

---

## NAME

**extract** — get specified spines from a Humdrum input

## SYNOPSIS

```
extract -f field1 field2,...,$-1,$ [inputfile ...]  
extract -i interp1,interp2,...,interpn [inputfile ...]  
extract -p spine#n [inputfile ...]  
extract -t field_trace_file.ftf [inputfile ...]
```

## DESCRIPTION

The **extract** command allows the user to select one or more spines from a Humdrum input. The command is typically used to extract parts (such as a tuba part) from some multi-part score; however **extract** can be used to isolate dynamic markings, musical lyrics, or any other stream of information that has been encoded as a separate Humdrum spine.

The **extract** command has four distinct modes of operation: (1) field mode (-f), (2) interpretation mode (-i), (3) spine-path mode (-p), and (4) field-trace mode (-t). The simplest mode is field mode; the most general-purpose mode is field-trace mode.

In *field mode*, the **extract** command operates in a manner similar to the UNIX `cut` command. The user may specify a given field (data column) or set of fields to be selected from the input stream. For example, the command:

```
extract -f 1,3,8
```

will extract the first, third, and eighth spines from the input stream. In field mode, field specifications may also be made with respect to the right-most field. For example, the expression "\$" refers to the right-most field in the input; the arithmetic expression "\$-1" refers to the right-most field minus one, etc. By way of example, the command:

```
extract -f '2,4-6,$' lassung
```

will extract the second, fourth, fifth, sixth, and last (right-most) spines in the file "lassung." The `extract -f` command differs from the UNIX `cut` command in that Humdrum global comments are properly preserved in the output.

In *interpretation mode*, the **extract** command outputs all spines containing the interpretation(s) specified by the user. By way of example, the command:

```
extract -i '**semits,**MIDI' hildegard
```

will extract all spines in the file `hildegard` containing `**semits` or `**MIDI` data.

In *spine-path mode*, the **extract** command follows a given spine starting at the beginning of the file, and traces the course of that spine throughout the input stream. If spine-path changes are encountered in the input (such as spine-exchanges, spine-merges, or spine-splits) the output adapts accordingly. By way of example, if the “third” spine is selected, the output consists of the third spine and follows the path of that spine through the input until it is terminated or the end-of-file is encountered. What begins as the third column, may end up as some other column (or columns) in the input.

In the *field-trace mode*, the **extract** command accepts a list of spine-column positions over the course of the file. In the *field-trace mode*, the user provides a file containing a list specifying the precise selection of spines through the file at various line numbers. The first column in this list specifies the line (record) number at which the field selection is defined. The second column lists the spine or spines to be extracted beginning at the specified line. For example, the following *field-trace file* specifies that the first spine from line 1 should be output; beginning at line 2, spines 3, 4, and 5 should be output; beginning at line 18, spines 1 and 4 should be output, and from line 78 to the end of the input, spine 9 should be output.

```
1      1
2      3-5
18     1, 4
78     9
```

The *field-trace* mode allows the user to select virtually any combination of data tokens from the input stream. Note that using the *field-trace* mode may produce output that no longer conforms to the Humdrum syntax. (See EXAMPLES below.)

## OPTIONS

The **extract** command provides the following options:

<b>-f</b> <i>fieldlist</i>	select field mode
<b>-h</b>	displays a help screen summarizing the command syntax
<b>-i</b> <i>interplist</i>	select interpretation mode
<b>-p</b> <i>spine#</i>	select spine-path mode
<b>-t</b> <i>fieldtrace file</i>	select field-trace mode

Options are specified in the command line.

A *fieldlist* consists of any set of integers separated by commas, or a range specification in which the lower and upper values are separated by a dash: e.g. 1, 4-8, 13. With the exception of range specifications, the order of the integers in the list is unimportant, hence 13, 1, 4-8 is equivalent to 1, 4-8, 13. The right-field anchor (\$) can be used only with the **-f** option.

An *interplist* consists of any set of tandem or exclusive interpretations separated by commas, e.g. '\*\*kern, \*C:'. Option arguments may require the use of quotation marks in order to prevent inadvertent expansion of shell metacharacter (\*, ?, etc.).

**EXAMPLES**

The following examples illustrate how **extract** may be used.

```
extract -f '1,3,$' holst
```

outputs the first, third, and last columns from the file `holst`.

```
extract -p 4 mossolov
```

outputs the spine *beginning* (but not necessarily continuing) in the fourth column of the file `mossolov`.

```
extract -t sibelius.fld sibelius
```

outputs the spines specified in the file `sibelius.fld` for the file `sibelius`. (See SAMPLE OUTPUTS below.)

```
extract -i '*F:,*f:' hendrix
```

outputs all spines that contain the tandem interpretations `*F:` or `*f:` (i.e. in the keys of F major or F minor).

**SAMPLE OUTPUTS**

The following examples illustrate the various **extract** options. Consider the following input file, dubbed `input1`:

```
!! 'extract' example #1
**ABC    **xyz    **123    **ABC    **foo
A        x        1        a        bar
B        y        2        b        .
C        z        3        c        #
*_       *_       *_       *_       *_
```

Executing either of the following commands:

```
extract -f '1,3,$-1' input1 > output
```

or

```
extract -i '**ABC,**123' input1 > output
```

would produce the following result:



```
!! 'extract' example #1
**ABC    **123    **ABC    ..
A        1        a
B        2        b
C        3        c
*-       *-       *-
```

Consider next the following sample input — dubbed `input2`:

```
!! 'extract' example #2
**ABC    **xyz    **123    **ABC    **foo
A        x        1        a        %
*        *        *^       *        *
B        y        2a       2b       b        &
C        z        3a       3b       c        #
*        *x       *x       *        *        *
A        4a       xyz      4b       d        %
*-       *        *        *        *-       *
5a       xyz      5b       &
*        *x       *x       *
6a       6b       xyz      #
*v       *v       *        *
7        xyz      %
8        xyz      &
*-       *-       *-
```

Executing the command:

```
extract -p 3 input2 > output
```

would produce the following result:

```
!! 'extract' example #2
**123
1
*^
2a        2b
3a        3b
*x        *
4a        4b
*        *
5a        5b
*        *x
6a        6b
*v        *v
7
8
*-
```

Notice that this output no longer conforms to the Humdrum syntax. (Output lines 7 and 11

contain only a single exchange-path interpretation.)

For the 'example #2' input file, the field-structure is as follows:

```

1      1-1  # Line 1 must appear in the file.
4      1-5  # *      *      *^      *      *      ....
5      1-6  # Line after path indicator record
7      1-6  # *      *x      *x      *      ....
8      1-6  # Line after path indicator record
9      1-6  # *-      *      *      *      *      ....
10     1-4  # Line after path indicator record
11     1-4  # *      *x      *x      *      ....
12     1-4  # Line after path indicator record
13     1-4  # *v      *v      *      *      ....
14     1-3  # Line after path indicator record
16     1-3  # *-      *-      *-      ....

```

(The above file may be generated via the `fields -s` command.) On the basis of this information a user might create the following field-trace file, dubbed `trace`:

```

1      1
4      3
5      3,5
7      2,3
8      3,2
9      1,3
10     4
14     3
15     2
16     1

```

Executing the following command:

```
extract -t trace input2 > output
```

would produce the following result:

```

!! 'extract' example #2
**ABC
A
*^
2a          b
3a          c
*x          *x
4a          xyz
*_          *
&
*
#
*
%
xyz
*_

```

Notice that in this case, data tokens have been selected from a variety of input spines.

## DIAGNOSTICS

In *field-trace* mode, if the specified field-structure does not correspond to the actual input file, then an ERROR message is issued.

## PORTABILITY

DOS 2.0 and up, with the MKS Toolkit. OS/2 with the MKS Toolkit. UNIX systems supporting the *Korn* shell or *Bourne* shell command interpreters, and revised *awk* (1985).

## WARNINGS

Note that only the *field* mode (-f) guarantees that the output will conform to the Humdrum syntax.

In *interpretation* mode, tandem interpretations may be specified in the command invocation, but mutually exclusive tandem interpretations (such as changes of key) will not force the program to stop outputting a specified spine.

## SEE ALSO

**assemble** (4), **cut** (UNIX), **fields** (4), **yank** (4)