## Hashing

## TOTAL POINTS 3

1. What is the minimum size of an array that can be used in the direct addressing scheme to store a map from 7-digit phone 1/1 point

10000000

20000000

0 1000000



Correct! 7-digit phone numbers correspond to integers from 0 to 9999999.

2. If it is guaranteed that the total length of all occurrences of a Pattern in a Text is at most L, which of the below estimates of the average running time of Rabin-Karp's algorithm to find all occurrences of the Pattern in the Text is the most tight out of the correct ones?

1/1 point

 $\bigcirc \ O\big(|Text||Pattern| + L\big)$ 

 $\bigcirc O(|Text||Pattern|L)$ 

 $\bigcirc \ O\big(|Text| + |Pattern| + L\big)$ 

 $\bigcirc O(|Text| + |Pattern|)$ 



Correct! Estimate from the lecture is O(|Text| + (q+1)|Pattern|), where q is the number of occurrences of the Pattern in the Text, and  $L=q\vert Pattern\vert$  in this case.

3. Let us slightly change the polynomial hash function for strings and set  $h(S) = \binom{|S|-1}{j=0} x^{|S|-1-j} S[j]) \mod p$ . Let us fix some Text and some Pattern. Denote by H[i] the hash function of the substrain Text[i..i+|Pattern|-1] of the Text starting from position i and having the same length as Pattern (for all appropriate positions i where the Pattern can occur in the Text). Which of the below formulas is the correct recurrence to compute H[i+1] given H[i]?

0 / 1 point

 $\bigcirc \ \ H[i+1] = \left(xH[i] + Text[i+|Pattern|-1] - x^{|Pattern|}Text[i]\right) \bmod p$ 

 $\textcircled{\textbf{0}} \ \ H[i+1] = \left(xH[i] + x^{|Pattern|}Text[i+|Pattern|] - Text[i]\right) \bmod p$ 

 $\bigcirc \ \ H[i] = \left(xH[i+1] + Text[i] - x^{|Pattern|} Text[i+|Pattern|]\right) \bmod p$ 

 $\bigcirc \ \ H[i+1] = \left(xH[i] + Text[i+|Pattern|] - x^{|Pattern|}Text[i]\right) \bmod p$ 

X Incorrect

Note that the first letter of the substring should be multiplied by the largest power of x, and the last letter of the substring should be multiplied by  $x^0=1$  while computing the hash function for it.