

多主体智能系统与自主计算

廖备水摇著

吉林人民出版社

图书在版编目(CIP)数据

多主体智能系统与自主计算/廖备水著.—长春:吉林人民出版社,2007.12
ISBN 978-7-206-05488-4

I.多… II.廖… III.人工智能—计算
IV.TP183

中国版本图书馆 CIP 数据核字(2007)第 202081 号

多主体智能系统与自主计算

著 者:廖备水

责任编辑:谷艳秋 封面设计:创意广告

吉林人民出版社出版 发行(长春市人民大街 7548 号 邮政编码:130022)

印 刷:长春永恒印业有限公司

开 本:850mm×1168mm 1/32

印 张:8.75 字数:220 千字

标准书号:ISBN 978-7-206-05488-4

版 次:2007 年 12 月第 1 版 印 次:2007 年 12 月第 1 次印刷

印 数:1-3 000 册 定 价:20.00 元

如发现印装质量问题,影响阅读,请与印刷厂联系调换。

目 录

第 1 章 绪论	(1)
1.1 多主体智能系统概述	(1)
1.1.1 智能主体	(1)
1.1.2 多主体的协调与协作	(2)
1.2 自主计算概述	(3)
1.2.1 自主计算的概念和定义	(3)
1.2.2 自主计算的研究现状	(4)
1.3 用多主体技术构造自主计算系统	(5)
第 2 章 预备知识	(6)
2.1 面向服务的体系结构和实现技术	(6)
2.1.1 面向服务的体系结构	(6)
2.1.2 实现技术	(7)
2.1.3 分析和小结	(8)
2.2 基于策略的管理	(8)
2.2.1 策略的定义	(8)
2.2.2 典型的策略方法	(9)
2.2.3 分析和小结	(9)
2.3 知识表示语言	(9)
2.3.1 本体表示语言	(9)
2.3.2 推理表示语言	(9)
2.3.3 策略表示语言 (悦 悦)	(9)

源缘	服务和合同的定义和表示	(缘原)
第猿章	自主计算概念模型与实现方法	(远园)
猿缘	自主计算概念模型	(远园)
猿缘	自主元素	(远园)
猿缘	自主计算系统	(苑员)
猿缘	自主计算系统与多主体系统的区别	(苑愿)
猿缘	两类主要的自主计算实现方法	(愿园)
猿缘	知识模型方法	(愿员)
猿缘	数学模型方法	(愿缘)
猿缘	比较与分析	(愿怨)
猿缘	结语	(怨园)
第源章	基于策略和合同的扩展主体模型	(怨原)
源缘	扩展主体的概念	(怨原)
源缘	用 ^身 谓逻辑	(怨苑)
源缘	语法	(怨愿)
源缘	语义	(怨怨)
源缘	基本公理	(员园)
源缘	谓词原 ^身 谓的体系结构模型	(员远)
源缘	体系结构	(员远)
源缘	相关元素的表示	(员员)
源缘	各模块工作机制	(员猿)
源缘	实例	(员园)
源缘	主体模型设计	(员缘)
源缘	主体决策	(员愿)
源缘	规划与动作执行	(员员)
源缘	结语	(员原)
第缘章	基于可废止逻辑的柔性主体	(员缘)
缘缘	柔性主体体系结构	(员缘)

基于可废止逻辑的柔性主体模型	(158)
预备知识：可废止逻辑	(158)
柔性主体理论	(159)
柔性主体基本模块工作原理	(159)
实例	(160)
结语	(161)
第 4 章 基于主体技术的自主计算系统	(162)
自主计算系统体系结构	(162)
策略的表示、细化和部署	(163)
面向服务的自主计算系统中的策略	
概念	(163)
策略管理框架	(165)
新型多主体智能系统	(166)
多主体系统的组成	(166)
多主体系统的体系结构	(167)
主体的协作机制	(167)
基于多主体智能系统的自主计算	(168)
系统自组织和服务自主合成机制	(168)
异常处理和系统自修复	(169)
基于授权和认证机制的系统自保护	(169)
结语	(170)
主要参考文献	(170)
附录	(170)
后记	(171)

图表目录

图 员原: 月网主体的一般结构	愿
图 圆原: 面向服务的体系结构协作关系	愿
图 圆原: 宰藻服务基础设施	猿
图 圆原: 基于阅粤蕴的语义宰藻服务体系结构	猿
图 圆原: 信息内容语义的建模框架	源
图 圆原: 协商服务及其特性的本体论描述	缘
图 猿原: 自主元素概念体系结构	远
图 猿原: 自主管理者工作机制	远
图 猿原: 粤酝概念体系结构	苑
图 猿原: 粤酝组织结构	苑
图 猿原: 粤酝编排模式	苑
图 猿原: 一种基于反应主体的自主元素	愿
图 猿原: 一种基于慎思主体的自主元素	愿
图 猿原: 一种基于自适应控制理论的自主元素	愿
图 猿原: 一种基于效用函数的自主元素	愿
图 源原: 孕悦原主体的决策过程	怨
图 源原: 信念世界和愿望世界 (例)	员
图 源原: 策略型义务世界和合同型义务世界 (例)	员
图 源原: 目标世界和意图世界 (例)	员
图 源原: 孕悦原粤酝体系结构	员
图 源原: 小型会议安排服务合成规划	员

图 源景范: 服务动态自组织规划	源匠
图 源景愿: 孕悦原粤藻城事件分类体系	源景
图 源景怨: 策略生命周期状态转换图	源苑
图 源景园: 合同引擎内部组成	源苑
图 源景员: 小型会议安排虚拟组织示意图	源景
图 缘景员: 基于可废止逻辑的柔性主体体系结构	源苑
图 缘景圆: 可废止理论的结论推理规则	源园
图 远景员: 柔性主体的自主计算系统层次结构图	源源
图 远景圆: 基于功能的策略概念分类	源苑
图 远景猿: 策略管理框架	源怨
图 远景原: 领域本体的例子	源景
图 远景缘: 基于处方和本体论的策略细化机制	源源
图 远景云: 策略细化过程树	源园
图 远景范: 自主管理层体系结构图	源怨
图 远景愿: 主体的协作关系	源园
图 远景怨: 应用型主体启动的协作过程	源景
图 远景园: 孕悦原粤藻城进行服务动态合成的自组织 过程	源猿
图 远景员: 服务的合成	源缘
图 远景圆: 服务异常本体	源远
表 猿景员: 自主计算系统与多主体系统的比较	源园
表 猿景圆: 基于各种不同技术的自主系统比较	源怨
表 源景员: 小型服务安排服务 瀑	源愿
表 源景圆: 意图的产生方式	源怨
表 源景猿: 意图的执行方式	源园
表 源景原: 酝粤原粤藻藻在时间点 燥拥有的子服务 列表	源园

表 源象: 合同 糟和 糟 (非形式化表示)	员蒙
表 源象: 小型会议安排服务 溪	员蒙
表 缘原: 柔性主体规则冲突处理策略	员怨
表 缘原: 自 赚至 赚, 主体 葬葬基本模块工作状况	员怨
表 远原: 角色指派信息	圆怨

第1章 绪论

目前，多主体智能系统与自主计算是两个相互关联的研究领域。智能主体的反应性、预动性和社会行为能力等特点使之成为构建自主计算系统的理想候选。不过，为了适应自主元素和自主计算系统的特性要求，需要对传统智能主体的体系结构和推理决策机制进行革新。

本章首先介绍多主体智能系统和自主计算的基本概念、理论和研究现状，然后提出运用多主体技术来构造自主计算系统的基本思想。

1.1 多主体智能系统概述

1.1.1 智能主体

1.1.1.1 基本概念

智能主体(Intelligent Agent, 简称为“主体”)技术作为人工智能的重要分支，已经得到了广泛而深入的研究。目前对主体的定义尚存在不同的看法。依据 Wooldridge 和 Jennings 的观点 (Wooldridge 1995)，主体定义为：

主体是处在某个环境中的计算机系统，该系统有能力在这个环境中自主行动以实现其设计目标。

一般来说，智能主体应该具有以下特性 (Jennings 2000)：

- ✓ 反应性：智能主体可以感知它们的环境，并可以对环境

发生的变化以即时的方式做出反应，以满足它们的设计目标；

- ✓ 预动性: 智能主体通过主动发起可以表现出目标引导的行为，以满足它们的设计目标可以主动发起一个动作；
- ✓ 社会行为能力: 智能主体可以与其他主体(也可能是人)交互，以满足它们的设计目标。

主体可以分为多种类型。从建造主体的角度出发，单个主体的结构通常分为慎思型主体、反应型主体和混合型主体 (Wooldridge 1995; Jennings 2000)。

慎思型主体的最大特点就是主体看作是一种意识系统 (Intentional System)。人们设计的基于主体系统的目的之一是把它们作为人类个体或社会行为的智能代理，那么主体就应该 (或必须) 能模拟或表现出被代理者具有的所谓意识态度 (Intentional Stance)，如信念、愿望、意图 (包括联合意图)、目标、承诺、责任等。Shoham 等人根据应用的特点将意识态度分为信息类 (刻画主体所具有的信息)、动机类 (主体的动作选择)、社会类 (与主体的社会、道德和理性行为有关) 和其他类 (如欢喜、恐惧等)。Wooldridge 等人则把单个主体应具有的意识态度 (或精神状态) 分为两大类: 信息态度和积极态度，前者是指主体所拥有的关于自己、环境及其他主体的信息和知识，如信念和知识等，后者是指那些能导致主体执行动作的状态，如愿望、目标、意图、承诺、责任、能力等。他们还指出，一个理性的主体总是基于信息态度去采取积极态度，如根据信念去选择目标和形成意图。在 Rao 和 Georgeff (Rao 1991; Rao 1995) 提出的 BDI (Belief-Desire-Intention: 信念-期望-意图) 模型中，用信念、愿望和意图这三类意识态度来刻画主体的结构，并最终通过规划库来研究 BDI 模型的抽象性

质和推理过程。其中，“信念”对应于主体具有的关于世界的信息；“愿望”表达主体想要（期望）达到的事件的状态；“意图”则表示主体承诺去完成的愿望。BDI 逻辑为主体体系结构以及主体系统的建立奠定了基础，比如 IRMA (Bratman 1988) 和 PRS (Fischer 1996) 就是典型的例子。

反应型主体是指主体的智能取决于感知和行动，主体不需要知识，不需要表示，也不需要推理，主体可以像人类一样逐步进化，主体的行为只能在现实世界与周围环境的交互作用中表现出来。反应型主体有着简单、经济、易计算性、预防失败的健壮性以及精巧性等优点，但其智能程度较低，也缺乏足够的灵活性。混合型主体综合了上述二者的优点，具有较强的灵活性和快速响应性。

混合结构的系统通常被设计成至少包括如下两部分的层次结构：高层是一个包含符号世界模型的认知层，它用传统符号 AI 的方式处理规划和进行决策；下层是一个能快速响应和处理环境中突发事件的反应层，它不使用任何符号表示和推理系统。反应层通常被给予更高的优先级。采用分层结构时要处理的主要问题是，各层应采用什么样的控制框架以及各层之间应如何交互。

在以上各类主体中，慎思型主体占据主导地位，同时 BDI 结构以其坚实的理论基础和方便的可操作性而成为目前研究和应用领域中使用最多的主体结构。

1.1.1.2 传统的 BDI 逻辑

传统的 BDI (信念-愿望-意图) 逻辑对主体技术的发展奠定了坚实的理论基础。在 BDI 逻辑中，*B* (信念) 是主体对环境及其自身的了解，*D* (愿望) 是主体希望达到的目标，*I* (意图) 是主体从当前的目标集中选择并承诺去执行的目标。BDI 逻辑通过

采用可能世界语义来反映客观世界的不确定性，并依此来自主决策主体在不同情境下应该实现的目标，以完成系统的设计任务。

BDI 概念的哲学观点源自于 Michael Bratman，他将意图看成部分行为计划，这些计划是主体承诺执行以完成其目标的计划。Bratman 的理论特别关注意图在推理过程中的作用。他指出意图之所以重要是因为意图限制了推理过程中行动的选择。例如，假定某人已经有去写一本书的意图，那么当他在决定做什么时，就不需要去考虑与该意图相冲突的动作（比如去度假或享受社会生活）。这样，他在决策时就减少了许多可能的选项，使得决策更加简单。由于任何现实的主体，尤其在计算机中实现的主体，都是资源受限的，因此基于意图的主体模型更有吸引力。基于 Bratman 的哲学观点，出现了许多关于 BDI 主体的形式化工作，其中最具有代表性的包括 Cohen 和 Levesque 的意图逻辑（Cohen 1990）以及 Rao 和 Georgeff 的 BDI 逻辑（Rao 1991；Rao 1995；Rao 1998），关于这些逻辑的具体介绍略，有兴趣的读者请查阅相关文献。

1.1.1.3 社会概念扩展的 BDI 逻辑

近年来，为了使得开放的环境中主体之间协作更加明确、稳定和可预见，规范(Norms)、合同(Contracts)和义务(Obligations)等概念被引入主体模型（Broersen 2003；Dignum1999；Fasli 2004 等）。经过扩展的 BDI 逻辑使得主体在决策时不仅考虑到内部的动机（依据环境的变化、为了达到设计目标而期望采取的动作），也考虑其它主体（通过合同）和社会（通过规范）的动机，这些动机以义务的形式体现。

许多研究者指出，在设计多主体系统时，为了支持丰富的协作行为，给个体主体提供各种形式的社会意识（Social awareness）是重要的。与社会意识相关的技术概念包括联合与共享规划

(Tambe 1997)、惯例和社会责任 (Jennings 1997)、社会承诺 (Castelfranchi 1998)、社会法律 (Shoham 1995)、用义务 (Obligation) 进行推理 (Barbuceanu 1998) 等。目前, 基于规范和义务的扩展 BDI 模型受到了很大的关注。把规范和义务整合进 BDI 主体推理过程, 可以提高推理的灵活性。在人类社会, 规范 (Norm) 有助于标准化个体的行为, 使得特定社会中的协作或交互变得更加容易。同样, 在主体社会, 把主体设计成趋向于遵守规范的实体, 主体之间的行为在某种程度上可以被相互预料, 因此将使得主体之间的协作变得简单。另一方面, 义务 (职责) 与具体的实施策略相关联, 违约者可以被处罚。在这种意义上, 义务是影响其他主体行为的显式工具, 在主体交互的过程中提高稳定性和可靠性。同时, 义务也提供了一定程度的自由选择权, 即对选择的结果事先做出明确的说明。规范如何影响主体的行为呢? 规范不能作为主体可能目标的过滤器, 因为如果这样的话, 主体将总要遵守可行的规范, 而我们希望的是: 是否遵守规范应该是有动机和有意识的决策。基于这些认识, 近年来许多学者提出了基于规范和义务的扩展 BDI 主体模型 (Dignum 2004; Broersen 2003; Boella 2004A 等)。

(Dignum 2002) 指出, 传统的 BDI 主体模型缺乏机制来明确表达 (推理) 义务和规范对主体目标和意图的影响。由于这个原因, 传统的 BDI 理论与各种实际主体结构存在差距。因此, 他们的目标是用社会概念 (诸如规范、义务等), 从理论和实际两个层次来扩展 BDI 模型, 提出了一种义务和规范扩展的 BDI 主体模型, 称为 B-DOING 主体。该模型区分愿望、义务、规范和目标。其中, 愿望是产生于主体内部的动机; 义务和规范是与其他主体交互时产生的动机; 而目标是从前面两类动机经过决策后生成的动机 (决定各种动机的优先次序)。被采纳的目标作为进一步

慎思以及手段一目的推理的基础。义务经常产生于一对主体交互的结果，比如承诺、协定、合同等等，因此它直接与其他个体相关，并与违约处罚相联系。义务作为清晰的机制来影响主体的行为，在主体之间相互作用时，义务提供了一定的稳定性和可靠性，同时允许一定的灵活性。义务用已知的明确后果给主体提供一种自由选择的权利。另一方面，规范可以帮助个体行为的标准化，使得主体社会中主体之间的合作和交互变得更容易。义务和规范的区别在于规范更加稳定和抽象，且是主体所在的组织所固有的，而义务通常是主体本身直接行动的结果。在 B-DOING 模型中，问题的核心是如何处理各种动机之间的冲突，即要在主体内部愿望、其他主体的兴趣（义务）以及社会的期望（规范）之间找到平衡。B-DOING 逻辑采用扩展的 Prohairesic 道义逻辑(PDL)(van der Torre 1999) 来表示义务和规范，形式化地， $N^x(p|q)$ 表示“它是某个社会或组织 x 的一条规范：如果 q 成立那么 p 应该成立”； $O_{a,b}^x(p|q)$ 表示“如果 q 成立，那么主体 a 有义务对 b 使得 p 成立”。其中， x 是社会或组织，负责违约处罚的实施。这两个模态算子的语义都是基于可能世界之上的优先排序。规范的优先排序基于情景的社会利益，而义务的优先排序则基于违约的处罚。这样，个体主体的行为至少受到三个方面的影响：个体主体的自身效用函数、社会所有成员的效用函数、社会的凝聚力程度。基于优先级排序，B-DOING 逻辑通过目标维持和意图维持两个阶段推进主体的决策。

(Broersen 2003) 提出一种扩展的 BDI_{CTL} 逻辑 ($BDIO_{CTL}$)。该逻辑以 Rao 和 Georgeff 的 BDI_{CTL} 逻辑为基础，引入标准道义逻辑模态算子 $O_i(\alpha)$ ：“主体 i 有义务使得公式 α 成立”，讨论了四个模态 (B 、 D 、 I 和 O) 之间可能具有的特性。除了 Rao 和 Georgeff 的实在性特性 ($B_i p \rightarrow D_i p$; $D_i p \rightarrow \neg B_i \neg p$; $D_i EF p \rightarrow B_i EF p$) 和承诺

策略 ($I_iAFp \rightarrow A(I_iAFpUB_i p)$; $I_iAFp \rightarrow A(I_iAFpU(B_i p \vee \neg B_i EFp))$) (式中的 E 、 A 、 F 和 U 是时态算子), BDI_{CTL} 逻辑定义了个体主体及多主体的义务特性。这些特性指出主体在不同的背景下, 义务算子与其他模态算子之间可能有不同的关系。例如, 对于“礼貌的”主体, 它可能具有特性: $O_i p \rightarrow D_i p$, 即如果主体 i 有义务使得 p 成立, 那么主体 i 期望 p 成立, 而对于“自私的”主体, 有特性: $D_i p \rightarrow O_i p$ 。另外, 对于多主体系统, 也有不同的特性, 例如, 对于主仆关系有特性: $B_i(I_j(\alpha)) \rightarrow O_i(\alpha)$, 即如果主体 i 相信其“主人”主体 j 有意图使得 α 成立, 那么主体 i 有义务使得 α 成立。

(Boella 2004A) 指出主体的组织结构不应该事先被完全固定, 这样有利于适应动态的环境。该文提出将合同作为多主体组织的法律制度, 使得主体在特定社会中交互与合作的变更可以变得清晰化: 主体创建职责 (Obligations)、许可 (Permissions) 以及新的交互可能性。从合同的观点, 组织可以看作是为了满足自利个体不同利益的协议的可能集合。(Boella 2004A) 叙述了合同的形式化表示以及如何通过合同来修改组织的行为。

1.1.1.4 BDI 主体结构

BDI 主体模型根源于人类的实用推理 (Practical Reasoning)。实用推理包含三个重要的过程: 决定实现什么目标 (慎思); 如何实现这些目标 (手段-目的分析); 以及当环境发生变化时, 如何调整所追求的目标以适应变化了的环境 (反应性)。基于这三个过程的 BDI 主体的一般结构可以表示为图 1-1 所示 (姚莉 2002)。

第一, 主体的慎思过程可以用两个函数来建模: 选择生成函数和过滤函数。假设用 B 表示主体当前信念的变元, Bel 表示所有信念的集合; 用 D 作为愿望的变元, Des 表示所有愿望的集合; 用 I 代表主体的意图, Int 表示所有可能的意图的集合。选择生成

函数 $options$ 可以表示如下（式中 \wp 表示幂集）：

$$options : \wp(Bel) \times \wp(Int) \rightarrow \wp(Des)$$

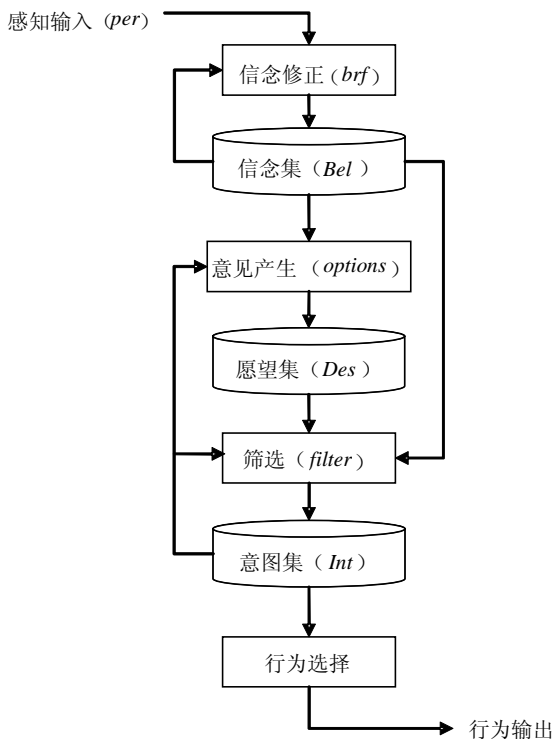


图 1-1: BDI 主体的一般结构

这个函数接受主体当前的信念和意图，以此为基础产生可能的选择或愿望的集合。另一方面，为了从竞争的选项中做出选择，主体使用过滤函数。直觉上，过滤函数必须做出“最佳”的选择供

主体做出承诺。过滤函数 *filter* 表示如下：

$$filter : \wp(Bel) \times \wp(Des) \times \wp(Int) \rightarrow \wp(Int)$$

另外，主体的信念更新过程通过信念修正函数建模（式中 *per* 表示感知输入）：

$$brf : \wp(Bel) \times per \rightarrow \wp(Bel)$$

第二，手段一目的推理是使用已有的手段（即可以采取的动作）决定如何实现目标（即意图）的过程。手段一目的推理的过程就是产生一个规划的过程。规划本质上是自动程序设计，它接受下列输入：意图、当前的环境状态（主体的信念）和主体可以采取的动作，其输出则是一个规划序列。最后，主体执行该规划序列，以实现特定的目标（意图）。

第三，当环境发生改变时，主体进行信念修正并实施新一轮的慎思和手段一目的推理。该过程的主要问题是在承诺一个意图和取消一个已承诺的意图之间怎样达到平衡。一般来说，没有足够的理由，主体不会放弃一个意图。意图会一直保持，直到相信已经成功实现了意图、相信不能实现这个意图，或者相信实现这个意图的理由已经不存在为止。

1.1.2 多主体的协调与协作

一般来说，多主体系统（MAS, Multi-agent System）是由多个主体组成的主体社会，是一种分布式的自治系统。MAS 系统的研究涉及的范围很广泛，包括主体通讯、组织形成以及各种协作机制等。其中，主体之间的协调(Nwana 1996)和协作(Doran 1997)是 MAS 研究的主要方面。

在开放、动态、异构的多主体系统中，为了避免出现混乱状