Name ______________________________

Computer Architecture Research Methods

EE 7700-1

Take-Home Final Examination

10 May 2006, 21:30- 15 May 2006, 9:00 CDT

Problem 1	________	(25 pts)
Problem 2	________	(25 pts)
Problem 3	________	(25 pts)
Problem 4	________	(25 pts)
Exam Total	________	(100 pts)

Alias ______________________________

Good Luck!

Problem 1: In the Alloyed predictor the PHT is indexed with a concatenation of global history, local history, and a part of the branch address. The rationale is to provide both local and global history, since some branches might correlate with only one kind of history or the other. However, the size of each history is much smaller than a PAs or GAs of the same size.

Consider a variation in which the global and local history bits are exclusive-ored together, call it MAx. Compare that to a MAs system with the same PHT size. (25 pts)

(*a*) Describe a case in which MAx would perform better than MAs. Illustrate using a code example.

(*b*) Describe a case in which MAx would perform worse than MAs.

(*c*) Given our experience with predictors so far, and using whatever other sources you like, do you think that overall MAx would do better or worse than MAs? Explain your answer.

Problem 2: In the taxonomy study certain branch mispredictions were classified as *training* (warmup) mispredictions. As we mentioned in class, one disadvantage of gshare (and GAs) is long warmup time, so we might expect to see a large number of warmup misses when those predictors are used. (25 pts)

(*a*) Consider a benchmark in which GAs predicts much more accurately than a bimodal predictor. Nevertheless, the system using GAs still suffers what we would call warmup (training) misses. Why wouldn't those warmup (training) mispredictions be counted as training mispredictions by the system described in the paper?

(*b*) Suggest a way to count the number of warmup mispredictions. That might be a modification of the technique in the paper, or describe it in terms of the simulators we used. Be sure to put sufficient detail in your answer. (It's okay to ask if an answer is detailed enough.)

Problem 3: Consider the code below, recycled from this semester's EE 4720 final exam, executing on a 32 k entries MAp system described in the paper. The value of the PC bits used to index the PHT are the same for branches B1 and B2. Branches B1 and B2 are taken with probability .5 and their outcomes are independent of everything, including the other random branch. (25 pts)

```
BIG:lb r3, 0(r4)
B1: beq r3, r0  S2  # Random.
    lb r3, 1(r4)
S2: nop
B2: beq r3, r0  S3  # Random.
    addi r4, r4, 2
S3: nop

    lw r1,0(r2)
B3: beq r1, 0  SK   # T N T T N  T N T T N  T N T T N  T N T T N  ...
    nop
    nop
SK: j BIG
    addi r1, r1, 4
```

(*a*) The data in Figure 6 shows the classification of mispredictions of all branches. It is also possible to classify mispredictions for an individual branch. Estimate what that would look like for branch B1 for the GAs predictor. (Do this by hand, don't write a simulator!) That is, show the percentage of mispredictions that would be classified as PHT conflict, training, etc.

Problem 4: In Figure 6 the number of PHT conflict misses is high for the local predictor. (25 pts)

(*a*) Why does the sharing of PHT entries by more than one branch have a much more negative impact on performance on GAs than on PAs? Why might PHT entry sharing improve the performance of PAg (local predictor as described in class, no PC bits in PHT index) over PAs?

(*b*) Extra Credit: If PHT entry sharing in PAs is not harmful, then why would a large proportion of the mispredictions be classified as PHT conflict for the local predictor? (See Figure 6.) *WARNING: I don't know the answer!*