Unix Tools

Stanley Yao
Computer Science Department
University of Arizona

awk – text processing

- Records (usually “line”): awk process the input text stream record by record.
- Fields (usually “word”): within a record, files are identified by $1, $2, … and $0 refers to the entire record
- Pattern-action statement
  - Pattern (actions)
  - Default pattern matches all records
  - Default action is copying the record onto the stdout
  - Special pattern BEGIN and END

awk – Regular Expression

- Any character
- Zero or more of the preceding character
- One or more of the preceding character
- Zero or one of the preceding character
- Or
- Grouping
- Character class (e.g. [a-z]). If the first character is ‘^’, it means not.
- Beginning of record
- End of record

awk - Variables

- No need to declare
- Special variables
  - NR: record number of the current record
  - NF: number of fields in the current record
  - FILENAME: name of the current input file
  - $1, $2, …: field 1, file2, …
  - $0: the current record

awk - Examples

- who | awk ‘/mike/ {printf("%s
", $1);}’
- who | awk ‘$1~/mike/ {printf("%s
", $1);}’
- who | awk ‘$1~/yao/’
- ps | awk ‘$/-/[0-9]/ {print $1}’
- who | awk ‘{for (i=1; i<=NF; i=i+1)
  printf("%s", $i); print("\n")}’
- who | awk ‘{printf("%d: %s
", NR, $0);}’

Further Readings

Acknowledgement