

交互式群范例学习问题的求解及应用

邢清华, 刘付显

(空军工程大学导弹学院防空指挥系, 三原 713800)

摘要: 针对传统基于范例推理方法存在的缺陷以及实际决策问题的需求, 提出了一种交互式基于群范例学习的问题求解技术, 给出了这一技术的问题求解过程、求解算法, 针对算法中涉及到的范例检索方法, 给出了一种基于 2 级索引的范例检索算法。弥补了传统 CBR 方法的缺陷, 满足了解决实际决策问题的需要, 展望了它的应用。

关键词: 基于群范例的推理; 范例检索; 交互式; 问题求解技术

Problem Resolving Method and Application of Conversation Group Case-based Reasoning

XING Qinghua, LIU Fuxian

(Department of Air Defence and Command, Missile Institute of Air Force Engineering University, Sanyuan 713800)

【Abstract】 The traditional case-based reasoning has some drawbacks, and it does not satisfy the demand of actual decision problem. This paper proposes a problem resolving method of conversation group case-based reasoning. It gives the resolving process and arithmetic in particularly and gives the structure of case base. A two level index based method is provided for the case index problem. The technique can overcome the limitation of traditional CBR, meet the requirement of decision problem. In the end, it gives a prospect to its application.

【Key words】 Group case-based reasoning; Case index; Conversation; Problem resolving method

1 概述

基于范例的推理(CBR)^[1,2]是由目标范例的提示而获得记忆中的源范例,并由源范例来指导目标范例求解的一种策略。该技术利用最合适的相似决策经验,不必依赖于基本的因果机制,能够有效地完成问题的求解。因而它非常适合于没有理论模型或领域知识不完全而经验丰富的决策领域^[3,4]。也正是这些优点,使得AI领域对它的研究兴趣越来越大。

然而,在应用范例推理方法解决实际的军事决策问题时,由于决策者本身的需求及问题复杂度的不同,使得利用传统的 CBR 方法不能有效解决这些问题。

首先,目前关于基于范例推理(CBR)的研究主要集中在范例的有效表示、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方面^[5]。一般来说,现有的CBR是以静态的方式来求解问题的,不能适应所要求解的问题及其环境的非本质变化;另一方面,在大部分CBR系统中,都要求用户一开始就要向系统输入所要求解问题的完整描述,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,这在实际上往往难以做到。从应用的角度看,传统CBR系统功能一般较为单一,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,在交互、实时等方面,还显得有所欠缺。其次,对于同样一个决策问题,由于决策者的风格和所处的环境不同,给出的决策结果也可能不同。这样,在基于范例的求解中,就需要参照同一源问题的多个决策结果给出目标问题的求解答案。再次,对于复杂的决策问题,决策者一般要对问题进行分解,使每一子问题映射到一个相似的范例上,然后继承范例学习的所有片段知识,形成决策问题的初步方

案。这时,范例库中存储的可以是若干子问题的解,由若干子问题的解综合成复杂问题的解方案。基于以上需求,本文提出一种交互式基于群范例学习的问题求解方法。

2 交互式基于群范例学习的问题求解过程

交互式基于群范例学习的问题求解过程如图 1 所示。

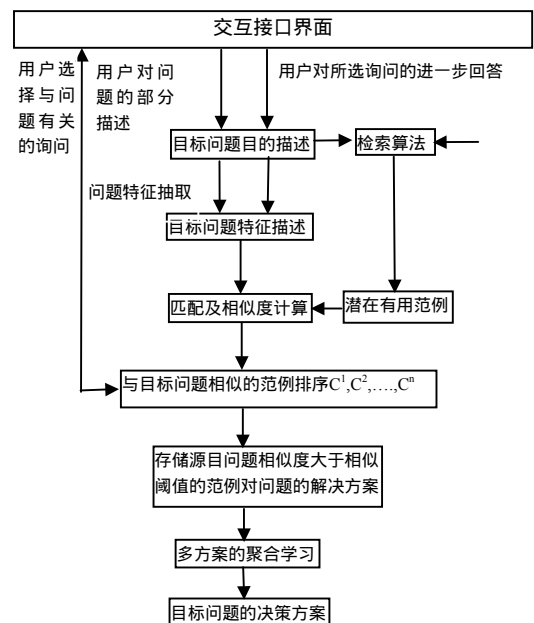


图 1 交互式基于群范例学习的问题求解过程

作者简介: 邢清华(1966—),女,副教授、博士,主研方向:防空作战智能决策,防空作战仿真建模;刘付显,博士、教授

收稿日期: 2005-08-15 **E-mail:** liuxqh@126.com

在交互式基于群范例学习的问题求解系统中，用户一开始可以只向系统输入一个所要求解的问题的简要描述，主要是题目的描述，然后，由检索算法控制从范例库中检索范例，在此过程中，一旦发现潜在有用的范例(即新问题的目的与范例中问题的目的相似)，就对其进行标识。当对范例库检索完毕后，若存在潜在有用范例，则对问题特征进一步抽取，形成问题特征或问题处境的进一步描述，并对特征进行存储。每个描述特征用一个特征描述槽表示。然后由检索算法调用匹配算法，对潜在有用范例的特征描述槽与目标问题(或新问题)的特征描述槽作相似性鉴别，进行相似度计算，并对这些潜在有用范例按照相似性程度进行排序并显示。对于源问题的相似度大于规定阈值的范例，存储范例对问题的解决方案，然后进行多范例方案的聚合学习和方案修正，最后获得目标问题的解决方案。若没有范例与目标问题的相似度大于规定的阈值，或暂不能由所检索出的范例解决的目标问题，用户可以根据已检索出的范例所解决的源问题的特征与目标问题特征的差别，通过与系统的交互式联系，选择范例提出的与目标问题有关的询问，并由用户对该询问进行回答，系统将询问及回答添加到用户问题的文本描述和特征描述之中。随着更多的询问被回答，相似性程度的计算和范例的排序也将更为准确，同时用户还可随时删除或修改他们先前对问题的回答，若经过上述循环，目标问题与源问题的相似度仍达不到要求，则没有找到目标问题的解。否则对满足相似性程度要求的范例进行与上面一样的处理，即对范例进行存储，将这些范例采用的解答作为解决新问题的起点，在此基础上进行多范例聚合学习，找出问题的解决方案。

当用户问题是复杂问题时，求解前先将问题分解为子问题，再找各子问题的相似范例，用与上面相似的步骤找到各子问题的相似范例后，再确定复杂问题的解决方案。

问题的分解有 2 种形式：串行分解和并行分解，具体是哪种分解，视具体问题而定。串行分解时，子问题 i 与子问题 $i+1$ 之间可能是一个继承关系，即子问题 i 的解决是子问题 $i+1$ 解决的前提条件，每一个子问题依次求解完毕后，最终的目标问题解决方案即被求出。并行分解时，各子问题之间可能没有直接的关系，各子问题被求解后，通过各子问题求解结果的综合得到目标问题的解答。

3 交互式基于群范例学习的问题求解算法

下面给出交互式基于群范例学习的问题求解算法。

输入：范例库 C ，其中的范例为 C^1, C^2, \dots, C^n ，题目的相似度函数 $S_c()$ 。

问题特征描述槽相似度函数 $S_c()$ ，范例排序函数 $\text{Rank_case}()$ 。

输出：问题解决方案。

```

BEGIN
Sub_Question_Flag=0;           // 问题分解标志
i=1,k=1;
KK: Qi=Question_Text();      // 新问题或其子题目的描述文本
REPEAT
    Si= Sc(Qi, Ci);        // Ci为范例Ci的问题目的
    IF Si>σ1 Then Ci为潜在有用范例，并对Ci进行标识；
//σ1为阈值
    i++;
UNTIL 范例库检索完毕；
IF 没有潜在有用范例，则问题不能用基于范例的方法求解，

```

返回。

```

Qa=φ;                          // 回答询问
Qc=Question_Character();      // 问题特征描述槽
Qc=Qc Qa;
REPEAT
    IF Ck是标识过的潜在有用范例 Then
        Sck=Sc(Qc, Ck);      // 注意k是变化的
        QCk=Rank_Question(Ck);
// 需要时，该语句下面也可有与用户的交互回答
        Rank_Case();
    ENDIF
    k++;
UNTIL 潜在有用范例检索完毕；
IF 潜在有用范例中不存在相似度大于σ2的范例 THEN
//σ2为阈值
    选择与目标问题相似度较大的一个范例Cm；
    REPEAT
        用户选择一个询问qi QCm进行回答，答案为a；
        Qa=Qa (qi,a);
        Qc=Qc Qa;
        Sc=Sc(Qc, Cm);
        Rank_Case();
    UNTIL 用户没有问题回答或存在相似度大于σ2的范例；
    ENDIF
    IF 不存在相似度大于σ2的范例 THEN
        没有找到相似范例，返回；
    ENDIF
m=0;
REPEAT
    选择源目问题相似度大于σ2的范例Ck；
    Store(Ck,S[m]);
    m++;
UNTIL相似度大于σ2的范例选择完毕；
    IF Sub_Question_Flag=0 THEN
        SQ=Combine(S);
    ELSE
        SQ=Combine_Subquestion(SQ,S);
    ENDIF
    IF SQ不满意 或 目标问题还有其它子问题未解决 THEN
        Sub_Question_Flag=1;
        GOTO KK;
    ENDIF
END
系统推理不成功有 2 种可能原因：
(1)范例库中问题的目的与新问题的目的不一样，它们的相似度为 0，或相似度小于预定的阈值。相似度为 0 说明范例库中不存在类似的范例，换句话说，系统无法通过范例学习获取问题求解的经验；
(2)潜在范例的问题特征描述槽与新问题特征描述槽的相似度小于范例抽取要求的阈值，并且通过用户交互式的回答之后，仍然不能满足要求。
需要说明的是，在上面算法的第 2 个重复(REPEAT...UNTIL)语句中也可以加上用户与系统的交互回

```

答，这样可以得到更满意的相似度。

4 范例库的建造及检索机制

范例库的建造通常是一项庞大的“范例工程”^[6]。范例的表示、索引和存储是范例库构造的关键技术。从上面的问题求解过程及算法中可以看出，这里的范例检索算法采用了2级索引机制。其基本思想是这样的：对于军事决策问题的范例，首先根据问题的特性，找出由若干关键词组成的对问题目的、特征的描述，把这些关键词以适当的方式分配给范例(把这些关键词作为范例的属性或特征)，然后将若干征求客户回答的问题或询问分配给范例，这样范例的基本组成即可确定。范例库可以组织为简单的线性列表或复杂的树状或图状结构，这些结构各有优缺点，应根据范例索引的特征和范例库规模来确定组织结构的选取。这里用基于关系数据库的方法来设计范例库，范例在关系型数据库中以记录的形式进行存储表达，每条记录就是一个范例，每个字段就是范例的一项属性或索引，属性类型可以是离散的或者连续的，也可以是定性的或定量的。范例库结构如表1所示。

表1 范例库结构

字段名	类型	长度
范例名	C	20
范例的目的描述	C	200
范例涉及的关键特征个数 N	Int	4
范例涉及的关键特征 1	C 或 Int 或 Real	X
-----	C 或 Int 或 Real	X
范例涉及的关键特征 N	C 或 Int 或 Real	X
范例涉及的一般特征个数 M	C 或 Int 或 Real	X
范例涉及的一般特征 1	C 或 Int 或 Real	X
-----	C 或 Int 或 Real	X
范例涉及的一般特征 M	C 或 Int 或 Real	X
范例询问个数 K	Int	50
范例询问 1	C	50
-----	C	50
范例询问 K	C	50
范例涉及的问题解决方案	C	100
方案后果	C	200
形成时间	Int	4

长度字段的 X 表示根据具体问题确定。

解决某一类决策问题的若干个范例存储于同一个库文件中，平时该库以文件的形式存储于外存中，运行时由范例装载机制将范例装载入内存区域，以便进行交互式问题求解。

范例的关键词实际上就是支持问题检索的索引。范例的索引是判定类似性所需要重点考虑的方面(描述槽)，它对匹配算法的设计起指导作用。一个范例可以有多个索引，它们分别对应于多个描述槽。索引应该以检索机制便于使用的方式表示，当检索算法依据范例库的组织结构搜索与新问题处境描述部分匹配的范例时，为了提高搜索效率，对范例建立2级索引：关键特征作为1级索引，一般特征作为2级索引。

只有与关键特征匹配的范例才可以是部分匹配的，否则不匹配。一般特征可以依据其相对重要性分级，并且可以不

匹配。这样检索算法对所有部分匹配的范例都进一步计算它匹配一般特征的程度，并加以排序，最后选取匹配程度较好的几个范例，并把这几个范例提供的解答作为产生新问题解答的参考和起点。

另外需要说明的是，范例库中存储的范例的解答未必全是正确的，我们可以有选择地对一些失败的范例也进行存储，这时需要在范例中提供该方案的后果，即该范例中的方案执行后的结果状态，包括解答执行是成功还是失败、成功的条件和失败的原因。通过关于后果的反馈信息，一方面可以警告和避免潜在的失败，另一方面可以评价直接使用(或改编)解答会在新问题中导致的后果。

5 总结

本文详细给出了交互式基于群范例学习的问题求解过程，设计了解决算法，针对算法中涉及到的范例检索方法，给出了一种基于2级索引的范例检索算法。归纳这一问题求解技术可得到以下几个特点：(1)在这一算法中，每个范例包含若干征求指挥员回答的问题或询问，这样为系统与指挥员的交互提供了基础，也使问题特征得到进一步抽取；同时用交互式方法还可充分获取变化的环境信息以及指挥员的主观能动性，指挥员不只是等待结果的出现，而是能通过对问题或询问的回答参与到问题求解过程中，不断寻找问题求解的相似范例，从中学习求解问题的经验。(2)将问题特征描述(或索引)分为2类：必须严格匹配的关键特征和一般特征。这时，当新问题的关键特征与范例的关键特征不匹配时，可以直接确定新问题与范例的相似度为0，不再需要进行与其它特征的匹配，这样可以提高问题求解的效率。(3)通过群范例决策方法可以综合不同指挥风格指挥员对同一决策问题的不同决策结果。通过对失败范例的存储，可以获得失败的原因，避免在新问题决策中出现类似的错误。(4)可以对复杂决策问题进行问题分解，由子问题相似范例的解答确定目标问题的解决方案。

这一求解技术已应用于多个军事决策问题之中，比如雷达辐射源识别及威胁等级判断，作战指挥中敌机主要来袭方向的判断、目标综合识别、目标威胁判断等，随着对其研究的进一步深入，这一问题求解技术的应用将更加广阔。

参考文献

- 1 Spori B B, Althoff K D. Learning from and Reasoning About Case-based Reasoning System[C]. Proceedings of the 4th German Conference on Knowledge-based Systems, 1997.
- 2 Aleven V. Teaching Case-based Argumentation Through a Model and Examples[D]. University of Pittsburgh, 1997.
- 3 Shin K S, Han I, A Case-based Approach Using Inductive Indexing for Corporate Bond Rating[J]. Decision Support Systems, 2001, 32(1): 41-52.
- 4 吴建林, 薛华成. 基于案例学习和目标驱动学习的观点生成框架[J]. 系统工程理论与实践, 1999, 19(2).
- 5 王清毅, 刘 钧. 交互式基于范例的推理及应用研究[J]. 计算机科学, 2001, 128(10): 49-52.
- 6 史忠植. 高级人工智能[M]. 北京: 科学出版社, 1998.