2-SAT 解法浅析

华中师大一附中 赵爽

SAT 理论基础

设 $B=\{b_1,b_2,\cdots,b_n\}$ 为一个有限布尔变量集，$\hat{B}=\{b_1,b_2,\cdots,b_n,\neg b_1,\neg b_2,\cdots,\neg b_n\}$。设 B' 是 $\hat{B}$ 的非空子集，定义 $\vee B'=\underset{b\in B'}{\vee} b$。对于给定的 $B'_1,B'_2,\cdots,B'_m\in\hat{B}$[①]，求 B，使得

$$(\vee B'_1)\wedge(\vee B'_2)\wedge\cdots\wedge(\vee B'_m)=1$$

成立的问题，称为**适定性(Satisfiability)问题**，简称 **SAT**。

特别的，对于给定的 $\{B'_m\}$，如果 $\max\limits_{i=1,2,\cdots,m}\{|B'_i|\}=k$，我们就把这个问题称为 **k-适定性问题**，简称 **k-SAT**。

可以证明，当 $k>2$ 时，k-SAT 是 NP 完全的。下面我们要讨论的，是 $k=2$ 时的情况。

2-SAT

在 2-SAT 中，B'_i 只有两种形式，一种是单个布尔变量 $x\in\hat{B}$，另一种是两个布尔变量的或：$x\vee y\left(x,y\in\hat{B}\right)$。为了方便，我们先分析只存在后一种形式的情况。

我们可以构造有向图 G。G 包含 $2n$ 个顶点，代表 $\hat{B}$ 中的 $2n$ 个元素。我们的问题转化为从 G 中选出 n 个顶点，使其满足 2-SAT 的条件——当然，代表 b_i 和 $\neg b_i$ 的顶点不能同时被选择。下面我们分析一下 $B'_i=x\vee y$ 在图对应什么。

显然，$x\vee y=\neg(\neg x\wedge\neg y)$。这也就是说，如果我们选中 $\neg x$，那么我们必须选择 y；同样的，如果我们选中 $\neg y$，那么我们必须选择 x。因此，对于 $B'_i=x\vee y$，我们可以在 G 中增加弧 $(\neg x,y)$ 和 $(\neg y,x)$[②]。

[①] 在下文中，$x_1,\cdots,x_n$ 简写作 $\{x_n\}$。

[②] 这里，$x,y,\neg x,\neg y$ 都表示 G 中代表它们的顶点。下同。

然后我们可以求 G 的所有强连通分量。很明显，如果我们选中强连通分量中的任何一点，那么该强连通分量中的所有其它的顶点也必须被选择。如果 b_j 和 $\neg b_j$ 同属于一个强连通分量，那么产生矛盾，该 2-SAT 无解。

如果没有产生矛盾，我们就可以把处在同一个强连通分量中的点和边缩成一个点，得到新的有向图 G'。然后，我们把 G' 中的所有弧反向，得到图 G''。

现在我们观察 G''。由于已经进行了缩点的操作，因此 G'' 中一定不存在圈，也就是说，G'' 具有拓扑结构。

我们把 G'' 中所有顶点置为“未着色”。按照拓扑顺序重复下面的操作：

1、选择第一个**未着色**的顶点 x。把 x 染成红色。

2、把所有与 x 矛盾的顶点 y（如果存在 $b_j, \neg b_j \in \hat{B}$，且 b_j 属于 x 代表的强连通分量，$\neg b_j$ 属于 y 代表的强连通分量，那么 x 和 y 就是互相矛盾的顶点）及其子孙全部全部染成蓝色。

3、重复操作 1 和 2，直到不存在未着色的点为止。此时，G'' 中被染成红色的点在图 G 中对应的顶点集合，就对应着该 2-SAT 的一组解。

那么，以上的操作中是否可能出现矛盾呢？答案是否定的。这是因为：首先，假如我们选定了 G'' 中的未着色顶点 p，那么和 p 矛盾的所有其它顶点及其后代均会被染成蓝色。因此，我们把一个顶点 p 染成红色时，任何一个和 p 矛盾顶点都不会也是红色。

同时，由于我们按照拓扑顺序检查，并且把一个顶点染成蓝色的时候，立刻把它的所有子孙也染成蓝色。也就是说，如果一个顶点 p 不可选，那么所有直接或间接满足条件“假如选择 q 就必须选择 p”的顶点也会被染成蓝色。这样一来，G'' 中不存在弧 (x, y)，其中 x 为红色而 y 为蓝色。因此我们得到的结论不可能和 B_i' 对应的条件矛盾。

综合这两条结论，我们就可能证明上面的操作中不会产生任何矛盾。

下面证明 G'' 对应的解就是该 2-SAT 的解，即：

1、我们得到的解不会同时选定 b_j 和 $\neg b_j$。

证明：首先，对于 b_j 和 $\neg b_j$，它们在 G 中一定属于不同的强连通分量（如果不满足这个条件，事先就会被判定为无解），因此在 G'' 中被不同的顶点所代表，不妨设为 p 和 q。显然 p 和 q 是相互矛盾的顶点。由于上面的着色操作不会产生矛盾（已证），因此 p 和 q 不会被同时染成红色，即不可能同时选定 b_j 和 $\neg b_j$。证毕。

2、我们得到的解对于任意的 $b_j, \neg b_j \in \hat{B}$，会至少包含其中的一个。

证明：注意到对于 G 中任何一条弧 (x, y)，一定存在一条与之“对偶”的弧 $(\neg y, \neg x)$。这

是由 B_i' 的形式决定的。这就意味着如果 x, y 处于 G 的同一个强连通分量内，那么 $\neg x, \neg y$ 必然也处于同一个强连通分量内；如果 x, y 之间有路，那么 $\neg y, \neg x$ 之间也有路。

假设 $b_j, \neg b_j$ 均未被选中，即它们所在的强连通分量 p, q 均被染成蓝色。下面分两种情况进行讨论:

(1) p,q 是在同一次染色操作中被染成蓝色

不妨假设把顶点 r 染成红色之后，将 p, q 同时染成蓝色。在 G'' 中，p, q 之间不存在路径。因为如果 p, q 之间有路，那么由 G 的结构可知 q, p 之间也有路，这和 G'' 的拓扑结构矛盾。如果 p, q 之间没有路，那么一定存在 $p', q' \in G''$，它们和 r 直接矛盾，且 p 是 p' 的后代，q 是 q' 的后代。即存在 $G(x) \in G''(r), G(\neg x) \in G''(p')$ [3]，以及 $G(y) \in G''(r), G(\neg y) \in G''(q')$。但是由 G 的结构，$\neg x, \neg y$ 应该处于同一个强连通分量内，这和 $G(\neg x) \in G''(p'), G(\neg y) \in G''(q')$ 矛盾！

因此 p, q 不可能在同一次染色操作中被染色。

(2) p,q 先后被染成蓝色

不妨假设 q 后被染色，并且在把 r 染成蓝色的时候，由于 q 和 r 矛盾，才把 q 染成蓝色。

如果 q 和 r 是直接矛盾的，即存在 $G(x) \in G''(r), G(\neg x) \in G''(q)$，又 $G(\neg b_j) \in G''(q)$，由 G 的结构可知 $G(b_j) \in G''(r)$，因此 $r = p$，这和 p 被染成蓝色矛盾。

如果 q 和 r 是间接矛盾的，即存在 $G(x) \in G''(r), G(\neg x) \in G''(r')$，且 q 是 r' 的祖先。又由 G 的结构可知 p 必然是 r 的祖先。而 p 已经被染成了蓝色，因此 r 根本不可能被染成红色，矛盾！

综合(1)，(2)，命题得证。

通过上面的证明，我们知道这种算法是正确的。而求有向图的强连通分量、拓扑排序的时间复杂度都是 $O(e)$，而染色操作的复杂度是 $O(e)$，因此整个算法的时间复杂度是 $O(e)$。

[3] 这里的 $G(x) \in G''(r)$ 表示 G 中的顶点 x 所在的强连通分量在 G'' 中是顶点 r。下同。