

交互式基于范例的推理及应用研究^{*}

Conversation Case-based Reasoning and its Applications

王清毅¹ 刘 钧² 唐守涛¹ 蔡庆生¹

(中国科技大学计算机系 合肥230027)¹ (皖西学院)²

Abstract The traditional case-based reasoning has three drawbacks at least. They are (1) that solves the problems with Static fashions; (2) requires user to input a complete description of the problems to be solved; (3) and the system function is somewhat single. Conversation case-based reasoning can solve the above problems. This paper conducts the research and discuss to the new research method from the concept, research congress, reasoning process, current research problems, applications and the future directions.

Keywords Conversation case-based reasoning, Artificial intelligence

1. 引言

近几年来,研究人员一直在致力于研究和解决基于范例推理(Case-based reasoning, CBR)的理论和应用课题,如更为有效的范例表示、索引、检索和修改方法、范例库的创建及维护方法,将 CBR 与其它人工智能技术集成等。一般说来,传统的 CBR 是以静态的方式求解问题的,范例是处于一种被动、等待被检索的状态,不能根据所要求解的问题及环境的变化,通过和用户的交互调整范例的内容、结构进行问题的求解。另一方面,在大部分传统的 CBR 系统中,都是要求用户一开始就要向系统输入一个所要求解的问题的完整描述,然后才开始求解,这就要求用户事先必须确定与问题求解有关的特征,并具有较为详尽的领域知识,这在实际上往往是难以做到的。再者,从应用的角度看,传统的 CBR 系统的功能一般较为单一,问题求解过程常常就是从范例库中检索出恰当范例的过程,在广泛用于完成更为复杂的决策支持任务(例如交互、实时及智能的)方面,还显得很欠缺。

为解决以上问题,研究人员提出了交互式基于范例的推理^[1](Conversational or Interactive Case-Based Reasoning, CCBR)。下面我们从概念及研究进展、推理过程、当前研究课题、应用以及进一步的研究方向等方面对这一新的推理方法进行研究和讨论。

在一个 CCBR 系统中,用户在一开始可以只向系统输入一个所要求解的问题的简要部分描述。接着系

统通过索引和检索从范例库中找出与问题的部分描述最为相似的范例,并对这些范例的解决方案按照与问题的部分描述的相似程度进行排序并显示,即解决方案显示;并且对于暂不能由所检索出的范例回答的问题(预先已在这些范例中定义好),系统也根据一定的策略(例如信息增益)对它们进行排序并显示,即问题显示。用户再和这两个显示进行交互,或者选择一个满意的解决方案,或者选择一个问题进行回答以修正原来的问题描述,再进行下一步的求解。这就是用户与系统的一次交互过程。这一过程不断循环,直到用户获得一个满意的答案。因此,整个 CCBR 过程是一个交互的、动态的、渐增的问题求解过程,在此过程中用户只需要回答系统所提出问题,而不必先具备详尽的与问题求解有关的领域知识。

美国的 Inference 公司在1995年推出了第一个 CCBR 系统,为 CCBR 做了开拓性工作。目前的研究进展和有代表性的研究事例例如:美国海军研究室人工智能应用研究中心的 David W. Aha 等人,他们的研究内容包括在 CCBR 中如何求精范例库^[2]、对话式推理^[3]和交互式基于范例的规划^[4]等;美国 Indiana 大学计算机系的 David Leake 等人提出将 CBR 与交互式工具集成利用概念映射来获取专家知识,并着手研究范例的获取、表示、检索和修改的交互式方法^[5];美国的 Coleraine Ulster 大学的 McSherry 等人研究将 CCBR 用于诊断任务^[6]等。从目前的国际市场来看,基于 CCBR 的工具已经占据了相当大的桌面决策支持

^{*} 本文得到国家自然科学基金资助(编号:60075015),王清毅 博士、讲师,研究方向为人工智能、交互式基于范例的推理、知识发现;刘 钧 讲师,中国科技大学访问学者,研究方向为交互式基于范例的推理;唐守涛 硕士研究生,研究方向为人工智能、基于范例的推理、知识发现;蔡庆生 教授、博士生导师,研究方向为人工智能、机器学习、知识发现。

软件的市场份额,这在很大程度上是由于 CCBR 可以递增和交互地建立起用户的查询,而对用户查询的内容又很少加以限制。

2. 推理的一般过程

CCBR 系统中一个范例 C 的最一般的表示形式如下:

(1) 问题 $C_p = C_d + C_{qa}$: 表示可由 C_c (见下) 求解的问题,其中: (d) 描述 C_d : 描述范例 C 的问题的文本; (a) 说明 C_{qa} : 一个 (问题, 回答) 序偶的集合。

(2) 解决方案 $C_s = \{C_{s1}, C_{s2}, \dots, C_{sn}\}$: 对应于问题 C_p 的行动 C_{si} 的序列。

当用户输入了一个问题的描述文本 Q_d 时,可以认为用户和系统的一次交互过程开始。换句话说,用户通过向系统呈交一个查询而和系统进行交互。系统接着计算 Q_d 和范例库中每个范例的问题描述 C_d 的相似

程度 $S(Q, C)$, 从而产生一个初步的范例相似排序。此后,系统提供两个显示:一个是按照 $S(Q, C)$ 的递减,确定一个范例的集合,并显示出它们各自的解决方案即解决方案显示 D_s ;另一个是对这些范例中目前未被回答的问题按照它们的重要程度进行排序和显示,即问题显示 D_q 。与这两个显示相应,用户也有两个选择:一是从问题显示 D_q 中选择一个问题 q ,并回答它。这时用户输入问题 q 的答案 a ,系统再将 $\langle q, a \rangle$ 添加到用户查询的问题说明 Q_{qa} 中,并重新计算所有范例的相似程度,修改 D_s 和 D_q 两个显示。随着更多的问题被回答,相似度 $S(Q, C)$ 的计算和范例的排序也将更为准确。同时,用户还可以随时删除或修改他们先前对问题的回答。最后,用户要么选择到一个满意的解决方案 $s \in D_s$,这时问题得以求解;要么无法确定恰当的相似范例或没有进一步的问题要询问,这时问题无法求解。交互式基于范例推理的一般过程如图1所示。

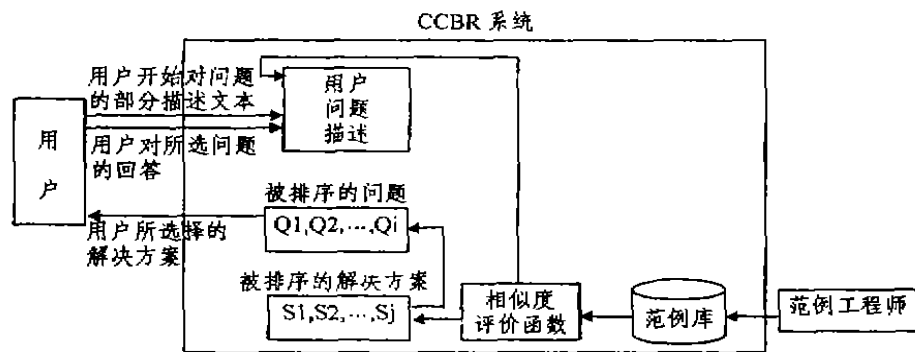


图1 CCBR 的一般过程

图2示出了 CCBR 中执行一次交互的算法,当用户得到一个满意的解决方案时,则认为一次交互结束。

```

算法: CCBR 中的一次交互。
输入: 范例库 L,  $D_s$  的大小 k,  $D_q$  的大小 n, 相似度函数  $S()$ , 问题排序函数  $r()$ 。
输出: 解决方案。
Begin
   $Q_d = user\_text()$ ; // 用户输入的问题描述文本
   $Q_{qa} = Q_d = \phi$ ;
  Repeat
     $D_s = rank\_solution(Q_d, L, S(), k)$ ; // 检索出并被排序的解决方案
     $D_q = rank\_questions(D_s, L, r(), n)$ ; // 目前未被回答的问题
    If 用户选择一个问题  $q \in D_q$  进行回答, 答案为 a
    Then  $Q_{qa} = Q_{qa} \cup \{q, a\}$ ;
  Until 用户选择一个解决方案  $S_s \in D_s$ , 或者系统推理不成功
  Return  $S_s$ ;
End
    
```

图2 CCBR 中的一次交互算法

3. 范例库的建造

传统 CBR 范例库的建造常常是一项庞大的“范例

工程”(Case Engineering)工作。尽管已有一些如何设计范例库的指南,但范例工程师最终还是发现他们所要建造的范例库的复杂性常常使得这些指南难以奏效。这一问题在 CCBR 范例库的建造过程中同样存在,并且,CCBR 的推理过程也增加了范例工程的复杂性。例如,范例工程师必须决定在一次交互过程中的不同阶段应向用户提供哪些问题和范例。如果所提供的范例过少,就会影响所要求解的问题最终形式;如果所提供的范例过少,则可能检索不到最适合的范例。另一方面,向范例工程师提供的设计指南常常很多,要想正确熟练地运用不是一件易事,这对于缺乏经验的范例工程师来说,就更是如此。为解决这一问题,美国海军研究室人工智能应用研究中心的 David W. Aha 等人提出采用软件工具对范例库进行修改以改善范例的质量。他们的方法分为三个阶段,如图3所示,即:(1)首先建立起一个范例的表示结构,这一表示结构应当使得对范例库结构的评价以及对范例的修改更易于进行;(2)其次根据恰当的范例设计指南对范例的表示结构

进行修改;(3)最后再从修改过的范例表示中抽取范例。

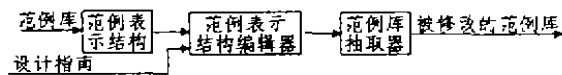


图3 范例库的修改过程

整个修改过程对用户来说是透明的,用户只需和被修改的范例库交互,而无需和范例的内部表示交互。

这一方法采用一个自顶向下的决策树归纳算法来建立范例的表示结构。算法递归地运用一个选择标准选择索引问题以划分一个节点的范例。以往大多数的选择标准都是要求范例被同样的问题集合所定义即范例是同质的,并且范例也已经被聚类,但这两项要求对CCBR的范例库就不成立。通常,CCBR的范例库中很少有问题在全部范例库中都给出回答即范例是异质的,而且每个范例的解决方案都是唯一的。因此,这里是从一个节点的范例中选择一个被回答频率最高的问题作为选择标准,这样,对于所选择的问题没有回答的范例就会被分枝到另一个节点中,再递归地进行划分。在建立了能够索引范例库中范例的表示即决策树后,就可以利用具体范例的特征来对范例进行修改,具体说来就是从范例C的 C_{in} 中,删除掉不出现在任何从决策树的根节点到含有范例C的叶节点的路径上的 $\langle q, a \rangle$ 序偶。这是因为这些 $\langle q, a \rangle$ 对决策树范例的划分没有任何作用。在这一阶段,CCBR系统顺序记录出现在从根节点到每个范例C的每条路径上的 $\langle q, a \rangle$ 即范例C的说明 C_{in} ,本方法的一个不足就是没有有效地利用领域知识,这样导致在决策树归纳过程中被删除的问题极有可能是—个很重要的问题。

4. 对话式推理

CCBR中的对话式推理就是在用户输入的问题描述和范例库之间进行推理,以保证系统和用户的交互具有一定的智能。一般说来,CCBR系统不应该重复显示在用户给出的问题初始描述中已经被回答或实际上被蕴涵的问题,我们来看一下图4。尽管前两个问题即 Q_1 和 Q_2 实际上已经在用户的问题描述即查询 Q_0 中给予了回答,但系统往往不能自动地推出这些答案,仍然需要重复显示这些问题,提请用户回答,这就降低了系统的推理效率。

用户的初始问题描述 Q_0 ,打印的文稿上有黑色条纹,系统的响应被排序的问题显示。

- (1) Q_1 :打印的质量如何?
- (2) Q_2 :打印质量有问题吗?
- (3) Q_3 :打印纸放置正确吗?
- (4) Q_4 :所显示的打印信息是什么?

图4 对话式推理的一个例子

因此,如果系统能够自动地推出被蕴涵的答案,则

它就能更加精确地定义问题的状态以便进行相似度的计算,从而减少在整个交互过程中所必须回答的问题的数量,提高系统的性能,Aha等人提出采用基于模型的方法来进行对话式推理^[1]。用户在和系统的对话过程中交互地输入一个库模型,这一库模型由对象模型(即反映了范例库中所涉及的对象)和问题模型(即范例库中的问题和对象模型之间的联系)系统结合当前的问题描述,导出问题答案,再将这些答案添加到用户的问题描述中。

5. 应用

Aha等人开发了一个交互式基于范例的规划系统(HICAP),并将它用于复杂的军事任务^[1]。

再以电子商务领域为例。目前大部分的电子商务系统都向顾客提供了搜索商品的供应商,或比较同种商品的价格的服务。但是当顾客在复杂的商品目录或庞大的数据库中漫游的同时,常常感到难以搜索和选择到他们真正所需要的商品或信息。再者,顾客在购买他们最需要的商品时所提出的要求要么过于具体,要么不够具体,因为他们往往不具备有关商品的知识,显然需要相应的支持,已有研究者提出利用基于范例的推理来解决这一问题^[4],但仍然是建立在传统的基于范例的推理基础之上,没有向顾客提供充分的交互性,也有研究者提出了在电子商务中利用软件Agent技术的解决方案,在一定程度上缓解了这一问题,但如何进一步提高顾客的满意程度仍然是下一代面向Agent的电子商务所追求的目标^[10]。这一问题的出现很大程度上是因为目前的电子商务系统仅仅向顾客提供了在线的商品目录或数据库(即提供了销售支持),而没有向顾客提供充分的智能支持,不能和顾客进行实时互动的信息交流,因此,完全可以考虑将CCBR引进电子商务(例如和软件Agent技术结合),解决智能销售支持问题。

交互式基于范例的推理不仅在电子商务,而且在规划、诊断、分类、调度、教育、虚拟现实、军事以及数据挖掘和知识发现等领域都有着广阔的应用前景。

6. 进一步的研究方向

(1)CCBR范例库的建造是一项需要大量专门知识知识工程任务,迄今已提出了一些设计范例库的指南。范例工程师在参照这些指南设计CCBR范例库时,首先必须精心地挑选范例,然后还必须确定系统在交互过程中应当向用户提供哪些范例和问题,由此引来了庞大的“范例工程”工作,问题是范例工程师最终还是发现这些指南对于复杂范例库的建造难以奏效,事实上一些人已经放弃了利用这些指南,而转向通过

代价高昂的咨询服务或利用可以在一定程度上解决问题的其它技术来设计复杂的范例库,也有人研究了改善手工建造范例库过程的方法,但不是面向 CCBR 范例库的。因此,有必要研究能为 CCBR 中高性能复杂范例库的设计提供自动支持的软件工具。

(2) 对话式推理过程就是在用户输入的问题描述和系统之间进行推理,这一过程更加强调领域知识的运用,以提高范例的检索效率。已有的解决方案是通过范例工程师将完整正确的推理规则集输入系统,但这样做一方面又带来了范例工程、可理解性和维护问题,另一方面也无法保证推理规则的正确性和领域完备性。因此,一些范例工程师已经尽量避免采用这一方法。再者,目前对话式推理所需的对象模型和问题模型还是由用户在设计范例库时人工建立的。近来,已有研究者开始研究在对话式推理过程中集成基于模型的方法以改善系统的性能。但总的看来,尚缺乏进行对话式推理的智能化方法。因此,必须提出新的对话式推理的智能化方法,具体内容例如如何从范例的描述、问题描述、解决方案中自动地提取对话式推理所需的对象模型和问题模型、反省学习(Introspective Learning)在对话式推理过程中的作用等。

(3) 目前的 CCBR 中的文本推理一般限于用户在一开始输入的问题描述文本和系统提供的问题进行交互,接着用户只能回答系统所提出的问题,用户无法再次输入文本以回答系统所提出的问题。因此,用户在问题求解过程中显得较为被动。系统也不能从用户所输入的文本中自动地推出问题的解决方案。另一方面,目前的 CCBR 系统在寻求问题解决方案的过程中还不能有效地和非结构化或半结构化的文本交互。因此,应该提出能和用户反复输入的文本(包括非结构化或半结构化的文本)进行交互推理的方法。

本文给出了交互式基于范例推理的概念及研究进展,讨论了它的一般推理过程,介绍了它的当前研究课题,探讨了它的应用前景和进一步的研究方向,期望本文能对在我国开展对交互式基于范例推理的进一步深入研究起到抛砖引玉的作用。

致谢 研究过程中,我们和美国海军研究室人工智能应用研究中心的 David W. Aha 博士进行了有益

的学术交流,得到了他的热情帮助,在此特表谢意。

参考文献

- 1 Aha D W, Breslow L A, Munoz-Avila H. Conversational case-based reasoning. To appear in *Applied Intelligence*, early in 2001
- 2 Aha D W, Breslow L A. Refining conversational case libraries. In: *Proc. of the Second Intl. Conf. on Case-Based Reasoning*, 1997. 267~278
- 3 Aha D W, Maney T, Breslow L A. Supporting dialogue inferencing in conversational case-based reasoning. In: *Proc. of the Fourth European Workshop on Case-Based Reasoning*, Dublin, Ireland: Springer, 1998
- 4 Munoz-Avila H, McFarlane D. Using guidelines to constrain interactive case-based HTN planning. In: *Proc. of the Third Intl. Conf. on Case-Based Reasoning*, Seon, Germany: Springer, 1999
- 5 Leake D B, Wilson D C. Combining CBR with interactive knowledge acquisition, manipulation and reuse. In: *Proc. of the Third Intl. Conf. on Case-Based Reasoning*, Seon, Germany: Springer, 1999
- 6 McSherry D. Interactive case-based reasoning in sequential diagnosis. To appear in *Applied Intelligence*, early in 2001
- 7 Yang Q, Wu J. Enhancing the effectiveness of interactive case-based reasoning with clustering and decision forests. To appear in *Applied Intelligence*, early in 2001
- 8 Aha D W. *Extual reasoning in the context of conversational case-based reasoning system*. [Technical Report] Navy Center for Applied Research in Artificial Intelligence, Naval Research Laboratory, Washington, DC USA, 1999
- 9 Wilke W. CBR and electronic commerce on the World Wide Web. Invited talk on the Intl. Conference on Case-Based Reasoning, Providence, Rhode Island, 1997
- 10 Guttman R, Moukas A, Maes P. Agent-mediated electronic commerce: A survey. *Knowledge Engineering Review*, 1998
- 11 王清毅, 陈恩红, 蔡庆生. 知识发现的若干问题及应用研究. *计算机科学*, 1997, 24(5): 73~77
- 12 王清毅, 蔡庆生. 基于范例系统的验证和有效性研究. *小型微型计算机系统*, 1997, 18(4): 37~42