# §1.6 可满足性问题与 Cook 定理

## 一．可满足性问题

1.例子：给出一个合取范式

$$(\bar{A} \vee B \vee C) \wedge (\bar{A} \vee C \vee D) \wedge (B \vee \bar{C}) \wedge (\bar{A} \vee C \vee D) \qquad (1.1)$$

A, B, C, D 为逻辑变量，取“真”和“假”两个值，用 T 和 F 表示。

$\bar{A}$为 A的补，$\bar{B}$为 B的补等等。

$(\bar{A} \vee B \vee C), (\bar{A} \vee C \vee D), (B \vee \bar{C}), (\bar{A} \vee C \vee D)$称为合取范式的子句。一个合取范式为 T 当且仅当每个子句取 T 值。上例中

$A = F, B = C = D = T$时，(1.1)式为 T。

判断(1.1)式是否可满足的问题相当于在每个子句中取一个文字组成集合，使得所选取的文字集合避免出现互补的一对。上例中，取

$\bar{A}, B, C, D$。

2. 可满足性问题：

设$L = \{A, B, \cdots, \bar{A}, \bar{B}, \cdots\}$

$C_1, C_2, \cdots, C_k$是 L的有限子集，称为子句。每个$C_i$中不出现 L 中互补的一对(即 $x \in C_i$, 则$\bar{x} \notin C_i$), $i = 1, 2, \cdots, k$。所谓可满足性问题，是确定是否存在一个集合$S \subseteq L$，满足以下两个要求：

$$\forall x \in S, \bar{x} \notin S;$$

$S \cap C_i \neq \emptyset, i = 1,2, \cdots, k$。

可满足性问题用 SAT 表示。

## 二．Cook 定理

### 定理 1.5(Cook 定理)：若$L \in NP$, 则$L \propto SAT$。

证：由于 SAT 问题属于 NP 类，故只要证任一属于 NP 类的问题 L 在多项式时间内可转换为 SAT 问题。

设已知一非确定型图灵机 M，在 T(n)多项式界内对输入 x 进行识别，只要找一有多项式界的方法把输入符号串 x 转换为一组子句 f(x)，而且 M 接受 x 当且仅当 f(x)是可满足的。

设 NDTM 状态集为$Q = \{q_1, q_2, \cdots, q_l\}$，其中，$q_1$为初始状态，$q_2$为$q_Y$，$q_3$为$q_N$。NDTM 的带字母表为$\Gamma = \{a_1, a_2, \cdots, a_m\}$，其中$a_1$为空格符#。

*NDTM 在不超过 N 步内接受 x，每一步，读写头最多左移或右移一格。故读写头的活动范围不超过以初始位置为中心的左、右各 N 格，共 2N+1 格。

*每行的 2N+1 个方格上的符号串，机器的当时状态以及读写头的位置，这些信息完全描绘了机器的瞬间图像。共有 N 行，描述了 N 个时刻的机器的瞬间图像。

t 从 1 到 N，每个时刻机器的瞬间图像是 x 被接受的全过程。

引入一$N \times (2N + 1)$的方格，第 t 行记录下 t 时刻带上的符号，第 t 行第 i 列的方格用(t, i)表示。令

$$A(t,i,j)=\begin{cases}T, & 若(t,i)的符号为a_j\\ F, & 其它\end{cases}$$

$$H(t,i)=\begin{cases}T, & 若\ t\ 时刻读写头的位置为第\ i\ 格\\ F, & 其它\end{cases}$$

$$S(t,k)=\begin{cases}T, & 若\ t\ 时刻机器的状态为q_k\\ F, & 其它\end{cases}$$

其中：$1\le t\le N, 1\le i\le 2N+1, 1\le j\le m, 1\le k\le l$。

初始输入符号串为$a_{k_1}a_{k_2}\cdots a_{k_n}$，带上符号为

$$\underbrace{\#\#\cdots\#}_{N}a_{k_1}a_{k_2}\cdots a_{k_n}\underbrace{\#\#\cdots\#}_{N+1-n}$$

可用下列子句叙述如下：

$$\bigwedge_{i=1}^{N}A(1,i,1)\bigwedge_{i=N+1}^{N+n}A(1,i,k_{i-N})\bigwedge_{i=N+n+1}^{2N+1}A(1,i,1) \qquad (1.2)$$

初始状态为$q_1$时，读头位于第 N+1 格，对应子句

$$S(1,1)\wedge H(1,N+1) \qquad (1.3)$$

(1)对于任一 t, $0<t\le N$，机器有一种状态，而且只有一种状态，对应子句为

$$\bigwedge_{t=1}^{N}\left(\left(\bigvee_{j=1}^{l}S(t,j)\right)\wedge\left(\bigwedge_{1\le j_1<j_2\le l}\overline{\left(S(t,j_1)\wedge S(t,j_2)\right)}\right)\right) \qquad (1.4)$$

由德·摩根律，可化为

$$\bigwedge_{t=1}^{N}\left(\left(\bigvee_{j=1}^{l}S(t,j)\right)\bigwedge_{j_1<j_2}\left(\overline{S(t,j_1)}\vee\overline{S(t,j_2)}\right)\right) \qquad (1.5)$$

(2)对于$1\le t\le N, 1\le i\le 2N+1$的任一方格，$(t,i)$有一个字符且仅有一个字符，和(3)相似，对应子句为

$$\bigwedge_{t=1}^{N}\bigwedge_{i=1}^{2N+1}\left(\left(\bigvee_{j=1}^{m}A(t,i,j)\right)\bigwedge_{1\le j_1<j_2\le m}\left(\overline{A(t,i,j_1)}\vee\overline{A(t,i,j_2)}\right)\right) \qquad (1.6)$$

(3)对于$1\le t\le N$的任何时刻，读写头在一个方格上且仅在一个方格上，对应子句为

$$\bigwedge_{t=1}^{N}\left(\left(\bigvee_{i=1}^{2N+1}H(t,i)\right)\bigwedge_{1\le i_1<i_2\le 2N+1}\left(\overline{H(t,i_1)}\vee\overline{H(t,i_2)}\right)\right) \qquad (1.7)$$

(4)从 t 时刻到$t+1$时刻，只有 t 时刻读写头所在的方格，它的符号才有所改变。对应子句为

$$A(t,i,j)\leftrightarrow\left(A(t+1,i,j)\vee H(t,i)\right) \qquad (1.8)$$

上式易于化为合取式。

(5)对于 t 时刻机器的状态$q\in Q\backslash\{q_Y,q_N\}$，则$t+1$时刻机器状态，读写头所在的位置，以及原来 t 时刻读写头所在位置的符号的改变，这三者均服从机器的控制功能。对应有

$$\bigwedge_{t=1}^{N}\bigwedge_{i=1}^{2N+1}\bigwedge_{k=1}^{l}\bigwedge_{j=1}^{m}\left(\overline{A(t,i,j)}\vee\overline{H(t,i)}\vee\overline{S(t,k)}\vee\right.$$

$$(\bigvee_{c\in f}(A(t+1,i,j_c)\wedge S(t+1,k_c)\wedge H(t+1,i_c))))\qquad(1.9)$$

后面括号中的$\vee_{c\in f}$指的是根据当前的状态$q_k$和当前读头读到的带符号$a_j$对转移函数的所有转移动作进行的。

(1.9)式有两种可能：一是 t 时刻机器状态与

$$A(t,i,j)\wedge H(t,i)\wedge S(t,k)\text{不符，即}$$
$$\overline{A(t,i,j)}\vee\overline{H(t,i)}\vee\overline{S(t,k)}=T；$$

二是 A,H,S 满足 t 时刻的状态，这时

$$\overline{A(t,i,j)}\vee\overline{H(t,i)}\vee\overline{S(t,k)}=F\text{，而}$$
$$\bigvee_{c\in f}(A(t+1,i,j_c)\wedge S(t+1,k_c)\wedge H(t+1,i_c))\qquad(1.10)$$

描述了$t+1$时刻机器的可能状态。

(6)NDTM 总有一个时刻处于停机状态

$$\bigvee_{t=1}^{N}S(t,2)\qquad(1.11)$$

表达式(1.9)可满足，当且仅当 NDTM 存在一个接受 x 的动作序列。

对以上所有子句作合取，经逻辑变换，可化为合取范式。该逻辑表达式可满足当且仅当 NDTM 接受 x。而把输入串 x 化为这一公式的过程的时间复杂度有多项式的上界。从而所有 NP 类问题可化为可满足性问题。证毕。

# §1.7 其它 NP 完全问题及归约

*上一节我们证明了可满足性问题是 NP 完全问题，本节我们根据定理 1.4，给出几个其它的 NP 完全问题，并介绍证明 NP 完全问题的归约技术。

## 一. 团问题：

若完全图$G_1$是图 G 的子图，则称$G_1$是图 G 的团。

团问题：已知图 G 和整数 k，试判定图 G 是否存在 k 个顶点的团。

### 定理 1.6：团问题属于 NPC。

证：可满足性问题可化为团问题。做法如下：令图的顶点分别对应于表达式中出现的文字。

两个顶点之间有边相连当且仅当

两个顶点对应的文字不出现在同一子句中；

两个顶点对应的文字不相互补。

故存在 k 个顶点的团(k 为合取范式中子句的个数)的充分必要条件

是子句集合是可满足的。显然，团的问题属于 NP 类，故团的问题是 NP 完全的。证毕。

## 二．3SAT

3SAT 问题：每个子句恰有 3 个文字的可满足性问题。

## 定理 1.7：3SAT 问题属于 NPC。

证：3SAT 问题显然属于 NP。对合取范式的每个子句

$$x_1 \vee x_2 \vee \cdots \vee x_k \tag{1.12}$$

代以 $(x_1 \vee x_2 \vee y_1) \wedge (\overline{y_1} \vee x_3 \vee \cdots \vee x_k)$ (1.13)

继续以上过程，直到每个子句都不超过 3 个文字为止，其中$y_i$是不出现在原合取范式中的文字。

下面证明(1.12)是可满足的充要条件是(1.13)是可满足的。

设$x_1 \vee x_2 \vee \cdots \vee x_k = T$，则存在$x_i$，使$x_i = T$。若$(x_1 = T) \vee (x_2 = T)$，则令$y_1 = F$，则有$(x_1 \vee x_2 \vee y_1) \wedge (\overline{y_1} \vee x_3 \vee \cdots \vee x_k) = T$。

若$x_3, \cdots, x_k$中有一个为 T，则令$y_1 = T$，上式依然成立。

故如果(1.12)可满足，则(1.13)也可满足。

反之，若(1.13)可满足，由于$y_1 \wedge \overline{y_1} = F$，故$x_1, x_2, \cdots, x_k$中至少有一个为 T。即有$x_1 \vee x_2 \vee \cdots \vee x_k = T$。

故如果(1.13)式可满足，则(1.12)式也可满足。

若 k = 1 或 2，则有以下两个等值式：

(1)$(x_1 \vee y_1 \vee y_2) \wedge (x_1 \vee \overline{y_1} \vee y_2) \wedge (x_1 \vee y_1 \vee \overline{y_2}) \wedge (x_1 \vee \overline{y_1} \vee \overline{y_2}) = x_1$

(2)$(x_1 \vee x_2 \vee y_1) \wedge (x_1 \vee x_2 \vee \overline{y_1}) = x_1 \vee x_2$

用上述等值式左边代替右边，可将 SAT 问题化为 3SAT 问题。以上将 SAT 问题转化为 3SAT 问题的过程所需时间显然为多项式时间。证毕。

## 三．独立集、顶点覆盖

设$G = (V, E)$是一个图，$S \subseteq V(G)$。若集合 S 中任意两个顶点都不相邻，则称 S 为图 G 的独立集。若$C \subseteq V$，且 G 的任一边至少有一个端点属于 C，则称 C 为图 G 的顶点覆盖。

独立集问题：已知图$G = (V, E), k \leq |V|$，问是否存在独立集$S \subseteq V$，使得$|S| \geq k$?

顶点覆盖问题：已知图$G = (V, E)$，正整数$k \leq |V|$，问是否存在顶点覆盖$C \subseteq V$，使得$|C| \leq k$?

## 定理 1.8：独立集问题属于 NPC。

证：因独立集问题属于 NP 类，只要证团问题可以多项式归约于独立集问题即可。

已知图$G = (V, E)$及 k，作图 G 的补图$\overline{G} = (V, \overline{E})$使得$\forall u, v \in V$，$(u, v) \in \overline{E}$当且仅当$(u, v) \notin E$。$S \subseteq V$ 是 G 的团的充要条件是：对于图$\overline{G}$，S 是独立集。证毕。

## 定理 1.9：顶点覆盖问题属于 NPC。

证：顶点覆盖问题显然属于 NP 类，只要证独立集问题可以多项式归约为顶点覆盖问题。

已知图$G=(V,E)$及整数 k，令$l=|V|-k$。

如果有一独立集 S，使得$|S|\geq k$, 显然$V\backslash S$是图 G 的顶点覆盖，$V\backslash S$的顶点数为$|V|-|S|\leq |V|-k=l$。

反之，如果 C 是图 G 的顶点覆盖集，$|C|\leq l$，则$V\backslash C$是独立集，且$V\backslash C$的顶点数为$|V|-|C|\geq |V|-l=k$。证毕。

## 四．哈密顿道路问题

已知有向图$G=(V,E)$及两个顶点$u,v$，问是否存在以 u 为始点，以 v 为终点的有向哈密顿道路？这个问题称为有向哈密顿道路问题。

## 定理 1.10：有向哈密顿道路问题属于 NPC。

证：显然有向哈密顿道路问题属于 NP 类。只要证顶点覆盖问题可以多项式归约为有向哈密顿道路问题。

已知无向图$G=(V,E)$, $k$是正整数，与$v\in V$点相关联的边(即以v 点为一端点的边)的数目设为$d_v$，这些边设为$e_1^v,e_2^v,\cdots,e_{d_v}^v$。现构造一有向图$G'=(V',E')$如下：

顶点集 V’有以下两个组成部分：

(a)新增加 k+1 个顶点：$a_0,a_1,\cdots,a_k$

(b)对于任一顶点$v\in V$，对应有$2d_v$个顶点：

$v_1^{(1)},v_1^{(2)},v_2^{(1)},v_2^{(2)},\cdots,v_{d_v}^{(1)},v_{d_v}^{(2)}$

有向边集 E’的组成部分有：

(a)$\left(a_i,v_1^{(1)}\right),0\leq i\leq k,v\in V$

(b)$\left(v_{d_v}^{(2)},a_i\right),0\leq i\leq k,v\in V$

(c)对于图 G 中相邻两个顶点 u 和 v，设$e_i^u=e_j^v$,则有边：

$\left(u_i^{(1)},v_j^{(1)}\right),\left(u_i^{(2)},v_j^{(2)}\right),\left(v_j^{(1)},u_i^{(1)}\right),(v_j^{(2)},u_i^{(2)})$

(d)对于每个$v\in V$，对应有边：

$\left(v_i^{(1)},v_i^{(2)}\right),1\leq i\leq d_v$,和 $\left(v_i^{(2)},v_{i+1}^{(1)}\right)$，$1\leq i<d_v$

故对应每个顶点$v\in V$，存在一个道路

$H_v: v_1^{(1)}\to v_1^{(2)}\to v_2^{(1)}\to v_2^{(2)}\to\cdots\to v_{d_v}^{(1)}\to v_{d_v}^{(2)}$

对应于边$e_i^u=e_j^v=(u,v)$的两端点 u 和 v，道路$H_u$和$H_v$之间存在边

$\left(u_i^{(1)}, v_j^{(1)}\right), \left(u_i^{(2)}, v_j^{(2)}\right)$

$\left(v_j^{(1)}, u_i^{(1)}\right), (v_j^{(2)}, u_i^{(2)})$

即边$(u, v)$对应一由 4 个顶点$u_i^{(1)}, u_i^{(2)}, v_j^{(1)}, v_j^{(2)}$组成的子图：(见图 1.6)

若有哈密顿道路进入$u_i^{(1)}$点，则必须从$u_i^{(2)}$点退出该子图。如果从$v_j^{(2)}$退出，则或者无法经过$u_i^{(2)}$，或者无法经过$v_j^{(1)}$。由$u_i^{(1)}$进入由$u_i^{(2)}$退出有两种可能，一是经过这 4 个顶点，另一个是走$u_i^{(1)} \to u_i^{(2)}$退出这个子图。

从图 1.6 可见，若图 G 中有$(u, v) \in E$，则 G’图中存在从$a_i$到$a_j$的道路：$a_i u_1^{(1)} u_1^{(2)} \cdots u_i^{(1)} v_j^{(1)} v_j^{(2)} u_i^{(2)} \cdots u_{d_u}^{(1)} u_{d_u}^{(2)} a_j$ 和 $a_i v_1^{(1)} v_1^{(2)} \cdots v_j^{(1)} u_i^{(1)} u_i^{(2)} v_j^{(2)} \cdots v_{d_v}^{(1)} v_{d_v}^{(2)} a_j$

图 G’存在有向哈密顿道路的充要条件是图 G 有 k 个顶点的顶点覆盖，设

$$C = \{u, v, \cdots, w\} \quad (|C| = k)$$

是图 G 的顶点覆盖，可构造图 G’的从$a_0$到$a_k$的哈密顿道路，步骤如下：

(a)构造一从$a_0$到$a_k$的哈密顿道路 P (见图 1.7)

(b)对于边$e = (u, z)$设点 u 属于顶点覆盖 C，但 z 不属于 C，则道路 P 中$H_u$的一段为绕道通过$z_j^{(1)}$和$z_j^{(2)}$两点一段所取代。其中$(u, z) = e_j^z = e_i^u$，即$H_u$中$u_i^{(1)} \to u_i^{(2)}$的一段改道为$u_i^{(1)} \to z_j^{(1)} \to z_j^{(2)} \to u_i^{(2)}$，从而吸收了$z_j^{(1)}$和$z_j^{(2)}$点。因 C 是顶点覆盖，故可以吸收 G’的所有顶点，正如 u 点和 z 点通过$e_j^z$边相邻，从而吸收$z_j^{(1)}$和$z_j^{(2)}$一样。

反之，可以从图 G’构造 G 的顶点覆盖集 S。由于哈密顿道路过所有的顶点，故过所有的$a_i$ $(0 \le i \le k)$。由于 E’中不存在从$a_i$到$a_j$的边，故哈密顿道路可以分解为若干段，每一段从某个$a_i$到另一个$a_j$，中间不存在其它的$a_k$。对于每一条从$a_i$到$a_j$的道路对应一顶点$v \in V$，使得道路的端点为$v_1^{(1)}$，这样的 k 个顶点便是图 G 的顶点覆盖集。 证毕。