Tecnológico de Monterrey

Integrated Activity 1

Integrantes:

Gerardo Ulises Sánchez Félix - A01641788

Alan Antonio Ruelas Robles - A01641426

Analysis and Design of Advanced Algorithms

(Gpo 603)

09 de septiembre del 2023

For this situation problem we chose three different algorithms in order to solve the 3 requested tasks corresponding to analyzing 2 transmission codes and 3 malicious codes and the relations between them.

The first problem required us to find if there were instances of the contents of the malicious codes in the transmission codes, and if so, where do they start within them. In order to do this, we used the KMP's algorithm because of its efficiency complexity-wise. It first preprocesses an array that stores the longest prefix substring in the pattern (malicious code) so that the algorithm can skip unnecessary comparisons when a mismatched has occurred, which translates in less computations compared to the naive algorithm. It has a time complexity of O(n + m); O(m) of the preprocessing plus the O(n) to iterate over the text that is being analyzed ,which is better than the O(nm) that the brute force algorithm has.

The second problem required us to find the longest mirrored (palindrome) code there is in the transmission codes. To solve this, we use the Manachor's algorithm because of its smart approach on using palindromes properties. It stores the expansion length of the palindrome centered at each index into an array. Because it uses the symmetry of a palindrome, it does not need to manually go to each index and expand to check the expansion length every time, so it is much more efficient than the naive implementation which takes O(n^3) time versus the O(n) of Manachor's.

We decided to implement the Longest Common Subsequence (LCS) algorithm to solve the third problem because we needed to display the longest common substring between both stream files, so the obvious implementation is that algorithm, the time complexity is O(m * n) and the space complexity: O(L). The letter L represents the length of the Longest Common Subsequence (LCS) between the two input strings X and Y. We decided to use a hashmap because the normal implementation took up a lot of memory space as it stored all 0 values in the matrix. This implementation saves a lot of memory since it only stores the necessary values.

Reflections

Alan Antonio Ruelas Robles - A01641426
With this activity, I learned some new algorithms related to strings and its uses in some real life problems. These approaches have given me some new and smart insights on more efficient implementations compared to the naive ones that are more easily grasped on anyone's mind.

With the KMP, Manachor and LCS algorithms, I was able to think out of the box and try to implement them based on the pseudocodes and explanations of them and I was amazed at how intelligent were the people that created them. That inspires me to make better and more readable code into the future with the upcoming activities and hopefully be useful in the next problems that arise as the technology evolves.

Gerardo Ulises Sanchez Felix - A01641788
Through this activity, I learned how to implement various algorithms to find if different smaller strings exist in a larger string. This has many applications in real life, for example when analyzing DNA chains to find different diseases or characteristics of a person.

These are some useful implementation in real life:

- Network Routing: KMP can be applied to find the longest prefix match in routing tables, helping determine the next hop for a given IP address.
- Data Compression: KMP can be used in various compression algorithms to find repeated patterns within the data.
- Palindromic Substring Detection: Manacher's algorithm has applications in DNA sequence analysis, and speech recognition.
- Image Processing: Manacher's algorithm can be used in image processing for detecting circular patterns, which can be useful in object recognition.
- Bioinformatics: LCS can be applied to compare DNA sequences or protein sequences to identify common subsequences, which can have important biological implications.

These are some useful applications of the content you saw in class.