P vs NP
Problem Hardness

张家琳
zhangjialin@ict.ac.cn
中科院计算所

Outline

- Definition of P, NP (计导)
- Definition of NP-hard, NP-complete（理论）
- Reduction（理论）
- More reductions

Millennium Prize Problems

P versus NP problem
Hodge conjecture
Poincaré conjecture (solved)
Riemann hypothesis
Yang–Mills existence and mass gap
Navier–Stokes existence and smoothness
Birch and Swinnerton-Dyer conjecture

Two types of problems

- Decision problem
  - The answer is YES/NO
  - Ex: s-t CONNECTIVITY problem
- Optimization problem
  - The answer is a “best” solution
  - Ex: SHORTEST-PATH problem
- Transfer an optimization problem to a decision problem

Definitions (intuitively)

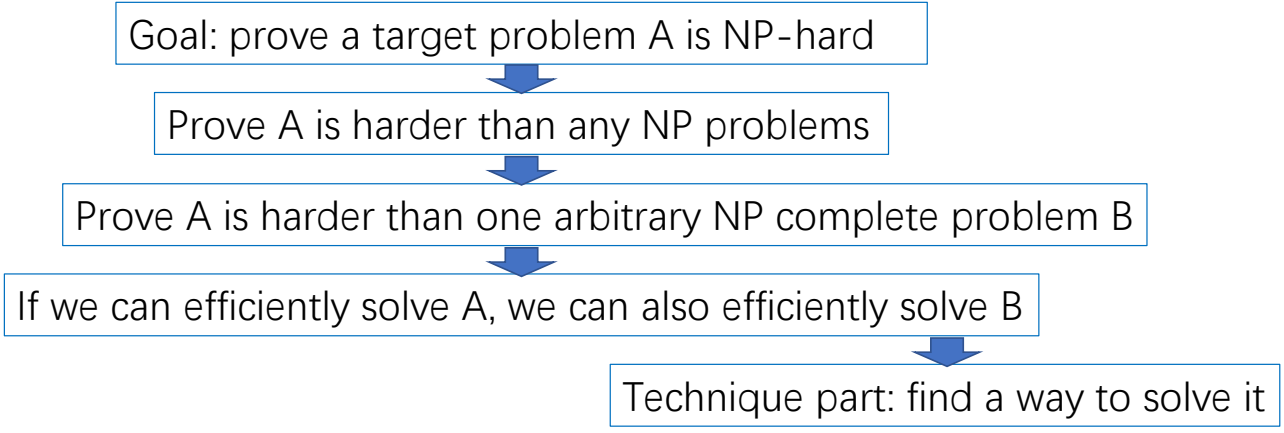
- P: can be computed fast
- NP: may be or may not be computed fast, but can be checked fast
- NP-hard: not easier than any NP problem
- NP-complete:
  - NP-hard
  - NP

Definitions (formally)

- A is a decision problem
- A is in P
  - A can be computed in polynomial time
- A is in NP
  - A can be checked in polynomial time
- A is NP-hard
  - If A can be computed in polynomial time, any problem in NP can be computed in polynomial time
- A is NP-complete:
  - A is NP-hard
  - A is in NP

Reduction

B is reducible to A: $B \leq_p A$



Go back to Lec3.pdf page 103

More reduction examples

- “求助外卖优惠券问题：假设我点了n个菜，该餐馆有k种满减优惠，那么应该怎样切分订单才能获得最大优惠？”
- First step: transform it to a decision problem

- Second step: reduction
- Subset sum: given a set of integers and an objective value *w*, is there a non-empty subset whose sum is *w*?
- A special case —— partition problem: given a set of integers whose sum is *s*, is there a non-empty subset whose sum is *s/2*?

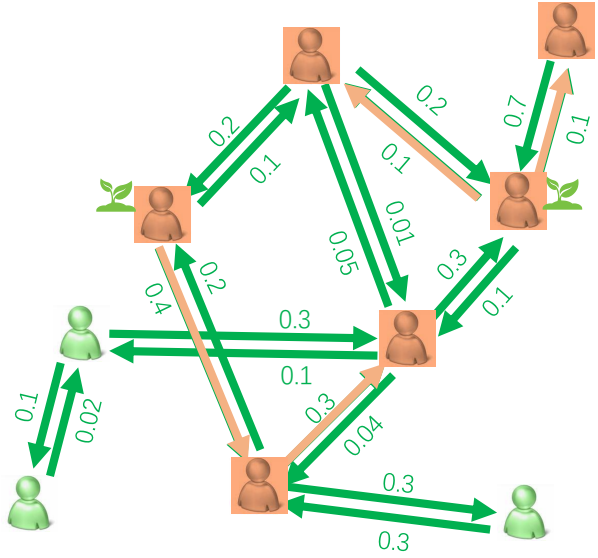
外卖优惠券

- The problem is NP-hard even if k=1, the number of coupon >=2

- What if k=1 and the number of coupon =1?
- What if k>=2, but the number of each coupon = 1?
- What if k=1 with unlimited number, but the coupon is fixed (not part of the input)?

Maximize influence in social network

- **Influence model: independent cascade model(IC)**
- Each edge (u, v) has a *influence probability* $p(u, v)$
- Initially seed nodes in S_0 are activated
- At each step t, each node u activated at step $t-1$ activates its neighbor v independently with probability $p(u, v)$
- The influence spread $f(S_0)$ be the expected number of activated nodes.



Maximize influence in social network

- **Problem definition**
- Given a social network in IC model, and a number k, find a seed set S of at most k nodes such that the influence spread of S is maximized.

- **Hardness**
- Given S, to compute $f(S)$ is #P-hard, but can be approximate to arbitrary precision by using Monte-Carlo simulation
- If we have a oracle to compute $f(S)$, is the problem hard?

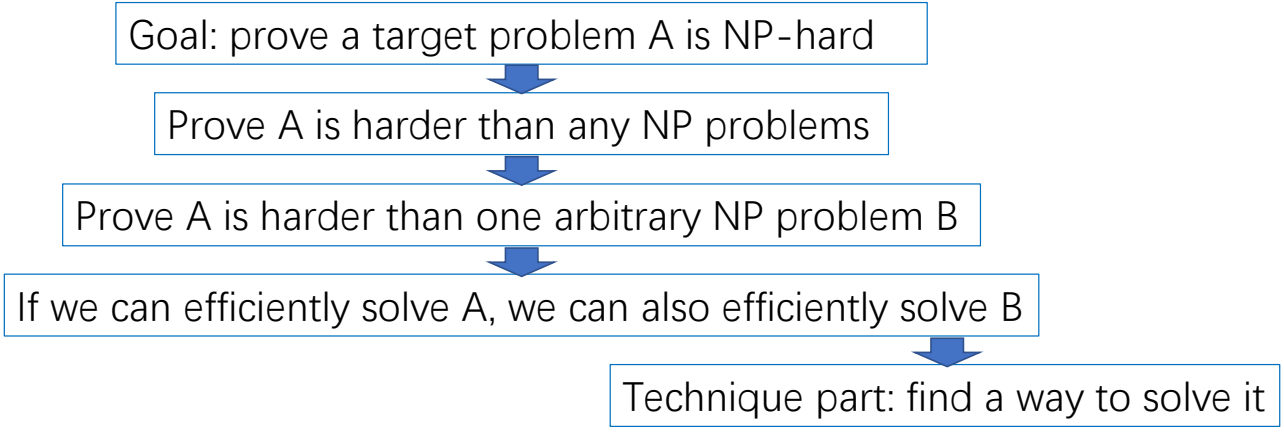
Maximize influence in social network

- **Reduction**
- First step: transform it into decision problem.
- Second step: Set cover
- **Observation**
- Hardness does not come from
  - Compute f(S)
  - Probability
- **Further problems**
  - What if the model is undirected graph?
  - What if the model is undirected graph + planar graph?

  - Ref：Influence Maximization in Undirected Networks, SODA 2014

If you want/need to prove NP-hard in the future……

- Method 1: find a friend to do it for you
- You need to understand your friend's proof
  - Understand the logic behind the NP-hard proof



If you want/need to prove NP-hard in the future……

- If you must do it yourself, you need to figure out the technique part
- Step 1: find enough existing NP-hard problem
  - Your textbook
  - Wikipedia: **List of NP-complete problems**
  - http://www.nada.kth.se/~viggo/problemlist/compendium.html (NPC Bible)
- Step 2: get familiar with the reduction
  - Practice, practice, practice

Thanks!

Question?