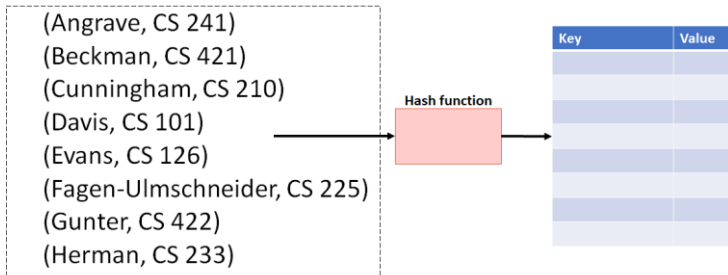


Every hash table contains three pieces:

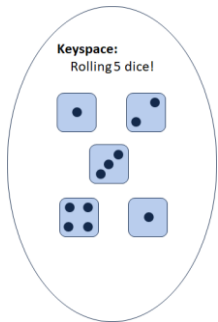
1. A **hash function, $f(k)$** . The hash function transforms a key from the keyspace into a small integer.
2. **An array**.
3. A third element that **handles chaos** when it occurs.

A Perfect Hash Function



...characteristics of this function?

A Second Hash Function



0	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

...characteristics of this function?

All hash functions will consist of two parts:

- A **hash**:
- A **compression**:

Characteristics of a good hash function:

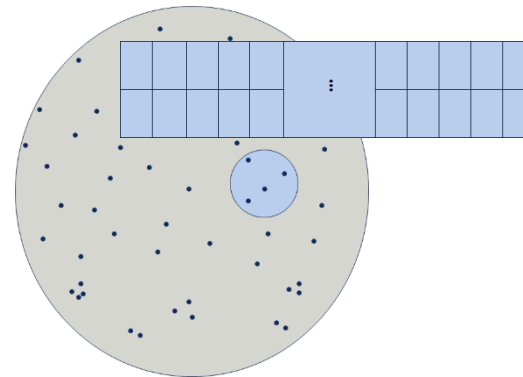
1. Computation Time:
2. Deterministic:
3. SUHA:

Towards a general-purpose hashing function:

It is easy to create a general-purpose hashing function when the keyspace is proportional to the table size:

- **Ex:** Professors at CS@Illinois
- **Ex:** Anything you can reason about every possible value

It is difficult to create a general-purpose hashing function when the keyspace is large:



My 40-character strategy:

Alice in Wonderland, Page 1	
1	Alice was beginning to get very tired of
2	sitting by her sister on the bank, and
3	of having nothing to do: once or twice s
4	he had peeped into the book her sister w
5	as reading, but it had no pictures or co
6	nversations in it, 'and what is the use
7	of a book,' thought Alice 'without pictu
8	res or conversations?' So she was consi
9	dering in her own mind (as well as she c
10	ould, for the hot day made her feel very
11	sleepy and stupid), whether the pleasur
12	e of making a daisy-chain would be worth
13	the trouble of getting up and picking t
14	he daisies, when suddenly a White Rabbit
15	with pink eyes ran close by her. There
16	was nothing so very remarkable in that;
17	nor did Alice think it so very much out
18	of the way to hear the Rabbit say to it
19	self, 'Oh dear! Oh dear! I shall be late
20	!' (when she thought it over afterwards,
21	it occurred to her that she ought to ha

...what is a naïve hashing strategy for this input?

...characteristics of this function?

What is an example of bad input data on this hash function?

Reflections on Hashing

We are starting the study of general-purpose hash functions. There are many other types of hashes for specific uses (ex: cryptographic hash functions).

Even if we build a good hash function, it is not perfect. What happens when the function isn't always a bijection?

Collision Handling Strategy #1: Separate Chaining

Example: $S = \{ 16, 8, 4, 13, 29, 11, 22 \}$, $|S| = n$
 $h(k) = k \% 7$, $|Array| = N$

[0]	
[1]	
[2]	
[3]	
[4]	
[5]	
[6]	
[7]	

Load Factor:

Running time of Separate Chaining:

	Worst Case	SUHA
Insert		
Remove/Find		

CS 225 – Things To Be Doing:

1. Programming Exam B starts tomorrow
2. MP5 has been released; EC⁷ deadline is Monday back from break
3. lab_btrees released today
4. Daily POTDs are ongoing!