

자료구조 과목
3차 과제 설명서
(단어 정렬 프로그램 구현)

2020. 6. 1

향후 강의 및 과제 추진 일정

주차	기간	강의 내용	과제 내용(일정)
9	5/11 ~ 5/15	9장 우선순위 큐(완료)	
10	5/18 ~ 5/22	10장 그래프 1	2차 과제 공지(5/18)
11	5/25 ~ 5/29	12장 정렬 1	2차 과제 제출(5/29)
12	6/1 ~ 6/5	2차 과제 발표 12장 정렬 2	2차 과제 발표(6/1) 3차 과제 공지(6/3)
13	6/8 ~ 6/12	12장 정렬 2 13장 탐색 1	
14	6/15 ~ 6/19	13장 탐색 2 3차 과제 발표 1	3차 과제 제출(6/15) 3차 과제 발표 및 시연 (6/17 ~ 6/22)
15	6/22 ~ 6/26	3차 과제 발표 2 학기말 고사(6/24)	

3차 과제 소개

■ 학생 단위로 진행

■ 3차 과제

- 파일에 들어있는 영어 단어 빈도수 기준으로 3가지 정렬 알고리즘을 사용하는 프로그램 1종 설계 및 구현하고,
- 프로그램 개발결과보고서 작성 후 제출하고,
- 프로그램 개발결과보고서 발표하면서 시연
 - 발표 희망자 및 발표 지명자에 한하여 강의 시간 중 발표 진행
 - 발표 참여 시 가점있음
 - 강의 참석이 어려울 경우 관련 내용 동영상 제출도 가능

■ 과제 내용 : 3가지 정렬 알고리즘 포함된 영어 단어 정렬 프로그램 구현

- 영어 단어가 들어있는 텍스트 파일에 포함된 영어 단어를 추출하고,
- 학습한 정렬 알고리즘을 이용하여 단어와 빈도수를 정렬하고 출력하는 영어 단어 정렬 프로그램 구현

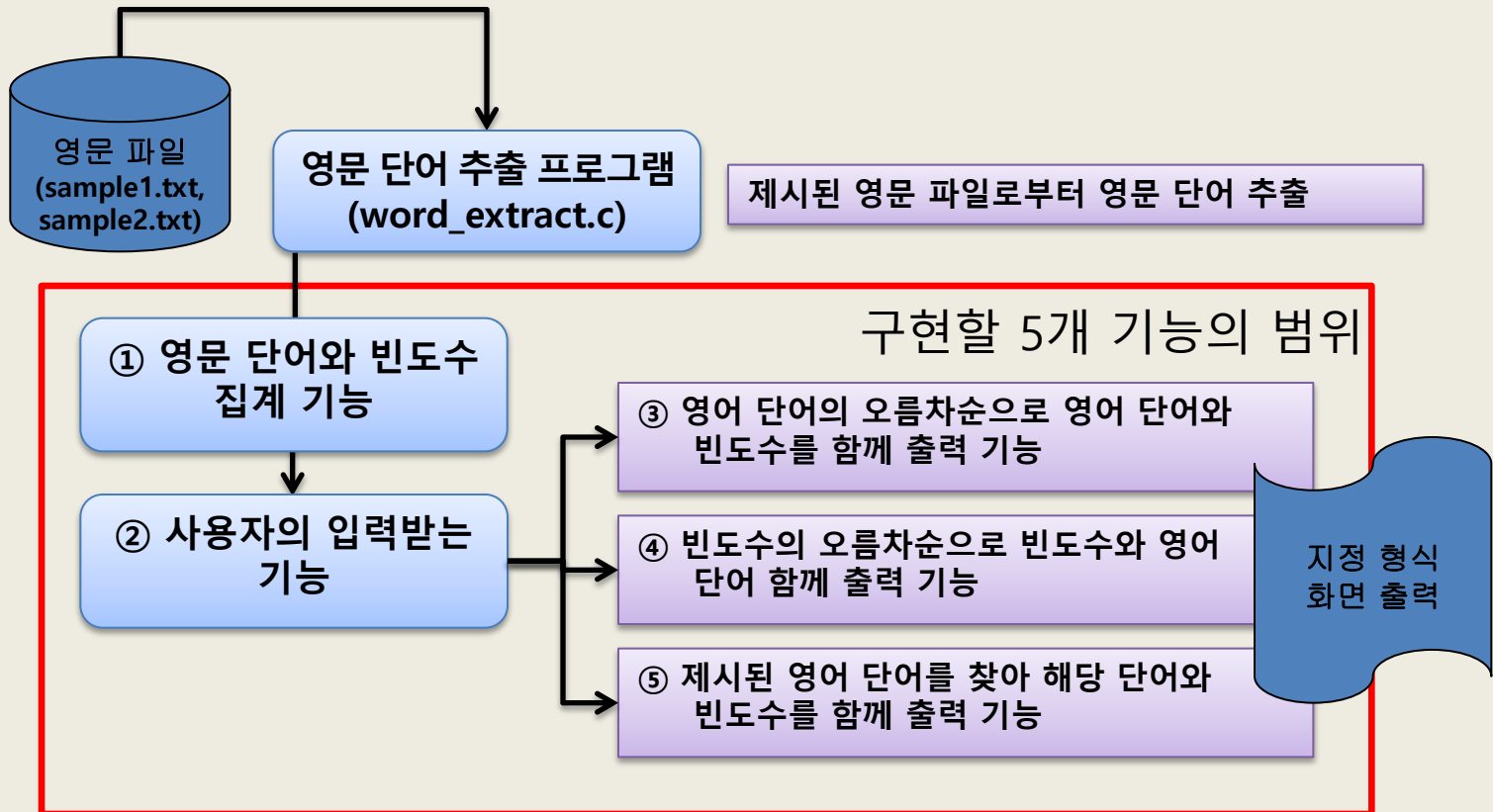
3차 과제 소개

■ 영어 단어 빈도수 기준 정렬 프로그램이란?

- 텍스트 파일에 들어 있는 영어 단어에 대하여 단어와 빈도수를 함께 표기하는 프로그램
- 프로그램의 목적은 데이터 파일에 포함되어 있는 빈도수가 높은 단어 추출하여 파일의 내용이 속하는 분야를 추정하는 목적으로 주로 사용됨

3차 과제 소개

■ 정렬 알고리즘 3종을 선택적으로 사용하는 프로그램의 동작 절차



3차 과제 소개

■ 정렬 알고리즘 3종을 선택적으로 사용하는 프로그램 구현

■ 동작 절차 및 기능

- ① 영문 단어와 빈도수 집계 기능
- ② 사용자의 입력받는 기능
- ③ 영어 단어의 오름차순으로 영어 단어와 빈도수를 함께 출력 기능 → 출력 1.
- ④ 빈도수의 오름차순으로 빈도수와 영어 단어 함께 출력 기능 → 출력 2.
- ⑤ 제시된 영어 단어를 찾아 해당 단어와 빈도수를 함께 출력 기능 → 출력 3.

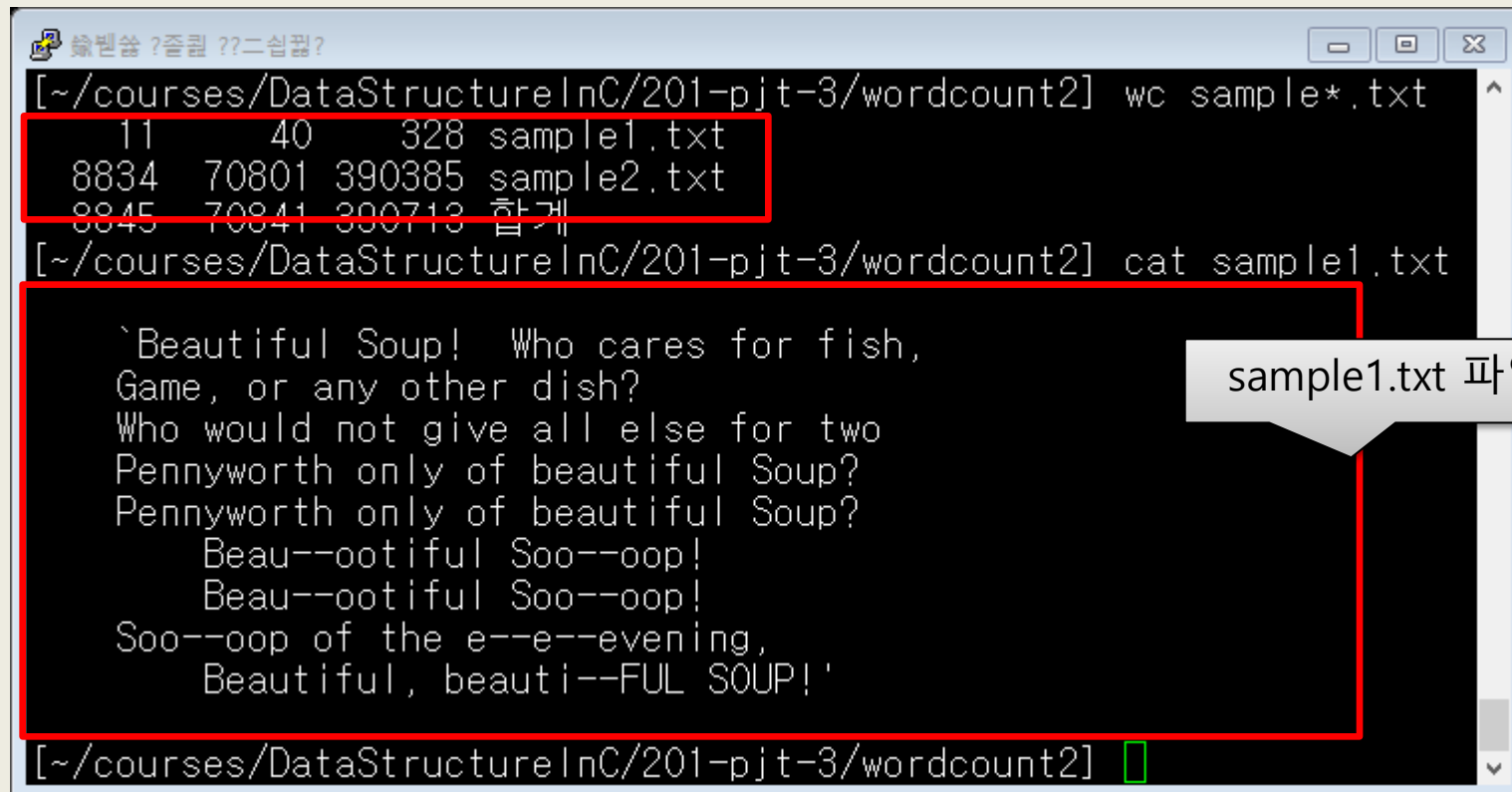
■ 구현된 프로그램은 사용자가 지정한 정렬방법에 따라 정렬 수행할 것

- 정렬 알고리즘은 "빠른" > "중간" > "느린" 방법들 중에서 각각 1종씩 선택하고,
- 선택한 정렬 알고리즘들을 3가지 출력 기능에 1대1로 매핑 후 구현
(적용 예) 출력 1. → (느린)삽입정렬, 출력 2. → (중간)셸정렬,
출력 3. → (빠른)퀵 정렬
- Unix 라이브러리 함수인 `qsort()` 사용을 권장하지 않습니다!

3차 과제 소개

■ 입력 파일 2종 제공

- 파일 1. sample1.txt → 시험용으로만 사용할 것
- 파일 2. sample2.txt → 보고서 제출과 시연용으로 사용할 것



A terminal window titled "터미널 ? 줄임 ??二침필?" showing the execution of two commands. The first command, `wc sample*.txt`, is followed by a red box highlighting the output: `11 40 328 sample1.txt`, `8834 70801 390385 sample2.txt`, and `8845 70841 390713 합계`. The second command, `cat sample1.txt`, is followed by a red box highlighting the output, which is the text of sample1.txt. A callout bubble points to this output with the text "sample1.txt 파일 내용".

```
[~/courses/DataStructureInC/201-pjt-3/wordcount2] wc sample*.txt
11 40 328 sample1.txt
8834 70801 390385 sample2.txt
8845 70841 390713 합계
[~/courses/DataStructureInC/201-pjt-3/wordcount2] cat sample1.txt
`Beautiful Soup! Who cares for fish,
Game, or any other dish?
Who would not give all else for two
Pennyworth only of beautiful Soup?
Pennyworth only of beautiful Soup?
    Beau--ootiful Soo--oop!
    Beau--ootiful Soo--oop!
    Soo--oop of the e--e--evening,
    Beautiful, beauti--FUL SOUP!'
[~/courses/DataStructureInC/201-pjt-3/wordcount2]
```

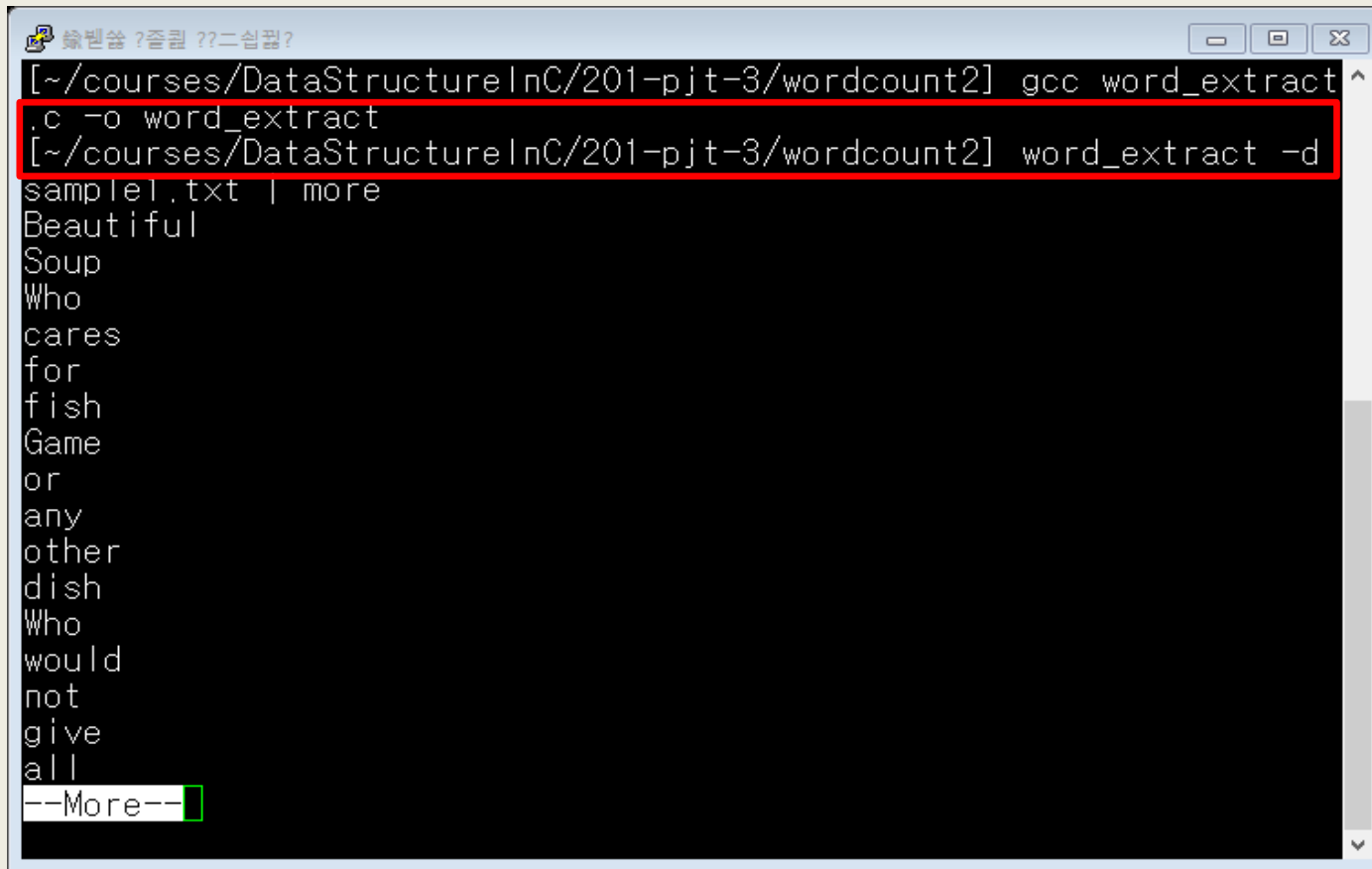
sample1.txt 파일 내용

3차 과제 소개

■ 입력 파일에 포함된 영문 단어를 추출하는 프로그램 제공

- 파일 이름 : word_extract.c

- ① 영문 단어와 빈도수 집계 기능 구현할 때 활용할 것



```
[~/courses/DataStructureInC/201-pjt-3/wordcount2] gcc word_extract.c -o word_extract
[~/courses/DataStructureInC/201-pjt-3/wordcount2] word_extract -d sample1.txt | more
Beautiful
Soup
Who
cares
for
fish
Game
or
any
other
dish
Who
would
not
give
all
--More--
```


3차 과제 소개

■ 사용자의 명령줄로부터 정렬 방법 입력받아 관련 기능 호출

② 사용자가 지정한 정렬 알고리즘 입력받아 처리하는 기능

→ 상세 내용은 12쪽 참조할 것

→ 관련 내용은 다음을 참조하세요.

■ 참조 : C언어본색(명강의가 일으키는 C언어의 기적)

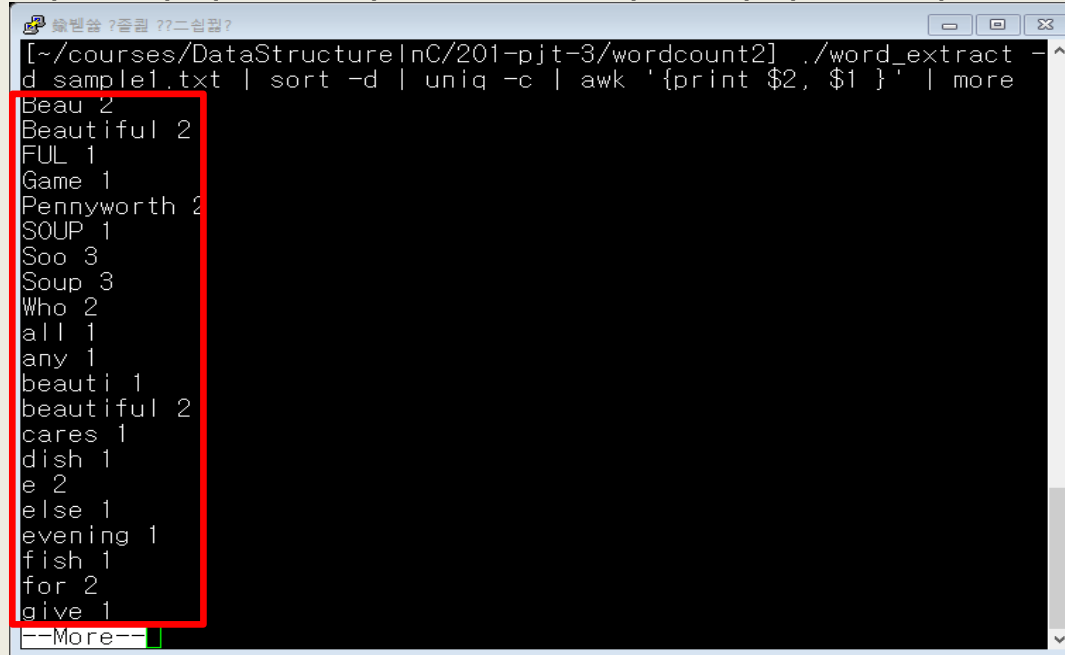
- 박정민 교수 저, C언어본색(명강의가 일으키는 C언어의 기적)

- Part 2, 제5장 포인터와 함수 그리고 void형 포인터, 5.3 main() 함수에 인자가 있을 때(p454)

3차 과제 소개

■ 단어 추출 프로그램을 사용한 출력 1. 예시

③ 영어 단어의 오름차순으로 영어 단어와 빈도수를 함께 출력 기능



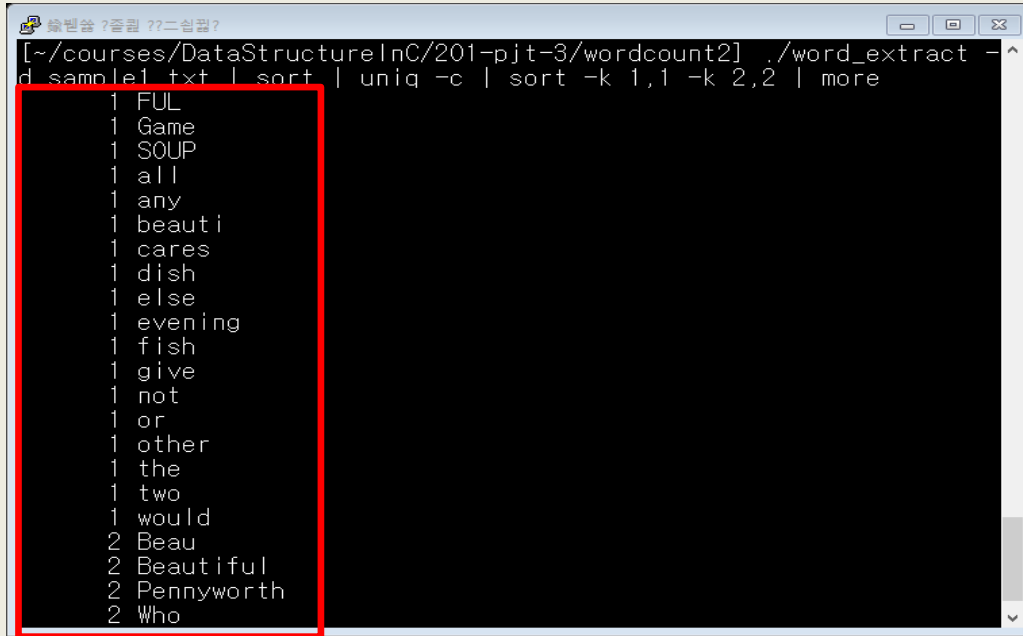
```
[~/courses/DataStructureInC/201-pjt-3/wordcount2] ./word_extract -^
d sample1.txt | sort -d | uniq -c | awk '{print $2, $1 }' | more
Beau 2
Beautiful 2
FUL 1
Game 1
Pennyworth 2
SOUP 1
Soo 3
Soup 3
Who 2
all 1
any 1
beauti 1
beautiful 2
cares 1
dish 1
e 2
else 1
evening 1
fish 1
for 2
give 1
--More--
```

- 출력 1.을 얻기 위한 셸 명령어 모음(참조만 하세요!)
 - `./word_extract -d sample1.txt | sort -d | uniq -c | awk '{print $2, $1 }'`
- 단어 추출 프로그램 실행 결과를 파이프로 입력받아 ③과 같이 출력되도록 밑줄에서 수행하는 기능만을 구현할 것

3차 과제 소개

■ 단어 추출 프로그램을 사용한 출력 2. 예시

④ 빈도수의 오름차순으로 빈도수와 영어 단어 함께 출력 기능



```
[~/courses/DataStructureInC/201-pjt-3/wordcount2] ./word_extract -d sample1.txt | sort | uniq -c | sort -k 1,1 -k 2,2 | more
1 FUL
1 Game
1 SOUP
1 all
1 any
1 beauti
1 cares
1 dish
1 else
1 evening
1 fish
1 give
1 not
1 or
1 other
1 the
1 two
1 would
2 Beau
2 Beautiful
2 Pennyworth
2 Who
```

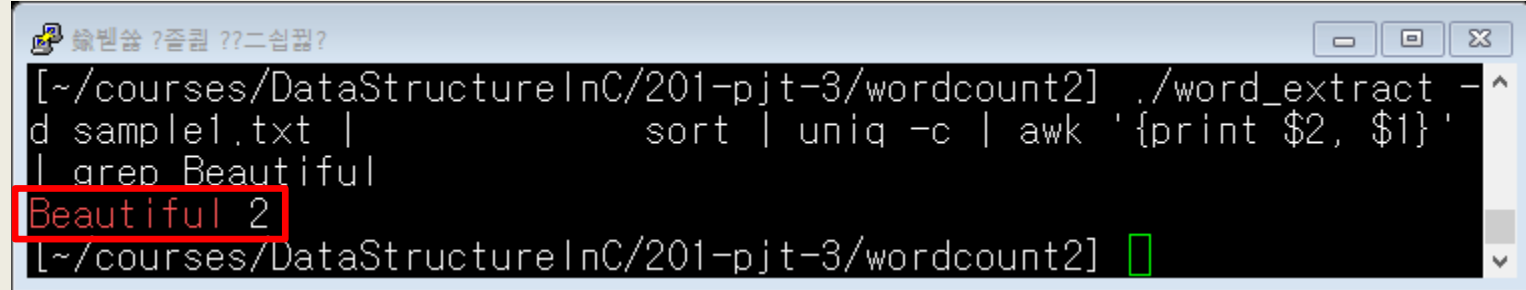
- 출력 2.을 얻기 위한 셸 명령어 모음(참조만 하세요!)
 - `./word_extract -d sample1.txt | sort | uniq -c | sort -k 1,1 -k 2,2`
- 단어 추출 프로그램 실행 결과를 파이프로 입력받아 ④와 같이 출력되도록 밑줄에서 수행하는 기능만을 구현할 것

3차 과제 소개

■ 단어 추출 프로그램을 사용한 출력 3. 예시

⑤ 제시된 영어 단어를 찾아 해당 단어와 빈도수를 함께 출력 기능

- 찾을 단어의 예 : Beautiful
- 실행 결과 : 찾을 단어(Beautiful)와 빈도수(2회) 출력



```
[~/courses/DataStructureInC/201-pjt-3/wordcount2] ./word_extract -d sample1.txt | sort | uniq -c | awk '{print $2, $1}' | grep Beautiful
Beautiful 2
[~/courses/DataStructureInC/201-pjt-3/wordcount2] █
```

■ 출력 3.을 얻기 위한 셸 명령어 모음(참조만 하세요!)

- `./word_extract -d sample1.txt | sort | uniq -c | awk '{print $2, $1}' | grep Beautiful`
- 단어 추출 프로그램 실행 결과를 파이프로 입력받아 ⑤와 같이 출력되도록 밑줄에서 수행하는 기능만을 구현할 것

3차 과제 소개

■ 다음과 같이 동작하는 코드 구현

- 단어 추출 프로그램을 사용하여 다음과 같이 동작하는 프로그램 1종을 구현
 - 구현된 실행 파일 이름을 **WordCount**로 가정
 - 초기 구현 단계에서 sample1.txt 파일 사용하고, 최종 검증 단계에서 sample2.txt 사용

■ 입력 파일로 sample2.txt 사용하여 실행할 경우 3개 명령어 내용

- ③과 같이 출력하고자 할 때 사용자의 명령줄 입력 내용
 - **./word_extract -d sample2.txt | WordCount -1**
- ④와 같이 출력하고자 할 때 명령줄 입력 내용
 - **./word_extract -d sample2.txt | WordCount -2**
- ⑤와 같이 출력하고자 할 때 명령줄 입력 내용
 - **./word_extract -d sample2.txt | WordCount -3 Beautiful**

3차 과제 소개

■ 3차 과제 평가 항목 및 배점

구분		평가항목 1	평가항목 2	평가항목 3	평가항목 4	점수 계
3차 과제	평가항목	동작 시나리오 적절성	구현 내용 설명 적절성	구현 결과 적절성	발표 참여 시	
	점수	4	3	4	4	15

3차 과제 소개

■ 3차 과제 작성 및 제출

- 파일 이름은 "자구-3차과제-학생이름.ppt"로 할 것
- 3차 과제 결과보고서 포함 내용
 - 설계 및 동작 시나리오
 - 구현 내용 설명, 구현 결과, 교훈 등
- 6/15(월) 오후 6시까지 e-class에 다음 항목 업로드할 것
 - 3차 과제 결과보고서
 - 12분 발표와 3분 Q&A 진행 대비하여 파워포인트로 작성(10쪽 이상)
 - 구현된 프로그램(Linux 서버에서 실행 가능할 것)

3차 과제 소개

■ 3차 과제 발표

- 6/17(수) ~ 6/22(월), 선착순으로 총 12명의 발표 희망자 모집 후 진행
 - 발표 참여 시 부가점수 있음
 - 참여 희망자는 참여 여부와 참여 형식(대면 또는 비대면 발표)을 꼭지로 발송할 것
 - 12명 명단 확정되는 대로 대면 및 비대면 발표자 명단 공지 예정
 - 비대면 발표자는 15분 발표 동영상을 과제와 함께 제출
 - » 동영상 상영 시간은 15분 이내로 할 것(시간 초과 시 강제 종료)
 - » 이클래스에 용량 문제로 업로드 안될 경우 제 이메일로 보내주세요!
 - 참여 희망자가 부족할 경우 발표자 지명
 - 지명되어도 발표에 참여하지 않을 수 있음
 - 발표가 진행되지 않는 시간은 일정 조정 후 추가 강의 진행함

Q & A

1. 의문 사항있으면 질문해 주세요.
2. E-class, e-mail, SMS 등 온라인으로,
연구실 방문 또는 전화 등 오프라인 모두 가능합니다.
3. 하지만, 가급적 온라인을 활용하여 문의바랍니다.