



NP完全问题

东南大学计算机学院 方效林

本章内容

- NP问题
- NP完全问题
- NP完全问题证明方法

NP问题

- P问题 (Polynomial)
 - 多项式时间可解的问题
 - 多项式时间: $O(n)$, $O(\log n)$, $O(n^c)$
 - 非多项式时间: $O(2^n)$, $O(n!)$
- NP问题 (Non-Deterministic Polynomial)
 - 多项式时间可验证一个解的问题

所有NP问题都是判定问题，回答yes或no

NP问题

■ 3CNF-SAT问题

- 给定n个布尔变元，子句是3个文字(变元或变元的非) 的析取操作

➤ $(\bar{x}_1 \vee x_2 \vee x_3)$

- 给定m个子句的合取操作

➤ $(\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_7 \vee x_9 \vee \bar{x}_{10}) \wedge \cdots \wedge (x_5 \vee \bar{x}_6 \vee x_4)$

- 问该合取范式的值可否为真？

- 给定n个变元的一个取值 x_1, x_2, \dots, x_n , $x_i=1$ 或 0 , 判断能否使合取范式为真, 时间复杂度 $O(n)$, 即多项式时间可判定

$$\begin{array}{ccc} \text{C1} & \text{C2} & \text{C3} \\ (\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \bar{x}_2 \vee \bar{x}_3) \end{array}$$

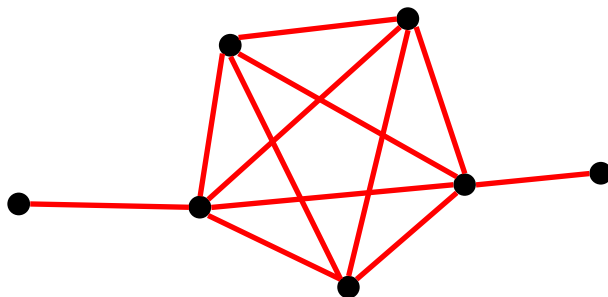
$x_1=0, x_2=1, x_3=0$, 则 $c_1 \wedge c_2 \wedge c_3=1$

NP问题

■ K团问题

□ 给定一个图G和常数k，G有没有k团？

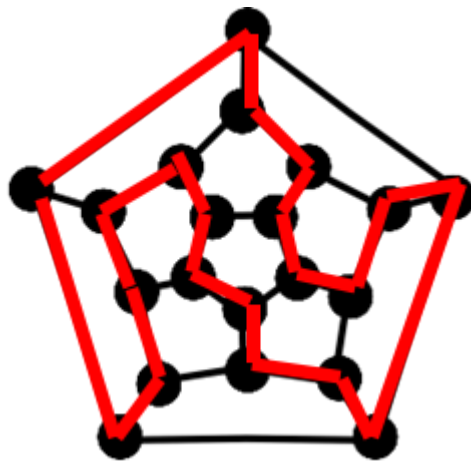
- 一个图G的k团是G的 k 个顶点的集合，使得这个集合中每对顶点之间都有边
- 任取 k 个顶点，判断是否两两之间有边，时间复杂度为 $O(n^2)$ ，即多项式时间可判定这k个点是否是一个团



NP问题

■ Hamilton路径问题

- 给定无向图G，问是否存在一条路径经过G中所有顶点一次且只经过一次？
 - 给定n个顶点的一个排列 $v_{i_1}, v_{i_2}, \dots, v_{i_n}$ ，依次判断 v_{i_k} 和 $v_{i_{k+1}}$ 之间是否存在边，时间复杂度 $O(n)$ ，即多项式时间可判定这个排列是否是一条路径



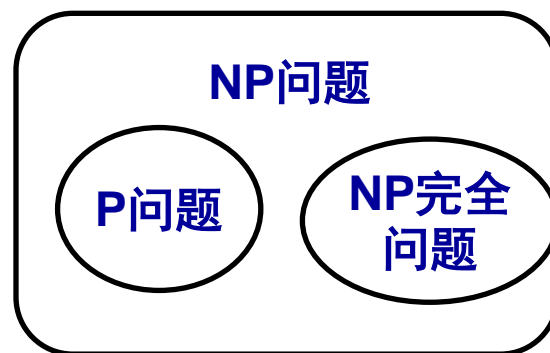
NP问题

- 最优化问题可与一个判定问题对应
 - 最优化问题：给定一个图 G ，寻找 最大团
 - 判定问题：给定一个图 G 和常数 k ， G 有没有 k 团？
- 如何对应？
 - 若判定问题多项式时间可解，则可采用二分搜索策略找到最优的 k
 - 反之，若已求解最优化问题，就已求解判定问题

NP完全问题

- 问题A是一个NP完全问题，则有
 1. $A \in \text{NP}$ 问题 (多项式时间可验证解是否正确)
 2. 且 $\forall B \in \text{NP}, B \leq_p A$
 - 任意NP问题B都可多项式时间归约到A
 - 构造一个多项式时间对应关系

P≠NP否还不知道，
现在普遍认为**P≠NP**



NP完全问题证明

- 证明问题A是NP完全问题分两步：
 1. A是一个NP问题 (多项式时间可验证解是否正确)
 2. 从一个已知的NP完全问题A', 多项式时间规约到A的一个实例A'', 证明A'有解当且仅当A''有解

第一个NP完全问题

■ CNF-SAT问题(可满足问题)

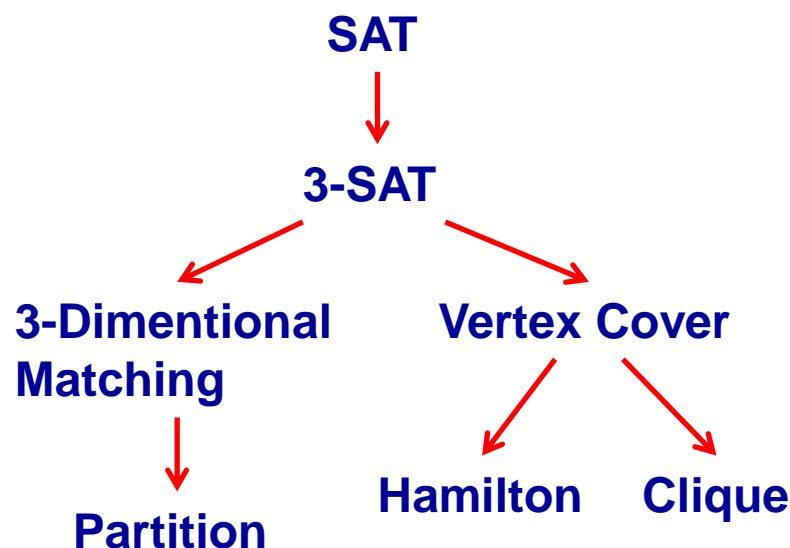
- 给定n个布尔变量 x_1, x_2, \dots, x_n , 和m个子句, 子句是1个或多个个文字(变量或变量的非)的析取操作
 - 如: $(\bar{x}_1 \vee x_2)$, $(\bar{x}_3 \vee x_1 \vee x_5)$
- 给定这m个子句的合取范式
 - $(\bar{x}_1 \vee x_2) \wedge (x_7 \vee x_9 \vee \bar{x}_{10}) \wedge \dots \wedge x_5$
- 问该合取范式的值可否为真?

第一个NP完全问题

■ Cook定理

- 可满足问题是NP完全问题 ($\text{CNF-SAT} \in \text{NP-C}$)

其证明过程用到非确定图灵机理论



3CNF-SAT问题是NP完全问题

■ 3CNF-SAT问题

- 给定n个布尔变元和m个子句，子句是3个文字(变元或变元的非)的析取操作

➤ $(\bar{x}_1 \vee x_2 \vee x_3)$

- 问这m个子句的合取范式的值可否为真？

➤ $(\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_7 \vee x_9 \vee \bar{x}_{10}) \wedge \cdots \wedge (x_5 \vee \bar{x}_6 \vee x_4)$

$\text{C1} \qquad \qquad \qquad \text{C2} \qquad \qquad \qquad \text{Cm}$

3CNF-SAT问题是NP完全问题

■ 证明

- 显然3CNF-SAT是NP问题(多项式时间可判定)
- 将CNF-SAT归约(多项式时间转换)到3CNF-SAT
 - CNF-SAT中各子句中文字个数可能为1, 2, 3或 ≥ 3 个
 - $|C_i| = 1$ 时, 设 $C_i = x_{i1}$ 。增加两个布尔变量 $U_i = \{y_{i1}, y_{i2}\}$, 子句变换为 $C'_i = (x_{i1} \vee y_{i1} \vee y_{i2}) \wedge (x_{i1} \vee y_{i1} \vee \bar{y}_{i2}) \wedge (x_{i1} \vee \bar{y}_{i1} \vee y_{i2}) \wedge (x_{i1} \vee \bar{y}_{i1} \vee \bar{y}_{i2})$
 - $|C_i| = 2$ 时, 设 $C_i = x_{i1} \vee x_{i2}$ 。增加一个布尔变量 $U_i = \{y_{i1}\}$, 子句变换为 $C'_i = (x_{i1} \vee x_{i2} \vee y_{i1}) \wedge (x_{i1} \vee x_{i2} \vee \bar{y}_{i1})$
 - $|C_i| = 3$ 时, 设 $C_i = x_{i1} \vee x_{i2} \vee x_{i3}$ 。 $C'_i = C_i$
 - $|C_i| > 3$ 时, 设 $C_i = x_{i1} \vee x_{i2} \vee \cdots \vee x_{i|C_i|}$ 。增加 $|C_i| - 3$ 个布尔变量 $U_i = \{y_{i1}, y_{i2}, \dots, y_{i(|C_i|-3)}\}$, 子句变换为 $C'_i = (x_{i1} \vee x_{i2} \vee y_{i1}) \wedge (\bar{y}_{i1} \vee x_{i3} \vee y_{i2}) \wedge \cdots \wedge (\bar{y}_{ik} \vee x_{i(k+2)} \vee y_{i(k+1)}) \wedge \cdots \wedge (\bar{y}_{i(|C_i|-3)} \vee x_{i(|C_i|-1)} \vee x_{i|C_i|})$

3CNF-SAT问题是NP完全问题

■ 证明

- 显然3CNF-SAT是NP问题(多项式时间可判定)
- 将CNF-SAT归约(多项式时间转换)到某3CNF-SAT实例
- 证明CNF-SAT可满足当且仅当构造的3CNF-SAT可满足

➤ 必要性： CNF-SAT可满足(每个子句都可满足)， 则3CNF-SAT可满足。

- $|C_i| = 1$ 时可满足, 即 $C_i = x_{i1} = 1$, 得 $C'_i = (x_{i1} \vee y_{i1} \vee y_{i2}) \wedge (x_{i1} \vee y_{i1} \vee \bar{y}_{i2}) \wedge (x_{i1} \vee \bar{y}_{i1} \vee y_{i2}) \wedge (x_{i1} \vee \bar{y}_{i1} \vee \bar{y}_{i2}) = 1$
- $|C_i| = 2$ 时可满足, 即 $C_i = x_{i1} \vee x_{i2} = 1$, 得 $C'_i = (x_{i1} \vee x_{i2} \vee y_{i1}) \wedge (x_{i1} \vee x_{i2} \vee \bar{y}_{i1}) = 1$
- $|C_i| = 3$ 时可满足, 即 $C_i = x_{i1} \vee x_{i2} \vee x_{i3} = 1$, 得 $C'_i = C_i = 1$
- $|C_i| > 3$ 时可满足, 即 $C_i = x_{i1} \vee x_{i2} \vee \cdots \vee x_{i|C_i|} = 1$, 得 $C'_i = (x_{i1} \vee x_{i2} \vee y_{i1}) \wedge (\bar{y}_{i1} \vee x_{i3} \vee y_{i2}) \wedge \cdots \wedge (\bar{y}_{ik} \vee x_{i(k+2)} \vee y_{i(k+1)}) \wedge \cdots \wedge (\bar{y}_{i(|C_i|-3)} \vee x_{i(|C_i|-1)} \vee x_{i|C_i|}) = 1$

3CNF-SAT问题是NP完全问题

■ 证明

- 显然3CNF-SAT是NP问题(多项式时间可判定)
- 将CNF-SAT归约(多项式时间转换)到某3CNF-SAT实例
- 证明CNF-SAT可满足当且仅当构造的3CNF-SAT可满足

➤ 充分性： 3CNF-SAT可满足， 则CNF-SAT可满足。

1. $C'_i = (x_{i1} \vee y_{i1} \vee y_{i2}) \wedge (x_{i1} \vee y_{i1} \vee \bar{y}_{i2}) \wedge (x_{i1} \vee \bar{y}_{i1} \vee y_{i2}) \wedge (x_{i1} \vee \bar{y}_{i1} \vee \bar{y}_{i2}) = 1$ ， 得 $C_i = x_{i1} = 1$
2. $C'_i = (x_{i1} \vee x_{i2} \vee y_{i1}) \wedge (x_{i1} \vee x_{i2} \vee \bar{y}_{i1}) = 1$ ， 得 $C_i = x_{i1} \vee x_{i2} = 1$
3. $C'_i = C_i = x_{i1} \vee x_{i2} \vee x_{i3} = 1$
4. $C'_i = (x_{i1} \vee x_{i2} \vee y_{i1}) \wedge (\bar{y}_{i1} \vee x_{i3} \vee y_{i2}) \wedge \cdots \wedge (\bar{y}_{ik} \vee x_{i(k+2)} \vee y_{i(k+1)}) \wedge \cdots \wedge (\bar{y}_{i(|C_i|-3)} \vee x_{i(|C_i|-1)} \vee x_{i|C_i|}) = 1$ ， 得 $C_i = x_{i1} \vee x_{i2} \vee \cdots \vee x_{i|C_i|} = 1$

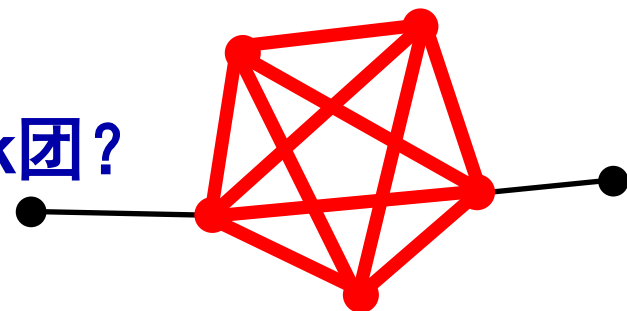
NP完全问题证明

- 假设已知3CNF-SAT问题是NP完全问题
- 现在要证
 - k团问题
 - 顶点覆盖问题
 - 子集和问题

k团问题是NP完全问题

■ k团问题

- 给定一个图G和常数k，G有没有k团？



■ 3CNF-SAT问题

- 给定n个布尔变元和m个子句，子句是3个文字(变元或变元的非)的析取操作

➤ $(\bar{x}_1 \vee x_2 \vee x_3)$

- 问这m个子句的合取范式的值可否为真？

➤ $(\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_7 \vee x_9 \vee \bar{x}_{10}) \wedge \cdots \wedge (x_5 \vee \bar{x}_6 \vee x_4)$

k团问题是NP完全问题

■ 证明

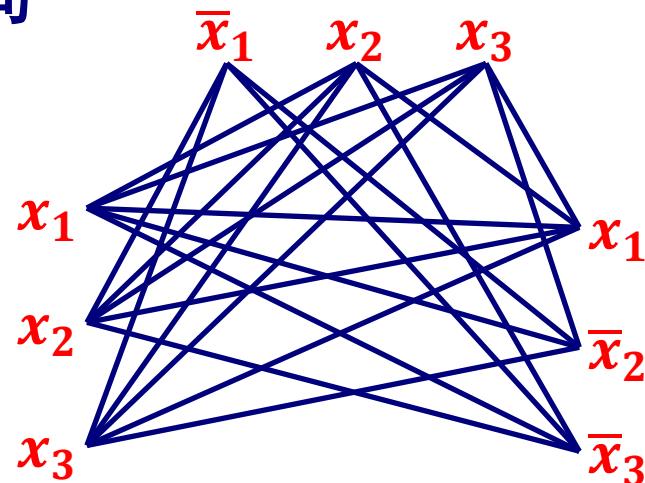
- 首先，k团问题是NP问题(多项式时间可验证)
- 规约方法：3CNF-SAT表达式中每一文字对应k团问题一个顶点，若两顶点同时满足下列两条则存在边：
 - 两个顶点中的文字不属于同一子句
 - 两个顶点中的文字不是互补的
 - 令 $k=m$

C1

C2

C3

$$(\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \bar{x}_2 \vee \bar{x}_3)$$



k团问题是NP完全问题

- 往证有k个子句的3CNF-SAT表达式A是可满足的当且仅当构造出来的无向图G中有一个k团
- 必要性证明：设表达式A是可满足的，必有 $A=1$ ，由于合取，必有 $c_1=c_2=\dots=c_k=1$ ，每个子句 c_i 中的文字至少有一个取值为1， c_i 中取一个值为1的文字所对应的顶点，可得一个由k个顶点所构成的子集V。V中任取两点，对应的文字来自不同的子句，两个文字必定不是互补的。
 - \therefore 两个顶点之间必定有边。
 - $\therefore V$ 是k团

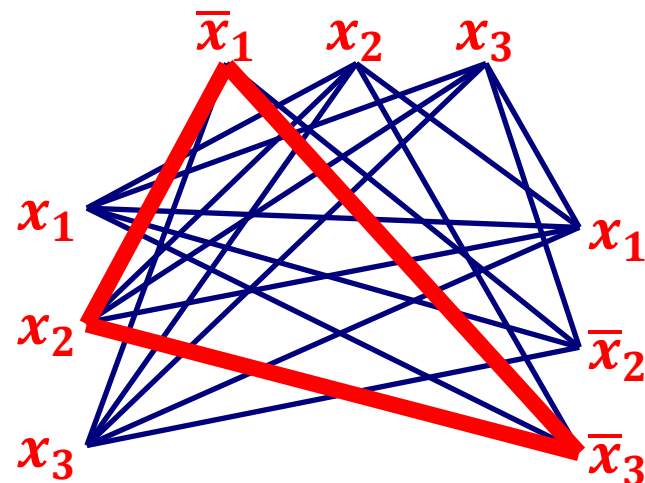
$c_1 \wedge c_2 \wedge c_3 = 1$ ，取 $x_1=0$ ， $x_2=1$ ， $x_3=0$

C1

C2

C3

$(\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \bar{x}_2 \vee \bar{x}_3)$



k团问题是NP完全问题

- 充分性证明：设V是构造出来的无向图中的k团，则V中任意两点间有边。由于同一子句中的顶点无边相连，故k个点恰好来自k个不同的子句。现对这k个顶点文字进行赋值：
 - 若顶点文字为 x_i ，则 x_i 赋值1，其所在子句的值为1；
 - 若顶点文字为 \bar{x}_i ，则 x_i 赋值0，其所在子句的值为1
- 由于 x_i 与 \bar{x}_i 无边相连，这k个顶点文字不会同时出现 x_i 与 \bar{x}_i ， \therefore 上述的赋值方法可使表达式为1

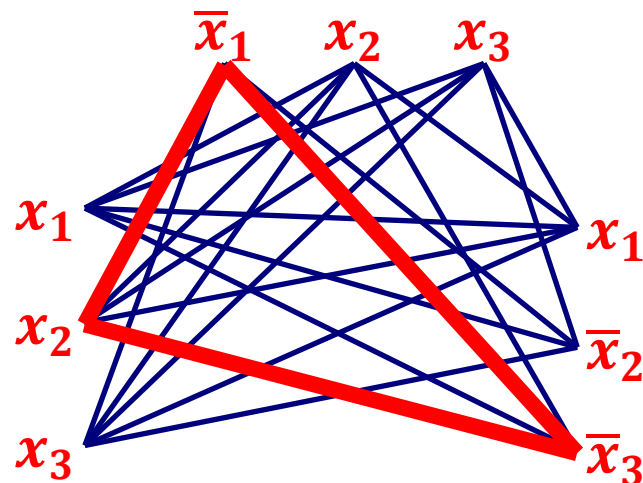
$x_1=0, x_2=1, x_3=0$ ，则 $c_1 \wedge c_2 \wedge c_3=1$

C1

C2

C3

$(\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \bar{x}_2 \vee \bar{x}_3)$



顶点覆盖问题是NP完全问题

■ 顶点覆盖问题

□ 无向图G中是否存在顶点数为k的顶点覆盖？

➤ G 中大小为 k 的顶点集合，使得 G 中任一条边的两个顶点至少有一个在此集合中

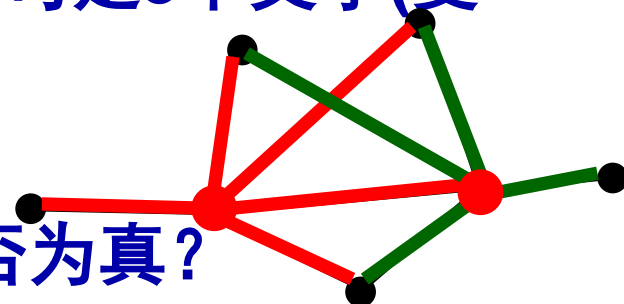
■ 3CNF-SAT问题

□ 给定n个布尔变元和m个子句，子句是3个文字(变元或变元的非)的析取操作

➤ $(\bar{x}_1 \vee x_2 \vee x_3)$

□ 问这m个子句的合取范式的值可否为真？

➤ $(\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_7 \vee x_9 \vee \bar{x}_{10}) \wedge \cdots \wedge (x_5 \vee \bar{x}_6 \vee x_4)$

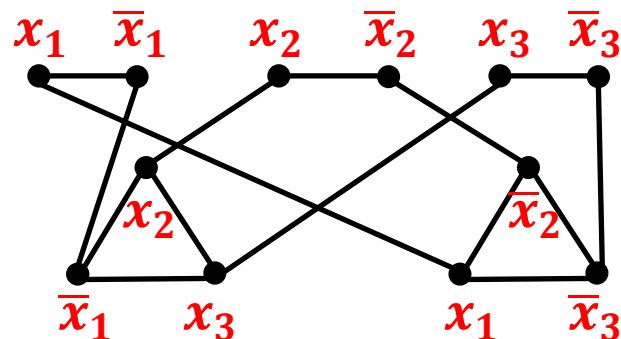


顶点覆盖问题是NP完全问题

■ 证明

- 首先, 顶点覆盖问题是NP问题(多项式时间可验证)
- 规约方法 :
 - n 个变元对应 $2n$ 个顶点 (x_i 和 \bar{x}_i , 共 $2n$ 个), x_i 和 \bar{x}_i 间有边
 - m 个子句对应 $3m$ 个顶点 ($3m$ 文字), 子句内文字两两有边
 - 子句内的文字与变元对应的顶点间有边
 - 令 $k=n+2m$

$$\begin{array}{cc} \text{C1} & \text{C2} \\ (\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \bar{x}_2 \vee \bar{x}_3) & \\ k=n+2m=3+2*2=7 & \end{array}$$



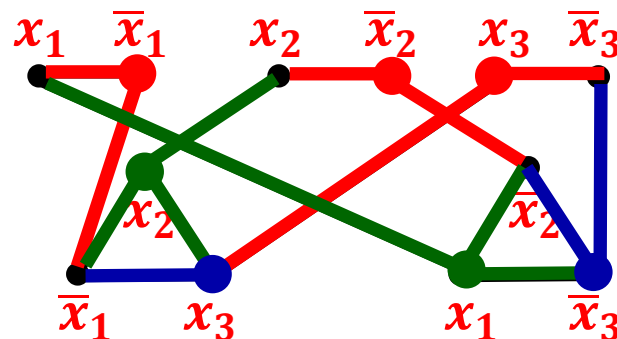
顶点覆盖问题是NP完全问题

- 往证3CNF-SAT表达式是可满足的当且仅当构造出来的无向图G中有一个大小为k覆盖集合，其中 $k=n+2m$
- 必要性证明**：若表达式可满足，则有一种变元的赋值方式，使得表达式为真。
- 变元对应的 $2n$ 个顶点选择方式：若变元 $x_i=1$ ，则将 x_i 加入集合；若变元 $x_i=0$ ，则将 \bar{x}_i 加入集合。共选择了 n 个
- 子句对应的 $3m$ 个顶点选择方式：表达式可满足，每个子句至少有一个文字值为1，该文字对应的顶点与变元对应顶点之间的边已被覆盖。选择剩下的两个顶点即可，共 $2m$ 个
- $\therefore G$ 中有大小为 $k=n+2m$ 的覆盖集合

$x_1=0, x_2=0, x_3=1$ ，则 $c_1 \wedge c_2=1$

$$(\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \bar{x}_2 \vee \bar{x}_3)$$

$k=n+2m=3+2*2=7$



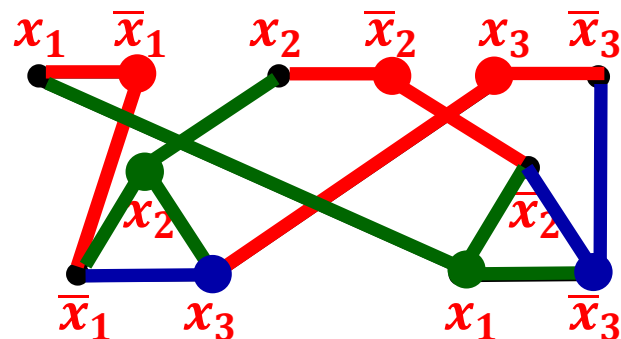
顶点覆盖问题是NP完全问题

- **充分性证明：** 若G中有大小为 $k=n+2m$ 的覆盖集合。
- 变元对应的 $2n$ 个顶点有 n 条边,要覆盖这 n 条边,每条边至少得选1个顶点, 至少共 n 个顶点。
- 子句对应的 $3m$ 个顶点是 m 个大小为3的团, 要覆盖每个大小为3的团至少得选2个顶点, 至少共 $2m$ 个顶点。
- 必然在变元对应的 $2n$ 个顶点选 n 个(每条边选一个);
- 子句对应的 $3m$ 个顶点选 $2m$ 个(每个子句选2个)
 - 剩哪一个不选? 剩和变元对应的顶点相连的不选
- 如何对变元赋值使表达式为真?
 - 若选择的变元对应的顶点为 x_i , 则 x_i 赋1;
 - 若顶点为 \bar{x}_i , 则 x_i 赋0

$$x_1=0, \quad x_2=0, \quad x_3=1$$

$$(\bar{x}_1 \vee \overset{C1}{x_2} \vee x_3) \wedge (\overset{C2}{x_1} \vee \bar{x}_2 \vee \bar{x}_3)$$

$$k=n+2m=3+2*2=7$$



子集和问题是NP完全问题

■ 子集和问题

$A = \{2, 5, 3, 6, 8, 9, 11\}, B=13$

$A' = \{2, 3, 8\} \subseteq A, 2+3+8=13$

- 有一个数集 $A = \{a_1, a_2, \dots, a_n\}$ 及一个目标数 B , 问 A 中是否能找出子集 $A' \subseteq A$, 使得 $\sum_{a_i \in A'} a_i = B$?

■ 3CNF-SAT问题

- 给定 n 个布尔变元和 m 个子句, 子句是3个文字(变元或变元的非)的析取操作
 - $(\bar{x}_1 \vee x_2 \vee x_3)$
- 问这 m 个子句的合取范式的值可否为真?
 - $(\bar{x}_1 \vee x_2 \vee x_3) \wedge (x_7 \vee x_9 \vee \bar{x}_{10}) \wedge \dots \wedge (x_5 \vee \bar{x}_6 \vee x_4)$

子集和问题是NP完全问题

■ 证明

□ 首先, 子集和问题是NP问题(多项式时间可验证)

□ 规约方法 :

➤ 3CNF-SAT有 n 个变元, m 个子句,

➤ 构造 $2(n+m)+1$ 个十进制数(每个数 $n+m$ 位):

□ 前 $2n$ 个数: x_i 与 \bar{x}_i 第 i 位为1, 出现在第 j 个子句, $n+j$ 位为1, 其他位为0

□ 中间 $2m$ 个数: g_j 与 h_j 第 $n+j$ 位为1

□ 最后一个数: 1.....13.....3(n 个1, m 个3)

x_i 与 \bar{x}_i 在
子集和问题中
表示十进制数

x_1	1	0	0	0	1
\bar{x}_1	1	0	0	1	0
x_2	0	1	0	1	0
\bar{x}_2	0	1	0	0	1
x_3	0	0	1	1	0
\bar{x}_3	0	0	1	0	1

g_1	0	0	0	1	0
h_1	0	0	0	1	0
g_2	0	0	0	0	1
h_2	0	0	0	0	1

B 1 1 1 3 3

$$\overset{C1}{(\bar{x}_1 \vee x_2 \vee x_3)} \wedge (\overset{C2}{x_1 \vee \bar{x}_2 \vee \bar{x}_3})$$

子集和问题是NP完全问题

- 往证3CNF-SAT表达式是可满足的当且仅当构造出的前 $2(n+m)$ 个十进制数中的若干个数之和等于B
- **必要性证明**：若3CNF-SAT表达式可满足，则有一种变元的赋值方式，使得表达式为真。
- 若变元 x_i 赋值为**真**，就把十进制数 x_i 放入子集
- 若变元 x_i 赋值为**假**，就把十进制数 \bar{x}_i 放入子集
 - 此时，选择的数的和前n位中每位均为1，
 - 后m位：第 $n+j$ 位可能为1，可能为2，可能为3。
为1则 g_j 和 h_j 均选择加入子集；为2则 g_j 和 h_j 任选一个加入子集；为3则均不选择
 - 此时选择数的和的后m位每位均为3

$$\overset{\text{C1}}{(\bar{x}_1 \vee x_2 \vee x_3)} \wedge \overset{\text{C2}}{(x_1 \vee \bar{x}_2 \vee \bar{x}_3)}$$

$$x_1=0, \quad x_2=0, \quad x_3=1$$

x_1	1	0	0	0	1
\bar{x}_1	1	0	0	1	0
x_2	0	1	0	1	0
\bar{x}_2	0	1	0	0	1
x_3	0	0	1	1	0
\bar{x}_3	0	0	1	0	1
g_1	0	0	0	1	0
h_1	0	0	0	1	0
g_2	0	0	0	0	1
h_2	0	0	0	0	1
B	1	1	1	3	3

子集和问题是NP完全问题

- **充分性证明：** 设 $2(n+m)$ 个数中有若干个加起来恰好为 $1.....13.....3$
 - 即使是所有数加和每位最多为5，不会产生进位
- 和中前 n 位中每位均为1，说明 x_i 与 \bar{x}_i 恰好只取了一个。
 - 若十进制数 x_i 在子集中，则令变元 x_i 为1(真)
 - 若十进制数 \bar{x}_i 在子集中，则令变元 \bar{x}_i 为0(假)
- 和中后 m 位中每位均为3
 - 即使所有 g_j 和 h_j (中间 $2m$ 个数)均在子集中，后 m 位中每位的和也不过为2，必然还要加上前面某个已在子集中的 x_i 或 \bar{x}_i ，和才为3

$$\overset{\text{C1}}{(\bar{x}_1 \vee x_2 \vee x_3)} \wedge \overset{\text{C2}}{(x_1 \vee \bar{x}_2 \vee \bar{x}_3)}$$

$$x_1=0, \quad x_2=0, \quad x_3=1$$

x_1	1	0	0	0	1
\bar{x}_1	1	0	0	1	0
x_2	0	1	0	1	0
\bar{x}_2	0	1	0	0	1
x_3	0	0	1	1	0
\bar{x}_3	0	0	1	0	1

g_1	0	0	0	1	0
h_1	0	0	0	1	0
g_2	0	0	0	0	1
h_2	0	0	0	0	1

B 1 1 1 3 3