解释学习结果的优化排序*

郝继刚　石纯一

(清华大学计算机系　北京　100084)

摘要　解释学习可以直接应用于知识库的推理加速.经解释学习学到的一组规则,其参加匹配的先后次序直接影响着学习的效用.本文提出并证明了在以规则的成功频率与本身匹配费用的比值为序自大至小排列时,可以使这些规则的整体匹配费用最小.文中给出学习结果的优化排序算法,对于训练例的任意给定次序,可求得最小费用排列.

关键词　解释学习,效用分析,知识库加速,规则排序.

1　解释学习及其效用

解释学习 EBL(explanation—based learning)[1,2]作为一种对领域知识进行重组的分析式学习方法,可以用于改善知识库的推理效率.EBL 以原有知识库作为领域理论,借助于训练例,将目标概念(即求解目标)的推理路径压缩为充分性判断规则.其目的在于把这些规则并入原知识库后,降低目标概念在扩充的知识库上的求解费用.对于这种费用降低程度的度量,称为 EBL 的效用.

EBL 对知识库扩充的一般方式为[3]:给定一训练例,EBL 对其解释、概括出一条规则(称为宏规则).若有后续训练例未被该规则覆盖,则从其学到另一条宏规则.若还有 2 条规则都未覆盖的例子,再学习第 3 条规则,并依此进行下去.这样增量式地学到表示目标概念的、相互为析取关系的规则集合,每条规则对应于目标概念的一种特殊情形.EBL 的学习结果就是这些宏规则的集合.宏规则先于原知识库中的规则匹配,宏规则间则按生成次序先后进行匹配.如图 1 所示.

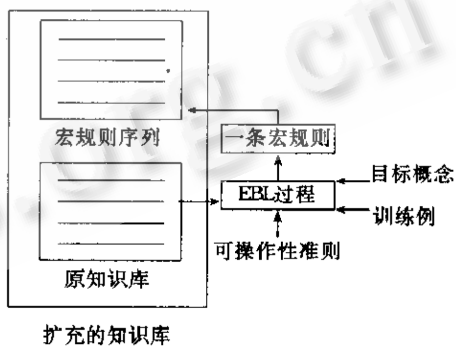

图1　EBL对知识库的扩充

宏规则能降低知识库中部分推理路径的费用,但同时其自身的匹配操作也使其他路径的推理费用增加.EBL 的效用即为因学习导致的推理费用的节省与学习结果本身匹配费用之差.文献[4]中给出 EBL 的整体效用估计公式:

---

*作者郝继刚,1967年生,博士生,主要研究领域为人工智能.石纯一,1935年生,教授,博士导师,主要研究领域为人工智能应用基础,知识工程.

本文通讯联系人:郝继刚,北京 100084,清华大学计算机系

本文 1995—01—17 收到修改稿

$$U = P * C_{DT} - C_{EBL} \tag{1}$$

其中 U 为效用，P 为目标概念可被学习结果解决的概率，C_{EBL} 为学习结果的匹配费用，C_{DT} 为目标概念被原知识库求解的费用. 且有：

$$C_{EBL} = \sum_{i=1}^{n} (p_i * \sum_{j=1}^{i} c_j) \tag{2}$$

其中 n 为宏规则数，p_i 为第 i 条宏规则成功的概率，c_i 为其匹配费用，通常以规则中谓词数表征.

及

$$P = \sum_{i=1}^{n} p_i \tag{3}$$

这里假设学习结果不存在冗余，宏规则间相互独立，即目标概念的实例至多只被一条宏规则成功识别.

我们发现，C_{EBL} 不仅与宏规则数 n，宏规则的匹配费用 c_i 和成功概率 p_i 有关，而且受规则匹配次序影响. 其中的某个次序，可以使(2)式中的 C_{EBL} 最小.

定义 1. n 条宏规则，对应着数偶集合 $\langle c_i, p_i \rangle$，其中 $c_i > 0, p_i > 0, 1 \leqslant i \leqslant n$. 令 n 组数偶的任一个排列 $\langle i_1, i_2, \cdots, i_n \rangle$ 下的目标函数为 $f(i_1, i_2, \cdots, i_n) = \Sigma_{j=1}^{n} (p_{ij} * \Sigma_{k=1}^{j} c_{ik})$，则称使 f 取最小值的排列为最小费用排列.

2 最小费用排列的确定

引理 1. 设 $\langle i_1, \cdots, i_n \rangle$ 是给定数偶集合 $\langle c_i, p_i \rangle, 1 \leqslant i \leqslant n, c_i > 0, p_i > 0$ 的任一个排列. 若存在 $s, 1 \leqslant s \leqslant n-1$，满足

$$\frac{p_s}{c_s} \leqslant \frac{p_{s+1}}{c_{s+1}} \tag{4}$$

则交换 i_s, i_s+1 后，有

$$f(i_1, i_2, \cdots, i_{s-1}, i_s, i_{s+1}, i_{s+2}, \cdots, i_n) \geqslant f(i_1, i_2, \cdots, i_{s-1}, i_{s+1}, i_s, i_{s+2}, \cdots, i_n) \tag{5}$$

且(5)式中"="成立 iff (4)式中"="成立.

证明：由定义 1，知

$$f(i_1, i_2, \cdots, i_{s-1}, i_s, i_{s+1}, i_{s+2}, \cdots, i_n) - f(i_1, i_2, \cdots, i_{s-1}, i_{s+1}, i_s, i_{s+2}, \cdots i_n)$$
$$= p_{s+1} * c_s - p_s * c_{s+1} = (p_{s+1}/c_{s+1} - p_s/c_s) * c_{s+1} * c_s \geqslant 0$$

定义 2. 数偶集合 $\langle c_i, p_i \rangle, 1 \leqslant i \leqslant n, c_i > 0, p_i > 0$ 的任意一个排列 $\langle i_1, \cdots, i_n \rangle$ 称为满足 p/c 降序条件或按 p/c 降序排列 iff$\forall\ 1 \leqslant k \leqslant n-1, p_{ik}/c_{ik} \geqslant p_{ik+1}/c_{ik+1}$.

引理 2. 设 $\langle i_1, \cdots, i_n \rangle$ 与 $\langle j_1, \cdots, j_n \rangle$ 是数偶集合 $\langle c_i, p_i \rangle, c_i > 0, p_i > 0, 1 \leqslant i \leqslant n$ 的任意 2 个满足 p/c 降序条件的排列，则 $f(i_1, \cdots, i_n) = f(j_1, \cdots, j_n)$.

证明：对数偶集合的基数 n 使用归纳法：

(1) $n=1$ 时，显然成立；

(2) 设 $n=m$ 时命题成立. 当 $n=m+1$ 时，分 2 种情况：

case 1：若 $p_{i_{m+1}} = p_{j_{m+1}}$ 且 $c_{i_{m+1}} = c_{j_{m+1}}$，

则 $\langle i_1, \cdots, i_m \rangle$ 与 $\langle j_1, \cdots, j_m \rangle$ 为同一数偶集合的按 p/c 降序排列. 由归纳假设，知 $f(i_1, \cdots, i_m) = f(j_1, \cdots, j_m)$.

$$f(i_1,\cdots,i_m,i_{m+1})=\Sigma_{i=1}^{m}(p_{ik}*\Sigma_{k=1}^{i}c_{ik})+p_{i_{m+1}}*\Sigma_{k=1}^{m+1}c_{ik}$$
$$=f(i_1,\cdots,i_m)p_{j_{m+1}}*\Sigma_{k=1}^{m+1}c_{jk}$$
$$=f(j_1,\cdots,j_m)+p_{j_{m+1}}*\Sigma_{k=1}^{m+1}c_{jk}$$
$$=f(j_1,\cdots,j_m,j_{m+1})$$

case 2:否则,必∃ $1\leqslant s\leqslant m$,$p_{i_{m+1}}=p_{j_s}$且 $c_{i_{m+1}}=c_{j_s}$,由 p/c 降序条件可知:

$$p_{i_{m+1}}/c_{i_{m+1}}=p_{j_s}/c_{j_s}=p_{j_{s+1}}/c_{j_{s+1}}=\cdots=p_{j_{m+1}}/c_{j_{m+1}}$$

由引理 1,排列$\langle j_1,\cdots,j_{m+1}\rangle$经 $j_s,j_{s+1},\cdots,j_{m+1}$相邻交换得排列$\langle j_1,\cdots,j_{s-1},j_{s+1},\cdots,j_{m+1},j_s\rangle$,仍满足 p/c 降序条件,且

$$f(j_1,\cdots,j_{m+1})=f(j_1,\cdots,j_{s-1},j_{s+1},\cdots,j_{m+1},j_s) \tag{6}$$

因 $p_{i_{m+1}}=p_{j_s}$且 $c_{i_{m+1}}=c_{j_s}$,由 case 1,知

$$f(i_1,\cdots,i_m,i_{m+1})=f(j_1,\cdots,j_{s-1},j_{s+1},\cdots,j_{m+1},j_s)$$

从而结合(6)知,$f(i_1,\cdots,i_m,i_{m+1})=f(j_1,\cdots,j_{m+1})$

因此,$n=m+1$ 时,命题成立.□

定理 1. 设$\langle i_1,\cdots,i_n\rangle$是给定数偶集合$\langle c_i,p_i\rangle$,$1\leqslant i\leqslant n$,$c_i>0$,$p_i>0$ 的任一个排列,则 $f(i_1,\cdots,i_n)$取得最小值 iff $\langle i_1,\cdots,i_n\rangle$满足 p/c 降序条件.

证明:(1)当$\langle i_1,\cdots,i_n\rangle$使 f 达到最小值.

若$\langle i_1,\cdots,i_n\rangle$不满足 p/c 降序条件,则必∃ $1\leqslant s\leqslant n-1$,$p_s/c_s<p_{s+1}/c_{s+1}$.由引理 1,交换第 s 和 $s+1$ 项后,$f(i_1,\cdots,i_{s-1},i_{s+1},i_s,i_{s+2},\cdots,i_n)<f(i_1,\cdots,i_n)$,与假设矛盾.

因此,$\langle i_1,\cdots,i_n\rangle$必满足 p/c 降序条件.

(2)设$\langle i_1,\cdots,i_n\rangle$满足 p/c 降序条件.

假设排列$\langle j_1,\cdots,j_n\rangle$使 f 达到最小值,由(1)可知,$\langle j_1,\cdots,j_n\rangle$满足 p/c 降序条件.由引理 2,有 $f(i_1,\cdots,i_n)=f(j_1,\cdots,j_n)$.所以,$\langle i_1,\cdots,i_n\rangle$也使 f 达到最小值.

推论. 给定数偶集合$\langle c_i,p_i\rangle$,$1\leqslant i\leqslant n$,$c_i>0$,$p_i>0$,求其最小费用排列,可通过排序算法得到,且复杂度最小.

证明:求最小费用排列,由定理 1,可以通过对 p_i/c_i 进行排序而得到,故复杂度至多为 $O(n*\log(n))$.同理,对 n 个正数 $c_1,c_2,\cdots,c_n$ 排序,可转化为求$(c_i,1)$,$1\leqslant i\leqslant n$ 的最小费用排列.故获得最小费用排列,其复杂度不小于排序算法.

据推论,EBL 学习结果以宏规则的 p_i/c_i 为序自大至小排列时,(2)式中 C_{EBL}取得最小值;同时 P 和 C_{DT}不变,从而(1)式中 U 取得最大值.

3　优化结果次序的 EBL 算法

根据以上结论,我们给出可以动态优化学习结果匹配次序的增量式学习算法 ORS-EBL,见图 2.算法中变量 N 为学到的宏规则个数;二维数组 A 第 i 个元素的 2 个单元中,分别存放第 i 个宏规则的成功次数和所含谓词数,反映该规则的成功概率和匹配费用.

算法中的 EBL 过程,根据训练例、知识库、目标概念和可操作性准则,将知识库中的一条推理路径压缩为宏规则,参见文献[1],记其复杂度为 $O(EBL)$.

算法将训练例与宏规则按序匹配,若成功则按 p_i/c_i 重新排序,这是由第 8~13 步完成

的.若不存在宏规则与例子成功匹配,则学习一条新的宏规则,并按序插入宏规则序列中.易知算法中排序操作的复杂度为$O(n)$,n为学得的宏规则数目;复杂度最大的操作为第5步的EBL过程;整个算法的复杂度为$O(n*O(EBL))$;当不再有新的宏规则产生时,降为$O(n*O(match))$,$O(match)$为成功匹配一条宏规则的复杂度.

输入:知识库DB,目标概念TC,可操作性准则λ,训练例集合$Example$;

输出:优化匹配次序的宏规则序列$R_1,R_2,\cdots,R_N$.

过程:(1)$N\leftarrow 0$;

(2)输入训练例$I\in Example$;

(3)$i\leftarrow 1$;

(4)若$i\leqslant N$ goto(8);

(5)调用过程$EBL(KB, TC, I, \lambda)$,从KB中学得一条宏规则R_i,含m个谓词;

(6)$N\leftarrow i$;$A_{[N][2]}\leftarrow m$;$A_{[N][1]}\leftarrow 1$;

(7)goto (11);

(8)I与第i条宏规则匹配;

(9)若失败,则$i\leftarrow i+1$; goto (4);

(10)否则,$A_{[i][1]}\leftarrow A_{[i][1]}+1$;

(11)若$i=1$,goto (2);

(12)若$A_{[i][1]}/A_{[i][2]}\leqslant A_{[i-1][1]}/A_{[i-1][2]}$,goto (2);

(13)调换第i条和第$i-1$条宏规则的次序;$t_1\leftarrow A_{[i][1]}$;$t_2\leftarrow A_{[i][2]}$;$A_{[i][1]}\leftarrow A_{[i-1][1]}$;$A_{[i][2]}\leftarrow A_{[i-1][2]}$;$A_{[i-1][1]}\leftarrow t_1$;$A_{[i-1][2]}\leftarrow t_2$;$i\leftarrow i-1$;goto (11).

图2 优化结果次序的解释学习算法ORS—EBL

4 结 论

解释学习对于知识库推理加速的贡献主要在于从中抽取出使用频率高、匹配费用小的推理路径,优先加以匹配,从而减小知识库的平均推理费用,这是一个对问题分布的适应过程.已有的工作往往着眼于将学习结果置于原知识库之前的好处,而未进一步考虑学到的宏规则相互间匹配次序对整体效用的影响.本文探讨各宏规则的匹配成功概率、费用及先后次序对整体效用的影响,确定学习结果的匹配次序原则,并给出实现算法.对于任意给定的训练例次序,都能达到最小费用排列,从而进一步加速知识库的推理.

致谢 成文过程中,得到了张旗、柳常青、贺斯敏和林成江同志的帮助,在此表示感谢!

参考文献

1 Mitchell T M, Keller R M, Kedar—Cabelli S T. Explanation—based generalization: a unifying view. J. of Machine Learning, 1986,**1**(1):47~84.

2 Dejong G F, Mooney R J. Explanation—based learning: an alternative view. J. of Machine Learning, 1986,**1**(2):145~176.

3 Minton S. Quantitative results concerning the utility of explanation—based approach. J. of AI, 1990,**42**:363~391.

4 Hao Jigang, Lu Yuchang, Shi Chunyi *et al*. Utility analysis and result organization in explanation—based learning. In: Proceedings of the 3rd Pacific—Rim Conference on Artificial Intelligence, 1994. 369~375.

THE RESULT ORGANIZATION IN EXPLANATION—BASED LEARNING

Hao Jigang Shi Chunyi

(*Department of Computer Science and Technology Tsinghua University Beijing* 100084)

Abstract EBL (explanation—based learning) can be applied to speed up reasoning on KB (knowledge bases). The match order of the macro rules, learned through EBL, affects its utility. The authors prove that, when the learned rules are ordered by the ratio of their probability of successive match to their match cost, from large to small, the overall match cost is minimal. An ordering algorithm on the learning result of EBL is presented. From all possible orders of training examples, it outputs macro rule sequence of minimal match cost.

Key words Explanation—based learning, utility analysis, knowledge base speedup, ordering of rules.