

Simulation Assignment III

240C Winter 2022

February 14, 2022

1 Adaptive Computing

In this assignment, you will explore methods to increase adaptivity of cache management through replacement policy and prefetching.

1.1 Survey

What are the four ways to improve the adaptability of cache management? Start with "Adaptive Insertion Policies for High Performance Caching by Qureshi et al". Conduct a short half-page survey.

1.2 Adaptive Replacement Policy

Merge the capabilities of the two replacement policies you studied in assignment II using *set dueling* technique explained in paper "Adaptive Insertion Policies for High Performance Caching by Qureshi et al". Your method should enable the last level cache to choose between two policies, depending on which policy incurs fewer misses for the current workload.

1.3 Adaptive Instruction Prefetching

Repeat the experiment from section 1.2 for two instruction prefetchers you studied in assignment II. Your method should enable the I-cache to choose between two prefetchers, depending on which has more benefit for the current workload.

1.4 Improve Speedup per Area Overhead

Perform sensitivity studies for section 1.2 and 1.3. Analyze your observation and explain the trends. Based on your observations, apply intelligent design decisions to modify the adaptive policy, and show that your dynamic design (come up with a name for it) can perform superior to a single replacement policy/prefetcher with equal hardware budget. Increase Speedup/Area!

Note: Your choice of sensitivity study should lead to valuable observation which can be used to improve the design. If you can't find valuable observations to improve the performance of your design, you won't receive any point for that [Motivation Section of your final paper].

1.5 Putting it all together

Measure the speedup provided by having both of your designs running: a dynamic replacement policy for LLC and dynamic instruction prefetching for I-cache. Show box plot of speedup over no dynamic base-line.

2 Deliverable

- Merge the reports of your Assignment I, II and III into one document with ACM/IEEE two column paper format. Make it look like a ISCA/MICRO/HPCA paper that we studied in class. Title it e.g, Exploring Adaptive Cache Management for Instruction & Data.
- Your paper should include abstract, introduction, background & Motivation, methodology, results, related works and references (page limit 15 pages).

Bonus point (10%): Show results for multi-core. All multi-core traces (120 traces) are on the DSMLP server in .gz format. The path is /datasets/cs240c-wi22-a00-public/data/Assignment2-multicore/ on DSMLP server.

- Submit by March 4th, 2022, at 11:59 PM on Canvas under HW3.