第 14 卷 第 5 期　　自 动 化 学 报　　Vol. 14, No. 5
1988 年 9 月　　ACTA AUTOMATICA SINICA　　Sept., 1988

用于特征选择的 BF* 算法及其与 B&B 算法的比较

徐 雷　阎平凡　常 迥
(清 华 大 学)

摘　要

本文将模式识别中的特征选择问题转化为有向图上最佳路径搜索问题，并应用 AI 中的 Best First（简记 BF*）策略搜索最佳路径，提出了特征选择 GBFF* 和 TBFF* 算法，证明了用它们可不穷举而一定找到最佳子集，同目前被认为最好的全局最佳算法——B & B 相比，TBFF* 搜索的特征子集数目优于 B&B.

一、引　言

设特征集 $S_N=\{x_1, x_2, \cdots, x_N\}$，特征选择问题为：选择一个 $m<Np$ 元素构成的子集 S_m^*，使 $F(S_m^*)=\mathrm{Max}F(S_m)$，对于所有 $S_m\subset S_N$。$F(\cdot)$ 为特征选择用的准则函数。

特征选择问题在统计 *P. R.* 中起着重要作用，早期发表的特征选择算法皆属于逐步优化的算法，例如，从底向上、从顶向下、序贯等算法以及它们的各种变形或组合[1,2]. Cover 的研究[3,4]表明，逐步优化的方法不能保证找到全局最优子集。要全局最优，只有穷举所有 C_N^m 个可能的子集，但这又导致计算量的"指数爆炸"。Narendra & Funkunaga[5] 提出了用于特征选择的 Branch and Bound（简记 B&B）算法，它可以既考虑所有可能的子集，而又通过定界方法在许多情况下并不真正穷举而得到最优的特征子集。这一结果目前被认为是求全局最优子集的最好算法[6]。

本文将特征选择问题转化为有向图上最佳路径搜索问题，并应用 AI 中的 Best First (BF*) 策略搜索最佳路径，提出了 GBFF* 算法和 TBFF* 算法，证明了用它们进行特征子集选择，可不穷举而一定找到最佳子集。且与目前被认为最好的全局最佳算法 B&B 相比，TBFF* 搜索的特征子集数目少于(甚至远少于) B&B 所搜索的特征子集数目，也就是说 TBFF* 算法优于 B&B 算法。

二、特征选择与有向图搜索

类似 [5]，从顶向下逐次从 S_N 中丢弃一个变量，则 $N-m$ 次后得一个 S_m 子集。考

本文于 1986 年 5 月 12 日收到.

虑 $\forall S_m \in S_N$，可得图 1 所示的有向图．仅有的一始节点对应 S_N，C_N^m 个终节点对应所有可能的 S_m 子集．每条有向边对应被丢弃的变量，每个节点对应此刻没有被丢弃的那些变量构成的子集．用每个节点对应的子集的准则值 $F(\cdot)$ 作为节点值．于是，特征选择问题可转化为：在有向图中，由始节点出发，搜索一条到达具有最大节点值的某终节点的最佳路径问题．如果把节点值做为启发函数，可用 AI 中的 Best First (BF) 策略[7]搜索这样的最佳路径．进一步分析可发现图 1 所示的有向图还有一些特殊性质，可用来由 BF 策略得到适于特征选择问题的搜索算法．

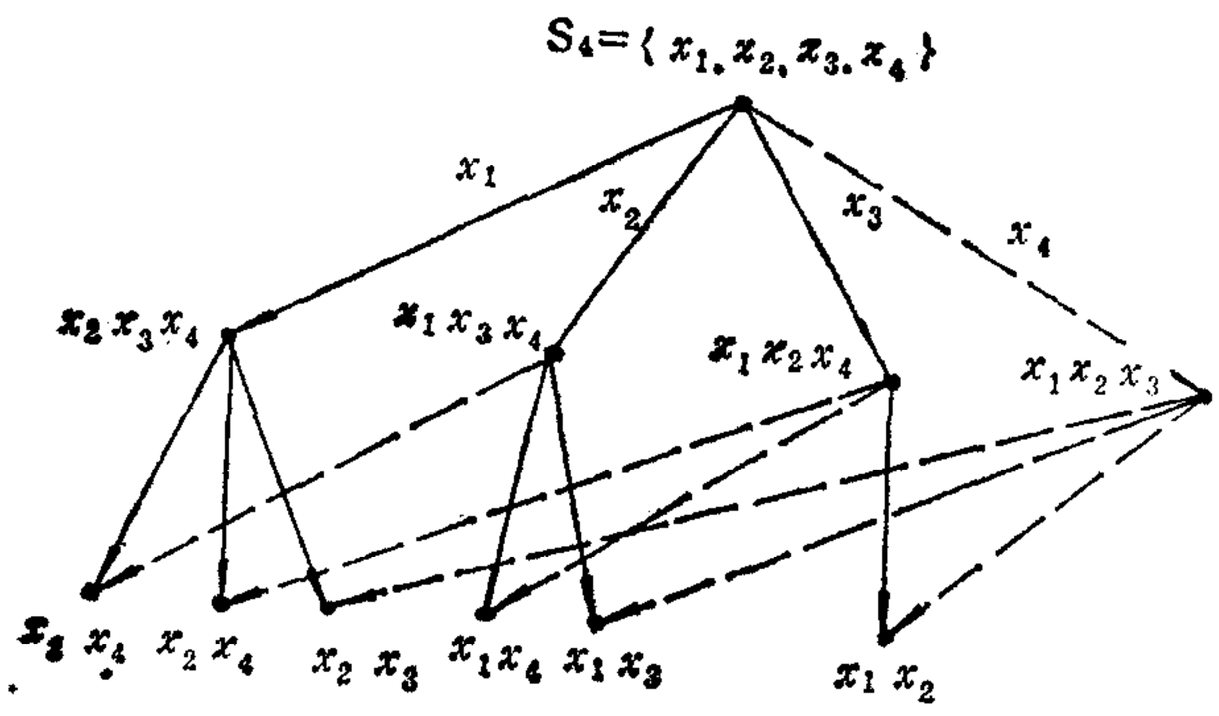

图 1 用于特征选择的加权有向图 $N=4$，$m=2$

下面先介绍一些术语和符号：

有向图 $G=\{V, E, S_0, \Gamma, f\}$：其中 V 为节点集．每个 $n \in V$ 对应一个子集 $S_l \subset S_N$，$m \leqslant l \leqslant N$，有时记 $n(S_l)$．E 为有向边集，每个 $e \in E$，e 从 $n(S_l)$ 指向 $n(S_{l-1})$，对应一个由 S_l 中取出丢弃的变量（结果得到 S_{l-1}）．$S_0 = n(S_N)$ 为始节点，Γ 为终节点集对应由所有 C_N^m 个子集 $S_m \subset S_N$ 构成的集，f 为启发函数，对于节点 $\lambda(S_l)$，$f(n)$ 表示其节点值，即准则值 $F(S_l)$．

路径：$P_{n_i-n_j}$ 指由 n_i 到 n_j 的某一条路径，$P^s = P_{s_0 \to \gamma}$ 为 S_0 到某个 $\gamma \in \Gamma$ 的一条解路．本文中 G 为有限图，解路可枚举，记 $\boldsymbol{P} = \{P^s / G$ 中所有 $P^s\}$．记 $f(P_{s_0 \to n})$ 为 $P_{s_0 \to n}$ 的路径值，并令 $f(P_{s_0 \to n}) = f(n)$．若对于 $f(P^s) = f(\gamma)$，$\gamma \in \Gamma$，有 $f(\gamma) = \mathrm{Max} f(\alpha)$，$\forall \alpha \in \Gamma$，则 P^s 称为最佳解路，记为 P^{s*}．

节点关系：SON(n) 表示 n 的子节点集，FAT(n) 表示 n 的父节点集，DES(n) 表示 n 的后代节点集；ANC(n) 表示 n 的先辈节点集．

算子 OPERATOR($\cdot\cdot$)：当用它作用于节点 n（记为 OPERATOR(n)），则产生 SON．本文定义二个这样的算子，分别在 GBFF* 和 TBFF* 中产生不同的 SON(n)．

其它术语和符号，例如 OPEN．CLOSED，可停止的、可采纳的、完备的等，含义见[7]．

现可将特征选择问题形式地表述为：

问题 A：在图 G 中，寻找 P^{s*}，使 $f(P^{s*}) = f^*$，$f^* = \mathrm{Max} f(P_i^s)$，$\forall P_i^s \in \boldsymbol{P}$，也即 $f^* = \mathrm{Max} f(\alpha), \forall \alpha \in \Gamma$．

三、GBFF* 算法及其有关定理

1. GBFF* 算法

步 1：将 $S_0 = n(S_N)$ 放入 OPEN 表．

步 2：从 OPEN 中的所有节点中选取一节点 n，其 $f(n)$ 最大．若有数个，则优先取终节点；若没有终节点，取深度最深的节点，其特征变量数最少；若这样的节点还不止一

个，任选其一．然后，将 n 放入 CLOSED 表中．

步 3：若 $n\in\Gamma$，结束．此时的 S_m 作为最佳 S_m^*．

步 4：OPER1(n)，得 SON(n)．即对于 $n=n(S_k)$，$S_k=\{x_1, x_2, \cdots, x_k\}$，得到 $\mathrm{SON}(n)=\{n_1, n_2, \cdots, n_k\}$，$n_1=n(S_k-\{x_1\})$，$n_2=n(S_k-\{x_2\})$，$\cdots$，$n_k=n(S_k-\{x_k\})$．

步 5：对于每个 $n'\in\mathrm{SON}(n)$，若 $n'\notin$ OPEN，且 $n'\notin$ CLOSED，则计算 $f(n')$，将 n' 放入 OPEN．转步 2．

说明：GBFF* 由 BF* 算法[7]并考虑到特征选择问题的特点修改得到．这里取最大代替了取最小，并不必考虑回溯和 reopen 问题．

2. 有关图 G 性质和有关算法的定理

用 OPER1($\cdot$)，由 S_0 起，逐次作用其产生的所有节点并以深度 $N-m$ 为限制，将产生图 G，即图 1 所示有向图，这个图具有如下性质：

引理 1. 任给 $n\in G$，设 $P'_{s_0\to n}$，$P''_{s_0\to n}$ 为由 S_0 到 n 的任意两条不同的路径，则有

$$f(P'_{s_0\to n})=f(P''_{s_0\to n})=f(n).$$

此引理由 $f(P_{s_0\to n})$ 的定义易得，它保证了 GBFF* 不用考虑 reopen 问题．另外，由准则 $F(\cdot)$ 的单调性可得：

引理 2. 对于每个 $n\in G$，有

1) 若 $n'\in\mathrm{FAT}(n)$，$n''\in\mathrm{SON}(n)$，则 $f(n'')\leqslant f(n)\leqslant f(n')$．

2) 若 $n'\in\mathrm{ANC}(n)$，$n''\in\mathrm{DES}(n)$，则 $f(n'')\leqslant f(n)\leqslant f(n')$．

3) 对于任一路径 $P_{s_0\to n}$，若 n' 为 $P_{s_0\to n}$ 上的一个节点，则有

$$f(n')\geqslant f(n). \tag{1}$$

引理 3. 设 $\gamma\in\Gamma$，$P^s=P_{s_0\to\gamma}=S_0, n_1, n_2, \cdots, \gamma$，

则有

$$f(\gamma)\leqslant\cdots\leqslant f(n_2)\leqslant f(n_1)\leqslant f(S_0). \tag{2}$$

有了上述引理，现在研究有关 GBFF* 算法的几个定理．因 G 为有限图，易知 GBFF* 一定可停止且是完备的，下面来讨论其可采纳性．

引理 4. 当 S_0 展开后，在任一 $P_{s_0\to n}$ 上，当 n 展开前，必有一个节点 n' 同时在 $P_{s_0\to n}$ 上和 OPEN 中．

证明：用反证法，假设存在一 $P_{s_0\to n}$，其上无节点在 OPEN 中，则一方面，可证 $P_{s_0\to n}$ 上没有一个节点曾被展开，这与假设矛盾．另一方面，由于在 $P_{s_0\to n}$ 上必有节点 $n''\in\mathrm{SON}(S_0)$，它在 OPEN 或已被展开，也与假设矛盾(详略)．证毕．

定理 1. GBFF* 是可采纳的(即算法停止时，找到的解一定是最优解)．

证明：假设矛盾，即 GBFF* 在 $B\in\Gamma$ 处停止，但 $f(\beta)<f^*=\underset{\forall\gamma\in\Gamma}{\mathrm{Max}}\, f(\gamma)$．设

$$f(P_{s_0\to\alpha})=f(\alpha)=f^*,\ \alpha\in\Gamma,$$

由引理 4，展开节点 α 前，必有一个 $n\in$ OPEN 且 n 在 $P_{s_0\to\alpha}$ 上，由于 $n\in\mathrm{ANC}(\alpha)$，由引理 2，有 $f(n)\geqslant f(\alpha)=f^*>f(B)$，于是，GBFF* 在步 2 将选择 n 展开，而不是在 B 处停止．矛盾．证毕．

定理 2. 设 $f^* = \underset{\forall \gamma \in \Gamma}{\mathrm{Max}} f(\gamma)$，对于任一个 $n \in G$，n 被 GBFF* 展开的必要条件为 $f(n) \geqslant f^*$，充分条件为 $f(n) > f^*$.

证明：1）必要性：假设有一个节点 n 被展开且 $f(n) < f^*$. 令 $f^* = f(P_{s_0 \to \alpha})$，$\alpha \in \Gamma$，由引理4，在 α 被展开前，必 $\exists n' \in$ OPEN 且 n' 在 $P_{s_0 \to \alpha}$ 上，由引理 2 有 $f(n') \geqslant f^* > f(n)$，于是在展开节点 n 前，GBFF* 将在步 2 选取 n' 而不是 n 展开. n' 展开后，又有 $n'' \in \mathrm{SON}(n')$ 且 n'' 在 $P_{s_0 \to \alpha}$ 上，类似地 n'' 将在 n 以前展开，…，如此下去，由引理 3，直到 GBFF* 展开 α 后停止，n 仍在 OPEN，与假设矛盾.

2）充分性：假设 GBFF* 在 $\gamma \in \Gamma$ 处停止，$f(\gamma) = f(P_{s_0 \to \gamma}) = f^*$，而存在一个节点 $n \in G$ 且 $f(n) > f^*$ 却未被展开.

如果 $n \in$ OPEN，则 GBFF* 在 γ 停止前，将在步 2 选择 n 而不是 γ 展开，这与假设矛盾；如果 $n \notin$ OPEN，由引理 3，在 $P_{s_0 \to n}$ 上有一个 $n' \in \mathrm{ANC}(n)$ 且 $n' \in$ OPEN，由引理 2，$f(n') \geqslant f(n) \geqslant f^* = f(\gamma)$，于是 n' 将在 γ 前被展开；一旦不被展开后，$n'' \in \mathrm{SON}(n')$ 必将在 $P_{s_0 \to n}$ 上和 OPEN 中，类似地 $f(n'') > f^* = f(\gamma)$，n'' 将在 γ 前被展开，…，如此下去，直到 n 被放入 OPEN 且在 γ 前展开. 与假设矛盾. 证毕.

定理 1 保证了 GBFF* 是全局最优算法，而定理 2 则指出了 G 中所有 $f(n) < f^*$ 的节点将不被枚举，GBFF* 不是穷举法. 进一步有：

定理 3. 在图 G 中，令 $f^* = \underset{\gamma \in \Gamma}{\mathrm{Max}} f(\gamma)$，$N(\mathrm{GBFF}^*)$ 表示 G 中被展开的节点数，$N_g(G)$ 和 $N_e(G)$ 分别表示 G 中所有 $n \in V-\Gamma$ 且 $f(n) > f^*$ 和 $f(n) = f^*$ 的节点数. $N(G)$ 表示 $V-\Gamma$ 中的所有节点数，则有：

$$N_g(G) \leqslant N(\mathrm{GBFF}^*) - 1 \leqslant N_g(G) + N_e(G) \leqslant N(G),$$

而且仅当对所有 $n' \in \mathrm{FAT}(\Gamma)$，$f(n') \geqslant f^*$ 时，可能有 $N_g(G) + N_e(G) = N(G)$，其中 $\mathrm{FAT}(\Gamma) = \{n /$ 对于所有 $\gamma \in \Gamma$，n 是 γ 的父节点$\}$.

四、TBFF* 算法及其与 B&B 算法的比较

1. TBFF* 算法

由图 1 可看出，去掉对应 $\{x_1, x_2, x_3\}$ 的节点及虚线边，便是[6]中的搜索树. GBFF* 算法虽然在步 5 避免了考虑虚线边，但没有避免象 $\{x_1, x_2, x_3\}$ 这样的点，还有改进的可能. 本文用 OPER2(n) 将特征选择问题转化为图 2 所示的树上搜索最佳路径的问题，此树的结构与[6]中 B & B 算法的搜索树相同，以后记此树为 $T_e = \{V, E, S_0, \Gamma, f\}$. 与 G 相比，E 中边的数目和 V 中节点的数目都减少了许多. 将 G 改成 T_e 的关键是 OPER1(n) 换成 OPER2(n)，这时，GBFF* 算法变成了 TBFF* 算法如下：

步 1：对 S_N 的 N 个变量标序，得 $S_N = \{x_1, x_2, \cdots, x_N\}$，$S_0 = n(S_N)$ 放入 OPEN，令 $d(S_0) = 0$，$I(S_0) = 1$.

步 2—3：同 GBFF* 的步 2、步 3.

步 4：OPER2(n)，即产生子节点 $n_1(S_1)$，$S_1 = S - \{x_k\}$，$k = I(n)$，令 $d(n_1) = d(n) + 1$，$I(n_1) = I(n) + 1$，再产生 $n = (S_2)$，$S_2 = S - \{x_k\}$，$k = I(n_1)$ 令

$$d(n_2)=d(n)+1,\ I(n_2)=I(n_1)+1,\cdots,$$

如此直到产生第 l 个子节点，$n_l(S_l)$，$S_l=S-\{x_k\}$，$k=I(n_l-1)$，令

$$d(n_l)=d(n)+1,\ I(n_l)=I(n_l-1)+1.$$

这里，$l=m+2+d(n)-I(n)$，得到子节点集 $\mathrm{SON}(n)=\{n_1,n_2,\cdots,n_l\}$。

步 5：计算 $f(n')$，$\forall n'\in\mathrm{SON}(n)$，将 SON($n$) 放入 OPEN，转步 2。

说明：$d(n)$ 表示 n 在 T_e 中的深度(即[6]中的 Level)，$I(n)$ 的作用对应于[6]中的 POINTER。

不难验证，第三节中的所有引理和定理对 TBFF* 均成立，只要相应地将“GBBF*”改为“TBFF*”及“图 G”改成“树 T_e”便可。

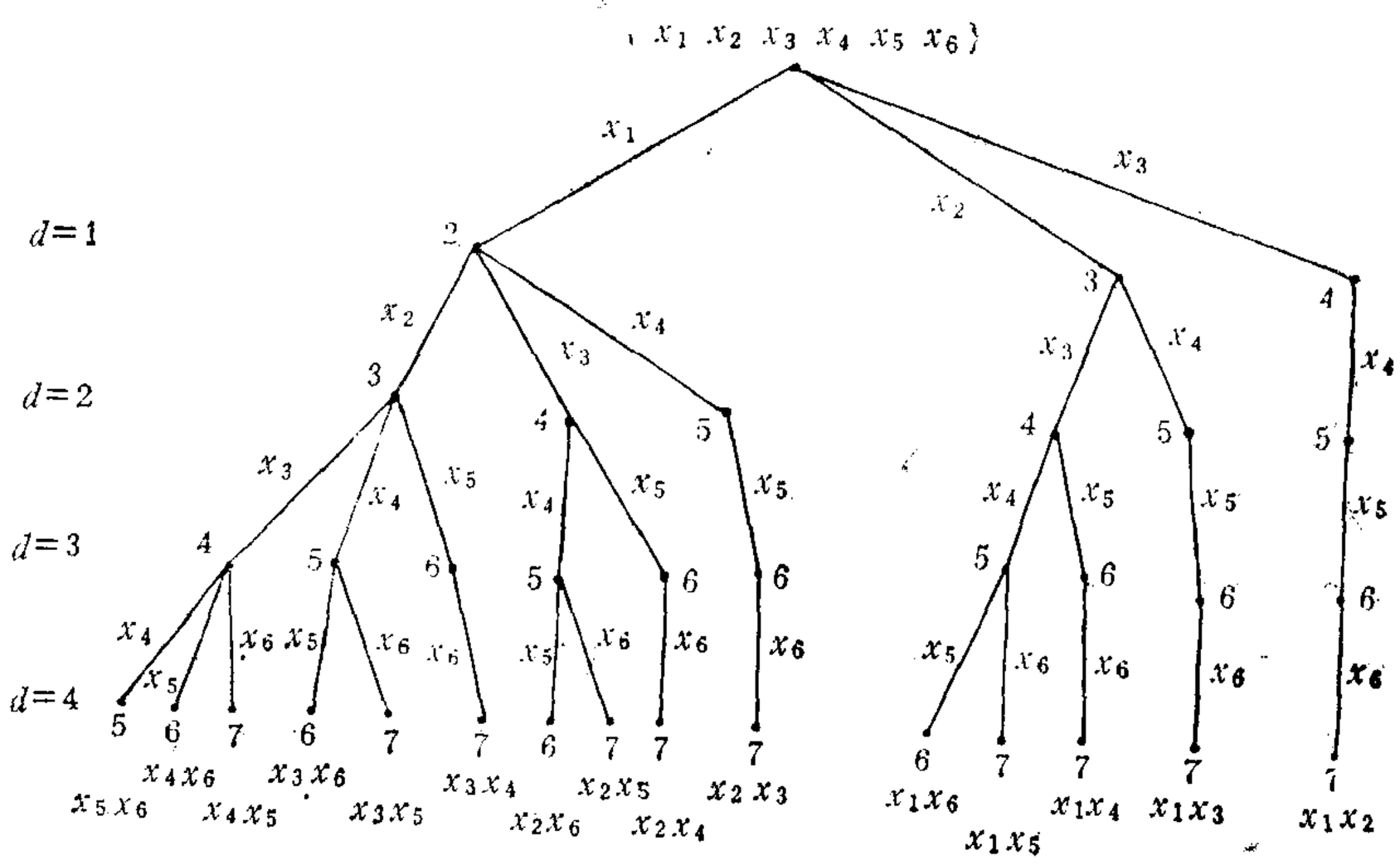

图 2　T_e 搜索树

$N=6$，$m=2$。图中每条边表示该次去掉的变量，节点旁的整数为 $I(n)$ 标号

图 2 为取 $N=6$，$m=2$，$S_N=\{x_1,x_2,x_3,x_4,x_5,x_6\}$，由 $S_0=n(S_N)$ 起，

$$\alpha(S_0)=0,\ I(S_0)=1,$$

依次使用 OPER2(n)，不考虑 $f(n)$，且以 $(N-m)$ 深度为限，所得到的穷举搜索树 T_e。最底排为所有可能的两变量特征子集。不难看出图 2 的结构同 [6] 中图 1 的结构一样，不同的是不象[6]用 $f(n)$ 对节点排序，这对 TBFF* 不必要。

2. TBFF* 与 B&B 的比较

由[6]，B&B 也是可采纳的，其搜索树与 TBFF* 的搜索树 T_e 结构相同，其起始节点 S_0 与目标集 Γ 也相同，且 $N(T_e')=N(T_e)$，故可比较两个算法的优劣。

由[6]的枚举过程，B&B 将不展开每个 $n\in T_e'$ 其 $f(n)<B$，且

$$0<B\leqslant f^*=\mathrm{Max}f(\gamma),\ \text{对于所有}\ \gamma\in\Gamma. \tag{3}$$

B&B 枚举开始时，初始估计 $B=B_0$，以后每当搜索达到一个节点 $\gamma\in\Gamma$ 时，B 被刷新一次。若以 $B_0,B_1,B_2,\cdots,B_i,\cdots$，表示 B 的变化，则由 B&B 的步 3 和步 5，有：

$$B_0\leqslant B_1\leqslant B_2\leqslant\cdots\leqslant B_i\leqslant\cdots\leqslant f^*, \tag{4}$$

其中 $f^* = \underset{\forall\gamma\in\Gamma}{\mathrm{Max}} f(\gamma)$. 由 B&B 的步 3 和(4)式,类似于定理 2 和定理3,易得:

定理 4. 对于任一个 $n\in T'_c$, 被 B&B 展开的充分条件为 $f(n) > f^*$, 必要条件为 $f(n) > B$.

定理 5. 令 $N(\mathrm{B\&B})$ 表示 T'_c 中被 B&B 展开的所有节点的数目, $N_{B_0}(T'_c)$ 表示 $V'-\Gamma'$ 中所有的 $f(n) > B_0$ 的节点数目,则有:

$$N_g(T'_c) + N_e(T'_c) \leqslant N(\mathrm{B\&B}) - 1 \leqslant N_{B_0}(T'_c), \tag{5}$$

在定理 5 中,仅当 $B_0 < B^* = \mathrm{Min}f(n)$, $n\in \mathrm{FAT}(\gamma)$, $\forall\gamma\in\Gamma$ 时,才有

$$N_{B_0}(T'_c) = N(T'_c).$$

由于实际中 B_0 常初始被置值为 0 或很小的值以避免 B&B 的错误拒绝行动. 故(5)式实际上可写成:

$$N_g(T'_c) + N_e(T'_c) \leqslant N(\mathrm{B\&B}) - 1 \leqslant N(T'_c) \tag{6}$$

为了进一步讨论,再补充一些符号和术语:

1) 令 i^* 表示(4)式中 $B_i = f^*$ 时的 i, i^* 可能取 $[0, C_N^m]$ 中的所有整数. 对于节点 $n\in T'_c$, 如果 n' 在 n 的同深度的左侧(或右侧), n' 称为 n 的左(或右)兄弟, n 的所有左(或右)兄弟构成集 LBRO(n) (或 RBRO(n)). 若 $n'\in$ LBRO(n) 且 n' 紧邻 n, n' 称为 n 的第一左兄弟, 记为 Lbro1(n). 类似地有第一右兄弟, 记为 Rbro1(n). 若 $n' =$ Lbro1(n''), $n'' =$FAT(n), 则 n' 称为 n 的第一左叔叔,记为 Lunc1(n). 进一步的,如果 $n' =$Lunc1(n) 且 $f(n') \geqslant f(n'')$, $\forall n''\in$ RBRO(n), 则 n' 称为 n 的强 Lunc1(n).

2) 由 S_N, 用从顶向下方式,逐一地从 S_N 丢弃一个特征 $x^*\in S_N$, 使

$$f[S_N - \{x^*\}] = \underset{\forall x\in S_N}{\mathrm{Max}} f[S_N - \{x\}], \cdots,$$

如此直到得到 S_m^* 且 $f(S_m^*) = \underset{\forall x\in S_{m-1}}{\mathrm{Max}} f[S_{m-1} - \{x\}]$. 如果对于这样得到的 S_m^*, 有 $f(S_m^*) = \underset{\forall S_m\subset S_N}{\mathrm{Max}} f(S_m)$ 成立,则称 S_N 为可逐一选择的特征集.

定理 6. 1) $N(\mathrm{B\&B}) = N(T'_c)$, 当 ($a$) $f(n') \geqslant f^*$, $\forall n'\in$ FAT(Γ), 或当(b)对于所有 $\gamma\in\Gamma'$, Lunc1(γ) 是强的.

2) $N(\mathrm{B\&B}) \geqslant N(\mathrm{TBFF^*})$, 且差 $\Delta N = N(\mathrm{B\&B}) - N(\mathrm{TBFF^*})$ 随 i^* 增加. 等号仅当以下两种情况时成立: (a) S_N 是可逐一选择的特征集. (b) $f(n') \geqslant f^*$, 对于所有 $n'\in$ FAT(Γ).

证明: 1) 当 (a) 类似定理 3 有 $N(\mathrm{B\&B}) = N(T'_c)$, 当 ($b$) 对任一 $\gamma\in\Gamma'$, 有 $n =$ Lunc1(γ), $f(n) \geqslant f(\alpha)$, $\forall\alpha\in$ RBRO(n); 则由引理 2 有 $f(n') \geqslant f(n) \geqslant B$, $n'\in$ ANC(n) 且 n' 在 $P_{s_0\to n}$ 上,由 B&B 的步 3 , n' 和 n 都不会被拒绝,于是有 $N(\mathrm{B\&B}) = N(T'_c)$.

2) 显然, $N(T'_c) = N(T_c)$, $N_e(T'_c) = N_e(T_c)$, $N_g(T'_c) = N_g(T_c)$, 由定理 3、5, $N(\mathrm{B\&B}) \geqslant N(\mathrm{TBFF^*})$ 成立,且 i^* 越大, $f^* - B$ 越大, 满足 $B \leqslant f(n) \leqslant f^*$ 而被展开的节点越多,于是 ΔN 越大,即 ΔN 随 i^* 增加而增加.

对于"="情形, (a) 由于 B&B 的节点排序方式,必有

$$B_1 = f^*,\ N(\text{B\&B}) = N_g(T_c') + N_c(T_c'),$$

但 $N(\text{TBFF*})$ 只是可能等于 $N_g(T_c) + N_c(T_c)$，故 $N(\text{B\&B}) = N(\text{TBFF*})$ 可能成立。

当 (b)，由 (i) 和定理 3，当 $f(n') > f^*$，$\forall n' \in \text{FAT}(T)$ 有 $N(\text{B\&B}) = N(T_c')$，$N(\text{TBFF*}) = N(T_c)$，故必有 $N(\text{B\&B}) = N(\text{TBFF*})$；但是，当 $f(n') = f^*$，$\forall n' \in \text{FAT}(T)$ 时，尽管仍有 $N(\text{B\&B}) = N(T_c')$，但 $N(\text{TBFF*}) = N(T_c)$ 只是可能成立，取决于 TBFF* 在步 2 中当若干个最大值节点出现时的处理情况。证毕。

总而言之，由定理 3、6，TBFF* 展开的节点数总是小于(甚至远小于，当 i^* 充分大时) B&B 展开的节点数。两算法效果相当的情况仅在 S_N 为可逐一选择的特征集或 $f(n') \geqslant f^*$，$\forall n' \in \text{FAT}(T)$ 皆不如简单的由顶向下的算法；而对于后者，若 $f(n') \geqslant f^*$，$\forall n' \in \text{FAT}(T)$，B&B 将穷举 T_c' 中的所有节点；若严格不等式 $f(n') > f^*$ 成立，TBFF* 也将穷举 T_c 中的所有节点。这种情形是两个算法的最坏情形，也是 TBFF* 仅有的最坏情形；但对于 B&B 还有另一种最坏情形，即对于所有 $\gamma \in \Gamma$，Lunc1(γ) 是强的，在此情形下，B&B 也得穷举 T_c' 中的所有节点。

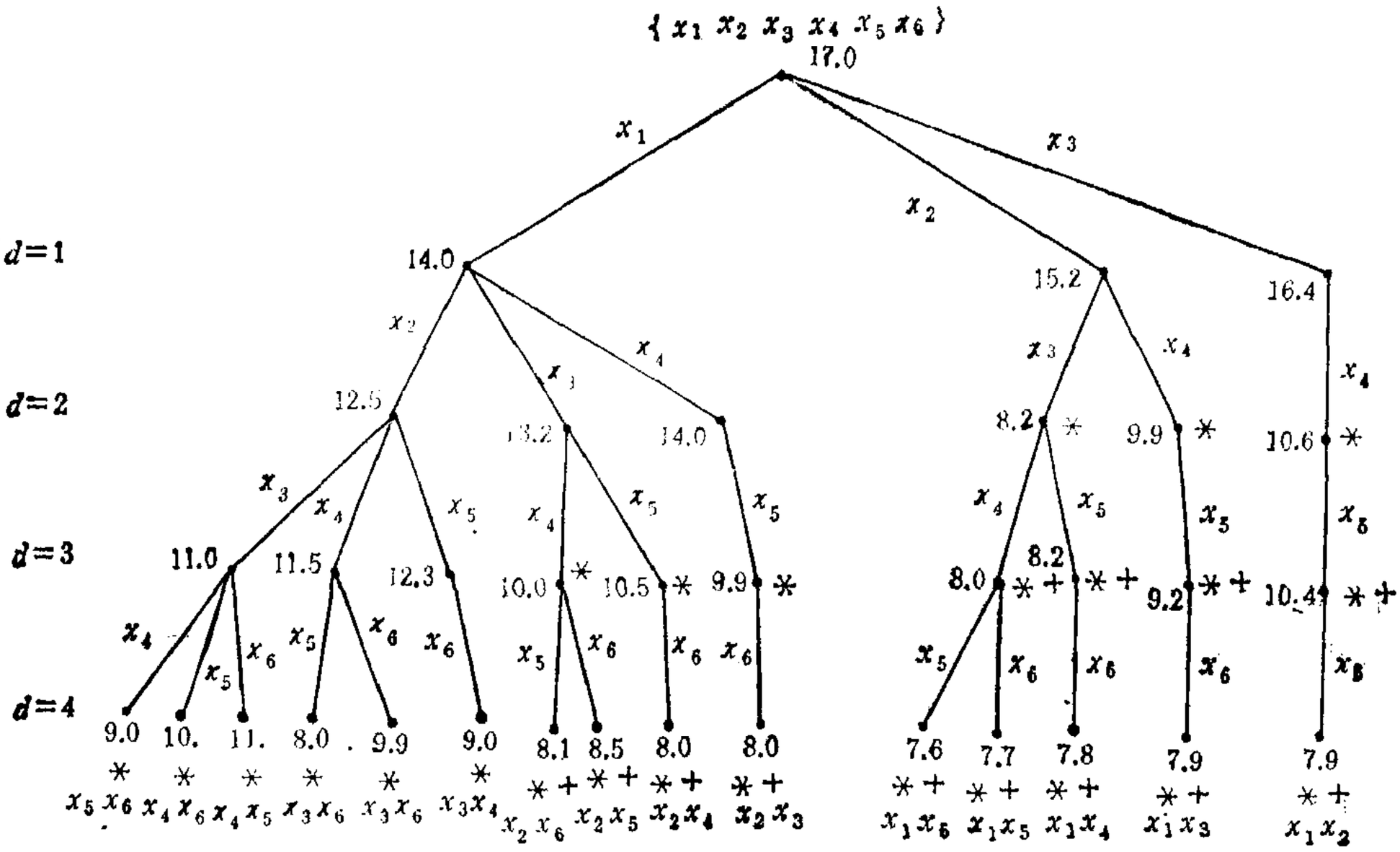

图 3　TBBF* 和 B&B 比较的一个例子。

图中每条边表示该次去掉的变量，节点旁为节点值。用 B&B 算法，为了找最佳目标节点 xx，图中全部节点值必须算出(穷举)，也即树中 $d \leqslant 3$ 的节点必被展开，用 TBFF* 算法，为了找到 xx，"+" 所示的节点值都不必算出，也即树中"*"所示的节点不必展开。图中，$N = 6$，$m = 4$。

图 3 给出了说明上述结论的一个例子。图中标志"+"的节点(对应有 9 个两变量子集，而两变量子集总共才 $C_6^2 = 15$ 个)，图中所有标志"*"的节点均不被 TBFF* 展开。

最后，值得指出的是 TBFF* 所要求的存贮量比 B&B 所要求的大，TBFF* 和 B&B 两算法在存贮量和运行时间上的关系类似于 AI 中的有信息的 Best First 算法和有信息的 Backtracing 算法[7]。

参 考 文 献

[1] Stearns, S D., On Selecting Features for Pattern Classifiers, Proc. of 3rd IJCPR Coronado, CA., Nov., 1976, 71—75.

[2] Kittle, J., Feature Set Search Algorithms, in Pattern Recognition and Signal Processing. Chen, C. H., Alphen ann den Rijn: Sythoff and Noordhoff, 1978, 41—46.

[3] Cover, T. M., The Best Two Independent Measurements Are Not The Two Best. *IEEE Tr* Vol. **SMC-4**, No. 1, 116—117, Jan, 1974.

[4] Cover, T. M. and Van Campenhout, J. M. On the Possible Orderings in the Measurement Selection Problem. *IEEE Tr.* Vol. **SMC-47**, No. 9, 657—661, Sep., 1977.

[5] Narendra, P. M. and Funkunaga, K., A Brach and Bound Algorithm for Feature Subset Selection *IEEE Tr* Vol **C-26**. No. 9, 1977, 917—922.

[6] Devijer, P. A., Advances in Nonparametric Techniques of Statistical Pattern Classification, In Pattern Recognition Theory and Applications, Pau, L. F., the Proc. of the NATO Advanced Study Institue, D. Reidel Pub. Company, 1982, 3—18.

[7] Pearl, J, Heuristics· Intelligent Search Strategies for Computer Problem Solving. Addision-Weslv. Reading, Mass. 1984

BF* STRATEGY FOR FEATURE SELECTION AND ITS COMPARISION WITH BRANCH AND BOUND ALGORITHM

XU LEI YAN PINGFAN CHANG TONG

(QingHua University)

ABSTRACT

In this paper, the problem of feature selection is converted into the optimal pathsearching problem in a weighted directional graph. Then by means of the so called informed Best First (BF*) search strategy for problem solving in A.I., Algorithms GBFF* and TBFF* are proposed to search the optimal path, i.e., the optimal feature subset. These algorithms guarantee optimality of the selected subset without exhaustive search. In compararison with the well known Branch and Bound (B & B) algorithm, it has been shown that the number of the expanded nodes by TBFF* is less (even much less) than that by B & B: In other words. TBFF* is superior to B & B.