

# More Regex Changes for UAX #38

John H. Jenkins

12 April 2018

In the process of synchronizing the Unihan Database with Michel's latest data for 10646, I found that three of the regular expressions for validation need to be changed:

## kIRG\_GSource:

*From:* `G(4K|BK|CH|CY|FZ|HC|HZ|((BK|CH|CY|DZ|GH|HC|RM|WZ|XC|XH|ZH)-[0-9]{4}\.[0-9]{2})|HZ-[0-9]{5}\.[0-9]{2}|(KX-[01][0-9]{3}\.1?[0-9]{2})|((CYY|FZ|JZ|ZFY|ZJW)-[0-9]{5})|([0135789ES]-[0-9A-F]{4})|(XHZ-[0-9]{3})|(PGLG-[0-9]{4})|(IDC-[0-9]{3})|(K-[0-9A-F]{4})|(OCD-\d{3})|(H-\d{4})|(H-\d{7})|(LGYJ-\d{4})|(ZYS-\d{5})|(Z-\d{4})|(Z-\d{7})|(G?F[CZ]-\d{3,6}))`

*To:* `G(4K|BK|CH|CY|FZ|HC|HZ|((BK|CH|CY|DZ|GH|HC|RM|WZ|XC|XH|ZH)-[0-9]{4}\.[0-9]{2})|HZ-[0-9]{5}\.[0-9]{2}|(KX-[01][0-9]{3}\.1?[0-9]{2})|((CYY|FZ|JZ|ZFY|ZJW)-[0-9]{5})|([0135789ES]-[0-9A-F]{4})|(XHZ-[0-9]{3})|(PGLG-[0-9]{4})|(IDC-[0-9]{3})|(K-[0-9A-F]{4})|((OCD|CE)-\d{3})|(H-\d{4})|(H-\d{7})|(LGYJ-\d{4})|(ZYS-\d{5})|(Z-\d{4})|(Z-\d{7})|(G?F[CZ]-\d{3,6}))`

## kIRG\_HSource:

*From:* `H(-[0-9A-F]{4,5}|B[012]-[0-9A-F]{4})`

*To:* `H(-[0-9A-F]{4,5}|(B[012]|D)-[0-9A-F]{4})`

## kIRG\_VSource:

*From:* `V[0-4]-[0-9A-F]{4}`

*To:* `V[0-4U]-[02]?[0-9A-F]{4}`