



Yihan Sun

Parallel Balanced Binary Trees Using Just Join

Friday, December 1, 2017 – 4:00 p.m. – GHC 4405

Tree data structures play an important role in almost every area of computer science. Nowadays the explosion of data puts forward a higher demand for the performance of the trees, which requires efficient parallel algorithms for processing large-scale data, as well as a full-featured interface for achieving more functionalities. Traditional algorithm designing on balanced binary trees based on insertions and deletions are plagued with challenges such as being hard to parallelize, single-operand oriented, and varying among different balancing schemes. This thesis proposes a parallel algorithmic framework that overcomes these challenges, which captures all balancing criteria in a single function JOIN. The JOIN-based algorithms are then used for supporting sequences, ordered sets, ordered maps and augmented maps (formally defined in this thesis). A wide variety of functions form a hierarchical interface for these four data types, which are all implemented by join-based algorithms as part of a library PAM (Parallel Augmented Maps). This library is parallel, work-efficient, generic across balancing schemes, persistent, safe for concurrency and applicable to a wide range of applications and queries by proper augmentations. The augmented map structure itself is an abstraction befitting many real applications. The PAM interface greatly simplifies the programming of many applications over direct implementations, such as interval trees, range trees, inverted indices, segment trees, and database snapshot isolations. Experiments show that the implementations of the functions in PAM, as well as all the applications using PAM are practically efficient. Sequentially the code achieves performance that matches or exceeds existing libraries designed specially for a single application, and the parallel implementation gets speedups ranging from 40 to 90 on 72 cores with 2-way hyperthreading.

Thesis Committee:

Guy E. Blelloch, Chair

Andrew Pavlo

Daniel D.K. Sleator

Michael T. Goodrich, University of California, Irvine