

4. 进一步讨论

下面我们将利用前面的结果来求反合一子的完备集。

由推论3.2可知反合一子 σ 与最一般合一子 θ 的关系,即:

$$\forall \rho_1 \rho_2 (\theta \circ \rho_1(Z(s,t)) \neq \sigma \circ \rho_2(Z(s,t)))$$

其中 θ 是可求的,现在要求的是 σ ,可麻烦在于 ρ_1, ρ_2 都是任意的,不等式两边都不确定,如何去求 σ 并保证尽管 ρ_1, ρ_2 随意变动但不等式始终成立是我们面临的问题。

由于在讨论替换间的关系以及构造新替换时,直接拿替换来讨论不如把它先转化为项的形式来讨论方便,所以在前文我们引入了与 θ 对应的函数 G_θ ,由定理3.3和定理3.4我们可以把 θ 与 σ 的关系转化成 G_θ 与 G_σ 的关系,这样一来我们希望通过构造 G_σ ,使得 $\forall \rho_1 \rho_2 (\rho_1 G_\theta \neq \rho_2 G_\sigma)$,也就使得 $\forall \rho_1 \rho_2 (\theta \circ \rho_1(Z(s,t)) \neq \sigma \circ \rho_2(Z(s,t)))$,从而得到与 G_θ 对应的 σ , σ 就是反合一子。

下面我们将通过一个小小的转化来摆脱 ρ_1 和 ρ_2 随意变动所带来的麻烦。

考察一下合一子就可以看出,最一般合一子 θ 相当于一个模式,其它的合一子只是在其上继续代入罢了,假如我们能找到另一些完全不同于最一般合一子 θ 模式的替换,则这些替换必为反合一子,这也正是定理3.1和推论3.2的含义。

既然最一般合一子 θ 是一种模式,我们就把它

对应的 G_θ 中的所有变量改写成一些完全不同于项集 T 的符号系统 S 的符号,如: S 中含有变量 ∇, \diamond ,我们可以把 $G_\theta(f(x), y)$ 改写成 $G_\theta'(f(\nabla), \diamond)$,而在构造其它替换 σ 所对应的 G_σ 时仍用项集 T 的符号系统,这样一来 $\text{Var}(G_\sigma') \cap \text{Var}(G_\theta) = \emptyset$ 。

我们只要使 G_θ 与 G_σ' 不能合一(合一过程是在符号系统 $T \cup S$ 下进行的),借助于定理3.5可得 $\forall \rho_1 \rho_2 (\rho_1 G_\theta' \neq \rho_2 G_\sigma)$,这样 ρ_1 和 ρ_2 随意变动所带来的问题就解决了,不过到目前为止,构造 G_σ 只找到了排列的方法,即在 G_θ 的所有位置代入项的所有可能情形,并与 G_σ' 进行合一检测,留下与 G_σ' 不可合一的 G_σ ,所有 G_σ 对应的 σ 就构成反合一子的完备集。

结束语 本文是以一阶项和字面合一为基础来讨论反合一问题的,至于以高阶项和方程式合一为基础来探讨反合一还是相当复杂的,有关这类问题我们将在以后的工作中进一步探索。

参考文献

- [1] J oRGH Siekmann, Unification Theory, J. Symbolic Computation vol. 7, 1989
- [2] John C. Mitchell, Foundations for Programming Languages, The MIT Press Cambridge, Massachusetts, London, England, 1996
- [3] Jean H. Galler et al, Complete sets of transformations for general E-unification, Theorem Computer Science, 67, 1989
- [4] 许满武,戴劲雯等,合一问题研究,南京大学计算机学术报告,1995
- [5] J. de Kleer, Local methods for localizing faults in electronic circuits, MIT Artificial Intelligence Laboratory, 1976
- [6] D. Sleeman, et al, Intelligent tutoring systems, Academic Press, 1982
- [7] R. Patil, et al, Causal understanding of patient illness in medical diagnosis, IJCAI, 1981
- [8] R. Reiter, A theory of first principles, Artificial Intelligence, 32(1)1987
- [9] J. de Kleer et al, Diagnosing multiple faults, Artificial Intelligence, 32(1)1987
- [10] D. Poole, Normality and faults in logic-based diagnosis, IJCAI, 1989
- [11] T. Bylander, The computational complexity of abduction, Artificial Intelligence, 49(1-3), 1991
- [12] K. Konolige, abduction versus closure in causal theories, Artificial Intelligence, 1992, 53
- [13] O. Dressler, et al, Model-based diagnosis with the default-based diagnostic engine: effective control strategies that works in practice, ECAI, 1994
- [14] L. Console, et al, Model-based Diagnoses, Science Publishers, 1994
- [15] K. R. Apt, et al, Meta-logic and logic programming, Morgan Kaufmann Publisher, 1995
- [16] W. Nejdl, et al, A formal framework for representing diagnosis strategies in model-based diagnostic system, IJCAI, 1995
- [17] F. V. Harmelen, Using domain knowledge to select solution in abductive diagnosis, ECAI, 1994
- [18] M. P. Feret, et al, A formal model for experience-aided diagnosis, Computational intelligence, 13(2)1997
- [19] J. I. J. Orlando, Future diagnostic technology, Expert System with Application, 1996, 11, 2

人工智能 模型诊断 诊断推理 (13)

计算机科学1998Vol. 25No. 6

54-57,3 | 基于模型诊断推理的回顾与展望*

An Retrospect and Prospect on Model-based Diagnostic Reasoning

李占山^{1,2} 姜云飞¹ TP18

(吉林大学计算机科学系 长春130023)¹(中山大学计算机软件所 广州510275)²

摘要 The paper surveys the origin and development of model-based diagnostic reasoning, briefly states its typical theories and methods, and analyzes its developing trends.

关键词 Model-based Diagnosis, Consistency, Abduction

诊断推理技术起源于美国,虽然沿用了医学领域中的名称,但其研究内容已远远超出了医学诊断。它在工程、社会和经济等领域有着广泛的应用前景。由于60年代初期航天和军工等高科技工程的迫切需要,诊断迅速发展起来成为一门新兴学科。特别是近十几年来,由于一些关键性现代化设备出现重大故障,引起灾难性后果的事件时有发生,为避免这些事件的发生,人们在设备诊断上投入了大量研究工作,在人工智能领域形成了诊断推理这一研究热点。作为智能诊断推理重要方法的基于模型诊断正是在此形势下,为了克服第一代诊断专家系统的严重缺陷而出现的新型诊断推理技术^[1,2,12]。本文对其发展过程进行了扼要回顾,简要介绍了具有代表性的理论与方法,并在此基础上对未来发展趋势进行了分析和展望。

1 基于模型诊断

基于模型诊断是人工智能领域近年发展起来的一个十分活跃的研究分支。其主要思想是:根据系统的组成元件与元件之间的连接建立起待诊断系统模型(结构,功能,行为),这种模型通常用一阶逻辑语句来描述。根据系统的逻辑模型以及系统的输入,我们能通过逻辑的推理理论推导出系统在正常情况下预期的行为,如果观测到的系统实际行为与系统预期的行为有差异,就说明系统存在故障。利用逻辑的推理理论,我们能够确定引发故障的元件集合,这就是基于模型的诊断。

基于模型诊断与传统诊断方法有着本质的区别,是一种全新的诊断方法,因而被一些人工智能专家誉为诊断理论和技术上的革命^[1]。传统诊断方法都是基于经验的诊断(也称为基于“浅知识”的诊断),这种诊断主要依靠领域专家对诊断对象的经验,对领域有极强的依赖性。对一个对象适用的诊断不能用在另外一个诊断对象上。采用这种方法的计算机诊断系统主要有两类:一类是基于规则的诊断系统,另一类是使用判别树的诊断系统。研制这种诊断系统的关键是知识获取,而问题就由此产生:一方面领域专家需要长时间积累才能获得诊断对象的知识,另一方面,从领域专家那里获得知识并转换成计算机系统可用的产生式规则也是一大问题。因此,开发计算机自动诊断系统所需的研制周期较长,跟不上设备更新换代的步伐,满足不了应用需要。对于不断采用新技术新设计的电子电器产品更是如此。此外,基于经验的诊断有极强的设备依赖性,并且不能对诊断给出解释。与其相比,基于模型诊断系统的优点是:它可以从产品设计或制造可利用知识开始构造,可避免从领域专家获取知识的困难,只要模型抽象得正确,就能诊断所有可能的故障,不但能处理新故障,而且还能给出令人信服的解释。并且由于它把系统的推理内核与系统模型分开,只要更换系统模型就可用于诊断另外一种设备,因而具有设备独立性,易于推广。

基于模型诊断研究的动机正是希望克服第一代诊断系统的缺陷。70年代中期到80年代初为此项研

*) 本课题得到国家自然科学基金、国家教委博士点基金资助,李占山 博士研究生,研究方向:基于模型诊断、智能决策。姜云飞 教授,博士生导师,研究方向:自动推理、基于模型诊断及规划。

究的开创性时期 (INTER, SOPHIE, ABEL, HT, DART 等^[1,2,3]), 之后逐渐发展起来成为新型诊断推理技术。因其具有上述诸多优点及很高的实用价值, 吸引了众多研究者, 到了八十年代中期这个领域逐渐热了起来, 到九十年代它在国际上已经成为一个十分活跃的研究分支。从推理的角度看, 基于模型的诊断可以分为两大类, 一类是基于一致 (consistency-based) 的方法, 另一类是溯因 (abduction) 方法。

1.1 基于一致方法

基于一致方法最初提出是为了处理待诊断系统的正常行为模型, 而所谓的“一致”是指系统的行为描述 (正常行为或异常行为) 与观测无矛盾, 这就是基于一致诊断系统理论。在此方法中诊断是一 (最小) 有关该系统某些元件异常行为的假设集合, 使得观测与所有其余元件行为正常的假设一致, 即诊断系统要求描述 SD, 观测 OBS 与诊断假设 HYP 一致, 这相当于解释的弱概念, 持这种观点的代表人物是 de Kleer, Williams 和 Reiter。

Reiter 的工作^[4]对基于一致诊断给出了很好的形式化描述。其基本思想就是: 诊断是系统中某些元件发生故障而其余元件行为正常的假设, 使得此假设与系统描述和观测无矛盾。而诊断的出发点是从系统输入和模型预测的输出与实际观测到的输出之间的差异着手, 首先识别冲突集合, 然后从该集合中找出与每一个冲突存在公共元素的碰集 (hitting set)。Reiter 证明了最小碰集就是对系统的诊断。与此同时, de Kleer 和 Williams 发表了基于 ATMS 的多故障诊断 GDE (general diagnostic engine)^[5]。利用了 ATMS 的特点和约束传播方法完成了冲突检测, 产生了冲突集, 这一点与 Reiter 的结果一致。并且利用诊断候选的故障概率指导进一步诊断测试, 提高了效率, 因而在工程上有一定实用性。这两篇文章奠定了基于一致诊断理论基础, 随后的许多文章都沿用了他们的术语和方法。

从 Reiter 和 de Kleer 等人的工作可以总结出, 基于一致诊断有以下特点:

(1) 系统通过一阶语言或与其等价工具建模, 并不要求与领域有多少联系, 扩展了系统应用范围, 突出了有关正常行为的“第一原则”, 增强了灵活性。(2) 诊断工作从设备设计时可利用信息开始, 所需模型不需时间积累即可获得, 提高了建模效率, 降低了使用成本。(3) 能够处理新故障, 可靠性高。(4) 但这些模型缺少对任意问题求解知识或模型识别能力,

不能识别熟悉的事例和利用专家经验, 导致效率差, 常要付出昂贵计算代价。理论上已证明 Reiter 的诊断求解问题是 NP 的。

1.2 溯因方法

溯因方法的提出最初是为了处理待诊断系统的故障行为模型。在此方法中诊断是一 (最小) 有关系统某些元件异常行为假设集合, 使得该集合覆盖 (蕴涵) 观测 OBS。基本思想是: 溯因诊断以系统故障与征兆间因果理论 Σ 为系统描述的基础, 由此实现了故障集合 C 和征兆集合 E 之间的映射。对一观测集 $O \subseteq E$ (征兆集的一子集) 的溯因诊断是 C 中一最小假设集 A, 使得 A 不但与 Σ 一致而且 Σ 与 A 共同蕴涵观测集 O。溯因诊断的着手点是系统所表现出来的异常行为 (症状), 然后利用已建立的因果理论找出那些蕴涵症状的原因集合, 最后利用启发式知识否定不可能候选诊断。持这种观点的代表人物是 Poole^[6]。因此溯因诊断相当于更强解释概念, 它不仅要求诊断与观测一致, 而且要求诊断蕴涵观测。因此相对于基于一致诊断给出了更少的候选解。在对知识的要求上, 溯因诊断模型要求记录部件发生故障时反映的异常现象, 并希望通过已知故障模型解释观测结果。总结 Poole 等人的工作我们可以得到以下结论:

(1) 系统通过一阶语言建模, 与领域关系密切, 突出了因果理论, 具有针对性强等特点。(2) 给出了从故障到观测的完整路径。(3) 已知系统的故障方式并以此解释观测结果的方法很直接, 效率高, 并能有效地区分各个故障。(4) 当具有完全设备故障模型时, 溯因诊断优于基于一致诊断。但实际应用中很难获得完全故障模型, 这会严重影响系统可靠性。(5) Bylander 已证明了大多数溯因问题是 NP 的^[9], 这使溯因诊断应用受到了限制。

2 基于一致模型的变种

从上面两种诊断方法的描述我们能够发现基于一致和基于溯因诊断方法的优缺点是互补的, 许多研究者已注意到这一问题并做了大量工作, 这些工作基本是按照两条技术路线进行的: 扩充基于一致诊断处理故障模型; 扩充溯因诊断处理正常行为模型。

1989年 de Kleer 和 Williams 提出了带有行为方式模型的诊断 Sherlock, 把元件故障行为方式加入到系统模型中, 试图利用故障方式排除不可能候选诊断, 缩小诊断空间。同年 Struss 和 Dressler 提

出把预测故障模型加入其中,用一可能故障集刻画每一元件且在系统描述中包含描述故障结果的公式,建立了GDE⁺系统,排除了逻辑上一致,物理上不可能故障。与GDE⁺方法相似,1990年Friedrich等人用物理不可能替换了GDE⁺中元件故障模型,导致了更高效策略。同年Friedrich等人又开发了MOMO系统,把不同类型的行为模型(如正常、故障、alibis、和物理不可能等)集成到系统模型中,并且改变了候选生成测试方式,每次只生成一个候选诊断,接下来就进行测试,这些措施大大提高了诊断效率。1991年Console和Torasso总结了前人的工作,给出了一诊断集成框架,把溯因诊断和基于一致诊断在一个定义中表示出来,基于一致和溯因诊断是该定义的两极端情况。1992年Konolige证明了在领域模型完备化情况下,这两种诊断结果完全相同^[10]。

1991年de Kleer提出了焦点策略,利用经验把诊断重点聚焦于最可能故障元件上,而忽略其它元件,并且分析了在此之前的诊断方法中预测、冲突识别、候选生成三个阶段产生计算复杂性原因。de Kleer为此把三个过程同时处理以使花费在其上的计算工作同时最小,另外每次仅生成一个候选,边生成边检测,这样节省了大量花费在不可能候选诊断上的时间,与非焦点GDE相比效率明显提高。1994年Dressler和Struss总结了各种焦点方法^[11],利用缺省诊断机DDE中元件故障方式优先序形式化描述了焦点原理。在文献[11]中诊断是工作假设w下系统元件的方式指派 Π 使得SD、OBS、w和 Π 无矛盾, Π 是w下优先诊断当且仅当在w下没有其余诊断严格优于它,而某一元件各行为方式间的优先权是在一特定环境下给定的。Dressler和Struss利用缺省理论表示了诊断的优先权,并证明了其在缺省逻辑框架中的有效性。

另一种处理复杂设备诊断方法是层次诊断。该方法依托于基于一致或溯因方法,其基本思想是把待诊断设备按其结构分层,这样每一层次处理的元件数量受到限制,而转换到详细层时,仅生成与上层诊断相关的候选,搜索空间大大减小,从而提高了诊断效率。

1991年Mozetic对层次诊断进行了较深入的研究,定义了三个抽象算子及其逆:值的细化/倒塌、映射的详细/简化和变量的引入/删除。1993年Friedrich提出了ab-理论诊断法,它是Reiter工作的抽象化处理。在该文中分层理论 Σ 是下列形式的

ab-子句集合:

$$ab(c_i) \equiv ab(c_1) \vee \dots \vee ab(c_k)$$

其中 $ab(c_i)$ 是每一个 $ab(c_i) \in \{ab(c_1), \dots, ab(c_k)\}$ 的直接祖先,左边仅包含一个文字,但右边至少包含一个文字,就是说元件 c_i 异常那么构成 c_i 的子元件 c_1, \dots, c_k 中至少有一个也异常。这种类型的等式相应于下列形式子句:

$$\neg ab(c_i) \vee ab(c_1) \vee \dots \vee ab(c_k)$$

$$ab(c_i) \vee \neg ab(c_1), \dots, ab(c_k) \vee \neg ab(c_k)$$

通过用ab-理论建立的抽象化描述 Σ 与(SD, COMPS, OBS)描述相结合使诊断问题描述具有结构性,易于理解 and 解释,ab-理论是对层次诊断的形式化处理,是对(SD, COMPS, OBS)进行的抽象,如分层理论抽象 Σ_a 。对问题求解先是求最小理论诊断(MT-诊断),而后求(SD, COMPS, OBS)的最小诊断,该方法具有一定的构造性和简明性,能够解决较复杂的问题。

1995年Damasio等人总结了前人的工作,针对诊断系统缺乏策略指导的问题,在系统中显式地表示了优先和策略,突出了优先和策略在诊断中的重要性^[12],并对层次间转换定义了谓词refined。这样把诊断策略直接编辑到诊断程序中,使系统知识表达能力更强更清晰。同年Nejdl等人在文献[13]基础上进一步明确了诊断策略表示问题^[14]。在该文中诊断问题是用(SD, STRAT, COMPS, OBS)表示的,其中STRAT是策略集,给出了一个基于模态逻辑框架的诊断策略表示,并描述了各种策略在诊断中的实际应用表示问题。

总的看来,上述的各种方法都是基于一致方法的扩充,下面介绍的方法是溯因诊断方法的扩充。溯因诊断通常对给定一观测集合会生成很多解,而对如何从中选择满意解的描述较少,这样就使问题呈现了NP性。为了解决这一问题,1994年Harmelen和Teije提出了用领域知识辅助选择溯因解^[15],它们把因果蕴涵分为四种情况:充分因果;可能因果;条件因果;条件可能因果。相应地溯因解也划分为四种,并且按照条件强弱进行排序,对求满意溯因解进行了尝试。同年McIlraith扩充了Console和de Kleer等人的工作,提出了用与最小基于一致诊断相对应的最小基于一致排除诊断刻画诊断空间的思想,并证明了在所有最小冲突为负的情况下,最小基于一致排除诊断假设成立且刻画基于一致诊断空间。另外他还扩充了核心溯因诊断的刻画问题,其优点是更直接地证明了核心基于一致诊断与核心溯因

诊断的关系,把核心溯因诊断与类-ATMS 算法可计算的本原蕴涵式直接联系起来。同年 Feret 和 Glasgow 提出了利用经验辅助诊断^[16],该文分析了溯因诊断存在的问题。为此他们扩充了 Console 和 Torasso 的定义,把溯因、演绎、归纳和基于事例推理集成到一统一框架中。利用演绎定理和故障因果模型排除不可能诊断,结合归纳与基于事例推理来处理溯因模型解决不了的问题,从而使经验辅助诊断模型能容纳各类诊断解:当系统模型完备时,可以得到基于一致或溯因解,当系统模型不完备或模型不可靠无法获得满意解时,可借助于事例库中事例与当前系统情况对比归纳得到诊断解,处理或部分地解决现有诊断方法存在的缺陷,具有一定实用性。

3 基于模型诊断研究的现状

经过二十多年的发展,基于模型诊断在理论上已逐步走向成熟,现在已成为一个十分活跃的人工智能研究分支。至今已出版了两本基于模型诊断论文集,一本是1992年由 Hamscher、Console 和 de Kleer 主编的 *Readings in Model-based Diagnosis*^[1],另一本是1994年由 Console 和 Friedrich 主编的 *Model-based Diagnosis*^[12]。这两本书中包含了62篇有关基于模型诊断的优秀论文。在著名的 *Artificial Intelligence* 杂志以及 *IJCAI*、*AAAI* 和 *ECAI* 等重要会议论文集中也经常能看到有关基于模型诊断的文章出现。另外,欧美一些国家从九十年代初期开始每年举行一次诊断原理会议。这些论文反映了各种各样的理论和方法对诊断知识的表示和描述,如 McIlraith 对诊断理论的进一步刻画。Eshghu 讨论了在待诊断系统模型中使用失败即否定(negation as failure)的重要性。Damasio 和 Nejd 等人对诊断策略和优先序的表示,强调了策略在诊断中的重要性。Geffner 和 Pearl 将系统行为的因果知识用信念网表示,并用 Bayes 推理来实现诊断。Friedrich 对 Reiter 理论的抽象刻画,为层次诊断提供了理论依据。正是由于这些理论和方法的不断出现,促进了基于模型诊断理论的蓬勃发展,并使其在实践中得到了更好的应用。

在应用上,一些开发的基于模型诊断系统已应用到实际中,满足了应用的需要。如在工程领域,Mosterman 和 Biswas 成功地把 Lackinger 和 Nejd 提出的诊断动态系统的监视与诊断相结合的方法应用到增殖反应堆辅助冷却循环系统的故障诊断上;在医学领域,帮助医生诊断与修复外伤治疗的

TraumAID 系统;在经济领域,帮助人们审计财会数据的决策系统 CROSBY 等。

4 结论与展望

通过上述各种系统的描述,我们可以得到以下结论:

(1)基于一致和溯因推理的集成化方向发展,既依托于第一原则又利用了因果理论,使诊断更接近于实际,能很好地解释故障产生的原因;

(2)由静态诊断向动态过渡,开始着手于实际领域和复杂设备诊断理论的研究,诊断策略在诊断过程中显得越发重要;

(3)把诊断推理与诊断模型分开,使诊断系统有很强的设备独立性,只要更换模型,就可用于诊断其它设备;

(4)诊断的事后性,基于模型诊断是在故障发生后进行的不具有实时诊断和预防能力^[17];

(5)在获得反映系统本质的完好诊断模型上存在困难。

一些研究人员已注意到这些问题并进行了一系列研究。一方面随着神经网络技术的发展,人们已开发出基于神经网络的诊断系统,试图利用神经网络技术的优点解决这些问题。但纯神经网络诊断系统也存在一定缺陷(如没有理论指导且不能解释诊断,训练样本的困难,忽视专家经验等)。另一方面随着模糊技术不断发展,人们正在研究基于模糊推理的诊断系统,试图用模糊技术解决诊断系统中征兆的随机性和间断性给基于模型诊断带来的困难。

基于上述情况,我们认为基于模型诊断未来的一个发展方向应是神经网络理解成一类知识源的表达与处理模式,并把它融入基于模型诊断模型中,实现逻辑推理与形象思维相结合;另一个发展方向是如何把模糊推理技术结合到基于模型诊断中,发挥它们各自的长处,这是我们目前正在研究的问题。此外我们也在尝试用逻辑对象表示及实现诊断问题且已取得了一定成果。

参考文献

- [1] W Hamscher, et al. *Readings in model-based diagnoses*, Morgan Kaufmann publisher, 1992
- [2] S. Bubln et al. *Diagnosing electronics system by integration models of normal and abnormal functioning*, IEA/AIE, 1989

(下转第31页)