

comm

按行比较两个已排序的文件。

概要

```
comm [OPTION]... FILE1 FILE2
```

主要用途

- 按行比较两个已排序的文件。
- 当FILE1或FILE2为-时，读取标准输入。
- 无选项时输出三列，第一列为FILE1独有的行，第二列为FILE2独有的行，第三列为FILE1与FILE2共有的行。

选项

-1	不输出第一列。
-2	不输出第二列。
-3	不输出第三列。
--check-order	检查输入行是否正确的排序，即使它们确实是已排序过的。
--nocheck-order	不检查输入行是否正确的排序。
--output-delimiter=STR	使用STR作为输出列之间的分隔符而不是默认的分隔符。
--total	额外地增加第四列输出概要。
-z, --zero-terminated	设置行终止符为NUL（空），而不是换行符。
--help	显示帮助信息并退出。
--version	显示版本信息并退出。

返回值

返回0表示成功，返回非0值表示失败。

例子

文本 aaa.txt 内容

```
[root@localhost text]# cat aaa.txt
aaa
bbb
ccc
ddd
eee
111
222
```

文本 bbb.txt 内容

```
[root@localhost text]# cat bbb.txt
bbb
ccc
```

```
aaa
hhh
ttt
jjj
```

比较结果

```
[root@localhost text]# comm --nocheck-order aaa.txt bbb.txt
aaa
      bbb
      ccc
    aaa
ddd
eee
111
222
      hhh
      ttt
      jjj
```

输出的第一列只包含在aaa.txt中出现的行，第二列包含在bbb.txt中出现的行，第三列包含在aaa.txt和bbb.txt中相同的行。各列之间以制表符\t作为分隔符。

比较排序过的文档

先通过 sort 将文件内容排序：

```
[root@localhost ~]# sort aaa.txt > aaal.txt
[root@localhost ~]# sort bbb.txt > bbb1.txt
```

比较结果：

```
[root@localhost ~]# comm aaal.txt bbb1.txt
111
222
      aaa
      bbb
      ccc
ddd
eee
      hhh
      jjj
      ttt
```

交集

打印两个文件的交集，需要删除第一列和第二列：

```
[root@localhost text]# comm aaa.txt bbb.txt -1 -2
bbb
ccc
```

差集

通过删除不需要的列，可以得到aaa.txt和bbb.txt的差集：

aaa.txt的差集

```
[root@localhost text]# comm aaa.txt bbb.txt -2 -3
aaa
ddd
eee
111
222
```

bbb.txt的差集

```
[root@localhost text]# comm aaa.txt bbb.txt -1 -3
aaa
hhh
ttt
jjj
```

注意

1. 该命令是GNU coreutils包中的命令，相关的帮助信息请查看`man -s 1 comm`或`info coreutils 'comm invocation'`

From:

<https://rd.irust.top/> - 学习笔记

Permanent link:

<https://rd.irust.top/doku.php?id=command:comm>

Last update: **2021/10/15 14:58**

