COMPUTER SCIENCE COLLOQUIUM

Java 7's Dual Pivot Quicksort — Analysis and Engineering

Markus Nebel

Department of Computer Science
University of Kaiserslautern, Germany

Department of Mathematics and Computer Science
University of Southern Denmark

Tuesday, 05 March, 2013 at 14:15

IMADA's Seminar Room

Abstract:

Recently, a new Quicksort variant due to Yaroslavskiy was chosen as standard sorting method for Oracle's Java 7 runtime library. The decision for the change was based on empirical studies showing that on average, the new algorithm is faster than the formerly used classic Quicksort. Surprisingly, the improvement was achieved by using a dual pivot approach, an idea that was considered not promising by several theoretical studies in the past. In this talk, we identify the reason for this unexpected success. Moreover, we present the first precise average case analysis of the new algorithm showing e.g. that a random permutation of length n is sorted using $1.9n \ln n - 2.46n + O(\ln n)$ key comparisons and $0.6n \ln n + 0.08n + O(\ln n)$ swaps.

Our analysis reveals a highly asymmetric nature of the algorithm. This suggest that asymmetric pivot choices are preferable to symmetric ones. From a theoretical point of view, this should allow us to improve on the current implementation in Oracle's Java 7 runtime library. We present results derived by means of our tool MaLiJAn to confirm this for asymptotic combinatorial cost measures such as the total number of executed instructions (Java bytecodes). However, the observed running times show converse behavior. With the support of data provided by MaLiJAn we are able to identify the profiling capabilities of Oracle's just-in-time compiler to be responsible for this unexpected outcome.