NAME join – relational database operator

SYNOPSIS **join** [–**a** *filenumber* | –**v** *filenumber*] [–**1** *fieldnumber*] [–**2** *fieldnumber*]
[–**o** *list*] [–**e** *string*] [–**t** *char*] *file1 file2*

join [–**a** *filenumber*] [–**j** *fieldnumber*] [–**j1** *fieldnumber*] [–**j2** *fieldnumber*]
[–**o** *list*] [–**e** *string*] [–**t** *char*] *file1 file2*

AVAILABILITY SUNWcsu

DESCRIPTION The **join** command forms, on the standard output, a join of the two relations specified by the lines of *file1* and *file2*.

There is one line in the output for each pair of lines in *file1* and *file2* that have identical join fields. The output line normally consists of the common field, then the rest of the line from *file1*, then the rest of the line from *file2*. This format can be changed by using the –**o** option (see below). The –**a** option can be used to add unmatched lines to the output. The –**v** option can be used to output only unmatched lines.

The default input field separators are blank, tab, or new-line. In this case, multiple separators count as one field separator, and leading separators are ignored. The default output field separator is a blank.

If the input files are not in the appropriate collating sequence, the results are unspecified.

OPTIONS Some of the options below use the argument *filenumber*. This argument should be a **1** or a **2** referring to either *file1* or *file2*, respectively.

–**a** *filenumber* In addition to the normal output, produce a line for each unpairable line in file *filenumber*, where *filenumber* is **1** or **2**. If both –**a 1** and –**a 2** are specified, all unpairable lines will be output.

–**e** *string* Replace empty output fields with *string*.

–**j** *fieldnumber* Equivalent to –**1** *fieldnumber* –**2** *fieldnumber*.

–**j1** *fieldnumber* Equivalent to –**1** *fieldnumber*.

–**j2** *fieldnumber* Equivalent to –**2** *fieldnumber* Fields are numbered starting with **1**.

–**o** *list* Each output line includes the fields specified in *list*. Fields selected by *list* that do not appear in the input will be treated as empty output fields. (See the –**e** option.) Each element of which has the either the form *filenumber*.*fieldnumber*, or **0**, which represents the **join** field. The common field is not printed unless specifically requested.

–**t** *char* Use character *char* as a separator. Every appearance of *char* in a line is significant. The character *char* is used as the field separator for both input and output. With this option specified, the collating term should be the same as **sort** without the –**b** option.

–**v** *filenumber* Instead of the default output, produce a line only for each unpairable line in *filenumber*, where *filenumber* is **1** or **2**. If both –**v 1** and –**v 2** are specified, all unpairable lines will be output.

−1 *fieldnumber* Join on the *fieldnumber*th field of file 1 . Fields are decimal integers starting with **1**.

−2 *fieldnumber* Join on the *fieldnumber*th field of file 2. Fields are decimal integers starting with **1**.

OPERANDS

The following operands are supported:

file1
file2 A path name of a file to be joined. If either of the *file1* or *file2* operands is −, the standard input is used in its place.

file1 and *file2* must be sorted in increasing collating sequence as determined by **LC_COLLATE** on the fields on which they are to be joined, normally the first in each line (see **sort**(1)).

EXAMPLES

The following command line will join the password file and the group file, matching on the numeric group ID, and outputting the login name, the group name and the login directory. It is assumed that the files have been sorted in ASCII collating sequence on the group ID fields.

example% join −j1 4 −j2 3 −o 1.1 2.1 1.6 −t: /etc/passwd /etc/group

The −**o 0** field essentially selects the union of the join fields. For example, given file **phone**:

!Name	**Phone Number**
Don	**+1 123-456-7890**
Hal	**+1 234-567-8901**
Yasushi	**+2 345-678-9012**

and file **fax**:

!Name	**Fax Number**
Don	**+1 123-456-7899**
Keith	**+1 456-789-0122**
Yasushi	**+2 345-678-9011**

(where the large expanses of white space are meant to each represent a single tab character), the command:

example% join -t "<*tab*>" -a 1 -a 2 -e ′(unknown)′ -o 0,1.2,2.2 phone fax

would produce:

!Name	**Phone Number**	**Fax Number**
Don	**+1 123-456-7890**	**+1 123-456-7899**
Hal	**+1 234-567-8901**	**(unknown)**
Keith	**(unknown)**	**+1 456-789-0122**
Yasushi	**+2 345-678-9012**	**+2 345-678-9011**

ENVIRONMENT

See **environ**(5) for descriptions of the following environment variables that affect the execution of **join**: **LC_CTYPE**, **LC_MESSAGES**, and **NLSPATH**.

EXIT STATUS

The following exit values are returned:

0 All input files were output successfully.

>0 An error occurred.

SEE ALSO

awk(1), **comm**(1), **sort**(1), **uniq**(1), **environ**(5)

NOTES

With default field separation, the collating sequence is that of **sort −b**; with **−t**, the sequence is that of a plain sort.

The conventions of the **join**, **sort**, **comm**, **uniq**, and **awk** commands are wildly incongruous.