

Built-in variables and awk Script

Spoken Tutorial Project

<http://spoken-tutorial.org>

National Mission on Education through ICT

<http://sakshat.ac.in>

Antara Roy Choudhury

Praveen S

IIT Bombay

1 January 2018



Learning Objectives

We will learn about



Learning Objectives

We will learn about

- Built-in variables



Learning Objectives

We will learn about

- Built-in variables
- `awk` script



System Requirements



System Requirements

- **Ubuntu Linux 16.04**



System Requirements

- **Ubuntu Linux 16.04**
- **gedit text editor 3.20.1**



Code Files

- The files used in this tutorial are available in the [Code Files link](#) on this tutorial page



Code Files

- The files used in this tutorial are available in the Code Files link on this tutorial page
- Please download and use them.



Pre-requisites

- Please go through the earlier awk tutorials on <http://spoken-tutorial.org/>



Built-in Variables



Built-in Variables

- RS:



Built-in Variables

- RS:
 - Record separator in an input file



Built-in Variables

- RS:
 - Record separator in an input file
 - **Default value is newline**



Built-in Variables

- RS:
 - Record separator **in an input file**
 - **Default value is** newline
- FS:



Built-in Variables

- RS:
 - Record separator **in an input file**
 - **Default value is** newline
- FS:
 - Field separator **in an input file**



Built-in Variables

- RS:
 - Record separator **in an input file**
 - **Default value is** newline
- FS:
 - Field separator **in an input file**
 - **Default value is** whitespace



Built-in Variables(Cont.)

- ORS:



Built-in Variables(Cont.)

- ORS:
 - Output Record separator



Built-in Variables(Cont.)

- ORS:
 - Output Record separator
 - **Default value is newline**



Built-in Variables(Cont.)

- ORS:
 - Output Record separator
 - **Default value is** newline
- OFS:



Built-in Variables(Cont.)

- ORS:
 - Output Record separator
 - **Default value is** newline
- OFS:
 - Output Field separator



Built-in Variables(Cont.)

- ORS:
 - Output Record separator
 - **Default value is** newline
- OFS:
 - Output Field separator
 - **Default value is** whitespace



How to reset value of FS variable?

- **By default, any number of spaces and/or tabs**



How to reset value of FS variable?

- By default, any number of spaces and/or tabs
- Reset using command line parameter `-F`



How to reset value of FS variable?

- By default, any number of spaces and/or tabs
- Reset using command line parameter `-F`
- **Reset using FS variable in BEGIN section**



Example:

Suppose I want

- Colon as output field separator



Example:

Suppose I want

- Colon as output field separator
- Double newline as output record separator



Built-in variables(Cont.)



Built-in variables(Cont.)

- **NR: Number of Records processed by awk**



Built-in variables(Cont.)

- **NR: Number of Records processed by awk**
- **NF: Number of fields in the current record**



Example on system variable



Example on system variable

- **How to find incomplete lines in this file?**



Example on system variable

- **How to find incomplete lines in this file?**
- **Incomplete line means it has less than 6 fields**



Example

Suppose two files

- **demo1.txt**
- **demo2.txt**



Example

Suppose two files

- **demo1.txt**
- **demo2.txt**

Print first 3 lines from each of these two files



Example

Suppose two files

- demo1.txt
- demo2.txt

Print first 3 lines from each of these two files

— Use NR



Use FNR



Use FNR

- Current record number **in the current file**



Use FNR

- Current record number **in the current file**
- Incremented **each time a new record is read**



Use FNR

- Current record number **in the current file**
- Incremented **each time a new record is read**
- **It is reinitialized to zero each time a new input file is started**



NR vs FNR

- **NR: number of records since starting of the program's execution**



NR vs FNR

- NR: number of records since starting of the program's execution
- It does to reset to zero with a new file



Built-in Variables(Cont.)



Built-in Variables(Cont.)

- **FILENAME:** Name of the file being read



Built-in Variables(Cont.)

- **FILENAME:** Name of the file being read
- **ARGC:** Number of arguments provided at the command line



Built-in Variables(Cont.)

- **ARGV: Array that stores the command line arguments**



Built-in Variables(Cont.)

- **ARGV**: **Array** that stores the **command line arguments**
- **ENVIRON**: **Array** of the **shell environment variables** and **corresponding values**



Example

- Find students passed and having stipend > 8000



Example

- Find students passed and having stipend > 8000
- Comma as output field separator



Example

- Find students passed and having stipend > 8000
- Comma as output field separator
- Footer: print “The data is shown for file” FILENAME



awk script

- More and more complex tasks



awk script

- More and more complex tasks
- Difficult to write command every time on the terminal



awk script

- More and more complex tasks
- Difficult to write command every time on the terminal
- Write awk program in separate file



awk script

- More and more complex tasks
- Difficult to write command every time on the terminal
- Write awk program in separate file
- File should have **.awk** extension



awk script execution

- Specify this awk program filename with awk command



awk script execution

- Specify this awk program filename with awk command
- **Need -f option**



Summary

We learnt about

- Built-in variables
- awk script



Assignment 1

- **Write an awk script to print the last field of the 5th line in awkdemo.txt file**



Assignment 2

- Open the system file `/etc/passwd`



Assignment 2

- Open the system file `/etc/passwd`
- Identify the separators therein



Assignment 2

- Open the system file `/etc/passwd`
- Identify the separators therein
- From the 20th line onwards



Assignment 2

- Open the system file `/etc/passwd`
- Identify the separators therein
- From the 20th line onwards
- For lines containing more than 6 fields



Assignment 2

- Open the system file `/etc/passwd`
- Identify the separators therein
- From the 20th line onwards
- For lines containing more than 6 fields
- Print line number, entire line and count of fields



About the Spoken Tutorial Project

- Watch the video available at http://spoken-tutorial.org/What_is_a_Spoken_Tutorial
- It summarises the Spoken Tutorial project
- If you do not have good bandwidth, you can download and watch it



Spoken Tutorial Workshops

The Spoken Tutorial Project Team

- Conducts workshops using spoken tutorials
- Gives certificates to those who pass an online test
- For more details, please write to contact@spoken-tutorial.org



Forum for specific questions

- Do you have questions in THIS Spoken Tutorial?
- Please visit <http://forums.spoken-tutorial.org>
- Choose the minute and second where you have the question
- Explain your question briefly
- Someone from our team will answer them



Acknowledgements

- Spoken Tutorial Project is a part of the Talk to a Teacher project
- It is supported by the National Mission on Education through ICT, MHRD, Government of India
- More information on this Mission is available at:

<http://spoken-tutorial.org/NMEICT-Intro>



Thank You

Thanks for joining
<http://spoken-tutorial.org>

