

The background features a dark blue gradient with faint, light blue concentric circles and degree markings (40, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260) on the left side. There are also several circular arrows indicating rotation.

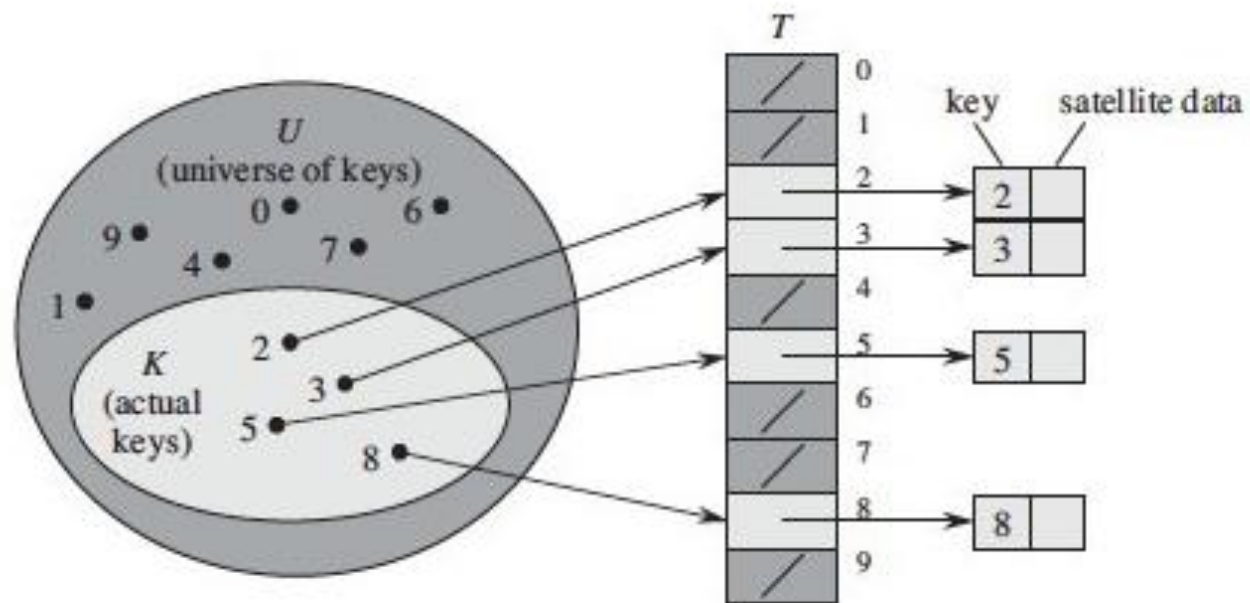
KT AIVLE 코딩 스터디 - HASH 알고리즘

성용제

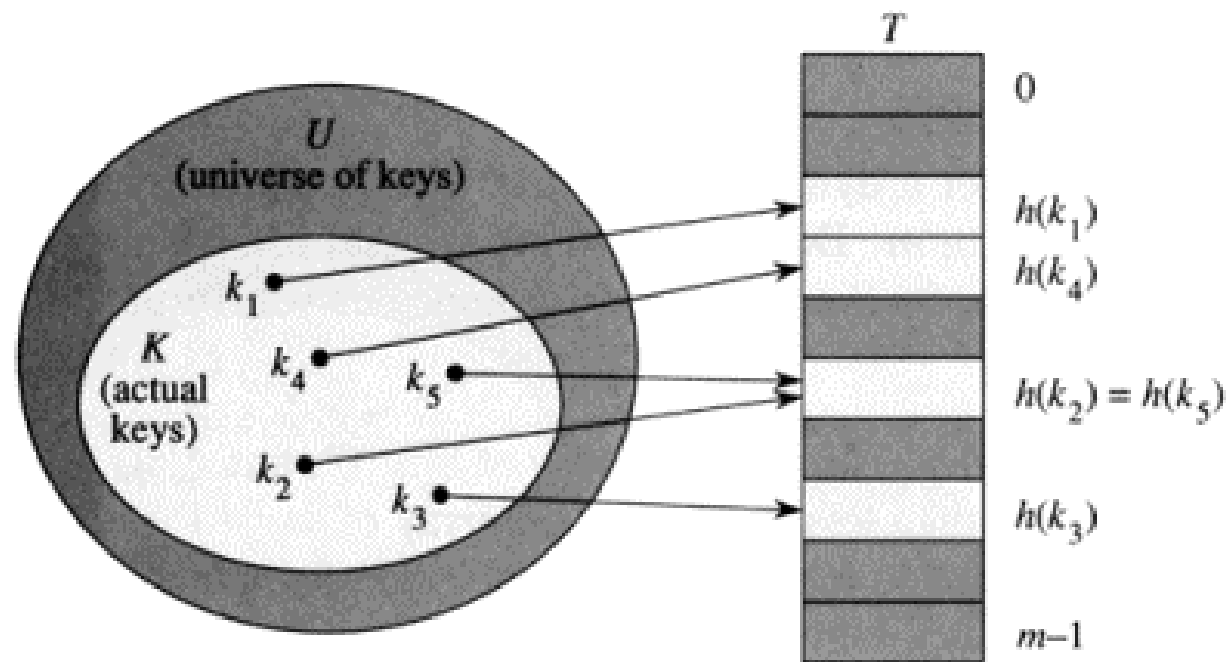
HASH 알고리즘

- 임의의 크기를 가진 데이터(Key)를 고정된 크기의 데이터(Value)로 변화시켜 저장하는 것
- 키에 대한 해시 값을 사용하여 값을 저장하고 Key값Value쌍의 개수에 따라 동적으로 크기가 증가하는 associate array
- 키에 대한 해시값을 구하는 과정을 hashing(해싱)이라고 하며 이때 사용하는 함수(알고리즘)를 해시 함수라고 함
- 해시 값 자체를 index로 사용하기 때문에 평균 시간 복잡도가 $O(1)$ 로 매우 빠름

1. DIRECT ADDRESSING TABLE



2. HASH TABLE

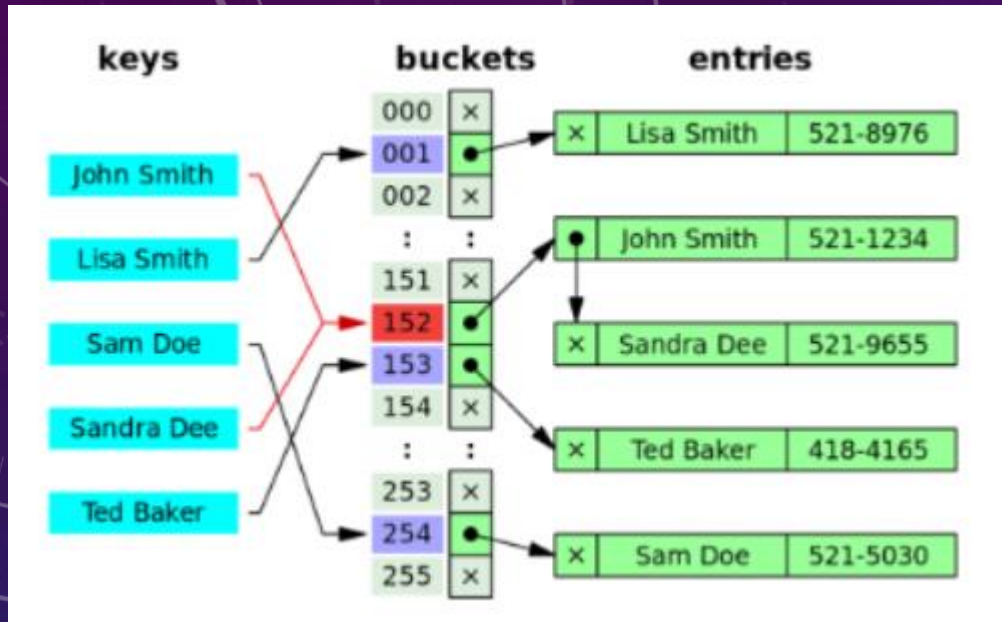


3. COLLUSION

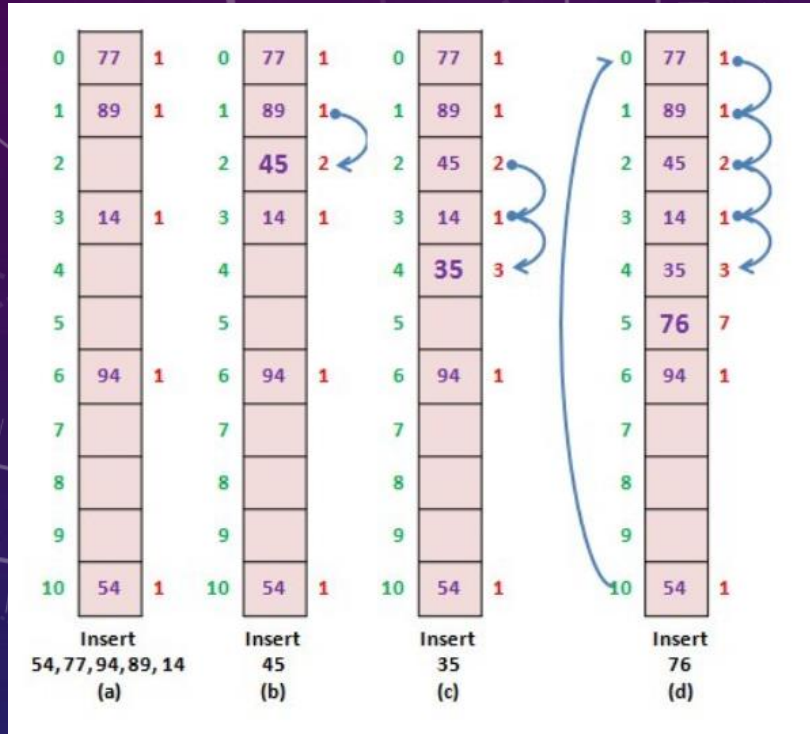
해쉬 충돌 : 서로 다른 key 값이 동일한 $h(k)$ 값을 갖는 것

효율성을 떨어뜨림 해시 테이블의 성능 저하

3.1 분리 연결법 (SEPARATE CHAINING)



3.2 개방 주소법 (OPEN ADDRESSING)



4. ALGORITHM

1. MD5 : key 값을 128비트 길이의 해쉬 값으로 출력

단방향 암호화로 입력 값 복원 불가 같은 입력 값은 매번 같은 출력 값을 갖는다.

충돌 확률은 매우 적다.

2. SHA : 미국 NSA(국가안보국 국방부)가 제작한 방법

한국에서는 SHA-256사용-> 256비트 결과를 출력

5. STL -> UNORDERED_MAP

```
#include <unordered_map>
```

함수

```
empty()
```

```
size()
```

```
find(key)
```

```
count(key)
```

```
Insert({ key, value }) : arr[key]=value
```

```
erase(key)
```

6.1 문제1) 듣보잡 – 1764 – 실버 4

문제

김진영이 듣도 못한 사람의 명단과, 보도 못한 사람의 명단이 주어질 때, 듣도 보도 못한 사람의 명단을 구하는 프로그램을 작성하시오.

입력

첫째 줄에 듣도 못한 사람의 수 N , 보도 못한 사람의 수 M 이 주어진다. 이어서 둘째 줄부터 N 개의 줄에 걸쳐 듣도 못한 사람의 이름과, $N+2$ 째 줄부터 보도 못한 사람의 이름이 순서대로 주어진다. 이름은 띄어쓰기 없이 알파벳 소문자로만 이루어지며, 그 길이는 20 이하이다. N , M 은 500,000 이하의 자연수이다.

듣도 못한 사람의 명단에는 중복되는 이름이 없으며, 보도 못한 사람의 명단도 마찬가지이다.

출력

듣보잡의 수와 그 명단을 사전순으로 출력한다.

예제 입력 1 복사

```
3 4
ohhenrie
charlie
baesangwook
obama
baesangwook
ohhenrie
clinton
```

예제 출력 1 복사

```
2
baesangwook
ohhenrie
```

6.2 문제2) 문자열 집합-14425 – 실버3

문제

총 N 개의 문자열로 이루어진 집합 S 가 주어진다.

입력으로 주어지는 M 개의 문자열 중에서 집합 S 에 포함되어 있는 것이 총 몇 개인지 구하는 프로그램을 작성하시오.

입력

첫째 줄에 문자열의 개수 N 과 M ($1 \leq N \leq 10,000$, $1 \leq M \leq 10,000$)이 주어진다.

다음 N 개의 줄에는 집합 S 에 포함되어 있는 문자열들이 주어진다.

다음 M 개의 줄에는 검사해야 하는 문자열들이 주어진다.

입력으로 주어지는 문자열은 알파벳 소문자로만 이루어져 있으며, 길이는 500을 넘지 않는다. 집합 S 에 같은 문자열이 여러 번 주어지는 경우는 없다.

출력

0 첫째 줄에 M 개의 문자열 중에 총 몇 개가 집합 S 에 포함되어 있는지 출력한다.

예제 입력 1 복사

```
5 11
baekjoononlinejudge
startlink
codeplus
sundaycoding
codingsh
baekjoon
codeplus
codeminus
startlink
starlink
sundaycoding
codingsh
codingsh
sondaycoding
startrink
icerink
```

예제 출력 1 복사

```
4
```

6.3 문제3) 패션왕 신해빈 – 9375 – 실버3

문제

해빈이는 패션에 매우 민감해서 한번 입었던 옷들의 조합을 절대 다시 입지 않는다. 예를 들어 오늘 해빈이가 안경, 코트, 상의, 신발을 입었다면, 다음날은 바지를 추가로 입거나 안경대신 렌즈를 착용하거나 해야한다. 해빈이가 가진 의상들이 주어졌을때 과연 해빈이는 알몸이 아닌 상태로 며칠동안 밖에 돌아다닐 수 있을까?

입력

첫째 줄에 테스트 케이스가 주어진다. 테스트 케이스는 최대 100이다.

- 각 테스트 케이스의 첫째 줄에는 해빈이가 가진 의상의 수 $n(0 \leq n \leq 30)$ 이 주어진다.
- 다음 n 개에는 해빈이가 가진 의상의 이름과 의상의 종류가 공백으로 구분되어 주어진다. 같은 종류의 의상은 하나만 입을 수 있다.

모든 문자열은 1이상 20이하의 알파벳 소문자로 이루어져있으며 같은 이름을 가진 의상은 존재하지 않는다.

출력

각 테스트 케이스에 대해 해빈이가 알몸이 아닌 상태로 의상을 입을 수 있는 경우를 출력하시오.

예제 입력 1 복사

```
2
3
hat headgear
sunglasses eyewear
turban headgear
3
mask face
sunglasses face
makeup face
```

예제 출력 1 복사

```
5
3
```


6.4 문제4) 추월 – 2002 – 실버1

문제

대한민국을 비롯한 대부분의 나라에서는 터널 내에서의 차선 변경을 법률로 금하고 있다. 조금만 관찰력이 있는 학생이라면 터널 내부에서는 차선이 파선이 아닌 실선으로 되어 있다는 것을 알고 있을 것이다. 이는 차선을 변경할 수 없음을 말하는 것이고, 따라서 터널 내부에서의 추월은 불가능하다.

소문난 명콤비 경찰 대근이와 영식이가 추월하는 차량을 잡기 위해 한 터널에 투입되었다. 대근이는 터널의 입구에, 영식은 터널의 출구에 각각 잠복하고, 대근이는 차가 터널에 들어가는 순서대로, 영식은 차가 터널에서 나오는 순서대로 각각 차량 번호를 적어 두었다.

N개의 차량이 지나간 후, 대근이와 영식은 자신들이 적어 둔 차량 번호의 목록을 보고, 터널 내부에서 반드시 추월을 했을 것으로 여겨지는 차들이 몇 대 있다는 것을 알게 되었다. 대근이와 영식을 도와 이를 구하는 프로그램을 작성해 보자.

입력

입력은 총 $2N+1$ 개의 줄로 이루어져 있다. 첫 줄에는 차의 대수 $N(1 \leq N \leq 1,000)$ 이 주어진다. 둘째 줄부터 N개의 줄에는 대근이가 적은 차량 번호 목록이 주어지고, N+2째 줄부터 N개의 줄에는 영식이 적은 차량 번호 목록이 주어진다. 각 차량 번호는 6글자 이상 8글자 이하의 문자열로, 영어 대문자('A'-'Z')와 숫자('0'-'9')로만 이루어져 있다.

같은 차량 번호가 두 번 이상 주어지는 경우는 없다.

출력

첫째 줄에 터널 내부에서 반드시 추월을 했을 것으로 여겨지는 차가 몇 대인지 출력한다.

예제 입력 1 복사

```
4
ZG431SN
ZG5080K
ST123D
ZG206A
ZG206A
ZG431SN
ZG5080K
ST123D
```

예제 출력 1 복사

```
1
```

6.5 문제5) 왕위 계승 – 5021 – 골드5

문제

유토피아의 왕이 사망했다. 왕은 자손을 남기지 않고 사망했기 때문에, 왕위를 계승할 사람을 지목하지 않았다. 왕실 귀족은 왕위를 주장하기 시작했다. 유토피아의 법에는 왕의 계승자가 없는 경우에, 나라를 세운 사람과 혈통이 가까운 사람이 유토피아를 통치한다는 조항이 있다.

나라를 세운 사람과 혈통이 가장 가까운 사람은 그의 자손이 아닌 사람과 피가 덜 섞인 사람이다. 모든 사람은 아버지의 혈통과 어머니의 혈통을 반 씩 받게 된다. 나라를 세운 사람의 자식은 1/2 왕족이며, 그 아들이 왕족이 아닌 사람과 결혼한 경우에는 아들의 자식은 1/4 왕족이 된다.

가장 나라를 세운 사람과 혈통이 가까운 사람을 찾는 프로그램을 작성하시오.

입력

첫째 줄에 N 과 M 이 주어진다. ($2 \leq N, M \leq 50$)

둘째 줄에 유토피아를 세운 사람의 이름이 주어진다.

다음 N 개 줄에는 가족 정보가 주어진다. 정보는 이름 세 개로 이루어져 있고, 공백으로 구분되어져 있다. 첫 번째 이름은 자식의 이름이고, 뒤의 두 이름은 부모의 이름이다.

다음 M 개 줄에는 왕위를 계승받기를 주장하는 사람의 이름이 한 줄에 하나씩 주어진다.

모든 이름은 1~10글자이며, 알파벳 소문자로만 이루어져 있다. 나라를 세운 사람이 왕위를 계승하는 경우나, 누군가의 자식인 경우는 없다.

출력

첫째 줄에 유토피아를 세운 사람과 가장 혈통이 가까운 사람의 이름을 출력한다. 항상 답이 유일한 경우만 입력으로 주어진다.

문제에 주어지는 가족 관계는 성별, 나이를 고려하지 않고 만들었기 때문에, 실제로는 말이 안되는 경우가 나올 수도 있다. 하지만, 모든 자식의 부모는 유일하며, 다시 자식이 부모의 부모가 되는 경우도 없다. 또, 한 사람이 여러 명의 자식이 될 수 도 없다.

예제 입력 1 복사

```
9 2
edwardi
charlesi edwardi diana
philip charlesi mistress
wilhelm mary philip
matthew wilhelm helen
edwardii charlesi laura
alice laura charlesi
helen alice bernard
henrii edwardii roxane
charlesii elizabeth henrii
charlesii
matthew
```

예제 출력 1 복사

```
matthew
```