

UNIX 및 실습

7장. 파일과 디렉토리 검색하기

학습목표

- ▶ 파일의 내용을 검색하는 방법을 익힌다.
- ▶ 조건에 맞는 파일과 디렉토리를 찾는 방법을 익힌다.
- ▶ 명령이 있는 위치를 찾는 방법을 익힌다.

01. 파일 내용 검색 - grep

grep [옵션] 패턴 파일명들

- ▶ global regular expression print
- ▶ 지정한 파일에 특정 문자열(패턴)이 들어있는지 검색
- ▶ 패턴
 - ▶ 문자, 문자열, 문장, 정규표현식(regular expression)
- ▶ 옵션

옵션	기능
-i	대소문자를 무시하고 검색
-l	해당 패턴이 들어있는 파일 이름을 출력
-n	각 라인의 번호도 함께 출력
-v	명시된 패턴과 일치하지 않는 줄을 출력
-c	패턴과 일치하는 라인수 출력
-w	패턴이 하나의 단어로 된 것만 검색

정규표현식(Regular Expression) (1)

- ▶ 검색에서 동일한 문자와 일치시키는 데 사용되는 문자 패턴
- ▶ 정규표현식에 사용되는 메타(meta) 문자
 - ▶ ^ 줄의 시작 지시자
 - ▶ \$ 줄의 끝 지시자
 - ▶ . 문자 한 개와 일치
 - ▶ * 0개 이상의 문자와 일치
 - ▶ [] 설정 중 한 문자와 일치
 - ▶ [a-z] 설정한 범위 안의 한 문자와 일치
 - ▶ [^] 설정에 속하지 않은 문자와 일치
 - ▶ \w 메타문자 기능 제거

정규표현식(Regular Expression) (2)

- ▶ 여러 유닉스/리눅스 프로그램에서 사용되는 추가 메타문자
 - ▶ $\backslash\wedge<$ 단어의 시작 지시자
 - ▶ $\backslash\wedge>$ 단어의 끝 지시자
 - ▶ $\backslash\wedge(..\wedge)$ 나중에 사용하기 위한 태그 일치 문자
 - ▶ $x\wedge\{m\wedge\}$ x문자를 m번 반복
 - ▶ $x\wedge\{m, \wedge\}$ x문자를 적어도 m번 이상 반복
 - ▶ $x\wedge\{m, n\wedge\}$ x문자를 m번 이상 n번 이하 반복
 - ▶ 예
 - ▶ $/\wedge<love/$: love로 시작하는 단어가 포함된 줄
 - ▶ $/love\wedge>/$: love로 끝나는 단어가 포함된 줄
 - ▶ $o\wedge\{5, 10\wedge\}$: o가 5번 이상 10번 이하 반복되는 문자열

grep 명령

▶ grep 의미

- ▶ ex 편집기 프로그램에서 지원
- ▶ 문자 검색을 위해
 - ▶ `:/pattern/p` (pattern이 있는 줄 출력)
- ▶ 보통 범위(전체 파일)를 지정하기 위해
 - ▶ `:g/pattern/p`
- ▶ 이 패턴이 정규표현식
 - ▶ `:g/RE/p`

파일 내용 검색 - grep

- ▶ grep 명령 사용 예제
 - ▶ 기본 데이터 (grep.dat)

```
UNIX 12345
unix+ 123
system admin
Network 5
root other sh
sjyoun prof ksh
jongwon prof KSH
ROOT other csh
ck07555 student ksh
CK08777 student bash
```

- ▶ 1) 기본 검색

```
ssh lily.mmu.ac.kr
$ grep unix grep.dat
unix+ 123
$
```

- ▶ 2) 대소문자 무시 -i

```
ssh lily.mmu.ac.kr
$ grep -i unix grep.dat
UNIX 12345
unix+ 123
$
```

파일 내용 검색 - grep

- ▶ grep 명령 사용 예제
 - ▶ 기본 데이터 (grep.dat)

```
UNIX 12345
unix+ 123
system admin
Network 5
root other sh
sjyoun prof ksh
jongwon prof KSH
ROOT other csh
ck07555 student ksh
CK08777 student bash
```

- ▶ 3) 파일 이름 출력 : -l

```
ssh lily.mmu.ac.kr
$ grep -l unix *.dat
grep.dat
$
```

- ▶ 4) 줄번호 출력 : -n

```
ssh lily.mmu.ac.kr
$ grep -n unix grep.dat
2:unix+ 123
$
```

파일 내용 검색 - grep

- ▶ grep 명령 사용 예제
 - ▶ 기본 데이터 (grep.dat)

```
UNIX 12345
unix+ 123
system admin
Network 5
root other sh
sjyoun prof ksh
jongwon prof KSH
ROOT other csh
ck07555 student ksh
CK08777 student bash
```

- ▶ 5) 불일치: -v

```
ssh lily.mmu.ac.kr
$ grep -v unix grep.dat
UNIX 12345
system admin
Network 5
...
$
```

- ▶ 6) 일치하는 줄 수 : -c

```
ssh lily.mmu.ac.kr
$ grep -c 123 grep.dat
2
$
```

파일 내용 검색 - grep

- ▶ grep 명령 사용 예제
 - ▶ 기본 데이터 (grep.dat)

```
UNIX 12345
unix+ 123
system admin
Network 5
root other sh
sjyoun prof ksh
jongwon prof KSH
ROOT other csh
ck07555 student ksh
CK08777 student bash
```

- ▶ 7) 단어 검색 : -w

```
ssh lily.mmu.ac.kr
$ grep -w 123 grep.dat
unix+ 123
$
```

- ▶ 8) 다중 옵션 사용

```
ssh lily.mmu.ac.kr
$ grep -ni unix grep.dat
1:UNIX 12345
2:unix+ 123
$
```

[실습하기] 파일 내용 검색 - grep

▶ 실습하기

▶ 데이터 파일 : h.txt

```
root  other  sh
ROOT  csh    user1
USER2 12root  ksh
user1  KSH    csh
```

결과를 예측해 봅시다

- 1) `grep sh h.txt`
- 2) `grep -w sh h.txt`
- 3) `grep -v root h.txt`
- 4) `grep -I root h.txt`
- 5) `grep -ci sh h.txt`

- 6) `grep user1 /etc/passwd`
- 7) `grep root /etc/passwd`
- 8) `grep root /etc/group`
- 9) `grep other /etc/group`

[실습하기] 파일 내용 검색 - grep

▶ 실습하기

▶ 데이터 파일 : h.txt

```
root  other  sh
ROOT  csh  user1
USER2 12root ksh
user  KSH  csh
```

- 1) 1만 검색하기 위한 명령은?
- 2) 1이 들어있지 않은 라인을 검색하기 위한 명령은?
- 3) User를 대소문자 구분 없이 검색하기 위한 명령은?
- 4) sh가 들어간 라인의 수를 구하는 명령은?
- 5) 정확히 sh만 들어간 라인의 수를 구하는 명령은?
- 6) root가 들어간 라인을 찾아 파일 h.out에 저장하는 명령은?

정규표현식

- ▶ 특정한 규칙을 가진 문자열의 집합을 표현하는데 사용하는 형식 언어
- ▶ 구성 요소
 - ▶ 앵커(Anchor) : 검색시 한 줄에서 패턴의 위치를 표현(^, \$)
 - ▶ 문자 집합(Character Set) : 하나 이상의 문자들을 표현(알파벳, 숫자,..[]..)
 - ▶ 변환자(Modifier) : 이전 문자 집합의 반복횟수 지정(*)
- ▶ 특수문자

문자	의미	예	결과
^	라인의 시작	'^문자열'	문자열로 시작하는 모든 행
\$	라인의 끝	'문자열\$'	문자열로 끝나는 모든 행
.	한 글자	'a...b'	한글자 대응, a로 시작해서 b로 끝나는 5글자 검색
?	없거나 한글자	'patter?'	patter 또는 patter과 한 문자 더 있는 문자열 검색 (pattern, pattera 등)
[]	괄호 안의 글자 중 하나	'[Pp]attern'	Pattern 또는 pattern이 나타나는 라인 검색
[^]	괄호 안에 있는 글자가 아닌 글자	'[^a-m]att'	att앞에 a부터 m까지 나오지 않는 라인 검색
*	앞의 항목이 없거나 여러 번 반복	'ab*'	a다음에 b가 없거나 반복적으로 나타나는 라인 검색

[실습하기] 파일 내용 검색 - grep

▶ 실습하기

▶ 데이터 파일 : g.txt

```
UNIX 12345
unix+ 123
system admin
Network 5
root other sh
sjyoun prof ksh
jongwon prof KSH
ROOT other csh
ck07555 student ksh
CK08777 student bash
```

주의!

패턴을 작은 따옴표 ' ' 로
꼭 감싸야 함!

- 1) grep `^root` g.txt
- 2) grep `sh\$` g.txt
- 3) grep `r..t` g.txt
- 4) grep `oo*` g.txt
- 5) grep `[0-9].*` g.txt
- 6) grep `[^c]sh` g.txt

결과를 예측해 봅시다

파일 내용 검색 - egrep

egrep [옵션] 패턴 파일명들

- ▶ extended regular expression을 사용하는 grep 명령
- ▶ 확장된 패턴 표현식

메타 문자	의미	예	결과
+	앞의 글자가 하나 이상 나온다.	'[a-z]+ark'	airport, dark, bark, shark
x y	x나 y중 하나가 나온다	'apple orange'	apple 또는 orange
()	문자열 그룹	'(1 2)+'	1또는 2가 하나이상
		'search(es ing)+'	searches 또는 searching

파일 내용 검색 - egrep

- ▶ grep 명령 사용 예제
 - ▶ 기본 데이터 (g.dat)

```
UNIX 12345
unix+ 123
system admin
Network 5
root other sh
sjyoun prof ksh
jongwon prof KSH
ROOT other csh
ck07555 student ksh
CK08777 student bash
```

- ▶ 1) +

```
ssh lily.mmu.ac.kr
$ egrep '[78]+' g.dat
ck07555 student ksh
CK08777 student bash
$
```

- ▶ 2) x|y

```
ssh lily.mmu.ac.kr
$ egrep 'csh|bash' g.dat
ROOT other csh
CK08777 student bash
$
```

[실습하기] 파일 내용 검색 - grep

▶ 실습하기

▶ 데이터 파일 : eh.txt

```
root sh user05567
ROOT csh user05777
root ksh user05666
ROOT ksh user05888
```

- 1) `egrep 'root|ROOT' eh.txt`
- 2) `egrep '(root|ROOT) ksh' eh.txt`
- 3) `egrep '[56]+' eh.txt`
- 4) `egrep 'csh|bash' eh.txt`
- 5) `egrep '[a-z]+body' /etc/passwd`

결과를 예측해 봅시다

파일 내용 검색 - fgrep

fgrep [옵션] 문자열 파일명들

- ▶ 문자열 검색
- ▶ 문자열 내의 모든 문자를 일반 문자로 해석
 - ▶ 정규 표현식 사용 안함
 - ▶ 모든 정규표현식 특수 문자를 문자 그대로 검색
 - ▶ * : 문자 * 로 인식하여 검색

파일 내용 검색 - fgrep

▶ fgrep 명령 사용 예제

▶ 기본 데이터 (fg.dat)

```
# fgrep data file
#
JAVA 2244
aix admin
* Unix admin
* Network 25
```

▶ 1) * 검색

```
ssh lily.mmu.ac.kr
$ fgrep '*' fg.dat
* Unix admin
* Network 25
$
```

▶ 2) # 검색

```
ssh lily.mmu.ac.kr
$ fgrep '#' fg.dat
# fgrep data file
#
$
```

파일 내용 검색 - grep과 파이프

- ▶ grep 명령은 파이프와 함께 자주 사용됨

```
ssh lily.mmu.ac.kr
$ ls -l | grep rw-
-rw-r--r-- 1 user1 other 50 4월30일 12:00 g.dat
-rw-r--r-- 1 user1 other 50 4월30일 12:00 g.dat
$
```

```
ssh lily.mmu.ac.kr
$ ps -ef | grep user1
user1 6683 6680 0 17:28:15 pts/1 0:00 -ksh
user1 6720 6683 0 17:50:53 pts/1 0:00 -ksh
$
```

grep 응용 과제 (1)

▶ 다음 datafile을 이용하여

northwest	NW	Charles Main	3.0	.98	3	34
western	WE	Sharon Gray	5.3	.97	5	23
southwest	SW	Lewis Dalsase	2.7	.8	2	18
southern	SO	Suan Chin	5.1	.95	4	15
southeast	SE	Patrick Hemenway	4.0	.7	4	17
eastern	EA	TB Savage	4.4	.84	5	20
northeast	NE	AM Main Jt.	5.1	.94	3	13
north	NO	Margot Weber	4.5	.89	5	9
central	CT	Ann Stephens	5.7	.94	5	13

1. NW가 포함된 모든 줄 찾기
2. n으로 시작하는 모든 줄 찾기
3. 4로 끝나는 모든 줄 찾기
4. TB Savage가 포함된 모든 줄 찾기

grep 응용 과제 (2)

5. 숫자 5와 소수점이 포함된 모든 줄 찾기
6. .5가 포함된 모든 줄 찾기
7. w나 e로 시작하는 모든 줄 찾기
8. north로 시작하는 단어가 포함된 모든 줄 찾기
9. south로 시작하는 단어가 포함되지 않은 모든 줄 찾기
10. 대소문자 구별 없이 so로 시작하는 모든 줄의 줄 번호와 찾은 패턴 보이기
11. east로 시작하는 단어가 포함된 줄 수 출력하기
12. 단어 자체가 north인 패턴이 포함된 모든 줄 찾기 (northwest는 불포함)

GNU grep

- ▶ Gnu grep은 -G, -E, -F, -P 사용 가능
 - ▶ grep -G : 기본 grep
 - ▶ grep -E 또는 egrep
 - ▶ grep -F 또는 fgrep
 - ▶ 재귀 grep (rgrep 또는 grep -R)
 - ▶ grep -P (Perl 정규표현식 지원)