

# U

[:// 246. /a /1662121500551](https://doi.org/10.24646/doi/10.24646/doi/1662121500551)

- 4.0 (CC BY-SA 4.0)

# 起因

UTF-8

11110000  
10011111  
10000111  
10111000  
11110000  
10011111  
10000111  
10101010

UTF-8

UTF-8

# 初步排查

UTF-8

:// . - . / / 3629# a -5  
2

OFFSET = 0 1F1A5

- a ( ):  
(OFFSET + ( [0] ))) + (OFFSET + ( [1] )))

CN US

0 1F1E6 0 1F1FF

U

# 新的问题

a a a  
a a

U

" "

+

0 1F3FB-0 1F3FF

ZAAAAAAAAAAAAAAAAAAAAAALGGOOOOOOOOOOO!

ZALGO!

- 0 20FF 0 FE20 - 0 FE2F

0 0300 - 0 036F 0 1DC0 - 0 1DFF 0 20D0

U

:// . . /P b / /14.0/ - .

:// . . /P b / /14.0/ - - .

U 14.0  
200D 0 FEOF

0 200D

0 200D

0 FEOF  
0E 0 FEOF

0 FE00-0 FEOF

0 F

U 5.0

ba \_ a ba  
\_ a ba \_ a

:// . a . / / /

U

## 总结

UTF-8

UTF-16

U

1.

ba \_ a ba \_ a  
ba \_ a ba \_ a

2.

0 FEOF 0 0300

3.

1

1-2

0 200D

U

1.

2. U