

Reporte: Empatamiento de Cadenas

Oswaldo de Luna

14 de noviembre de 2019

Problema

Una cadena es una secuencia de caracteres sobre un alfabeto finito. Por ejemplo $SASASASADEDEDESA \in \{A, D, E, S\}$.

El problema de emparejamiento de cadenas es encontrar todas las ocurrencias de una cadena, llamada patrón, en una cadena más grande del mismo alfabeto. Esto implica la implementación de algoritmos de búsqueda de subcadena para buscar la existencia de una subcadena dentro de una cadena.

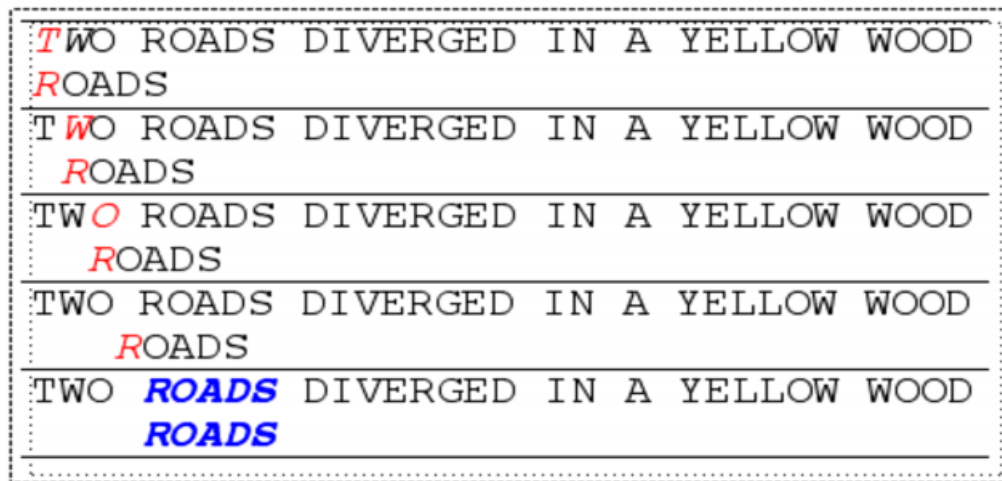


Figura 1: Ejemplo de Empatamiento de Cadenas

Solución propuesta

Para fuerza bruta, se recorre la cadena principal hasta encontrar que la letra en la que nos encontramos sea la misma con la que inicia la subcadena, una vez encontrada se hace una segunda comparación recorriendo la

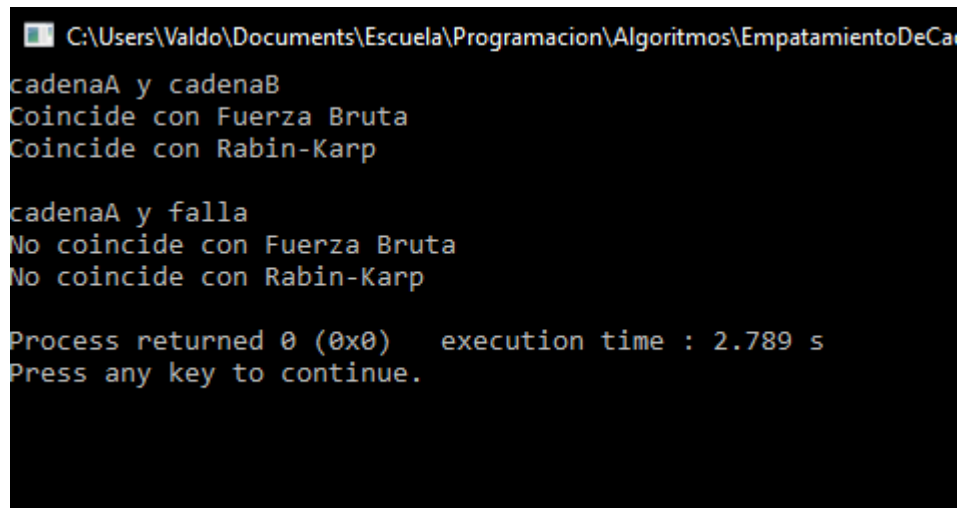
subcadena y contabilizando el número de coincidencias que se hayan encontrado, si el número de coincidencias es igual al número de caracteres de la subcadena, entonces la subcadena se encuentra en la cadena y sino se sigue con el recorrido de la cadena principal.

Para Rabin Karp, se calcula un valor *hash* para los n caracteres al comienzo de la cadena y para la subcadena a encontrar. Este valor se calcula con la siguiente fórmula: $h = \sum_0^n x_n p^n$, donde n = Tamaño de la subcadena, x_n = Valor numérico del carácter según su posición n y p = Un número primo cualquiera. Este hash no será modificado para la subcadena, mientras que para la cadena se irá recalculando, ya que se estará comparando los n caracteres siguientes (se restaría el primero y se agregaría el siguiente). Para dicha operación no es necesario repetir la fórmula, se propone utilizar la siguiente:

$$h_r = \frac{(h - x_0)}{p} + x_n p^{n-1}$$
, donde h_r = Hash recalculado, h = Hash anterior de la cadena principal, x_0 = Primer carácter de la subcadena de la cadena a la que se le calculó el hash y n y p son los mismos que la fórmula anterior. Cuando hay una coincidencia de hashes se comparan ambas subcadenas, de ser las mismas hay un empatamiento de cadena.

Implementación

Con CadenaA = "Mund Mundo!"
CadenaB = "Mundo"
falla = "Mund Mun."



```
C:\Users\Valdo\Documents\Escuela\Programacion\Algoritmos\EmpatamientoDeCa
cadenaA y cadenaB
Coincide con Fuerza Bruta
Coincide con Rabin-Karp

cadenaA y falla
No coincide con Fuerza Bruta
No coincide con Rabin-Karp

Process returned 0 (0x0)   execution time : 2.789 s
Press any key to continue.
```

Figura 2: Ejemplo de ambos métodos implementados