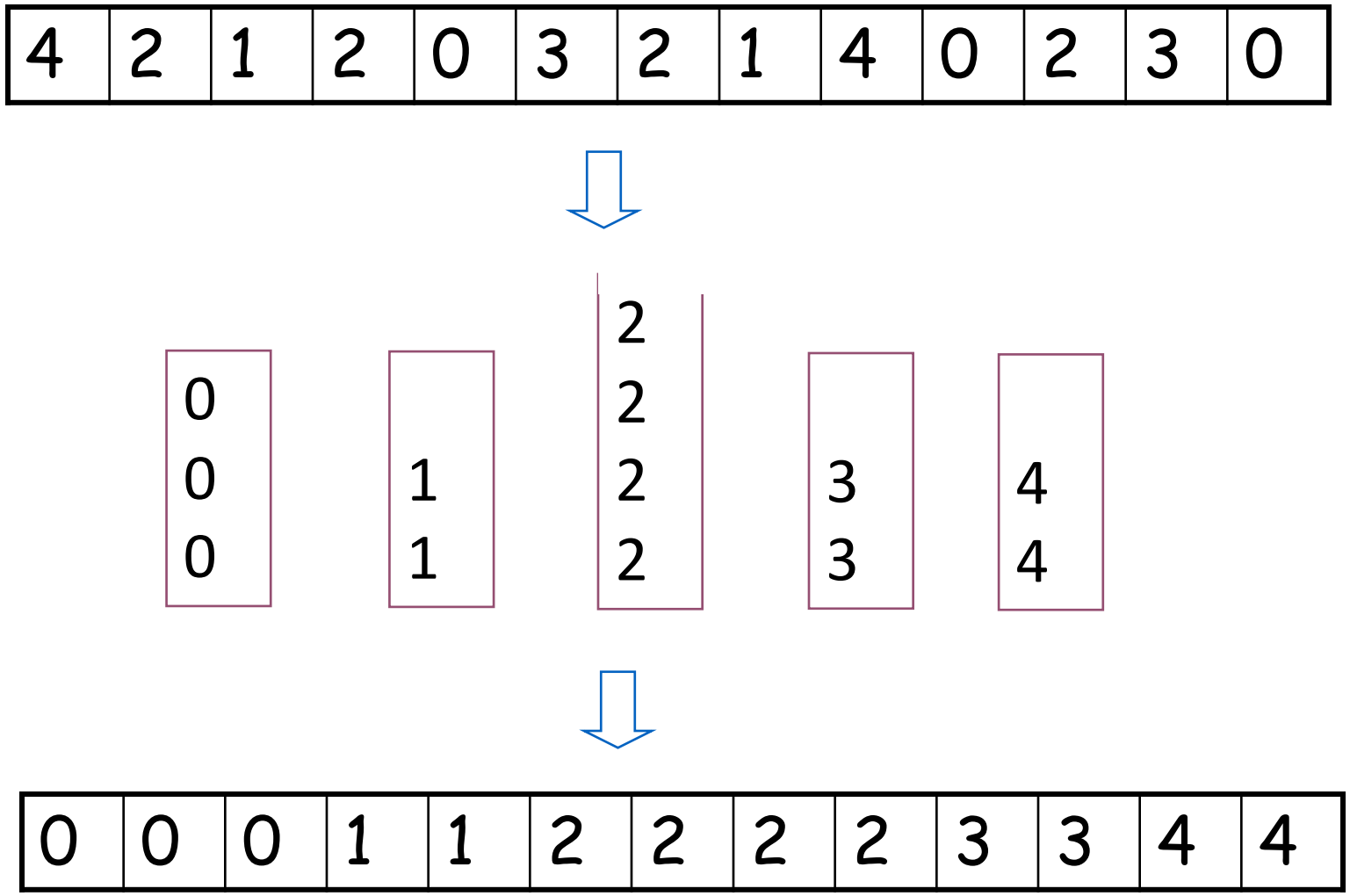
Bucket & Radix Sort

IIITS

Bucket Sort

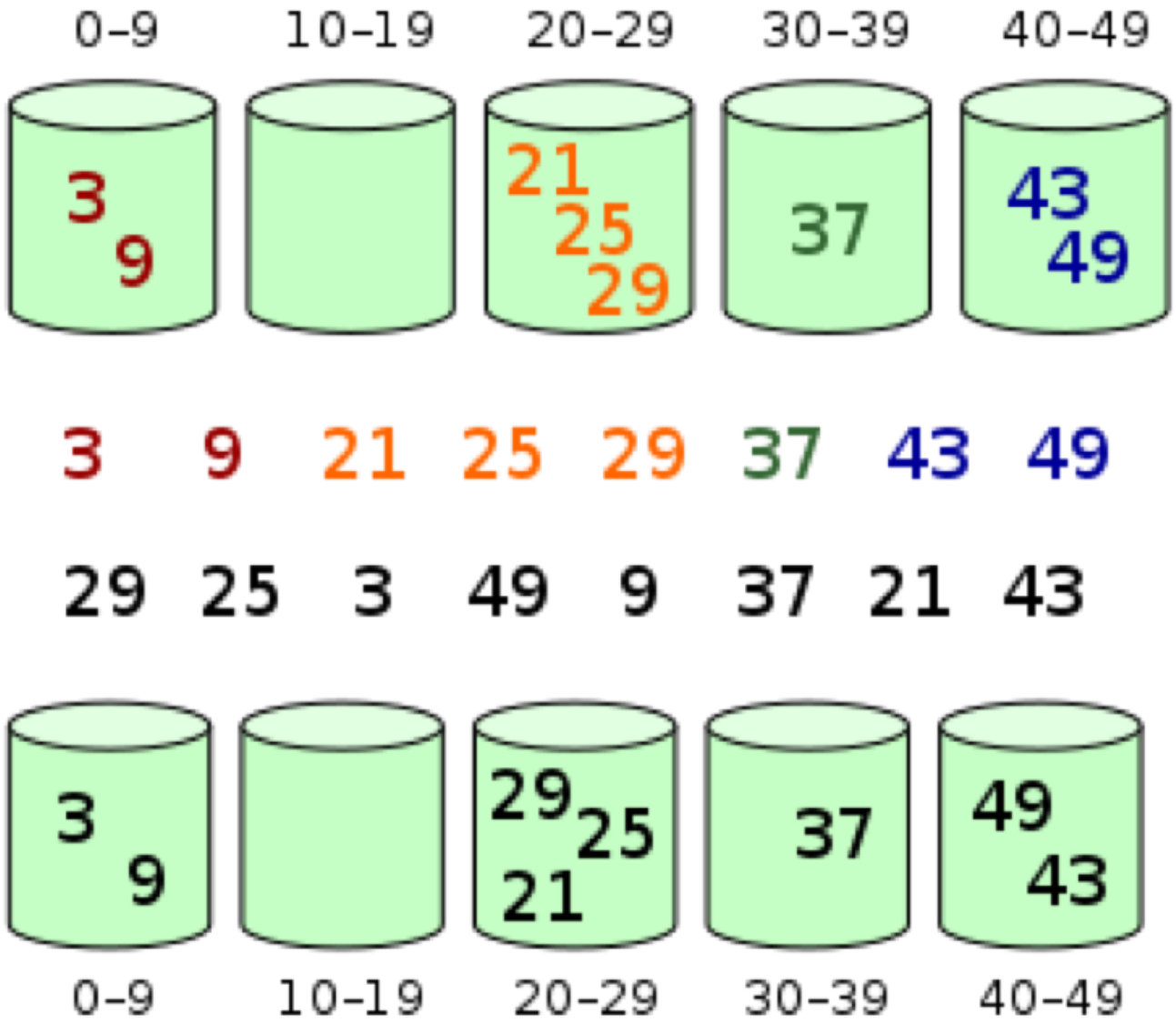
- Bucket sort
  - Assumptions: the keys are in the range [0, N], and there are repetitions.
  - Basic idea:
    1. Create *N* linked lists (*buckets*) to divide interval [0,N] into subintervals of size 1
    2. Add each input element to appropriate bucket
    3. Concatenate the buckets
  - Expected total time is O(n + N), with n = size of original sequence
    - if N is O(n) → sorting algorithm in O(n) !

Example



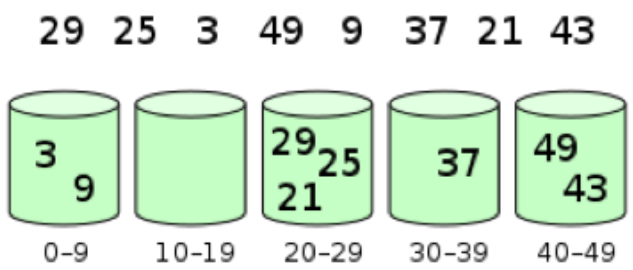
```
void bucketSort(int a[], int n) {
    int i, j, k, buckets[SIZE];

    for(i = 0; i < SIZE; ++i)
        buckets[i] = 0;

    for(i = 0; i < n; ++i)
        ++buckets[a[i]];

    for(i = 0, j = 0; j < SIZE; ++j)
        for(k = buckets[j]; k > 0; --k)
            a[i++] = j;
}
```

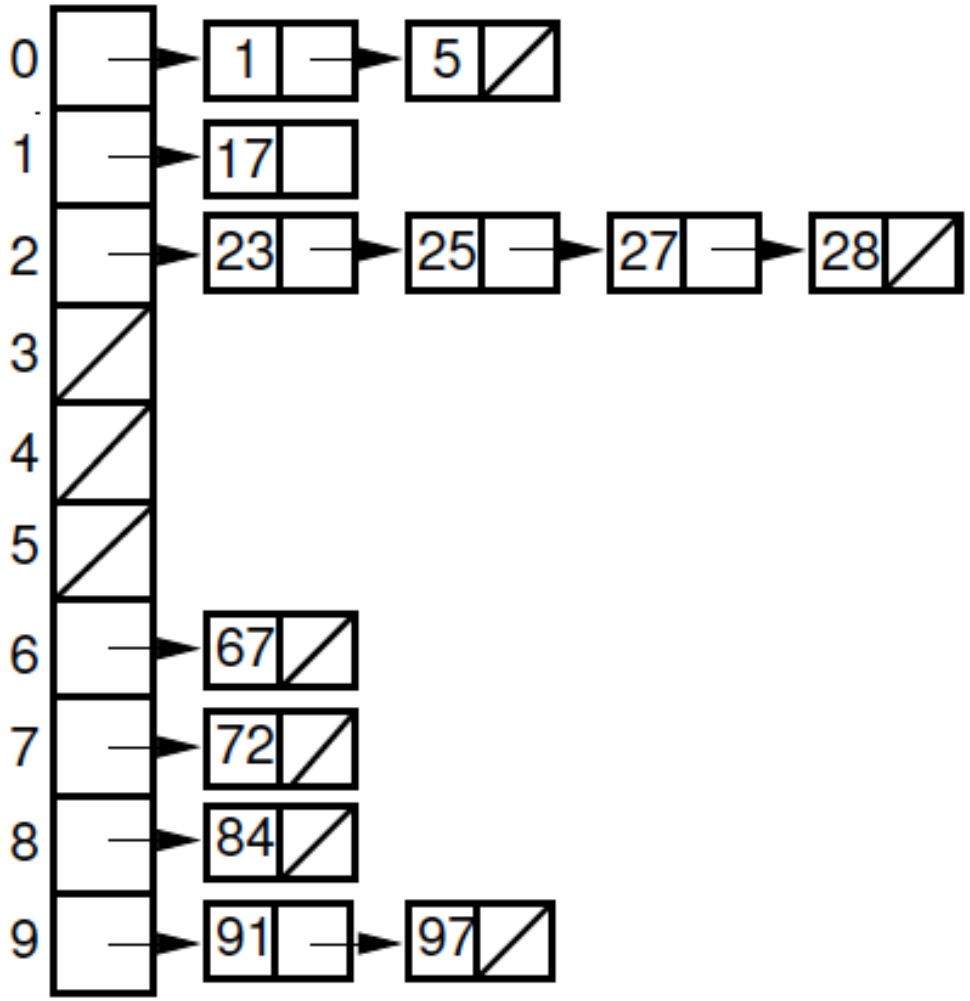
Bucket– Linked List



Bucket– Linked List Assignment

27	91	1	97	17	23	84	28	72	5	67	25





Does it Work for Real Numbers? Assignment

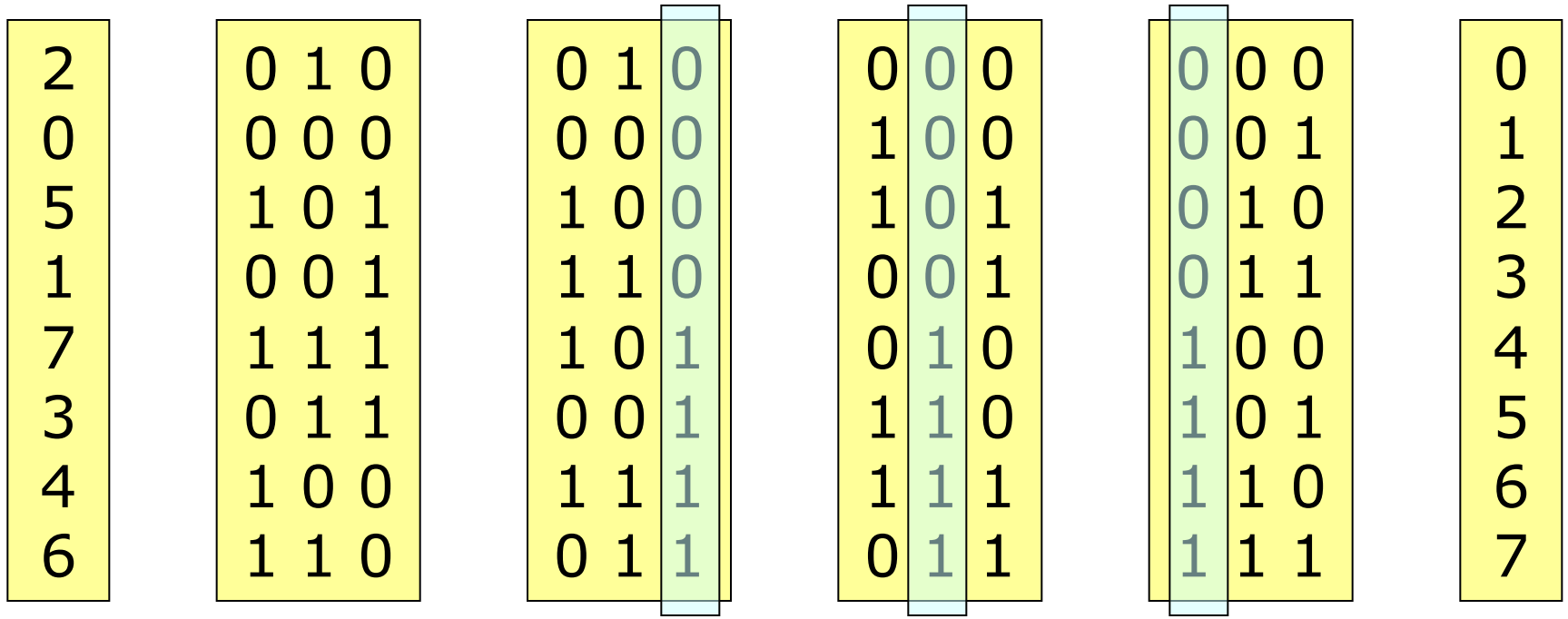
- What if keys are not integers?
  - Assumption: input is *n* reals from [0, 1]
  - Basic idea:
    - Create *k* linked lists (*buckets*) to divide interval [0,1] into subintervals of size n/k
    - Add each input element to appropriate bucket and sort buckets with insertion sort

RadixSort

- Radix = “The base of a number system” (Webster’s dictionary)
- History: used in 1890 U.S. census
- Idea: Bucket Sort on each digit, bottom up.

Radix sort

- Example:

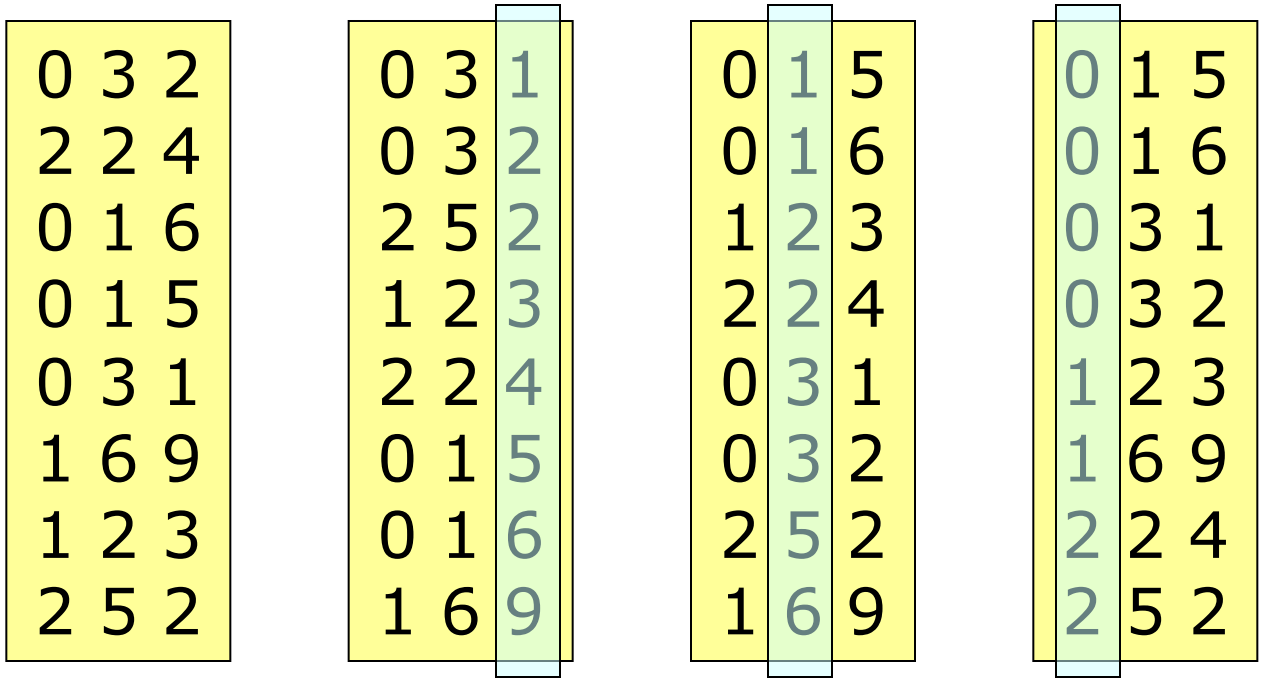


Radix sort characteristics

- Each sorting step can be performed via bucket sort, and is thus $O(N)$.
- If the numbers are all b bits long, then there are b sorting steps.
- Hence, radix sort is $O(bN)$.

What about non-binary?

- Radix sort can be used for decimal numbers and alphanumeric strings.



RadixSorting Strings

- Break strings into characters.
- Need to know length of biggest string (or calculate this on the fly).
- The size of the data structure would be as the longest string.

RadixSorting Strings example

	5th pass	4th pass	3rd pass	2nd pass	1st pass
String 1	z	i	p	p	y
String 2	z	a	p		
String 3	a	n	t	s	
String 4	f	l	a	p	s

NULLs are just like fake characters

Radix Sort - Assignments

- For string
- For floating point numbers