

创造性归纳 人工智能 机器学习 符号学习

62-64

# 创造性归纳综述\*

An Introduction of Constructive Induction

葛翔 周志华 陈兆乾 TP18

(南京大学计算机软件新技术国家重点实验室 南京210093)

**Abstract** In this paper we introduced the concepts and research ideas of constructive induction in machine learning, and showed the panorama of the research of constructive induction by summarizing and analyzing the existed research fruits

**Keywords** Machine learning, Induction learning, Constructive induction

## 1. 引言

机器学习是人工智能领域的一个长期的研究热点,其研究目标,就是利用计算机来模拟、实现各种形式的学习行为。学习涉及两方面的内容:知识获取和技能获取。前者是从外界获得新的信息,后者是完善已得到的知识。对于这两方面在机器学习领域中有两种主要的学习典范,一种称为符号经验式学习——SEL(Symbolic Empirical Learning)——研究知识获取的本质问题,这种方法试图从事先提供的概念例子中分析出概念相应的规则,或称为描述,这是一种从具体到一般的推理过程,我们称之为归纳,因此SEL也可以称为归纳学习。另一种是解释学习(Explanation-Based Learning——EBL)研究改善知识,获取技能的方法。EBL是一个分析、演绎式的学习过程,它运用领域知识来解释,推广学习例,期望在维持正确性和概括性的同时提高系统的效率。近年来,对各种有效的归纳学习算法的研究是机器学习的一个重要的问题。一般的归纳学习算法可以分为两类:选择性归纳学习和创造性归纳学习。

选择性归纳学习是指观察陈述中包含了描述概念所需的全部描述子,学习过程仅是对学习对象进行分类,归纳和描述。

当观察陈述中包含的描述子不适合所描述的概念时,选择性归纳学习就显得比较困难了,这时一个好的解决方法就是构造出新的描述子以适合所描述的概念,以进行创造性归纳学习。创造性归纳学习就是指观察陈述中未包含描述概念所需的全部描述子,有些描述子是在学习过程中产生出来的。学习过程包含了生成新描述子的过程。

创造性归纳这一名词是在1978年由Michalski首先提出的<sup>[1]</sup>,并且于1977年在由Michalski和Larson开发的INDUCE-1系统<sup>[2]</sup>中首次得到了实现。相对而言,创造性归纳学习的研究比较薄弱,但这并非是由于创造性归纳学习不重要,主要原因是其研究难度较大。本文将总结创造性归纳自提出以来的研究发展情况,并针对过去研究中存在的缺陷提出一点创造性归纳技术未来发展的看法。

## 2. 创造性归纳概述

在人工智能领域中,好的描述方法经常是解决困难问题的关键,这一点在机器学习里尤为突出。当被提供的训练例采用合适的描述时,学习相对要容易一些。但如果被用来描述训练例的属性对目标概念来说是不适合的,学习就会变得十分困难。选择性归纳有一个根本上的局限性,即只有原始描述和所学习的概念的特征直接相关时,学习任务才能够顺利完成,创造性归纳通过创建所需的新描述子的机制克服了这一局限,同时创造性归纳也通过修改、删除原有描述的方法使原始实例空间发生面向任务(problem-oriented)的转换。

由创造性归纳引起的实例空间的变化在某些情况下是非常重要的,首先,许多问题如果不改变表示方法是无法解决的,如果概念描述在原始的实例空间之外,那么只有通过转换或扩展原始空间才能够学习这些概念。其次,领域理论和创造性归纳减轻了示教工作的负担,教师不必要在训练例子中提供所有与任务相关的,可能对学习过程有帮助的属性,系统可以根据需要构造有用的属性在概念描述中使用。第三,实际应用中,很难保证用于描述示例的属

\* 本课题获得国家自然科学基金与江苏省自然科学基金资助。  
• 62 •

性包含了所有必需的特征,所以为了完成学习任务有时必须进行构造性归纳。

构造性归纳的基本目的是为了从整体上提高学习的效率和降低概念描述的整体复杂度。归纳学习过程是一个在概念空间中进行搜索的过程,构造性归纳转换了原始描述空间,使搜索在一个较小的范围内进行,因而提高了学习的效率。同时,新描述子的引入,特别是用原始描述子的组合形成的新描述子的引入,简化了学习结果的表示形式。所以在归纳学习系统中运用构造性归纳技术,会使得系统的学习能力得到加强,学习效率和精度都得到提高,并且也会使得知识表示变得简单。

构造性归纳按生成的新描述子对背景知识的依赖程度可分为两类:

1)弱构造性归纳:学习过程中产生的新描述子可以是必要的,但是这些新描述子实际上来源于背景知识:如 AQ 系统中的 L 规则<sup>[3]</sup>,CIGOL 中的 absorption 规则<sup>[4]</sup>。

2)强构造性归纳:学习过程中构造的是必要描述子,既不在观察陈述(即原有知识)中,也不在背景知识中:如递归逻辑程序的强构造分层学习算法中新生成的谓词 P<sup>[5]</sup>。

构造性归纳系统可以采取不同的策略来转换表示空间,根据其采取的策略可分为:数据驱动(data-driven),假设驱动(hypothesis-driven),知识驱动(knowledge-driven)和多策略(multistrategy)四类构造性归纳技术。

1)数据驱动的构造性归纳(DCI),这类系统分析处理输入数据(例子),尤其是例子中各属性之间的相互关系,在此基础上形成对描述空间的改变。如 BACON 系统<sup>[6]</sup>、AQ17-DCI 系统<sup>[7]</sup>。

2)假设驱动的构造性归纳(HCI),这类系统对描述空间进行增量式转换。每次都分析上一个过程中形成的假设,检测到可用的模式后将其作为下一次过程的属性。如 BLIP 系统<sup>[8]</sup>、CITRE 系统<sup>[9]</sup>。

3)知识驱动的构造性归纳(KCI),这类系统依靠专家提供的领域理论来构造新描述。如 AM 系统<sup>[10]</sup>、AQ15 系统<sup>[11]</sup>。

4)多策略构造性归纳(MCI),顾名思义,这类系统就是综合运用多种构造性归纳策略,如 INDUCE-1 系统<sup>[2]</sup>、STABB 系统<sup>[12]</sup>、AQ17 系统<sup>[13]</sup>。

### 3. 构造性归纳的研究成果

虽然构造性归纳的研究有很大的困难,但是由于许多从事构造性归纳的研究者不懈的努力近年来也产生了许多具有构造性归纳能力的归纳学习系统。这些构造性归纳学习系统往往要用不同的策略

来生成新的描述子,或者更一般地说是改变问题的描述空间。

我们将按这些系统使用的策略分别介绍构造性归纳的主要研究成果。

#### 3.1 数据驱动的构造性归纳系统

这类构造性归纳一般是利用一些操作符机械地把原始属性连接起来生成新的属性,这样的构造不能较有效地扩展原有的描述空间,因此往往使得学习效果没有明显的改善,但此类构造的优点是算法简单,实现起来比较容易。

Langley 等人于 1983 年开发了 BACON 系统<sup>[6]</sup>,此系统中产生的新属性是原始属性的数学函数的形式,在生成新属性的过程中采用了基于原始属性之间相互独立性的启发式策略。

Rendell 于 1985 年研制了 PLISO 系统<sup>[14]</sup>,此系统中采用在对象、结构和群关系三级抽象上进行的聚类形式从原始属性生成新的属性。

Flann 和 Dietterich 于 1986 年研制的 Wyl,IOE 系统<sup>[15]</sup>学习在棋类游戏中选择的概念的结构的描述,首先把训练例子从性能级的表示(一个棋盘)映射成学习级的表示(刻画游戏状态的概念),然后在表示上泛化它们,再把学习生成的概念转换成性能级的表示来进行有效的识别。

Schlumberger 在 1987 年开发了 STAGGER 系统<sup>[16]</sup>,它靠产生一系列描述元素(属性-值对)的布尔连接并使用一个统计函数离散化连续属性来改变描述空间。

在 Michalski 和 Falkenhainer 于 1990 年研制的 ABACUS 系统<sup>[17]</sup>中采用把数据划分成多个子群的方法,并对每个子群指定一个方程式,运用符号归纳的方法来确定各个方程式的适用条件。

Michalski 和 Bloedorn 于 1991 年开发了系统 AQ17-DCI<sup>[7]</sup>,此系统使用许多不同的逻辑和数学操作符作用于原始属性上来产生新的候选属性,这些新生成的候选属性要经过一个属性质量评价函数的评估,满足某个规定的阈值后才能加入到原始属性集里去,然后使用这个经过扩展后的属性集来进行归纳学习。此系统在实验中取得了比较好的效果,使用此系统生成的规则比只使用原始属性的系统生成的规则要来得好。在文本识别领域中,使用此系统生成的包含新属性的规则无论在测试,规则的复杂性和单个覆盖的质量等方面都要优于只使用原始属性的系统生成的规则。

#### 3.2 假设驱动的构造性归纳系统

这类构造性归纳对描述空间进行增量式转换,使描述空间能更好地不断满足学习要求,使得学习效果会有明显的改善,但这类构造中选择较好的假

设和对它进行有效的处理则是难点,这也正是构造中的关键部分。

在 Matheus 于 1989 年开发的 CITRE 系统<sup>[9]</sup>中,首先利用原始属性生成一棵二叉判定树,然后借助于分析这棵判定树来构造新的属性。值得一提的是, CITRE 系统中域理论知识的引入,从实验结果看,域理论知识的使用对于进行有效的构造性归纳是很重要的,而一些现存的系统往往忽视了它。

1990 年 Pagallo 和 Haussler 研制的 FRINGE 系统<sup>[11]</sup>依靠避免在判定树中进行重复的测试来改善判定树的质量。靠判定树的“FRINGES”(边缘)构造新的属性,并表示成布尔属性连接的形式。

1991 年 Kietz 和 Morik 开发的系统 KLUSTER<sup>[15]</sup>在有概念不能被表示或刻画时引进新的关系或概念,接着使用原有的例子集来学习新的关系或概念。

### 3.3 知识驱动的构造性归纳系统

这类构造性归纳使用领域理论来生成新的属性,也就是通过分析所提供的背景知识来转换描述空间,这样生成的新的描述空间能较好地满足学习要求,往往会有较好的学习效果,但构造中背景知识的获取、表示以及分析方法的选取往往有一定的难度。

Lenat 于 1977 年研制开发的系统 AM (Automated Mathematician)<sup>[12]</sup>依靠对下列三个方面采用预先确定的启发式策略来改变描述空间:(1)定义新的概念,新的概念表示成框架的形式;(2)为概念框架产生新的插件和相应插件值;(3)使在一个领域中开发出来的概念框架适应另一个领域。

Kokar 于 1985 年研制的系统 COPER<sup>[20]</sup>依靠对链接变元成非空间单项式运用空间分析规则来生成新的函数变元。

Michalski, Mozetic, Hong 和 Lavrac 于 1986 年开发的系统 AQ15<sup>[13]</sup>运用算术运算符(又叫 A 规则)和逻辑运算符(又叫 L 规则)来构造新的属性。

Drastal, Czako 和 Raatz 于 1989 年研制开发的系统 MIRO<sup>[21]</sup>首先运用专家定义的规则(又称为域理论)来构造一个抽象空间,然后在这个抽象空间上进行构造性归纳学习。

### 3.4 多策略构造性归纳系统

这类构造性归纳由于综合运用多种构造性策略,因此一般都会获得较好的学习效果,但由于采用

多种策略,协调好各种策略之间的关系是难点。此类构造往往有构造算法复杂,不易实现的缺点。

Larson 和 Michalski 于 1977 年开发的 INDUCE-1 系统<sup>[2]</sup>把知识驱动策略和数据驱动策略结合起来进行构造性归纳,它使用一些由用户在预先确定的规则集中选出的规则和一些固定的过程来产生新的属性。此系统是基于对训练例子的结构化描述进行分析(KCI)和对输入描述中的数字属性之间的性质上的相互独立性进行分析(DCI)的。

Muggleton 于 1987 年开发的 Duce 系统<sup>[22]</sup>提供域特征给用户,这个系统是基于一个例子对象描述集和六个转换操作符的(HCI),这些转换操作符要经过测试并由用户来确定它们的有效性(KCI)。

Knoblock 于 1990 年研制开发的 ALPINE 系统<sup>[23]</sup>使用一个领域中的操作符来构造一个固定的层次结构,或其他的固定结构,或者是抽象空间(DCI),此系统使用一些和操作符之主要影响有关的域公理和知识来避免产生不必要的限制(KCI)。

Bloedorn, Michalski 和 Wnek 于 1993 年开发的系统 AQ17<sup>[13]</sup>以相互协作配合的方式集成了 AQ15 系统, INDUCE 系统, AQ17-DCI 系统和 AQ17-HCI 系统的构造性归纳能力,也使得系统具有了这些系统的优点。

讨论 构造性归纳作为一项应用于归纳学习中的新技术,以其在归纳学习领域中表现出来的众多的优点,已证实了其对于改善学习系统的学习性能、知识表示等方面都有相当重要的作用,但是在已进行的构造性归纳研究中也暴露出一些缺点,这些缺点是进一步改进的目标,也是发展的动力。例如:数据驱动构造性归纳中,可以从学习的目标概念出发来考虑各个属性之间的关系,从而构造出能较好满足学习要求的新属性;知识驱动构造性归纳中,背景知识可采用专家提供的方法事先获取,应针对不同的背景知识而选用适合的表示方法。

从总的发展趋势来看,今后构造性归纳技术在运用策略方面会从运用单个策略向综合运用多个策略方向发展,在构造能力方面会从弱构造性归纳向强构造性归纳方向发展。对该技术的深入研究将取得丰硕的成果,并使机器学习的质量和产生质的飞跃。

(参考文献共 23 篇,略)