

基于案例推理的中医诊疗专家系统

杨 健¹, 马小兰², 杨邓奇¹

(1. 大理学院数学与计算机学院, 大理 671003; 2. 大理学院临床医学院, 大理 671003)

摘要:把基于案例推理技术用于中医诊疗专家系统的知识表示和推理。提出一个系统模型, 介绍案例推理的基本结构, 案例提取网以及案例提取算法, 对系统中案例的学习和修正机制进行了说明, 给出用基于案例的解释来生成诊断结果的解释性说明, 以及辅助构建基于案例的中医诊疗辅助教学系统。对系统的特点进行总结, 指出进一步研究发展方向。

关键词:人工智能; 专家系统; 中医; 基于案例推理; 案例提取网

Diagnosis and Treatment Expert System of Chinese Traditional Medicine Using Case-based Reasoning

YANG Jian¹, MA Xiao-lan², YANG Deng-qi¹

(1. Faculty of Mathematics and Computer, Dali University, Dali 671003; 2. Faculty of Clinical Medicine, Dali University, Dali 671003)

【Abstract】This paper integrates the CBR technique into the knowledge's representation and reasoning of the expert system of Chinese traditional medicine, puts forward the system's model, introduces the basic framework of case reasoning: case retrieval nets and case extraction algorithm. It explains the mechanism of case study and revise, and puts forward a method to explain the diagnosis result and to build an assistant tutoring system of Chinese traditional medicine by using case-base explanation. This paper makes a summary of the advantages of the system, and puts forward the further research and development directions.

【Key words】Artificial Intelligence(AI); expert system; Chinese traditional medicine; case-based reasoning; case retrieval nets

1 概述

中医药现代化是国家中长期科技发展规划中具有战略意义的研究课题, 为了适应信息时代发展要求, 促进祖国传统医学的传承和发展, 使中医中药在国际竞争中更具优势和特色, 利用先进的、智能的信息技术来解决中医诊断信息化过程中的关键问题, 无疑是一项迫在眉睫的基础性工作。然而, 我国在这方面的整体研究水平还较低, 多数中医诊疗专家系统缺乏足够的智能性, 离临床使用要求还有很大距离。归结其原因主要有:

(1)没有从中医整体性、系统性来分析和解决问题, 仅针对具体病症分散开发一些小系统^[1]。

(2)系统结构固定, 多采用基于规则推理, 准确性依赖于初始化时专家知识库的建立, 难以适应多变的实际应用环境。

(3)没有很好实现中医专家知识的自动获取和学习, 对于半结构化和非结构化诊疗知识无法很好地表示, 存在知识获取瓶颈。

(4)直觉性经验知识是专家经验知识的重要组成部分, 现有系统知识获取和表示形式难以适应中医的经验性知识。

基于案例推理(Case-Based Reasoning, CBR)是将以前解决问题的经验以案例形式存储, 作为以后问题解决参考的一种机器学习和推理方法^[2]。它在非结构化知识表示上有很大优势, 并且在知识获取上, 优于基于规则的表达。在基于CBR的医学诊疗专家系统中, 知识的主体是病案案例, 在知识获取和学习上有大量现成的来源。本文将CBR技术引入中医诊疗专家系统的建模中, 使用案例表示中医专家的诊疗经验, 用案例提取网(Case Retrieval Nets, CRN)作为案例提取结构,

实现案例的提取、学习。系统根据几千年来众多中医名家的诊疗经验建立案例库, 以案例推理方法为基础, 模拟中医专家看病推理过程, 针对病人的“望、闻、问、切”四诊症状, 推断出病人几种症候的可能性, 并由此提出建议处方, 实现智能化的中医专家诊断过程。

2 系统框架和案例提取算法

2.1 系统框架

中医诊断治疗过程如下:(1)建立病人信息模型, 通过“望、闻、问、切”四诊来获取疾病症状、体征等临床信息;(2)根据患者个人信息和症状信息, 结合中医理论, 给出患者的证型症候信息及对患者症状的解释, 即“辨证”; (3)根据辨证和药物特征, 结合前人方药经验, 开出药方, 提出相应的治疗方案^[3]。可以看出, 在这个过程中包含2个推理阶段: 由疾病的表现得到症候信息和由症候信息推理得到相应的治疗方案和方剂。

有经验的中医专家在这2个阶段中除了能够充分结合多方面的中医药理论知识, 最重要的一点就是他们还能将各种诊疗经验结合在当前疾病的诊疗上。而这种诊疗经验可以转化为案例的表述, 从而利用案例推理实现智能的中医诊疗系统。因此, 提出将案例推理技术用于中医诊疗专家系统开发的思想。

基于案例推理的中医诊疗专家系统诊断模型如图1所

作者简介:杨 健(1976-), 男, 讲师、硕士, 主研方向: 数据库, 人工智能; 马小兰, 博士; 杨邓奇, 硕士

收稿日期:2007-12-20 **E-mail:** sbjc1215@126.com

示。图中病人信息包括病人的既往病史、日常生活习惯、体征等内容，CBR 推理机根据病人信息和四诊症状信息进行推理，从诊断案例库提取的最近似案例送入案例修正模块，案例修正后的输出即为对当前病案的诊断结果。诊断并治疗取得一定效果后，该病案可以被输入到案例学习模块，根据病人反馈及诊断案例库现有案例的情况判断是否可以作为经验案例存入诊断案例库，即经验学习。

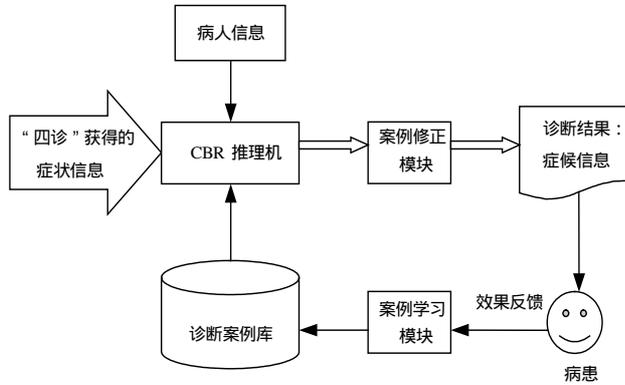


图1 基于 CBR 的诊断模块

诊断模块的输出是症候辨证信息，这也是治疗模块的输入信息。因为基于案例推理的治疗模块与上述诊断模块结构近似，这里不再赘述。

2.2 案例的表示和案例提取网

由于中医学的学科特性，案例的描述具有极大的不确定性，因此不能以固定属性描述的案例结构来表达诊疗经验。本文用案例提取网(Case-Retrieval Nets, CRN)作为诊疗案例的描述和提取结构，以动态属性结构描述案例，有效地解决了上述问题。根据不同的查询问题，这个网状结构在内存中动态生成^[4]。

图2是一个CRN的结构示意图。CRN中的案例通过若干个信息实体(Information Entities, IEs)来描述。IEs类似于传统CBR中案例描述的“属性-值”对，但它是一个原子结构，是CRN中知识的最小单元。通常情况下使用多个IEs来描述一个案例，不同案例由不同的IEs集合来描述。不同IEs之间有相似性关联，且每个IEs与它要描述的对应案例之间有相关性关联。

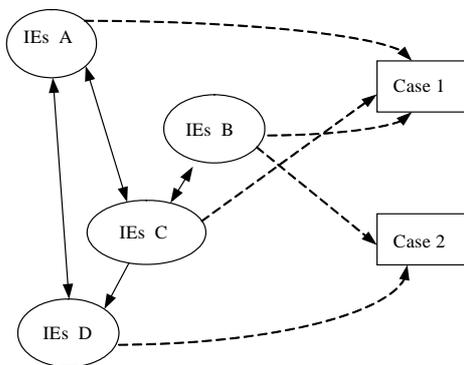


图2 CRN结构示意图

利用CRN解决问题包括以下3个基本步骤：

(1)与待解决问题相关的IEs的激活。在中医诊断过程中，可以把病人的症状信息解析为IEs；在治疗过程提出方案，把具体的症候描述作为IEs。

(2)CRN中IEs间相似性计算。IEs间的相似性计算提供

了类似于传统CBR中相同属性的不同值之间的比较。通过IEs之间的相似性，可以得到与问题描述相关的其他IEs的集合，这是一种传播激活的方式。通过这种方式，扩大了搜索的范围。

(3)案例相关性计算。根据集合中的各个IEs与各案例的相关度，计算出案例与当前问题的关联程度。而最终的案例提取就以此为依据。

2.3 CRN的建立和案例提取

CRN是根据具体问题实时建立的一个网状结构，然而，要根据问题生成CRN，需要案例库中存在如下2个关系：

(1)IEs e_1 和 IEs e_2 之间的相似性，用 $\sigma(e_1, e_2)$ 来表示。

(2)IEs e 和案例描述 c 之间的相关性，用 $\rho(e, c)$ 来表示。

这2个关系取值可以是[0, 1]间的一个小数。根据上面CRN解决问题的基本步骤，相应的CRN构建过程也有3个步骤：

(1)诊断过程中的各种输入信息(症状表现等)可以解析为一个IEs子集。对于案例库中所有IEs： $e \in E$ ，症状描述中若出现该IEs，则有 $\alpha_0(e) = 1$ ，否则，有 $\alpha_0(e) = 0$ 。由于可能有同义词出现，可以建立一个同义词表，多个同义词对应一个IEs。而匹配过程则是先以单个词为基础对问题描述进行简单的字符串匹配或是自然语言理解(NLP)分析，得到一个词语集合以后，再根据同义词与IEs的对应关系，将其转化成为信息实体全集 E 的一个子集。

这一步也是问题的解析过程，得到解析后IEs的集合，也就是 $E_1 = \{e | \alpha_0(e) = 1, e \in E\}$ ，是通过存储在内存中 $\alpha_0(e)$ 的值来表示它是与问题相关的。

(2)对于案例库中每个IEs： $e \in E$ ，计算：

$$\alpha_1(e) = \pi_e(\sigma(e_1, e) \cdot \alpha_0(e_1), \sigma(e_2, e) \cdot \alpha_0(e_2), \dots, \sigma(e_s, e) \cdot \alpha_0(e_s)) \quad (1)$$

其中， $e_1 \sim e_s \in E_1$ 是问题解析出来的IEs的集合； π_e 则是一个加权函数，一般情况下可以采用求最大值或是求和的算法。这一步完成后，针对当前问题的CRN中各个IEs之间的相似性关联就建立起来了。

通过 E_1 以及上述公式，激活另一些IEs： $e \in E_2$ ，它们满足 $\alpha_1(e) > 0$ 。这个激活过程可以通过以下公式进行无限的扩展：

$$\alpha_2(e) = \pi_e(\sigma(e_1, e) \cdot \alpha_1(e_1), \sigma(e_2, e) \cdot \alpha_1(e_2), \dots, \sigma(e_s, e) \cdot \alpha_1(e_s)) \quad (2)$$

其中， $e_1 \sim e_s \in E_{t-2}$ ，且 $E_{t-2} = \{e | \alpha_{t-1}(e) > m\}$ ； m 是计算 $\alpha_{t-1}(e)$ 时的一个阈值。

当这种激活过程只进行2步(即只计算到了 $\alpha_1(e)$)时，这种迭代的计算只进行了1次，因此，此时就是CRN的基础模型BCRN(Basic CRN)。

在传统CBR中，需要计算案例库中每个案例与当前问题对应的描述属性的两个属性值之间的相似程度，以反映问题描述与该案例的相似度。在CRN中，计算的 $\alpha_1(e)$ 值就反映了这样一种属性值的相似度，反映出案例库中案例的“属性-值”对与所提问题的相关度。

(3)对于案例库中每一个案例 c ，计算：

$$\alpha_2(c) = \pi_c(\rho(e_1, c) \cdot \alpha_1(e_1), \rho(e_2, c) \cdot \alpha_1(e_2), \dots, \rho(e_s, c) \cdot \alpha_1(e_s)) \quad (3)$$

其中， π_c 函数与上述的 π_e 类似，是一个加权函数。这时，IEs之间存在相似性关联(即 $\alpha_1(e)$)，而IEs与案例之间存在相关性关联(即 $\alpha_2(c)$)，图2所示的CRN网络构建完成。

CRN 中每个 $\alpha_2(c) \neq 0$ 的案例都是跟当前问题有关联的案例,只是它们与当前问题关联程度不同,而 $\alpha_2(c)$ 值就是案例 c 与当前问题关联程度的数值表示。这个数值可以称为关联值。可以提取出关联值最大的一个或几个案例作为当前问题解决的参考。在基于案例推理的中医诊疗专家系统中,对于诊断过程,根据具体病例信息建立 CRN,然后提取出与当前症状最相似的以前的案例(存储着诊断经验),采用其诊断结果经过案例修正,作为当前病例的建议诊断结果,实现了基于案例推理的诊断。中医“论治”的治疗过程与此类似,这里不再赘述。

3 案例修正、学习和解释

3.1 案例的修正和学习

案例的修正是 CBR 系统中一个重要的环节。因为提取出来的案例不可能完全与待解决问题吻合,所以就要根据待解决问题以及一些修正规则,对提取的案例进行改编,以期能更加符合用户的要求。案例的修正也是基于案例推理系统的一个难点,在中医专家系统中,中医学基础理论是案例修正规则的主要来源,也可以利用人工智能方法提取修正规则,作为案例修正的依据。

在中医诊疗专家系统案例的初始化过程中,领域专家要将各种疾病的典型中医病案案例以标准术语进行描述,形成规范的 IEs 全集,并初始化相似度和相关度量,然后知识工程师将这些典型病例输入系统,构造两个案例库:诊断案例库和治疗案例库。症状表现等信息是诊断案例库的案例描述,而辨证的症候描述则是诊断案例库的案例解决;同理,症候描述是治疗案例库的案例描述,而相应的治疗方案和药方则是其问题解决描述。

在系统使用过程中,具体案例的诊疗又可以以标准化术语(IEs)描述,根据其治疗效果决定是否加入案例库,成为以后诊疗的经验,实现案例的学习。并且,在学习过程中还要避免过于相似的案例存入案例库,造成案例库的冗余。

3.2 案例的解释

案例的解释在中医诊疗专家系统中有 2 个重要的目的:(1)面向病患的解释,用于说明病患症状的成因;(2)面向系统使用者的解释,这种解释通过进一步的改造可以成为基于案例的中医辅助教学系统。

可以使用中医学理论规则作为病案案例的解释,然而,已有研究表明,基于规则的解释在分类/诊断类型应用中并不优于基于案例的解释^[5]。基于案例的解释(Case-Based Explanation, CBE)是基于案例的推理与解释技术的结合,这种结合主要在 3 个层次上:使用解释支持 CBR 的内部过程,利用 CBR 来生成解释,使用案例为外部用户解释系统的推理结果。在中医专家系统中,可以提供以前案例诊断过程和结果作为当前病案诊断的解释,这对于解释的第 2 个目的来说很有效果。然而面向病患的解释还需要有专家规则的辅助才能生成。

基于案例的解释在诊断过程中还有重要的辅助作用。当患者提供的初始信息及症状描述并不完备时,需要医生对某些可能症状进行询问以进一步明确症状表现。这时,需要根据已有的信息找出最应该询问的问题。参考 McSherry 在其

First Case 和 Top Case 系统中采用增量最近邻方法,在诊断过程中,根据当前已提供信息进行最相似案例查找,然后对这最相似的若干个案例进行比较,找出没有提供的且差异最大的“属性-值”对(也即信息实体),据此向病患提出问题,以明确具体症状。这其实就是实现了“望闻问切”中的问诊。通过这种交互反馈方式,进一步减小了系统误诊的几率,提高了系统的精度,也有助于使用系统的医生水平的提高。

4 总结和展望

将基于案例推理技术用于中医诊疗专家系统是中医智能诊疗系统建设的一个新思路,能够有效解决中医专家系统在构建过程中的若干问题,其优势主要有以下几点:

(1)对于不确定、不完全和不一致的病患信息有较强的适应能力,并可以利用基于案例的解释实现问诊,以进一步明确症状信息,提高诊断精度。

(2)案例提取网的结构决定了病案案例的知识获取非常方便,并且不会影响以前的案例,有效解决了知识获取瓶颈。大量的古今病案案例也是很好的案例来源。

(3)通过基于案例的解释进一步发展,可以生成中医诊疗案例教学系统,帮助年轻医师迅速提高诊疗经验。

(4)随着系统的使用时间增长,案例不断增加,系统也可以逐渐改善推理性能,适应各种病证而不仅限于某一种或一类疾病。

要建成实际可应用的系统,还应在以下方面进一步完善和发展:

(1)望闻问切四诊的定量化、数字化研究的进一步发展,有助于促进中医诊疗专家系统的发展。

(2)利用数据挖掘和人工神经元网等人工智能方法从大量病例中挖掘有效的专家规则,用于指导案例的学习和修正。

(3)会话式 CBR(conversational CBR)和交互式 CBR 的研究发展有助于帮助中医问诊的智能化实现。

(4)需要找到合适的知识表示方法来表达病案案例中病情发展的时序关系。

(5)2.3 节中的各加权函数常采用最大值或求和的形式,可以考虑利用人工神经网络技术求取和表示这些函数及权值信息,以求得更加准确的表达。

参考文献

- [1] 周昌乐, 张志枫. 智能中医诊断信息处理技术研究进展与展望[J]. 中西医结合学报, 2006, 4(6): 560-565.
- [2] Aamodt A, Plaza E. Case-based Reasoning: Foundational Issues, Methodological Variations, and System Approaches[J]. Artificial Intelligence Communications, 1994, 7(1): 39-59.
- [3] 胡东红, 李德华, 关景火, 等. 中医的四诊特征空间与辨证特征空间[J]. 北京生物医学工程, 2003, 22(4): 286-289.
- [4] Lenz M, Burkhard H D. Case Retrieval Nets: Basic Ideas and Extensions[C]//Proc. of Advances in Artificial Intelligence. [S. 1.]: Springer, 1996: 227-239.
- [5] Cunningham P, Doyle D, Loughrey J. An Evaluation of the Usefulness of Case-based Explanation[C]//Proceedings of the 5th International Conference on CBR. [S. 1.]: Springer, 2003: 122-130.