

과제 #06

데이터사이언스를 위한 컴퓨팅 1 (2022년도 1학기, M3239.005500)

Due: 2022년 05월 25일(수) 23시 59분

1 Binary Search Tree (5점)

배포된 과제 파일의 bst 디렉토리에 binary search tree가 구현되어 있다. 함께 제공되는 Makefile을 이용하여 make 한 후 실행한 결과 화면을 첨부하라.

2 Red-black Tree (5점)

배포된 과제 파일의 rbt 디렉토리에 red-black tree가 구현되어 있다. 함께 제공되는 Makefile을 이용하여 make 한 후 실행한 결과 화면을 첨부하라.

3 Set container

3.1 MySet 구현 (20점)

STL의 set과 유사한 container를 binary search tree 와 red-black tree를 사용하여 구현할 것이다. 배포된 과제 파일의 container 디렉토리에 있는 bst.cpp (binary search tree 구현 파일) 와 rbt.cpp (red-black tree가 구현 파일)를 사용할 수 있다. MySet은 다음과 같은 인터페이스를 가진다.

- void Insert(int data)
Set에 data를 값으로 가지는 원소를 삽입, 주의할 점은 STL set과 마찬가지로 MySet은 중복된 원소의 삽입을 허용하지 않는다
- void Erase(int data)
Set에서 data와 같은 값을 가지는 원소를 제거
- bool Find(int data)
Set에 data와 같은 값을 가지는 원소가 있는지 검색, 있으면 true를 반환하고 없으면 false를 반환
- size_t Size()
Set에 들어있는 모든 원소의 개수를 리턴

myset.h 파일에 위의 인터페이스가 정의되어 있으며, myset.cpp 파일에 인터페이스의 내용을 구현한다. MySet 구현을 위해 binary search tree 와 red-black tree 중 어떤 것을 선택할 지는 자유이다. 함께 제공되는 Makefile을 이용하여 소스코드를 컴파일하고 실행할 수 있다. MySet을 구현한 후 실행한 결과가 아래와 같으면 된다.

```
Size of the set is 6
Size of the set after inserts and erases is 5
Is number '32' exists in the set ? 0
Is number '22' exists in the set ? 0
Is number '100' exists in the set ? 1
```

3.2 MySet 성능 측정 (70점)

MySet에서 제공하는 4가지 함수에 대해 binary search tree를 사용하여 구현했을 때와 red-black tree를 사용하여 구현하였을 때의 성능을 비교해보고 그 이유를 분석하여 상세히 서술하라. 성능이 잘 나오는 상황, 성능이 잘 나오지 않는 상황 등의 시나리오를 자유롭게 구성하여 실험을 설계하고 분석하라.

- Insert 함수 성능 측정 및 분석
- Erase 함수 성능 측정 및 분석
- Find 함수 성능 측정 및 분석
- Size 함수 성능 측정 및 분석

4 참고 사항

성능 측정 시에는 다음의 함수를 코드에 삽입하여 사용할 수 있다.

```
#include <time.h>    // 헤더 파일

struct timespec t_begin, t_end;
clock_gettime(CLOCK_MONOTONIC, &t_begin);    // 성능 측정 시작 부분

... 성능 측정할 코드 ...

clock_gettime(CLOCK_MONOTONIC, &t_end);    // 성능 측정 종료 부분

double diff = (t_end.tv_sec - t_begin.tv_sec) * 1000.0 +
    (t_end.tv_nsec - t_begin.tv_nsec) / 1000.0 / 1000.0;
printf("Processing time: %.3lfms\n", diff);
```

5 Submission Instruction

- myset.cpp, myset.h, report.pdf를 한 파일로 압축하여 ETL에 제출한다.
- myset.cpp와 myset.h는 문제 3.1에 대한 구현만 제출하면 되며, 문제 3.2에 관련된 코드는 제출하지 않는다.
- 문제 3.1 해결 과정에서 bst.cpp, bst.h, rbt.cpp, rbt.h 파일을 수정할 경우 함께 제출할 수 있다.
- 첨부 파일명은 계정이름 HW06.zip으로 한다. (e.g., cfd999.HW06.zip)
- 본인의 계정을 모를 경우 ETL 자료실에 업로드 되어있는 게시글에서 확인한다. 게시글에 본인의 이름이 적혀있지 않은 경우 조교에게 이메일로 문의한다.
- Grace day를 사용하려면 본인이 과제를 제출한 날에 조교에게 메일(cfdsta@aces.snu.ac.kr)로 알려야 한다. 메일 없이 제출만 한 경우 다음 과제를 위해 아낀 것으로 판단, 미제출 처리된다. 또한, grace day 사용 시에도 과제 제출은 이메일이 아닌 ETL을 통해 해야한다.
- 채점은 프로그램에 의해 기계적으로 처리되므로 위 사항을 지키지 않은 경우 채점 대상에서 누락되거나 감점을 받을 수 있다. 자세한 감점 기준은 ETL 공지사항을 참조한다.