18.433 组合最优化

NP-完全问题

Nov 18 *授课教师：Santosh Vempala*

到目前为止，我们已经找到了匹配、流、线性规划和凸规划问题的许多有效算法，它们都是多项式时间算法。

但是，仍然有很多问题还没有找到多项式算法。NP-完全理论把这些问题统一起来。粗略地说，一个 NP 完全问题是与一大类问题同等难度的问题。例如，旅行售货员问题(TSP)，整数规划问题(IP)，最长圈问题和适定性(SAT)问题都是这类问题。NP 完全性告诉我们，在准确意义上它们都是同等难度的。让我们详细地看一下每个问题。

1、 旅行售货员问题：
   假设有某推销员，要走遍 n 个城市，城市 i 和 j 之间的距离为 $w_{i,j}$，他想经过每个城市恰好一次而且旅行路线（时间）最短。换句话说，有一个结点 i 和 j 之间长度为 $w_{i,j}$ 的完全图 $G=(V,E)$，我们的问题是：什么是恰好经过每个顶点的最短圈？

2、 整数线性规划
   假设有下面的线性规划问题：

$$\min c^T x$$

$$Ax \le b \text{ for } x_i \ge 0$$

   这是一个典型的线性规划问题，现在如果要在 x_i 上增加一个整数约束，即要求 x_i 必须为正整数，那我们就得到了一个整数线性规划(ILP)。

3、 适定性
   适定性问题(SAT)用如下的布尔表达式：

$$f=(x_1 \vee x_2 \vee x_4) \wedge (x_3 \vee \bar{x}_4) \ldots.$$

   其中，$x_i = \{\text{True,False}\}$，且利用我们都知道的布尔恒等式。F 有一个令人满意的分配吗？我们能否找到 x_i 的取值使 F 中每个子句取值都为 1 呢？

4、 最长圈问题
   给定一个图 $G=(V,E)$，寻找最长的圈。

5、 团
   团是一个完全子图。给定图 $G=(V,E)$，找一个基数最大（顶点数最多）的团。

上面的例子看起来各不相同，但他们有一些共同的属性：

A) 任何一个问题迄今都未发现多项式时间算法。

B) 如果其中一个找到了多项式算法，那么其它的也都有。

1 最优化问题 vs 判定问题

对上述的每个问题，性质 A 比较容易得到，而性质 B 则不那么容易。为了理解这个性质，我们首先给出这些最优化问题的判定形式。

在费用函数为 c 的可行解集合 F 中找一个最优解
Vs
是否有一个费用小于等于 L 的可行解？

如果最优(OPT)问题解决了，那么判定(DEC)问题也就解决了。即，DEC 可归约到 OPT。那么，OPT 可以归约到 DEC 吗？以 TSP 为例，我们要问：是否有一个途径小于等于 L？接下来，我们进行二进制搜索来寻找长度最短的途径，不妨设为 S。但是，我们仍然不知道 S 是哪一条。找出它的一个算法如下：

拿出一条边 e。
问同样的图是否仍有一个途径小于等于 S。
如果是，那说明我们不需要边 e，删掉它。
如果不是，那么我们保留边 e，因为它将是我们途径的一部分。
对所有的边重复这个算法。

在例子 ILP 中，我们要问：是否存在一个费用 $\leq L$ 的 x？一个方法就是令 $x_i = 0$，如果最优值保持不变，那么我们可以固定这个特殊的 x_i 为 0。

对于最大团问题，OPT 问题是：找到最大团。而 DEC 问题是：是否存在一个大小 $\leq k$ 的团？要找最优大小 k^*，我们再来做二进制搜索。然后考虑 v_i 和它所有的相邻点形成的图，如果这个图中的最优值保持不变，那么保留这个顶点，而删掉其它的顶点。否则，删掉 v_i，因为它不在我们的最大团中。

2 P 和 NP

2.1 定义

P：判定问题可以在多项式时间内解决的一类问题。
NP：可以简短的证明判定问题的答案为“是”，即证明长度是输入规模的多项式函数，且验证算法可以在多项式时间内完成。

注意 P 问题对“是”和“否”答案都有简短证明，这就意味着 $P \subseteq NP$。让我们来看一个 P 问题：
线性规划：最小值小于某个 c？
　　是：给出一个 $\leq c$ 的可行解

否：利用其对偶问题给出一个更小的下界

接下来看一下上述的 NP 问题：

1、TSP，是否有一个途径 $\leq L$？

是：给出一个途径

否：？

2、SAT，是否存在一个令人满意的分配？

是：给出一个满意的分配

否：？

3、最小 ILP，最小值 $\leq c$？

是：给出一个 $\leq c$ 的可行解

否：？

这就引导我们考虑这个问题：是否有 $P = NP$？

2.2 归约

问题 A 到 B 的一个归约是一个函数 $f : A \rightarrow B$ 满足对所有的 x 有

$$x \in A \Leftrightarrow f(x) \in B.$$

如果其中函数 f 可以在多项式时间内计算出来，那么称其为一个多项式归约。它的一个意义如下：

如果 B 存在一个多项式算法，那么 A 也存在。

如果 NP 类中的每个问题都有一个到 B 的多项式归约，那么称 B 是 NP-hard 的。另外，如果 B 也在 NP 中，那么它是一个 NP 完全问题。

如果 A 是一个 NP 完全问题，且其有一个到 NP 类中 B 问题的一个归约，那么 B 也是 NP 完全问题。

2.3 归约的例子

SAT 是一个 NP 完全问题（在本课中不予证明）。

1、ILP 是 NP 完全问题。取下面的 SAT 问题，看它是否能用一个 ILP 问题解决。

$$F = (x_1 \vee x_2 \vee \ldots \vee \bar{x}_i) \wedge (x_4 \vee \bar{x}_5) \wedge \ldots \wedge (x_a \vee x_b \vee \ldots \vee x_c)$$

这个 SAT 问题可以写成如下形式：

$$x_1 + x_2 + \ldots + \bar{x}_i \geq 1$$

$$x_4 + \bar{x}_5 \geq 1$$

$$x_a + x_b + \ldots + \bar{x}_c \geq 1$$

$$\cdots$$

$$x_i = \begin{cases} 1, & \text{then true} \\ 0, & \text{then false} \end{cases}$$

既然 SAT 可以归约到一个 ILP 问题，所以 ILP 是 NP 完全问题。

2、 最大团是 NP 完全问题

用下面的构造 SAT 可被归约到团问题。假设我们有一个具有 m 个子句的公式 F。

1） 定点转化为形式 $< x_a, i >$，这里 x_a 是出现在子句 C_i 中的字母。

2） 边转化为形式 $\{< x_a, i >, < x_b, j >\}$ 对于所有的 $x_a \neq \bar{x}_b$ 和 $i \neq j$。

通过这种方式来定义顶点和边，可以保证所有连通的顶点是相容的，因为他们的真实值不交叠。如果我们在这个图中找到了大小为 m 的团，那 F 是适定的。见图 1。

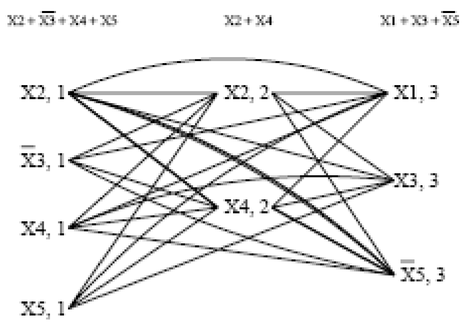

图 1：团是 NP 完全问题