Out-of-order Processing and Memory Operations

Victor Ying

(slides adapted in part from prior 6.823 offerings)

Previously: Complex Pipelines

- Scoreboarding
- Variable-latency execution units
- Out-of-order (OoO) processing
  - OoO Issue, OoO completion, in-order retiring (commit)
  - Register renaming

Out-of-Order (OoO) Summary

- OoO Processor: Restricted “data-flow” machine
  - Dynamically builds the data-flow graph
- Tolerates long latency operations by executing independent instructions in parallel
- The dynamically constructed data-flow graph is limited to the instruction window

OoO memory operations

- Can the ROB (Issue Queue) track memory dependences?
- Must respect all dependences:
  - WAW, WAR
    - buffer store in **store queue** until commit
  - RAW
    - If store executed before load:
      - Store-to-load forwarding (from store queue/buffer)
    - If load executed before store, need to disambiguate. Three solutions:
      - **Load queue** search when store executes
      - Re-execute the load at commit-time
      - Stall the load

When to bring data into the cache?

- Reads
- Writes?
  - Write allocate vs. no-write allocate
- Prefetching
- Advantages and disadvantages of bringing more data into the cache?

Branch Prediction

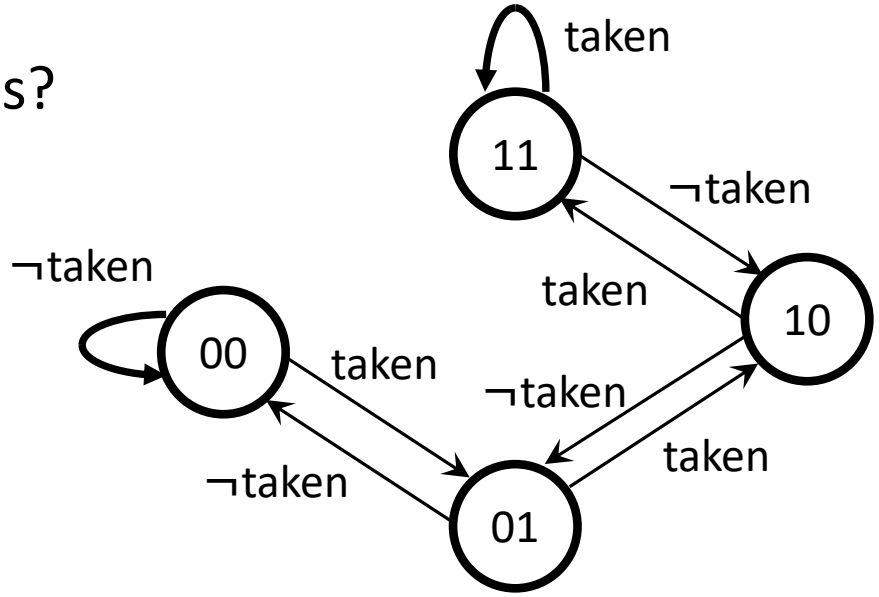
Control Flow Dependences. How to handle them?

- Stall: Delay until we know the next PC
- Speculate: Guess next value
- Do something else: Multi-threading

Why is branch prediction crucial for out-of-order and superscalar processors?

Branch predictor (BHT) entries:

- 1-bit (bimodal) predictor
- 2-bit predictor
  - Counter
  - Other options?

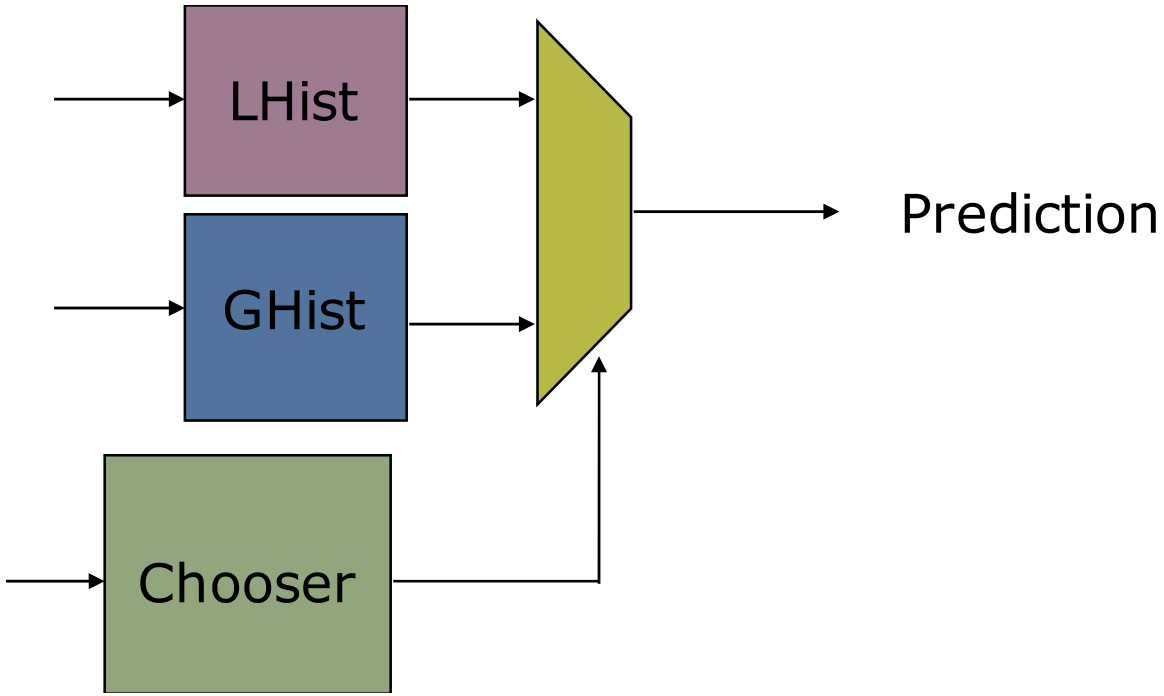


- More bits?
  - Perceptrons

How to index into the BHT?

- Some bits from PC
- Two-level predictor uses recent branch outcomes to compute index
  - Global history register
  - Local history table
- Combination of the above
  - Gshare uses XOR of bits from PC and global history

Tournament Predictors



Reminders

- Lab 2 due at 11:59pm Eastern Daylight Time (UTC+4)
- Review session for Quiz 2
  - 6pm on Tuesday
  - Using the same Zoom URL as recitation