11

CST 405
Algorithm Analysis & Design

Al Lake
Oregon Institute of Technology

Chapter 11

Hash Tables

Hash Table

- A hash table directly addesses into an ordinary array.
- The usage of examining an arbitrary position in an array is *O*(1) time.

Direct Addressing

- **A direct-address table is an array in which each position or slot corresponds to a key.**

Hash Tables

- A hash function *h* is used to compute the slot from key *k*.
- If two keys hash to the same slot a collision occurs.
- To correct the collision chaining is done. Chaining shifts the slot *n* spaces to the left or right.

Hash Functions

- **A good hash function provides a uniform hashing: each key is equally likely to hash to an of the *m* slots, independent of where any other key has hashed to.**

Hashing Methods

- Division method
  - Map a key *k* into one of *m* slots by taking the remainder of *k* divided by *m*.
  - h(k) = k mod m
- Multiplication method
  - Multiply the key *k* by a constant A in the range 0 < A < 1 and extract the fractional part of kA.
  - Multiply this value by *m* and take the floor of the result.
  - h(k) = $\lfloor m\ (k\ A \bmod 1) \rfloor$
- Universal method
  - Choose the hash function ar random from a carefully designed class of functions.
  - Provably good performance on average, no matter what keys are chosen.

Double Hashing

- **Double hashing uses two functions, to position successive keys.**
  - **$h(k,i) = (h_1(k) + ih_2(k)) \bmod m$**

Perfect Hashing

- **A hash table that has no collisions.**

Perfect Minimal Hash

- A hash table that has no collisions and no empty slots.
- There is a 1-1 relationship between the hash table and the key.