My Google Experience
which might help you

Shouyuan Chen
Outline

• Interview

• Machine Learning Research
Interview

• Data collected
  – from my own interviews
  – from other interns
  – from employees’ websites (not encrypted)
Interview at Google

• coding problems
  – implement some standard algorithms
    • all programming languages are acceptable.

• “algorithmic” problems
  – most of them are extremely simple,
  – interviewers are not that familiar with “theory”.

• system/design problems
  – system oriented designing problems

• never heard of any brain teasers
Coding problems

• ASCII art
  
  1
  2 3
  4 5 6

• high-precision arithmetic
  – BigInteger

• graph algorithms
  – traversal
  – shortest path (with negative edges)
Algorithmic problems (1)

• Streaming algorithm
  – take K uniform samples from a data stream

• Algorithmic tricks
  – find the median of 2 sorted arrays
  – point in a polygon
Algorithmic problems (2)

• Dynamic programming
  – counting (closed-form or generating function)
  – tree-related problems
    • find farthest pair of vertices in a tree
  – optimization, game theory and a lot more.
    • 2 stone piles, 2 players, take 3 stone each turn, who win
    • further: NIM game seems to be a popular problem
System/Design problem

• From the lowest level
  – how to implement malloc?
  – what happens when keystroke?

• To the highest level
  – how to implement Facebook?
Machine learning research

• ML research at Google

• ML research folklores
ML Research at Google

• Focus
  – Practical and Simple
    • Logistic regression is among the most popular algorithms!
  – Efficiency
    • Parallelization (MapReduce)
      – Spectral Clustering, SVM and LDA
    • Online algorithms

• Different from other research labs (e.g. MSRA)
  – Every research are product oriented
    • Have immediate applications
How Googler do ML-related research
(so far as I observed)

1. Motivated from real products/features
2. Collect tremendous amount of data
3. Design a simple (but reasonable) algorithm
4. Run it on hundreds of machines
5. Change the algorithm and try again
ML Research Frontier and Folklores

• Once a year, Michael Jordan had 10 papers been rejected from NIPS.
• Bayesian nonparametric is a private club.
• Deep learning is considered as the next big thing.
Thank you!