

智能 Agent 自我的统一与分离

李金厚*

(华东冶金学院计算机科学系 安徽 马鞍山 243002)

【摘要】 在以自我表示 Agent 精神中心思想的基础上, 进一步提出在智能 Agent 行为与活动过程中, 对外其自我表现出统一, 而在其内部却常表现出分离的新观点, 揭示出智能 Agent 自我需要和自我愿望等多重性特征, 并对 Agent 自我分离的特点和性质进行了分析和讨论, 为将智能 Agent 行为与其结构联系起来奠定了基础。

关键词 智能 Agent; 统一与分离; 表达连接; 表达替换; 意识状态

中图分类号 TP18

智能 Agent 是 MAS 研究和基于 Agent 的智能办公网络研究的基础^[1,2], 其研究重点是 Agent 的理论和结构。一般认为将 Agent 作为意识系统进行研究是合理且有用的。Haddadi 认为可将通常应用于人类自身行为描述的一些认知概念如信念、意图、承诺等赋予 Agent 系统, 称为意识系统^[1]。Wooldridge 等认为意识态度和意识系统是一种抽象工具, 为人们描述、解释和预测复杂系统的行为提供了一种方便而熟悉的工具^[2]。Rao 和 Georgeff 提出了 BDI 模型, 用信念、愿望和意图来刻画 Agent 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^[3]。符号 AI 虽然从某种程度上可以解释 Agent 行为与条件的逻辑关系和推理过程, 但缺乏内部实现结构和机制的支持。

Kiss 将意识属性分为三类^[4]: 1) 认知类: 指通常意义下的认知概念, 如信念、知识、意识等; 2) 意图类, 指在某种企图驱动下去执行某一动作或控制, 如意图、承诺、规划等; 3) 偏好类: 指与 Agent 动态行为相关的意识属性, 如目标、期望、偏爱等。这一分类方法将 Agent 行为与其内部意识属性联系起来。Kiss 又利用动力学系统规范行为 Agent 的目标、喜好等基本概念, 给出了一个理论模型^[5], 为形式化描述 Agent 和 MA 提供了基础, 但未能提供具体的实现过程。本文用 EGO 表示 Agent 自我的所在, 突出强调了 Agent 自我在其行为与活动过程中所起的不可替代的作用。指出 Agent 自我对外是统一的, 对内却是分离的。在此基础上, 给出了基于表达的 Agent 自我协商模式, 对进一步了解 Agent 内部结构、行为及其实现过程具有重要的意义。

1 智能 Agent 自我的统一、分离及其形式化描述模型

1.1 Agent 自我的多重性特征

从心理学角度来看, 人的心理、自我常常表现出多面性, 即变化的一面。美国心理学家 A. H. 马斯洛将人类的需要分为五个层次。他认为需要满足是人类发展的一个简单原则。不过, 在人类的复杂行为中, 除了需要, 欲望和利益等在行为决策中也起着重要的作用。人类自我需要和自我愿望等多重性特征决定了人的行为具有取舍的特点, 即满足自我的某种要求而放弃其他。这种取舍需要一个过程, 即自我需要、愿望或利益等思想争夺过程。这种争夺过程在精神上会有所体现, 经常还会伴随情绪的变化。激烈的思想争斗会导致精神的紧张状态, 紧张、难以调和的思想争斗在超过个体承受力时甚至会破坏自我的统一。

同样, 对具有复杂行为的智能 Agent 来说, 具有自我需要和自我愿望等多重性是合理的。对外 Agent 自我是统一的, Agent 将朝着有利于生存的方向进行行为的取舍^[6]; 而在 Agent 内部, 多

2000年7月3日收稿

* 男 35岁 硕士 讲师

重需要、愿望之间存在协商与取舍的过程,在这一过程中, Agent 自我必然表现出分离的一面。

1.2 Agent 自我的统一与分离及其形式化描述模型

在 Agent 正常精神状态下, Agent 对外自我是统一的,其所作的行为决策有利于生存发展的方向。但由于 Agent 自我多重性特征,在自我需要、愿望等发生分歧时,其内部要有一个自我的协商过程,从中选择出最适合生存的决策作为行为依据,在这一过程中势必要求 Agent 自我是分离的。

文献[7]提出了一种基于 Agent 内部各系统环节功能的形式化描述模型。在该模型中, Agent: $=\langle S_i, F_s, V_L(F_s), I_r, I_s \rangle$ 。其中 S_i 表示 Agent 内部物理构成; F_s 是 S_i 依据各个系统环节功能的抽象,表示 Agent 的逻辑结构; $V_L(F_s)$ 是取值函数,可以对 F_s 内部各个不同系统环节进行取值,从而了解该环节的状态、行为等; I_r 是 Agent 对自我行为与活动的表述,即自知的活动记录, I_r 是按时间展开的序列,该序列中既有有意识的行为、活动表述,又有一些自我的无意识状态,与意识的转换过程等相对应; I_s 是在其他 Agent 角度观察到的该 Agent 的行为与活动。对 F_s 有 $F_s:: =\langle EGO, Obj, Psh, FL, SS_L, PC_L, RS_L, AR_L, SC_L, MD_s, IT_L, AcT_L, A_s, A_p, A_b, RF_L, \dots \rangle$ 。其中 EGO 表示一个假想的精神中心,代表 Agent 自我的所在; AR_L 表示协商环节,虽然 Agent 的自我中心是 EGO 统一的,但 Agent 的自我会有多种不同的需要、愿望等, AR_L 环节允许 Agent 在不同需要、愿望之间相互协调以作出更合理的决策; SC_L 表示意识状态转换环节,因为 Agent 所处的思维状态是经常变化的,在下一个时刻到来前, Agent 都面临着这样的选择:是继续当前思维状态不变还是转换到别的思维状态,这种变与不变便是通过 SC_L 环节来实现的; A_s 表示感觉信息投射区; A_p 表示知觉信息投射区。感觉的信息要让 EGO 注意到必须经过知觉环节将有关信息输出到 A_p ; A_b 表示行为信息投射区,是集中各类行为指令的地方,与动作有关的信息会经过 A_b 交 AcT_L 执行,而与思维活动有关的信息会经过 A_b 投射到 Agent 内部相应的行为环节上。

由于 Agent 自我分离的需要, EGO 内部不再是一个统一的整体,而是各个相对独立的 $ego_i \in E (i=1,2,\dots,n)$, E 表示所有的可能自我集。

定义 1 EGO 表示 Agent 自我的所在,是抽象的,可以理解为各个具体 ego 的集合。 ego 是 Agent 可以独立活动的主体,允许有多种类型与层次。

定义 2 ego 创建是指 Agent 因某种行为或活动需要而建立有关 ego 的过程。 ego 退化是指创建的 ego 因为活动少或活动受到抑制而导致功能衰退的过程。

以 C 表示各个 Agent 自我是否被创建的属性,取值 1 或 0 分别表示创建或未创建。有

$$V_L(EGO)/C = (V_L(ego_1)/C, V_L(ego_2)/C, \dots, V_L(ego_n)/C) \quad (1)$$

显然 $V_L(EGO)/C$ 是取值 0 与 1 的序列。令

$$V_L(EGO) = (V_L(ego_{i_1}), V_L(ego_{i_2}), \dots, V_L(ego_{i_m})) \quad (2)$$

式中 i_1, i_2, \dots, i_m 在 1 与 n 之间取整数值且互不相等, $V_L(ego_{ij})/C = 1 (j=1, 2, \dots, m)$, 而 $V_L(ego_{ik})/C = 0 (k=1, 2, \dots, m)$ 。表明 Agent 的当前行为总与已经创建的 ego 有关, ego 只有被创建才能参与 Agent 行为。将 $ego_{i_1}, ego_{i_2}, \dots, ego_{i_m}$ 构成的集合记为 E_N , $E_N \subseteq E$ 。为了反映 Agent 具有学习和进化的能力, E_N 必须是可变的,一方面新的 ego 创建后加入会导致 E_N 的扩充;另一方面,已存在的 ego 在活动过程中其功能得到增强或减弱,与 Agent 的最大利益相一致。

定义 3 E_N 不断扩充的过程称为 Agent 自我的生长; ego 因为生存需要其活动得到增强或抑制的过程称为进化。

从生理学上可找到不同时期人类自我变化的实例,比如儿童自我恐惧及保护意识等比较强,到了成人阶段则逐步减弱;儿童几乎没有独立意识,到了成人阶段则会树立起较强的独立意识等。解剖学提供了一个自我分离的有力且有趣的实例。早期有医生为了减缓癫痫病人病症,采用外科手术将病人大脑胼胝体剥离。手术初期病人病情显示好转迹象,但不久病人却出现另外一种奇怪

病症, 其左手好象不受思想支配, 表现出独立的意志和行为。这一病例表明自我分离是有物理基础的, 即自我实际以独立的形式存在, 对外其自我表现出统一性, 所以难以观察到, 当医生通过手术把病人左右大脑联系割裂时, 这种自我分离的实质便显现出来。

因 *ego* 有类型之分, 一个 *ego* 的活动往往需要其他 *ego* 的配合与支持, 而为了得到其他 *ego* 的认同与支持, 该 *ego* 必须清楚地表达自己的需要或愿望及拟采取的行动等, 即具有表达的属性。

性质 1 *ego* 具有多种不同类型的属性, 如自我意识的属性(意识到自己的存在及需要、知道自己区别于其他个体等)、活动的属性(自身活动以及支配系统其他环节活动)和表达的属性(为了让其他 *ego* 了解自己的需要, *ego* 总是不断地表达自己)等。

性质 2 Agent 需要、愿望等的满足是通过各个 *ego* 需要、愿望等的满足来具体实现的。

ego 集体的需要、愿望代表和体现了 Agent 的需要和愿望。不同 *ego* 对 Agent 存在整体利益上的一致性, 这是统一的基础, 但这种一致性的程度有高有低。某些情况下, 一个 *ego* 因某种需要或愿望拟采取的行动可能对其他 *ego* 需要、愿望造成不利影响而出现冲突, 此时, 存在一个 *ego* 之间自我协商和争夺的过程(因为每个 *ego* 都希望自己能代表 Agent 的统一意志), 最终 Agent 通过自我协商来进行需要和愿望等取舍。不过, 一旦 Agent 作出了选择, 该决定便代表了 Agent 的统一自我。为了表明这一点, 在式(2)中增加一项变为

$$V_L(EGO) = (\overline{V_L}(ego), V_L(ego_{i1}), V_L(ego_{i2}), \dots, V_L(ego_{im})) \quad (3)$$

式中 $\overline{V_L}(ego)$ 为 Agent 统一自我的所在, 代表 Agent 的统一意志。正常情况下 Agent 行为总是建立在统一意志基础之上的。

以下将 *ego* 的需要、愿望等统称为要求。要求的最根本来源不在 *ego*, 而在 Agent 的各个系统环节。不同的系统环节需要不同的 *ego* (*ego* 组) 为其服务。

定义 4 系统各环节唤起 *ego* 为之服务的过程称为唤醒。

定义 5 利益交叉点是指不同 *ego* 对 Agent 所提的要求。

定义 6 意识状态下 *ego* 表达自己的要求, 根据其他 *ego* 意见及表决结果决定下一步行为的过程称为自我协商过程。

定义 6 说明了自我的协商是通过某 *ego* 表达自己的要求, 其他 *ego* 就此要求根据自身利益发表意见进行磋商来实现的。该磋商可以是单轮次的, 也可以是多轮次的, 因此其原则可以由性质 3 来表述。

性质 3 Agent 行为取舍总是朝着可以获得最大利益交叉点的方向进行。

作为局部活动的主体, *ego* 与外界有着多种联系。一方面输入来自要求信息投射区(是知觉信息投射区 A_p 的一部分)的信息, 并据此表达自己的要求; 另一方面又输出信息到 A_B (行为信息投射区) 以支配 Agent 行为环节进行活动; 另外还要接收来自其他 *ego* 的意识信息等。*ego* 构成了 Agent 行为的基础, 事实上, Agent 行为就是多个 *ego* 的协调活动。尤其是当 Agent 有多种行为同时存在时, 每种行为分别由一组 *ego* 协调活动组成。

综上所述, 从内部看 Agent 自我是分离的, 存在为各个 $\overline{ego}_i \in E_N$, 而对外 Agent 自我又是统一的, 各 \overline{ego}_i 的要求必须成为 *ego* 意志的一部分才能决定 Agent 行为。

2 基于表达的 Agent 自我协商模式

智能 Agent 自我的行为应该是复杂的, 一方面可以根据其意识属性对其行为方向作出一定的预测; 另一方面, Agent 的具体行为过程又具有某种不确定性。由于 Agent 行为变化与其意识变化有着密切的关系, 本文对意识状态及无意识状态作如下定义:

定义 7 Agent 的意识状态, 是指 Agent 对自我及其行为自知(即有意识)的状态。Agent 的无意识状态是在 Agent 的意识状态间隙存在的一些非意识状态。

Agent 的某类行为总与一种特定的意识状态有关, 并且是有一定目的的。Agent 的意识状态经常变化, 导致 Agent 行为的变化。定义无意识状态是为了说明在意识状态的切换过程中存在意识不可跟踪的过程。

定义 8 Agent 的完整意识过程是指一个连贯的达成目的的意识状态变化过程(除了无意识间隙之外, 没有其他意识过程插入)。

在 Agent 进入某种意识状态后, 会力争保持该状态(行为), 称为专注力。外界信号对 Agent 行为有一定吸引作用, 外界信号引起 Agent 关注的力量, 称为吸引力。专注力和吸引力定义如下:

定义 9 专注力 g 表示 Agent 维护当前意识状态的能力。其值越大, 当前意识状态越不易改变, 反之则易改变。

定义 10 吸引力 h 是指外界信号吸引 Agent 关注的力量(与强度、信号类型和变化率等有关), 当 $h > g$ 时, Agent 会暂停当前意识活动而关注外界信号; 当 $h < g$ 时, 当前意识活动不会受外界信号的影响。

专注力 g 在一个完整意识过程中并不是一成不变的, 会随着预期、兴趣等发生变化而变大或变小。外界信号吸引力 h 的大小也不是一成不变的, 导致 Agent 行为易于改变, 如图1所示, 这种改变可能是因环境变化引起, 也可能因它对当前行为失去了兴趣。

Agent 自我协商的实现过程如图1所示。图中略去了其他信息连接的细节^[8]。 AR_L 表示协商环节, SC_L 表示思维状态转换环节, A_p 表示知觉信息投射区, 各种要求信息输入要求投射区, 各 ego 表达的信息送入统一自我。 T_b 时刻起, 作用于 Agent 的外界环境信息为 $I(t)=(I_1(t), I_2(t), \dots, I_n(t))$, 其中信号 $I_d(t)$ 会引起 ego 的分歧。假设 T_b 起, Agent 处在一个 $I_d(t)$ 引起的自我协商意识过程 A_p , Agent 对其专注力为 $g(t)$, 为时间的函数。如果正常进行, 该过程将到时刻 T_e 结束。

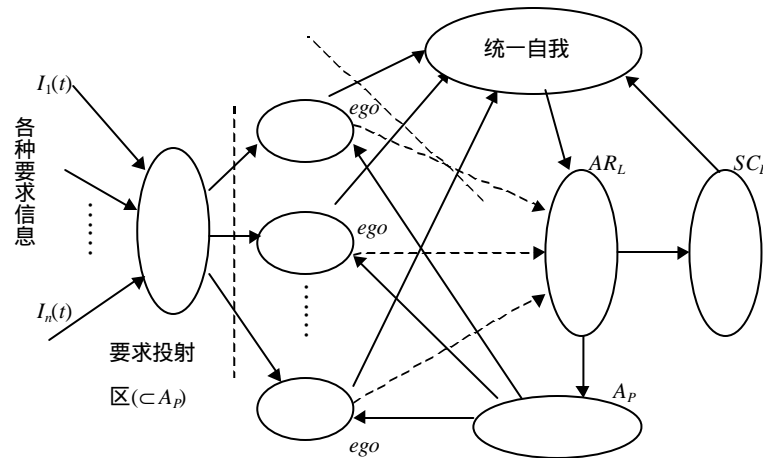


图1 基于表达的自我协商过程示意图

来自要求投射区的信息要经过选择才会到达 ego 。被连接的 ego 唤醒后便进行自我表达。这种表达是从该 ego 角度出发的, 而意识表达又要求建立在统一自我立场上。所以, 该 ego 表达的信息需由统一自我将有关信息进行转换后表达, 本文分别定义为如下的表达连接和表达替换:

定义 11 来自要求投射区的信息要经过选择才会到达 ego , 这一过程称为表达连接; 表达替换是指把局部信息转换为站在 Agent 统一自我立场上表达的普遍可以理解的信息的过程。

结合前面的定义, 可对协商环节工作作如下结论:

结论 1 如果 $\max(h|I_i(t)) < g(t)(T_b = t < T_e, i \neq d)$, 则 Agent 自我协商的过程会持续进行而不被打断。其中, $h|I_i(t)$ 表示外界信号 $I_i(t)$ 的吸引力。

结论 2 如果到达某时刻 t 有 $\max(\mathbf{h}I_i(t)) > g(t)$ (t 是 T_b 和 T_e 之间的某一时刻), 则 Agent 自我协商的过程进行到时刻 t 将被打断。

针对图1的情况, 具体的自我协商过程如下:

1) 前提: Agent 其他意识活动专注程度降低; Agent 关注各外界信息 $I_i(t)$; 由于 $I_d(t)$ 具有最大吸引力, 可实现表达连接; $I_d(t)$ 引起 ego 之间的分歧, 导致 Agent 进入自我协商的意识状态。

2) 在自我协商状态, 有关 ego 表达自己的要求, 此信息经过表达替换后, 以统一自我的角度表达该要求, 由此产生的信息会经一定渠道告知其他 ego 。

3) 其他 ego 根据该要求对自己有利或不利的情况进行表决(表决在 AR_L 中进行)。如果表决产生结果, 则将表决结果送至 SC_L , 从而实现相应意识状态的切换:

(1) 如果通过, Agent 切换到实现该要求的意识状态;

(2) 如果表决暂未通过, 但也未放弃, 会导致新一轮表决过程, 直致产生表决结果(如果不被其他意识过程打断);

(3) 如果表决未通过, 且决定放弃, 则 Agent 不进入实现该要求的意识状态;

(4) 根据前面的假设可知, 到达 T_e 时刻时, Agent 结束协商状态, 但具体是否按信号 $I_d(t)$ 的要求实施行为取决于自我协商的结果。

3 结束语

在讨论智能 Agent 自我多重性特征的基础上, 本文提出了自我的统一与分离是智能 Agent 内在本质特征之一的观点。并给出了一种基于表达的 Agent 自我协商模式, 可以看出在 Agent 行为决策过程中, Agent 自我具有分离的特点, 而最终作出的决策又是建立在一个统一自我的基础之上, 这对了解 Agent 内部结构及其运行机制具有十分重要的意义。

参 考 文 献

- 1 Mizoguchi F, Nishiyama H, Ohwada H, *et al.* Smart office robot collaborating based on multi-agent programming. *Artificial Intelligence*, 1999, (114): 57~94
- 2 马 争, 魏险峰. 智能网——网络智能化的关键技术. *电子科技大学学报*, 2000, 29(1):1~4
- 3 Haddadi A. Reasoning about cooperation in agent systems: a pragmatic theory. [PhD Thesis]. Manchester: University of Manchester Institute of Science and Technology (UMIST), 1995
- 4 Wooldridge M J, Jennings N R. Intelligent agents: theory and practice. *Knowledge Engineering review*. 1995, 10(2): 115~152
- 5 Rao A S, Georgeff M P. BDI agents: from theory to practice. In: Georgeff M P ed. *Proceedings of the 1st International Conference on Multi-Agent System (ICMAS-95)*, San Francisco: ACM Press, 1995: 312~319
- 6 Kiss G. Variable coupling of agents to their environment: combining situated and symbolic automata. In: Werner E, Demazeau Y, eds. *Decentralized Artificial Intelligence 3*, Elsevier/Holland, Amsterdam, 1992: 231~248
- 7 Kiss G, Reichgelt H. Towards a semantics of desires. In: Werner E, Demazeau Y eds. *Decentralized AI 3, Proc of the third European Workshop on Modeling Autonomous Agents in a Multi-Agent World*, Kaiserslautern, Germany, Amsterdam: Elsevier Science Publishers, 1992: 115~127

(下转第86页)

An Examination Base and Its Realization Using DELPHI Tool

Yang Hong

(Dept.of Electronic Engineering, UEST of China Chengdu 610054)

Abstract This paper introduces an examination base, in which DELPHI method is applied to realize its design structure. The examination base is developed by the use of visual tool and can be used in wide range of applications.

Key words examination; in front of user; visual; DELPHI tool

(上接第81页)

Unification and Separation of Agent's *EGO*

Li Jinhou

(Department of Computer, East China Univ. of Metallurgy Anhui Maanshan 243002)

Abstract Though intelligent Agent can be considered as an entity from outside, it shows the property of separation because of its need inside. The conclusion is put forward based upon the thought that *EGO* is used to indicate the center of intelligent Agent's spirit. The multiple characteristics of Agent's need and desire are revealed in this paper, and the qualities and characteristics of Agent's separation are analyzed and discussed further. The knowledge about Agent's structure lays a foundation to analyze its behavior according to its structure.

Key words intelligent agent; unification and separation; express connection; express replace; intention state

2000年第6期学报谏误表

页码	行数	误	正
675	倒9行	1994年“微机编排的技术问题与分析”获四川省好稿评选三等奖,省科委、宣传部、新闻出版局	1995年“四川省科技期刊审读分析”获国家教委优秀论文一等奖
677	5行	(1995年~1988年···)	(1959年~1988年···)
679	右栏10行	1999 28(4) ⁸⁾	1999 28(4) ⁹⁾
679	倒8行	1988 卅周年校庆论文集(上)	1986 卅周年校庆论文集(上)
679	倒7行	1988 卅周年校庆论文集(下)	1986 卅周年校庆论文集(下)

· 本刊编辑部 ·