

基于多议题协商的贝叶斯学习

王娟, 柴玉梅

(郑州大学 信息工程学院, 河南 郑州 450052)

摘要:随着 Internet 的日益完善和电子商务的普及, 如何快速、高效地进行 agent 协商学习是必须面对和解决的一个重要问题。文中从买方 agent 的观点出发, 在协商过程中采用贝叶斯学习机制进行在线更新对方 agent 的信念, 从而缩短了协商时间, 提高了协商效率, 并实验说明了其可行性。

关键词:agent 协商; 贝叶斯; 在线学习; 多议题

中图分类号:TP18

文献标识码:A

文章编号:1005-3751(2006)02-0154-03

A Bayesian Learning Based on Multi-Issues Negotiation

WANG Juan, CHAI Yu-mei

(School of Information Engineering, Zhengzhou University, Zhengzhou 450052, China)

Abstract: With the development of Internet and the popularization of EC, how to solve the agent negotiation learning problem is an important problem. In this paper, stated the viewpoints from the respective of the buyer agent firstly. In negotiation process, used Bayesian learning to update initial belief of the opponent through the online learning. Sequentially could shorten negotiation time and advance negotiation efficiency. Moreover, testified the feasibility by experiment.

Key words: agent negotiation; Bayesian; online learning; multi-issues

0 引言

从 20 世纪 80 年代末、90 年代初开始, 多 agent 学习逐渐成为一个研究课题, 引起了多 agent 系统(MAS)、分布式人工智能(DAI)、机器学习(ML)等多个领域研究者的广泛兴趣^[1]。学习是使计算机系统具有智能的根本途径。研究多 agent 系统的学习行为对于智能系统的适应性是至关重要的。在多 agent 系统中, 由于功能和结构的复杂性, 难以甚至无法预先确定其行为准则, 需要通过学习来适应动态变化的环境。

多 agent 系统中的学习, 已成为多 agent 系统研究的一个重要内容^[2]。一方面, 由于单个智能主体能力有限, 很难完成大规模的复杂任务。多个 agent 的组合, 通过协作、协调与协商将会大大地提高系统的智能。另一方面, 随着 Internet 的逐渐普及与迅速扩大, 网络上的智能主体自然地形成一个 MAS 系统, 因此, 研究基于多智能主体的学习方法显得分外迫切。虽然多 agent 学习的技术甚至规范还都不成熟, 但其应用的前景非常广阔, 如在智能网络、智能通信、机器人及多类博弈应用等方面。

当前, 多 agent 多议题协商学习的研究重点在于采用

贝叶斯(Bayesian)学习、强化学习等学习机制进行在线学习。一般情况下, 每个 agent 所代表的个体的目标、知识、偏爱等不同, 所以多个 agent 就某一事情要达成协议就必须进行协商, 而如何提高协商效率则是 agent 协商中的一个重要研究方面。文中从买方 agent 的观点用形式化语言描述了协商过程, 并利用贝叶斯学习对卖方 agent 的信念进行了在线学习。

1 多 agent 学习

1.1 agent 协商理论基础

协商(negotiation)是 MAS 实现协同、协作、冲突消解和矛盾处理的关键环节, 是建立在 agent 通信语言之上的一种 agent 交互机制, 通过协商对某些问题达成一致意见。

多 agent 协商的研究内容主要包括协商协议、协商策略和协商处理^[3]3 方面。协商协议主要处理协商过程中 agent 间的交互; 协商策略主要涉及 agent 内的决策和控制过程; 而协商处理则侧重于对单个 agent 和多 agent 协商社会的整体协商行为的描述和分析^[4]。

协商的特点是先对所有项目在每个项目的大范围约束下进行协商以便求得共同满意的初步协议, 此时项目的约束范围变得相对较小, 然后再在共同认可的小范围约束下进行协商以便求得更进一步共同满意的协议, 重复进行直到求得共同满意的最后协议。

收稿日期: 2005-05-23

基金项目: 河南省自然科学基金资助项目(0211050110)

作者简介: 王娟(1981-), 女, 河南人, 硕士研究生, 主要研究方向为分布式人工智能与数据挖掘; 柴玉梅, 硕士生导师, 副教授, 主要研究方向为分布式人工智能与数据挖掘。

从本质上讲,协商过程是一个双方不断妥协、放宽限制的过程,参与者都希望用最小的代价获得最大的效用。在协商进行过程中,可能要经过多次的“讨价还价”才能达成最终的协定。

多议题协商是指两方或多方 agent 个体就某一个协商命题的多个相关因子进行交互,从而达成一致。协商参与者关注代价、质量等多个方面的效用。一方面,通过协商使这些议题的整合效用最大化,另一方面为了达成一致的协商结果,agent 个体也可以接受整合效用较低的结果^[5]。

1.2 agent 学习策略分析

多 agent 学习是指多个 agent 在追求一个共同的目标过程中彼此相互通讯、合作,由于获取信息而改变自身状态和周围环境,每个 agent 在学习过程中受到其他 agent 的知识、信念和意图的影响。

学习并不仅指目标的实现,同时也意味着 agent 自身以及所处环境的自我更新。一般情况下,每个 agent 在协商前都拥有初始的信念,在协商进行过程中,agent 通过在线学习不断感知周围环境的变化,更新初始的信念,及时调整自身的策略,从而达到最大的收益^[6]。

与单 agent 学习相比,多 agent 学习提出了一些新的问题,围绕着学习与交互的关系可以分为两类:为学习进行的交互和为交互而进行的学习。目前,人们更注重为学习而进行的交互。交互的 agent 通过交换信息、改变它们所处的环境,可以显著地影响其它 agent 的个体学习。特别当多个 agent 试图以团队的形式去完成单个 agent 不能完成的学习任务,即集体学习时,交互更是关键问题。在它们不同的知识、能力、经验、喜好的基础上,通过分配学习任务,追求一个共同的学习目标^[7]。

1.3 贝叶斯学习

在协商过程中若能较准确地掌握对方的信念,则协商效率会大大提高。由于 agent 和环境本身所固有的不确定性,而贝叶斯技术为不确定性问题的解决提供了理论框架,使得贝叶斯技术和 agent 技术自然地走向融合,成为当今研究的一个热点。

贝叶斯公式中蕴含有一种学习机制,使用贝叶斯公式的学习机制来改善系统的功能,就称为贝叶斯学习。在贝叶斯学习中,后验信息是由先验信息与样本数据综合得到的,可以通过样本信息使先验信息得到改善。

贝叶斯学习方法的突出特点是其学习机制可以综合先验信息和后验信息,既可避免只使用先验信息可能带来的主观偏见,和缺乏样本信息时的大量盲目搜索与计算,也可避免只使用样本信息带来的噪音的影响,只要合理地确定先验,就可以进行有效的学习。

2 agent 协商的形式化描述

2.1 协商模型的形式描述

agent 个体向对方发送一个提议(offer)并收到回应为

一次交互,而当 agent 发送 {accept, quit} 时无需回应也称为一次交互。整个协商中双方的提议以 {accept} 或 {quit} 结束。文中针对买方 agent 的协商环境使用一个五元组模型 $M = \langle G, V_a^X, F_a^X, P_a^X, T \rangle$, 其中 $G: G = \{a_1, a_2, a_3, \dots, a_n\}$ 表示协商 agent 集合, n 为 agent 的个数。文中限定 $n = 2$, 即为两方协商。协商命题 $X = \{X_m, X_s\}$ 是一个二元组, 表示参与者共同协商的命题(任务、资源), 其中 X_m 为命题的主题部分, 包括命题名称、命题标识等; $X_s = \langle x_1, x_2, \dots, x_n \rangle$ 表示该命题在协商中涉及到的 n 个议题。

$V_a^X: \forall a_i \in G, \forall t, 0 \leq t \leq k$, 其中 k 表示协商中的交互次数, $V(a_i, t, X_s) = \langle v(a_i, t, x_1), v(a_i, t, x_2), \dots, v(a_i, t, x_n) \rangle$, 表示 agent 个体 a_i 第 t 次提出的议题向量, 向量中元素 $v(a_i, t, x_m)$ 表示议题 x_m 的值。协商议题保留值向量: $\forall a_i \in G, \bar{v}_i^X = \langle \bar{v}_{a_i}^1, \bar{v}_{a_i}^2, \dots, \bar{v}_{a_i}^n \rangle$ 表示 agent 个体 a_i 在协商中保留的底线值, 协商的最终结果不能比该保留值差, 其中, $\forall x \in X, \bar{v}_i^x$ 表示议题 x 的底线值。

F_a^X : 对卖方 agent 的预测保留值向量 $F_a^X = \langle f_{a_i}^1, f_{a_i}^2, \dots, f_{a_i}^n \rangle$, 其中 $\forall a_i \in G, i \in N$, 表示 agent 个体 a_i 在开始时对对方 agent 每个议题保留值的预测结果向量。 $f_{a_i}^x$ 表示 agent 个体 a_i 在议题 x_i 上的一次预测结果。

$P_a^X: \forall a_i \in G, \forall x_m \in X$, 单议题效用函数 $P_{a_i}^{x_m}$ 表示 agent 个体 a_i 在议题 x_m 上的效用函数, 是 $v_{a_i}^{x_m}$ 到实值空间的映射, $P_{a_i}^{x_m}: |v_{a_i}^{x_m}| \rightarrow R$, 其中 $v_{a_i}^{x_m}$ 表示 agent 个体 a_i 在议题 x_m 上的取值。议题效用距离: $\forall a_i \in G, \forall v_{a_i}^x, v_{a_i}'^x \in V_a^X, P_{a_i}^X$ 为 a_i 在议题 X 上的单议题效用函数, $V(P_{a_i}^X(v_{a_i}^x), P_{a_i}^X(v_{a_i}'^x)) = P_{a_i}^X(v_{a_i}^x) - P_{a_i}^X(v_{a_i}'^x)$ 表示 agent 个体 a_i 在议题集 X 上的两个取值向量 $v_{a_i}^x, v_{a_i}'^x$ 之间的效用函数的距离。

T : 表示协商的限制时间。若协商时间超过最大限制时间 T , 则认为协商失败, 从而结束整个协商过程。

$P(f_j)$ 就是预测 f_j 成立的概率, 其初始值是由用户的一般常识所确定的。

在协商开始前的初始信念可能来自于 agent 所代表的用户的常识信息或者是通过分析学习而获得的先验知识。这些信念与 agent 的执行策略相结合, 在 agent 后续协商过程中具有重要的指导作用。

2.2 协商的一次交互过程

假定买方 agent 为 agent A, 卖方 agent 为 agent B。文中从买方 agent A 的观点出发, 在 agent B 给出一个反向提议后, 利用贝叶斯学习在线更新对 agent B 的预测保留值向量, 从而提高整个协商的效率。

假定在协商前 agent A 已经对 agent B 的协商历史使用离线(off-line)学习方法, 对 $f_{a_i}^x$ 计算了其概率 $P(f_{a_i}^x)$, 选取概率值较大的一个或多个作为初始信念。

假定 agent A 和 agent B 正在进行第 t 次交互, $0 < t < T$, T 为协商的限制时间, 则一次交互过程如下:

(1) A 收到 B 发送的提议, 如果是结束提议, 则结束整个协商过程。

(2) 如果是新的提议(议题向量值为 $v_{a_b}^x$), 则 agent A 首先利用朴素贝叶斯学习机制的概率公式计算 $P(v_0^x, v_1^x, \dots, v_n^x | f_{j_i}^x)P(f_{j_i}^x)$, 选取概率较大的值采用线性协商策略^[8]对 $F_{a_b}^x$ 和 $\bar{v}_{a_b}^x$ 进行更新, 更新后的预测保留值向量为 $F_{a_b}^x$, 议题保留值向量为 $\bar{v}_{a_b}^x$, 然后判断 $\Delta(P_{a_b}^x(v_{a_b}^x), P_{a_b}^x(\bar{v}_{a_b}^x)) \geq (P_{a_b}^x(v_{a_b}^x), P_{a_b}^x(\bar{v}_{a_b}^x))$ 成立, 若成立则协商成功, 最终的协商议题值结果值是 $v_{a_b}^x$, 否则表示对方提议不能满足 agent A 对协商命题效用的期望, 需进一步协商。

(3) 判断 $t > T$ 是否成立, 如果是则表示协商超时, 此时判断 $\Delta(P_{a_b}^x(v_{a_b}^x), P_{a_b}^x(\bar{v}_{a_b}^x)) \geq 0$, 如果成立则协商成功, 最终的协商议题值结果值是 $v_{a_b}^x$, 否则发送结束提议, 协商结束。若 $t < T$ 成立, 则需进一步协商。

(4) 进入提议推理, 根据协商推理策略放宽各议题的取值, 得到更新后的 $v_{a_b}^x$, 向 agent B 发送新的提议。

3 实验分析

为了验证该在线学习方法的效率, 根据上述模型, 设计实现了一个验证系统, 并与没有进行在线学习的协商模型进行了比较。假定 agent A 就时间和价格两个议题与 agent B 进行协商, 采用的一组试验数据如表 1 所示。

表 1 一组实验数据

Agent	议题	初始向量值	保留向量值	权值
A	TC	<47, 80>	<55, 100>	$W_{a_A}^{Time} = 0.5$
				$W_{a_A}^{Cost} = 0.5$
B	TC	<65, 115>	<50, 95>	$W_{a_B}^{Time} = 0.5$
				$W_{a_B}^{Cost} = 0.5$

其中 T 为议题 Time, C 为议题 Cost。

假定 agent B 没有进行在线学习, 则具体协商过程中议题值的变化如图 1 所示。

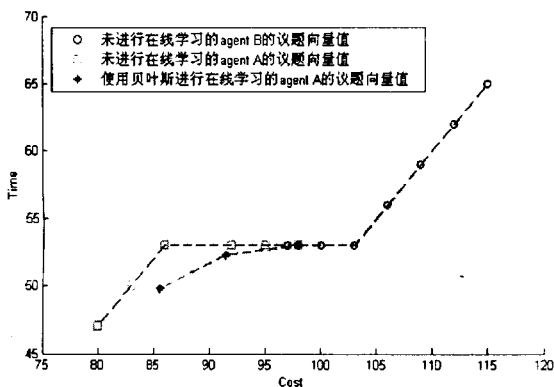


图 1 议题值变化曲线图

从图中各议题值的变化曲线可以看出, 在双方协商过程中, 采用贝叶斯在线学习策略与在协商过程中没有进行在线学习相比, 在线学习可以使协商双方在不降低整体效用的前提下尽快达成一致, 缩短了协商时间, 提高了协商效率。

4 结束语

给出了一种在协商过程中利用贝叶斯进行在线学习的策略, 它有助于提高多 agent 多议题的协商效率, 缩短协商时间。算法从对方 agent 发送的提议中分析、提取、计算, 最后利用文献[8]中的线性策略更新了对对方 agent 可能的保留价格和自己发送的 offer, 从而使协商时间缩短, 协商效率提高, 并实验证明了该算法的可行性。

虽然贝叶斯学习能有效地综合先验知识和后验知识, 但在多 agent 系统中, 贝叶斯学习仍有其局限性, 例如需要可操作的完整的先验理论, 在许多情况下先验分布的合理性和准确性难以评价等。在此基础上, 深入研究多 agent 多议题协商过程中采用贝叶斯网络学习机制进行在线学习将是下一步的研究方向。

参考文献:

- [1] Weiss G, Sen S. Adaption and learning in multi-agent systems[Z]. Lecture Notes in Artificial Intelligence 1042, Berlin: Springer-Verlag, 1996. 22-39.
- [2] Karnik N M, Tripathi A R. Design Issues in Mobile-Agent Programming Systems[J]. IEEE Concurrency, 1998, 6(3): 52-61.
- [3] Kraus S, Sycara A, Evenchik A. Reaching agreements through argumentation: a logical model and implementation[J]. Artificial Intelligence journal, 1998, 104: 1-69.
- [4] Woodriddle M, Jennings N R. Pitfalls of agent-oriented development[A]. In: Proc. of Second Intl. Conf. on Autonomous Agents (Agent98)[C]. Minneapolis, MN: [s. n.], 1998.
- [5] Fatima S S, Wooldridge M J, Jennings N R. Multi-Issue negotiation under time constraints[A]. In Proceedings of 1st International Joint Conference on Autonomous Agents and Multi-Agent Systems[A]. Bologna, Italy: [s. n.], 2002. 143-150.
- [6] Oprea M. An Adaptive Negotiation Model for Agent-Based Electronic Commerce[D]. Romania: Department of Informatics, University of Ploiesti, 2000.
- [7] Stone P. Layered learning in multi-agent systems[D]. Pittsburgh, PA: Carnegie Mellon University, 1998. 98-187.
- [8] Faratin P, Sierra C, Jennings N R. Negotiation Decision Functions for Autonomous Agents[Z]. London Dept. Electronic Engineering, Queen Mary and Westfield College, University of London, 1997.