NAME

uniq - report or filter out repeated lines in a file

SYNOPSIS

uniq [**-c** | **-d** | **-D** | **-u**] [**-i**] [**-f** *num*] [**-s** *chars*] [*input_file* [*output_file*]]

DESCRIPTION

The **uniq** utility reads the specified *input_file* comparing adjacent lines, and writes a copy of each unique input line to the *output_file*. If *input_file* is a single dash (‘**-**’) or absent, the standard input is read. If *output_file* is absent, standard output is used for output. The second and succeeding copies of identical adjacent input lines are not written. Repeated lines in the input will not be detected if they are not adjacent, so it may be necessary to sort the files first.

The following options are available:

-c, **--count**

Precede each output line with the count of the number of times the line occurred in the input, followed by a single space.

-d, **--repeated**

Output a single copy of each line that is repeated in the input.

-D, **--all-repeated** [*septype*]

Output all lines that are repeated (like **-d**, but each copy of the repeated line is written). The optional *septype* argument controls how to separate groups of repeated lines in the output; it must be one of the following values:

none	Do not separate groups of lines (this is the default).
prepend	Output an empty line before each group of lines.
separate	Output an empty line after each group of lines.

-f *num*, **--skip-fields** *num*

Ignore the first *num* fields in each input line when doing comparisons. A field is a string of non-blank characters separated from adjacent fields by blanks. Field numbers are one based, i.e., the first field is field one.

-i, **--ignore-case**

Case insensitive comparison of lines.

-s *chars*, **--skip-chars** *chars*

Ignore the first *chars* characters in each input line when doing comparisons. If specified in conjunction with the **-f**, **--unique** option, the first *chars* characters after the first *num* fields will be ignored. Character numbers are one based, i.e., the first character is character one.

-u, **--unique**
Only output lines that are not repeated in the input.

ENVIRONMENT

The LANG, LC_ALL, LC_COLLATE and LC_CTYPE environment variables affect the execution of **uniq** as described in environ(7).

EXIT STATUS

The **uniq** utility exits 0 on success, and >0 if an error occurs.

EXAMPLES

Assuming a file named cities.txt with the following content:

```
Madrid
Lisbon
Madrid
```

The following command reports three different lines since identical elements are not adjacent:

```
$ uniq -u cities.txt
Madrid
Lisbon
Madrid
```

Sort the file and count the number of identical lines:

```
$ sort cities.txt | uniq -c
      1 Lisbon
      2 Madrid
```

Assuming the following content for the file cities.txt:

```
madrid
Madrid
Lisbon
```

Show repeated lines ignoring case sensitiveness:

```
$ uniq -d -i cities.txt
madrid
```

Same as above but showing the whole group of repeated lines:

```
$ uniq -D -i cities.txt
madrid
Madrid
```

Report the number of identical lines ignoring the first character of every line:

```
$ uniq -s 1 -c cities.txt
      2 madrid
      1 Lisbon
```

COMPATIBILITY

The historic **+***number* and **-***number* options have been deprecated but are still supported in this implementation.

SEE ALSO

sort(1)

STANDARDS

The **uniq** utility conforms to IEEE Std 1003.1-2001 ("POSIX.1") as amended by Cor. 1-2002.

HISTORY

A **uniq** command appeared in Version 3 AT&T UNIX.