

分治技术

智能推理

人工智能

类比推理

(6)

计算机科学1999Vol 26No. 3

利用分治技术的智能推理^{*}

Artificial Reasoning Using the Divide-and-Conquer Technique

陈建国 潘云鹤

(浙江大学人工智能研究所 杭州 310027)

Abstract This paper introduces the divide-and-conquer technique for the artificial reasoning. The effectiveness and the flexibility of the divide-and-conquer technology are demonstrated in the field of the artificial reasoning by showing the technology applications in analogy-based reasoning, case-based reasoning and prototype-based reasoning.

Keywords Divide-and-conquer technique, Decomposition, Combination, Intelligent reasoning, AI

一、引言

“分治”技术是一种分解-合成技术,即对于复杂的不能直接求解的问题,将该问题分割成几个子问题,分别求解子问题,然后合成子问题的解,以获取整个问题的解,若分割出的子问题仍不易求解,则继续分解子问题,一直到可以直接求解为止。因此,这是一个递归模式。

同样,人工智能在求解难题时,常常将难题分解为较为简单的子问题,然后合成各个子问题的解决方案,形成一个整体方案。本文通过描述分治技术在基于类比推理、基于事例推理以及基于原型推理的应用,表明了该技术在智能推理中的有效性和灵活性。

二、类比推理中源的选择

类比被定义为一种推论,即如果两个或多个事例在某些方面相符,则它在其它方面也可能相符。类比推论的有效程度取决于已知的共享性质和推断的共享性质之间关系的类型和匹配程度。

2.1 类比设计

类比设计是一种强大的设计方法,当其它设计方法无所作为时,类比设计常常是一种有效的问题求解手段,原因是它无需大量先例和丰富的领域理论。

人们在陌生领域中求解问题时,很少盲目从事。他们依靠在熟悉领域中的解题经验,改编已知的技

术,将已求解问题的约束映射至新的问题,调整现有方案以适应当前问题的需要。针对一个新问题,人们选择一个类似的已求解的问题(源),计算源和目标问题描述之间的映射,调整源方案,并将其映射到当前领域。

类比设计生成的三个主要步骤是源选择、映射形成和推断生成。针对一个目标问题,类比设计挑选一个已有成功设计方案且与目标问题共享主要性质的源,找到了适合的源,系统就进行映射构造。一旦形成了映射,类比系统使用描述源和源→目标映射的信息推断出目标设计。但是,通常类比设计过程开销大且不灵活,大多数系统的限制因素在于驱动类比形成过程的源的选择。类比的成败很大程度上取决于它选取相关类比源的能力^[1]。

当前的大多数的类比研究主要集中在源→目标映射的建立上,在复杂的源选择上的研究较少。本文提出的相似源事例合成算法大大增加了类比设计的灵活性和实用性,方法是:当没有一个源设计能够提供足够的信息时,合成多个源设计。

这里,类比设计系统是基于图匹配算法,即比较两种用图表示的两个设计,以确定设计的相似性。对于一个以图形式表示的目标问题指标,系统的图匹配技术生成一个满足目标的设计。每一个设计在数据库均以直接的非循环的图表示。在图中的节点表示了对象的名称和属性。链接表示了对象间的关系。如果一个源图的结构与目标图的结构相匹配,则选取该源。

^{*} 国家自然科学基金资助课题。陈建国 博士生。潘云鹤 中国工程院院士,浙江大学校长,博士生导师。

系统的输入是两个子图,即目标问题的初始状态和终止状态的描述。系统搜索数据库,寻找这两个子图的最佳匹配。系统使用这两个子图的匹配输出,将源设计映射至目标领域,以产生目标问题的设计方案,如果得到的设计不成功,或找不到足够相似的源,系统试图合并多个源。这几个源相互间是相似的,且与目标问题相似。

2.2 相似源事例的合成

类比系统若要有有效工作,首先必须找到一个与目标足够相似的源,然而,在实际情况下,由于设计的多样性,找到一个足够相似的源的机会很小。尽管类比设计无需大量例子,但是众多的源事例将有助于类比^[2]。

使用多个源事例的例子有渐进类比,一个源提供目标所需的部分信息,而不是全部信息,另一个源提供剩余的所需信息。目标与这些源中的任何一个都不足以形成类比,而这两个独立的类比结果的合并将导致完整的有用的类比。

使用多个源事例的另一种方法是合并相似的源,形成一个“虚”的源。虚源比单一的事例更有利于类比,原因是它消除了异常,综合了可选的操作,并覆盖目标领域。另外,因为被合并在一起的对目标有利的有关内容保存在虚源中,所以合并源事例的类比是集中了源事例相关内容的类比。我们把合并在一起的结构相似的源事例称为相似源事例。

需要进行相似源事例的合成,即需要使用图合成算法的情况有:

1) 在目标的初始/终止状态与源的初始/终止状态之间存在匹配,但是在源中的中间步骤不与目标匹配或不能用于目标领域。

2) 源不能与目标完全匹配,但多个源与目标的匹配相当接近,并且,非匹配部分可以用一种映射至目标设计的方式进行综合,或者非匹配部分部分不重叠。

当选择用于合并的源时,算法根据综合的方便程度选择源事例。需合并图的类型有(按优先权顺序排列):

- 合并具有明显源/目标差异的图。最简单且最有利的合并是在候选的源图相互匹配,而它们与目标的差异并不重叠。如图1所示。每个源的匹配部分保留在虚拟的源图之中。

- 放松顺序约束。图匹配算法仅仅寻找图中对应层中节点之间的匹配。一个操作符的重新排序常常使源正好符合目标。比较仅仅是操作顺序不同的众多设计,很明显对操作符的调动有可能解决目

标问题。

- 子图的分离/综合。当比较具有重叠缺陷部分的子图时,可以对重叠的子图进行综合。综合的目的在于抽象非映射子图,直至经综合的子图覆盖了目标。随着在综合中源的增加,源子图的合成方法的行为类似于纯归纳。

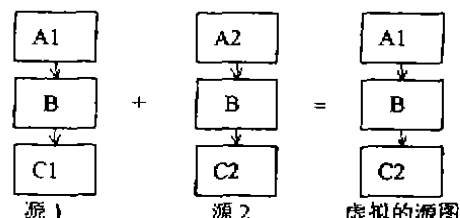


图1 源事例的合成

三、事例的分解与合成

基于事例推理(CBR)利用先例求解新问题。当以前的经验适用时,CBR方法可以有效地解决问题。传统意义上的CBR问题求解通常包括检索相关的先例,调整先例的方案。必要时,将当前情节作为一个新事例存贮,以便将来用于解决新的问题^[1]。

本文提出一个新的CBR方法,它通过有效组合多个事例,利用多个先例的片断求解新问题。在这种方法中,事例的每个片断可直接访问,同时保留了片断之间的链接。为了使多个事例在合成过程中得到有效使用,必须考虑以下几个问题:①当一个事例的片断有用时,如何检索事例。②如何寻找和分割在当前情况下有用的先例的片断。③如何综合片断以反映问题求解经验的共性。

3.1 事例的表达

当多个事例用于求解问题时,每个事例中通常只有部分片断在综合中有效^[4]。这些片断可能埋藏在为许多目标服务的扩展序列之中。例如,当需要为一个特定的假设寻找试验时,则事例中有关试验的片断才值得注意。此时,所采取的行动和其它假设并不重要。因此,事例的表达必须使它们的片断能够有效地访问。

习惯上,CBR使用的事例是作为单一实体,即一个情节是作为一个知识结构的一个范例。事例的各方面被表示成表达中的槽。一般来说,索引将事例作为一个整体来检索。将一个情节作为一个单一的事例并将所有信息都嵌入进去的做法对使用多个事例片断求解问题存在以下两个问题:

1) 检索有用的事例片断必须分两步。首先访问有关的事例,然后找到隐藏在事例内的对当前情况

有用的片断,即使事例拥有正确的索引,要找到事例内的片断还需费不少努力。如果对于当前情况,能够单步直接访问事例的有用片断,则效率肯定会提高。

2) 对一个需要进行有用综合的系统来说,单一事例包括了太多的信息。因为事例拥有许多的片断,其中一些片断可以与其它事例的片断进行综合;而其它片断是独特的,共性的综合必须推迟,等到与其它有共性的片断出现且当需要时再作综合。

在问题求解过程中,为了方便地访问事例的有用片断,最好的方法是将事例划分成片断且以片断方式保存。这样就可以使用事例的较小的片断进行推理,而无需费力地绕过较大的单一事例。片断的组织均围绕特定的目标,片断还包括了实现目标的有关信息。每个片断包括了目标实现初始时的当前问题的求解环境,其中有初始问题描述和迄今为止所采取的措施。在片断之间存在链接,从而保持了问题求解的结构。

3.2 片断的内容

每个片断被认为是较大情节的一个场合,它有三种主要信息。首先是片断发生时问题求解的环境,其次是片断围绕的目标信息,最后是片断与其它片断的链接信息。

3.2.1 环境 问题求解的环境包括在问题求解中早期采取的行为及其结果,也包括问题的属性。片断的环境有全局环境和内部环境之分。全局环境是为全局的问题情节所设的属性,内部环境是指作为问题求解的一部分的已采取的行为所建立的环境状态或知识。全局环境对事例中的每个片断都是相同的,而内部环境却是变化的。

片断的环境用于检索匹配。全局环境和内部环境对维护行动的相关性和一致性是很有必要的。因为片断包含了全局环境和内部环境,因此检索的结果是有用的相似的事例片断。

这种环境形式为组合多个片断寻找解决方案创造了有利条件。当需要访问另一个事例的片断时,拥有内部环境允许对先前的问题求解结果进行匹配。采取相似的步骤和具有类似结果的片断将受到重视。

对单一的事例,在问题求解中每一内部环境不能方便地用于片断访问。将环境保存到每个片断,即可方便地访问片断,这里是以空间换取灵活性。为达到灵活,可选的方法有:在每一点都表示内部环境,或在检索时重新计算内部环境。另外,表达内部环境还可进行分析推理,以确定当前环境是否与先例中的片断已做的工作兼容。

3.2.2 目标的实现 每个片断围绕一个目标的实现,即在片断中记录了为实现目标而采取的行为及其结果。在问题求解过程中,当检索到一个片断时,如果当前情况与以前一样,这些槽的数值将建议应采取的行为和期望的结果。

3.2.3 链接 将事例分成片断的方法使系统可以方便地直接访问事例的片断,同样,根据片断来重构事例也很重要。为了重构,片断包括与所围绕目标的源处和出处相对应的片断的链接。

片断链接在一起保持了事例中目标实现的基本结构。保持事例内部的链接对将来使用事例是很重要的。保持这样的链接使事例能够指导下一步该做什么,因为只有先例的期望与新情节的期望一致,跟随先例的推理步骤是有益的。这种指导不受特殊的临时顺序的影响,因此临时顺序并不反映先例的基本结构。

3.2.4 片断的访问 片断的访问有两种方式:直接式和顺序式。直接式访问是通过检索,将当前情节与片断的目标和环境相匹配。顺序式访问是跟随片断间的链接。通常,直接访问的检索首先是限于围绕当前目标的片断,然后使用一个加权相似度,对内部环境和全局环境内的所有属性进行匹配。顺序访问的检索很方便。片断与跟随它们的其它片断之间存在链接。当使用的片断提出一个行为时,如果结果与片断执行的结果相同,则沿着链接继续推理,跟踪至下一个目标及其行为。当顺序访问检索合适时,可以不考虑直接访问,原因是顺序访问保持了一致性,避免了不必要的处理。

3.3 优势

这种事例分解和合成方法的优点在于:

- 可方便地访问先例的所有片断,因此当需要某一片断时,可直接访问它。

- 事例的结构得以保留,因此当需要时可重构整个或部分事例。只有当前情况与先例的情况的一致,即当前步骤得到的结果与对应先例的步骤的结果相同,先例采取的行为将连续提供指导。

- 实现某一特定的目标和实现事例片断的综合不受事例中其它目标实现的影响。

这种表达可以灵活且很方便地求解问题。若不采取这种表达方法,有些问题求解可能很困难。对于一个单一的事例,可能没有合适的索引。即使有正确的事例索引,要找到事例中能应用于当前情节的合适片断也很费力。如果事例是根据内部环境的重要属性进行索引,则当事例的片断与当前情节相关时,在访问事例时有必要找到事例问题求解的某一环

节,该处的环境与当前环境最匹配。该过程应包括事例拥有的所有相同信息,即事例索引过程中每一环节上的环境直接与事例中的步骤相联系。

然而,这种方法可能引起几个问题。首先,当片断变得越来越小时,事例与情节/行为规律之间的区别有些模糊。片断表达在许多重要方面不同于决策规则。首先,前者保留了行为之间的链接,因此只要行为的结果与先例预测的一样,便可跟踪事例。顺序访问片断将导致针对目标的一致性,这在个别的决策规则集合中是不存在的。其次,片断在它们的环境中包括了初始问题描述和该环节以前所采取行为的结果,这意味着当使用直接访问时,需执行的行为选择直接受到已发生的问题求解的影响。第三,事例片断即使对当前情节仅仅只有部分匹配,也能够提出有用的建议,而此时决策规则不可用。

另一个问题是对问题求解而言,多大的片断尺寸最合适。这需要在匹配和使用单一事例作为问题方案的效率和在当前情况下可用片断的匹配综合之间达到均衡。每个片断应关注一个目标,即一个片断与一个目标叶节点相连。一个高层目标通过使用目标和子目标知识分裂成多个子目标。最低层的目标正是在实现过程中需寻求指导的目标。片断构造就是围绕这一层目标。

当使用多个事例解决一个问题的不同部分时,这种基于要实现的目标的事例划分成小片断的方法是很重要的。

四、原形推理求解复杂问题

基于原型推理(PBR)是一种集成基于模型推理(MBR)和基于事例推理(CBR)的推理方法^[5]。MBR适合于求解新问题,而CBR适合于解决经常遇到的问题。对于大型的新问题,MBR和CBR均不能有效求解,但PBR却往往大有作为。

4.1 基于模型推理和基于事例推理

MBR使用领域模型描述领域中的基本实体及其相互关系,对有关的情节根据第一原则进行推理,因此特别适合于新问题的求解。CBR从以前的问题求解情节的记忆中检索合适的相关问题,调整其解决方案以适合当前问题,所以特别适合于求解经常遇到的问题。但是,这两种方法均不能有效求解大型的新问题。

随着新问题的复杂性和规模的增加,MBR很快就变得难以处理(原因正是它根据第一原则出发进行推理);而CBR找到已保存的类似事例的可能性急剧减少,即使找到相似事例,寻找中的匹配成本是

巨大的,从而使CBR效率很低。为充分利用MBR和CBR的各自长处,克服各自的弱点,这里介绍一种结合MBR和CBR的PBR方法。

4.2 基于原型推理(PBR)

在PBR中,大型的新问题被认为是多个常见问题的新组合。CBR被用于将问题分解为几个熟悉的子问题,而MBR将各个方案组合起来以生成整体方案。PBR由三个主要阶段组成:

1)寻找熟悉的子问题。系统运用CBR寻找包含在给定问题中的所有熟悉的子问题。CBR使用根据特征索引的以前问题求解情节的组织记忆。CBR的第一步是将特征与问题的一部分进行匹配,检索出经索引的事例以达到成功匹配。因为问题中的子问题不太可能正好与一个事例精确匹配,所以必须使用部分匹配检索潜在的候选事例。与传统的CBR不同,这里的CBR方法必须处理子问题边界的识别问题。因此,这里的CBR系统将特征相匹配的问题的一部分作为种子,生成与被检索事例同等大小的子问题。第二步,针对子问题,调整被检索事例的方案。使用MBR推理部分匹配的子问题,在可能时渐进地根据被检索事例进行推理。

2)构造方案。因为可能存在部分匹配和没有清晰定义子问题边界,所以几个子问题有可能重叠。一个子问题包含了其它子问题,或一个子问题已被其它子问题覆盖等。PBR对完全覆盖了整个给定问题的子问题寻求方案的相容组合。

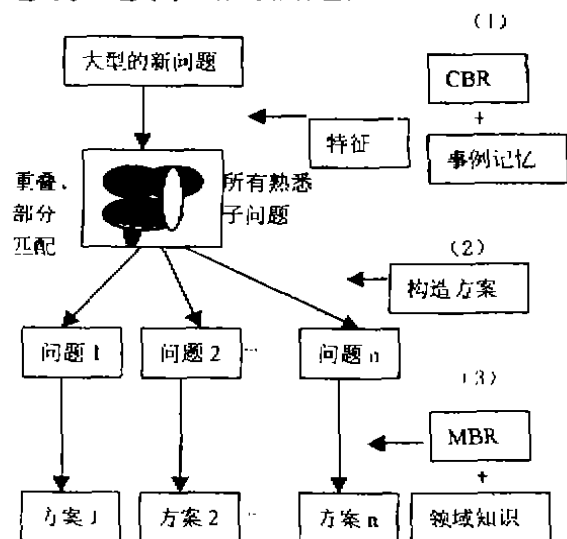


图2 基于原型推理的三个主要阶段

3)求解已构造的问题。每个相容组合被认为是一个新问题。与原来的问题相比,新问题要简单得

多,将新问题提交给 MBR,MBR 根据第一原则并使用领域模型获得新问题的解决方案。

图2表达了 PBR 的三个阶段,但是,PBR 中的分解和组合引入了几个新的难题。例如,定义子问题的边界和各个事例方案的相容组合,CBR 和 MBR 可能不得不求解许多小问题,而有些小问题根本用不着去求解。因此,对某些类型的领域和推理任务,直接使用集成方法(即 PBR)是无效的。为了使 PBR 有效地求解大型的新问题,必须严密控制这些搜索,为此,在方法、推理任务和领域方面提出了几点要求:

1)原始事例:在记忆中的事例必须是原始的,它是一个常用于构造较大问题的“构造块”(原型)。通过维护这样的原型库,找到相似子问题以及随后的分解大型问题的可能性将大大提高。另外,原型的尺寸必须比在领域中遇到的典型问题要小得多。

2)稳定的记忆:在记忆中的原型个数应保持相对稳定。在正常的问题求解中,几乎不学习新的原型。但是,PBR 允许学习更精密的观点,细致的变化或原型的改进,这些都在该原型下进行索引。对熟悉的子问题的搜索并不因学习而增加很多,因为必须先找到原型,然后当且仅当有必要时才检查其变形,视察其变形是否匹配得更好。

3)有限的组合。每个原型与其它原型的组合必须只限于一定的方式。这个局部约束用来控制对原型与问题未匹配部分部分的相容组合。

4)拓扑约束。大型问题必须提供连接信息或用于组合原型的其它信息。这进一步约束了相容组合的数量。

5)抽象。原型方案必须强调它在较大问题上的功能或行为作用,消除无关的细节。这个约束极大地限制了 MBR 指导的为寻找全局方案的搜索。

这些附加要求极大地增加了集成方法的有效性,但限制了它的应用性。然而,通过 PBR,仍可以有效解决一些有趣的且有重要意义的几类问题,例如确定大型化合物的化学结构,发现复杂化学反应的机制,找到蛋白质的遗传成分,计算复杂规则物体的质量中心等。

结论 分治技术递归地将问题分解成较小的子问题,通过解决各个子问题来达到整个问题的求解。本文分别将该技术使用于类比设计中类比源的选择、事例的分解和合成以及原型系统中。

在类比源的选择中,当没有一个事例完全与目

标匹配时,则合成多个相似的源事例。一种方法是渐进类比:一个源事例提供目标所需的一部分信息,另一个源事例提供剩余的信息。这两个类比结果的合并与目标足够匹配。另一种合成方法是组合相似源事例以形成一个与目标匹配的“虚”源事例。这两种合成方法都要求访问事例的一部分。因此,自然地要求直接或间接地访问事例的片断。

在事例的分解和合成中,事例被分解成片断,片断间存在链接,片断可以直接访问,也可顺序访问。需要时,合成多个事例的片断求解问题。由于片断间的链接关系,保存了事例的结构,因此必要时可重构整个或部分事例。因为采取了这样的事例表达方式,系统可以挑选和合成不同类型的事例片断,所以合成的结果往往具有较高的创造性。

在原型系统中,CBR 系统用于将大型的复杂问题分解成多个熟悉的子问题,而 MBR 将各个子方案组合起来形成全局方案。正是利用了分治技术,使 PBR 系统具有求解大型复杂问题的能力。

因此,分治技术在基于类比推理、基于事例推理和基于原型推理中是有效的和灵活的。我们也许可以推断,在其它智能推理乃至其它领域中,对于不能直接求解的复杂问题,该技术同样具有有效性和灵活性。

另外,因为运用分治技术分解后的子问题通常相互独立且可以分别求解,所以分治技术将引出子问题的并行处理,从而可加快整个问题的求解。

参考文献

- 1 Strublefield W A, Luger G F. Source selection for analogical reasoning an empirical approach. AAAI-96. 696~702
- 2 Cook D J. Application of analogical planning to engineering design. In: Sixth Conference on Artificial Intelligence Applications. 1990. 224~249
- 3 Watson I, Marir F. Case-based reasoning: A review The Knowledge Engineering Review, 1994, 9(4): 327~354
- 4 Redmond M. Distributed Cases for Case-Based Reasoning; Facilitation Use of Multiple Cases. AAAI-90. p304~309
- 5 Rajamoney S A, Lee H Y. Prototype-Based Reasoning An Integrated Approach to Solving Large Novel Problems. AAAI-91. p34~39