

# 多智能体系统联合行动的形式化研究及其应用\*

刘金琨 王树青

(浙江大学工业控制技术国家重点实验室 杭州 310027)

(中国科学院自动化研究所模式识别国家重点实验室)

**摘要** 针对多智能体协作的需要,从智能体的群组思想、社会规划、群组意图及联合意图几方面,探讨了多智能体联合行动的形式描述,并对每种形式做了具体分析,提出利用多智能体的联合行动实现复杂工业过程联合监控的思想。最后阐述了该项研究进一步的发展方向。

**关键词** 多智能体,联合行动,联合意图,联合监控

**分类号** TP 18

## Fomalism Research on the Joint Action for Multiagent System and Its Application

Liu J inkun, Wang Shuqing

(Zhejiang University)

**Abstract** According to the needs of multiagent corporation, the formalism for multiagent joint action is discussed from the points of group mind, social plan, group intention and joint intention. Moreover, concrete analysis is given for each formalism. The multiagent joint supervision control is put forward by adopting multiagent joint action technology. At last, further research directions are pointed out.

**Key words** multiagent, joint action, joint intention, joint supervision

## 1 引言

对于一个多智能体系统,每个智能体都有共同的目标<sup>[1]</sup>,智能体为了达到该目标,就需要有一个联合行动。这是由于在多智能体的联合问题求解中,单个智能体没有能力及信息完成整个任务,必须通过多智能体协作,产生联合行动去求解。为此,每个智能体都应有各自的意图<sup>[2]</sup>,根据该意图,智能体相互协作,并采取相应的行动。

为探讨智能体联合行动的形式,本文从多智能体的群组思想、社会规划、群组意图及联合意图的角度加以分析。

## 2 多智能体系统的联合行动描述

### 2.1 群组思想

群组思想是为智能体进行问题求解而提供的一

种有益的指导思想,它对智能体的行动有着重要影响<sup>[3]</sup>。群组思想来源于智能体,但不是任何一个单个智能体的一部分。群组思想应用于多智能体的协调活动中,多智能体的联合行动需要群组思想来指导。

### 2.2 社会规划

为了描述多智能体系统的联合行动,应对其进行社会规划<sup>[4]</sup>。规划由规划符号和智能体成对表达,规划符号是一种抽象结构,当被智能体执行时,它就会产生行动。复杂的规划由算子和动态逻辑相结合来表达。

**例 1** 有三个智能体  $x_1, x_2, x_3$ , 分别要进行  $e_1, e_2, e_3$  三项行动,则可规划为

$(e_1, x_1); (e_1, x_2); (e_1, x_3) \quad (e_2, x_1); (e_2, x_2); (e_2, x_3)$   
 $(e_3, x_1); (e_3, x_2); (e_3, x_3)$

**例 2** 三个智能体  $x_1, x_2, x_3$ , 分别完成同一任务  $e$ , 则可由一个行动序列规划为

\* 国家自然科学基金项目(69874037)和浙江大学工业控制技术国家重点实验室开放课题基金项目(K97M 05)

1998 - 07 - 17 收稿, 1998 - 09 - 28 修回

$(e x_1); (e x_2); (e x_3)$

### 2.3 群组意图

群组意图是指一个智能体组为达到目标所采取的那些策略<sup>[5]</sup>,即一个智能体组欲达到目标  $p$ , 如果它遵从能使  $p$  成“真”的必要的手段, 则该组能达到  $p$ 。

群组意图可用于基于多智能体的分布式问题求解系统, 该系统由问题、智能体及智能体组织构成。每个智能体都有自己内部的表示形式, 并且智能体间相互作用, 共同协作解决问题。文献[6]以 4 个智能体 ( $B$ ) 搜索一个智能体  $M$  为例, 给出群组意图的以下几种控制形式。其中, 集中式控制中  $B_1$  作为集中式控制器, 其余智能体为  $B_k (1 < k \leq 4)$ 。

1) 全能的集中式控制器: 智能体  $B_1$  作为全能控制器, 拥有对整个系统状态的描述, 负责做出最优决策, 收集各智能体位置及环境信息, 并对其它三个智能体  $B_k$  发布命令。

2) 智能体感知自己位置的集中式控制器: 各智能体能感知各自的位置, 并且  $B_k$  能将各自的位置信息传递给  $B_1$ 。

3) 搜索智能体  $M$  的集中式控制器: 各智能体均能检测到智能体  $M$  的位置,  $B_1$  能够估算  $B_k$  的行动范围, 一旦智能体  $B$  探测到智能体  $M$ , 它将不会失去。

4) 抽象命令的集中式控制器: 各智能体能做出一些局部决策, 并能执行一些抽象命令, 使智能体从事更复杂的行为。

5) 在利它的同等级位的智能体中进行分布式控制: 各智能体是平等的, 利它的, 均能做出全局决策, 采用总体最优的方法确定每个智能体的行动, 并以广播方式传递信息。

6) 在自我利益的智能体中进行分布式控制: 智能体存在自我利益的考虑, 能局部估算各自行动的费用, 并在总体寻优。联合行动时, 每个智能体将各自的估计值传递给其它智能体。

对群组意图进行描述, 需要从下述几方面来进行: 1) 智能体的外界环境; 2) 外界环境的转变; 3) 外界环境的约束; 4) 智能体的行动能力; 5) 智能体的推理能力; 6) 智能体的感知能力; 7) 智能体的通讯能力。

### 2.4 联合意图

一些智能体有了群组意图, 只是有了联合行动的基础, 尚不能构成联合行动。智能体之间应为达到一个共同行动目标而建立一定的联系, 即建立对联

合行动的承诺机制。

联合意图是在确定的共同精神状态下完成一个集体行动的联合承诺<sup>[7]</sup>, 是在群组意图之上的更高级的意图, 即基于承诺机制的群组意图, 因此它具备群组意图的功能。联合意图建立在联合恒定目标的基础上, 而联合恒定目标的形式化描述又是基于可达目标的定义。因此, 智能体联合意图的描述涉及到联合恒定目标和可达目标<sup>[3]</sup>。

#### 2.4.1 可达目标

可达目标 (attainable goal) 定义了在一一定的动机下, 为共同完成一个目标的智能体组中的智能体状态。假如存在下述条件之一, 则智能体  $x$  以  $p$  为可达目标。

1)  $x_1$  不相信  $p$  为“真”, 并将最终为“真”的  $p$  作为一个目标;

2)  $x_1$  相信  $p$  为“真”, 并让所有智能体都相互相信这一事实;

3)  $x_1$  相信  $p$  永远不能为“真”, 并让所有智能体都相互相信这一事实。

以两个智能体  $x_1$  和  $x_2$  为达到目标  $p$  为例, 可达目标可描述为

$$\begin{aligned} (AG x_1 x_2 p) = & \\ [\neg BEL(x_1 p) \quad GOAL(x_1 p)] & \\ [BEL(x_1 p) \quad GOAL(x_1 MB(x_1 x_2 p))] & \\ [BEL(x_1 \neg p) & \\ GOAL(x_1 MB(x_1 x_2 \neg p))] & \end{aligned}$$

其中 MB (mutual belief) 为互信念, 它是关于其它智能体信念的无限次合取<sup>[8]</sup>。例如,  $x_1$  和  $x_2$  对命题  $p$  有一个互信念, 则要求:  $x_1$  和  $x_2$  相信  $p$ ,  $x_1$  相信  $x_2$  相信,  $x_2$  相信  $x_1$  相信  $x_2$  相信, ...。即定义互信念为

$$MB(x_1, x_2, p) = BEL(x_1, p \quad MB(x_2, x_1, p))$$

#### 2.4.2 联合恒定目标

联合恒定目标 (joint persistent goal) 意味着参与协作的智能体, 如果有了联合恒定目标将不会放弃, 直到已达到或它们全部 (或部分) 认为目标永不可达。联合恒定目标是建立在可达目标基础上的。设一组智能体 ( $x_1$  和  $x_2$ ), 为达到联合恒定目标  $p$ , 当且仅当: 1) 它们都相信当前  $p$  是“假”的; 2) 它们都想让  $p$  最终成为“真”; 3) 它们都相信每一个智能体都是以  $p$  作为可达目标, 直到相信  $p$  为“真”或  $p$  永远不能为“真”为止。上述描述用公式可表达为

$$\begin{aligned} (JPG x_1 x_2 p) = & \\ MB(x_1 x_2 \neg p) \quad MB(x_1 x_2 GOAL(x_1 p)) & \end{aligned}$$

GOAL ( $x_2 \ p$ ) (BEFORE (MB ( $x_1 \ x_2 \ p$ ))  
 MB ( $x_1 \ x_2 \ \neg p$ )) $\Rightarrow$ MB ( $x_1 \ x_2$ (AG  $x_1 \ x_2 \ p$ ))  
 (AG  $x_2 \ x_1 \ p$ )))

### 2.4.3 智能体的联合意图

智能体的联合意图(joint intention)即多个智能体为完成行动  $a$  而对集体行动做出的承诺。联合意图中应用了联合恒定目标,标明了智能体必须跟踪目标的有效性,并为智能体提供了完成任务所需的条件。智能体的联合意图可描述为

$$\begin{aligned} (JI \ x_1 \ x_2 \ a) = & \\ (JPG \ x_1 \ x_2 \ (DONE \ x_1 \ x_2 & \\ [BEFORE \ (DONE \ x_1 \ x_2 \ a) & \\ \Rightarrow (MB \ x_1 \ x_2 \ (DONE \ x_1 \ x_2 \ a)) & ]?; a)) \end{aligned}$$

## 3 基于多智能体联合行动的复杂工业过程联合监控

### 3.1 多智能体联合监控原理

研究表明,通过采用多智能体的联合行动可实现分布式问题求解<sup>[9]</sup>。对于一个多智能体系统,为了实现联合监控任务,需要进行联合行动。联合意图作为多智能体系统对行动目标的一种共同承诺的描述,一旦所有智能体都同意一个共同目标,且每个智能体承诺去达到目标,则智能体就有了联合意图。多智能体通过联合意图来实现联合行动,从而实现对复杂工业过程的联合监控。

### 3.2 多智能体联合行动的建立

多智能体在联合行动前应满足下列条件<sup>[10]</sup>: 1) 有一个共同的目标和意图; 2) 有愿为达到该目标而进行协作的智能体; 3) 承认一个共同协定; 4) 制定实现共同目标的协定。

#### 3.2.1 多智能体联合行动的规则

R1: IF 一个智能体局部诊断任务完成 AND 该任务是联合行动所要求的目标 THEN 联合行动的目标已满足。

R2: IF 联合行动目标已达到 THEN 将该信息传递给多智能体的协作模块 AND 将所有智能体的局部活动挂起。

R3: IF 某一智能体的个体行动受到阻碍 THEN 将该信息通知各智能体 AND 对联合行动采取适当的修正。

R4: IF 网络信息 agent 不再采集信息 THEN 撤消承诺 AND 停止联合行动。

#### 3.2.2 多智能体联合行动的建立

多智能体联合行动的建立分为5个步骤,如图1所示。

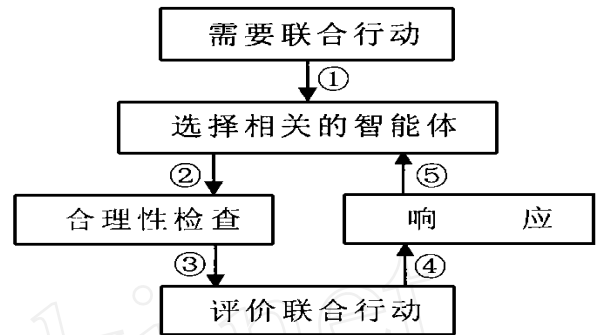


图1 建立联合行动

### 3.2.3 多智能体联合行动的实现

多智能体系统联合行动主要分为三个过程,即匹配、选择和执行。具体实现措施如图2所示。三个过程描述如下:

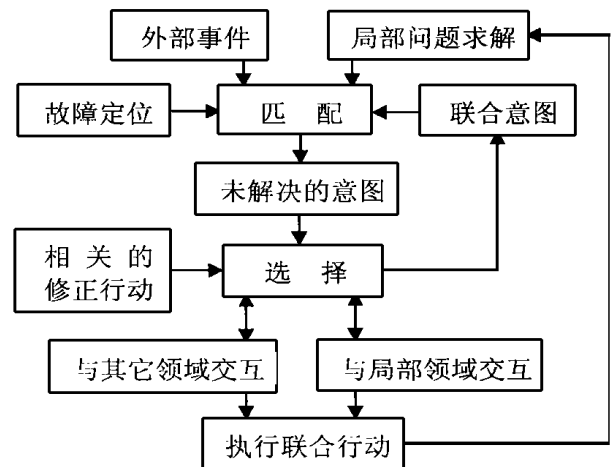


图2 多智能体系统联合监控的实现

#### 1) 匹配过程

R1: IF 任务  $t_1$  延迟 AND  $t_1$  是联合行动方案  $R$  的一部分 AND 任务  $t_1$  需与  $t_2$  同步 THEN 方案  $R$  冲突。

R2: IF 执行方案  $R$  完成 AND  $R$  所期望的结论未产生 AND 替代方案存在 THEN  $R$  无效。

#### 2) 选择过程

R1: IF 联合行动目标达到 THEN 放弃所有相关的局部活动 AND 通知协作模块联合行动已完成。

R2: IF 方案  $R_1$  无效 AND 可替代方案  $R_2$  存在 THEN 放弃所有与  $R_1$  相关的局部活动 AND 通知协作模块  $R_1$  无效 AND 将  $R_2$  作为联合行动的方案。

#### 3) 执行联合行动

R1: IF 联合行动方案 R1 有冲突 AND 可替代方案 R2 存在 THEN 执行联合行动方案 R2。

### 3.3 多智能体联合监控工作过程

多智能体有了联合意图、共同目标和联合行动,便可进行联合问题求解,从而实现联合监控。其联合监控机理如图 3 所示。多智能体联合监控的工作过程描述如下:

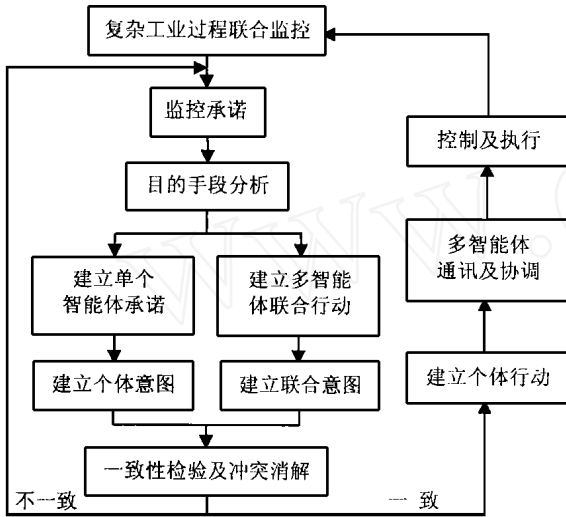


图 3 多智能体联合监控工作过程

1) 联合监控过程完成下述任务: 建立新的多智能体联合监控目标, 完成多智能体联合监控任务, 放弃或修改智能体承诺。

2) 目的手段分析: 确定新的目标应如何实现及具体实现方法, 即目标是局部完成还是联合完成以及各种目标或方法的选择。其输出有三个: 不开始新的活动, 决定局部执行, 决定以联合行动方式执行。

3) 如果经过目—手段的分析与推理, 认为新的目标需要联合实现, 则建立联合行动。这意味这一组智能体同意一起协作, 通过联合意图达到共同的目标。

4) 联合意图是对多个智能体联合行动进行问题求解的一种描述, 不能直接执行, 只有智能体本身才有控制能力, 因而个体意图是和行动联系在一起的。

5) 联合意图和个体意图间有关联, 在联合行动中, 智能体作为参与者至少执行了一个个体意图。联合意图和个体意图之间应保持一致性, 以保证局部和整体目标之间无冲突。

6) 一致性检验: 如果所建立的意图与现有的意图不一致, 则需要重新回到监控过程中, 进一步做目—手段分析, 对现有智能体意图进行修改或建立新的目标。

7) 联合意图与个体意图相结合, 变为智能体的个体行动, 完成两个任务: 通过黑板结构实现与其它智能体的信息共享与协调; 完成对任务的控制与执行。

8) 监控过程接受来自控制及执行过程的反馈, 以决定是否修改或放弃智能体的承诺, 重新检验联合意图和联合行动。

## 4 进一步的研究工作

多智能体的联合行动比较复杂, 应用于实际工业中还有许多问题有待于探索。进一步的研究工作主要表现在以下几方面:

1) 在一个复杂的动态环境中, 多智能体相互协调与合作受环境资源的限制<sup>[11]</sup>。因此, 要保持智能体组联合行动的动态协调性, 联合行动的描述不仅要考虑达到目标的状态, 而且应对多智能体的行动进行合理规划。

2) 时间是限制多智能体相互协作进行资源分配的关键<sup>[12]</sup>, 尤其在实时环境下, 多智能体联合行动需要在一定的时间下完成, 这就有必要研究在时间限制下的多智能体联合行动。

3) 在复杂的动态环境中, 智能体必须适应外界环境的变化<sup>[13]</sup>。要保持智能体组联合行动的动态协调性, 应考虑联合行动中智能体的鲁棒性、自适应性和自学习能力。

4) 分布式系统中通常存在不确定性, 尤其是当各智能体分别采用不同的推理模型时, 它们之间的不确定性推理就显得更加重要<sup>[14]</sup>。因此, 在多智能体的联合行动中, 应探讨智能体间的不确定性推理问题。

5) 各智能体对参加联合行动并非都是自愿的, 在智能体的联合行动中应体现出智能体的意愿性, 并应考虑联合行动失败的情形。

## 5 结 语

多智能体的联合行动需要形式化的描述。本文提出了多智能体形式化描述的几种方法, 并结合实例探讨了群组意图, 建立了多智能体联合意图模型和联合监控机制, 阐述了多智能体联合行动进一步的研究工作。本研究为复杂工业过程分布式智能控制提供了新的途径。目前, 分布式问题求解方法尚处于理论研究阶段, 还有许多实际问题有待于进一步研究。

## 参 考 文 献

- 1 D M Lane, A G M cfadzean Distributed problem solving and real-time mechanisms in robot architectures Engineering Application in Artificial Intelligence, 1994, 7 (2): 105\_ 117
- 2 P R Cohen, H T Levesque Intension is choice with commitment Artificial Intelligence, 1990, 42(2): 213 \_ 261
- 3 H J Levesque, P R Cohen, J H T Nunes On acting together In: Proc of 8th National Conference on AI (AAAI-90). Boston, 1990 94- 99
- 4 A S Rao, M P Georgeff, E A Sonenberg Social plans: A preliminary report In: Proc of Modelling an Autonomous Agent in a Multi-agent World Washington, 1991 40\_ 46
- 5 E H Durfee, T A Montgomery. MICE: A flexible testbed for intelligent coordination experiments In: Proc of 9th Workshop on Distributed Artificial Intelligence Washington, 1989 25\_ 40
- 6 M P Singh, M N Huhns, L M Stephens Declarative representations of multiagent systems IEEE Trans on Knowledge and Data Engineering, 1993, 5(5): 721\_ 739
- 7 B Chaib Ddraa, P M ilot A framework for cooperative work: An approach based on the intentionality. Artificial Intelligence in Engineering, 1990, 5(4): 199\_ 205
- 8 J Barwise Three views of common knowledge In: Proc of 2nd Conference on Reasoning about Knowledge Pacific Grove, 1988 60\_ 65
- 9 N R Jennings Controlling cooperative problem solving in industrial multiagent systems using joint intentions Artificial Intelligence, 1995, 75(2): 195\_ 240
- 10 B Chaib Ddraa, P M ilot A framework for cooperative work: An approach based on the intentionality. Artificial Intelligence in Engineering, 1990, 5(4): 199\_ 205
- 11 N V Findler Multiagent coordination and cooperation in a distributed dynamic environment with limited resources Artificial Intelligence in Engineering, 1995, 9: 229\_ 238
- 12 N V Findler, U K Sengupta Multiagent collaboration in time-constrained domains Artificial Intelligence in Engineering, 1994, 9: 39\_ 52
- 13 J Doulamis, M A Gray. Multiagent reason maintenance and group adaptability. IEEE Trans on Systems, Man and Cybernetics, 1996, 26(6): 850\_ 856
- 14 C Q Zhang Cooperation under uncertainty in distributed expert systems Artificial Intelligence, 1992, 56 (1): 21\_ 69

## 作 者 简 介

刘金琨 男, 1965年生。1997年获东北大学博士学位, 现在浙江大学工业控制技术研究所从事博士后研究。研究方向为智能控制理论及应用。

王树青 男, 1940年生。浙江大学工业控制技术国家重点实验室主任, 教授, 博士生导师。研究方向为自动控制理论及应用, 工业生产过程模型化与优化控制等。