

意图逻辑^{*}陈小平 刘贵全[✓] 蔡庆生 刘 洁

(中国科学技术大学计算机系 合肥 230027)

A

摘 要 The formalized study of intention is becoming an important issue in artificial intelligence, computational linguistics and philosophical logic in recent years. In this paper, we analyze some well-known intention logics and point out their drawbacks. We propose that we should focus on the dynamic character of intentions in the further research. A set of primitive dynamic constraint is given.

关键词 Intention logic, Limited self-government system, Dynamic constraint

近年来,关于意图的形式化研究已成为哲学逻辑、人工智能和计算语言学共同关心的重要课题,在实用推理^[1],认知建模^[2],规划与行动推理^[3]和自然语言处理^[4]等领域有着广泛的应用。意图的基本内涵是对未来行动的“合理”选择,正是通过这种选择,一个有限自治系统才能维持信念、目标、规划和行动的合理平衡。

一个人是一个有限自治系统。其“有限性”表现在人的知识、推理能力和计算资源是不完备和不充足的,其“自治性”表现在人能够在一定条件下依据自己的知识、信念和其他因素对自己的行动进行设计、选择和实施。尽管对于一个计算机系统能否获得完全自治/自主性存在争论,但人工智能的实践表明,建造有自治的智能系统不仅是必要的,而且是可能的。事实上,一个处理复杂任务的智能系统必须是一个有限自治系统,即该系统必须具有在一定范围内合理选择未来行动的能力。目前,这种选择/规划的机制主要通过过程性方式表达,对其性质(性能)的研究主要采用实验方法。这种状况已不能适应人工智能的发展和应用的需要。

对意图概念的哲学研究已有很长的历史,这些研究往往试图将意图归结为信念和愿望。但 Bratman^[5]指出,意图不能归结为信念和愿望,一是资源有限性决定了建立在对信念、愿望的“充分权衡”之上的“正确”决策实际上是办不到的,只能做到“合理”决策;二是行动必须加以协调,而协调又必须依赖于已做出的决策。因此,将意图做为一个独立心灵

状态是必要的。

一、有限自治系统的合理性

意图的形式化研究首先必须澄清意图与信念、目标和行动之间的基本约束关系。这些约束构成有限自治系统的合理性必要条件。Bratman 提出“反对称论题”和“无副作用原理”做为合理性必要条件^[1]。Bratman 认为,允许一个有限自治系统以某一行动为其意图,同时又相信该行动无法实现是不合理的,即不允许意图-信念不相容性(BI-ICN)。另一方面,一个有限自治系统采纳一个行动为意图,但并不相信它能实现该行动是合理的,即应允许意图-概念的不完全性(BI-ICM)。这种不相容性与不完全性的不对称性称为反对称论题,形式表述如下:

$$(BI-ICN) \forall M: M \models INTEND(\alpha) \wedge BEL(\neg\alpha)$$

$$(BI-ICM) \exists M: M \models INTEND(\alpha) \wedge \neg BEL(\alpha)$$

无副作用原理是说,以 A 为意图的有限自治系统,不应被迫以 A 的“后承”B 为意图,而不论它相信行动 A 与行动 B 之间的关联多么强。例如,一个意欲拔牙的正常患者,绝不会意欲享受拔牙时的疼痛,尽管他十分清楚拔牙引起的疼痛是不可避免的。无副作用原理记为,BI-SE,其形式表述在下节给出。

A. A. Rao 和 M. P. Georgeff 建议将反对称论题推广到意图与目标,以及目标与信念之间^[6],可表述如下:

$$(GI-ICN) \forall M: M \models INTEND(\alpha) \wedge GOAL(\neg\alpha)$$

$$(GI-ICM) \exists M: M \models INTEND(\alpha) \wedge \neg GOAL(\alpha)$$

^{*} 本文的研究得到了国家自然科学基金、863 计划和国家教委博士点基金的资助。

(BG-ICN) $\forall M: M \models \text{GOAL}(a) \wedge \text{BEL}(\neg a)$

(BG-ICM) $\exists M: M \models \text{GOAL}(a) \wedge \neg \text{BEL}(a)$

此外,他们还提出了与副作用相关的非迁移原理做为合理性约束,在目标与信念之间的非迁移性意指:无论一个有限自治系统对基本命题 a 的信念多么强,它不应采纳 a 为其目标。非迁移原理的形式描述如下:

(BI-NT) $\exists M: M \models \text{BEL}(a) \wedge \text{INTEND}(a)$

(GI-NT) $\exists M: M \models \text{GOAL}(a) \wedge \neg \text{INTEND}(a)$

(BG-NT) $\exists M: M \models \text{BEL}(a) \wedge \neg \text{GOAL}(a)$

以上诸约束条件包括了当前意图形式化研究的主要考虑,构成意图逻辑系统构造的主要目标,据此,Rao 和 Georgeff 提出下述合理性必要条件论题^[4]:

一个有限自治系统是合理的(有理性的),仅当它至少满足下述全部条件:a)反对称条件论题 BI-ICN, BI-ICM, GI-ICN, GI-ICM, BG-ICN 和 BG-ICM 成立;b)非迁移原理 BI-NT, GI-NT 和 BG-NT 成立;c)无副作用原理 BI-SE 的全部条款成立。

此论题包含许多含糊与欠缺,我们将在以下几节的讨论中加以澄清和弥补。

二、意图与承诺

P. R. Cohen 和 H. J. Levesque 率先开展了意图的形式化研究,他们构造了一个有四个基本模态词 BEL, GOAL, HAPPENS 和 DONE 的模态逻辑系统(以下记为 CL 系统)^[5],其中意图被定义为导出概念。CL 系统的形式语言如下:

〈行动变元〉 ::= $a, b, a_1, b_1, a_2, b_2, \dots, c, c_1, c_2, \dots$

〈规则变元〉 ::= $i, j, \dots$

〈变元〉 ::= 〈行动变元〉 | 〈规则变元〉

〈谓词〉 ::= $\langle \langle \text{谓词符号} \rangle \langle \text{变元} \rangle_1, \dots, \langle \text{变元} \rangle_n \rangle$

〈公式〉 ::= $\langle \text{谓词} \rangle | \neg \langle \text{公式} \rangle | \langle \text{公式} \rangle \vee \langle \text{公式} \rangle$

且 〈变元〉 〈公式〉 | 下列之一:

〈变元〉 = 〈变元〉₁

HAPPENS 〈行动表达式〉

DONE 〈行动表达式〉

BEL 〈公式〉

GOAL 〈公式〉

〈时间表达式〉

〈行动变元〉 $\leq$ 〈行动变元〉

〈时间命题〉 ::= 〈数字〉

〈行动表达式〉 ::= 〈行动变元〉 | 下列之一:

〈行动表达式〉; 〈行动表达式〉

〈行动表达式〉 | 〈行动表达式〉

〈公式〉?

〈行动表达式〉*

一个模型是一个结构 $\langle \Theta, E, T, B, G, \Phi \rangle$, 其中 Θ 为一集合, E 是基本事件类型的集合, $T \subseteq [Z \rightarrow E]$ (其中 Z 为整数集), $B \subseteq T \times Z \times T$ (信念可达关系), $R \subseteq T \times Z \times T$ (目标可达关系), $\Phi \subseteq [\langle \text{谓词} \rangle_k \times T \times Z \times D_k]$ (谓词的解释, 其中 $D = \Theta \vee E^*$)。

对任何模型 M , 事件序列 σ , 整数 n 和变元赋值, 其公式的真值条件如下。

(1) $M, \sigma, v, n \models P(x_1, \dots, x_k)$ iff $\langle v(x_1), \dots, v(x_k) \rangle \in \Phi(P, \sigma, n)$;

(2) $M, \sigma, v, n \models \neg a$ iff $M, \sigma, v, n \not\models a$;

(3) $M, \sigma, v, n \models (a \vee b)$ iff $M, \sigma, v, n \models a$ 或 $M, \sigma, v, n \models b$;

(4) $M, \sigma, v, n \models \exists x a$ iff 对某 $d \in D$ $M, \sigma, v(d/x), n \models a$;

(5) $M, \sigma, v, n \models (x_1 = x_k)$ iff $v(x_1) = v(x_k)$;

(6) $M, \sigma, v, n \models \langle \text{时间表达式} \rangle$ iff $v(\langle \text{时间表达式} \rangle) = n$;

(7) $M, \sigma, v, n \models (c_1 \leq c_2)$ iff $v(c_1)$ 是 $v(c_2)$ 的初段;

(8) $M, \sigma, v, n \models (\text{HAPPENS } a)$ iff $\exists n \leq m: M, \sigma, v, n \models [a]_m$;

(9) $M, \sigma, v, n \models (\text{DONE } A)$ iff $\exists m \leq n: M, \sigma, v, m \models [a]_n$, 其中 $n \models [a]_m$ 解释为行动序列 a 在时间 n 至 m 之间在 σ 中发生(详见文[5])。

(10) $M, \sigma, v, n \models (\text{BEL } a)$ iff $\forall \sigma', \langle \sigma, n \rangle B a', M, \sigma', v, n \models a$;

(11) $M, \sigma, v, n \models (\text{GOAL } a)$ iff $\forall \sigma', \langle \sigma, n \rangle G \sigma', M, \sigma', v, n \models a$;

为了保证模型的合理性, CL 系统引入两条约束:(1)相容性, B 是欧几里德的、传递的和序贯的, G 是序贯的;(2)现实性: $\forall \sigma, \sigma'$, 如果 $\langle \sigma, n \rangle G \sigma'$, 则 $\langle \sigma, n \rangle B \sigma'$ 。

在 CL 系统中, 可以定义一系列时序模态词:

$\Diamond a \equiv \exists x (\text{HAPPENS } x, a?)$

$\Box a \equiv \neg \Diamond \neg a$

$\text{LATER } p \equiv \neg p \wedge \Diamond p$

$\text{BEFORE } pq \equiv \neg \forall c (\text{HAPPENS } c, q?) \supset \exists a (a \leq c) \wedge (\text{HAPPENS } a, p?)$

以上构成 CL 系统的“原子层”。在这个层次上可以分析信念、目标和行动的关系。在原子层上可进一步构建用以研究意图概念的“分子层”。分子层的

基本概念是“持续目标”,定义为:

$$P\text{-GOAL}_p \equiv_{\text{def}} (GOAL(LATER_p)) \wedge (BEL \rightarrow p) \wedge [BEFORE((BEL \ p) \vee (BEL \ \Box \neg p)) \rightarrow (GOAL(LATER \ p))]$$

其含义是, p 是持续目标当且仅当, (a) p 在当前被相信为真; (b) p 最终将被接受为目标; (c) 在放弃 p 之前或者确信 p 已实现, 或者确信 p 永远无法实现。

注意持续目标表达了一种特殊的承诺, 即所谓“狂热主义”承诺。另外, (a)、(b) 表达了意图的“客观必要性”。

在 CL 系统中, 意图被定义为特殊的持续性目标。分为两类, 即关于行动的和关于状态的。对任何行动表达式 a , 行动的意图定义为:

$$INTEND_a \equiv_{\text{def}} P\text{-GOAL}[DONE(BEL(HAPPENS_a))]; a]$$

这就是说, 关于行动的意图是一种(狂热主义的)承诺, 但并非仅仅承诺做 a , 而是要求在行动之前相信自己将要去做 a , 这就在意图的客观必要性上又增加了“主观必要性”要求。因此可以说, 在 CL 系统中某行动成为意图当且仅当它是一种有客观必要性和可能性的有意识的行动。

关于状态的意图定义为:

$$INTEND_p \equiv_{\text{def}} P\text{-GOAL}(\exists e (DONE(\exists e' (HAPPENDSe', p?)) \wedge \neg(GOAL \rightarrow (HAPPENDe, p?))))]; e, p?]$$

其含义如下, 首先, 关于状态 p 的意图是一项执行一系列最终引起 p 的行为的承诺。其次, 要求这些行动是有意识的(事先相信将要发生)。最后, 要求这些行动与目标相容。

由现实性假定, CL 系统满足相容性条件 BG-ICN, 但不满足独立性条件 BG-IND。由持续目标的定义, CL 系统不满足相容性条件 GI-ICN 和 BI-ICN, 然而真正成问题的是无副作用原理。在 CL 系统中, 可以表达 p 与 q 之间的六种不同的关联。假设在所有情况下有 $P\text{-GOAL}_p$ (在某个非平凡模型中成立), 表 1 列出了相应的副作用情况。可见在后三种情况下出现副作用。

表 1 BI-SE1~6

情况	P 与 Q 的关系	$(P\text{-GOAL}_q)?$	副作用条款
1	$p \supset q$	N	BI-SE1
2	$BEL(p \supset q)$	N	BI-SE2
3	$BEL \Box(p \supset q)$	N	BI-SE3
4	$\Box BEL \Box(p \supset q)$	Y	BI-SE4
5	$\models p \supset q$	Y	BI-SE5
6	$\models p \equiv q$	Y	BI-SE6

我们将后三种分别称为“必须信念副作用”, “分析蕴含副作用”和“逻辑等值副作用”。Cohen 和 Levesque 将前两种副作用问题的解决寄希望于比可能世界语义学粒度更细的语义模型。

三、意图与现实性

A. S. Rao 和 M. P. Georgeff^[6]着重探讨了通过减弱现实性假设以达到合理性必要条件的途径。他们给出了两个意图逻辑系统, 以下分别记为 LRG 和 BRG。

LRG 是一个线性时间意图逻辑系统。LRG 的形式语言与 CL 系统类似, 但多一个基本模态词 INTEND, 相应地, LRG 的语义模型比 CL 模型多一个分量 I , 即意图可达关系。LRG 采用下述“弱现实性假设”:

$$(WC\text{-}BG) \ G[\sigma, n] \cap B[\sigma, n] \neq \emptyset;$$

$$(WC\text{-}GI) \ I[\sigma, n] \cap G[\sigma, n] \neq \emptyset;$$

$$(WC\text{-}BI) \ I[\sigma, n] \cap B[\sigma, n] \neq \emptyset;$$

其中, $G[\sigma, n] = \{\sigma', \langle \sigma, n \rangle G \sigma'\}$, $I[\sigma, n]$ 和 $B[\sigma, n]$ 类似定义。

在上述假设之下, Rao 和 Georgeff 得出结论: LRG 系统满足合理性必要条件中除 I-NEC 和 BI-SEG 之外的全部条款。然而这一论断实际上建立在混淆时态的基础上。显然上述若现实性假设的三项条款均只规定了认知态度算子语义之间的共时性关系, 而在合理性必要条件中, 条款 BI-ICN 中出现的算子 BEL 的语义应理解为“对未来的信念”。由于 LRG 系统未建立 BEL 自身及其与 GOAL 和 INTEND 的语义历时性关系, 合理性必要条件中凡涉及对 BEL 做“未来信念”理解的条款在 LRG 系统中均无法得到。

将 $B[\sigma, n]$ 直接理解为“未来信念”并不能解决上述问题, 因为合理性必要条件中的某些条款实际上要求对 BEL 做“当前信念”理解。例如, 在 BG-INC 中, 将 BEL 理解为“未来信念”是不恰当的。

上述分析表明, 在合理性必要条件中, 应进一步区分“当前信念”与“未来信念”。事实上, 意图的必要性条款只有在对信念做时态区分的基础上才能形式地刻画。由此推出下述重要结论:

一个恰当的意图逻辑系统必须能够(显式地)区分“当前信念”与“未来信念”。

BRG 是一个分支时间意图逻辑系统。弱现实性约束的表述依赖于“子世界”概念。称 W' 是 W 的子世界, 记为 $W' \subseteq W$, 当且仅当下述条件成立:

i) W' 的索引集是 W 索引集的子集;

ii) 若两个索引同时出现在 W' 和 W 中, 规定它们之间发生的事件相同;

iii) 谓词符号 W' 和 W 中的真值赋值相同;

iv) W' 和 W 上的可达关系相同。

BRG 中的弱现实性约束表达为:

$(WC_BG_B) \exists W' \in B[w, n] \text{ 使 } \exists W'' \in G[w, n] : W'' \subseteq W'$

$(WC_GL_B) \exists W' \in G[w, n] \text{ 使 } \exists W'' \in I[w, n] : W'' \subseteq W'$

$(WC_BL_B) \exists W' \in B[w, n] \text{ 使 } \exists W'' \in I[w, n] : W'' \subseteq W'$

这实际上是对 LRG 中的弱现实性约束未提供历时性关联所进行的修补。假设有 M, w, v , 和 n 使 $M, w, v, n \models \text{INTEND}(a)$, 则依 WC_BI_B , 存在 $W' \in B[w, n]$ 和 $W'' \subseteq I[w, n]$ 使 $W'' \subseteq W'$, 但并不保证对每 $W' \in B[w, n]$ 都有 $W'' \in I[w, n]$ 使 $W'' \subseteq W'$ 。这就是说, a 是否可信依赖于具体选择 (是否“进入 W'' ”)。因此, 即使信念不发生变化, 即 $M \models \text{inevitable}$ ($\text{BEL}(\text{inevitable}(a \supset \beta))$)。其中 inevitable 是 BRG 系统的模态词, 解释为“对所有未来路径”, 都不会产生必然信念副作用。

然而问题并未彻底解决, 与 CL 和 LRG 系统相同, 在 BRG 中有 $a \supset \beta \vdash \text{INTEND}(a) \supset \text{INTEND}(\beta)$ 。显然, 这种分析蕴含副作用在 CL 和 LRG, BGR 系统中是不可能消除的, 因为它是这些系统的形式载体——正规模态逻辑——的基本性质。

排除分析蕴含副作用的理由可由下例说明。考虑一个定理证明的有限自治系统。假设为了证明某定理, “决定”先证明该定理的一条引理 a , 于是形成意图 $I(a)$ 。显然, 选择以 a 为意图决不意味着采纳 a 的所有逻辑后承 β 为意图。这里的意图是 (未来) 行动的计划或方案, 该计划的确定 (选择) 和实施一般会导致确立该计划的全部逻辑后承作为系统的计划。另一个更为明显的例子是, 打算做 a 不应蕴含打算做 $a \vee b$ 。

四、意图与副作用

针对上述分析蕴含副作用问题, K. Konolige 和 M. E. Pollack 提出以非正规模态逻辑作为描述工具^[2]。KP 系统的语义模型是三元组 (W, Σ, I) , W 是可能世界的集合, $w \in W$ 是一个状态序列。假定所有解释均相对于一个特定的时刻。模态算子 $\Diamond$ 解释为:

$W, w \models \Diamond a \text{ iff } \exists w' \in W, w, w' \models a$

$\Box$ 定义为 $\Box a = \neg \Diamond \neg a$ 。 $\Sigma \subseteq W$ 用来解释信

念:

$(W, \Sigma, I) \models \beta(a) \text{ iff } \forall w \in \Sigma, w \models a$ 。

$I \subseteq P(W)$ 用来解释意图:

$(W, \Sigma, I) \models I(a) \text{ iff } \exists M \in I, M = \{w \in W, w \models a\}$

I 中的每个 M 对应于一个基本意图, 其作用相当于一个模态算子。与正规模态逻辑语义解释的不同是在上述解释规则中使用 $=$ 而不是 $\subseteq$, 这使 KP 系统排除了分析蕴含副作用。假设 $\vdash a \supset \beta$ 但是 $\not\vdash \beta \supset a$, 则 $\{w \in W : w \models a\} \neq \{w \in W : w \models \beta\}$, 故当 $(W, \Sigma, I) \models I(a)$ 时未必有 $(W, \Sigma, I) \models I(\beta)$ 。但是, 逻辑等值副作用仍然存在。

KP 系统对其语义模型有两条约束。a) 可达性: $\exists w \in \Sigma : \forall M \in I : w \in M$; b) 非平凡性: $\forall M \in I : \exists w \in \Sigma, w \notin M$ 。在此二约束下有下述性质

$\neg \text{INTEND}(a \wedge \neg a)$

$\text{INTEND}(a) \wedge \text{INTEND}(\beta) \supset \Diamond(a \wedge \beta)$

$\text{INTEND}(a) \supset \neg \text{BEL}(\neg a)$

$\text{INTEND}(a) \supset \neg \text{BEL}(a) \wedge \neg \text{BEL}(\neg a)$ (认识独立性)

若将 BEL 理解为“未来信念”, 则上述第 3 条性质即 BI-ICM。但依此理解, 认识独立性显然比 BI-ICM 强, 而且强出的部分含义十分可疑。凡想做的事一定不属于“未来信念”。这显然是不合理的。其实, 作者的本意可能是想用此性质表达意图的客观必要性, 即想做的事一定是既不属于当前信念, 其否定又不属于未来信念 (因此是必要的和可能的)。但这一含义要求区分信念的时态, 而 KP 系统的基本假定排斥了这种可能性。另外, KP 系统不包含 GOAL。

KP 系统的另一特点是考虑了意图之间的目的-手段关系的刻画。为此在语义模型中增加一个解释成分——嵌入图 $\rightarrow$, 在语言中增加一个算子 $B_\rightarrow$, 定义:

$(W, \Sigma, I, \rightarrow) \models B_\rightarrow(a, \beta_1, \dots, \beta_n) \text{ iff } \beta_1, \dots, \beta_n \rightarrow a$

在一定的条件下, 下述性质成立:

$(W, \Sigma, I, \rightarrow) \models B_\rightarrow(a, \beta_1, \dots, \beta_n) \supset B(\beta_1 \wedge \dots \wedge \beta_n \supset a)$

KP 系统是一个静态理论, 不可避免地带有静态理论的局限性。此外, 该系统无法简单地扩充为动态理论。例如, 可达性约束是说, 相信所有基本意图的“交”是可能的。如果说在静态情况下这还只是一种过强的限制的话, 那么在动态情况下就是一种不合理的限制, 因为它排斥了意图演化的“自我纠错”能力。例如, 意图 $\text{INTEND}(a)$ 的实施可以导致有限自治系统发现这是一个错误的决策, 而正确的决策是 $\text{INTEND}(\neg a)$ 。

五、有限自治系统的动态合理性

上面所分析(以及未分析)的意图逻辑系统可在两个维度上加以划分: I) 按照语义的时态性质分为静态描述与动态描述两类,前者只提供有关认知态度算子语义上的共时性规定,LRG 系统和 KP 系统属前一类,CL 系统和 BRG 系统属后一类。II) 按照模态算子的解释原则分为正规模态逻辑系统与非正规模态逻辑系统,CL 系统、LRG 系统和 BRG 系统属前一类,KP 系统属后一类。

这样就出现四种可能的组合。

a) 静态描述的正规模态逻辑系统,例如 LRG 系统。这类系统的局限性主要包括两个方面:(1) 由于不能区分“关于当前的信念”与“关于未来的信念”,因而无法同时澄清之。意图-信念相容性 BI-ICN 和目标-信念相容性 BG-ICN;(2) 意图算子的解释原则决定了不可能消除分析蕴涵副作用与逻辑等值副作用。

b) 动态描述的正规模态逻辑系统,例如 BRG 系统和 CL 系统。这类系统能够描述历时性关系,但从现有工作看,它们的动态特征描述能力及逻辑性质与语义模型的复杂度有关。例,在 BRG 系统中弱现实性约束使 BRG 系统能避免 CL 系统存在的必然信念副作用。这就引出了系统的逻辑性能与语义复杂度之间的折衷问题。另外,模态正规性决定了无法消除分析蕴涵副作用。

c) 静态描述的非正规模态逻辑系统,例如 KP 系统。这类系统最突出的优点是可以消除直到分析蕴涵的所有副作用,但仍无法消除逻辑等值副作用。此外还包含静态描述系统的共同缺陷。

d) 动态描述的非正规模态逻辑系统,目前尚无先例。此类型是最为开放的,可以有多种不同的动态描述方式和不同的非正规模态。

鉴于意图形式化研究必然涉及意图、信念和目标等算子及其相互关系,模态逻辑提供了一个较为自然、直接的表示框架。因此,用上述 d 类系统去解决前三类系统难以或无法解决的问题是一条值得探索的道路,这就要求首先抓住现存问题的关键。依据以上分析,我们认为,意图形式化研究的关键性问题有二:

1) 历时性关系的适当表达能力。所谓“适当”指可表达性与表达方式复杂性的平衡,而可表达性是第一位的。静态描述的局限性已充分证明了这种表达能力的必要性。

2) 逻辑等值副作用问题。对于有限自治系统来说,逻辑等值不等于计算等价。系统知识的不完备性和计算资源的有限性可能使目标在某一种表达形式下是可解的,而在另一种逻辑等值的表达形式下却是不可解的。

对于上述问题,从不同的观点出发可以得出不同的解决方法。我们的基本观点是:有限自治系统中的推理是一种依赖于计算环境的活动。由此得出本文的基本结论:有限自治系统的合理性条件应当建立在动态概念之上。此结论的依据并非仅仅来自上面的分析。事实上,从设计的角度看,有限自治系统的动态特征更为重要。

我们提出有限自治系统的动态合理性条件如下(BEL_i 和 BEL_j 分别表“关于未来的信念”和“关于当前的信念”,i, j, k 为时间变量,且 i < j < k)。

• 相容性

$$\begin{aligned} (BI-ICN) & \forall M \cdot M \models \text{INTEND}(a) \wedge \text{BEL}_i(\neg a) \\ (BI-ICN) & \forall M \cdot M \models \text{INTEND}(a) \wedge \text{GOAL}(\neg a) \\ (BG-ICN) & \forall M \cdot M \models \text{GOAL}(a) \wedge \text{BEL}_i(\neg a) \end{aligned}$$

• 独立性

$$\begin{aligned} (BI-IND) & \exists M \cdot M \models \text{INTEND}(a) \wedge \neg \text{BEL}_i(a) \\ & \exists M \cdot M \models \text{BEL}_i(a) \wedge \neg \text{INTEND}(a) \\ (BI-IND) & \exists M \cdot M \models \text{INTEND}(a) \wedge \neg \text{GOAL}(a) \\ & \exists M \cdot M \models \text{GOAL}(a) \wedge \neg \text{INTEND}(a) \\ (BG-IND) & \exists M \cdot M \models \text{GOAL}(a) \wedge \neg \text{BEL}_i(a) \\ & \exists M \cdot M \models \text{BEL}_i(a) \wedge \neg \text{GOAL}(a) \end{aligned}$$

• 必要性

$$(BI-NEC) \models \text{INTEND}(a) \supset \neg \text{BEL}_i(a)$$

• 无副作用原理

$$\begin{aligned} (BI-SE4) & \exists M \cdot M \models \text{INTEND}(a) \wedge \Box \text{BEL} \Box (a \supset \beta) \\ & \wedge \neg \text{INTEND}(\beta) \\ (BI-SE5) & \exists M \cdot M \models \text{INTEND}(a) \wedge \Box (a \supset \beta) \wedge \\ & \neg \text{INTEND}(\beta) \\ (BI-SE6) & \exists M \cdot M \models \text{INTEND}(a) \wedge \Box (a \equiv \beta) \wedge \\ & \neg \text{INTEND}(\beta) \end{aligned}$$

求精-演化原理

$$\begin{aligned} (BI-ELA) & \forall M \exists C_j \cdot \text{Byj}(a, \beta_1, \dots, \beta_n) \wedge C_j(\beta_1, \dots, \\ & \beta_n) \models \neg \text{INTEND}_k(a) \\ (II-ELA) & \forall M \exists C_j \cdot \text{INTEND}_i(a) \wedge C_i(a, \beta) \models \Diamond \text{IN-} \\ & \text{TEND}_{i+1}(\neg a) \end{aligned}$$

$$(II-EVO) \exists M \cdot M \models \text{INTEND}_i(a) \wedge \text{INTEND}_j(\neg a)$$

结论 知识和推理能力不完备且计算资源有限的系统称为有限系统,有能力识别自身局限和行动目的

ADA, 软件精英

ADA 前途如何?

技术顾问

——TRI-Ada'96大会印象记

学术会议

王振宇

徐宝文

(中船总七院七〇九研究所) (东南大学计算机系)

TP311.52

一、一般情况

TRI-Ada 是由美国政府、学术界、产业界三方联合召开的、一年一度的学术大会。TRI-Ada'96 于 1996 年 12 月 3 日-7 日在美国东部费城召开。这次大会的主题是：“有纪律的 Ada 软件开发”(Disciplined Software Development in Ada)。大会程序委员会主席 A. Gargaro 说：选定这个主题是因为整个世界都在关心着可靠而安全的程序设计，当年推动设计 Ada 语言的这个目标已经被其它语言所接受，商业界和工业界象政府一样，面临着软件的复杂性、可维护性和可靠性的类似问题。

出席这次大会的有来自世界各国(大部分是美国)的 Ada 同行约 600 人。安排有大会特邀报告、技术报告和分会场报告。

大会的三个特邀报告是：

B. Boehm(美国国防部顾问、DARPA 信息科学技术局主任)：美国科学院国家研究局关于 Ada 在国防部使用情况的调查结果。

D. Turner(美国陆军通信电子司令部研究中心主任)：围绕 Ada 的技术和商业两分法观点。

W. Humphrey(美国软件工程研究所研究员)：依赖软件的生活是什么样子的？

与大会同时进行的还有一个 Ada 产品展览会，集中展示了近年来新开发的 Ada 产品，这些产品都是国际最高水平的。

二、Ada 前途如何？

我研究 Ada 多年，在国内，也宣传 Ada 多年，我也知道，美国军方使用 Ada 的决心很大，取得了不少成功。但在国内一直有一部分人对 Ada 的应用持怀疑态度。听说美国国防部放弃了对 Ada 的商标控

并据此合理选择未来行动的有限系统称为有限自治系统。意图形式化研究的基本目标是澄清有限自治系统合理性的必要条件。

本文分析了四个有代表性的意图逻辑系统，着重指出了它们的缺陷及其原因，在此基础上，依据语义约束的时态性质和模态算子的解释原则划分出意图逻辑的四种类型，进而归纳出现有工作存在的两个关键问题。我们认为，这两个问题的解决有赖于对现有静态理论框架的突破，为此我们提出了“有限自治系统中的推理是一种依赖于计算环境的活动”的观点，并给出了有限自治系统动态合理性的必要条件。

我们将在动态描述的非正规模态逻辑的框架内尝试构造满足这些条件的意图逻辑系统，并解决相关的各种问题

参考文献

- [1] M. Bratman, Intention, Plans and Practical Reason, Harvard Univ. Press, Cambridge, MA, 1987
- [2] J. Doyle, Rational control of reasoning in artificial intelligence, in: LNAI465, Springer-Verlag, 1991
- [3] M. P. Georgeff, Actions, processes, and causality, in: Procs. of PRA, OR, 1986
- [4] P. R. Cohen et al., Intentions in Communication, MIT Press, 1990
- [5] P. R. Cohen et al., Intention is choice with commitment, Art. Intell. 42, 1990
- [6] A. S. Rao and M. P. Georgeff, Asymmetry thesis and side-effect problems in linear-time and branching-time intention logics, in: IJCAI-91, 1991
- [7] K. Konolige and M. E. Pollack, A representationalist theory of intention, in: IJCAI-93, 1993

王振宇 研究员、博士生导师，从事 Ada 语言、工具与环境研究。徐宝文 教授，博士生导师，系副主任，从事 Ada 语言、工具与环境、软件工程研究。