

UNIX 및 실습

11장 유닉스 유틸리티 이용하기

학습목표

- ▶ 유닉스 시스템이 제공하는 다양한 유틸리티의 사용 방법을 익힌다.
- ▶ 파일의 행 수, 단어 수, 문자 수를 찾는 방법을 익힌다.
- ▶ 파일을 정렬하고 내용의 중복을 제거하는 방법을 익힌다.
- ▶ 파일을 분할하거나 원하는 부분을 잘라내어 붙여서 새로운 파일을 생성하는 방법을 익힌다.

Section 01 파일 정보 수집 - wc

wc [옵션] 파일

- ▶ 파일의 라인수, 단어수, 바이트, 문자수 출력
- ▶ 옵션
 - ▶ -c : 바이트 수
 - ▶ -m : 문자수
 - ▶ -C : -m과 동일
 - ▶ -l : 라인 수
 - ▶ -w : 화이트 스페이스나 새로운 행으로 구분된 단어수

파일 정보 수집 - wc

▶ 사용법

```
telnet hanbitbook.co.kr
```

```
$ ls -l
-rw-r--r-- 1 user1 other 0 4월 15일 13:25 t1
-rw-r--r-- 1 user1 other 0 4월 15일 13:25 t2
-rw-r--r-- 1 user1 other 0 4월 15일 13:25 t3
-rw-r--r-- 1 user1 other 0 4월 10일 13:00 t4
$ ls -l > test_wc
$ wc test_wc
      6      47      320  test_wc
$
```

```
telnet hanbitbook.co.kr
```

```
$ wc -c test_wc
      320 test_wc
$ wc -m test_wc
      309 test_wc
$ wc -cl test_wc
      6      320 test_wc
$ wc -mw test_wc
      7      309 test_wc
$ who | wc -l
      16
```

옵션의 중복
사용

현재 시스템
접속자수

[실습하기] 파일 정보 수집 - wc

▶ 실습하기

- 1) `wc /etc/passwd`
- 2) `wc /etc/hosts`
- 3) `wc -l /etc/services`
- 4) `cat /etc/hosts | wc`
- 5) `ls -l /usr/bin | wc`
- 6) ???

- 1) /etc/passwd 파일의 정보를 확인
- 2) /etc/hosts 파일의 정보 확인
- 3) /etc/services 파일의 라인 수 확인
- 4) /etc/hosts 파일의 라인수를 확인하는 다른 방법
- 5) /usr/bin 디렉토리 파일의 라인 수가 의미하는 것은 무엇인가?
- 6) 현재 디렉토리에 파일이 몇개 있는지 알려면 어떻게 해야 하는가?

Section 02 파일 정렬 - sort

sort [옵션] 파일

- ▶ 아스키 코드값을 기준으로 파일의 텍스트 내용을 정렬하여 화면에 출력
 - ▶ 옵션
 - ▶ -b: 앞에 붙는 공백 무시
 - ▶ -c: 정렬이 되지 않은 상태로 출력
 - ▶ -d: 사전식 순서로 정렬. 숫자, 문자, 공백만 비교
 - ▶ -f: 대소문자 구분 안함
 - ▶ -m: 정렬된 파일을 통합
 - ▶ -n: 숫자를 산술 값으로 전환해 정렬
 - ▶ -r: 역순 정렬
 - ▶ -t 문자: 지정한 문자를 필드 구분자로 사용
 - ▶ +번호: 번호+1 필드를 기준으로 정렬
 - ▶ +pos1 -pos2: 정렬하고자 하는 필드의 열을 지정, pos1부터 pos2열까지 정렬

파일 정렬 - sort

▶ 사용 방법

▶ 기본 데이터 (s.dat)

```
$ more test_sort1
This is a test file.
1234

unix
Unix
_love
~love
!abc
 700
50
@_@
#include
love
^love

<love>
`love`
hi
&987
  abcd
{Love
$
```

▶ 기본 정렬

```
telnet hanbitbook.co.kr
$ sort test_sort1
700
abcd
!abc
#include
&987
1234
50
<love>
@_@
This is a test file.
Unix
^love
_love
`love`
hi
love
unix
{Love
~love
```

공백행들

첫 문자가 공백

숫자보다 먼저 나오는 특수문자들

숫자들

숫자와 대문자 사이에 나오는 특수문자들

대문자들

대문자와 소문자 사이에 나오는 특수문자들

소문자들

소문자 다음에 나오는 특수문자들

파일 정렬 - sort

▶ 사전식으로 정렬 : -d

```
telnet hanbitbook.co.kr
$ sort -d test_sort1
@_@
 700
abcd
1234
50
&987
{Love
This is a test file.
Unix
!abc
hi
#include
<love>
^love
_love
`love`
love
~love
unix
$
```

공백행들

첫문자가 공백문자

&는 무시하고 987로 간주

{ 는 무시하고 Love로 간주

!는 무시

앞의 특수문자들 무시

▶ 역순으로 정렬 : -r

```
telnet hanbitbook.co.kr
$ sort -r test_sort1
~love
{Love
unix
love
hi
`love`
_love
^love
Unix
This is a test file.
@_@
<love>
50
1234
&987
#include
!abc
abcd
700
$
```


파일 정렬 - sort

▶ 필드 정렬하기 : +n

```
telnet hanbitbook.co.kr
```

```
$ cat test_sort2  
001 Hong Gil-Dong 80 M  
002 Park Ji-Soo 100 M  
003 Lee Na-Young 54 F  
004 Kim Chan-Sook 60 F  
005 Han Ju-Hyun 75 M  
006 Jyun Doo-Ri 49 F  
007 Lee Mi-Ra 59 F  
$
```

```
telnet hanbitbook.co.kr
```

```
$ sort +1 test_sort2  
005 Han Ju-Hyun 75 M  
001 Hong Gil-Dong 80 M  
006 Jyun Doo-Ri 49 F  
004 Kim Chan-Sook 60 F  
007 Lee Mi-Ra 59 F  
003 Lee Na-Young 54 F  
002 Park Ji-Soo 100 M  
$
```

-k: 필드번호를
그대로 사용

```
telnet hanbitbook.co.kr
```

```
$ sort -k 2 test_sort2  
005 Han Ju-Hyun 75 M  
001 Hong Gil-Dong 80 M  
006 Jyun Doo-Ri 49 F  
004 Kim Chan-Sook 60 F  
007 Lee Mi-Ra 59 F  
003 Lee Na-Young 54 F  
002 Park Ji-Soo 100 M  
$
```

파일 정렬 - sort

▶ 특정 필드를 기준으로 정렬

```
telnet hanbitbook.co.kr
$ sort +1 -2 test_sort2
005 Han Ju-Hyun 75 M
001 Hong Gil-Dong 80 M
006 Jyun Doo-Ri 49 F
004 Kim Chan-Sook 60 F
003 Lee Na-Young 54 F
007 Lee Mi-Ra 59 F
002 Park Ji-Soo 100 M
$
```

2번 필드만을 기준으로 정렬
2열이 같으면 원래 파일 순서대로 정렬

▶ 숫자를 기준으로 정렬

```
telnet hanbitbook.co.kr
$ sort -k 4 test_sort2
002 Park Ji-Soo 100 M
006 Jyun Doo-Ri 49 F
003 Lee Na-Young 54 F
007 Lee Mi-Ra 59 F
004 Kim Chan-Sook 60 F
005 Han Ju-Hyun 75 M
001 Hong Gil-Dong 80 M
$
```



-n : 숫자처리

```
telnet hanbitbook.co.kr
$ sort -n -k 4 test_sort2
006 Jyun Doo-Ri 49 F
003 Lee Na-Young 54 F
007 Lee Mi-Ra 59 F
004 Kim Chan-Sook 60 F
005 Han Ju-Hyun 75 M
001 Hong Gil-Dong 80 M
002 Park Ji-Soo 100 M
$
```

파일 정렬 - sort

▶ 2차 정렬하기

```
telnet hanbitbook.co.kr
$ sort +2 -3 +1 -2 test_sort2
004 Kim Chan-Sook 60 F
006 Jyun Doo-Ri 49 F
001 Hong Gil-Dong 80 M
002 Park Ji-Soo 100 M
005 Han Ju-Hyun 75 M
007 Lee Mi-Ra 59 F
003 Lee Na-Young 54 F
$
```

▶ 정렬결과 저장하기

```
telnet hanbitbook.co.kr
$ sort -n +3 -4 -o sort.out test_sort2
$ cat sort.out
006 Jyun Doo-Ri 49 F
003 Lee Na-Young 54 F
007 Lee Mi-Ra 59 F
004 Kim Chan-Sook 60 F
005 Han Ju-Hyun 75 M
001 Hong Gil-Dong 80 M
002 Park Ji-Soo 100 M
$
```

3번 필드만을 기준으로 정렬
3열이 같으면 2열을 기준으로 정렬

파일 정렬 - sort

▶ 필드 구분자 지정하기 : -t

```
telnet hanbitbook.co.kr
```

```
$ sort -t: /etc/passwd
```

```
(중략)
```

```
jongwon:x:1089:1000:이종원:/expoxrt/home/jongwon:/bin/ksh
```

```
root:x:0:1:Super-User:/:/bin/ksh
```

```
sjyoun:x:1090:1000:윤소정:/export/home/sjyoun:/bin/ksh
```

```
user1:x:10001:2006:김삼순:/export/home/user1:/bin/ksh
```

```
user2:x:10002:2006:박나라:/export/home/user2:/bin/ksh
```

```
user3:x:10003:2006:이나영:/export/home/user3:/bin/ksh
```

```
$
```

[실습하기] 파일 정렬 - sort

▶ 실습하기

```
1) mkdir unix/ch11
2) cd unix/ch11
3) cp /tmp/s.dat .
4) cat s.dat
5) sort s.dat
6) sort -f s.dat
7) sort -r s.dat
8) sort +2 s.dat
9) sort -o sort.out s.dat
10) cat sort.out
11) sort +2 +3 s.dat
12) sort -m s.dat sort.out
```

- 1)
- 2)
- 3) 기본 데이터 파일을 현재 디렉토리로 복사
- 4) 기본 파일 내용을 확인
- 5) 기본 정렬
- 6) 대소문자 구분없이 정렬
- 7) 역순으로 정렬
- 8) 3번째 열을 기준으로 정렬
- 9) 정렬결과를 파일로 저장
- 10) 결과파일을 확인
- 11) 3, 4번째 컬럼을 기준으로 정렬
- 12) 기본 파일과 정렬파일을 합병

Section 03 파일 분할 - split

split [옵션] [파일]

- ▶ 큰 파일을 일정한 크기의 여러 개 작은 파일로 분할
- ▶ 분할된 파일 이름은 xaa, xab, ... 순으로 생성
 - ▶ 옵션
 - ▶ -b n : 크기가 n 바이트인 파일로 분할
 - ▶ -n : n줄씩 분할
- ▶ 옵션을 지정하지 않으면 1000줄씩 분할
- ▶ 파일을 지정하지 않으면 표준입력내용을 분할 저장

파일 분할 - split

▶ 사용 방법

▶ 기본데이터(test_split)

```
telnet hanbitbook.co.kr
$ cp /etc/services test_split
$ wc -l test_split
   98 test_split
$
```

▶ 행을 기준으로 파일 분할

```
telnet hanbitbook.co.kr
$ split -30 test_split1
$ ls
test_split1 xaa xab xac xad
$ wc -l x*
 30 xaa
 30 xab
 30 xac
  8 xad
 98 total
$
```

▶ 바이트를 기준으로 파일 분할

```
telnet hanbitbook.co.kr
$ wc -c test_split1
3035 test_split1
$ split -b 512 test_split1
$ ls
test_split1 xaa xab xac xad xae xaf
$ wc -c x*
512 xaa
512 xab
512 xac
512 xad
512 xae
475 xaf
3035 total
$
```

Section 04 중복 삭제 - uniq

uniq [옵션] [입력파일 [출력파일]]

- ▶ 파일 또는 표준입력으로 입력된 내용 중 중복된 내용의 줄이 연속으로 있으면 하나만 남기고 삭제
- ▶ 파일을 지정하지 않으면 표준입력내용을 처리
- ▶ 입출력 파일 이름은 달라야 함
- ▶ 옵션
 - ▶ -u : 중복되지 않는 줄만 출력
 - ▶ -d : 중복된 줄 중 1줄만 출력
 - ▶ -n : 원하는 필드 앞의 공백과 문자열 무시. n은 필드 번호

중복 삭제 - uniq

▶ 사용 방법

▶ 기본데이터(test_uniq1)

```
$ cat test_uniq1
aaaaa
abcde
abcde
bbbbb
abcde
bbbbb
bbbbb
abcde
ccc
$
```

▶ 중복제거

```
telnet hanbitbook.co.kr
$ uniq test_uniq1
aaaaa
abcde
bbbbb
abcde
bbbbb
abcde
ccc
$
```

중복 삭제 - uniq

- ▶ 정렬과 중복제거를 동시에

```
telnet hanbitbook.co.kr
$ sort test_uniq1
aaaaa
abcde
abcde
abcde
abcde
bbbbbb
bbbbbb
bbbbbb
ccc
$ sort test_uniq1 | uniq
aaaaa
abcde
bbbbbb
ccc
$
```

- ▶ 중복없는 행 보기

```
telnet hanbitbook.co.kr
$ uniq -u test_uniq1
aaaaa
ccc
$
```

- ▶ 중복행과 중복 횟수 보기

```
telnet hanbitbook.co.kr
$ uniq -d test_uniq1
abcde
bbbbbb
$ uniq -c test_uniq1
  1 aaaaa
  4 abcde
  3 bbbbb
  1 ccc
$
```

Section 05 필드 잘라내기 - cut

cut [옵션] [파일]

- ▶ 파일의 각 행에서 선택된 필드를 잘라냄
- ▶ 옵션
 - ▶ -c 리스트 : 각 줄에서 잘라낼 문자 위치 지정
 - ▶ -f 필드수 : 지정한 필드 잘라냄
 - ▶ -d 문자 : 필드 구분자

필드 잘라내기 - cut

▶ 사용 방법

▶ 기본 데이터 (test_cut)

```
$ cat test_cut
001 Hong Gil-Dong 80 M
002 Park Ji-Soo 100 M
003 Lee Na-Young 54 F
004 Kim Chan-Sook 60 F
005 Han Ju-Hyun 75 M
006 Jyun Doo-Ri 49 F
007 Lee Mi-Ra 59 F
$
```

▶ 필드추출 : -f 옵션

▶ 문자추출 : -c 옵션

```
telnet hanbitbook.co.kr
$ cut -c5-8 test_cut
Hong
Park
Lee
Kim
Han
Jyun
Lee
$
```

```
telnet hanbitbook.co.kr
$ cut -f 2,3 test_cut
Hong Gil-Dong
Park Ji-Soo
Lee Na-Young
Kim Chan-Sook
Han Ju-Hyun
Jyun Doo-Ri
Lee Mi-Ra
$
```

필드 잘라내기 - cut

▶ 필드 구분자 지정하기 : -d

```
telnet hanbitbook.co.kr
$ cut -d: -f 1 /etc/passwd | more
root
daemon
bin
sys
adm
lp
smtp
uucp
listen
nobody
noaccess
nobody4
mysql
(중략)
--계속--
```

```
telnet hanbitbook.co.kr
$ cut -d' ' -f 2 test_sort2
Hong
Park
Lee
Kim
Han
Jyun
Lee
$
```

Section 06 두 파일 연결하기 - paste

```
paste [옵션] [파일1 파일2 ...]
```

- ▶ 사용자가 지정한 두 개 이상의 파일 내용 중 같은 줄을 붙이거나 한 파일의 끝에 다른 파일의 내용을 추가
- ▶ 옵션
 - ▶ -s : 파일의 끝에 추가 (split로 나눈 파일을 원래대로 붙일 때)
 - ▶ -d문자 : 필드 구분자
 - ▶ - : 파일 대신 표준 입력 사용

두 파일 연결하기 - paste

▶ 사용 방법

▶ test_paste1

```
001  
002  
003  
004  
005  
006  
007
```

▶ test_paste2

```
Hong Gil-Dong  
Park Ji-Soo  
Lee Na-Young  
Kim Chan-Sook  
Han Ju-Hyun  
Jyun Doo-Ri  
Lee Mi-Ra
```

▶ 파일 붙이기

```
telnet hanbitbook.co.kr  
$ paste test_paste2 test_paste1  
Hong Gil-Dong 001  
Park Ji-Soo 002  
Lee Na-Young 003  
Kim Chan-Sook 004  
Han Ju-Hyun 005  
Jyun Doo-Ri 006  
Lee Mi-Ra 007  
$
```

두 파일 연결하기 - paste

- ▶ 두 파일의 행이 같지 않을 경우

- ▶ test_paste3

```
001 002 003 004
005 006 007
```

```
telnet hanbitbook.co.kr
```

```
$ paste test_paste2 test_paste3
Hong Gil-Dong 001 002 003 004
Park Ji-Soo 005 006 007
Lee Na-Young
Kim Chan-Sook
Han Ju-Hyun
Jyun Doo-Ri
Lee Mi-Ra
$
```

```
telnet hanbitbook.co.kr
```

```
$ paste -s test_paste1 test_paste2
001 002 003 004 005 006 007
Hong Gil-Dong Park Ji-Soo Lee Na-Young Kim Chan-Sook Han Ju-Hyun
Jyun Doo-Ri Lee Mi-Ra
$
```

- ▶ 필드구분자 지정 : -d

```
telnet hanbitbook.co.kr
```

```
$ paste -d: test_paste1 test_
paste2
001:Hong Gil-Dong
002:Park Ji-Soo
003:Lee Na-Young
004:Kim Chan-Sook
005:Han Ju-Hyun
006:Jyun Doo-Ri
007:Lee Mi-Ra
$
```

- ▶ 파일수평 붙이기 : -s

두 파일 연결하기 - paste

▶ cut과 paste의 복합 사용

▶ s.dat

```
Kim Ji-Soo    10  F
Lee Gil-Dong  20  M
Lee Gil-San   15  M
Park Ji-Soo   21  F
Choi Na-Na    20  F
```

▶ u.dat

```
aaaaa
abcde
abcde
bbbbb
```

```
telnet hanbitbook.co.kr
```

```
$ cut -d' ' -f1 s.dat | paste - u.dat
Kim aaaaa
Lee abcde
Lee abcde
Park bbbbb
Choi
$
```

Section 07 파일 덤프 - dd

```
dd [옵션] [if=입력파일] [of=출력파일]
```

- ▶ 지정된 입력 파일을 지정된 옵션에 따라 변환하여 출력파일로 저장하는 유틸리티
- ▶ 옵션
 - ▶ `bs=n` : 입출력 블록의 크기를 n 바이트로 지정 (기본 1블록=512바이트)
 - ▶ `conv=lcase` : 알파벳을 소문자로 변환
 - ▶ `conv=ucase` : 알파벳을 대문자로 변환

파일 덤프 - dd

▶ 사용 방법 : 대소문자 전환하기

```
telnet hanbitbook.co.kr
```

```
$ dd conv=lc case if=test_cut of=test_dd1
```

```
0+1 레코드 인:
```

```
0+1 레코드 아웃
```

```
$ more test_dd1
```

```
001 hong gil-dong 80 m
002 park ji-soo 100 m
003 lee na-young 54 f
004 kim chan-sook 60 f
005 han ju-hyun 75 m
006 jyun doo-ri 49 f
007 lee mi-ra 59 f
```

```
$
```

파일 덤프 - dd

▶ 파일 지우기

```
telnet hanbitbook.co.kr
$ dd if=/dev/null of=test_dd2
0+0 레코드 인:
0+0 레코드 아웃
$ more test_dd2
$ ls test_dd2
test_dd2
$
```

[실습하기]

```
1) mkdir Practice
2) cd Practice
3) pwd
4) more /etc/passwd
5) cut -f 1 -d: /etc/passwd > login_list
6) cut -f 5 -d: /etc/passwd > name_list
7) paste -d: name_list login_list > user_list
8) sort -o user_list user_list
```

- 1) 실습디렉토리 생성
- 2) 실습디렉토리로 이동
- 3) 현재 경로 파악
- 4) /etc/passwd파일 구조 파악
- 5) 첫번째 필드인 로그인ID추출
- 6) 사용자 이름 추출(5번째 필드)
- 7) 파일 합치기
- 8) 정렬

[실습과제]

- ▶ 실습 각 단계 화면 캡처하여 pdf 파일로 정리하여 과제 제출 (cms.mmu.ac.kr/bear)
- ▶ 제출기한 : 5월 14일 자정