Data Structures
How To Design A Good HAsh Function

Cain Susko

Queen's University
School of Computing

March 11, 2022

Algorithms

The djb2 algorithm was first reproted by Dan Berstien many years ago.

```
djb2(s): # s is a string
  a = 5381
  for x in s:
    a = a*33 + x
  return a
```

a=32*a + a +x
multiplying by a power of 2 is
just a left shift of the bits.

Python Hash In python, a hhash function could look like so:

```
hash(object)

object - the object whose hash value is to be returned
(integer, string, float)


# hash for integer unchanged
print('Hash for 181 is:', hash(181))

# hash for decimal
print('Hash for 181.23 is:',hash(181.23))

# hash for string
print('Hash for Python is:', hash('Python'))
```

```
Hash for 181 is: 181
Hash for 181.23 is:
530343892119126197
Hash for Python is:
2230730083538390373
```

Designing Your Own Hash Function

When Creating a Hash Function a large part of the process is figureing out what data to use from the input it hash the data. This is because is is common for the Hash Table to contain a besspoke dataa type for storing a specific class of data. The ocnvention is to design your algorithm, *analyze it*, and **iterate it**

Tip 1 Make Use of All Data, as this help avoid collisions

Tip 2 Try to Spread Out Values so that they are more Evenly Distributed.

In practice, is is very hard to derive the hash values independent of any patterns (which is what were trying to do with the quadratic probing and double hashing).
There are a few methods we can use in order to try and achive this even distrubution.

Division Method The idea is to map a key k to one of the$m$ slots by taking the remainder of k divided by m.

$$h(k) = k \mod m$$

Although this method is fast, prime values of m cannot be used.

Multiplication Method The algorithm for this method is generally as follows:

1. Multiply each key k by a constant A, where $0 < A < 1$
2. extract the fractional part of A, kA
3. multiply fractional part kA by m
4. take the floor of the above result:

$$h(k) = \lfloor m(kA - \lfloor kA \rfloor n(kA \mod 1) \rfloor$$

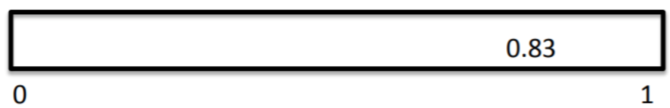
Examples

for Division

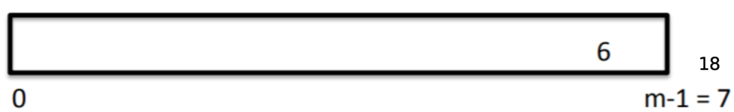
Key k = 3; m = 8 slots

(1) A = .61

(2) kA = 3 x .61 = 1.83 fractional part



(3) Floor (f m) = Floor(.83 x 8)
= Floor(6.64) = 6



18

and Multiplication

- The value of m is not critical now (e.g., $m = 2^p$)

assume $m = 2^3$

.101101 (A)
110101 (k)

1001010.0110011 (kA)

discard: 1001010

shift .0110011 by 3 bits to the left

011.0011

take integer part: 011

thus, h(110101)=011