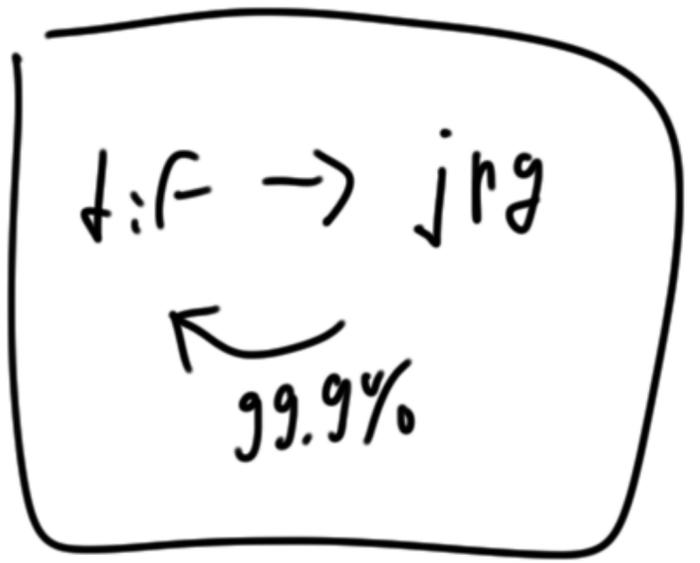




# BloomFilter: Core Idea

Lossless: RLE, WAH, decomposed bitmaps

Lossy: approximate bitmaps, bloom filter

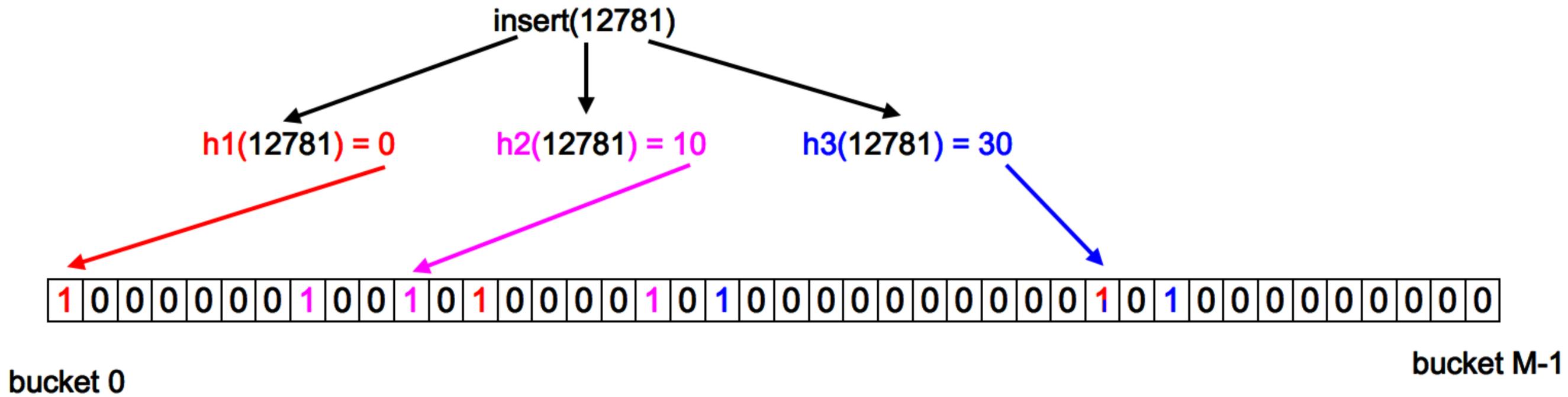


Bloom filter  $\approx$  hashing + bit lists





# setBit(12781), which color is bit 30?

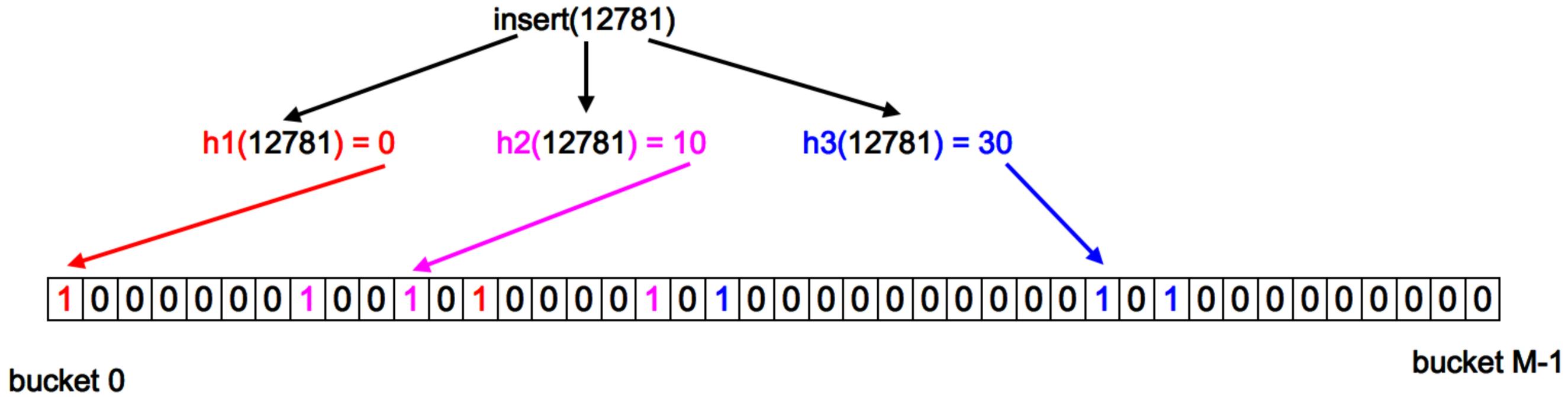


# of inserts so far,  
N=3

k hash functions,  
here k=3

bit list of size M,  
here M=42

# setBit(12781), it is yellow!



# of inserts so far,  
N=3

k hash functions,  
here k=3

bit list of size M,  
here M=42

# BloomFilter Lookup Semantics

for one hash-function  $h_{\langle i \rangle}(\text{Key}) == 0$ :

key **not** in bitmap

guaranteed!

# BloomFilter Lookup Semantics

① for one hash-function  $h_{\langle i \rangle}(\text{Key}) == 0$ :

key **not** in bitmap

② for all hash-function  $h_{\langle i \rangle}(\text{Key}) == 1$ :

key **maybe** in bitmap

2a

key is in bitlist

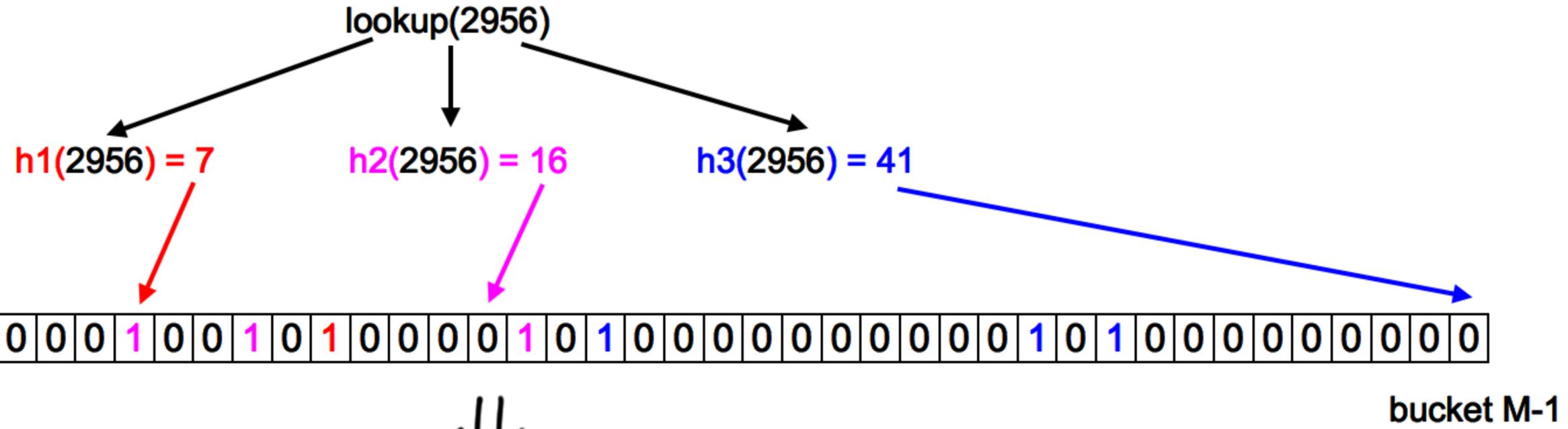
2b

key is not in bitlist

= false positive



# lookup(2956)



# of inserts so far,  
N=3

k hash functions,  
here k=3

bit list of size M,  
here M=42

⇓  
2956 not in bloom filter

