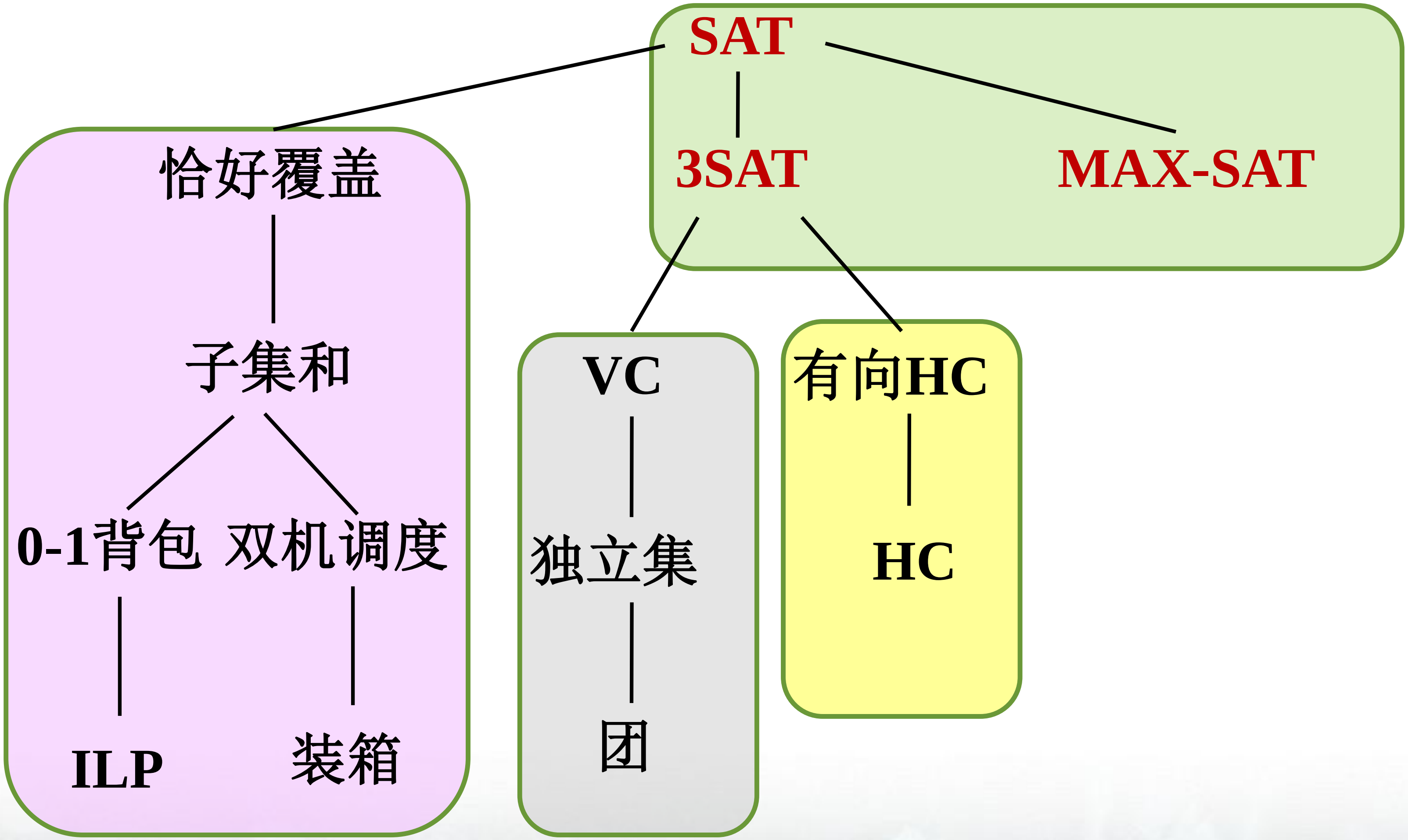
NP完全性证明



顶点覆盖,团,独立集概念

设无向图$G=\langle V,E\rangle$, $V'\subseteq V$. V'是G的一个

顶点覆盖: G的每一条边都至少有一个顶点在V'中.

团: 对任意的$u,v\in V'$且$u\neq v$, 都有$(u,v)\in E$.

独立集: 对任意的$u,v\in V'$, 都有$(u,v)\notin E$.

引理　对任意的无向图$G=\langle V,E\rangle$和子集$V'\subseteq V$, 下述命题是等价的:

(1) V'是G的顶点覆盖,

(2) $V-V'$是G的独立集,

(3) $V-V'$是补图 $G^c=\langle V,E^c\rangle$的团.

顶点覆盖,团,独立集的问题

顶点覆盖(VC): 任给一个无向图G=<V,E>和非负整数K≤|V|, 问G有顶点数不超过K的顶点覆盖吗?

团: 任给一个无向图G=<V,E>和非负整数J≤|V|, 问G有顶点数不小于J的团吗?

独立集: 任给一个无向图G=<V,E>和非负整数J≤|V|, 问G有顶点数不小于J的独立集吗?

根据引理, 很容易把这3个问题中的一个问题多项式时间变换到另一个问题.

VC∈NPC

定理 顶点覆盖(VC)是NP完全的.

证: VC的非确定型多项式时间算法: 任意猜想一个子集$V' \subseteq V$, $|V'| \le K$, 检查V'是否是一个顶点覆盖.

要证$3SAT \le_p VC$. 任给变元 $x_1, x_2, \ldots, x_n$ 的3元合取范式

$$F = C_1 \wedge C_2 \wedge \ldots \wedge C_m,$$

其中 $C_j = z_{j1} \vee z_{j2} \vee z_{j3}$, z_{jk} 是某个 x_i 或 $\neg x_i$.

如下构造VC的实例 $f(F)$:

$$G = \langle V, E \rangle \text{ 和 } K = n + 2m,$$

其中

变换：构件设计法

$V = V_1 \cup V_2$, $\quad E = E_1 \cup E_2 \cup E_3$,

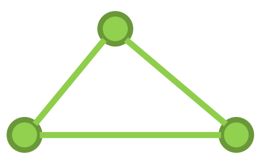
$V_1 = \{ x_i, \bar{x}_i \mid 1 \le i \le n \}$

$E_1 = \{ (x_i, \bar{x}_i) \mid 1 \le i \le n \}$

变元构件，$\forall x_i$, $i = 1, 2, \dots , n$

$V_2 = \{ [z'_{jk}, j] \mid k=1,2,3,\ 1 \le j \le m \}$,

$E_2 = \{ ([z'_{j1}, j],[z'_{j2}, j]),\ ([z'_{j2}, j],[z'_{j3}, j]),\ ([z'_{j3}, j],[z'_{j1}, j]) \mid 1 \le j \le m \}$.



析取式构件，$\forall C_j$, $j=1,2,\dots, m$

$E_3 = \{ ([z'_{jk}, j], z'_{jk}) \mid k=1,2,3,\ 1 \le j \le m \}$ 联络边

$C_j = z_{j1} \vee z_{j2} \vee z_{j3}$, 当$z_{jk}=x_i$时，$z'_{jk}=x_i$; 当$z_{jk}=\neg x_i$时，$z'_{jk}=\bar{x}_i$

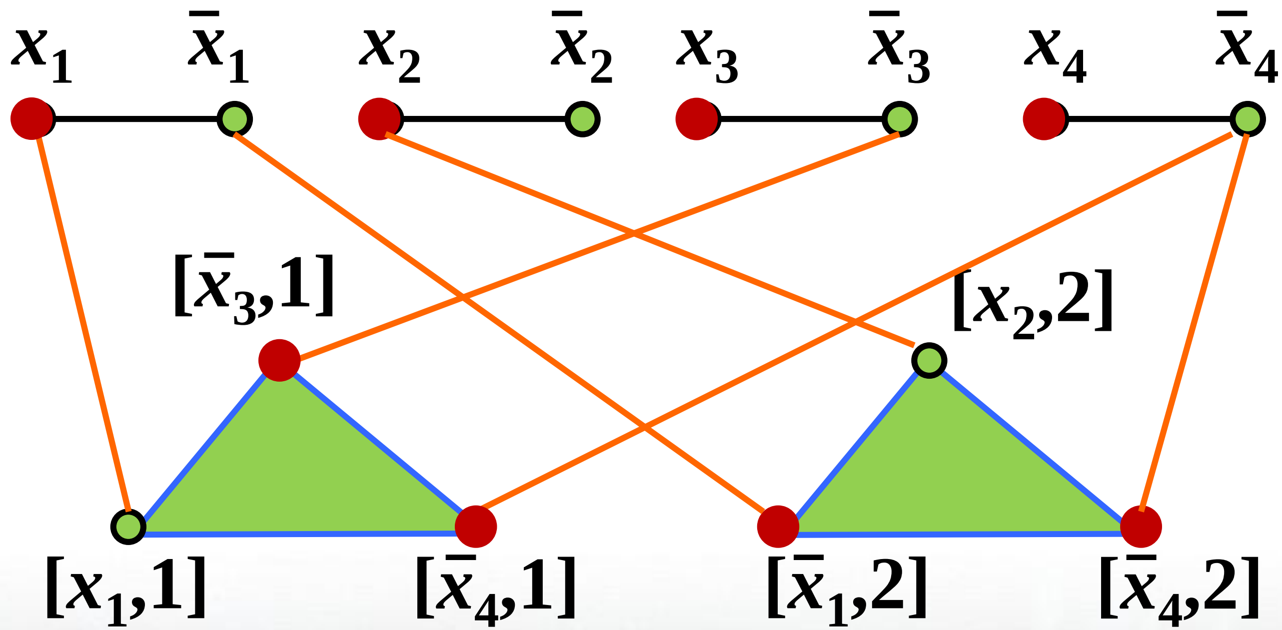
$K = n + 2m$

变换实例

$U = \{ x_1, x_2, x_3, x_4 \}$,

$C = (x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee x_2 \vee \neg x_4)$

$K = 4 + 2 \times 2 = 8$



证明

要证：F 是可满足的 $\Leftrightarrow$ G 恰好有K个顶点的顶点覆盖.

⇐ 任何顶点覆盖V'至少有$n+2m$个顶点，故任何不超过K的顶点覆盖V'恰好含K个顶点, 且在x_i和$\bar{x}_i$中取一个, 这恰好对应x_i的赋值, 取x_i对应$t(x_i)=1$，取$\bar{x}_i$对应$t(x_i)=0$. 三角形C_j的顶点$[z'_{j1},j]$、$[z'_{j2},j]$和$[z'_{j3},j]$中取2个，剩下顶点对应的变量满足C_j

⇒ 设t是F的成真赋值, 对每一个$i(1\le i\le n)$, 若$t(x_i)=1$, 则取x_i; 若$t(x_i)=0$, 则取$\bar{x}_i$. 这n个顶点覆盖E_1. $\forall j(1\le j\le m)$, C_j至少有一个文字z_{jk}的值为1. 于是, 从对应的三角形顶点$[z'_{jk},j]$引出的边$([z'_{jk},j], z'_{jk})$已被覆盖. 取该三角形的另外2个顶点即可.

$2n+3m$个顶点和$n+6m$条边, 显然在多项式时间内构造G和K,

定理 独立集和团是NP完全的.

其他基本NPC问题

有向哈密顿回路: 任给有向图D, 问:D中有哈密顿回路吗?

恰好覆盖: 给定有穷集$A=\{a_1, a_2,\dots, a_n\}$和$A$的子集的集合$W = \{S_1, S_2,\dots, S_m\}$, 问: 存在子集$U\subseteq W$ 使得U中子集彼此不交且它们的并集等于A? 称W这样的子集U是A的恰好覆盖(划分).

实例　设$A=\{1,2,3,4,5\}$, $S_1=\{1,2\}$, $S_2=\{1,3,4\}$, $S_3=\{2,4\}$, $S_4=\{2,5\}$, 则$\{S_2, S_4\}$是A的恰好覆盖. 若把S_4改为$S_4=\{3,5\}$, 则不存在A的恰好覆盖.

子集和: 给定正整数集合$X=\{x_1, x_2,\dots, x_n\}$及正整数$N$, 问存在$X$的子集$T$, 使得$T$中的元素之和等于$N$吗?

北京大學

其他基本NPC问题(续)

装箱: 给定n件物品, 物品 j 的重量为正整数 w_j, $1\le j\le n$, 以及箱子数K. 规定每只箱子装入物品的总重量不超过正整数B, 问能用 K 只箱子装入所有的物品吗?

双机调度: 有2台机器和 n 项作业 $J_1, J_2,\ldots, J_n$, 这2台机器完全相同, 每一项作业可以在任一台机器上进行, 没有先后顺序, 作业 J_i 的处理时间为 t_i, $1\le i\le n$, 截止时间为D, 所有 t_i 和 D 都是正整数, 问能把 n 项作业分配给这 2 台机器, 在截止时间 D 内完成所有的作业吗?

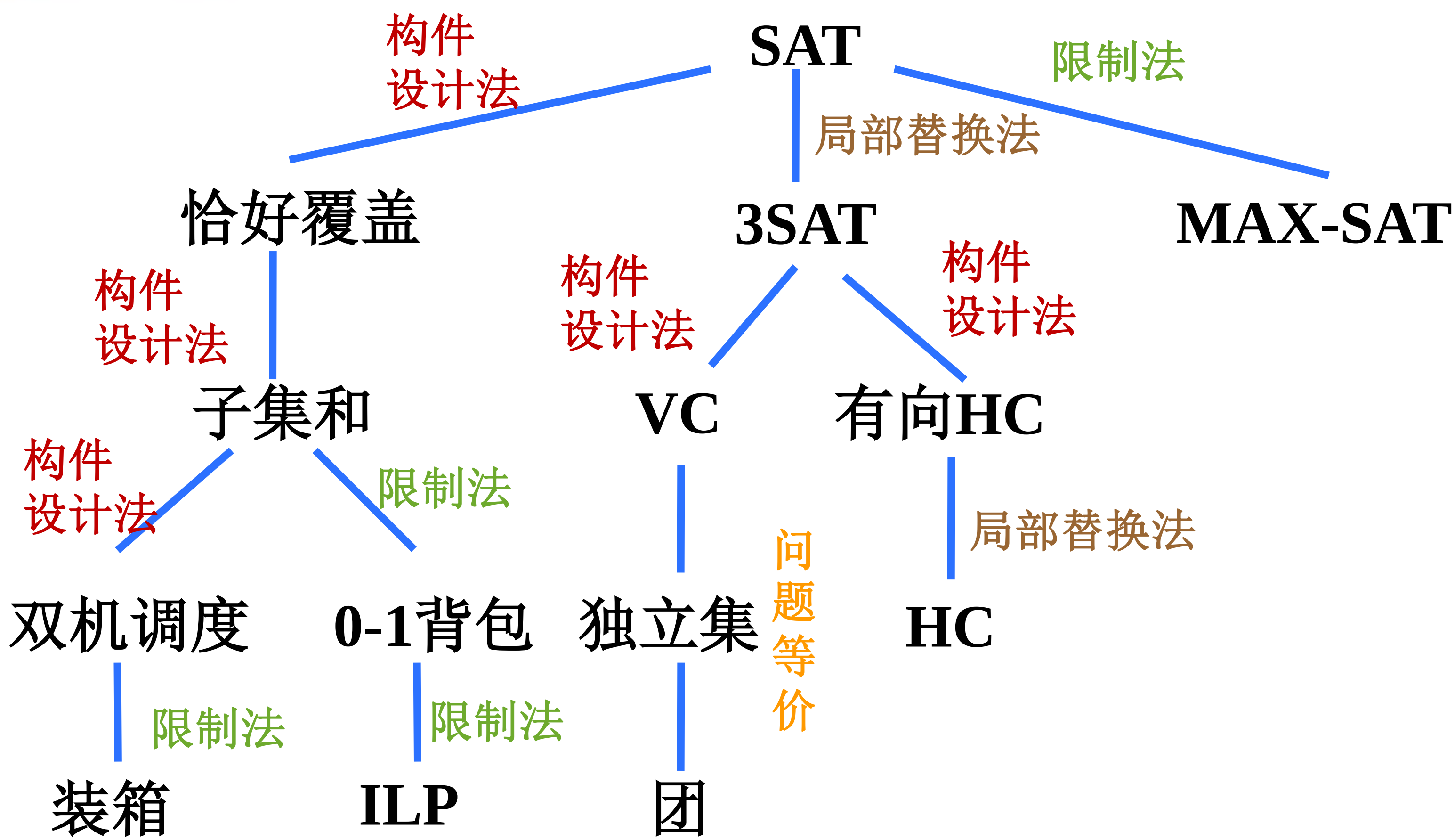
整数线性规划（ILP）：任给$m\times n$维整数矩阵和m维向量b，问下述问题：

$$AX\ge b$$

$$x\ge 0，x\text{为整数}$$

是否有解？

证明方法小结



NPC证明方法:

- 选好一个已知的NP完全问题.
- 使用限制法, 局部替换法和构件设计法.

NP完全性理论的应用

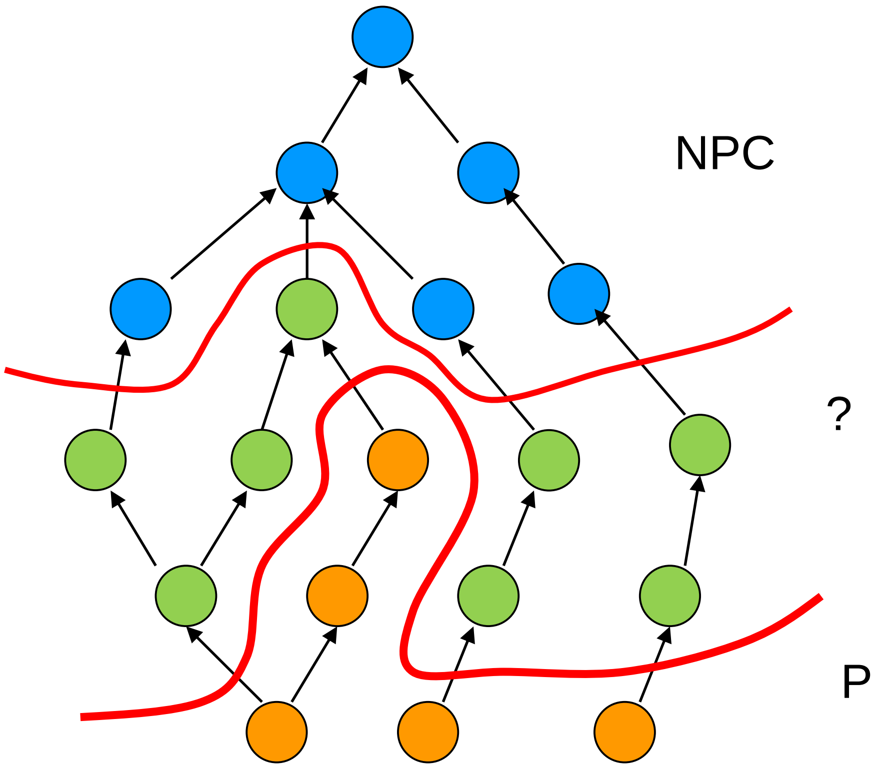
用NP完全性理论进行子问题分析

- 子问题的计算复杂性
- 子问题的NP完全性证明

推广到搜索问题与优化问题

- 搜索问题与优化问题
- **Turing**归约
- **NP-hard, NP-easy**

子问题的计算复杂性



努力扩大已知区域，缩小未知区域

当P≠NP时，存在不属于NPC也不属于P的问题

有先行约束的多处理机调度问题

优化问题:

给定任务集 T, m 台机器, $\forall t \in T$, $l(t) \in Z^+$, T上的偏序 $\prec$.

若 $\sigma: T \rightarrow \{0, 1, \ldots, D\}$满足下述条件, 则称 σ为可行调度.

(1) $\forall t \in T,\ \sigma(t)+l(t) \le D$

(2) $\forall i, 0 \le i \le D, |\{ t \in T : \sigma(t) \le i < \sigma(t)+l(t) \}| \le m$

(3) $\forall t, t' \in T,\ t \prec t' \Leftrightarrow \sigma(t) + l(t) \le \sigma(t')$

求使得 D 最小的可行调度.

条件说明:

任务在截止时间前完成

同时工作的台数不超过 m

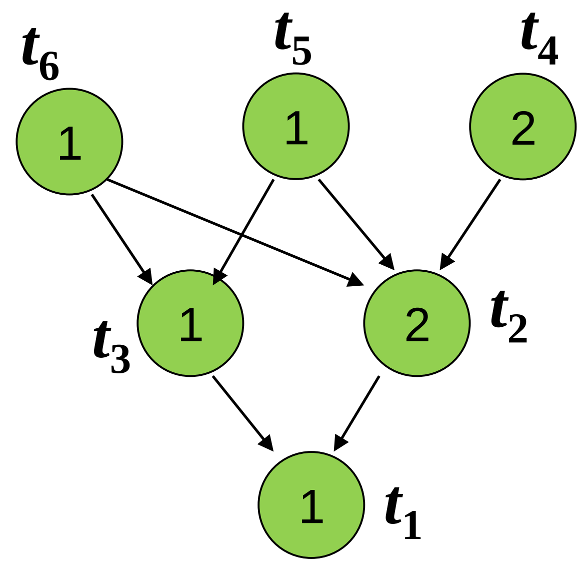
有偏序约束的任务必须按照约束先行

实例

任务集如图所示

m=2

求使得 D 最小的可行调度.



D=6

D=5

判定问题

实例：任务集T，m台机器，$\forall t \in T$，$l(t) \in Z^+$，T上的偏序 $\prec$，
截止时间 $D \in Z^+$.

问：是否存在小于等于 D 的可行调度？

子问题通过限制参数(机器数、工作时间、偏序)构成

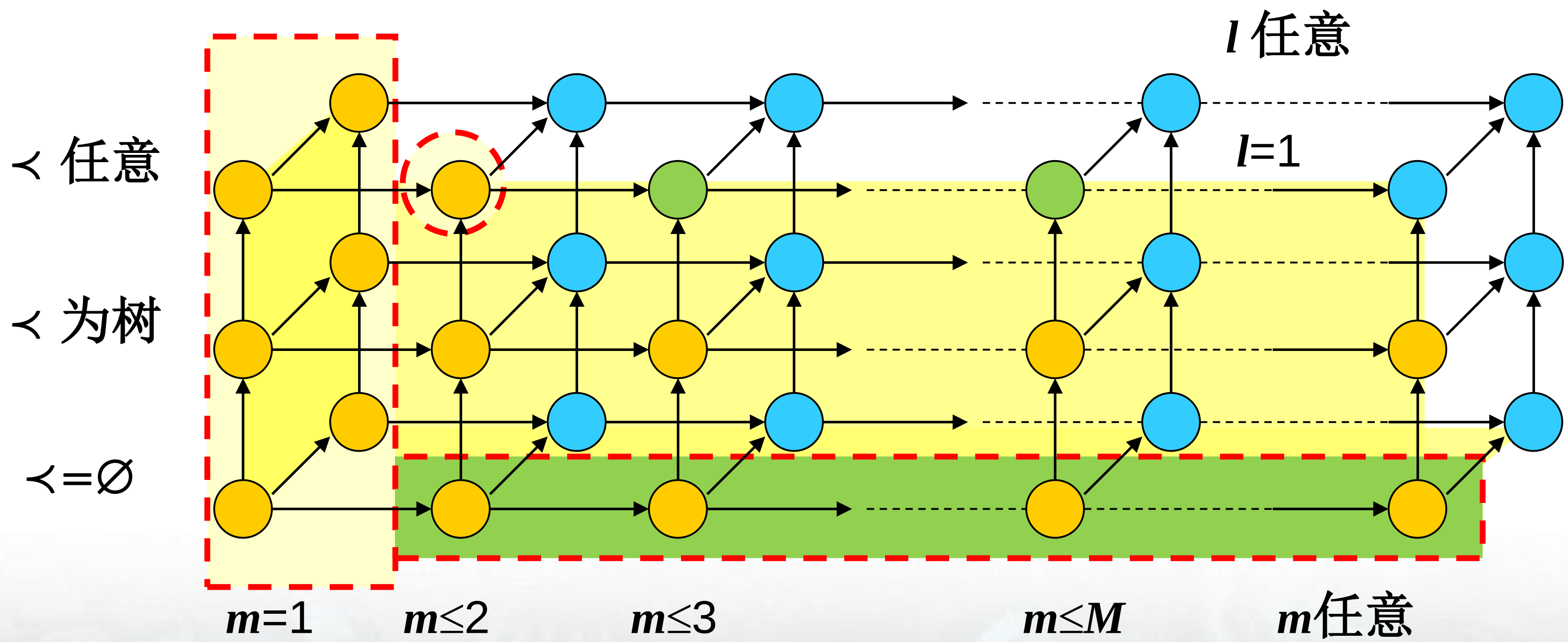
参数	限制		
偏序	$\varnothing$	树	任意
m大小	$m \le 1, 2, \ldots$, 某个常数		m 任意
l 大小	l 为常数		l 任意

调度问题的子问题结构

从上到下：偏序任意、树形偏序、无偏序约束

从左到右：处理器台数限制逐步放大

从前到后：各任务等长工作时间、任意工作时间



Turing归约

Turing归约

设π_1, π_2是搜索问题或优化问题，A 是利用解 π_2的假想子程序 s 解π_1的算法，且只要 s 是多项式时间的，A也是多项式时间的，则称算法 A 是从π_1到π_2的多项式时间的 **Turing归约**. 这时也称π_1 Turing归约到 π_2，记作 $\pi_1 \propto_T \pi_2$.

注意

- 多项式变换是特殊的Turing归约.
- **NP难**： 设π是搜索问题，如果存在 NP完全问题π' 使得 $\pi' \propto_T \pi$，则称π是**NP-hard**. 这意味着在多项式可计算的角度看，π至少和 NPC问题一样难. 许多NP完全问题对应的优化问题都是NP-hard. 反之也可以证明许多优化问题可以Turing归约到对应的判定问题, 称为**NP-easy**.
- NP-hard + NP-easy 称为 NP等价.

货郎问题的NP等价性

例1 货郎问题（TSO）是 NP等价的.

证：易证 TSO是 NP-hard. 下面证明 TSO 是NP-easy.

引入中间问题：货郎问题的延伸问题（TSE）

TSE

实例：有穷城市的集合 $C = \{ c_1, c_2, \ldots , c_m\}$

距离 $\forall c_i, c_j \in C, d(c_i, c_j) \in Z^+$,

长度限制 $B \in Z^+$,

部分旅行路线 $\vartheta = < c_{\pi(1)}, \ldots , c_{\pi(k)} >$

问：ϑ是否可以延伸成全长不超过 B 的全程旅行

$< c_{\pi(1)}, \ldots , c_{\pi(k)}, c_{\pi(k+1)}, \ldots , c_{\pi(m)} >$?

易证 TSE属于NP.

TSO到TSE的Turing归约

设 $s(C, d, \vartheta, B)$ 是解 TSE 的子程序，其中 C 为城市集，d 为距离函数，ϑ为部分旅行，B为长度限制.

下面构造解TSO的算法.

思路:

用二分法确定最短路旅行长度 B^*

旅行长度界于$m \to m \times d$, $d = \max\{ d(c_i, c_j)\}$

每次取中点值验证是否存在能延伸到此长度的旅行

根据最小长度值 B^*确定旅行路线

从c_1开始，依次检查$<c_1,c_2>,<c_1,c_3>\ldots$是否能延伸到 B^* 长度的旅行，选择第一个可延伸的顶点c_i.

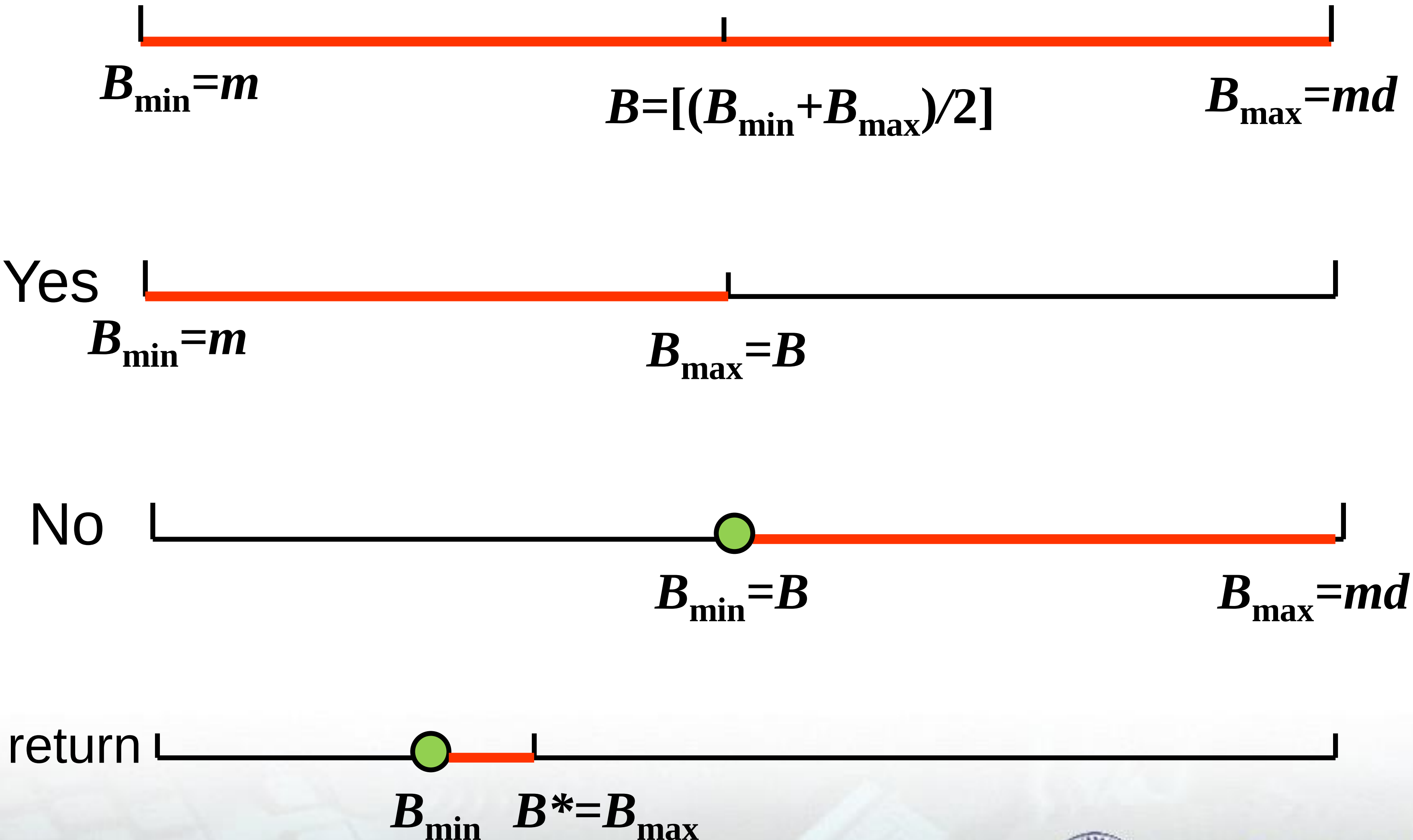
按照上面方法确定后面的其他顶点.

算法 MinLength

设 $s(C,d,\vartheta,B)$是解 TSE 的子程序，其中 C 为城市集，d为距离函数，ϑ为部分旅行，B为长度限制.

算法 Minlength （二分法确定最短旅行长度 B^*）

1. 令 $Bmin \leftarrow m$， $Bmax \leftarrow m \cdot \max\{ d(c_i,c_j): c_i, c_j \in C \}$
2. 若$Bmax-Bmin=1$, 则 $B^* \leftarrow Bmax$, 结束
3. $B \leftarrow [(Bmin+Bmax)/2]$
4. $s(C, d, <c_1>, B)$
5. 若回答” Yes” 则 $Bmax \leftarrow B$， 否则 $Bmin \leftarrow B$
6. 转2.


$B_{min}=m$
$B=[(B_{min}+B_{max})/2]$
$B_{max}=md$
Yes
$B_{min}=m$
$B_{max}=B$
No
$B_{min}=B$
$B_{max}=md$
return
B_{min}
$B^{*}=B_{max}$

算法 FindSolution

算法 FindSolution (找解)

1. $i \leftarrow 2$, $M \leftarrow \{2, 3, \ldots, m\}$
2. $j \leftarrow M$ 中的最小值
3. $\vartheta = \langle c_1, c_j \rangle$
4. $s(C, d, \vartheta, B^*)$
5. if 回答 “Yes”
6. then $i \leftarrow i+1$, $M \leftarrow M - \{j\}$
7. else
8. 　　从 ϑ 中去掉 c_j
9. 　　从 M 中选择大于 j 的最小值 k
10. 　　将 c_k 加入到 ϑ 的最后项
11. 如果 $i \le m$, 转4；否则停机

复杂度分析

至多m–2次调用s可以找到第2个城市，至多m–3次调用s可以找到第3个城市，…，至多1次调用s可以找到第m–1个城市。调用s的总次数至多为

$$(m-2)+(m-3)+\ldots+1=\frac{(m-1)(m-2)}{2}$$

为 m 的多项式，而 TSO 的实例规模为 m+log$B$max，所以是从 TSO 到 TSE 的 Turing归约.

而 TSE 是 NPC 问题，因为判定问题 TSP 是 TSE的子问题，相当 $\vartheta = \langle c_1 \rangle$的情况. 因此，TSO Turing归约到 NP 问题 TSE，从而证明了TSO 是 NP-easy，即TSO是NP等价的.

同样可以证明六个基本NPC问题对应的优化问题都是NP-等价的.

总结

- NPC证明
- NP完全理论的应用
  - 用NP完全性理论进行子问题分析
  - 推广到搜索问题与优化问题