

刻画基于模型的中心诊断*

欧阳彤¹ 姜云飞²

¹(吉林大学计算机科学系 长春 130023)

²(中山大学计算机软件所 广州 510275)

E-mail: jtuao@mail.interilow.com.cn

摘要 虽然对基于模型的诊断存在一系列不同的逻辑定义,但所幸的是存在一个统一的抽象定义,它概括了以往的不同定义.在该定义基础上提出了基于模型的中心诊断的概念.通过刻画基于模型的中心诊断过程,论证了基于模型的中心诊断与本原蕴含/蕴含式的直接关系,从而将其理论结果与 ATMS (assumption-based truth maintenance system) 这类算法联系起来.进一步指出,对基于一致性中心诊断的刻画仅仅是文中所给出的刻画的一个特殊情形.

关键词 基于模型的诊断,基于模型的中心诊断,本原蕴含/蕴含式.

中图法分类号 TP18

使用传统的专家系统方法建造起来的诊断系统依赖于具体的设备,而且如何从专家那里获取必要的诊断知识是维护和发展这些系统的瓶颈.基于模型的诊断是一项新的 AI 技术,它可以克服传统方法的缺点.

基于模型的诊断可以看成是在已知上下文中对观察集产生“解释”(Explanations)的过程.但是,在已有文献中至少存在两个不同的关于“解释”的逻辑概念.一种是解释的弱概念.在这种情形下,称一诊断解释一观察 m ,若它与 m 相一致,参见文献[1~4].另一种是解释的强概念.在这种情形下,称一诊断解释一观察 m ,若它逻辑推出 m ,参见文献[5,6].显然,逻辑推出比一致性要求强得多.那么给出一个观察集,应选择哪些观察必须被诊断逻辑推出呢?这是一个引起争议的地方.对需要被逻辑推出的观察子集的不同选择产生了基于模型诊断的不同逻辑定义.比如,文献[7]认为所有的观察必须被诊断逻辑推出;而文献[8,9]则认为一诊断只与所有观察相一致即可,不要求任何观察被诊断逻辑推出.

Console L 和 Torasso P^[10]分析了已有文献中出现的基于模型诊断的不同逻辑定义,提出了一个统一的定义.本文即在该定义下,进一步给出了能够刻画基于模型诊断空间的基于模型的中心诊断的概念,并对它的诊断过程进行了刻画.该刻画论证了基于模型的中心诊断与本原蕴含/蕴含式的直接关系,从而可使用 ATMS^[11] (assumption-based truth maintenance system) 算法来计算,因而将我们的理论结果与实现联系起来.

1 基于模型的诊断问题

定义 1. 基于模型的诊断问题 DP 是一个三元组 $\langle \langle SD, COMPS \rangle, CXT, OBS \rangle$, 其中 SD 为系统描述,是一阶语句的集合; $COMPS$ 为系统组成部件,是一个有限的常量集; CXT 为表示上下文数据的基原子集合, OBS 为待解释的观察集的基原子集合.

CXT 或 OBS 中每一原子 $f(a)$ 的含义为:关于参数 f 的观察值为 a .

定义 2. 已知一诊断问题 $DP = \langle \langle SD, COMPS \rangle, CXT, OBS \rangle$, 相应于 DP 的溯因问题 AP 是一个三元组

* 本文研究得到国家自然科学基金和国家教委博士点基金资助.作者欧阳彤,女,1968年生,博士生,讲师,主要研究领域为基于模型的诊断,定理机器证明.姜云飞,1945年生,教授,博士生导师,主要研究领域为自动推理,基于模型均诊断及规划.

本文通讯联系人:欧阳彤,长春 130023,吉林大学计算机科学系

本文 1997-12-01 收到原稿,1998-01-23 收到修改稿

$\langle \langle SD, COMPS \rangle, CXT, \langle \Psi^+, \Psi^- \rangle \rangle$, 其中 $\Psi^+ \subseteq OBS, \Psi^- = \{ \sim f(x) \mid f(y) \in OBS, \text{对于 } f \text{ 的每一非 } y \text{ 的允许值 } x \}$.

一元谓词 AB 意味着“abnormal” (反常), 当部件 $c \in COMPS$ 反常时, $AB(c)$ 为真.

定义 3. 已知两个部件集 C_p, C_n , 定义 $D(C_p, C_n)$ 如下: $[\bigwedge_{c \in C_p} AB(c)] \wedge [\bigwedge_{c \in C_n} \sim AB(c)]$.

定义 4. 已知一诊断问题 DP 以及相应于 DP 的溯因问题 AP 如定义 2, 称 $D(\Delta, COMPS - \Delta)$ 为关于 DP 的一个基于模型的诊断, 当且仅当

- (1) $D(\Delta, COMPS - \Delta)$ 逻辑推出 Ψ^- , 即对每一 $m \in \Psi^+$, 都有 $SDUCXTU\{D(\Delta, COMPS - \Delta)\} - m$.
- (2) $D(\Delta, COMPS - \Delta)$ 与 Ψ^- 一致, 即 $SDUCXTU\{D(\Delta, COMPS - \Delta)\} \cup \Psi^-$ 可满足.

原则上, Ψ^+ 可为 OBS 的任一子集. 若 OBS 形如 $\{O_1, \dots, O_n\}$, $\Psi^+ \subseteq OBS$, 且形如 $\{O_{11}, \dots, O_{1r}\}$. 下面用 $\sim OBS$ 表示 $\{\sim O_1 \vee \dots \vee \sim O_n\}$, $\sim \Psi^+$ 表示 $\{\sim O_{11} \vee \dots \vee \sim O_{1r}\}$.

由定义 4 不难得到如下定理.

定理 1. 已知 DP 和 AP 如上, 若存在关于 DP 的基于模型的诊断, 则 $SDUCXTUOBS$ 是可满足的, 且 $SDUCXTU \sim OBS$ 也是可满足的, 特别地, $SDUCXTU \sim \Psi^+$ 也是可满足的.

2 刻画基于模型的中心诊断

2.1 基于模型的中心诊断

定义 5. 称一些文字的合取式 C 覆盖另一些文字的合取式 D 当且仅当 C 的每一个文字都出现在 D 中.

定义 6. 已知 DP 和 AP 如上, 关于 DP 的基于模型的部分诊断为一可满足的 $D(\Delta_1, \Delta_2)$, 使得对每一被 $D(\Delta_1, \Delta_2)$ 覆盖的可满足的 $D(\Delta'_1, \Delta'_2)$, 都有: (1) $SDUCXTU\Psi^- \cup \{D(\Delta'_1, \Delta'_2)\}$ 一致, (2) 若 $\Psi^- \neq \emptyset$, 则 $SDUCXTU\{D(\Delta'_1, \Delta'_2)\} \vdash \Psi^+$ (注意, $SDUCXTU\{D(\Delta'_1, \Delta'_2)\} \cup \Psi^-$ 一致等价于 $SDUCXTU\{D(\Delta'_1, \Delta'_2)\} \cup OBS$ 可满足).

基于一致性部分诊断与部分溯因诊断是该定义两种特殊情形. 前者相当于在该框架中取 $\Psi^+ = \emptyset$. 后者相当于在该框架中取 $\Psi^+ = OBS$. 由该定义可直接得到如下引理.

引理 1. 已知 DP 如上, 关于 DP 的每一个基于模型的部分诊断都是关于 DP 的基于一致性的部分诊断. 反过来不成立.

定义 7. 基于模型的中心诊断是具有如下性质的基于模型的部分诊断: 即能覆盖它的唯一的基于模型的部分诊断就是它本身.

不难得出如下结论: $D(\Delta, COMPS - \Delta)$ 为一基于模型的诊断, 当且仅当存在一个基于模型的中心诊断来覆盖它. 也就是说, 使用基于模型的中心诊断可以刻画基于模型的诊断空间. 那么, 如何计算基于模型的中心诊断呢? 在此之前, 先回顾一下与本文相关的一些工作.

定义 8. AB 文字为 $AB(c)$ 或 $\sim AB(c), c \in COMPS$. AB 子句为 AB 文字的析取式.

定义 9. 关于 $\langle \langle SD, COMPS \rangle, CXT, OBS \rangle$ 的一个冲突是从 $SDUCXTUOBS$ 出发推出的一个 AB 子句. 称一冲突为最小的, 当且仅当该冲突的任一真子子句都不是一冲突.

定义 10. 设 Σ 是命题公式的集合, 称一可满足的基文字合取式 Π (即不含互补对的文字的合取式) 是 Σ 的一个蕴含 (Implicant) 当且仅当 Π 能推出 Σ 中的每一个语句. 称 Π 是 Σ 的一个本原蕴含 (Prime Implicant) 当且仅当能覆盖 Π 的唯一的 Σ 的蕴含是 Π 本身.

定义 11. 设 Σ 是命题公式的集合, 称基文字的析取式 C 为 Σ 的蕴含式 (Implicate) 当且仅当 Σ 能推出 C . 称 C 为 Σ 的本原蕴含式 (Prime Implicate) 当且仅当不存在 C 的真子子句 C' , 使得 C' 为 Σ 的蕴含式.

定理 2^[12] (刻画基于一致性的中心诊断). 关于 $\langle \langle SD, COMPS \rangle, CXT, OBS \rangle$ 的基于一致性的中心诊断是 $SDUCXTUOBS$ 的所有最小冲突的本原蕴含.

2.2 刻画基于模型的中心诊断

引理 2^[12] 关于 DP 的基于一致性部分诊断是关于 DP 的所有最小冲突的蕴含.

下面,我们用 \prod 表示 $SDUCXTUOBS$ 的所有形如 AB 子句形式的本原蕴含式的合取式,形如 $\sim D(\Delta_1, \Delta_j) \wedge \dots \wedge \sim D(\Delta_m, \Delta_m)$, 对其中每一 $\sim D(\Delta_p, \Delta_p)$, 都有 $SDUCXTUOBS \vdash \sim D(\Delta_p, \Delta_p)$; 用 Ω 表示关于 $SDUCXTU \sim \Psi^+$ 的所有形如 AB 子句形式的本原蕴含式的合取式,形如 $\sim D(\Delta_1, \Delta_1) \wedge \dots \wedge \sim D(\Delta_m, \Delta_m)$, 对其中每一 $\sim D(\Delta_q, \Delta_q)$, 都有 $SDUCXTU \sim \Psi^+ \vdash \sim D(\Delta_q, \Delta_q)$, 特别地, 若 $\Psi^- = \emptyset$, 则令 $\Omega = \text{false}$.

定理 3(刻画基于模型的中心诊断). 已知 DP 及 AP 如上, 设 \prod 为 $SDUCXTUOBS$ 的所有 AB 子句形式的本原蕴含式(即关于 DP 的所有最小冲突)的合取式, Ω 为 $SDUCXTU \sim \Psi^+$ 的所有 AB 子句形式的本原蕴含式的合取式, 则 $D(\Delta_1, \Delta_2)$ 为关于 DP 的一个基于模型的中心诊断, 当且仅当 $D(\Delta_1, \Delta_2)$ 为 $\prod \wedge \sim \Omega$ 的一个本原蕴含.

证明: 必要性. 设 $D(\Delta_1, \Delta_2)$ 是关于 DP 的一个基于模型的中心诊断. 首先证明 $D(\Delta_1, \Delta_2) \vdash \prod \wedge \sim \Omega$. 由引理 1 知, $D(\Delta_1, \Delta_2)$ 也是关于 DP 的基于一致性部分诊断. 因此, 由引理 2, $D(\Delta_1, \Delta_2) \vdash \prod$. 现在证明 $D(\Delta_1, \Delta_2) \vdash \sim \Omega$, 用反证法. 若不然, 则对 $\sim \Omega$ 中的任意一个 $D(\Delta_p, \Delta_p)$, 都有 $D(\Delta_1, \Delta_2) \not\vdash D(\Delta_p, \Delta_p)$. 又由 $D(\Delta_1, \Delta_2)$ 是基于模型的中心诊断知, $D(\Delta_1, \Delta_2)$ 亦为基于模型的部分诊断. 故由定义 6 知, $SDUCXTU \{D(\Delta_1, \Delta_2)\} \vdash \Psi^+$, 即 $SDUCXTU \sim \Psi^+ \vdash \sim D(\Delta_1, \Delta_2)$. 这就是说, $\sim D(\Delta_1, \Delta_2)$ 为 $SDUCXTU \sim \Psi^-$ 的一个蕴含式. 因此, Ω 中一定存在某一本原蕴含式 $\sim D(\Delta_q, \Delta_q)$, 使得 $\sim D(\Delta_q, \Delta_q)$ 为 $\sim D(\Delta_1, \Delta_2)$ 子析取式. 亦即 $\sim \Omega$ 中存在某一 $D(\Delta_q, \Delta_q)$, 使得 $D(\Delta_q, \Delta_q)$ 为 $D(\Delta_1, \Delta_2)$ 的子合取式. 当然就有 $D(\Delta_1, \Delta_2) \vdash D(\Delta_q, \Delta_q)$. 这样就与上面的对于 $\sim \Omega$ 中任意一个 $D(\Delta_p, \Delta_p)$, 都有 $D(\Delta_1, \Delta_2) \not\vdash D(\Delta_p, \Delta_p)$ 产生了矛盾. 故原假设 $D(\Delta_1, \Delta_2) \not\vdash \sim \Omega$ 不对, 即 $D(\Delta_1, \Delta_2) \vdash \sim \Omega$. 综上, 有 $D(\Delta_1, \Delta_2) \vdash \prod \wedge \sim \Omega$. 现在假设存在一个 $D(\Delta_1', \Delta_2')$, 它覆盖 $D(\Delta_1, \Delta_2)$ 且是 $\prod \wedge \sim \Omega$ 的蕴含, 则 $D(\Delta_1', \Delta_2') \vdash \prod \wedge \sim \Omega$. 由 $D(\Delta_1', \Delta_2') \vdash \prod$ 及引理 2 知, $D(\Delta_1', \Delta_2')$ 是关于 DP 的基于一致性部分诊断. 由定义 6 知, $SDUCXTU \{D(\Delta_1', \Delta_2')\} \cup OBS$ 是可满足的. 因为 $SDUCXTU \sim \Psi^+ \vdash \Omega$, 因此, $SDUCXTU \{\sim \Omega\} \vdash \Psi^+$, 又由 $D(\Delta_1', \Delta_2') \vdash \sim \Omega$, 故有 $SDUCXTU \{D(\Delta_1', \Delta_2')\} \vdash \Psi^+$. 因而, 由定义 6 知, $D(\Delta_1', \Delta_2')$ 是基于模型的部分诊断. 又因为它覆盖了基于模型的中心诊断 $D(\Delta_1, \Delta_2)$, 所以, $D(\Delta_1', \Delta_2')$ 就是 $D(\Delta_1, \Delta_2)$. 这就是说, $D(\Delta_1, \Delta_2)$ 是 $\prod \wedge \sim \Omega$ 的本原蕴含.

充分性. 假设 $D(\Delta_1, \Delta_2)$ 是 $\prod \wedge \sim \Omega$ 的一个本原蕴含. 由上面证明知, 它一定是基于模型的部分诊断. 设存在一个 $D(\Delta_1', \Delta_2')$, 它覆盖了 $D(\Delta_1, \Delta_2)$ 且是基于模型的部分诊断, 那么由必要性部分证明知, $D(\Delta_1', \Delta_2')$ 一定是 $\prod \wedge \sim \Omega$ 的蕴含. 又因为它覆盖了 $\prod \wedge \sim \Omega$ 的本原蕴含, 所以, $D(\Delta_1', \Delta_2')$ 就是 $D(\Delta_1, \Delta_2)$. 这就是说, $D(\Delta_1, \Delta_2)$ 是基于模型的中心诊断. □

可以看出, 该定理有如下优点: (1) 将求基于模型的中心诊断的过程清晰地分为与领域有关的冲突识别及与领域无关的候选产生两大步骤. (2) 当 Ψ^+ 取作 \emptyset 时, 可直接得到文献[12]中对基于一致性中心诊断的刻画, 这就是说, 它仅仅是该定理的一个特殊情形. 直观上看, 基于模型的中心诊断是在基于一致性中心诊断上附加了最小数目的 AB 文字, 保证了它不仅与 OBS 一致, 而且能逻辑推出 Ψ^+ . 这样, 在实际应用中, 我们可以首先从 \prod 计算出限制弱的基于一致性中心诊断空间, 然后通过逐步增加需被逻辑推出的观察数目(扩大 Ψ^-), 逐步在该诊断上附加最小数目的 AB 文字, 以便找出真正有用的解. (3) 尤其重要的是, 我们证明了基于模型的中心诊断与本原蕴含和本原蕴含式的直接关系, 从而将我们的理论结果与 ATMS 类算法联系起来.

3 实例

考虑汽车发动机的一个部件油杯的简化行为模型. 设该部件有两种行为方式 correct 及 holed(相应于 $\sim AB$ ($oil-cup$) 及 $AB(oil-cup)$), 它的模型 SD 包括如下公式:

- $holed(oil-cup) \rightarrow oil\ below\ car(present)$
- $holed(oil-cup) \rightarrow oil-level(low)$
- $correct(oil-cup) \rightarrow oil-level(normal)$
- $oil-level(low) \wedge engine(on) \rightarrow engine-temperature(high)$
- $oil-level(normal) \wedge engine(on) \rightarrow engine-temperature(normal)$

下面考察如下诊断问题. $DP = \langle \langle SD, COMPS \rangle, CXT, OBS \rangle$. 其中 $COMPS = \{oil-cup\}$, $OBS = \{oil-below-car(present)\}$, $CXT = \{engine(on)\}$. 设与 DP 相应的溯因问题 $AP = \langle \langle SD, COMPS \rangle, CXT, \langle \Psi^+, \Psi^- \rangle \rangle$, 其中 $\Psi^- = \{\sim oil-below-car(absent)\}$.

若该模型是完备的(完备是指每一观察的所有原因都在 SD 中), 则取 $\Psi^+ = \{oil\ below-car(present)\}$. 应用定理 3, 第 1 步: 求出 \prod 为 $holed(oil-cup) \vee correct(oil-cup)$, Ω 为 $\sim holed(oil-cup)$. 第 2 步: 求出 $\prod \wedge \sim \Omega$ 的本原蕴含为 $holed(oil-cup)$. 因此, 基于模型的中心诊断为 $holed(oil-cup)$. 这一解是合理的. 因为在 $oil-below-car(present)$ 的原因的完备模型中, $holed(oil-cup)$ 是 $oil-below-car(present)$ 的唯一原因.

若该模型是不完备的, 即 $holed(oil-cup)$ 不是 $oil-below-car(present)$ 的唯一原因, 还有与部件 $oil-cup$ 状态不相关的额外原因可能导致 $oil-below-car(present)$, 这时, 只给出唯一诊断 $holed(oil-cup)$ 就不合理了. 因此, 这时若取 $\Psi^+ = \emptyset$, 以削弱限制, 则可从 \prod 求出两个中心诊断 $correct(oil\ cup)$ 及 $holed(oil-cup)$, 这显然是合理的.

4 结 论

本文提出了一个能够刻画基于模型诊断空间的基于模型的中心诊断的概念, 并对基于模型的中心诊断过程进行了刻画. 我们将该诊断过程清晰地分为与领域有关的冲突识别及与领域无关的候选产生这两大步骤, 并论证了基于模型的中心诊断与本原蕴含/蕴含式的直接关系, 从而将我们的理论结果与实现联系起来.

参 考 文 献

- 1 Davis R. Diagnostic reasoning based on structure and behavior. *Artificial Intelligence*, 1984, 24(1-3): 347~410
- 2 Genesereth M R. The use of design descriptions in automated diagnosis. *Artificial Intelligence*, 1984, 24(1-3): 411~436
- 3 Reiter R. A theory of diagnosis from first principles. *Artificial Intelligence*, 1987, 32(1): 57~96
- 4 de Kleer J, Williams B C. Diagnosing multiple faults. *Artificial Intelligence*, 1987, 32(1): 97~130
- 5 Poole D, Goebel R, Aleliunas R. THEORIST: a logical reasoning system for defaults and diagnosis. In: Cercone M, Calla G MC eds. *The Knowledge Frontier*. New York: Springer Verlag, 1987. 331~352
- 6 Console L, Dupré D T, Torasso P. A theory of diagnosis for incomplete causal models. In: Sridharan N S ed. *Proceedings of the 11th International Joint Conference on Artificial Intelligence*. San Mateo, CA: Morgan-Kaufmann Publishers, 1989. 1311~1317
- 7 Poole D. Normality and faults in logic based diagnosis. In: Sridharan N S ed. *Proceedings of the 11th International Joint Conference on Artificial Intelligence*. San Mateo, CA: Morgan-Kaufmann Publishers, 1989. 1304~1310
- 8 de Kleer J, Williams B C. Diagnosis with behavioral modes. In: Sridharan N S ed. *Proceedings of the 11th International Joint Conference on Artificial Intelligence*. San Mateo, CA: Morgan Kaufmann Publishers, 1989. 1324~1330
- 9 Struss P, Dressler O. Physical negation—integrating fault models into the general diagnostic engine. In: Sridharan N S ed. *Proceedings of the 11th International Joint Conference on Artificial Intelligence*. San Mateo, CA: Morgan-Kaufmann Publishers, 1989. 1318~1323
- 10 Console L, Torasso P. A spectrum of logical definitions of model-based diagnosis. *Computational Intelligence*, 1991, 7(3): 133~141
- 11 de Kleer J. An assumption-based tms. *Artificial Intelligence*, 1986, 28(2): 127~162
- 12 de Kleer J, Mackworth A K, Reiter R. Characterizing diagnosis and systems. *Artificial Intelligence*, 1992, 56(2-3): 197~222

Characterizing Kernel Model-based Diagnosis

OU-YANG Dan-tong¹ JIANG Yun-fei²

¹(Department of Computer Science Jilin University Changchun 130023)

²(Computer Software Institute Zhongshan University Guangzhou 510275)

Abstract Although there is a spectrum of different logical definitions of model based diagnosis, fortunately, there exists a unified abstract definition which generalizes the definitions proposed in the past. Based on this definition, the concept of kernel model-based diagnosis is proposed. Then by characterizing the process of kernel model-based diagnosis, the direct relationship between it and the prime implicants/implicates is also demonstrated. Therefore, the theoretical results in this paper can be linked with algorithms such as the ATMS (assumption-based truth maintenance system). Furthermore, it is pointed out that the characterization of kernel consistency-based diagnosis is only a special case of the characterization in this paper.

Key words Model-based diagnosis, kernel model-based diagnosis, prime implicant/implicate.