_from_fwd_BUT_jvm_length_two_bytes_01.png` added

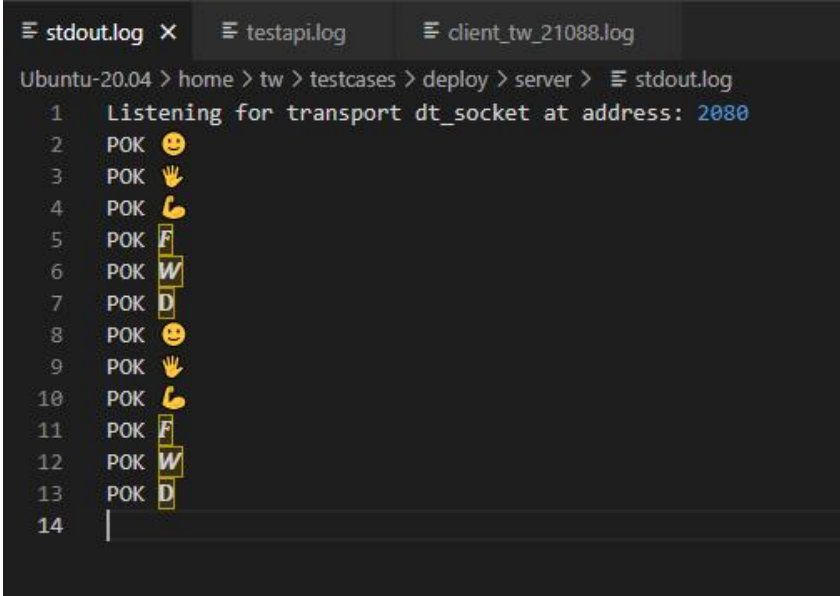
Small pic showing the end status for now. Please attach to the new bug issue along with this comment.

Using the `gcd_fwd-3821c-13749__refs_6193.v1.bzr.patch` patch, the xfer testcases pass for string equality comparison, but fail for length comparison.

Reproduction:

```
cd ~/testcases/deploy/client-chui && ./client-terminal.sh client:cmd-line-option:startup-procedure=field_io/test_fieldio.p
```

Pic:



```
≡ stdout.log × ≡ testapi.log ≡ client_tw_21088.log
Ubuntu-20.04 > home > tw > testcases > deploy > server > ≡ stdout.log
1 Listening for transport dt_socket at address: 2080
2 POK 😊
3 POK 🙌
4 POK 👍
5 POK F
6 POK W
7 POK D
8 POK 😊
9 POK 🙌
10 POK 👍
11 POK F
12 POK W
13 POK D
14 |
```

#4 - 04/09/2022 08:20 AM - Tijs Wickardt

- Start date deleted (04/08/2022)

#5 - 04/11/2022 02:44 PM - Eugenie Lyzenko

Greg,

What is our plan here? Do we need to work on this now or can be deferred?

#6 - 04/11/2022 03:32 PM - Greg Shah

Defer.

#7 - 04/12/2022 03:04 AM - Tijs Wickardt

The inconsistency regarding emoji length of the JVM versus other runtimes is nicely explained at [Emojis, Java and Strings](#)

Files

gcd_fwd-3821c-13749__refs_6193.v1.bzr.patch	1.84 KB	04/08/2022	Eugenie Lyzenko
emojis_confusables_working_from_fwd_BUT_jvm_length_two_bytes_01.png	1.0 KB	04/08/2022	Eugenie Lyzenko