Two-Level Adaptive Branch Prediction

Read Sections 1 through 3 of "Two-Level Adaptive Training Branch Prediction" by Tse-Yu Yeh and Yale Patt, then answer the questions below. Please complete this assignment in groups of two or three.

- The paper is available from the ACM Digital Library through the GVSU library. (You will need to access the ACM Digital Library through the GVSU library site to get free access.)
- Some of the questions will refer to sections 4 through 6. You need not read these sections in their entirety — just skim them as necessary.

1. When averaged across all the programs used by the study, approximately what percent of branches were taken? (This would be the accuracy of a static "predict taken" branch predictor.)

2. What are the pros and cons of using profiling to establish a set of static predictions for a program?

3. Which benchmark was used for the experiments in this paper?

4. Which of the automata from Figure 2 were discussed in detail in lecture?

5. Explain the difference between automata A3 and A4 using non-technical language (i.e., explain it like you would explain it to your grandmother).

6. Explain the difference between an Associative History Register Table (AHRT) and a Hash History Register table (HHRT). What are the pros and cons of each?

7. What makes the Ideal History Register Table (IHRT) ideal?

8. How do the authors avoid needing to make two lookups during one cycle?

9. For context: Figures 5 and 6 examine different configurations of the author's proposed branch predictor (i.e., different table sizes and automata). According to Figure 5:

   (a) How much better is the best four-state automaton than the one-bit automaton (LT)?

   (b) How much difference is there between the four-state automata?

10. According to Figure 6, approximately how much better (as a percent) is AHRT than HHRT? (The horizontal lines are at .84, .88, .92, and .96.)

11. According to Figure 6 and section 5.1.2,

    (a) how many rows do the authors put in the history table? .

    (b) How many rows in the pattern table?

    (c) Assuming 32-bit MIPS instruction address (that are always word-aligned), how much bigger (as a percentage) is the history table than the pattern table?

12. Figure 10 compares the best two-level adaptative predictor to other, existing branch prediction algorithms.

   (a) Which existing algorithm comes closest to the new algorithm's performance?

   (b) How much better is the proposed algorithm's accuracy?

   (c) Assuming 11% of instructions are branches and a 17-cycle branch penalty, what is the speedup of a CPU using a two-level adaptive predictor over the same CPU using the next-best predictor?