

Abstract

Cache-conscious Hierarchical Clustering with a Lisp Garbage Collector

MASAHIRO YASUGI^{1,a)} SHINGO YUASA^{2,†1}

Presented: June 19, 2014

With increasing processor-memory performance gap, improving cache locality is as important as improving virtual memory locality. In many applications, especially in search algorithms on large pointer-based data structures, breadth-first copying algorithms increase cache misses, TLB misses, and page faults. To improve locality at both the cache and virtual memory levels of the memory hierarchy, “hierarchical clustering,” which groups data objects at multiple hierarchical levels, was proposed. In this study, we implemented hierarchical clustering in a commercial Common Lisp system; we considered various implementation issues, particularly in generational GC. Our garbage collector automatically improves data locality at multiple levels; it also allows us to employ a simple tuning method for further improvement. Evaluations with two microbenchmark programs, an XML application, and a tree-based ray-tracing application show that hierarchical clustering provides good overall performance; more precisely, it serves as an insurance that covers miss-induced performance degradation.

¹ Department of Artificial Intelligence, Kyushu Institute of Technology, Iizuka, Fukuoka 820–8502, Japan

² Graduate School of Informatics, Kyoto University, Sakyo, Kyoto 606–8501, Japan

^{†1} Presently with Google Inc.

^{a)} yasugi@ai.kyutech.ac.jp