

COMPUTATIONAL RESEARCH in BOSTON and BEYOND SEMINAR

On Optimal Partitioning for Variable Block Row Format

PETER JAMES AHRENS

Massachusetts Institute of Technology

ABSTRACT:

The Variable Block Row (VBR) format is an influential blocked sparse matrix format designed for matrices with shared sparsity structure between adjacent rows and columns. VBR groups adjacent rows and columns, storing the resulting blocks that contain nonzeros in a dense format. This reduces the memory footprint and enables optimizations such as register blocking and instruction-level parallelism. Existing approaches use heuristics to determine which rows and columns should be grouped together. We show that finding the optimal grouping of rows and columns for VBR is NP-hard under several reasonable cost models. In light of this finding, we propose a 1-dimensional variant of VBR, called 1D-VBR, which achieves better performance than VBR by only grouping rows. We describe detailed cost models for runtime and memory consumption. Then, we describe a linear time dynamic programming solution for optimally grouping the rows for 1D-VBR format. We extend our algorithm to produce a heuristic VBR partitioner which alternates between optimally partitioning rows and columns, assuming the columns or rows to be fixed, respectively. Our alternating heuristic produces VBR matrices with the smallest memory footprint of any partitioner we tested.

BIO:

Peter Ahrens is a DOE CSGF (Computational Science Graduate Fellow) at MIT working with Professor Saman Amarasinghe. Peter focuses on problems of efficiently allocating machine resources to sparse computations.

FRIDAY, MAY 7, 2021

12:00 PM – 1:00 PM

<http://math.mit.edu/crib/>

ZOOM MEETING info:

<https://mit.zoom.us/j/96155042770>

Meeting ID: 961 5504 2770