

Awk Command

awk is a text processor whose basic function is to search through an input file line by line, and execute actions if particular patterns are detected.

```
awk '/<regular expression>/' {<action on match>; <another action>;}' <file to parse>
```

- Lines within the body of an **awk** script consist of *rule-action* pairs.
- If the *rule* (`.`) is omitted, **awk** performs the *action* `{.}` for each line.
- If the *action* `{.}` is omitted, **awk** simply prints matching lines.

1.1 Field Indexing

Suppose the pattern specified in `<regular expression>` is matched. The action `print $0` prints the entire line for the identified record. The action `print $1` prints the first field of the identified record, and likewise for subsequent column indices.

1.2 Regular Expressions

Regular expressions are always bracketed by slashes `/<regular expression>/`.

1.3 Field Restricted Search

By default, **awk** search all fields for the target pattern. To search only the first field, use the syntax:

```
awk '($1 ~ /<regular expression>/)'
```

To search for records whose first field does **not** match the `<regular expression>`, use:

```
awk '($1 !~ /<regular expression>/)'
```

1.4 Internal Variables

awk recognizes the following internal variables:

- **FILENAME**: References the current input file.
- **NF**: Number of fields in the current record.
- **NR**: Overall number of the current record. Does not reset.
 - **FNR**: Record number within the *current file*. Resets when processing of a new file begins.

1.5 Beginning and Ending

If the line `BEGIN{<action 0>;}` precedes the body of an **awk** script, `<action 0>` is executed before any lines are read from the input. If the line `END{<action 1>;}` follows the body of an **awk** script, `<action 1>` is executed after all lines are read from the input.

Input

2.1 Records

The input file is read one *record* at a time. By default, one record corresponds to one line. Each record is *parsed* into *fields*. The default record separator is the newline character. To change the record separator to `/`, for instance, specify `RS="/"` in the `BEGIN` action.

2.2 Fields

By default, fields are separated by one or more spaces and/or tabs. Fields are referred to by the `$` sign, followed by the number of the field.