

# 任务级行动序列问题中的定性偏好研究

王炎娟<sup>1,2</sup>, 姚莉<sup>1</sup>, 刘斌<sup>1</sup>

(1. 国防科技大学 信息系统工程重点实验室, 湖南 长沙 410073; 2. 北京航天飞行控制中心, 北京 100094)

**摘要:**关注了一类典型行动序列, 研究如何在动作集合上存在定性偏好, 且偏好集合存在不一致性时开展规划。所考虑的行动序列问题称为任务级 COA, 以抽象层次的动作作为基本要素, 所考虑的定性偏好包括静态偏好和时序偏好, 所讨论的规划目的是获得最大满意度的 COA 方案。首先建立了偏好与约束的归一化形式描述, 在此基础上形成了 COA 方案设计算法; 进一步, 使用计算辩论技术排除偏好集合中的不一致性, 形成用户接受度最高的 COA 方案。文中建立的以定性推理为基础的规划框架, 实现了偏好解耦, 能够适应不同的领域问题, 是以定量计算为基础的传统规划算法的有效补充。通过快速响应卫星成像的 COA 案例, 演示了算法的可行性。

**关键词:**行动序列; 规划; 定性偏好; 时序偏好; 计算辩论技术; 偏好解耦

**中图分类号:** TP391 **文献标志码:** A **文章编号:** 1673-4785(2014)05-0551-09

**中文引用格式:** 王炎娟, 姚莉, 刘斌. 任务级行动序列问题中的定性偏好研究[J]. 智能系统学报, 2014, 9(5): 551-559.

**英文引用格式:** WANG Yanjuan, YAO Li, LIU Bin. Research on qualitative preference in planning of task-level Course-of-action [J]. CAAI Transactions on Intelligent Systems, 2014, 9(5): 551-559.

## Research on qualitative preference in planning of task-level Course-of-action

WANG Yanjuan<sup>1,2</sup>, YAO Li<sup>1</sup>, LIU Bin<sup>1</sup>

(1. Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha 410073 China; 2. Beijing Aerospace Control Center, Beijing 100094, China)

**Abstract:** This paper focuses on a special type of course-of-action. Specifically, performing study on planning with the existence of qualitative preferences and functions on the actions and owns the inner inconsistency. The course-of-action that is taken into consideration is called ‘task-level’ course-of-action (COA), with abstracted action as basic element. The qualitative preferences in discussion include static preferences and temporal preference. The objective of planning is a COA plan with satisfaction. Firstly, a unified formulated description is established for constraints and preferences, based on which an algorithm for COA planning is developed. Furthermore, computational argumentation is utilized to exclude inconsistency in the set of preferences, to maximize the user’s satisfaction for COA planning. The planning framework based on qualitative deduction is an effective add-in for conventional planning scheme based on quantitative computation. The property of preference-decoupling makes itself adaptable to applications in different domain. A case study on scheduling responsive imaging satellites is proposed to demonstrate the effectiveness of the scheme.

**Keywords:** course-of-action; planning; qualitative preference; temporal preference; computational argumentation; preference decoupling

者的广泛关注<sup>[1-2]</sup>。COA 的基本要素是动作 (action), 动作可以由子动作组合得到, 由此构成了动作的分层结构。如军事领域的“完成阵地转移”, 医学领域的“完成一个阶段的化疗”, 日常生活领域的“吃晚饭”, 所描述的动作都对应了一种状态的转换, 而其中不涉及动作的组织方式、实施流程、具体的资源分配和管理。与动作的分层结构对应的, COA 问题也可以具有分层结构。很多情况下, 为了完成一定层次战略规划, COA 不需要在最底层的动作上展开。本文关注一类典型的“任务级 COA 问题”, 即任务级的动作随时间安排实现预定目标的问题, 重点是在多种偏好并存时如何获得高质量的 COA 方案。决策理论中, 偏好一般是指每个决策者在面对几个事件或结果时选择其中一个事件或结果的倾向性<sup>[3]</sup>。COA 的偏好问题也已经得到了研究者的很多关注<sup>[4-6]</sup>。偏好来自 COA 的参与者或者发起者, 虽然不构成强制约束力, 但对于 COA 生成方案的质量至关重要。有研究者指出, 规划方案的质量最优, 除了要满足所有的约束条件之外, 一定也是对各种软约束 (包括用户的偏好) 有了最大程度的满足<sup>[7]</sup>。如何描述规划方案的质量, 或者说, 如何定义对用户满意度的最大化, 因偏好表示方法不同而变化。可以知道, 在使用定量偏好描述时, 可以使用定量的效能函数作为衡量。除常见的权重表示偏好, 研究者后来发展出了区间数<sup>[8]</sup>、模糊数<sup>[9]</sup>等定量表示方式。关于偏好的研究指出<sup>[10]</sup>, 定性偏好比定量偏好更具有普遍性, 很多时候用户使用自然语言描述的复杂偏好关系, 定量偏好是不能胜任表达的; 随之而来的, 定性偏好的使用使得定量的效能函数不再有效。

任务级别 COA 问题与任务调度有相似之处, 都涉及到多方面的约束, 如资源约束、能量约束等<sup>[11]</sup>。偏好关系也是软约束之一。从满足尽可能多的限制条件、尽可能好地完成的角度来考虑, 任务级 COA 与任务调度问题类似, 同样具有约束满足问题 (CSP)<sup>[12]</sup> 的特征。但是任务级 COA 对时间资源的调度具有突出的需求, 何时开始一个动作, 何时结束一个动作, 如何合理安排动作之间的时序和因果关系。任务级 COA 的约束满足不能通过一个单一的效能函数来表述和分析。

已经有研究者对定性偏好相关的规划开展了研究。文献<sup>[13]</sup>从满意度的角度设计规划策略, 对每一个偏好设置偏好强度值, 对一个规划序列中的所有偏好进行运算生成衡量值, 通过该衡量值选择最优序列。这里探讨的偏好不包括时序相关的偏好。

文献<sup>[14]</sup>对时序软约束下的规划展开研究, 根据软约束内容判定两个动作之间的可能关系: 紧随、超前或不相关, 为软约束设置偏好程度参数, 根据以上因素在每个时间点上搜索选择可用动作, 随着时间点的推移完成规划。这两种研究存在共同点, 即依赖偏好或软约束的强度值作为判定依据, 在动态搜索过程中获得满意度或排除过约束。与本文最相关的是 Blom<sup>[6]</sup> 的研究。在存在多种偏好, 且偏好之间有冲突时, Blom 首先使用非单调推理技术排除其中的冲突, 获得最大一致性集合, 在该集合基础上再进行决策或规划, 决策或规划的搜索都是传统方式, 通过这种设计实现了“偏好解耦”, 即规划过程与偏好的表示方式或偏好内容无关, 这对于规划算法在不同问题领域的通用性意义重大。遗憾的是, Blom 并没有关注时序偏好。Blom 使用的逻辑工具为计算辩论。计算辩论技术是自 1990 年以来逐渐发展成熟并得到众多关注的逻辑工具, 其在人工智能中的应用见于文献<sup>[15]</sup>。

针对带有定性偏好的任务级 COA 规划问题, 本文首先将使用比较格式给出任务级 COA 中的约束和偏好定义, 基于比较格式的形式化, 给出 COA 方案的生成算法; 然后使用计算辩论技术消除约束/偏好集合中的不一致性, 给出偏好集合上的最大满意度定义, 在此基础上完成 COA 方案的生成。最大化任务的收益也不是任务级 COA 的主要目的, 它只需要保证高质量地完成预设目标, 不需要定量优化的能力, 因此获得的是一个可行解集合, 而非确定的单个解。本文的研究是传统任务调度问题和约束问题研究的有效补充, 同时也探讨了定性偏好关系的形式化表达方式和推理方式。

## 1 定性偏好表示与任务级 COA 规划

以任务级动作为基本元素的 COA 问题称为任务级 COA。下面首先给出任务级动作的定义及其形式化表示。

**定义 1** (任务级动作) 给定 COA 问题  $P$ , 满足以下条件的动作称为任务级动作:

- 1) 动作的结果是一种确定的状态。
- 2) 动作可以拆分为多个子动作或者元动作实现, 子动作集合及其排序方式可能有多种
- 3) 动作之间没有目标的重叠, 即: 动作  $A$  引起的状态变化, 没有可能是可以由动作  $B$  实现的。

**定义 2** (任务级动作的形式化表示) 动作由一个三元组  $(s, g, t)$  表示。其中,  $s$  是动作主体,  $g$  是动作受体,  $t$  是某种时间标记的集合, 可以是一个时

间点、时间间隔或某种时间标记。所有的动作集合构成三元组集合  $(s, g, t)$ 。

为了表达逻辑的完备性,元素  $s, g, t$  设置通配符  $\sim$ , 表示对应的  $s, g, t$  位置可以任意取值。形如  $(\sim, g, t)$  的表达式包含了多个元动作:  $(s_1, g, t)$ ,  $(s_2, g, t)$ ,  $(s_3, g, t)$  等。 $(s, \sim, t)$ ,  $(s, g, \sim)$  与此类似。任务级 COA 问题的定性偏好可以划分为两类。第1种是与时序无关的静态偏好,或者称为动作间的属性偏好,即:因为动作某种属性被偏好,即:用户倾向于选择动作  $A$  还是动作  $B$ 。第2种是动作间的时序或因果顺序偏好,即:用户偏好于动作  $A$  在动作  $B$  之前,还是动作  $B$  在动作  $A$  之前,本文从另一个角度表述为:倾向于先执行使用动作  $A$ , 还是先执行动作  $B$ 。下面给出了定性偏好的归一化表示。

**定义3** (定性偏好的归一化表示)定性偏好由三元组  $(S, G, T)$  中的两个元素和一个二元操作符 ( $>$  或  $\triangleright$ ) 组成,形式化表达如下:

$$(s_i, g_i, t_i) >_{s(g,t)} (s_j, g_j, t_j)$$

$$(s_i, g_i, t_i) \triangleright_{s(g,t)} (s_j, g_j, t_j)$$

式中:  $>$  为连接两个三元组的支配关系,给出了两个选项的静态比较;  $\triangleright$  为连接两个三元组的一个“在前面”的语义关系,给出了一种因果关系或者时间的比较。每个表达式只对三元组中的一个元素进行比较,  $>$  或  $\triangleright$  都可以配备有下标  $s, g$  或  $t$ , 表示引起比较关系的有效元素。

为了表达逻辑的完备性,为三元组定义一个“空”状态  $\Theta$ , 用以表征一个“不被偏好”的虚拟动作,意为任意一个非空动作均可以对它构成二元关系。如自然语言的偏好表述:“用户偏好于动作  $(s_1, g_1, t_1)$ ”, 没有特定的比较对象,可以表示为  $(s_1, g_1, t_1) >_s \Theta$ 。

偏好的表示方式多种多样。除了经典的定量表示之外,定性表示可以划分为 C-P 网络<sup>[16]</sup>、时序逻辑<sup>[4]</sup>、比较格式<sup>[17-18]</sup>等。定义3给出的是一种比较格式的表达式,秉承了 Blom 的理念:偏好本质上是一种二元关系,因而本文选择的是一种相对偏好表示。这是实现偏好解耦设计的关键。

通过引入  $\triangleright$  符号,时序或因果偏好的形式化表述可以与静态偏好统一到一个框架内;而空状态  $\Theta$  和任意状态  $(\sim, \sim, \sim)$  的存在,可以把非典型的二元关系也赋予一个二元比较关系方式。如,动作  $(\sim, \sim, \sim) \triangleright_{s(g,t)} (s_1, g_1, t_1)$  表示用户偏好于最后执行动作  $(s_1, g_1, t_1)$ 。

在展开后续推导之前,给出3个断言。

**断言1** 任务级层次的定量偏好可以转换到定

性偏好。

定量表示的偏好值,一般用某属性的被偏好程度表示,为  $0 \sim 1$  的某个正数,数字越大表示偏好程度越强。举例来说,  $p(A) = 0.3$  表示动作  $A$  的偏好值为  $0.3$ , 如果有  $p(B) = 0.6$ , 那么显然  $B$  比  $A$  更受偏好。在进行选择时,只要知道  $p(B) > p(A)$  这一事实即可,至于  $B$  的受偏好程度是  $0.6$  还是  $0.7$ , 都可以得到  $B > A$  这一结论,不影响决策者的判断。定性偏好恰好具有这种相对偏好的表达能力。特殊地,如果没有能够与动作  $A$  进行比较的动作,那么可以知道:  $A > \Theta$ 。

当然,定性偏好能够表达的并不局限于这种数值基础上的相对关系,比如时序偏好。从语义覆盖性的角度,定性偏好的表达是涵盖定量偏好的。

**断言2** 定性偏好表达不能替代定量表达。

定性表达的优势在于复杂语义的阐述,不能像传统的定量优化算法那样得到具有优化指标的唯一最优解,一般只能得到可行解集合。本文的思路是,定性偏好的推理过程服务于定量偏好。比如,给定一组偏好集合,首先使用定性表达确定它们内部是不是有冲突,即:这个偏好集合是不是能够共存的,如果可以,那么就继续进行数值规划,给出优化指标;否则就继续寻找下一组偏好,不在没有可能的偏好集合上浪费定量优化的计算资源。

**断言3** 任务级层次的偏好与软约束具有相同的结构。

软约束是指不具有强制力的约束关系,自然语言理解为“建议满足”而非“必需满足”,这与偏好的语义有相通之处。更进一步,在把偏好和约束并列考察时,约束也应该具有与偏好对等的表达方式,因此后文将对偏好和约束进行归一化表示。软约束与偏好的区别仅在于:偏好具有人的主观特性,而软约束可能来自事物本身的客观特征,文中不明确区分软约束和偏好,统称为偏好;而约束则是指有强制约束力的硬约束。

**定义4** (约束和偏好的归一化表示)约束和偏好可以统一为四元组表示:  $\langle R, u_i, u_j, k \rangle$ 。其中,  $R$  为二元关系  $>$  或  $\triangleright$ ,  $u_i, u_j$  为前述定义的三元组,  $k$  为  $0$  表示硬约束,  $k$  为正数表示偏好,数字越大,偏好程度越弱。

假设偏好之间,以及偏好和约束之间没有相互矛盾,那么可以执行算法1获得 COA 方案。

**算法1** 由约束/偏好集合生成 COA 方案的两阶段算法 2DOF-SLCOA。

输入: 静态约束/偏好集合  $Ex_{\sim}$ , 时序约束/偏

好集合  $Ex_{\triangleright}$

输出: 动作序列(COA 规划方案)

步骤1 从静态约束集合  $Ex_{\triangleright}$  中选择可用动作集合。

1) 基于  $Ex_{\triangleright}$  中的元素构建有向图  $H_{\triangleright}$ 。对于  $Ex_{\triangleright}$  中每个约束/偏好关系, 连接两个三元组的  $\triangleright$  关系对应导向图中的一个边, 相关的两个三元组则是图中的节点. 从不同约束/偏好关系中提出的同一个三元组对应同一个节点, 因此, 某一个节点可能会有边界指向多个节点, 也可能有多个节点的边界指向它。

2) 选择有向图  $H_{\triangleright}$  中的优势节点, 即: 每个支链的起始节点, 置于一个集合  $U_{\triangleright}$  中。

步骤2 结合时序约束, 得到调度方案。

1)  $U_{\triangleright}$  中的每一个元素  $u$  都被替换为一个时序约束关系  $u \triangleright \Theta$ , 构成时序约束/偏好关系集合  $U_{\triangleright}$ 。

2) 基于  $U_{\triangleright}$  和  $Ex_{\triangleright}$  生成有向图  $H_{\triangleright} \cdot U_{\triangleright}$  或  $Ex_{\triangleright}$  中的  $\triangleright$  优势关系被翻译成有向图的边, 而与之相关联的三元组被放置在节点位置。对等的三元组只作为一个节点出现。

3) 寻找  $H_{\triangleright}$  中的可用序列。从一个根节点出发, 追踪连续的链接关系形成一个节点序列。

4) 节点序列经解释或翻译得到动作序列。

## 2 非单调偏好的处理和最大满意度

对于一个进行 COA 规划的智能体来说, 只要约束和偏好之间没有任何的冲突, 偏好的作用与约束是相同的。然而, 经常出现的情况是, 约束/偏好集合中会出现一个以上的冲突, 此时就需要首先排除不一致性, 获得具有最大一致性的约束/偏好集合, 才能继续推理得到规划方案。冲突关系数目较多时, 约束或偏好之间相互的优势关系错综复杂, 简单的逻辑判断不能确定最大一致性。本文使用计算辩论<sup>[19]</sup>这一非单调推理工具来执行逻辑推理。

### 2.1 辩论基础

这里简要介绍了 Dung 的辩论框架与相关的定义, 读者可以参考文献<sup>[19]</sup>以获取更多关于辩论的内容。

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 辩论框架及其相关性质。

辩论框架通常定义为二元组  $AF = (A, R)$ , 其中  $A$  表示论据集,  $R$  是定义在论据集  $A$  上的二元(攻

击)关系, 即有  $R \subseteq A \times A$ 。

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

论据集  $S (S \subseteq A)$ , 若  $\forall a, b \in S, (a, b) \notin R$ , 称  $S$  是无冲突的(conflict-free)。论据集  $S$  是无冲突的, 满足  $\forall a \in R - (S) \rightarrow (\exists b \in S). bRa$ , 称  $S$  为可容许集(admissible set)。

$AF = (A, R)$  可以表示为有向图, 称为攻击图。如图 1 所示的攻击图中  $A = \{a, b, c, d\}$ ,  $R = \{(a, c), (b, a), (c, b), (c, a), (d, c)\}$ 。由此可知,  $R_+(c) = \{a, b\}$  及  $R_-(c) = \{a, d\}$ 。

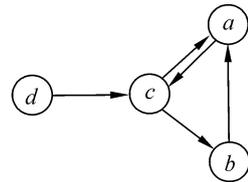


图1 辩论框架

Fig.1 Argumentation framework

根据 Dung 的扩展语义, 该实例中的首选扩展集合为  $S_{\text{pref}} = \{d, b\}$ 。该集合满足以下条件, 也就是首选扩展语义的定义:

- 1)  $S_{\text{pref}}$  中的论据相互之间不冲突;
- 2)  $S_{\text{pref}}$  中的论据不会被其他论据工具; 或者,
- 3) 如果论据  $x \in S_{\text{pref}}$  被另一个论据  $y \notin S_{\text{pref}}$  攻击, 那么必然有第 3 个论据  $z \in S_{\text{pref}}$  满足  $zRy$ 。

概括来讲, 辩论就是一个信念萃取的过程, 排除那些内含不一致性的选项, 达成一个最大一致性。在后续章节里, 将其直接用于逻辑推理, 不再涉及背后的逻辑过程。

### 2.2 论据定义

为了进行辩论, 首先是构造论据。每一个约束或偏好自然地形成一个论据。

定义5 任务级 COA 中的论据定义。一个约束或偏好相关的论据被定义一个 4 元结构:

$$(k^{\text{type}}, u^l, u^r, k^{\text{rank}})$$

式中:  $u^l, u^r$  分别为约束定义中的三元组。  $k^{\text{type}}$  表征了关系类型, 即  $\triangleright$  或  $\triangleright$ 。  $k^{\text{rank}}$  对应了  $\triangleright$  或  $\triangleright$  的上标, 即 0 或者一个表征软约束强度的正整数。

不包含通配符~的论据称为元论据。与此相对的,含有通配符~的论据可以通过把~对应的元素进行枚举,分解为多个元论据。两个元论据之间的攻击关系定义如下。

**定义 6** (论据间的攻击关系)给定两个元论据

$$a_i = (k_i^{\text{type}}, u_i^f, u_i^r, k_i^{\text{rank}}), a_j = (k_j^{\text{type}}, u_j^f, u_j^r, k_j^{\text{rank}})$$

它们之间的关系定义如下。

1) 两论据  $a_i$  和  $a_j$  是可比的,当且仅当  $k_i^{\text{type}} = k_j^{\text{type}}$ 。

2) 冲突关系:两论据被称为是冲突的,当且仅当它们是可比的,而且  $u_i^f = u_j^r, u_i^r = u_j^f$ 。

3) 攻击关系:论据  $a_i$  攻击  $a_j$ , 记为  $a_i Ra_j$ , 当且仅当  $a_i$  与  $a_j$  冲突, 并且  $k_i^{\text{rank}} \leq k_j^{\text{rank}}$ 。

### 2.3 任务级 COA 的辩论框架

接下来建立任务级 COA 问题的辩论框架。

**定义 7** 任务级 COA 辩论框架 TCOA-ARG。对于一个 TCOA-ARG 问题,其辩论框架为一个三元组  $\langle A, R, SL \rangle$ , 其中  $A$  为来自约束集合和偏好集合的论据,  $R$  为论据间的攻击关系,  $SL$  为获得辩论结果的辩论语义。

该框架与 Blom 的偏好解耦框架是一致的<sup>[6]</sup>, 区别仅在于,针对静态和时序约束存在两种二元关系,因此构造攻击关系时每种论据都是局限在自己的二元关系归类中。结果自然是每种攻击关系操作符都有一个独立的辩论结果。本文使用的辩论语义  $SL$  与 Blom<sup>[6]</sup> 相同,即  $SL$  为 Dung<sup>[19]</sup> 首选扩展语义。

**断言 4** TCOA-ARG 的辩论结果是一个静态(时序)约束的首选扩展集,也就是说,没有内部冲突的静态(时序)约束的最大集,记为

$$A = A_{>} \cup A_{\triangleright}$$

$$A_{>} \xrightarrow{SL} Ex_{>}$$

$$A_{\triangleright} \xrightarrow{SL} Ex_{\triangleright}$$

式中:  $A_{>}$  和  $A_{\triangleright}$  分别为基于  $>$  或  $\triangleright$  的论据子集,  $Ex_{>}$  和  $Ex_{\triangleright}$  分别为从  $A_{>}$  和  $A_{\triangleright}$  推导出来的首选扩展集合。

需要指出的是,每一个论据均对应了一个约束或偏好,表达了 2 个动作之间的优势关系。在辩论

中,一些偏好会被约束或其他一些更高优先级的偏好关系击败,因而被剔除,留在集合  $Ex_{>}$  和  $Ex_{\triangleright}$  中的就是最大兼容性的约束子集。

**断言 5** 论据的首选扩展集合给出的用户最大满意度集合。

### 2.4 主要结果

本节的主要结果总结在算法 2 和算法 3 中。

**算法 2** 基于辩论解决偏好不一致性。

**输入:**一个任务级 COA 问题,已建立约束/偏好集合。

**输出:**首选扩展集  $Ex_{>}$  与  $Ex_{\triangleright}$ 。

1) 构造论据集合,包含通配符的论据拆分为元论据。

2) 在论据间进行搜索,构造攻击关系。

3) 执行 DUNG 的首选辩论语义,获得首选扩展集合  $Ex_{>}$  和  $Ex_{\triangleright}$ 。

如果  $Ex_{>}$  或  $Ex_{\triangleright}$  是空集,那就意味着静态或时序约束中存在一些不能解决的冲突。因而不能获得调度方案。否则,就可以进行下一阶段的工作,引用算法 1 构造 COA 可行解。

**算法 3** 带定性约束/偏好的任务级 COA 规划 ARG-COA。

**输入:**一个任务级 COA 问题。

**输出:**COA 规划方案。

1) 建立定性约束/偏好集合。

2) 执行算法 2 获取一致性约束/偏好集合  $Ex_{>}$  与  $Ex_{\triangleright}$ 。

3) 执行算法 1 得到规划结果。

## 3 案例研究

文中以一种卫星的突发性任务观测为背景展开案例研究。卫星设计如目前正在运行的 LAPAN-TUBSAT<sup>[20]</sup>。卫星装有视频摄像机,提供实时目标监控,无板载存储。

研究次日有 2 颗卫星过境,兴趣目标有 3 个,每个卫星均有 2 个可操作时间窗口。表 1 给出了卫星的技术状态描述,表 2 给出了卫星的可见窗口描述。

表 1 卫星技术状态描述

Table 1 Description of satellites' technical status

卫星编号	可见弧段/min	分辨率	成像波段	与测控站距离
S <sub>1</sub>	15	2 048×2 048	红外/可见光	2 000
S <sub>2</sub>	8	762×576	可见光	1 000

表2 卫星可操控时间列表  
 Table 2 List of satellite's time widow for operation

卫星	操作时间窗口	
$S_1$	[15:40 15:55] am	[17:10 17:30] am
$S_2$	[9:40 9:55] am	[11:00 11:10] am

把观测任务视为 COA 问题,考察以下类型偏好。

- 1) 目标的价值有优先级,用户在白天要首先观察到优先级高的目标。
- 2) 根据地面测控站的要求,对同一个目标执行观测动作时,希望使用能提供较长可见弧段的卫星。
- 3) 考虑到卫星的能量储备,在进行观测时,优先考虑姿态机动小、耗费能量少的卫星。
- 4) 对同一目标成像,用户喜欢高分辨率的图像。
- 5) 对某目标成像时,用户希望首先得到低分辨率图像,再获得高分辨率图像,方便图像判读。
- 6) 从用户角度,监视运动目标具有更高优先级。
- 7) 从传输信道的稳定性来讲,希望使用距离测控站距离近的卫星进行观测。
- 8) 进行森林火灾目标的检测时,更希望使用配备红外相机的卫星。

下面根据以上规则给出各偏好的形式化表达。

目标优先级偏好:目标  $g_1$  的优先级最高,而目标  $g_2$  的优先级最低,因而有

$$(P1) (\sim, g_1, \sim) \triangleright_g (\sim, \#, \sim)$$

$$(P2) (\sim, \#, \sim) \triangleright_g (\sim, g_2, \sim)$$

可用测控弧段偏好:

$$(P3) (s_2, \sim, \sim) >_s (\neg s_2, \sim, \sim)$$

姿态机动能量偏好:

$$(P4) (s_1, g_3, \sim) >_g (s_1, g_1, \sim)$$

成像分辨率偏好:

$$(P5) (s_1, \sim, \sim) >_s (\neg s_1, \sim, \sim)$$

低、高分辨率切换顺序偏好:

$$(P6) (s_2, \sim, \sim) \triangleright_s (s_1, \sim, \sim)$$

移动目标监视偏好:

$$(P7) (\sim, g_1, \sim) >_g (\sim, \neg g_1, \sim)$$

信道传输质量偏好:

$$(P8) \Theta >_s (s_1, g_1, \sim)$$

输出帧率偏好:

$$(P9) (s_2, g_1, \sim) >_s (\neg s_2, g_1, \sim)$$

成像谱段偏好:目标  $g_2$  处于森林火险等级较高的区域,优先考虑使用红外相机观测,即

$$(P10) (s_2, g_2, \sim) >_s^0 (\neg s_2, \sim, \sim)$$

下面给出3个约束关系。据气象预报,上午10时前目标  $g_3$  被云层覆盖,不能被观测到。

$$(C1) (\sim, g_3, t_{10:00AM}) \triangleright_t^0 (\sim, g_3, \sim)$$

由于卫星自身的限制,次日下午5:00后卫星  $s_1$  不应该被用于目标  $g_3$  的观测,因为那时卫星会马上

进入阴影区,在下一个任务之前不能充分充电。

$$(C2) (s_1, g_3, \#) \triangleright_t^0 (s_1, g_3, t_{5:00PM})$$

现在可以综合 (P1)~(P10) 和 (C1)~(C2), 执行辩论并完成任务级 COA 方案。

从偏好构建论据时,还需要确定偏好自身的优先级。在本案例中,根据偏好的提出者来确定 Arank 这个值。卫星运营商、卫星设计方和用户提出的偏好关系将分别有 3、2 和 1 的优先级。当然,从约束构造的论据都具有 0 的优先级。

1) 论据构建

$$(P1): a_1 = [\triangleright, (\sim, g_1, \sim), (\sim, \#, \sim), 1]$$

$$(P2): a_2 = [\triangleright, (\sim, \#, \sim), (\sim, g_2, \sim), 1]$$

$$(P3): a_3 = [ >, (s_2, \sim, \sim), (\neg s_2, \sim, \sim), 3]$$

$$(P4): a_4 = [ >, (s_1, g_3, \sim), (s_1, g_1, \sim), 2]$$

$$(P5): a_5 = [ >, (s_1, \sim, \sim), (\neg s_1, \sim, \sim), 1]$$

$$(P6): a_6 = [\triangleright, (s_2, \#, \sim), (s_1, \#, \sim), 1]$$

$$(P7): a_7 = [ >, (\sim, g_1, \sim), (\sim, \neg g_1, \sim), 1]$$

$$(P8): a_8 = [ >, \Theta, (s_1, g_1, \sim), 1]$$

$$(P9): a_9 = [ >, (s_2, g_1, \sim), (\neg s_2, g_1, \sim), 1]$$

$$(P10): a_{12} = [ >, (s_2, g_2, \#), (\neg s_2, g_2, \#), 0]$$

$$(C1): a_{10} = [\triangleright, (\sim, \sim, t_{10:00AM}), (\sim, g_3, \neg t_{10:00AM}), 0]$$

$$(C2): a_{14} = [\triangleright, (s_1, g_3, \#), (s_1, g_3, t_{5:00PM}), 0]$$

2) 攻击关系构建

记

$$A_{>} = \{a_3, a_4, a_5, a_7, a_8, a_9, a_{12}\}$$

$$A_{\triangleright} = \{a_1, a_2, a_6, a_{10}, a_{14}\}$$

$A_{\triangleright}$  内部没有冲突,因此,  $Ex_{\triangleright} = A_{\triangleright}$ . 而对于  $A_{>}$  则需要执行一个辩论过程以获得  $Ex_{>}$ 。

论据  $a_5$  拆分为

$$a_5^1 = [ >, (s_1, \neg g_2, \sim), (\neg s_1, \neg g_2, \sim), 1]$$

$$a_5^2 = [ >, (s_1, g_2, \sim), (\neg s_1, g_2, \sim), 1]$$

$$a_5^3 = [ >, (s_1, \neg (g_1 | g_2), \sim),$$

$$(\neg s_1, \neg (g_1 | g_2), \sim), 1]$$

论据  $a_7$  拆分为

$$a_7^1 = [ >, (s_1, g_1, \sim), (s_1, \neg g_1, \sim), 1]$$

$$a_7^2 = [ >, (s_2, g_1, \sim), (s_2, \neg g_1, \sim), 1]$$

得到攻击关系如下:

$$a_5 Ra_3, a_7 Ra_4, a_{12} Ra_5^2, a_9 Ra_5^1, a_8 Ra_7^1$$

3) 辩论结果

依据攻击关系构成的导向图如图2所示。

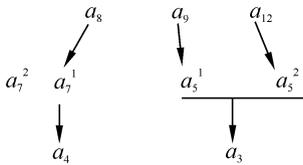


图 2 攻击关系导向图

Fig.2 Directed graph on attack relation

$A_{>}$  的首选扩展容易从图中获得:

$$EX_{>} = \{a_3, a_4, a_7^2, a_8, a_9, a_{12}\}$$

4) 构造有向图  $H_{>}$

$H_{>}$  绘制如图 3。

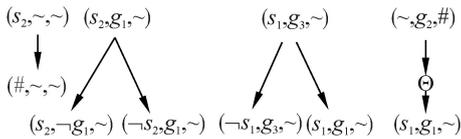


图 3 有向图  $H_{>}$

Fig.3 Directed graph  $H_{>}$

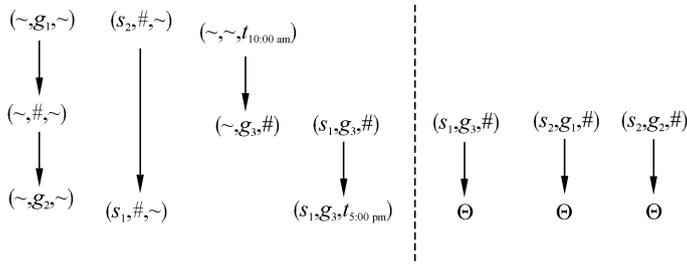


图 4 有向图  $H_{>}$

Fig.4 Directed graph  $H_{>}$

6) 构造 COA 方案

接下来的求解步骤是通过合并节点获得动作序列片段。对于虚线左侧的节点,尝试从虚线右侧找到匹配的节点,使得一个攻击链上相同位置的通配符可以由相同的确定元素替代。例如,  $(~, g_1, ~)$ ,  $(~, g_2, ~)$  处于同一个攻击链中,而在右侧,  $(s_2, g_1, ~)$ ,  $(s_2, g_2, ~)$  能够与他们匹配;而下一个攻击链  $(s_2, \#, ~)$ ,  $(s_1, \#, ~)$  则不能从右侧找到一组元素来匹配与之匹配。最后得到匹配结果如图 5 所示。

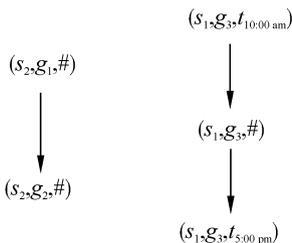


图 5 动作序列匹配结果图

Fig.5 Matched directed graph

选择  $H_{>}$  中的优势节点(攻击其他节点,而自己不被攻击的节点)构造集合  $U_{>}$

$$u_1 \quad u_2 \quad u_3 \quad u_4$$

$$(s_2, \sim, \sim)(s_2, g_1, \sim)(s_1, g_3, \sim)(\sim, g_2, \#)$$

$u_1, u_2, u_3, u_4$  等元素实际上均对应了多个三元组,相互之间有交叉重合的部分,这些三元组被认为是真正能够起作用的部分,定义为  $\tilde{U}_{>}$  :

$$\tilde{u}_2 \quad \tilde{u}_3 \quad \tilde{u}_4$$

$$(s_2, g_1, \#) \quad (s_1, g_3, \#) \quad (s_2, g_2, \#)$$

5) 构造有向图  $H_{\triangleright}$

基于  $EX_{\triangleright}$  和  $\tilde{U}_{>}$  构造  $H_{\triangleright}$  如图 4 所示。

这意味着,操作员应该使用卫星  $s_2$  在观测  $g_2$  之前观测  $g_1$ ;同时,应使用卫星  $s_1$  观测目标  $g_3$ ,时间在上午 10 时至下午 5 时之间。

结合操控时间窗口分布,得到了调度方案如下所示:

$$(s_2, g_1, [9:40, 9:50] \text{ am}) \rightarrow$$

$$(s_2, g_2, [11:00, 11:10] \text{ am}) \rightarrow$$

$$(s_1, g_3, [15:40, 15:55])$$

在解决过程中可以看到从模糊语义经过标定和校正逐步细化的过程。

下面将本文的 ARG-COA 算法与文献<sup>[21]</sup>中的 WSI 数值优化算法进行对比。

基于本文的背景案例生成 3 个测试集,分别包含不同数目的约束和偏好,如表 3 所示。由于 WSI 算法不支持时序关系求解,因此偏好只设定静态偏好。使用 WSI 算法时,把偏好也作为约束统一求解,以无冲突约束和偏好的强度值之和为指标函数值,指标大于 0 的约束/偏好集合均为可行解。在同样的处理平台上进行对比计算,结果如表 3 所示。

表3 ARG-COA 和 WSI 算法时的数值仿真对比

Table 3 Comparing the simulating results of ARG-COA and WSI

	测试集 1	测试集 2	测试集 3
约束数目	3	5	9
偏好数目	10	12	16
冲突数目	6	9	12
可行解数目	4	2	1
WSI 获得可行解的时间/s	0.9,1.4	1.6,1.8	3.1
ARG-COA 获得可行解的时间/s	1.1	1.2	1.3
WSI 对约束/偏好的遍历次数	149	403	1 413
ARG-COA 对约束/偏好的遍历次数	91	153	325

仿真计算结果表明:

1) 随着约束/偏好数目的增加, 获得可行解的时间越来越长, 但是 WSI 算法的时间增量更大。

2) WSI 算法比 ARG-COA 算法对约束/偏好进行遍历次数要多, 而且随着冲突数目的增加, 两者的差距越来越大。

3) WSI 算法能够更早获得可行解, 但是所有的可行解是要依次获得的, 这是由搜索过程决定的; 而 ARG-COA 算法则是同步获得全部可行解的, 花费较少时间。

前文指出, 定性偏好也并不能替代定量偏好, 同理, 基于辩论的算法并不能替代定量优化算法。ARG-COA 算法所获得的可行解, 都是用户满意度意义下的最优解。在算法 1 步骤 2 的 3) 子步骤, 可能会在一些节点处出现分支, 此时任何一个分支都对应一个可行解, ARG-COA 算法并不能从中选出唯一的最优解, 而这种“优中选优”是 WSI 优化算法可以胜任的。因此, 两种算法的相互结合可达到互补的效果: 首先使用辩论算法排除确实不可能有解的偏好集合, 然后再应用定量优化算法完成 COA 规划。这将是后续研究关注的重点。

计算辩论工具不但能够给出推理结论, 还能够把推理过程展现给用户, 因而进行 COA 的软件智能体可以向用户表明, 哪一个偏好被排除了, 或者由于用户过分强调哪一个偏好导致 COA 无法求解。由于用户往往不是专业人员, 对领域 COA 认识存在不足, 提出的偏好很可能不尽合理。当辩论过程给出反馈意见时, 用户可以根据自己的需求进行调整。如案例中用户提出的某种偏好优先级默认为 1, 在发生冲突时可能会排除掉一些优先级为 2, 或 3 的关系, 这可能是用户意料之外的, 用户可以把自己偏好的优先级更改为 2 或 3, 重新进行推理, 从而允许某些动作。这种人机交互式的推理将是未来研究方向之一。

## 4 结束语

本文对任务级的 COA 问题开展研究, 针对其中的定性偏好进行了建模, 闭环完成了 COA 方案生成算法设计, 以及在偏好与约束关系存在不一致性时如何通过非单调推理获得最大一致性, 最后形成了任务级 COA 问题求解的两阶段算法。辩论工具的使用, 一方面实现了偏好解耦, 算法的设计具有了领域无关的推广能力; 另一个潜在好处是能够提供用户交互参与规划过程的可能性。本文的研究结果是以定量优化为主的传统任务规划研究的有效补充。将来的研究将着力于与传统定量规划算法的结合, 并逐步把任务级 COA 向动作级 COA 扩展, 形成用户满意度和资源约束满足均达到最优、具有用户交互能力的 COA 规划方法。

## 参考文献:

- [1] FERGUSON R W, RASCH R A, TURMEL W, et al. Qualitative spatial interpretation of course-of-action diagrams [C] // Proceedings of the National Conference on Artificial Intelligence. AAAI Press, 2000: 1119-1120.
- [2] FORBUS K D, USHER J, CHAPMAN V. Sketching for military courses of action diagrams [C] // Proceedings of the 8th International Conference on Intelligent user Interfaces. ACM, 2003: 61-68.
- [3] SLOVIC P. The construction of preference [J]. American Psychologist, 1995, 50(5): 364-371.
- [4] BIENVENU M, MCILRAITH S. Qualitative dynamical preferences in the situation calculus [C] // Multidisciplinary IJ-CAI-05 Workshop on Advances in Preference Handling. 2005: 30-35.
- [5] BIENVENU M, FRITZ C, MCILRAITH S A. Planning with qualitative temporal preferences [C] // Proceedings of the 10th International Conference on Principles of Knowledge Representation and Reasoning (KR). Lake District, UK, 2006: 134-144.

- [6] BLOM M. Arguments and actions: decoupling preference and planning through argumentation [D]. Melbourne: University of Melbourne, 2011: 15-60.
- [7] 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.
- [8] 吴江, 黄登仕. 多属性决策中区间数偏好信息的一致化方法[J]. 系统工程理论方法应用, 2003, 12(4): 359-362.  
The uniform methods for interval number preference information in multi-attribute decision making[J]. Systems Engineering- Theory Methodology Applications, 2003, 12(4): 359-362
- [9] 张凤华. 模糊决策中决策偏好的情景依赖性[D]. 重庆: 西南大学, 2010: 30-55.  
Zhang Fenghua. The scenario-dependent of decision preference in fuzzy decision[D]. Chongqing: Southwest University. 2010: 30-55.
- [10] BRAFMAN R, DOMSHLAK C. Preference handling—an introductory tutorial[J]. AI Magazine, 2009, 30(1): 58-95.
- [11] MARINELLIA F, NOCELLAB S, ROSSIB F, et al. A Lagrangian heuristic for satellite range scheduling with resource constraints[J]. Computers & Operations Research, 2011, 38(11): 1572-1583.
- [12] KNIGHT R, SMITH B. Optimally solving nadir observation scheduling problems[C]//Proceedings of the 8th International Symposium on Artificial Intelligence, Robotics and Automation in Space(i-SAIRAS2005). Munich, Germany: 2005: 33-41.
- [13] GIUNCHIGLIA E, MARATEA M. Planning as satisfiability with preferences [C]//National Conference on Artificial Intelligence. Boston: MIT Press, 2007, 22(2): 987-992.
- [14] BADALONI S, FALDA M, GIACOMIN M. Solving temporal over-constrained problems using fuzzy techniques [J]. Journal of Intelligent and Fuzzy Systems, 2007, 18(2): 255-265.
- [15] BENCH-CAPON T J M, DUNNE P E. Argumentation in artificial intelligence[J]. Artificial Intelligence, 2007, 171(10-15): 619-641.
- [16] BOUTILIER C, BRAFMAN R I, DOMSHLAK C, et al. CP-nets: a tool for representing and reasoning with conditional ceteris paribus preference statements[J]. Journal of Artificial Intelligence Research, 2004, 21: 135-191.
- [17] DELGRANDE J P, SCHAUB T, TOMPITS H. Domain-specific preferences for causal reasoning and planning [C]// Proceedings of the Ninth International Conference on Principles of Knowledge Representation and Reasoning (KR.2004). AAAI Press, 2004: 673-682.
- [18] 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.
- [19] 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.
- [20] TRIHARJANTO R H, HASBI W, WIDIPAMINTO A, et al. LAPAN-TUBSAT: micro-satellite platform for surveillance & remote sensing [C]//Proceedings of the 4S Symposium: Small Satellites, Systems and Services. La Rochelle, France: 2004: 66-70.
- [21] 贺川, 朱晓敏, 邱涤珊. 面向应急成像观测任务的多星协同调度方法[J]. 系统工程与电子技术, 2012, 34(4): 726-731.  
HE Chuan, ZHU Xiaomin, QIU Dishan. Cooperative scheduling method of multi-satellites for imaging reconnaissance in emergency condition[J]. Systems Engineering and Electronics, 2012, 34(4): 726-731.

#### 作者简介:



王炎娟,女,1984年生,博士研究生,主要研究方向为人工智能、信息系统与智能决策。



姚莉,女,1965年生,教授,博士生导师,主要研究方向为人工智能、知识管理、信息系统与智能决策、计算辩论技术。



刘斌,男,1989年生,博士研究生,主要研究方向为人工智能、信息系统与智能决策。