Dennis Kuzminer
CSCI-UA 310-001 PS8

4.

a. If two strings will have the same hash this means that they are (most likely) the same. The algorithm creates a "window" for each substring of length t in a and b and hashes the substrings. Given the same seed and hash function, if some $r_i == s_i$ then the hash will be the same.

b. As Karp/Rubin uses ε-universal hashing, meaning ε=(t-1)/m
$|r_i|, |s_i| = n - t + 1$
Therefore, the maximum $Pr[h_\lambda(a) = h_\lambda(b),\ a \neq b] \leq (n - t + 1)(n - t + 1)(\frac{t-1}{m}) \rightarrow (n - t + 1)^2(\frac{t-1}{m})$

c. If we store the value of all r_i in a hashmap/hashtable, we can get an expected constant lookup time. So the only work that would need to be done is to assign the key-value pairs/create a hashtable for all r_i. This means we can simply loop through s_i and lookup each value in the hashtable with expected $O(1)$ time.