

行动选择的研究方法

刘雪亮 陈小平

(中国科学技术大学计算机科学与技术系 合肥230027)

摘要 “行动选择”是人工智能、博弈论和经济学等领域共同关心的一个基本问题,具有极其重要的科学意义和及其广阔的应用背景。本文分析、比较两种有代表性的行动选择理论——Maes 行动选择方法和学习分类器系统,指出它们的优劣和适用环境,为进一步研究行动选择提供了基础。

关键词 行动选择, 学习分类器, LCS, MASM, 自主体

Research Methods for Action Selection

LIU Xue-Liang CHEN Xiao-Ping

(Department of Computer Science and Technology, USTC, Hefei, 230027)

Abstract “Action Selection” is an elemental problem concerned about in the area of Artificial Intelligence, game theory and economics, which has an very important meaning of science and very broad application background. This paper will analyze and compare two typical action selection theory——Maes action selection method and learning classifier system, then we will point out their advantage and disadvantage, finally come to the conclusion.

Keywords Action selection, LCS, MASM, Agent

1 引言

“行动选择”是人工智能、博弈论和经济学等领域共同关心的一个基本问题,具有极其重要的科学意义和极其广阔的应用背景。从人工智能的观点看,智能主体需要在复杂动态环境中,利用有限的时间和资源,合理地选择自己的动作。因此,智能主体的性能在很大程度上是由行动选择机制的性能决定的。

行动选择机制的研究已经有了大量成果,提出了多种不同的思路和方法,比如基于符号主义的传统规划方法;联接主义采用的神经网络,强化学习方法^[5~8];Brooks 的反应式系统^[1~4];学习分类器系统(LCS)^[13,14,17];MASM(Maes Action Selection Mechanism)方法^[9~12]等等。根据不同的角度,现有行动选择机制可以有不同的分类标准。例如有基于时间的行动选择机制和基于状态的行动选择机制;层次式结构的行动选择机制和非层次式结构的行动选择机制;带学习功能的行动选择机制和不带学习功能的行动选择机制;反应式机制和慎思式机制等等。然而从作者掌握的文献来看,现有研究基本上局限于每种方法的内部,缺乏对各种不同方法的超越各自基本假定的分析和对比。我们相信,这样的分析对比有助于从更深层次上把握和推进现有成果。

本文具体分析和比较两种有代表性的行动选择理论:一种是结合了符号主义和联接主义方法优点的行动选择机制——MASM,另一种是 Holland 首先提出的学习分类器系统^[13]。本文第二节分析 MASM 方法;第三节分析 LCS;第四节对二者加以比较;最后给出结论。

2 MASM(Maes 行动选择机制)

MASM 理论被认为是对传统的规划理论和反应式系统的特点的融合, MASM 中行动选择的目标是建立一个智能系统——即一个主体间相互交互、合作并且从整体上系统呈

现某种全局性质的多主体系统。这里所指的主体的概念很简单,主体只是具备一定的能力的功能性模块,没有思维和意念,这样、具体不同的功能的模块可以被看作不同的主体,拓展和简化了主体的概念。针对动物行动选择问题, Tyrell 通过计算机模拟实现了该模型。

2.1 MASM 的组成

MASM 系统由大量节点组成,节点之间以及节点和外部环境、目标之间用不同类型的联接构成一个非层次式的网络结构。

主体(或称为能力模块、节点)可以描述为一个四元组 $(c_x, a_x, d_x, \alpha_x)$, c_x 表示主体的前提条件集,即 c_x 是命题逻辑中命题公式的集合,只有前提集合中的每一个条件都成立时主体才可能执行该能力模块代表的动作; a_x 和 d_x 称作增加列表和删除列表,表示当执行完能力模块代表的动作以后产生的效果,即该能力模块导致的新成立的命题集合和不能够继续成立的命题集合; α_x 表示该能力模块的激活阈值。这样,能力模块 x 在 t 时刻被选择的条件需要满足前提集为真并且这个时刻该模块的激活能力已经超过了该模块的激活阈值。

1) 网络的联接机制

·外部联接。外部环境输入:外部环境采用命题逻辑表示,如果某个环境命题 p 是能力模块 x 的前提集成员,则该命题有一个指向能力模块 x 的外部环境输入联接。目标输入:如果能力模块 x 的增加列表包含目标 g ,则目标 g 有一条指向能力模块 x 的目标输入联接。冲突目标输入:如果能力模块 x 的删除列表中包含目标 g ,则从目标 g 到能力模块 x 有一条冲突目标输入联接。

·内部联接。前驱联接:如果(i)命题 p 是能力模块 X 的前提集中的命题,并且(ii)命题 p 是能力模块 Y 的增加列表中的成员(也就是说 Y 能够帮助 X 执行的可能),那么就有一条从 X 到 Y 的前驱联接。后继联接:如果(i)命题 p 是能力模块 X 的增加列表中的命题,并且(ii)命题 p 是能力模块 Y 的前

图2a)中A的能量的增加量为 $\frac{\gamma}{\gamma} \frac{\alpha}{NM}$;b)中A的能量的增加量为 $\phi \frac{\alpha}{NM}$ 。

2.4 MASM 的特点

1)面向目标性和面向环境性 面向目标性体现在激活能量通过目标输入联接和前驱联接的反向传递上,而面向环境则体现在能量通过外部环境输入和后继联接的前向传递上。面向目标和面向环境本身是一对矛盾体,适当地调节全局参数 ϕ 和 γ 的比值可以改变系统面向目标或环境的强弱程度。

2)规划性和反应性 主体所处的外部环境如果不断地发生变化,这些变化会通过环境输入联接和系统的内部联接立即反映到系统中,主体可以实时地做出反应。所以 MASM 体现出了反应式系统的部分优点。

同时,系统也可以体现出一定的慎思能力,由于系统中的能量是逐步积累起来的,这样当主体执行某个行动序列时会有一种倾向性继续完成这个序列,即使这时有新的目标引入,但是由这个目标传递的能量还没有积累到一定的程度,因此系统体现出一定的规划性和持久性。

当然,规划性和反应性也是一对矛盾体,也可以通过调节全局参数 ϕ 和 γ 以及 π 进行折衷。

2.5 关于 MASM 的讨论

尽管 MASM 体现出了很多很好的性质,但是它本身还有很多问题没有解决。下面具体讨论 MASM 机制的一些限制和具体问题所在。Tyrrell 针对动物的行动选择机制建立一个模拟试验平台,并且检验了 MASM 机制,主要分析了 MASM 机制中“除法规则”的使用,指出了“除法规则”的不可取之处。这里不讨论 Tyrrell 的具体工作。

1)规划性:MASM 尽管具有一定的规划能力,但是这种规划性和传统人工智能中的规划理论不同。MASM 中的规划只能作为系统整体的系统体现出来,而不能用这种所谓的规划能力解决传统人工智能中的典型的规划问题,比如说猴子摘香蕉、搬积木等问题。所以,MASM 不是一步一步地解决问题,而是更适合于作为复杂自适应系统的一种动态平衡机制。

2)参数选择问题:MASM 中全局参数的选取对系统造成很大的影响,但是没有一套系统的方法或经验来告诉使用者如何选取合理的参数,这样针对具体的问题需要做大量的调试工作,直到参数满足解决问题的需要,这就为 MASM 的推广使用造成了困难。

3)最优性:MASM 中每一步选择的行动不能保证是全局最优的。即使设置了一个很大的激活能量阈值,系统还会陷入局部最优。当然,局部最优在很多情况下也是可以接受的,但是目前 MASM 还没有一套很好的证明其局部最优的方法,而通过模拟实验平台来检验又失之于准确性。

4)记忆性:MASM 中通过能力模块的当前激活能量值反映了系统对过去信息的纪录,但是这样简单的“记忆”能力功能十分有限,尽管有时避免了类似于传统的搜索树式的冗余式搜索,但当一个过去已完成的目标重新需要执行时,仍不能避免需要激活能量的重新积累过程,行动的重新选择过程。

5)学习性:MASM 没有学习机制。尽管没有学习机制,但是 MASM 可以加入我们需要的学习机制。不过,系统如何学习,如何把加入的学习机制和系统原有的能量激活或禁止传递协调统一,是今后研究的一个方向。

6)解释性:MASM 中每一个时间片行动选择网络都确定了系统的一个状态,所以也可以说是基于时间片的网络状态

转换方式,这种方式和基于逻辑推理的状态转换机制的一个根本缺陷是每一步的状态转换很难有完全合理的解释,所以从每一步来看是非理性的(或者说部分理性)。而一些使用强化学习的方法或者逻辑的方法可以做到每一步的状态转换是理性的,即每一步状态转换都有一个合理的解释。

7)除法规则:关于除法规则的缺陷可参看 Tyrrell 的工作。

3 学习分类器系统(LCS)

学习分类器系统是由 John Holland 引入的一种领域无关的自适应学习系统^[13]。本质上,它是带有参数和一些属性值的条件-行动规则。根据适应度(fitness)检测方法,学习分类器系统可以分为两大类:基于强度的学习分类器(传统的学习分类器)和基于精度的学习分类器。

学习分类器系统也被用于行动选择,尽管目前大部分的研究主要集中在设计的小环境中,也有许多研究者把 LCS 应用到自主机器人、动物行为选择的研究中。下面首先介绍学习分类器系统,然后结合 MASM 方法讨论其在行动选择方面的特点。

3.1 基于强度的学习分类器

强度是分类器对系统的有用程度的度量标志。分类器的强度应该根据系统对某一领域任务的学习过程所获得经验而不断调整。系统输入表示为三元比特串,构成元素集合是 $\{0, 1, \#\}$,其中 $\#$ 表示通配符,表示输入条件可以为1或0。输出动作表示为0,1字符串。

我们可以把传统的学习分类器系统分成三个部分:

(1)执行系统,包括输入和输出接口,消息列表和分类器列表。输入接口通过检测环境生成消息然后存储到消息列表中,分类器保存在分类器列表中并且应用到消息上。输出接口输出施用于环境的行动。

(2)学习系统,也叫信用分配系统,它使用一些学习算法比如 Q-学习等从环境中获得强化(payoff 或 reward),检查现有的分类器的有用性,然后赋予每个分类器一个数值,称作强度。

(3)规则发现系统,使用一些搜索算法生成有用的规则,删除无用或用处不大的规则。遗传算法和其他一些进化算法通常被用于规则发现中。

3.2 基于精度的分类器系统(XCS)

XCS 是目前研究得比较热的一类分类器系统,它首先是由 Wilson 引入的,XCS 区别于传统的学习分类器系统的主要特点在于,首先,XCS 对传统的分类器系统结构做了简化,它没有内部消息列表,其次 XCS 使用修改过的 Q-学习算法而不是桶队列算法^[13],最重要的是 XCS 中分类器的适合度是建立在分类器回报预测的精度上。传统的分类器系统中,强度不仅用于行动选择中的回报预测,还用于遗传算法和其他进化算法的适合度量中。但是在 XCS 中会有三个参数而不是一个:回报预测,预测误差和用于检验遗传算法的适合度。基于精度的适合度可以让 XCS 进化生成高效、精简的问题解决方案,也就是说解决方案由精确预测、最大程度通用化的分类器组成^[14~16]。

关于 XCS 的具体算法描述具体见文[14,15]。

3.2.1 XCS 的工作机理 XCS 中,分类器由条件、行动和三个主要的参数组成:条件 $C \in \{0, 1, \#\}^L$ 指明分类器可以匹配的环境输入状态($\#$ 表示通配符)。行动表示在该分类器

算法,自动完成系统的不断优化。所以说这两种方法都具有复杂自适应系统的演化特征。但是 MASM 中系统的参数的选择对系统性能的优劣影响很大,目前只能通过实验-反馈-调整参数-实验-反馈-调整参数这样循环反复的过程来优化系统的性能,需要人为地判断优劣,从这一层意义上来说,LCS 具有更强的适应性。

4. 知识表示:LCS 的本质是学习一些规则,这些规则最终用于行动选择。从最终的结果来看,LCS 能够概括总结行动选择的规律或本质,并且以直观的描述表现出来(通过规则),而 MASM 采用网络式结构,系统的行动选择完全体现在网络结点之间的连接和系统参数上,系统的整体性行动选择规律不能概括、总结出来。从这一点上来看,LCS 优于 MASM。

5. 最优性:MASM 不能保证行动选择的全局最优,而且关于 MASM 的理论研究比较难以入手。XCS 方法可以保证系统找到最优解——系统找到最少的分类器形成问题的解^[16]。不过,XCS 所找到的最优解是针对全局的,但是,在某个环境状态下,系统采用分类器所建议的行动是否就是该环境状态下的最佳行动,XCS 并不能保证。这时,XCS 能否保证选择最佳行动,需要其首先在学习过程中遍历所有的环境状态和行动组合,而这一点通常是做不到的。

6. 可扩展性:XCS 结构之上,可以包涵自适应主体的认知结构^[16],比如说规划。而 MASM 只能通过调整系统参数,从而在规划和反应式之间进行权衡。也就是说,LCS 是一个通用的行动选择方法,而系统的结构可以自行设计,系统是反应式还是慎思式体现在设计者的设计方法中,而 MASM 网络结构本身蕴含了这些特点。

7. 通用性:LCS 方法不只是行动选择方法,还可以用于其他领域,比如数据挖掘、问题求解等。而 MASM 的则局限于行动选择领域。

总结 本文介绍了两种典型的行动选择理论 MASM 和 LCS,并且进一步对比分析,讨论了它们的优缺点。从我们的分析中可以总结出,最根本上,两种方法的区别体现在学习上。抛开学习功能,LCS 在系统的适应性、知识表示、可扩展性、通用性方面都优于 MASM。所以,MASM 的应用具有很大的局限性,目前的研究,MASM 只是专门应用于动物的行动选择问题,而 LCS 则应用领域广泛,比如说问题求解、数据挖掘、机器人领域。为了改进 MASM 的缺陷,拓展 MASM 的应用,笔者认为 MASM 首先应该解决的问题是全局参数如

何选取的问题,只有这个问题解决了,MASM 才可能真正地便于应用。这个问题可以从两个方面着手:加入学习的功能,或者从模型论的角度给出全局参数间的制约关系。尽管 LCS 具有很好的特性,目前对 LCS 的很多研究还是从“小问题”(比如网格世界)研究 LCS 本身的特性,为了检验 LCS 在复杂行动选择问题环境中的真正效果,需要采用复杂模拟实验的方法检验其有效性。

参考文献

- 1 Brooks, Rodney A. A robust layered control system for a mobile robot. IEEE Journal of Robotics and Automation, 1986, 2: 14~23
- 2 Brooks, Rodney A. Intelligence without Representation. Artificial Intelligence, 1991, 47: 139~160
- 3 Brooks, Rodney A. Intelligence without Reason. In: Proc. of the 12th Intl. Joint Conf. on Artificial Intelligence (IJCAI-91), 1991
- 4 Brooks, Rodney A. Coherent Behavior from Many Adaptive Processes. In: Proc. of the Third Intl. Conf. on Simulation of Adaptive Behavior, 1994
- 5 Humphrys, Mark. W-learning: Competition among selfish Q-learners. [technical report no. 362]. University of Cambridge, Computer Laboratory, 1995
- 6 Humphrys, Mark. Towards self-organising Action Selection. In: Papers of the 14th Workshop of the UK Planning and Scheduling Special Interest Group, 1995
- 7 Humphrys, Mark. Action Selection in a hypothetical house robot: Using those RL numbers. In: Proc. of the First Intl. ICSC Symposium on Intelligent Industrial Automation (IIA-96) and Soft Computing (SOCO-96), 1996
- 8 Humphrys, Mark. Action Selection methods using Reinforcement Learning. [PhD thesis (first version)]. University of Cambridge, Computer Laboratory, 1996
- 9 Maes, Pattie. How To Do the Right Thing. Connection Science, 1989, 1: 291~323
- 10 Maes, Pattie. The dynamics of action selection. In: Proc. of the 11th Intl. Joint Conf. on Artificial Intelligence (IJCAI-89), 1989
- 11 Tyrrell, Toby. Computational Mechanisms for Action Selection. [PhD thesis]. University of Edinburgh, Centre for Cognitive Science, 1993
- 12 Tyrrell, Toby. The Use of Hierarchies for Action Selection. University of Edinburgh, Centre for Cognitive Science, 1993
- 13 约翰·H·霍兰. 隐秩序——适应性造就复杂性. 上海科技教育出版社, 2000
- 14 Wilson S W. Classifier Fitness Based on Accuracy. Evolutionary computation, 1995, 3(2): 149~175
- 15 Wilson S W. Generalization in the XCS classifier system. 1998. 665~674
- 16 Kovacs T. Evolving Optimal Populations with XCS Classifier Systems. Master's thesis, School of Computer Science, University of Birmingham, Birmingham, U. K., 1996
- 17 Wilson S W. A critical review of classifier systems. In: Proc. of the Third Intl. Conf. on Genetic Algorithms. San Mateo, CA: Morgan Kaufmann, 1992. 244~255

(上接第82页)

```

        保存 { (k-j)+1; /* j 表示当前视图的长度 */
        break; /* 退出循环 */
    }
}
3. f(k) = min { f(k-1)+1, f(j)+1 } /* f(k-1) 为上一步的结果 */
}
end if
输出 f(n), 由 f(n) 可以找出 MPRe(Q)

```

算法2对于路径查询时间复杂度为 $n \log(m)$, n 为查询的长度, m 为视图的个数。

对于树查询其最坏时间复杂度为 $(n-1) \log(m) + n * m * \log(m)$, 这种情况只能出现在视图集合中的所有视图的最后一个元素都与当前查询的最后一个元素相同, 且这些视图中没有与当前查询匹配, 这种情况是很少见的。

总结 在这篇论文里, 我们讨论了使用视图对查询进行优化, 而且与前面提到的文章不同的是我们增加了对树查询的处理能力, 使用动态规划的方法实现了视图重写, 而且我们的算法可以在 PTIME 完成。通过重写一方面可以减少搜索整个文档的次数, 另一方面对于结构相同的查询我们不必进行重复搜索, 从而有效地提高了查询效率。

参考文献

- 1 Wang Haixun, et al. ViST: A Dynamic Index Method for Querying XML Data by Tree Structures, 2003
- 2 Abiteboul S, et al. Incremental Maintenance for Materialized Views over Semistructured Data
- 3 Alstrup S, Rauhe T. Improved labeling scheme for ancestor queries. In: Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA), 2002