



Universidad de Concepción  
Dirección de Postgrado  
Facultad de Ingeniería -Programa de Magíster en Ciencias de la Computación

**Un algoritmo de descomposición de tareas para el  
reconocimiento de patrones de texto  
(A task decomposition algorithm for string matching)**

Tesis para optar al grado de Magíster en Ciencias de la Computación

ERICK ELEJALDE SIERRA  
CONCEPCIÓN-CHILE  
2014

Profesor Guía: Dr. Leo Ferres  
Dpto. de Ingeniería Informática y Ciencias de la Computación, Facultad de Ingeniería  
Universidad de Concepción

## Abstract

Previous attempts to parallelize string matching algorithms have focused on data decomposition; that is, splitting the text on the number of available cores. In this work, we demonstrate that it is possible to use task decomposition efficiently, using a multi-processor model. We introduce the Task Decomposition String Matching algorithm, a new algorithm based on bit parallelism and finite automata theory. We also present a producer/consumer strategy which may be applied to the traditional algorithms as an alternative way to parallelize them. We run experiments that compare both our approaches against traditional algorithms (such as Boyer-Moore, KMP and Bitap) by varying the length of the pattern and the properties of alphabets, two critical variables that affect performance in string matching. The Domain Decomposition strategy proved to be very efficient for this problem. Although task decomposition is asymptotically optimal, it is practically affected by cache contingencies.

