

sort

对文本文件中所有行进行排序。

概要

```
sort [OPTION]... [FILE]...
sort [OPTION]... --files0-from=F
```

主要用途

- 将所有输入文件的内容排序后并输出。
- 当没有文件或文件为 - 时，读取标准输入。

选项

排序选项：

-b, --ignore-leading-blanks	忽略开头的空白。
-d, --dictionary-order	仅考虑空白、字母、数字。
-f, --ignore-case	将小写字母作为大写字母考虑。
-g, --general-numeric-sort	根据数字排序。
-i, --ignore-nonprinting	排除不可打印字符。
-M, --month-sort	按照非月份、一月、十二月的顺序排序。
-h, --human-numeric-sort	根据存储容量排序(注意使用大写字母，例如 2K 1G)
-n, --numeric-sort	根据数字排序。
-R, --random-sort	随机排序，但分组相同的行。
--random-source=FILE	从FILE中获取随机长度的字节。
-r, --reverse	将结果倒序排列。
--sort=WORD	根据WORD排序，其中：general-numeric 等价于 -g
human-numeric 等价于 -h	month 等价于 -M
numeric 等价于 -n	random 等价于 -R
version 等价于 -V	
-V, --version-sort	文本中(版本)数字的自然排序。

其他选项：

--batch-size=NMERGE	一次合并最多NMERGE个输入；超过部分使用临时文件。
-c, --check, --check=diagnose-first	检查输入是否已排序，该操作不会执行排序。
-C, --check=quiet, --check=silent	类似于 -c 选项，但不输出第一个未排序的行。
--compress-program=PROG	使用PROG压缩临时文件；使用PROG -d解压缩。
--debug	注释用于排序的行，发送可疑用法的警报
到stderr	
--files0-from=F	从文件F中读取以NUL结尾的所有文件名称；如果F是 - ，那么从标准输入中读取名字。
-k, --key=KEYDEF	通过一个key排序KEYDEF给出位置和类型。
-m, --merge	合并已排序文件，之后不再排序。
-o, --output=FILE	将结果写入FILE而不是标准输出。
-s, --stable	通过禁用最后的比较来稳定排序。
-S, --buffer-size=SIZE	使用SIZE作为内存缓存大小。

-t, --field-separator=SEP	使用SEP作为列的分隔符。
-T, --temporary-directory=DIR	使用DIR作为临时目录，而不是 \$TMPDIR 或 /tmp。多次使用该选项指定多个临时目录。
--parallel=N	将并发运行的排序数更改为N。
-u, --unique	同时使用 -c 严格检查排序；不同时使用 -c 输出排序后去重的结果。
-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。

KEYDEF的格式为[F[.C][OPTS]][,F[.C][OPTS]] 表示开始到结束的位置。

F表示列的编号

C表示

OPTS为[bdfgiMhnRrV]中的一到多个字符，用于覆盖当前排序选项。

使用 --debug 选项可诊断出错误的用法。

SIZE 可以有以下的乘法后缀：

% 内存的1%；

b 1

K 1024（默认）；

剩余的 M, G, T, P, E, Z, Y 可以类推出来。

参数

FILE（可选）：要处理的文件，可以为任意数量。

返回值

返回0表示成功，返回非0值表示失败。

例子

sort将文件/文本的每一行作为一个单位相互比较，比较原则是从首字符向后依次按ASCII码值进行比较，最后将他们按升序输出。

```
root@mail text]# cat sort.txt
aaa:10:1.1
ccc:30:3.3
ddd:40:4.4
bbb:20:2.2
eee:50:5.5
eee:50:5.5

[root@mail text]# sort sort.txt
aaa:10:1.1
bbb:20:2.2
ccc:30:3.3
ddd:40:4.4
eee:50:5.5
```

```
eee:50:5.5
```

忽略相同行使用 -u 选项或者 uniq

```
[root@mail text]# cat sort.txt
aaa:10:1.1
ccc:30:3.3
ddd:40:4.4
bbb:20:2.2
eee:50:5.5
eee:50:5.5

[root@mail text]# sort -u sort.txt
aaa:10:1.1
bbb:20:2.2
ccc:30:3.3
ddd:40:4.4
eee:50:5.5

[root@mail text]# uniq sort.txt
aaa:10:1.1
ccc:30:3.3
ddd:40:4.4
bbb:20:2.2
eee:50:5.5
```

sort 的 -n -r -k -t 选项的使用：

```
[root@mail text]# cat sort.txt
AAA:BB:CC
aaa:30:1.6
ccc:50:3.3
ddd:20:4.2
bbb:10:2.5
eee:40:5.4
eee:60:5.1

# 将BB列按照数字从小到大顺序排列：
[root@mail text]# sort -nk 2 -t: sort.txt
AAA:BB:CC
bbb:10:2.5
ddd:20:4.2
aaa:30:1.6
eee:40:5.4
ccc:50:3.3
eee:60:5.1

# 将CC列数字从大到小顺序排列：
# -n 是按照数字大小排序 -r 是以相反顺序 -k 是指定需要排序的栏位 -t 指定栏位分隔符为冒号
[root@mail text]# sort -nrk 3 -t: sort.txt
eee:40:5.4
```

```
eee:60:5.1
ddd:20:4.2
ccc:50:3.3
bbb:10:2.5
aaa:30:1.6
AAA:BB:CC
```

关于 -k 选项的解读和例子：

-k 选项深度解读：

```
FStart.CStart Modifier,FEnd.CEnd Modifier
-----Start-----,-----End-----
FStart.CStart 选项 , FEnd.CEnd 选项
```

这个语法格式可以被其中的逗号,分为两大部分, **Start** 部分和 **End** 部分。Start部分由三部分组成, 其中的Modifier部分就是我们之前说过的选项部分; 我们重点说说Start部分的FStart和C.Start。C.Start是可以省略的, 省略的话就表示从本域的开头部分开始。FStart.CStart, 其中FStart就是表示使用的域, 而CStart则表示在FStart域中从第几个字符开始算排序首字符。同理, 在End部分中, 你可以设定FEnd.CEnd, 如果你省略.CEnd或将它设定为0, 则表示结尾到本域的最后一个字符。

例子：从公司英文名称的第二个字母开始排序：

```
$ sort -t ' ' -k 1.2 facebook.txt
baidu 100 5000
sohu 100 4500
google 110 5000
guge 50 3000
```

解读：使用了 -k 1.2, 表示对第一个域的第二个字符开始到本域的最后一个字符为止的字符串进行排序。你会发现baidu因为第二个字母是a而名列榜首,sohu和google第二个字符都是o,但sohu的h在google的o前面,所以两者分别排在第二和第三,guge只能屈居第四了。

例子：只针对公司英文名称的第二个字母进行排序, 如果相同的按照员工工资进行降序排序：

```
$ sort -t ' ' -k 1.2,1.2 -nrk 3,3 facebook.txt
baidu 100 5000
google 110 5000
sohu 100 4500
guge 50 3000
```

解读：由于只对第二个字母进行排序, 所以我们使用了 -k 1.2,1.2 的表示方式, 表示我们只对第二个字母进行排序 (如果你问我使用 -k 1.2 怎么不行? 当然不行, 因为你省略了End部分, 这就意味着你将对从第二个字母起到本域最后一个字符为止的字符串进行排序)。对员工工资进行排序, 我们也使用了 -k 3,3, 这是最准确的表述, 表示我们只对本域进行排序, 因为如果你省略了后面的3, 就变成了我们对第3个域开始到最后一个域位置的内容进行排序了。

注意

1. [关于-g和-n选项的区别](#)[stackoverflow](#)

2. 关于这个复杂命令的学习，建议您阅读info文档及参考博客、问答网站等。
3. 该命令是GNU coreutils包中的命令，相关的帮助信息请查看`man -s 1 shuf`或`info coreutils 'shuf invocation'`

From:

<https://rd.irust.top/> - 学习笔记

Permanent link:

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

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

