算法基础 2022 春 Homework 5

任课教师：陈雪 due: May 10, 23:59

（1）只准讨论思路，严禁抄袭

（2）只能阅读 bb 上的材料和教材算法导论。严禁网上搜寻任何材料，答案或者帮助

以下所有算法的设计请给出伪代码，证明算法正确性以及时间复杂度。

问题 1 (15 分). (a) 我们在课堂上展示了切比雪夫的素数分布定理：存在 $C_1, C_2 = O(1)$ 使得 $\pi(n) \in [C_1\frac{n}{\log n}, C_2\frac{n}{\log n}]$。证明该定理能导出以下结论：存在常数 C，使得对所有足够大的正整数 n，区间 $[n, Cn]$ 中都至少包含一个质数。

(b) 费马小定理告诉我们，对于任意一个质数 p，若整数 a 满足 $(a,p) = 1$，则 $a^{p-1} \equiv 1 \mod p$。但是它的逆命题却不一定成立。我们称满足费马小定理结论却不是质数的数 n 为伪质数。(1) 证明 $n = 561$ 是一个伪质数；(2) 计算有多少 $a \in \{1, 2, \cdots, n-1\}$ 满足 $(a, n) = 1$。

Answer:

(a) What we need to prove is equivalent to $\pi(Cn) \geq \pi(n) + 1$. Let C be $2 \cdot \frac{C_2+1}{C_1}$. Then

$$
\begin{aligned}
\pi(Cn) &\geq C_1 \cdot \frac{Cn}{\log Cn} \\
&= \frac{(2C_2+2)n}{\log(2C_2+2)n} \\
&= \frac{2C_2 n}{\log(2C_2+2) + \log n} + \frac{2n}{\log(2C_2+2) + \log n} \\
&\geq \frac{C_2 n}{\log n} + \frac{n}{\log n} \\
&\geq \pi(n) + 1
\end{aligned}
$$

So there exists a prime at least in $[n, Cn]$.

(b) As $561 = 3 \cdot 11 \cdot 17$, any integer a that $(a, n) = 1$ is also coprime with $3, 11, 17$.

By Fermat's little theorem, it implies

$$\begin{cases} a^{560} \equiv a^2 \equiv 1 \mod 3 \\ a^{560} \equiv a^{10} \equiv 1 \mod 11 \\ a^{560} \equiv a^{16} \equiv 1 \mod 17 \end{cases}$$

By Chinese remainder theorem, $a^{560} \equiv 1 \cdot 1 \cdot \frac{561}{3} + 1 \cdot 8 \cdot \frac{561}{11} + 1 \cdot 16 \cdot \frac{561}{17} \equiv 1123 \equiv 1$ mod 561.

(c) Assume $n = p_1^{t_1} p_2^{t_2} \dots p_k^{t_k}$, where all of p_i are primes. $(a, n) \neq 1$ iff $p_i|a$. So there are $n \cdot \prod_{1 \leq i \leq k} (1 - \frac{1}{p}_i)$ numbers that is coprime with n in $\{1, 2, \cdots, n-1\}$.

问题 2 (25 分)**.** 给定一个二分图 $G = (L \cup R, E)$，设计算法找到 G 中的最小顶点覆盖 (vertex cover)。

请给出伪代码（可以直接调用最大流过程），证明算法正确性以及时间复杂度。

提示：考虑顶点覆盖与割的关系。

Answer: Without loss of generality, we assume that G is a connected graph.

The maximum match contains independent edges in bipartite graph. To cover these edges, one of the edge's endpoints must be included in the vertex cover. So the vertex cover is minimum if it equal to the maximum match in a bipartite graph.

Let U be the set consist of all unmatched vertex in L after we found a maximum match.

And $Z = \{v | v \in L \cup R, v \textit{ has a cross path to some vertx in } U\}$. $S = L \cap Z, T = R \cap Z$.

By definition of Z, any neighborhood of vertex in S is belong to T. So T can cover all edges between T and S, and $L - S$ covers the rest of edges. Then $(L - S) \cup T$ is a vertex cover of G.

On the other hand, as Z is generated by cross paths in a maximum match, all vertexes in T and $L - S$ is matched. Otherwise, We can get a larger match by cross path. Therefore, $T + L - S$ reachs the infimum of vertex cover, namely a minimum vertex cover.

And the maximum match problem can be transform into a maximum flow problem.

Algorithm 1 mininum cover algorithm

---

function MININUM-COVER($G = (L \cup R, E)$)
    find the maximum match by maximum flow
    Start at all unmatched vertexes in L to get cross paths, mark each passing vertexes
    outputs the sets contains all marked vertexes in R and all unmarked vertexes in L

The complexity is same as the algorithm of maximum flow.

问题 3 (20 分)**.** 我们可以将一般图 $G = (V, E)$ 上的顶点覆盖 (vertex cover) 问题转换为一个线性规划问题。

(a) 使用变量 $x_1, \ldots, x_n \in [0, 1]$ 来指示每个顶点的状态（严格来说，我们希望 $x_i = 1$ 或 0 来指示在覆盖集中与否—但是线性规划需要 x_i 的范围组成凸包）。请设计出线性规划的目标和约束。

(b) 解出此线性规划后，得到 $x_1, \ldots, x_n \in [0, 1]$ 的实数解，然后令 $S = \{v_i | x_i \geq 1/2\}$。证明 S 是一个顶点覆盖，且同最小的顶点覆盖相比，S 的近似比不大于 2。

Answer:

(a) Given a graph $G = (V, E)$, $|V| = n, |E| = m$. Let x_i indicates vertex v_i whether or not be included in a vertex cover set. Then our goal is minimize $\sum x_i$. And the basic constraints is $x_i \geq 0$. To make sure every edge is covered, we add constraints $x_i + x_j \geq 1$ for all $(v_i, v_j) \in E$.

$$\min x_1 + x_2 + \cdots + x_n$$
$$\begin{cases} x_i + x_j \geq 1 \\ \vdots \\ x_1, x_2, \cdots, x_n \geq 0 \end{cases}$$

And the constraints that $x_i \leq 1$ is implicit in the linear programming. As our goal is minimizing and the only additional constraints is $x_i + x_j \geq 1$, the part beyond 1 is useless.

(b) The constraints $x_i + x_j \geq 1$ imply $\max\{x_i, x_j\} \geq 1/2$, which means S constains one of them at least. So S cover every edge in G, and is a vertex cover.

And It's obvious that the minimum vertex cover S_{opt} is also a solution of the above linear programming. So $\sum x_i \le |S_{opt}|$. As x_i is nonnegative, it infers that $|S| = |\{v_i|x_i \ge 1/2\}| \le \sum x_i/(1/2) \le 2 \cdot |S_{opt}|$.

So S is a vertex cover, and it's approximation ratio isn't more that 2.

问题 4 (15 分)**.** EQ-SUM 问题：给定 n 个正整数 $s_1, s_2, \cdots, s_n$，找到一个划分将它们分为两部分 P_1, P_2，并满足两部分的和相等，即 $\sum\limits_{s_i \in P_1} s_i = \sum\limits_{s_j \in P_2} s_j$。证明 EQ-SUM 问题是一个 NPC 问题。

注: 可以从任何已知的 NPC 问题作规约（包括 3SAT，CLIQUE，Vertex-Cover，Subset-Sum）。

Answer:

At fisrt, the result of EQ-SUM problem can be verified in poly-time as we just need to compare the sums of two parts.

Moreover, SUBSET-SUM problem, one of NPC problem that had been mentioned in our course, has a reduct to EQ-SUM problem in poly-time.

$L_{SUBSET-SUM} = \{< S = (s_1, s_2, \cdots, s_n), t >: \exists S' \subset S \ s.t \sum\limits_{s_i \in S'} s_i = t\}$.

Let $X = \sum\limits_{s_i \in S} s_i$, namely the sum of input.

The case 1 is $X = 2t$, then we direct take $S = (s_1, s_2, \cdots, s_n)$ as input of EQ-SUM problem. Then both partitions EQ-SUM outputs are the result of SUBSET-SUM problem.

The case 2 is $X < 2t$, EQ-SUM problem's input is $(s_1, s_2, \cdots, s_n, 2t - X)$. Then one of partitions is the result of SUBSET-SUM problem.

And the case 3 is $X > 2t$, EQ-SUM problem's input is $(s_1, s_2, \cdots, s_n, X - 2t)$. Then one of partitions' complementary set is the result of SUBSET-SUM problem.

Therefore, EQ-SUM problem is a NPC problem.

问题 5 (25 分)**.** 染色问题：给定一个图 $G = (V, E)$，找到最少的颜色数量 k，使得 G 是可以被 k-染色的（图 G 中每个顶点都被染上 k 种颜色之一，且相邻顶点的颜色都不同）。

(a) 染色问题是 NP 难 (NP-hard) 问题吗？证明你的结论。

提示：通过证明此优化问题的判定形式为 NP-complete 来得到 NP 难的结论。

(b) (5 分) 染色问题是 NPC 问题吗？

(c) (5 分) 对于 $k = 2$ 的特殊情况，设计一个多项式时间算法找到图 G 的一个 2-染色实例。请给出伪代码，证明算法正确性以及时间复杂度。

Answer:

(a) Yes, coloring problem is a NP-hard problem. The proof is as follows. By definition of NP-hard problem, we just need to prove that a NPC problem has a reduction to it. And we choose the 3-SAT problem to finish the proof. the proof is in CLRS 34-3. So coloring problem is a NP-hard problem.

(b) No, coloring problem isn't a NPC problem. Because we cann't verified a coloring instance has the smallest number of color in poly-time.

(c) When $k = 2$, coloring problem is equivalent to verified whether the graph is a bipartite graph. Specific algorithm is omitted. And the complexity of the algorithm is $O(|V|+|E|)$.