人工智能 溯因问题 溯因推理 (4)

17-20

溯因问题求解方法*)

陈保平 孙吉贵

（吉林大学计算机科学系 吉林大学符号计算与知识工程开放实验室 长春130023）

TP18

摘 要 Abduction is an important research field in AI, but it is a very difficult problem. In this paper, we first characterize the abductive explanations based on logic, and discuss the relation between minimal abductive explanations and abductive explanations. We give some methods to get all the abductive explanations or the best explanations, such as by resolution, circumscription, and AND/OR graph.

关键词 Abduction, Abductive explanation, Computing abductive explanation.

1. 引 言

溯因推理（Abduction）是已知某一结果已经发生，去寻找这个结果的原因。具体地说，给定原因集C，结果集E和背景理论T，即原因和结果之间的关系，对于观察 $d\in E$，溯因问题就是寻找合适的假设集H，满足 $H\subseteq C$，使H在背景理论T中能够解释d。

溯因问题可以定义成四元组⟨D,H,e,pl⟩，其中D是待解释的观察的集合，H是所有独立的解释的集合，e是从H的幂集到D的幂集的映射，即对 $H'\subseteq H$，e(H')可由H解释，pl(H')代表H'的信度，有关进一步的概念，请参见[1,2,4,6]，其中文[4]是国内介绍溯因推理及其应用的一篇非常好的综述性文章。

七十年代，Pople 指出了溯因在 AI 上的重要性，并同 Miller 和 Myers 实现了一个最早的溯因系统 INTERNIST-I，应用于医疗诊断领域，这个程序包含一个显式的疾病和症状的表，在病和症状之间有显式的因果关系的链，以及和链有关的概率信息[5]。

溯因推理不依赖于知识的表示和推理方法，可以使用诸如基于逻辑推理的因果关系表示，缺省推理上的概率表示等等。

2. 基于逻辑推理的溯因问题

假设 Th、h 和 f 是封闭公式，Th 是背景理论，h 是假设，f 是观察结果，我们假定 f 与 Th 相容，即 $Th\wedge f$ 可满足，且 f 不是 Th 的逻辑结果。

定义1[7] 设 h 是文字的合取。如果 $Th\wedge h\models f$，则 h 是 f 在 Th 上的溯因解。如果 $Th\wedge h$ 是可满足的，称 h 是一致的；称 h 是最小的，若 h 满足下面的条件：如果 h' 是解释，且 h' 是 h 的逻辑结果，则 h 和 h' 等价。

2.1 命题逻辑中的溯因解的描述

我们首先考虑最小解释和 prime implicant/implicate 的关系，并提供一些方法取得一致的最小解释，最后指出最小解可以覆盖所有解释。

2.1.1 最小解。根据上面的定义，h 是解释，则 $Th,h\models f$，从而 h 是解释当且仅当 $h\models(Th\rightarrow f)$。

定义2[7] 短语 C 是公式 A 的 implicant，如果 $C\models A$；子句 C 是 A 的 implicate，如果 $A\models C$；由此可知，f 在 Th 上的溯因解是 $(Th\rightarrow f)$ 的 implicant。

定义3[7] 短语 C 是公式 A 的 prime implicant，如果 $C\models A$ 且对任意短语 C'，$C'\models A$，$C\models C'$，则 $C'\models C$。子句 C 是公式 A 的 prime implicate，如果 $A\models C$ 且对任意子句 C'，$A\models C'$，$C'\models C$，则 $C\models C'$。

定理1[7] C 是 A 的 prime implicant 当且仅当 $\sim C$ 是 $\sim A$ 的 prime implicate

由上可知，最小解释是 $(Th\rightarrow f)$ 的 prime implicant。由于 $Th,h\models f$ iff $Th,\sim f\models\sim h$，则 f 在 Th 上的溯因解 h 是 $(Th\wedge\sim f)$ 的 implicate 的否定，最小解释是 $(Th\wedge\sim f)$ 的 prime implicate 的否定。

2.1.2 计算最小解。我们看规则：$\dfrac{C_1,C_2}{C_1\cup C_2-\{l,\sim l\}}$（$l$ 是 C_1 中文字，$\sim l$ 是 C_2 中文字），当 C_1,C_2 是短语时，上述规则是一致（consensus）规则，当 C_1,C_2 是子句时，上述规则是归结规则。设 A 为公式，C 为一致规则，R 为归结规则。

设 $C^i(DNF(A))$ 是对 $C^{i-1}(DNF(A))$ 实施了使用 C 规则的水平浸透法操作所得到的短语集，若存在 i 使 $C^i(DNF(A))=C^{i-1}(DNF(A))$，并且 max 是满足下述性质的短语 X 的集合：$X\models C^i(DNF(A))$，

*）国家自然科学基金，863计划资助项目

且若有其它短语 Y，Y⊨C'(DNF(A))，有 X⊨Y 不成立，则 max 是 A 的 prime implicant 的集合，记为 PI(A)。同理 IP(A)是与 C 规则相对应的使用 R 规则得到的 prime implicate 的集合，记为 IP(A)。

定理2[7]：$C_1,\cdots,C_n$ 是 DNF(A)的一个一致演绎当且仅当 $\sim C_1,\cdots,\sim C_n$ 是 DNF(A)的一个归结演绎。

由2.1.1的讨论，最小解可以通过计算 PI(A)和 IP(A)来取得，计算结果必须去掉那些与 Th 不一致的项。由于命题逻辑的满足性是可判断的，所以此问题可以解决。

定理3[7]　PI(Th→f)\PI(～Th)是 f 在 Th 上的所有最小、一致的溯因解。

上面的定理提供了一个取得溯因解的办法，但在一般情况下并不有效，主要由于 prime implicant 的数目可能很多，而且溯因的复杂性并不完全取决于 prime implicant 的数目。

2.1.3　最小解与解释。在命题逻辑中，PI(A)和 IP(A)可以看成是 A 的基底，它们具有下面的性质：

命题1[7]　C 是一个短语，则 C⊨A 当且仅当存在一个 A 的 prime implicant C'，使 C⊨C'；且 $A=\bigvee_{\{C,C\in PI(A)\}}C$。

C 是一个子句，则 A⊨C 当且仅当存在 A 的一个 prime implicate C'，使 C'⊨C；且 $A=\bigwedge_{\{C,C\in IP(A)\}}C$。

由此命题可知，如果 h'是 Th 上的一个溯因解，那么至少存在一个最小溯因解 h，使 h 是 h'的逻辑结果。

2.2　一阶逻辑中的溯因问题

在一阶逻辑中，f 在 Th 上的最小溯因解仍然可以被看成是(Th→f)的 prime implicant，但在一阶逻辑中，skolem 范式的转换会改变公式的 implicant/implicate 集，若我们只考虑纯公式(Th→f)，即 f 是纯存在公式，Th 是纯全称公式，则范式问题可以避免，可是由于一阶逻辑公式的等价和包含问题不可判定，所以不能用上面的方法产生最小一致的溯因解。

定理4[7]　一阶公式 S 不总等价于它的 prime implicant 的析取，即虽然 S 是短语 C 的逻辑结果，不一定存在 S 的 prime implicant C'，使 C⊨C'。

由上面的定理，最小解和解释之间不存在2.1中的对应关系，也就是给定一溯因解 h，但不一定存在最小解 h'，使 h⊨h'。

3. 使用归结求解溯因解

溯因解可以用扩展了的 SLD 归结来求解。首先考虑 SLD 归结，给定 definite 子句集 T 和一个目标子句 $\leftarrow G_0$(definite 子句是恰有一个正文字的 Horn 子句，目标子句是子句头为空的 Horn 子句)，一个 SLD 反驳是序列 $\leftarrow G_0,\cdots,\leftarrow G_n$，$\leftarrow G_0$是目标子句，$\leftarrow G_n$ 是空子句，每个 $\leftarrow G_{i+1}$由 $\leftarrow G_i$ 和 T 中某子句归结得到的。

若存在 $\leftarrow G_i$，其中选择的文字 g 不能与 T 中任何子句归结，通常意味着演绎失败，不做下去了，但如果我们要寻找一个单元子句集 Δ，使 T∪Δ⊨G_0，那么 Δ 中包含单元子句可与 g 归结，我们可以继续查找 $\leftarrow G_{i+1}$，$\leftarrow G_{i+1}$是 $\leftarrow G_i$删除 g 的结果。

如果找到单元子句集 Δ，使 T∪Δ⊨G_0，Δ 中仅包含可溯因文字，则 Δ 是 G_0在 T 中的溯因解。构造序列$\{\leftarrow G_0,\Delta_0\},\cdots,\{\leftarrow G_n,\Delta_n\}$，其中 $\leftarrow G_0$是目标子句，Δ_n 是可溯因文字组成的单元子句集，$\leftarrow G_n$ 是空子句，Δ_0是空集。每个$\{\leftarrow G_{i+1},\Delta_{i+1}\}$是$\{\leftarrow G_i,\Delta_i\}$使用如下规则得到的：

设 g 是从 $\leftarrow G_i$ 选择的进行归结的文字。(1)若 g 可与 T 中子句归结，那么 $\leftarrow G_{i+1}$是 $\leftarrow G_i$ 与 T 中子句以 g 为归结文字的归结式，$\Delta_{i+1}=\Delta_i$；(2)若 g 可溯因且不能与 T 中任何子句归结，则 $\leftarrow G_{i+1}$是 $\leftarrow G_i$ 删除 g 所得，Δ_{i+1}是 Δ_i，并入 g'←，g'是 g 的一个基子句。

上述方法可以用来计算溯因解，但只对 definite 子句集成立，下面我们讨论在子句体中加入负文字的方法。

定义4[8]　一个子句形如 $a\leftarrow l_1,\cdots,l_n$ 的公式，其中 $l_1,\cdots,l_n$ 是文字，a 是正文字。一个正规逻辑程序 P 是有限的子句的集合。一个目标是形如 $\leftarrow l_1,\cdots,l_n$ 的子句。

定义5[8]　一个正规程序是非循环的(acyclic)，如果对程序中任一子句的基例 $A\leftarrow B_1,\cdots B_n,\sim C_1,\cdots,\sim C_m$，存在一个原子到正整数的映射 level，使 level(B_i)<level(A)，level(C_i)<level(A)。

定义6[6]　P 是正规程序，G 是目标，一个子句是可接受的，如果子句中出现的每个个体变量都出现在头部或子句体的正文字中；一个子句是允许的，如果子句中出现的变量都出现在子句体的正文字中，即负文字中无新变量，P∪{G}是允许的，如果：(1)P 中每个子句是可接受的；(2)对于 P 或 G 的子句体中出现的正文字，定义这些文字的子句是允许的；(3)G 是允许的。

我们说 P 证明 G 当且仅当 Comp(P)⊨C，Comp(P)代表在封闭世界假定下的程序 P 的闭包。

我们在 SLD 归结中加入失败即否定(Negation as Failure)，称为 SLDNF 归结，即对于～g，如果 g 不能由当前程序推出，则～g 为真，对于 SLDNF 归结，其结果返回一个基替换 θ，使 P 证明 G'，θ 是 P∪{G}的一个回答。

定理5[6]　P 是非循环的正规程序，G 是目标，P

$\cup\{G\}$是允许的，则如果存在替换θ使$Comp(P)\models G^{\theta}$，那么$\theta$是由$P\cup\{G\}$使用SLDNF归结所得到的回答。

我们可以把溯因问题表示为$\langle P,G,\Delta\rangle$，P是一正规程序，代表背景理论，G是一目标，代表观察结果，Δ是单元子句集，是可溯因集。设P不能证明G，G在P上的溯因解定义为：$\Delta_0\subseteq\Delta$且$P\cup\Delta_0$可以证明G。一个关于P和G的SLDNF-Abduction反驳是序列$(G_0,\Delta_0,N_0),\cdots,(G_n,\Delta_n,N_n)$，其中$G_0=G$，$\Delta_0=N_0$为空集，$G_n$是空子句，每个$(G_i,\Delta_i,N_i)$由如下规则得到：对于$G_{i-1}$中选择文字g

(1)若g为正，且有P中子句C与之归结，G_i是C与G_{i-1}的归结式，$\Delta_i=\Delta_{i-1}$，$N_i=N_{i-1}$，θ_i是归结替换；

(2)若g为负，且$P\cup\Delta_{i-1}$不能证明$\sim g$，则$N_i=N_{i-1}\cup\{g\}$，$\Delta_i=\Delta_{i-1}$，G_i是G_{i-1}去掉g；

(3)若g为正，且无P中子句与之归结，g是可溯因的，如果$P\cup\Delta_{i-1}\cup N_{i-1}\cup\{g\}$一致，则$G_i$是$G_{i-1}$去掉g得到的子句，$\Delta_i=\Delta_{i-1}\cup\{g'\}$，g'是g的一个基例，$\theta_i$是其替换。

这样取得的Δ_n是G在P中的一个溯因解，$\theta=\prod_{i=1}^{n}\theta_i$是得到的回答。

4. 基于费用的溯因问题

由于溯因解往往很多，求取全部溯因解是很困难的，而且有时也不必要，所以溯因也可以看成是求取最优解的问题。很多时候，即使是最不可能的解释也可以推导出证据，这就要求给出解释的度量。传统的方法是使解释中的元素数最少，Hobbs和Stickle提供了加权溯因[9]，对每个独立的假设赋一个数字的费用，选择费用最小的解释。Charniak的基于费用的溯因[10]是前一种方法的变形，可以用WAODAG图(weighted AND/OR directed acyclic graph)来描述。

定义7[11] 一个WAODAG图是四元组$\langle G,c,r,S\rangle$，其中：①G是有向非循环图，$G=(V,E)$，V是点集，E是边集；②c是从$V\wedge\{true,false\}$到实数集的函数，叫费用函数；③r是从V到{and,or}的函数，叫标记；④S是子集，叫证据节点。V中入度为0的节点叫做假设节点，记为V_H。

定义8[11] 对WAODAG图的一个真值指定是从V到{true,false}的一个函数e，此函数有效当且仅当①对与节点q，e(q)=true iff q的所有父节点为真；②对或节点q，e(q)=ture iff 存在q的父节点为真。

如果e有效且使S中节点的真值为true，则称e是一个解释，实际上是将V_H看成是原因集，S看成是结果集，把背景理论转化为WAODAG图，利用图的特性求解的一种方法。

费用$C(e)$是由如下公式计算的：$C(e)=\sum_{q\in V}c(q,e(q))$。使C最小的解释e叫最优解，于是溯因问题的求解转化成了这种约束可满足问题的求解。

5. 基于限定方法的溯因解计算

本节提供一个使用限定和一致性方法来计算溯因解的算法，详细内容见文[13]。我们考虑这样一类溯因问题$\langle C,E,T\rangle$，其中C是原子集，叫做原因集；E是原子集，叫做结果集；T是任意子句集叫做背景理论。一个观察$g(g\in E)$的溯因解定义如下：①$A\subseteq C$，且A与T一致；②$T\cup A\models g$；③A是满足上述条件的最小集合。

定义9[12] 三元组$\langle C,E,T\rangle$定义同上，观察g是E的元素。关于g的否定集D定义如下：(1)$D\subseteq C$；(2)$T\cup\{g\}\cup\{\sim d\mid d\in D\}$一致。

当否定集最大时，即$T\cup\{g\}\cup\{\sim d\mid d\in D\}\models C\text{-}D$，我们把C-D叫做关于g的一致性解释。

对于这一类溯因问题，我们做如下两条假设：

(1)完备性公理：若$A_1,\cdots,A_n$是g在T中的所有溯因解，则对于任意公式S，$S\vdash g$在T中成立，存在A_i，使$A_i\vdash S$在T中成立。即g在T中成立当且仅当存在某个A_i在T成立。

(2)正原因有效：可溯因集(即原因集C)是正原子集，且任意C中原子的负出现是无效的。也就是说，对任意$c_1,\cdots,c_n,\sim b_1,\cdots,\sim b_m\vdash g$，$c_i$，$b_i$是T中原子，则$c_1,\cdots,c_n\vdash g$。

完备性公理是指对任意观察结果，总可以在原因集中得到解释，正原因有效是指一些观察是因为某些原因发生而产生的，而不是由于某些原因不发生而产生的，因此上述两个假设是合理的。

定理6 A是g在T中的溯因解，当且仅当A是g在$CIRC(T;E;Q)$中的溯因解。其中$CIRC(T;E;Q)$是T的限定，Q是T中C、E以外的谓词集合。

定理7 对于溯因问题$\langle C,E,T\rangle$，C、E、T、Q同上，且满足上面的两个假设；$T\nvdash\vee(L_i)$(其中L_i是C中的文字)，则A是g在T中的某个独立的溯因解当且仅当A是g在$CIRC(T;E;Q)$中的某个一致性解释。

对于非循环的definite子句集，我们给出计算溯因解的算法。

算法1 (计算T的溯因解)

(1)计算$CIRC(T;E;Q)$；

(2)若$CIRC(T;E;Q)\cup\{g\}$不一致，则无解；

(3)否则计算$CIRC(T;E;Q)$的一致性解释，其一致性解释就是溯因解。

算法2 (计算$CIRC(T;E;Q)$的一致性解释)

```
D:={};N:={};S:={};
while(S 不为空)
 {从 S 中选取文字 L;
 S:=S-{L};
 if({~L}∪CIRC(T;E;Q)∪{g}一致)
  then D:=D∪{L};
  else N:=N∪{L};
}
```

此算法结束后,N 就是 CIRC(T;E;Q)的一致性解释,也就是 T 的溯因解。

6. 结束语

通常情况下,求解溯因问题是非常困难的,比如说对于给定的观察和背景理论,溯因解往往会有很多,其中很多是不一致的.在许多情形下,求解溯因问题是 NP 难度的。

如果一个假设解释了一个数据当且仅当假设中某一元素解释了这个数据,这类溯因问题称为是独立的;比独立的溯因问题更普遍的一类溯因问题是多个假设可以解释一些不能由单个假设解释的数据,且一个假设不会丢失任何可以由它的子集解释的数据,即组合假设可以解释更多的数据,称此类问题为单调的溯因问题;上面两种情况都认为任意假设集是可能的,但在许多领域不是这样的,H 的某些子集可能不一致,这就是不相容的溯因问题,不相容的溯因问题要比独立的和单调的溯因问题复杂得多,下表总结了上面几类问题的计算复杂性.

	寻找所有解	寻找一个解	寻找最优解
独立的	NP	P	?
单调的	NP	P	?
不相容的	NP	NP	NP

P 代表存在多项式级的算法,NP 代表 NP 难度的,?是还没有结论。

由此可见,溯因解的计算是一件很困难的事情。但对于具体的实际问题,依赖于实际背景,给出有效的算法也是可能的。溯因推理的研究与人工智能的许多领域密切相关,如 Reiter 的诊断理论,Pearl 的信念修正问题等都可以转化为溯因问题进行研究[6],因此,溯因问题求解方法的研究成果将会有力地推动相关领域的研究,并将更广泛地应用于实际问题之中。

参考文献

[1]D. Poole, Explanation and prediction: an architecture for default and abductive reasoning, Comput. Intell. 5(1989)

[2]J. R. Hobbs 等. Interpretion as abduction, AI, 63 (1993)

[3]Raedt L. D., Belief update from integrity constraints and queries, AI, 53(1992)

[4]马绍汉、张宏,诱导推理及其应用,计算机科学 1993 Vol 20. №. 5

[5]R. A. Myller 等, INTERNIST-I: an expermental computer-based diagnostic consultant for general internal medicine, New England J, Med, 307 (1982)

[6]T. Bylander 等, The computational complexity of abduction, AI, 49(1991)

[7]P. Marquis, Extending Abduction from Propositional to First-Order Logic, Internal Workshop FAIR'91

[8]L. Cavedon, Acyclic logic program and the completeness of SLDNF-resolusion, Theorical computer science 86(1991)

[9]D. E. Appelt, A theory of abduction based on model preference, Proc. AAAI Symposium on abduction

[10]E. Charick 等, Probabilistic semantics for cost-based abduction, Proc. AAAI-90, (1990)

[11]E. Santos Jr, A liner constraint satisfaction approach to cost-based abduction, 同[3]

[12]Konolige, Abduction versus closure in causal theory, 同[3]

[13]陈保平,孙吉贵,刘叙华,一类溯因问题的求解,(待发表)

[14]M. Shanahan, Prediction is deduction but explanation is abduction, IJCAI-89

更正 我刊1996No.1所发表论文"图形化对象式需求定义语言 NDORL"中,p55的"聚合关系"、"子类关系"和"状态"之图形形式等价表示的插图,实有错位和遗漏,请读者参考 p50 § 3.2图符集中关于〈聚合关系图符〉、〈子类关系图符〉和〈状态图符〉之定义。在此,特向作者和广大读者深表歉意。