UNIVERSIDAD DE CONCEPCIÓN
Facultad de Ingeniería
Departamento de Ingeniería Informática y
Ciencias de la Computación

Profesor Patrocinante:
Dr. Leo Ferres

Búsqueda exacta en secuencias de caracteres UTF-8

Algoritmos paralelos versus algoritmos secuenciales palelizados mediante descomposición de dominio.

Rodrigo Antonio Reyes Cabrera

Tesis de Magister
para optar al título de
Magister en Ciencias de la Computación

16 de Diciembre del 2011

Abstract

This work focuses on comparing the efficiency of exact string matching algorithms in UTF-8 encoded strings. Specifically, we will compare the efficiency of sequential algorithms which have been later parallelized using domain decomposition techniques against the efficiency of originally parallelized algorithms. To this end, we will use the UTF-8 international string encoding standard which provides a big enough encoding space for all present known languages.

We prove that sequential algorithms with non original parallelization are more efficient than originally parallelized algorithms. For this purpose, we implement the Naive, Knuth-Morris-Pratt, Boyer-Moore and Bitap algorithms, both in their sequential and domain decomposition based parallelized versions. Then we compare them with the originally parallelized algorithm from Vishkin. As our final result, we found out that parallelized sequential algorithms are more efficient than originally parallelized algorithms.