

Programming Assignment#1

Due : 4th Oct. (Thur),5:59 PM

1. Introduction

이번 과제를 통해 해시테이블 자료구조에 익숙해지고 Key-Value 데이터베이스에 대해 이해해 본다.

2. Problem specification

Key-Value(KV) 데이터베이스는 간단한 Key-Value 함수를 사용하여 데이터를 저장하는 비관계형 데이터베이스이다.

이번 과제에서 구현해야할 KV 데이터베이스(DB)의 API와 역할은 다음과 같다.

```
db_t* db_open(int size);
```

- 데이터베이스에 필요한 자료구조를 생성하고 데이터베이스 구조체의 포인터를 반환한다.
- 해시 테이블의 엔트리 개수는 주어진 size로 만든다.

```
void db_close(db_t* db);
```

- 데이터베이스에서 사용했던 자원들을 모두 해제한다.

```
void db_put(db_t*db, char* key, int keylen, char* val, int vallen);
```

- 주어진 key와 value를 데이터 베이스에 저장한다.

```
char* db_get(db_t* db, char* key, int keylen, int* vallen);
```

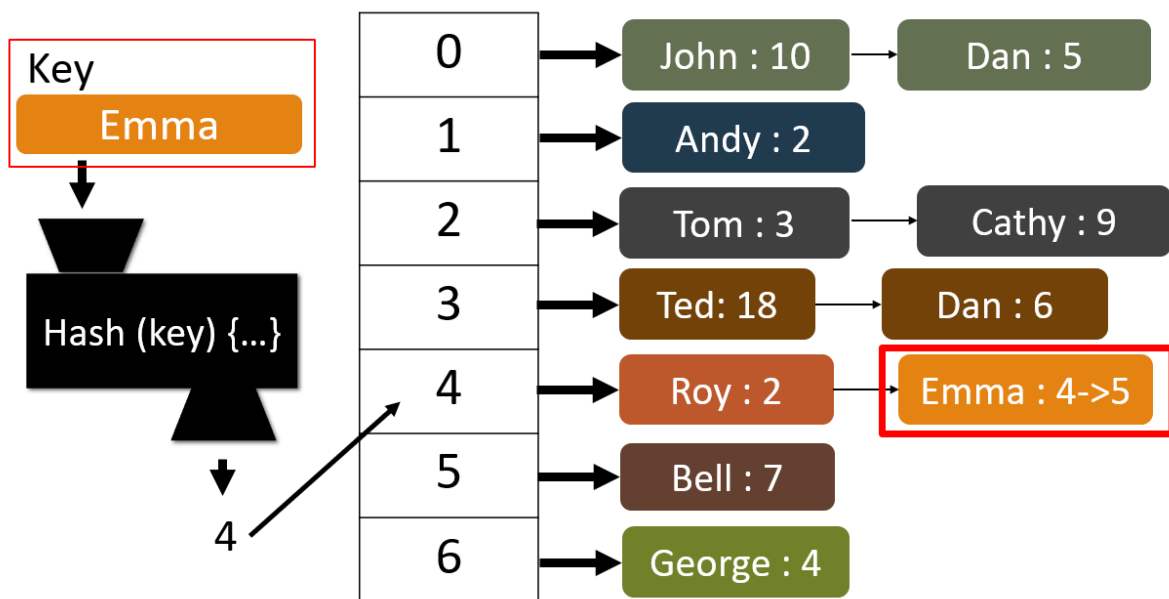
- key값에 해당하는 value를 char*의 형태로 반환한다.
- 주어진 key가 존재하지 않는다면 NULL
- 존재한다면 malloc()함수로 메모리 공간을 할당하고, 해당하는 value를 복사해 반환한다.

```
void db_put(db_t*db, char* key, int keylen, char* val, int vallen);
```

db_t* db	db_open()의 호출로 반환받은 데이터베이스 구조체 포인터
char* key	저장하려는 key
int keylen	주어진 key값의 길이
char* val	저장하려는 value
int vallen	주어진 value의 길이

```
char* db_get(db_t* db, char* key, int keylen, int* vallen);
db_t* db      db_open()의 호출로 반환받은 데이터베이스 구조체 포인터
char* key     탐색하려는 key
int keylen   key값의 길이
int *vallen  DB에 key값이 존재한다면 해당하는 value값의 길이를 반환
```

2-1. Hash Table



- Input data를 hash function을 거쳐 index를 얻어낸 후, 이를 테이블 내의 index로 활용하여 알맞은 entry에 key-value data를 저장하는 방식.
- Index가 서로 충돌하는 key-value pair가 발생할 경우, 이들은 서로 linked list로 연결된다.
- 본 과제에서 Database에 들어갈 자료들은 hash table로 관리되어야 한다.
- Hash function은 자유롭게 구현한다.

Reference : <https://www.hackerearth.com/practice/data-structures/hash-tables/basics-of-hash-tables/tutorial/>

4. Hand in instructions

- ✓ 홈페이지에 업로드 된 skeleton 파일은 아래의 파일들로 이루어져 있다.

Makefile:	GNU make도구를 위해 필요한 파일
main.c:	표준입력으로 단어를 읽어서 개수를 카운트
db.c:	데이터베이스의 오퍼레이션을 구현하는 파일
db.h:	데이터베이스의 오퍼레이션 API

- ✓ 작성한 프로그램 코드 상단에 이름과 학번을 적는다.
- ✓ 과제는 제출 시 "학번.tar.gz"로 압축한다.
 - 압축 파일은 Makefile, main.c, README.pdf로 이루어져 있어야 하며, 압축파일의 이름과 확장자는 학번.tar.gz 여야 한다.
- ✓ 프로그램 코드와 별도로, 구현에 대한 내용을 담은 보고서를 함께 제출한다. 보고서의 파일 포맷은 pdf로 제한하며, 형식에는 제한이 없다. 제목은 README.pdf로 한다.
- ✓ 과제는 icampus의 과제란에 제출한다.
- ✓ 과제 제출 시간은 icampus 제출 시간을 기준으로 하며, 기한 이후엔 10%씩 감점되고, 과제 제출은 추가 제출기간이 마감되기 전까지 제출 가능합니다.
- ✓ 본 과제는 혼자서 한다.
- ✓ GNU make 도구는 큰 프로그램을 만들 때 유용하게 사용되는데, 프로그램을 제작하기 위해 어떤 코드를 (재)컴파일해야 하는지 결정해준다. 일단 Makefile이 준비되면, 어떤 소스 코드를 변경하던지 셸에서 단순히 make란 명령을 실행시키는 것으로 재 컴파일이 필요한 모든 파일을 알아서 찾아 다시 컴파일한다.
- ✓ **Copy 할 경우, 연구실 자체 규정에 따라 처벌하며, 상당한 불이익이 있을 수 있다.**

5. Examples

INPUT	OUTPUT
\$/wordcount 128	DB opened
EMMA	GET [EMMA] [NULL]
	PUT [EMMA] [1]
EMMA	GET [EMMA] [1]
	PUT [EMMA] [2]
JOHN	GET [JOHN] [NULL]
	PUT [JOHN] [1]
	DB closed

Have fun!

컴퓨터시스템연구실