

Linux 基础入门之实验十一简单的文本处理

学习实验十一简单的文本处理, 内容涉及常见文本处理命令和如何简单处理文本, 包括 `tr` 命令、`col` 命令、`join` 命令、`paste` 命令。

具体的学习和实验情况如下所示:

1. 常用命令

(1) `tr` 命令

删除一段文本信息中的某些文字或者将其进行转换。

`tr [option] set1 [set2]`

```
Terminal 终端 - shiyanlou@bd8742cf3391: ~
文件(F) 编辑(E) 视图(V) 终端(T) 标签(A) 帮助(H)

shiyanlou:~/ $ echo 'hello shiyanlou' | tr -d 'o','l','h' [23:23:51]
e siyanu
shiyanlou:~/ $ echo 'hello' | tr -s 'l' [23:24:59]
helo
shiyanlou:~/ $ echo 'input some text here' | tr '[:lower:]' '[:upper:]'
INPUT SOME TEXT HERE
shiyanlou:~/ $ [23:26:43]
```

(2) `col` 命令

将 `tab` 换成对等数量的空格键或者反转操作。

`co [option]`

```
shiyanlou:~/ $ cat -A /etc/protocols | tail -n 5 [23:26:43]
manet^I138^I^I^I# MANET Protocols [RFC5498]$
hip^I139^IHIP^I^I# Host Identity Protocol$
shim6^I140^IShim6^I^I# Shim6 Protocol [RFC5533]$
wesp^I141^IWESP^I^I# Wrapped Encapsulating Security Payload$
rohc^I142^IROHC^I^I# Robust Header Compression$
shiyanlou:~/ $ cat /etc/protocols | col -x | cat -A | tail -n 5 [23:30:44]
manet 138 # MANET Protocols [RFC5498]$
hip 139 HIP # Host Identity Protocol$
shim6 140 Shim6 # Shim6 Protocol [RFC5533]$
wesp 141 WESP # Wrapped Encapsulating Security Payload$
rohc 142 ROHC # Robust Header Compression$
shiyanlou:~/ $ [23:31:33]
```

(3) `join` 命令

将两个文件中包含相同内容的那一行合并在一起。

`join [option] file1 file2`

```
shiyanlou:~/ $ cd /home/shiyanlou
shiyanlou:~/ $ echo '1 hello' > file1
shiyanlou:~/ $ echo '1 shiyanlou' > file2
shiyanlou:~/ $ join file1 file2
1 hello shiyanlou
shiyanlou:~/ $ sudo join -t ' ' /etc/passwd /etc/shadow
root:x:0:0:root:/root:/bin/bash:*:17590:0:99999:7:::
daemon:x:1:1:daemon:/usr/sbin:/usr/sbin/nologin:*:17590:0:99999:7:::
bin:x:2:2:bin:/bin:/usr/sbin/nologin:*:17590:0:99999:7:::
sys:x:3:3:sys:/dev:/usr/sbin/nologin:*:17590:0:99999:7:::
sync:x:4:65534:sync:/bin:/bin/sync:*:17590:0:99999:7:::
games:x:5:60:games:/usr/games:/usr/sbin/nologin:*:17590:0:99999:7:::
man:x:6:12:man:/var/cache/man:/usr/sbin/nologin:*:17590:0:99999:7:::
lp:x:7:7:lp:/var/spool/lpd:/usr/sbin/nologin:*:17590:0:99999:7:::
mail:x:8:8:mail:/var/mail:/usr/sbin/nologin:*:17590:0:99999:7:::
news:x:9:9:news:/var/spool/news:/usr/sbin/nologin:*:17590:0:99999:7:::
uucp:x:10:10:uucp:/var/spool/uucp:/usr/sbin/nologin:*:17590:0:99999:7:::
proxy:x:13:13:proxy:/bin:/usr/sbin/nologin:*:17590:0:99999:7:::
www-data:x:33:33:www-data:/var/www:/usr/sbin/nologin:*:17590:0:99999:7:::
backup:x:34:34:backup:/var/backups:/usr/sbin/nologin:*:17590:0:99999:7:::
list:x:38:38:Mailing List Manager:/var/list:/usr/sbin/nologin:*:17590:0:99999:7:::
irc:x:39:39:irc:/var/run/ircd:/usr/sbin/nologin:*:17590:0:99999:7:::
gnats:x:41:41:Gnats Bug-Reporting System (admin)/var/lib/gnats:/usr/sbin/nologin:*:17590:0:99999:7:::
nobody:x:65534:65534:nobody:/nonexistent:/usr/sbin/nologin:*:17590:0:99999:7:::
systemd-timesync:x:100:102:systemd Time Synchronization,,:/run/systemd:/bin/false:*:17590:0:99999:7:::
systemd-networkd:x:101:103:systemd Network Management,,:/run/systemd/netif:/bin/false:*:17590:0:99999:7:::
systemd-resolved:x:102:104:systemd Resolver,,:/run/systemd/resolve:/bin/false:*:17590:0:99999:7:::
systemd-bus-proxy:x:103:105:systemd Bus Proxy,,:/run/systemd:/bin/false:*:17590:0:99999:7:::
apt:x:104:65534:/nonexistent:/bin/false:*:17590:0:99999:7:::
messagebus:x:105:108:/var/run/dbus:/bin/false:*:17691:0:99999:7:::
colord:x:106:112:colord colour management daemon,,/var/lib/colord:/bin/false:*:17691:0:99999:7:::
sshd:x:107:65534:/var/run/sshd:/usr/sbin/nologin:*:17691:0:99999:7:::
pulse:x:108:113:PulseAudio daemon,,/var/run/pulse:/bin/false:*:17691:0:99999:7:::
rtkit:x:109:115:RealtimeKit,,/proc:/bin/false:*:17691:0:99999:7:::
usbmux:x:110:46:usbmux daemon,,/var/lib/usbmux:/bin/false:*:17691:0:99999:7:::
shiyanlou:x:5000:5000:/home/shiyanlou:/usr/bin/zsh:$6i6np2chc$8Gcupf08Fcaoc/N7IL7BoJYJ1RbxjUdJcVfZCdsyJSZbSj7MoBzkJPwtN6HnbaV1RgkvIavbSCs
xQjSjly0:18050:0:99999:7:::
labex:x:6000:6000:/home/labex:/usr/bin/zsh:$6i6np2chc$8Gcupf08Fcaoc/N7IL7BoJYJ1RbxjUdJcVfZCdsyJSZbSj7MoBzkJPwtN6HnbaV1RgkvIavbSCs
:17691:0:99999:7:::
```

```

shiyanolou:~/$ sudo join -t ':' -1 4 /etc/passwd -2 3 /etc/group
0:root:x:0:root:/root:/bin/bash:root:x:
1:daemon:x:1:daemon:/usr/sbin:/usr/sbin/nologin:daemon:x:
2:bin:x:2:bin:/bin:/usr/sbin/nologin:bin:x:
3:sys:x:3:sys:/dev:/usr/sbin/nologin:sys:x:
join: /etc/passwd:6: is not sorted: games:x:5:60:games:/usr/games:/usr/sbin/nologin
7:lp:x:7:lp:/var/spool/lpd:/usr/sbin/nologin:lp:x:
8:mail:x:8:mail:/var/mail:/usr/sbin/nologin:mail:x:
join: /etc/group:11: is not sorted: uucp:x:10:
9:news:x:9:news:/var/spool/news:/usr/sbin/nologin:news:x:
10:uucp:x:10:uucp:/var/spool/uucp:/usr/sbin/nologin:uucp:x:
13:proxy:x:13:proxy:/bin:/usr/sbin/nologin:proxy:x:
33:www-data:x:33:www-data:/var/www:/usr/sbin/nologin:www-data:x:
34:backup:x:34:backup:/var/backups:/usr/sbin/nologin:backup:x:
38:list:x:38:Mailing List Manager:/var/list:/usr/sbin/nologin:list:x:
39:irc:x:39:ircd:/var/run/ircd:/usr/sbin/nologin:irc:x:
41:gnats:x:41:Gnats Bug-Reporting System (admin):/var/lib/gnats:/usr/sbin/nologin:gnats:x:
65534:nobody:x:65534:nobody:/nonexistent:/usr/sbin/nologin:nogroup:x:
102:systemd-timesync:x:100:systemd Time Synchronization,,,:/run/systemd:/bin/false:systemd-timesync:x:
103:systemd-network:x:101:systemd Network Management,,,:/run/systemd/netif:/bin/false:systemd-network:x:
104:systemd-resolve:x:102:systemd Resolver,,,:/run/systemd/resolve:/bin/false:systemd-resolve:x:
105:systemd-bus-proxy:x:103:systemd Bus Proxy,,,:/run/systemd:/bin/false:systemd-bus-proxy:x:

```

(4)paste 命令

与 join 命令类似，在不对比数据的情况下将多个文件合并在一起，并以 tab 键隔开。

paste [option] file...

```

shiyanolou:~/$ echo hello > file1
shiyanolou:~/$ echo shiyanolou > file2
shiyanolou:~/$ echo www.shiyanolou.com > file3
zsh: command not found: edho
shiyanolou:~/$ echo www.shiyanolou.com > file3
shiyanolou:~/$ paste -d ':' file1 file2 file3
hello:shiyanolou:www.shiyanolou.com
shiyanolou:~/$ paste -s file1 file2 file3
hello
shiyanolou
www.shiyanolou.com
shiyanolou:~/$

```

2.作业

```

shiyanolou:~/$ file shiyan.txt
shiyan.txt: cannot open 'shiyan.txt' (No such file or directory)
shiyanolou:~/$ touch file shiyan.txt
shiyanolou:~/$ file shiyan.txt
shiyan.txt: empty
shiyanolou:~/$ tr -d '\r' <shiyan.txt> ushiyan.txt
shiyanolou:~/$ file ushiyan.txt
ushiyan.txt: empty
shiyanolou:~/$

```