

计划中的时序偏好表示和非单调推理*

王炎娟,郝智勇,姚莉

(国防科技大学 信息系统工程重点实验室,湖南长沙 410073)

摘要:针对行动序列计划中的时序偏好问题开展研究,重点关注时序偏好集合内部存在不一致性时完成行动计划的方法。论文的难点和创新点体现在两个方面。针对时序偏好的表示问题,建立了一种二元描述结构,能够实现对时序约束和时序偏好的归一化描述。使用计算辩论技术建立推理框架,用以排除约束/偏好集合中的冲突关系,获得具有最大一致性的约束/偏好子集。在此基础上建立了一种新的计划生成策略,通过映射的方法得到行动序列片段,组合这些片段后形成完整的行动序列计划。通过实例验证了方法的可行性。

关键词:计算辩论;偏好;时序逻辑;行动序列;计划生成

中图分类号:TP181 **文献标志码:**A **文章编号:**1001-2486(2014)04-0177-07

Temporal reference representation and nonmonotonic deduction in planning

WANG Yanjuan, HAO Zhiyong, YAO Li

(Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha 410073 China)

Abstract: A research on the temporal preference in course-of-action planning is presented, focusing on the scheme to implement planning with inconsistent temporal preference. The difficulties and innovations lie in two aspects. Firstly, considering the representation issue, a unified binary format is established for temporal preference and temporal constraints. Secondly, argumentation is utilized to develop a deduction framework in order to sort out the conflicts inside the constraint/preference set, achieving a subset of constraints/preferences with maximum consistence. A plan generation strategy was then proposed to acquire segments of plan via mapping, and to combine the segments into a complete plan. A case study was done in detail to show the effectiveness of the scheme.

Key words: computational argumentation; preference; temporal logic; course of action; plan generation

定性偏好在近十几年来是人工智能领域的研究热点之一,由于其在社会学、医学乃至军事领域广泛的应用背景,研究者对定性偏好的表示及推理^[1-4]、基于定性偏好的决策^[5-7]、基于定性偏好的规划^[8-10]等展开了研究。但正如研究者指出的^[9],基于定性偏好的规划并没有得到充分的研究,一个重要的原因是,基于偏好的推理与偏好的表示方式是紧密关联的,而由于定性偏好的表示本身就是一个难点,难以形成通用规划器,因而阻碍了规划推理方面的深入研究。本文针对行动规划中不可避免的时序关系,考察规划中时序偏好基础上的规划,重点关注偏好关系中冲突时,借助非单调推理实现规划的过程。

举例来说,用户希望动作 A1 最好在动作 A2 之前完成,希望动作 A3 在上午 10 点之后进行,

建议动作 A4 安排在所有动作之后,这都属于时序偏好。如果增加一个偏好,要求动作 A4 在上午 10 点之前结束,那么就出现了偏好间的冲突:规划器不可能同时把动作 A3 安排在上午 10 点之后、A4 安排在上午 10 点之前,而保证 A4 是最后执行的动作。显然,如果要满足最后加入的偏好,就必须推翻前面的推理结果,这是非单调推理才能胜任的。

时序偏好是定性偏好的一种。定性偏好的表示方式很多,包括:C-P网^[11]、时序范式^[9]、优先目标^[12]等,但并不是每一种表达方式都能够支持时序偏好。比如,C-P网中的每个偏好关系的默认语义是“选择”,即: x 偏好于 y ,意味着决策者更喜欢选择 x 不是选择 y ;而时序偏好包含“时序语义”,如: x 先于 y 发生,比 y 先于 x 发生更受

* 收稿日期:2013-03-20

基金项目:国家自然科学基金资助项目(71371184)

作者简介:王炎娟(1984—),女,江苏大丰人,博士研究生,E-mail:nudtwyj@gmail.com;

姚莉(通信作者),女,教授,E-mail:liyao6522@gmail.com

偏好。因此,尽管 C-P 网表示方式被广泛使用,但是它不能支持时序偏好。与此相对的,时序范式引入 until, next 等时态谓词,因而能够清晰表达复杂时序关系上的偏好。时序范式的表达能力在规划推理中得到了验证^[13]。

偏好是一种软约束,这意味着,如果偏好与传统意义上的约束(强制性的约束,硬约束)产生语义冲突时,它是可以被舍弃的;类似地,如果不同偏好之间出现冲突,需要使用一定的规则进行取舍。

针对定性偏好集合和约束集合中可能存在的冲突关系,Blom 在其研究中提出了“偏好解耦”的处理框架^[14]。所有的约束和偏好组成一个有冲突的信念集合,通过计算辩论方式排除其中的冲突关系,此过程中有些偏好关系被删除,最后获得的约束/偏好子集形成了信念上的“最大公约”,在此基础上再开展决策。Blom 选用的计算辩论就是一种流行的非单调推理工具。与辩论框架相适应,她选用了比较格式作为偏好的表达方式。比较格式来源于 Ouerdane 关于论据结构的研究^[15],体现了偏好是一种二元关系的本质特点,其结构化的特点使之非常适合于辩论过程,容易生成辩论所需的论据并建立论据间的攻击关系。Blom 所讨论的偏好局限在静态属性偏好,没有涉及时序偏好。

在面向多用户的辅助规划系统中,约束与偏好共存、多种偏好共存的现象是普遍存在的。各用户对计划的制定分别提出自己的偏好,辅助规划的软件智能体则应协调各方的期望,处理各方所提定性偏好与定性约束之间的冲突关系,使所有参与者或尽可能多的参与者满意,实现“最大一致性”。此类案例可见于军事领域中的多方协调作战、医学领域的多专家会诊等。在此背景下,本文把 Blom 的研究扩展到时序偏好基础上的偏好解耦规划,借鉴时序逻辑语言的偏好表示^[13]和 Blom 使用的比较格式,为时序偏好和时序约束引入一种新的二元表示方式,使用计算辩论获得最大一致性,然后由偏好和约束关系通过映射的方式生成可用规划片段,最后借鉴规划空间规划中的细化(refinement)方式^[16],链接、补齐规划片段完成规划方案。

1 辩论基础

辩论框架为辩论提供了形式化建模方法,建立一种处理不确定、不一致信息的基础。下面简要介绍 Dung 辩论框架^[17]及其相关性质。

辩论框架通常定义为二元组 $AF = (\mathcal{A}, \mathcal{R})$,其中 \mathcal{A} 表示论据集, \mathcal{R} 是定义在论据集 \mathcal{A} 上的二元(攻击)关系,即有 $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ 。

在 $AF = (\mathcal{A}, \mathcal{R})$ 中,若 \mathcal{A} 是有限集,称 AF 有限的,否则称为无限的。 $\forall a, b \in \mathcal{A}, a \mathcal{R} b$ 或 $(a, b) \in \mathcal{R}$ 表示 a 攻击 b ; a, b 间不存在攻击关系,表示为 $a \not\mathcal{R} b$ 或 $(a, b) \notin \mathcal{R}$, 亦称 a, b 是无冲突的。 $\mathcal{R}_+(a) = \{b \in \mathcal{A} | b \mathcal{R} a\}$ 表示被 a 攻击的论据集, $\mathcal{R}_-(a) = \{b \in \mathcal{A} | b \mathcal{R} a\}$ 表示攻击 a 的论据集;对于论据集 $S \subseteq \mathcal{A}$ 和论据 $b \in \mathcal{A}$, 若 $\exists a \in S$, 使得 $a \mathcal{R} b$, 则称 S 攻击 a 。同理,有 $\mathcal{R}_\omega(S) = \cup_{a \in S} \mathcal{R}_\omega(a)$ ($\omega \in \{+, -\}$)。

- (1) 论据集 $S(S \subseteq \mathcal{A})$, 若 $\forall a, b \in S, (a, b) \notin \mathcal{R}$, 称 S 是无冲突的(conflict-free)。
- (2) a 相对于 S 是可接受的, 当且仅当任意对 a 的攻击论据, S 中均有反驳论据, 即 $(\forall b \in \mathcal{A}. b \mathcal{R} a \rightarrow (\exists c \in S). c \mathcal{R} b)$
- (3) S 是无冲突的, 且 $\exists a \in S$ 相对于 S 是可接受的, 称 S 为可容许集。
- (4) 称可容许集 S 为 AF 的首选扩展, 当且仅当 S 是 AF 的极大可容许集, 写作 S_{pref} 。
- (5) 称可容许集 S 为 AF 的稳定扩展, 当且仅当所有不属于 S 的论据都被 S 攻击。
- (6) 称可容许集 S 为 AF 的完全扩展, 当且仅当所有相对于 S 可接受的论据都包含于 S 中。
- (7) 称可容许集 S 为 AF 的基础扩展, 当且仅当 S 是 AF 的极小完全扩展(集包含)。

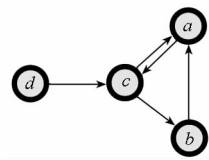


图 1 辩论框架

Fig. 1 Argumentation framework

$AF = (\mathcal{A}, \mathcal{R})$ 可以表示为有向图,称为攻击图。如图 1 所示的攻击图中,

$$\mathcal{A} = \{a, b, c, d\}$$

$$\mathcal{R} = \{(a, c), (b, a), (c, b), (c, a), (d, c)\}$$

由此可知,

$$\mathcal{R}_+(c) = \{a, b\} \text{ 及 } \mathcal{R}_-(c) = \{a, b\}$$

根据上述定义,图 1 中的首选扩展集合为 $S_{pref} = \{d, b\}$ 。

2 行动计划中的时序偏好表示

2.1 行动计划建模

本文对行动计划问题的描述参考了情景演算

(situation calculus) 语言^[18-19]的方式,详见定义1和定义2。

定义1(行动计划) 行动计划问题为一个三元组 (D, I, G) ,其中 I, G 分别表示初始状态和目标, D 为动作理论。

定义2(动作理论) 给定行动计划问题 $\Pi = (D, I, G)$,动作理论 D 为一个五元组 (L, \circ, S, do, K) ,其中 L 为动作理论的一阶逻辑语言; S 为一阶语言描述的情景集合,满足 $I, G \in S$; \circ 为动作的集合; do 是状态转换函数; K 为情景演算公理系统。

给定情景 $s_0, s_1 \in S, o \in \circ$,情景演算中的基本推理格式为

$$do[o, s_0] \supset s_1 \quad (1)$$

表示情景 s_0 在动作 o 的作用下到达情景 s_1 。

为了表述的方便,定义一种特殊的动作,成为“空动作”,它是以时间为变量的一阶谓词 $\Theta(t)$,表示在某时间执行的虚拟动作,该动作的执行不引起世界状态的变化,即

$$(\forall s \in S)(\forall t) s = do[\Theta(t), s]$$

情景 s 由多个谓词的合取进行刻画:

$$s \Leftarrow r_1(x_1) \wedge r_2(x_2) \wedge \dots \wedge r_n(x_n) \quad (2)$$

其中 r_1, r_2, \dots, r_n 为一阶逻辑语言 L 实例化后形成的一阶谓词表达式,意味着在情景 s 上 $r_i(x_i)$ 取值均为TRUE。

记 $G \Leftarrow g_1(x_1) \wedge g_2(x_2) \wedge \dots \wedge g_k(x_k)$,则规划目标对应了包含多个子目标 $g_i(x_i)$ 的集合。在确定了所有的子目标之后,计划的过程就是要获得实现这些子目标所需的动作集合,并使这些动作具有合理的时序安排。动作计划的方案具有如下形式:

$$I, o_1, s_1, o_2, s_2, \dots, o_i, s_i, \dots, o_n, G \quad (3)$$

表示初始情景在动作序列的作用下,逐步演化到达最终目标。

行动计划中的时序关系统一在时序关系论域上进行讨论。

定义3(时序关系论域) 给定动作理论 $D = (L, O, S, do, K)$,时序关系论域是最小集合 \mathcal{B} 里的一个表达式,满足

- (1) 空元素 $\emptyset \in \mathcal{B}$;
- (2) $\forall s \in S, s \in \mathcal{B}$;
- (3) 如果 $o \in O$,则 $occ(o) \in \mathcal{B}$;
- (4) 如果 $o \in O$,则 $fin(o) \in \mathcal{B}$ 。

谓词 occ 用于指示一个动作的发生,谓词 fin 则表示一个动作的结束。

定义4(时间函数) 给定时序关系论域 \mathcal{B} ,

时间函数 T 是从 \mathcal{B} 到时间的映射函数。时间函数的取值是可以比较大小的。

例如,

$$T(occ(eat(pizza))) = 20:00$$

表示吃比萨饼这个动作的发生时间是晚上20:00。

论域 \mathcal{B} 中元素间的时序关系是复杂多样的,可能的关系包括了顺序、并列、重叠等。作为时序偏好应用于行动计划的一个开端,本文对动作计划问题提出如下简化假设:

- (1) 一个动作导致的结果是唯一确定的。也就是说,如果 $do(o_i, s) \supset s_i$ 且 $do(o_i, s) \supset s_j$,则 $s_i = s_j$ 。
- (2) 引起某种情景迁移的动作也是唯一确定的。也就是说,如果 $do(o_i, s) \supset s_i$ 且 $do(o_j, s) \supset s_i$,则 $o_i = o_j$ 。
- (3) 动作是顺序发生的,没有时间上的重叠。也就是说,如果有 $T(occ(o_i)) < T(occ(o_j))$,则 $T(fin(o_i)) < T(occ(o_j))$ 。

2.2 时序约束与时序偏好

借鉴Blom的偏好比较格式,可以在论域 \mathcal{B} 上给出时序约束和时序偏好关系的描述。

定义5(时序约束关系) 时序约束关系 \gg 是定义在论域 \mathcal{B} 内两个元素上的二元偏好关系,记为

$$b_k: b_i \gg b_j$$

表示前提条件 b_k 成立时, $T(b_i) \leq T(b_j)$ 成立。当约束不需要前提条件时,约束简写为:

$$b_i \gg b_j \quad (4)$$

定义中用符号 \gg 表达了“……在……之前”的语义,描述了一种时间轴上的约束关系。当 $b_j = \emptyset$ 时,式(4)简写为 b_i ,意为:在没有可比较选项时,用户指定 b_i 。

关系符 \gg 还可以带有数字下标,如

$$b_i \gg_{\Delta t} b_j \quad (5)$$

表示 $T(b_i) \leq T(b_j) - \Delta t$ 成立。 Δt 称为时延。易知,式(5)成立,必然有式(4)成立。

对于物理世界的逻辑系统,前提和结论必然在时间轴上具有前后关系。因此情景演算中的基本推理式可以归纳到时序约束的形式:

$$s_0: occ(o_1) \gg s_1 \quad (6)$$

时序偏好在结构上与时序约束类似,不同之处在于其语义不是强制性的。

定义6(时序偏好) 给定论域,时序偏好具有以下形式:

$$b_i \triangleright b_j \quad (7)$$

意为:用户希望 b_i 发生在 b_j 之前,而不是 b_j 发生在 b_i 之前。

与时序约束类似,可以定义带有时延参数的时序偏好:

$$b_i \triangleright_{\Delta t} b_j \quad (8)$$

表示用户偏好于 b_i 发生在 b_j 前,且间隔至少 Δt 时间。

给出如下的偏好表述:

- 决策者偏好于最后执行动作 o
- 决策者喜欢在条件 X 满足之后执行 o

偏好表达式分别为:

$$(\forall t) \text{occ}(\Theta(t)) \triangleright \text{occ}(o)$$

$$X \triangleright \text{occ}(o)$$

可见,时序范式中需要使用 final, next 等时态谓词才能描述的关系可以用式(7)给出的比较格式来表达,实际上,时态谓词不易表述的关系也可以用比较格式进行描述,例如:

- 希望在下午 5:00 之前执行动作 o
- 希望动作至少执行 10 分钟

对应的偏好表达式分别为

$$\text{occ}(o) \triangleright \text{occ}(\Theta(5:00pm))$$

$$\text{occ}(o) \triangleright_{10} \text{fin}(o)$$

至此,我们用 \triangleright 和 \triangleright 实现了时序约束和时序

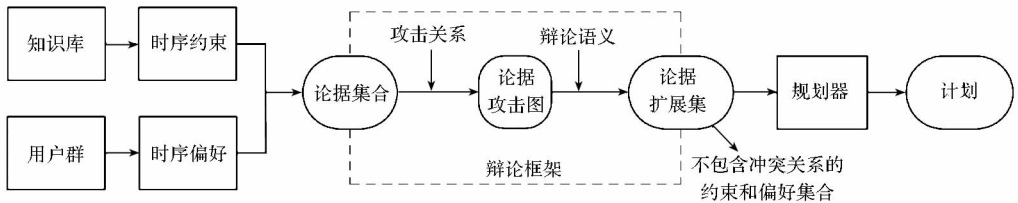


图 1 基于辩论的偏好解耦规划实施流程图

Fig. 1 Flowchart for preference-decoupling planning using argumentation

按照该流程,约束和偏好形成论据,进入辩论框架进行辩论,得到的结果是论据的扩展集合,对应了一个无内部冲突的偏好/约束集合。下面要依次明确论据的构造方法、攻击关系的构造方法、辩论语义和规划器。

每个时序约束对应一个认知论据:

$$a_e = \langle [b_k], b_i \gg b_j, [\Delta t] \rangle$$

其中前提条件 b_k 和时延数值 Δt 为可选项。 a_e 的前提和结论分别记为

$$\text{pre}(a_e) = b_k, \text{con}(a_e) = b_i \gg b_j$$

所有的认知论据集合记为 A_e 。

每一个时序偏好对应一个实用论据:

$$a^p = \langle b_i \triangleright b_j, [\Delta t] \rangle$$

其中时延数值 Δt 为可选项。该论据前提条件为空,论据的结论记为:

偏好的结构化定义。

下面给出时序约束/偏好间的优先级比较关系 \geq_p :

$$P_i \geq_p P_j \quad (9)$$

式(9)表示在对等的条件下, P_i 的优先级不低于 P_j 。

需要指出的是,时序约束的优先级必然比时序偏好优先级高。时序偏好之间的优先级关系则需要单独给出定义。

最后给出时序偏好行动计划问题的定义。

定义 7(时序偏好行动计划问题) 给定行动计划问题 Π ,时序偏好行动计划问题可定义为一个四元组 (Π, C, P, W) ,其中 C, P 分别为定义在问题 Π 上的时序约束集合和时序偏好集合, W 为定义在 P 上的优先级比较关系集合。

3 定性偏好的辩论推理和行动计划生成

3.1 非单调推理的辩论框架

本文参照 Blom 的偏好解耦决策框架设计行动计划的辩论框架。在文献[14]中找到更为详细的介绍。图 1 给出了偏好解耦规划的实施流程图。

$$\text{con}(a_p) = b_i \triangleright b_j$$

所有的实用论据集合记为 A_p 。

定义 8(时序偏好解耦的辩论框架 P-ADM) P-ADM 辩论框架是一个 3 元组 (A_e, A_p, X) ,其中

- (1) A_e 是一个认知论据的有限集合,
- (2) A_p 是一个实用论据的有限集合,
- (3) $X = (R, \geq, Sem)$, R 是一个非自反性的论据冲突关系, \geq 是一个具有自反且传递性的论据强度比较关系, Sem 为某种辩论语义。

定义 9(冲突关系 R) R 是一个非自反性的二元冲突关系。给出两个论据 $a_i, a_j \in A_e \cup A_p$, $(a_i, a_j) \in R$ 的条件为:对于 $a_i \in A_e, a_j \in A_e \cup A_p$ 或 $a_i, a_j \in A_p$,

$$\text{Conc}(a_i) = b_m \triangleright b_n, \text{Conc}(a_j) = b_n \triangleright b_m。$$

其中 $b_m, b_n \in \mathcal{B}$ 。

如果两个论据对应的偏好(约束)满足偏好强度关系 \geq_p , 这两个论据满足论据强度比较关系 \geq 。于是在冲突关系 R 和强度比较关系 \geq 的基础上,可以定义论据间的攻击关系。

记 $>$ 为 \geq 的严格关系, $A_e \cup A_p$ 上的二元攻击关系记为 w , 称 $(a_i, a_j) \in w$ 当且仅当 $(a_i, a_j) \in R$ (a_i 与 a_j 冲突) 且 $(a_i, a_j) \notin >$ (即, a_j 并不是严格偏好于 a_i 的)。

辩论语义 Sem 将沿用 Blom 使用的首选扩展语义 (preferred extension)。使用首选扩展集合定义最大满意度, 其内涵是: 优选出来的偏好关系数量是最多的。如果选用稳定扩展语义 (stable extension) 或者基础扩展语义 (ground extension), 优选偏好关系数量可能会减少, 但会获得其他意义上的满意度, 比如偏好提出者会感到基础扩展语义给出的偏好集合更具有公平性。限于篇幅, 本文不再做深入探讨。

时序约束和时序偏好经过了辩论框架之后, 应用 Dung 扩展语义之后即可得到扩展集合, 记为

$$EX = \{a_1, a_2, \dots, a_n\}$$

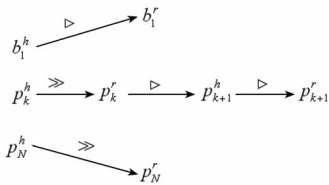
其中 $con(a_k) = b_k^h \triangleright b_k^r, k = 1, 2, \dots, n$ 。

EX 对应了一个无冲突的约束/偏好集合。下面将在其基础上设计规划器以获得行动计划。

3.2 规划器

规划器给出一个计划生成策略, 在 EX 的限定下, (1) 判定 \mathcal{B} 的哪些元素是行动计划所需要的; (2) 确立它们的时序关系, 最后确立行动计划。

首先把扩展集合用有向图表示出来。有向图的边是时序关系 \gg 或偏好关系 \triangleright , 优势图的顶点为 b_k^h, b_k^r 等论域元素。



前后衔接的多个边链接在一起就组成一个支链, 构成一个谓词序列

$$\dots b_k^h, b_k^r, b_{k+1}^h, b_{k+1}^r \dots \quad (10)$$

序列 (10) 蕴含了一个计划片段, 其中有动作和情景相关的谓词表达式, 通过翻译和解释, 就可以生成形如序列的计划序列。如果优势图可以划分为 K 个支链, 那么就对应了 K 个计划方案片段。

计划生成策略的目的, 就是组合已有的计划

片段, 建立完整的计划结果。

(1) 优势图中已经存在的节点, 自然是 COA 序列中的节点;

(2) 能够补足 COA 序列的语法结构, 而又没有在被淘汰的论据中出现的情景或动作, 是可接受的新节点, 加入计划片段;

(3) 已进入计划片段的节点, 其自然延伸导致的时序状态也是可接受的节点。

所谓自然延伸, 是指以下几种情况。

① 动作开始必然在动作结束之前:

$$occ(o) \gg fin(o) \quad (11)$$

② 在一段时间只能做一件事情, 两个动作执行的时间不能出现重合, 约束用形式化语言表达如下:

$$\{occ(o_i) \gg occ(o_j)\}; fin(o_i) \gg occ(o_j) \quad (12)$$

可以看到, 这里的计划生成策略与规划空间的细化阶段相类似^[16]。规划空间规划的细化阶段, 主要目的是处理顺序和因果关系上的矛盾或瑕疵; 而在本文中, 冲突消解在得到规划序列片段之前已经完成了。

4 案例研究

本节基于一个生活场景中的行动计划问题来展示行动计划的生成过程。

例 (球赛问题) 晚饭后, John 有三件事情要做: 出去散步, 给远方的朋友打个电话, 看电视上的足球比赛。时序约束条件包括:

(C1) 球赛 20:30 开场, 上下半场各 45 分钟, 中场休息 30 分钟。

(C2) 21:00 点之前天气不好。

John 有如下偏好:

(P1) 能够完整看完全场球

(P2) 能够看完完整的下半场球

(P3) 散步最好是在晴天进行

(P4) 散步希望在 9 点之前完成

(P5) 散步时间超过 30 分钟

(P6) 在开始看球赛之前打完电话

(P7) 打电话的时间超过 10 分钟

最后, 在 John 的心目中, 偏好 (P2) (P3) 比其他偏好更重要。

对 John 的生活安排进行计划, 就是要决定他去散步、打电话和看球赛的开始时间和持续时间。行动计划的目的是在时间轴上安排以下时序项:

$occ(walk), occ(watchTV), occ(phone),$

$fin(walk), fin(watchTV), fin(phone)$

时序约束可表示为

$$\text{occ}(\Theta(21:00)) \gg \text{GoodWeather} \quad (13)$$

7 个时序偏好的形式化表达如下:

$$P1: \begin{matrix} \text{occ}(\text{watchTV}) \triangleright \text{occ}(\Theta(20:30)) \\ \text{occ}(\Theta(22:30)) \triangleright \text{fin}(\text{watchTV}) \end{matrix} \quad (14)$$

$$P2: \begin{matrix} \text{occ}(\text{watchTV}) \triangleright \text{occ}(\Theta(21:45)) \\ \text{occ}(\Theta(22:30)) \triangleright \text{fin}(\text{watchTV}) \end{matrix} \quad (15)$$

$$P3: \text{GoodWeather} \triangleright \text{occ}(\text{walk}) \quad (16)$$

$$P4: \text{occ}(\text{walk}) \triangleright \text{occ}(\Theta(21:00)) \quad (17)$$

$$P5: \text{occ}(\text{walk}) \triangleright_{30} T(\text{fin}(\text{walk})) \quad (18)$$

$$P6: \text{fin}(\text{phone}) \triangleright \text{occ}(\text{watchTV}) \quad (19)$$

$$P7: \text{occ}(\text{phone}) \triangleright_{10} \text{fin}(\text{phone}) \quad (20)$$

John 对偏好关系进行了优先级排序:

$$W1: P2 >_p P1 (P4, P5, P6, P7)$$

$$W2: P3 >_p P1 (P4, P5, P6, P7)$$

约束(C2)对应一个认知论据,记为 a_1 。偏好(P1)(P2)分别对应 2 个实用论据,其余偏好各自对应一个实用论据,依次记为 $a_2 \sim a_{10}$ 。

根据 John 的偏好之间的强度关系为

$$a_1 > a_k, k = 2, 3, 4, 5, 6, 7, 8, 9, 10$$

$$a_3 > a_k, a_4 > a_k, a_5 > a_k, k = 2, 6, 7, 8, 9, 10$$

应用定义 9、定义 10 建立攻击关系,然后应用首选扩展语义进行辩论,最后形成的首选扩展集合为

$$\{a_1, a_4, a_5, a_6, a_8, a_9, a_{10}, a_{11}, a_{12}\}$$

辩论过程排除了论据 a_4, a_5, a_6 , 对应偏好表达式和。

把扩展集合中的元素布置到一个有向图中,得到三个规划支链:

$$(1) \text{occ}(\text{watchTV}) \triangleright \text{occ}(\Theta(21:45))$$

$$(2) \text{occ}(\Theta(22:30)) \triangleright \text{fin}(\text{watchTV})$$

$$(3) \text{occ}(\Theta(21:00)) \gg \text{GoodWeather} \triangleright$$

$$\text{occ}(\text{walk}) \triangleright_{30} \text{fin}(\text{walk})$$

$$\text{occ}(\text{phone}) \triangleright_{10} \text{fin}(\text{phone})$$

按照其中涉及的时间点顺序进行排列,引用时序约束,自动补齐其中的缺项,删除冗余信息,即可得到如下行动计划:

$$\begin{aligned} &\text{occ}(\Theta(21:00)) \gg \text{occ}(\text{walk}) \triangleright_{30} \text{fin}(\text{walk}) \\ &\triangleright \text{fin}(\text{walk}) \triangleright \text{occ}(\Theta(21:00 + (30 + T_1))) \\ &\triangleright \text{occ}(\Theta(21:45 - (10 + T_2))) \triangleright \text{occ}(\text{phone}) \\ &\triangleright_{10} \text{fin}(\text{phone}) \triangleright \text{fin}(\text{phone}) \\ &\triangleright \text{occ}(\text{watchTV}) \triangleright \text{occ}(\Theta(21:45)) \\ &\triangleright \text{fin}(\text{watchTV}) \end{aligned} \quad (21)$$

也就是说,John 应该晚上 21:00 出门散步,散

步回来之后,给朋友打电话,在 21:45 之前坐下来准备看电视。散步和打电话的时间分别为 $30 + T_1$ 分钟和 $10 + T_2$ 分钟, T_1, T_2 可以在 1, 2, 3, 4, 5 (分钟)内取值,但是要满足 $T_1 + T_2 < 5$ 。也就是说,表达式给出了用户最满意的行动计划,该计划中动作的时间安排具有 5 分钟的可调配裕量。

可以看到,应用本文建立的计划策略时,从时序约束和时序偏好过渡到计划片段时不依赖于经典规划中常用的搜索。这种处理约束和偏好的简洁方式首先要得益于时序约束和时序偏好的二元描述方式,而辩论工具则是其基础。

5 相关研究

针对一般性偏好满意度的研究出现较早^[20-21],但早期的实现方式都是某种搜索策略,伴随着决策或者规划得过程实现的。本文给出的最大满意度是基于不一致性定义的,在无冲突的偏好数量最多时得到最大满意度。不一致性是一个经典问题,其研究开始于 1990 年代^[22-23],其后研究者尝试了多种推理方式^[24-25]。最新的研究成果是使用计算辩论工具^[26],该文献通过辩论实例化的过程中把不一致性问题纳入抽象辩论体系,为智能体中的不一致性研究提供了较完备的框架。

本文的研究也是 Dung 抽象辩论体系的一种实例化。在实例化的过程中使用了论据结构化的技巧^[15, 27]。论据的结构是一种简单的比较格式 (comparison scheme),有两个选项参与比较。比较式是论据的结论,而论据的前提不是必需的。本文的特殊之处在于,比较是一种时序比较,表征的是前后关系。论据的攻击关系也比较简洁,是针对比较式结论正确性的攻击,实际上是一种“驳斥 (rebutting)”,没有使用 Prakken 定义的“未定 (undermining)”和“削弱 (undercutting)”这两种攻击关系。辩论实例化的另一个重要方面是形式化表达^[28-29]。本文使用使用了一阶谓词逻辑和确定性推理规则,并没有使用可废止推理规则,相对于典型的辩论实例是一种简化。

与本文研究最接近的是 Blom 关于偏好解耦的研究。Blom 在其博士论文中重点研究了如何使用辩论在偏好集合的基础上做出决策。其特点是,偏好是论据的内容,而以前偏好在辩论中的角色是参与论据攻击关系的构造。这种方式下,辩论并不能直接获得决策结果,需要借助于额外的推理器(如 Blom 使用的社会决策理论),但获得的好处是:偏好的具体形式和内容与决策推理过

程实现了解耦。本文把偏好内容扩展到了时序偏好,把决策扩展到了动作规划。

6 总结与展望

本文讨论了行动计划中时序偏好与时序约束共存,且存在冲突关系时的计划生成方法。首先为时序约束和时序偏好的表示建立了结构化的描述方式,然后使用辩论工具排除约束/偏好集合中的不一致性,并最终建立了一种不依赖于节点搜索的规划方案。本文给出的成果是使用偏好指导规划的一个尝试,也是 Blom “偏好解耦”思想在行动计划中的扩展。所建立的时序偏好表示方式对于一般定性偏好的研究也具有借鉴意义。

本文的研究结果还是比较初步的,比如还不能处理带有前提条件的偏好和组合偏好,在辩论语义上仅使用了首选扩展语义。未来的研究将进一步丰富偏好关系的表达方式,同时从攻击关系、辩论语义等方面对辩论框架进行完善。

参考文献 (References)

- [1] Lang J. Logical preference representation and combinatorial vote[J]. *Annals of Mathematics and Artificial Intelligence*, 2004, 42 (1): 37 - 71.
- [2] 张志政. 定性偏好的逻辑表示和推理[D]. 南京:东南大学,2007.
ZHANG Z Z. Logic representation and deduction of qualitative preferences[D]. Nanjing: Southeast University, 2007. (in Chinese)
- [3] Delgrande J P, Schaub T, Tompits H. A general framework for expressing preferences in causal reasoning and planning[J]. *Journal of Logic and Computation*, 2007, 17 (5): 871 - 907.
- [4] Brafman R, Domshlak C. Preference handling—an introductory tutorial[J]. *AI Magazine*, 2009, 30(1): 58 - 95.
- [5] Son T, Pontelli E. Planning with preferences using logic programming [J]. *Logic Programming and Nonmonotonic Reasoning*, 2004: 247 - 260.
- [6] Rudova H. Constraint satisfaction with preferences[D]. Brno: Masaryk University, 2001.
- [7] Giunchiglia E, Maratea M. Planning as satisfiability with preferences[C]//National Conference on Artificial Intelligence. Menlo Park, CA, Cambridge, MA, London, AAAI Press, MIT Press, 2007: 987 - 992.
- [8] Badaloni S, Falda M, Giacomini M. Solving temporal over-constrained problems using fuzzy techniques [J]. *Journal of Intelligent and Fuzzy Systems*, 2007, 18 (2): 255 - 265.
- [9] Biennu M, Fritz C, McIlraith S A. Planning with qualitative temporal preferences[J]. *KR*, 2006, 6: 134 - 144.
- [10] Rossi F, Venable K B. Uncertainty in soft temporal constraint problems: A general framework and controllability algorithms for the fuzzy case [J]. *Journal of Artificial Intelligence Research*, 2006, 27(1): 617 - 674.
- [11] Boutilier C, Brafman R I, Domshlak C, et al. CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements [J]. *J. Artif. Intell. Res. (JAIR)*, 2004, 21: 135 - 191.
- [12] Sardiña S, Shapiro S. Rational action in agent programs with prioritized goals [C]//Proceedings of the second international joint conference on Autonomous agents and multiagent systems. ACM, 2003: 417 - 424.
- [13] Biennu M, Fritz C, McIlraith S A. Specifying and computing preferred plans[J]. *Artificial Intelligence*, 2011, 175 (7): 1308 - 1345.
- [14] Blom M. Arguments and actions: decoupling preference and planning through argumentation [D]. Melbourne, University of Melbourne, Department of Computer Science and Software Engineering, 2011.
- [15] Ouerdane W, Maudet N, Tsoukias A. Argument schemes and critical questions for decision aiding process [C]//Proceedings of Computational Models of Argument (COMMA). IOS Press, 2008: 285 - 296.
- [16] Kambhampati S. Refinement planning as a unifying framework for plan synthesis[J]. *AI Magazine*, 1997, 18 (2): 67 - 72.
- [17] Dung P M. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games[J]. *Artificial Intelligence*, 1995, 77 (2): 321 - 357.
- [18] 武云鹏. 基于效果的行动过程建模与优化方法研究[D]. 长沙:国防科技大学,2006.
WU Y P. Research on the modeling and optimization of effect-based course-of-action [D]. Changsha: National University of Defense Technology, 2006. (in Chinese)
- [19] Biennu M, McIlraith S. Qualitative dynamical preferences in the situation calculus [C]//Multidisciplinary IJCAI - 05 Workshop on Advances in Preference Handling, 2005.
- [20] Haralick R M, Elliott G L. Increasing tree search efficiency for constraint satisfaction problems [J]. *Artificial Intelligence*, 1980, 14 (3): 263 - 313.
- [21] Kumar V. Algorithms for constraint-satisfaction problems: A survey[J]. *AI magazine*, 1992, 13 (1): 32 - 44.
- [22] Benferhat S, Cayrol C, Dubois D, et al. Inconsistency management and prioritized syntax-based entailment [C]//Proceedings of the 13th International Joint Conference on Artificial Intelligence. Chambéry, France, 1993. 640 - 645.
- [23] Finkelstein A C W, Gabbay D, Hunter A, et al. Inconsistency handling in multiperspective specifications[J]. *IEEE Transactions on Software Engineering*, 1994, 20 (8): 569 - 578.
- [24] Sichman J S, Demazeau Y. Exploiting social reasoning to deal with agency level inconsistency [C]//Proceedings of the First International Conference on Multiagent Systems. San Francisco, California, USA, 1995: 352 - 359.
- [25] Batens D. Towards the unification of inconsistency handling mechanisms[J]. *Logic and Logical Philosophy*, 2004, (8): 5 - 31.
- [26] Liao B, Huang H. ANGLE: an autonomous, normative and guidable agent with changing knowledge [J]. *Information Sciences*, 2010, 180 (17): 3117 - 3139.
- [27] Prakken H. An abstract framework for argumentation with structured arguments[J]. *Argument and Computation*, 2010, 1(2): 93 - 124.
- [28] Caminada M, Amgoud L. On the evaluation of argumentation formalisms[J]. *Artificial Intelligence*, 2007, 171 (5 - 6): 286 - 310.
- [29] Garcia A J, Simari G R. Defeasible logic programming: an argumentative approach [J]. *Theory Practical Logic Programming*, 2004, 4 (2): 95 - 138.