Why do you predict?

• So you can fetch more stuff instead of waiting
• Pipeline depth more = longer wait
• Memory dependencies more = longer
• Application dataflow graph
• Hardware structural constraints
How do you predict?

• Whatever you want:
  – Prior knowledge
    • Prior per-branch behavior
    • Branch history
    • More global characteristics
  – Heuristics
  – Branch address
  – Branch type
When do you predict?

- In fetch, which is a little late

- What is branch history?
  - Just some bits in a shift register
  - When do you update?
    - When you execute the branch
    - At completion
    - At prediction time
What is PAt table?

• 2 bit saturating counter
How do multiple predictors get used?

- You can vote
- Predictor predictor
- Confidences
- Choose at random
What resource does it take to speculate?

• Global history shift register has finite size

• Rollback:
  – Restore Architectural registers
  – Flush the pipeline