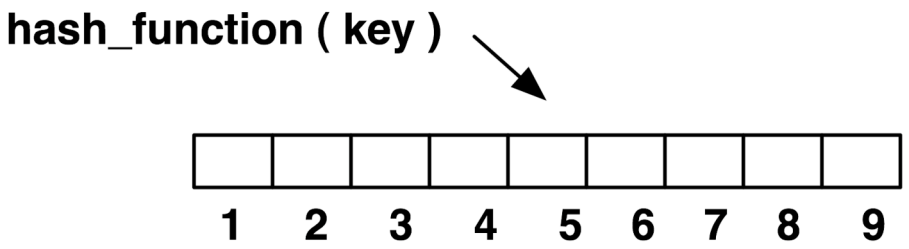
Slides Credits:
Gunnar Gotshalks, Michael Langer, Mark Redekopp, David Kempe, Sandra Batista

Definition

- A hash table has the following components
  - **An array – called a table – of size N**
  - **A mathematical function – called a hash function – that maps keys to valid array indices**

    hash_function: key → 0 .. N – 1

- Table entries are stored and retrieved by applying the hash function to the entry key to get the index used to probe the table.



Hash Function Basic Properties

- A hash function consists of two parts
  - » **Hash code function**

    $$\text{hash\_code}(\text{key}) \rightarrow \text{integer}$$

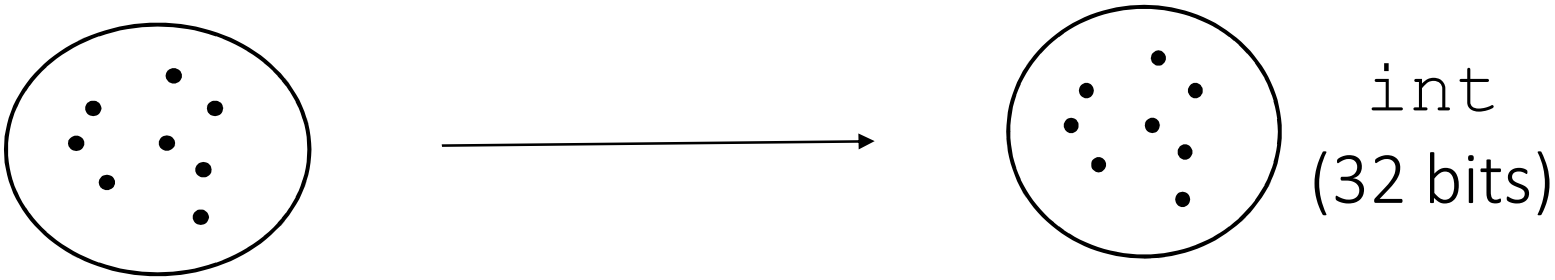
  - » **Compression function**
    - > **Maps the integer from the hash code function to the integer interval [0 , N – 1]**

      $$\text{compress}(\text{integer}) \rightarrow 0 .. N - 1$$

    - > **Program function composition**

      $$\text{compress}(\text{hash\_code}(\text{key})) \rightarrow 0 .. N - 1$$

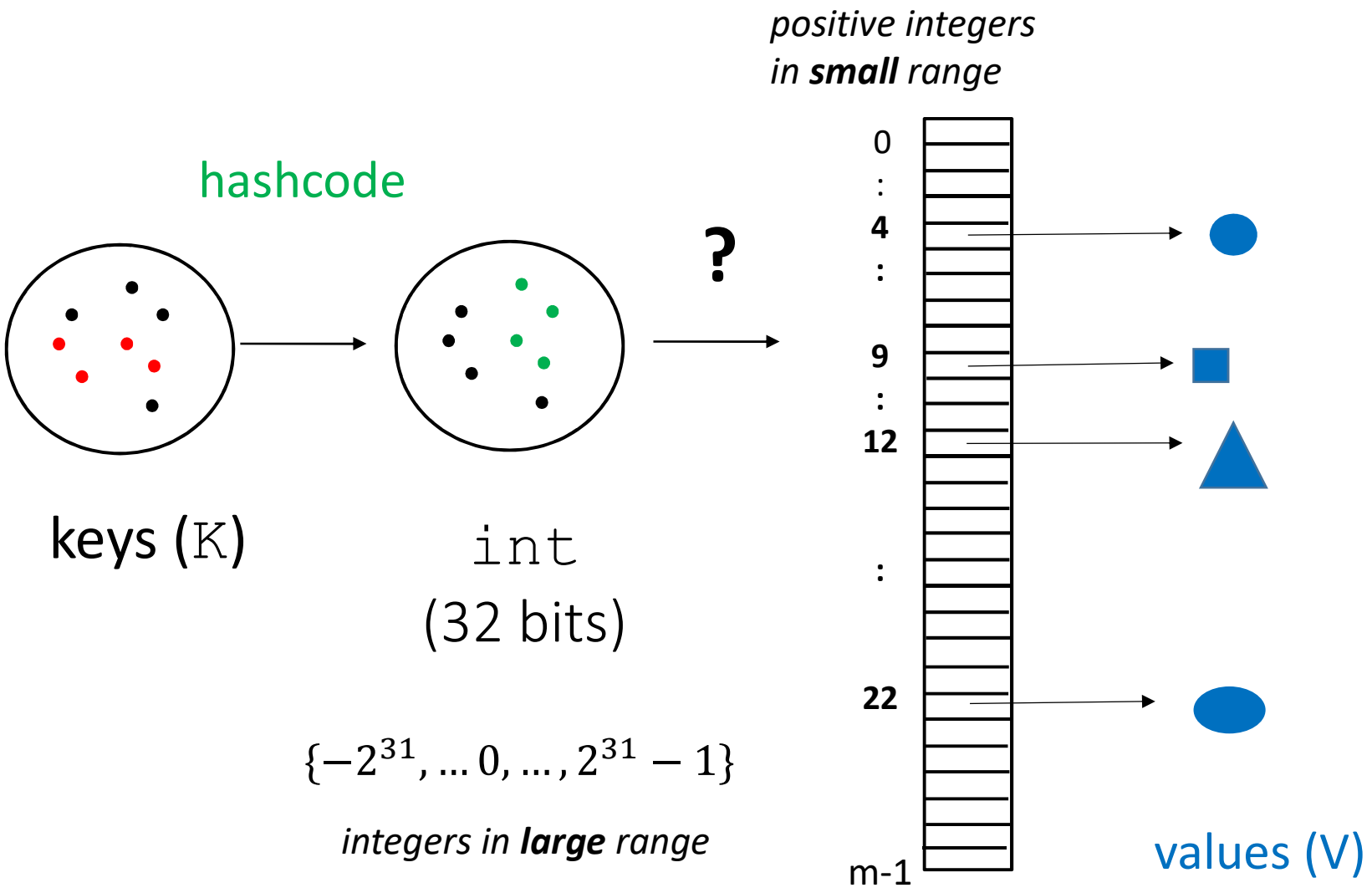
Hash Code



Example: Strings

- Use ASCII codes for each character and add them or group them
- "hello" => 'h' = 104, 'e'=101, 'l' = 108, 'l' = 108, 'o' = 111 = 532
- Hash function is then applied to the integer value 532 such that it maps to a value between 0 to M-1 where M is the table size

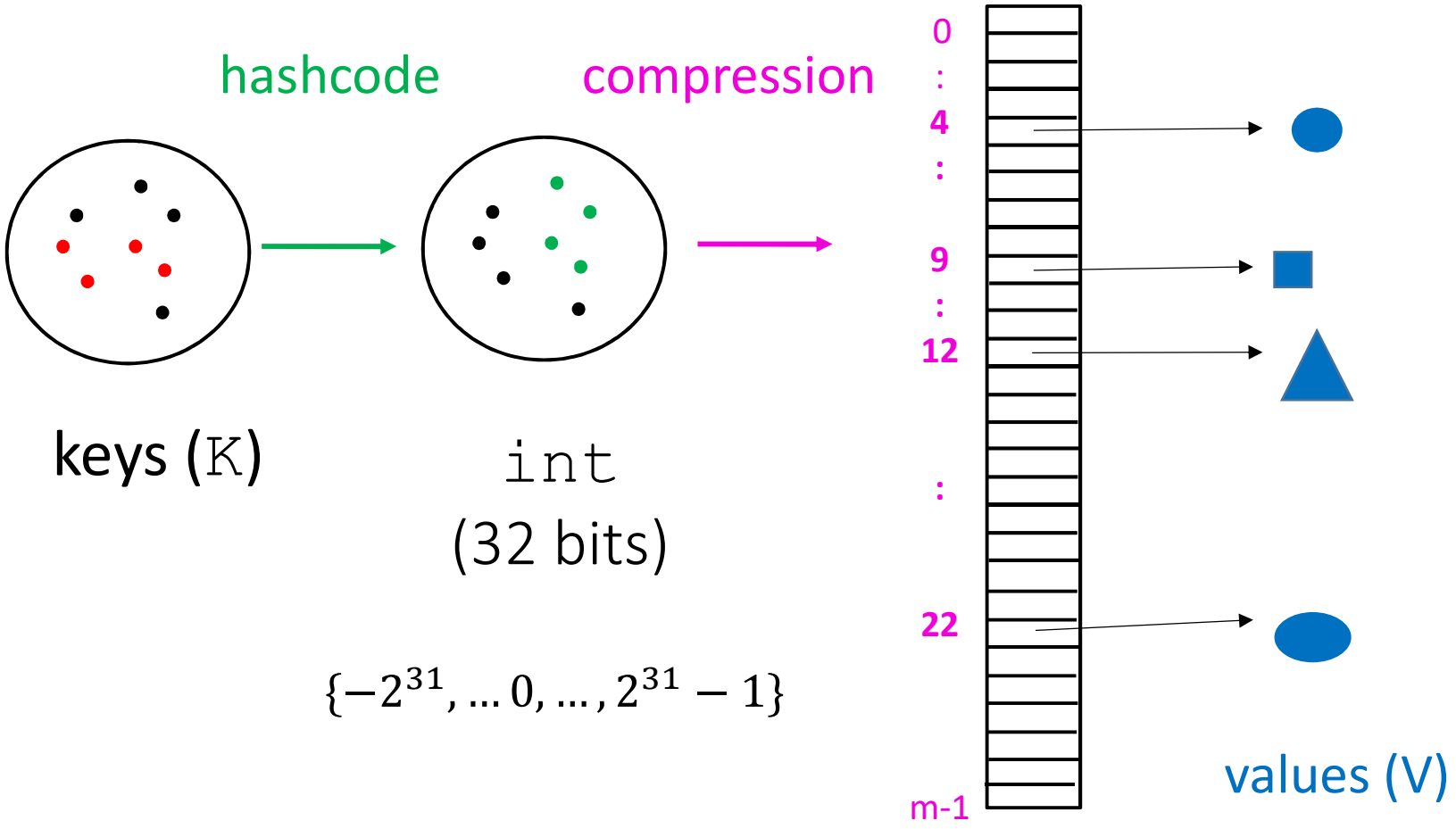
We will try to combine these two ideas as shown below....



... using a many-to-one "compression" map.

compression : $i \rightarrow |i| \ mod \ m,$

where m is the length of the array.



Universal Hash Example

- Suppose we want a universal hash for words in English language
- First, we select a prime table size, **m**
- For any word, w made of the sequence of letters w_1 w_2 … w_n we translate each letter into its position in the alphabet (0-25).
- Consider the length of the longest word in the English alphabet has length z
- Choose a random key word, K, of length z, K = k_1 k_2 … k_z
- The random key a is created once when the hash table is created and kept

- Hash function: $h(w) = \left(\sum_{i=1}^{len(w)} k_i \cdot w_i\right) mod\ \boldsymbol{m}$