

〈용어체크〉

해시 테이블

해시 테이블은 데이터가 저장되는 버깃들의 배열로 만들어지고 한 버깃은 하나 이상의 레코드를 수용할 수 있으며 해시 테이블을 이용해서 탐색한다.

해시 함수

하나의 문자열을 보다 빨리 찾을 수 있도록 주소에 직접 접근할 수 있는 짧은 길이의 값이나 키로 변환하는 알고리즘을 수식으로 표현한 것이며 주어진 키 값으로부터 레코드가 저장되어 있는 주소를 산출해 내는 수식이다.

충돌

해시 테이블의 한 주소를 놓고 두 개 이상의 원소가 자리를 다투는 것으로 한 원소를 해싱해서 저장하려는데 다른 원소가 이미 그 자리를 차지한 상황이다.

〈학습내용〉

해시 테이블의 개념

해시 함수

해시 테이블의 충돌

〈학습목표〉

해시 테이블이 무엇인지 설명할 수 있다.

해시 함수의 설계 원리를 설명할 수 있다.

해시 테이블에서 충돌이 무엇인지 설명할 수 있다.

Q. 자료구조의 대표적인 기능은 자료를 저장하고 필요할 때 찾는 것, 즉 검색입니다. 가능하면 검색 시간이 빠른 것이 좋은데 저장된 자료의 양이 많더라도 원소 하나를 저장하고 검색하는데 한번에 찾을 수 있게 하는 방법은 어떤 것이 있을까요?

: 우리가 이전에 배운 검색 트리는 상당히 효율적으로 자료를 저장하고 검색을 할 수 있습니다. 그러나 검색 트리는 검색하려는 원소와 검색 트리에 있는 원소를 루트부터 비교하여 해당 원소의 자리를 찾아 내려가므로 일정한 시간이 걸립니다. 그렇다면 저장된 자료의 양에 상관없이 항상 상수 시간에 검색이 가능하게 하는 방법은 없을까? 이 꿈을 실현하는 것이 해시 테이블입니다. 해시 테이블은 자료의 저장과 검색에 극단적인 효율을 보이는 자료구조입니다. 해시 테이블은 검색 트리와는 발상이 완전히 다르며 원소끼리 비교해서 자리를 찾는 것이 아니라 자신의 값이 자신의 자리를 바로 결정하므로 매우 빠른 응답이 가능합니다.

해시 테이블의 개념

해싱

- ✓ 해시 함수를 이용하여 레코드가 저장되어 있는 주소를 직접 구하여 검색한다.
- ✓ 키를 비교하지 않고 계산에 의해 검색하는 방법이다.
- ✓ 삽입과 삭제가 빈번한 자료에 적합하다.
- ✓ 키(key) 값을 가지고 있어서 데이터를 찾기 편리한 구조이다.
- ✓ 평균 상수 시간에 삽입, 삭제, 검색한다.

해시 테이블

- ✓ 레코드를 한 개 이상 보관하는 버킷들의 집합이다.
- ✓ 해시 테이블은 데이터가 저장되는 버킷들의 배열로 만들어지며 한 버킷은 하나 이상의 레코드를 수용할 수 있다.

해시 함수

주어진 키 값으로부터 레코드가 저장되어 있는 주소를 산출해 내는 수식이다.
나누기 방법, 곱하기 방법, 중간제곱법, 접지법 등이 있다.

해시 테이블의 충돌

해시 테이블의 한 주소를 놓고 두 개 이상의 원소가 자리를 다투는 것이다.
한 원소를 해싱해서 저장하려는데 다른 원소가 이미 그 자리를 차지한 상황이다.