

不确定层次任务网络规划研究综述

王红卫¹ 刘典¹ 赵鹏¹ 祁超¹ 陈曦¹

摘要 层次任务网络 (Hierarchical task network, HTN) 规划作为一项重要的智能规划技术被广泛应用于实际规划问题中, 传统的 HTN 规划无法处理不确定规划问题. 然而, 现实世界不可避免地存在无法确定或无法预测的信息, 这使许多学者开始关注不确定规划问题, 不确定 HTN 规划研究也成为 HTN 规划研究的前沿. 本文从 HTN 规划过程出发分析了不确定 HTN 规划问题中涉及的三类不确定, 即状态不确定、动作效果不确定和任务分解不确定; 总结了系统状态、动作效果和任务分解等不确定需要扩展确定性 HTN 规划模型的工作, 以此对现有不确定 HTN 规划的研究工作加以梳理和归类; 最后, 对不确定 HTN 规划研究中仍需要解决的问题和未来的研究方向作了进一步展望.

关键词 智能规划, 不确定规划, HTN 规划, 不确定 HTN 规划

引用格式 王红卫, 刘典, 赵鹏, 祁超, 陈曦. 不确定层次任务网络规划研究综述. 自动化学报, 2016, 42(5): 655–667

DOI 10.16383/j.aas.2016.c150198

Review on Hierarchical Task Network Planning under Uncertainty

WANG Hong-Wei¹ LIU Dian¹ ZHAO Peng¹ QI Chao¹ CHEN Xi¹

Abstract As an important automated planning technique, hierarchical task network (HTN) planning has been widely used in practical planning problems, but traditional HTN planning cannot deal with a planning problem with uncertainty. However, there inevitably exist uncertain or unpredictable information in the real world. As a result, many scholars began to focus on planning under uncertainty, and HTN planning under uncertainty has become the forefront research of HTN planning. In this paper, uncertainties of state, action effects and task decomposition are analysed, followed by a summary of the expansion of HTN planning model for treating these three types of uncertainties. Also, existing research works about HTN planning under uncertainty are reviewed and categorized, based on which some unsolved problems in HTN planning with uncertainty and future research directions are brought forward.

Key words Automated planning, uncertain planning, hierarchical task network (HTN) planning, HTN planning under uncertainty

Citation Wang Hong-Wei, Liu Dian, Zhao Peng, Qi Chao, Chen Xi. Review on hierarchical task network planning under uncertainty. *Acta Automatica Sinica*, 2016, 42(5): 655–667

智能规划^[1] 是人工智能研究的一个重要领域, 已经得到国内外学者的广泛关注, 产生了许多规划技术. 层次任务网络 (Hierarchical task network, HTN) 规划是众多规划技术中的一种, 因其规划过程与人类求解复杂问题的思维过程类似而被广泛应用于实际问题. HTN 规划的基本思想是利用领域知识递归地将复杂抽象的任务分解成越来越小的子任务, 直到出现的任务都可以通过执行规划动作就

能完成为止. 它最初由 Sacerdoti^[2] 于 1975 年的研究中提出, 但直到 20 世纪 90 年代人们才对它有了充分的理论认识, Yang^[3] 和 Kambhampati 等^[4] 对 HTN 规划的理论模型进行了研究, Erol 等^[5] 提出了 HTN 规划的形式化模型, 进行了复杂性分析^[6].

自 HTN 规划方法提出以来, 不断地有 HTN 规划器问世. 最早地, NOAH (Nets of action hierarchies) 规划器于 1975 年与 HTN 规划思想被一同提出, 随后 Nonlin^[7] 规划器于 1977 年被提出. 其后出现的规划器包括 SIPE (System for interactive planning and execution)^[8] (1984) 及其继承者 SIPE-2^[9] (1990)、O-Plan^[10] (1991) 及其继承者 O-Plan2^[11] (1994)、第一个被证明是可靠且完备的规划系统 UMCP (Universal method-composition planner)^[6] (1995)、SHOP (Simple hierarchical ordered planner)^[12] (1999) 及其继承者 SHOP2^[13] (2003)、SIADEx^[14] (2005). 文献 [15] 对上述

收稿日期 2015-04-20 录用日期 2016-02-26
Manuscript received April 20, 2015; accepted February 26, 2016
国家自然科学基金 (71371079), 国家杰出青年基金 (71125001), 安全预警与应急联动技术湖北省协同创新中心资助
Supported by National Natural Science Foundation of China (71371079), National Science Fund for Distinguished Young Scholars (71125001), and Hubei Collaborative Innovation Centre for Early Warning and Emergency Response Technology
本文责任编辑 赵千川
Recommended by Associate Editor ZHAO Qian-Chuan
1. 华中科技大学系统工程研究所图像信息处理与智能控制教育部重点实验室 武汉 430074
1. State Key Laboratory of Education Ministry for Image Processing and Intelligence Control, Systems Engineering Institute, Huazhong University of Science and Technology, Wuhan 430074

规划器的领域知识形式化体系、表达能力、问题求解能力、性能和适用性从理论和实践两方面进行了分析和比较。

因为 HTN 规划方法是依据问题求解“处方”进行分层规划,这与领域专家求解规划问题的思考方式非常类似,使得 HTN 规划技术在实践中得到广泛应用^[16-17],其应用领域包括生产线调度^[18]、危机管理和后勤规划^[10-11, 19]、航天器规划和调度^[20-21]、装备配置^[22]、加工流程规划^[23]、紧急疏散规划^[24]、桥牌^[25]、机器人^[26-27]、网络服务组合^[28]、森林灭火^[14]、医疗护理^[29]和应急响应决策^[30]等。在这些实际的规划问题中,客观世界系统被描述为确定性模型而符合确定性假设和完全可观性假设,因而规划解执行时系统状态的演化过程与规划时预测的系统状态变化过程相一致,保证了规划解的正确性。然而,现实世界不可避免地存在无法确定或无法预测的信息,这迫使人们开始研究带有不确定的规划问题。

现有文献中已有很多关于不确定规划问题的研究工作,而与 HTN 规划相关的不确定规划研究则相对较少。本文考虑对不确定 HTN 规划的研究现状进行分析和总结,以发现现有研究中不足的地方作为未来研究方向,希望可以推进 HTN 规划在不确定规划问题中的应用。本文首先阐述了不确定 HTN 规划问题中三类典型的不确定,即系统状态不确定、动作效果不确定和任务分解不确定;其次,在 HTN 规划概念模型的基础上,针对上述三类不确定总结了 9 种扩展模型,并以此对现有不确定 HTN 规划研究工作进行了梳理;最后,在分析现有研究工作不足的基础上提出了未来研究的建议。

1 不确定 HTN 规划问题

HTN 规划以给定的初始状态、任务目标和领域知识为输入,输出动作方案。初始状态描述了规划初始时刻系统的状态;任务目标是描述了应完成的任务集合及其逻辑关系的初始任务网络;领域知识包括操作符集合和方法集合,操作符描述了完成原子任务的动作执行的前提条件及其对系统状态产生的执行效果,方法以“处方”的形式描述了如何将非原子任务分解成更小子任务,包括分解前提条件和子任务集合。规划过程即在初始状态下,根据方法集合,对初始任务网络进行分解,直到任务网络中只含有可由实例化操作符(即动作)实现的原子任务为止,具体过程如图 1 所示。

HTN 规划中,动作用来完成原子任务并改变当前状态,而任务分解决定了进入任务网络的原子任务,状态决定了当前可用的动作和可以进行的任务分解。确定性 HTN 规划问题往往假设动作有唯一

确定的效果,使当前状态朝着确定的方向演化,同时系统状态可以完全确定,这样系统可以沿着一条完全可以预测的路径演化。然而,现实问题中,规划可能因为无法获得完整的状态信息而面临系统状态不确定,还可能因为系统中不可控的因素而面临动作效果不确定。这两方面的不确定构成基本的不确定规划问题,而利用 HTN 规划方法求解不确定规划问题会带来新的不确定,因为 HTN 规划依赖领域知识进行层次任务分解,领域知识中指导任务分解的方法可能存在不确定。为此,我们称不确定 HTN 规划问题是面临系统状态和(或)动作效果不确定,还需要考虑任务分解不确定的规划问题,并且这三类不确定之间会相互作用。动作效果不确定导致动作执行后的状态不确定,而不确定的状态会导致任务分解得到的子任务不确定,不同的子任务则会产生不同的可选动作集合。下面分别介绍上述三类不确定。

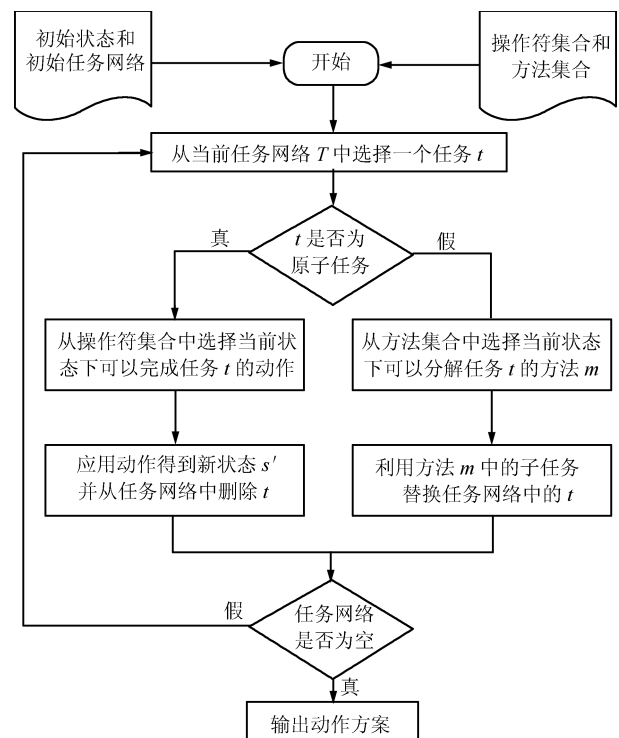


图 1 HTN 规划过程示意图

Fig. 1 An illustration of HTN planning

1) 系统状态不确定

系统状态描述了规划智能体所关注的环境的全部事实。规划智能体是可以设定目标和当前环境利用规划器进行动作规划并通过执行器执行动作对环境产生作用的东西,而规划器是规划算法的程序实现。根据规划智能体可以获得的系统状态信息的完整程度,将系统状态的可观性分为完全可观、部分可观和完全不可观。具体地,如果规划智能体在每个观察时刻都能获得系统全部的状态信息,那

么它可以根据获取的信息唯一确定系统所处的状态, 这种情况被称为系统状态完全可观. 如果规划智能体在每个观察时刻仅获得系统的部分状态信息, 那么它可以根据获得的信息确定系统处于哪些可能的状态但无法区分这些状态, 这种情况被称为系统状态部分可观. 如果规划智能体任何时刻都无法获得系统的状态信息, 那么它只知道初始时处于几个可能的状态之一, 而每个动作会因此导致几个可能的后继状态之一, 这种情况被称为系统状态完全不可观. 系统状态部分可观或完全不可观时, 规划智能体无法确定系统当前所处的状态, 这被称为系统状态不确定. 此时, 规划智能体只知道系统处于几个可能的状态之一, 这几个可能的状态构成的状态集合被称为信念状态^[31], 用来表示规划智能体对当前可能处于的实际状态的信念.

2) 动作效果不确定

动作的执行导致系统状态发生转移, 根据动作执行后状态的转移是否朝着确定的方向将动作效果分为确定和不确定两种情况. 如果动作执行后只有唯一的结果, 那么动作执行后状态朝着唯一确定的方向的转移, 则称动作具有确定的效果. 如果动作执行后有两个或更多的效果, 那么动作执行后状态的转移有两种或更多可能, 则称动作具有不确定的效果.

以上两类不确定的不同组合构成不同的不确定规划问题. 表 1 给出了 5 类不确定规划问题, 同时指明了各类不确定规划问题考虑的不确定类型, 还列举了一些最新的求解各类不确定规划问题的 HTN 规划器和非 HTN 规划器¹. 特别指出, P2 是 P3 的一种简化, 但两者的本质在于系统状态部分可观, 而现有文献研究通常假设动作效果确定, 认为问题 P2 的求解方法可以被扩展以处理动作效果不确定的情况. 问题 P4 和 P5 的情况类似.

3) 任务分解不确定

HTN 规划根据当前状态选择前提条件被满足的方法分解非原子任务, 并利用方法中子任务替代非原子任务在任务网络中的位置. 当前状态下, 一个非原子任务可以同时有多个前提条件被满足的方法对其进行分解, 但规划时只能不确定地选择其中之一, 而不同的方法将导致后续任务网络不同并影响最终的动作方案. 如果考虑这种方法选择不确定的随机性, 以及动作的代价和状态的奖励, 不同的方法选择会带来动作方案期望效用的不同. 因此, 在随机环境下, 不确定 HTN 规划问题需要考虑这种因方法选择不确定带来的任务分解不确定.

正因为 HTN 规划方法能够利用领域知识进行任务分解, 使规划朝着更有可能完成任务的方向搜索, 很好地减少了需要搜索的状态和路径. HTN 规划的这种优势使其能在一定程度上克服不确定带来的状态空间和搜索路径爆炸性增长的困难, 是一种值得研究的不确定规划方法.

2 HTN 规划模型的扩展

为了下文说明的方便, 下面先给出确定性 HTN 规划的概念模型^[1, 32-33].

一个 HTN 规划问题是一个 3 元组 $P = \langle D, s_0, tn_0 \rangle$, 其中 s_0 是初始状态, tn_0 是初始任务网络, D 是一个 HTN 规划领域.

一个任务网络是一个 2 元组 $tn = \langle T, C \rangle$, 其中 T 表示任务集合, C 表示 T 中任务之间的约束关系. 任务有原子任务和复合任务之分, 原子任务可由动作直接完成, 复合任务需要分解为更细致的子任务才能完成. 如果一个任务网络只含有原子任务则称它为原子任务网络.

一个 HTN 规划领域是一个 4 元组 $D = \langle S, O, M, \gamma \rangle$, 其中, S 是一个有限的状态集; O 是一个有限的操作符集, 操作符与经典规划相同, 可表示为三元组 $\langle head(o), pre(o), effect(o) \rangle$; M 是一个

表 1 不确定规划问题分类及代表性规划器

Table 1 Classification of planning problem with uncertainty and representative planner

不确定 规划问题	动作效果		状态可观性			规划器	
	确定	不确定	完全 可观	部分 可观	完全 不可观	非 HTN 规划器	HTN 规划器
P1		√	√			GRENDL ^[34] 、 NDP2 ^[35] 、PROST ^[36]	ND-SHOP2、YoYo、 LRTDP ^{SHOP2}
P2	√			√		PO-PRP ^[37] 、HCP ^[38] 、 CORPP ^[39]	Cond-SHOP2
P3		√		√			C-SHOP、PC-SHOP
P4	√				√		-
P5		√			√	RBPP ^[40] 、GC[LAMA] ^[41]	-

¹文献 [42-43] 中列举了一些早期求解各类不确定规划问题的规划器.

有限的方法集, 一个方法 $m \in M$ 是一个 2 元组 $m = \langle head(m), tn(m) \rangle$, 其中 $head(m)$ 表示方法可以求解的复合任务, $tn(m) = \langle T_m, C_m \rangle$ 表示方法 m 分解复合任务 $head(m)$ 得到任务网络; $\gamma : S \times A \rightarrow S$ 是状态转移函数, A 表示动作集合, 由操作符的实例化组成, $\gamma(s, a)$ 表示在状态 s 下应用动作 a 的后继状态.

一个 HTN 规划解是一个动作序列 $\langle a_0, a_1, \dots, a_n \rangle$, 满足: 1) 对应的任务序列 $\langle t_0, t_1, \dots, t_n \rangle$ 与一个原子任务网络 tn 一致, tn 可由初始任务网络 tn_0 分解得到; 2) $\gamma(s_i, a_i) = s_{i+1}$, $\gamma(s_{i+1}, a_{i+1}) = s_{i+2}, i = 0, \dots, n-1$. 因此, HTN 规划问题的求解是一个任务分解过程, 不断利用方法分解任务网络中的复合任务直到得到一个原子任务网络.

针对第 1 节的三类不确定, 我们需要在上述确定性 HTN 规划模型上进行不同的扩展²; 而不同的不确定表示方式也会造成规划问题形式及求解的不同. 现有的不确定规划研究中, 不确定的表示有两种方式: 逻辑表示和概率表示^[42]. 下面将从动作效果、系统状态和任务分解等 3 个方面讨论 HTN 规划中不确定及其在不同表示方式下的扩展.

2.1 动作效果不确定

确定性 HTN 规划模型通过操作符的效果文字 $effect(o)$ 来表示唯一的动作效果, 而为了表示不确定的动作效果, 必须对 $effect(o)$ 进行扩展以表示多个互斥的动作效果. 逻辑表示下, 扩展需要列举互斥的效果, 而在概率表示下, 扩展还需要给出每个效果的概率. 另一方面, 确定性 HTN 规划以完成任务为目标, 将复合任务逐步分解为原子任务并通过动作执行来完成, 而动作效果不确定意味着动作执行后会得到多个结果状态, 系统可以朝不同的方向演化, 无法保证任务完成时系统处于预想的最终状态. 因此, 对确定性 HTN 规划问题中的初始任务网络进行扩展, 明确初始任务网络完成时系统应该处于的目的状态, 作为规划是否完成的判断条件.

此外, 传统 HTN 规划解中, 前一个动作执行的结果状态一定满足后一个动作的前提条件, 而一个具有不确定动作效果的动作执行后有多多个结果状态, 无法都满足后一个动作的前提条件, 这将导致规划解无法继续执行. 因此, 需要扩展传统的 HTN 规划解以明确不同不确定动作效果出现时应该执行的后继动作. 再者, 规划解除了形式上的扩展, 在逻辑表示和概率表示下还需要满足不同的条件. 逻辑表示下, 规划解需要保证从初始状态到目标状态的可达性; 而在概率表示下, 规划解需要保证其成功概率或

具有最大期望效用. 基于上述分析, 提出下面 3 种扩展:

扩展 1 (不确定动作效果 (逻辑表示)). 将确定的动作效果 $effect(o)$ 扩展为 $\langle effect_1(o), effect_2(o), \dots, effect_n(o) \rangle$, 每个 $effect_i(o)$ 表示一种可能的动作效果. 相应地, 状态转移函数扩展为 $\gamma : S \times A \rightarrow 2^S$, $\gamma(s, a)$ 表示在状态 s 下应用动作 a 的后继状态集合.

扩展 1' (不确定动作效果 (概率表示)). 将确定的动作效果 $effect(o)$ 扩展为 $\langle (effect_1(o), p_1), (effect_2(o), p_2), \dots, (effect_n(o), p_n) \rangle$, 其中 p_i 表示不确定效果 $effect_i(o)$ 的概率, 且满足 $\sum_i p_i = 1$. 状态转移函数扩展为概率分布 $P : S \times A \times S \rightarrow [0, 1]$, $P_a(s'|s)$ 表示在状态 s 下应用动作 a 得到的结果状态是 s' 的概率, 利用 $\gamma(s, a) = \{s' | P_a(s'|s) > 0\}$ 表示在状态 s 下应用动作 a 的结果状态集合.

扩展 2 (目标状态). 将确定性 HTN 规划问题 $P = \langle D, s_0, tn_0 \rangle$ 扩展为 $P' = \langle D, s_0, tn_0, G \rangle$, 其中 G 表示目标状态集合. 只有将初始任务网络 tn_0 分解得到原子任务网络且系统状态满足目标状态, 规划问题才算求解成功.

扩展 3 (状态规划解 (逻辑表示)). 状态策略是一个表示从状态到动作映射的策略 $\pi : S \rightarrow A$. 策略 π 在规划领域 D 中从初始状态 s_0 出发, 所有可能的执行情况表示为 π 的执行结构 Σ_π ^[43]. 策略 π 是规划解的条件是其执行结构 Σ_π 中存在从 s_0 到 $g \in G$ 的路径, 根据执行结构中 s_0 到 $g \in G$ 的可达情况将规划解划分为弱规划解、强规划解和强循环规划解^[43].

扩展 3' (状态规划解 (概率表示)). 针对有奖励、代价函数和没有奖励、代价函数两种情况, 定义概率表示下的状态规划解如下:

1) 设有状态策略 $\pi : S \rightarrow A$, $\pi(s)$ 表示在状态 s 下选择执行的动作; 奖励函数 $R : S \rightarrow R$, $R(s)$ 表示达到状态 s 获得的奖励值; 代价函数 $C : S \times A \rightarrow R$, $C(s, a)$ 表示在状态 s 下应用动作 a 的代价值. 令 $E_\pi(s)$ 表示在策略 π 下状态 s 的期望效能值, 满足 $E_\pi(s) = R(s) - C(s, \pi(s)) + \beta \sum_{s' \in S} P_{\pi(s)}(s'|s) E_\pi(s')$, 其中 β 是折扣因子. 如果不存在其他策略 π' 使得 $E_\pi(s_0) < E_{\pi'}(s_0)$, 则称策略 π 是规划解.

2) 设有状态策略 $\pi : S \rightarrow A$, 存在从初始状态 s_0 达到目标状态 g 的路径 $path = \langle s_0, \pi(s_0), s_1, \pi(s_1), \dots, s_k, \pi(s_k), g \rangle$, 令 $P_{path} = \prod_{i=0}^k P_{\pi(s_i)}(s_{i+1}|s_i)$, $s_{k+1} = g$ 表示路径 $path$ 的概率, 则 $P_\pi = \sum_i P_{path_i}$ 表示策略 π 的成功概率,

²这里的扩展指一些基本的概念扩展, 具体到不同求解方法上可能不同的实现方式和其他扩展.

其中 $path_i$ 表示策略 π 的执行结构 Σ_π 中从初始状态可到达目标状态的执行路径. 如果不存在其他策略 π' 使得 $P_\pi < P_{\pi'}$, 则称策略 π 是规划解.

2.2 系统状态不确定

系统状态不完全可观导致初始状态不确定, 且无论动作效果确定与否, 其执行后的结果状态也不确定. 为此, 定义信念状态扩展表示不确定的状态集合, 并且信念状态在逻辑表示和概率表示下有不同形式. 相应地, 将系统状态完全可观的规划问题的初始状态和目标状态替换为初始信念状态和目标信念状态. 此时, 规划问题的解是信念状态与动作之间的对应关系, 而非状态与动作间的对应关系, 为此, 定义信念状态规划解扩展表示信念状态与动作之间的对应关系, 并指明规划解在逻辑表示和概率表示下分别应该满足的条件.

此外, 系统状态不完全可观有部分可观和完全不可观两种情况. 对于系统状态部分可观的规划问题, 规划智能体可以根据动作执行后的观察值以及观察值与状态间的对应关系推断系统当前的信念状态. 规划智能体通过感知获得系统的观察值, 而感知有自动感知和主动感知两种方式^[44]. 自动感知意味着每次动作执行后, 智能体即可获取所有可利用的感知信息; 主动感知意味着智能体在需要获得感知信息时只有执行特定的感知动作才能获得. 在主动感知下, 动作被区分为执行动作和感知动作^[38], 前者会改变系统状态, 后者只获取感知信息而不改变系统状态. 为此, 定义感知动作扩展表示规划智能体获得观察值的行为, 同时定义观察值函数表示观察值与状态之间的对应关系. 基于上述分析, 提出下面 4 种扩展:

扩展 4 (信念状态 (逻辑表示)). 将规划领域扩展为 $D = \langle S, B, O, M, \gamma \rangle$, 其中 $B \subseteq 2^S$ 是有限的信念状态集, $s \in b$ 表示信念状态 b 中状态 s . 相应地, 不确定 HTN 规划问题需要扩展为 $P' = \langle D, b_0, tn_0, b_G \rangle$, 其中 b_0 表示初始信念状态, $b_G = \{s | s \in G\}$ 表示目标信念状态.

扩展 4' (信念状态 (概率表示)). 将规划领域扩展为 $D = \langle S, B, O, M, \gamma \rangle$, 其中 $B \subseteq 2^S$ 表示信念状态集合, $b(s)$ 表示状态 s 属于信念状态 b 的概率, 满足 $b(s) \in [0, 1]$ 且 $\sum_{s \in S} b(s) = 1$, 则在信念状态 b 下应用动作 a 得到状态 s' 的概率为 $b_a(s') = \sum_{s \in S} P_a(s'|s)b(s)$. 不确定 HTN 规划问题也扩展为 $P' = \langle D, b_0, tn_0, b_G \rangle$.

扩展 5 (感知动作). 1) 对于自动感知, 感知动作利用扩展动作效果 $\langle effect(o), obs(o) \rangle$ 表示, 其中 $obs(o)$ 表示应用操作符 o 得到效果 $effect(o)$ 对应的观察值. 2) 对于主动感知, 感知动作被表示为

二元组 $o = \langle pre(o), obs(o) \rangle$, 其中 $obs(o)$ 表示执行感知动作后可能的观察值集合. 概率表示下的感知动作在上述定义的基础上. 类似扩展 1' 添加概率信息即可.

扩展 6 (观察值函数 (逻辑表示)). 令 Ω 表示观察值集合; $OB : A \times S \rightarrow \Omega$ 是观察值函数, $ob(a, s)$ 表示应用动作 a 后得到状态 s 时的观察值. 在信念状态 b 下应用动作 a 得到状态集合为 $S(b, a) = \{s' | s' \in \gamma(s, a), \forall s \in b\}$, 则对应观察值 ob 的信念状态为 $b_a^{ob} = \{s | s \in S(b, a) \wedge ob(a, s) = ob\}$.

扩展 6' (观察值函数 (概率表示)). 令 Ω 表示观察值集合; $OB : \Omega \times A \times S \rightarrow [0, 1]$ 是观察值函数, $P_a(ob|s)$ 表示在应用动作 a 后得到状态 s 时有观察值 ob 的概率. 在信念状态 b 下应用动作 a 后得到观察值 ob 的概率为 $b_a(ob) = \sum_{s \in S} P_a(ob|s)b_a(s)$.

扩展 7 (信念状态规划解 (逻辑表示)). 设有信念状态策略 $\pi : B \rightarrow A$, 策略 π 是规划解的条件为其执行结构 Σ_π 中存在从初始信念状态到目标信念状态的路径. 类似扩展 3, 规划解也有强规划解、强循环规划解和弱规划解之分.

扩展 7' (信念状态规划解 (概率表示)). 类似扩展 3', 针对有无奖励、代价函数两种情况, 分别定义概率表示下的信念状态规划解.

1) 设有信念状态策略 $\pi : B \rightarrow A$, 令 $E_\pi(b)$ 表示在策略 π 下信念状态 b 的期望效能值, 满足 $E_\pi(b) = R(b) - C(b, \pi(b)) + \beta \sum_{ob \in obs(b, \pi(b))} b_a(ob)E_\pi(b_a^{ob})$, 其中 $R(b) = \sum_{s \in b} R(s)b(s)$, $C(b, a) = \sum_{s \in b} C(s, a)b(s)$. 如果不存在其他策略 π' 使得 $E_\pi(b_0) < E_{\pi'}(b_0)$, 则称策略 π 是规划解.

2) 设有信念状态策略 $\pi : B \rightarrow A$, 存在从初始信念状态 b_0 达到目标信念状态 b_G 的路径 $path = \langle b_0, \pi(b_0), b_1, \pi(b_1), \dots, b_k, \pi(b_k), b_G \rangle$, 令 $P_{path} = \prod_{i=0}^k P_{\pi(b_i)}(b_{i+1}|b_i)$, $b_{k+1} = b_G$ 表示路径 $path$ 的概率, 其中 $P_{\pi(b_i)}(b_{i+1}|b_i) = b_{\pi(b_i)}(ob_{i+1})$, ob_{i+1} 表示信念状态 b_{i+1} 对应的观察值. 令 $P_\pi = \sum_i P_{path_i}$ 表示策略 π 的成功概率, 其中 $path_i$ 表示策略 π 的执行结构 Σ_π 中从初始状态可到达目标状态的执行路径. 如果不存在其他策略 π' 使得 $P_\pi < P_{\pi'}$, 则称策略 π 是规划解.

2.3 任务分解不确定

领域知识中的方法是 HTN 规划进行任务分解的依据, 需要领域专家提供. 对于熟悉的规划领域, 经验丰富的领域专家可以给出完备的 HTN 方法, 但面对复杂的规划领域可能无法提供完备的 HTN 方法, 甚至无法给出 HTN 方法. 这种领域知识的缺失会阻碍 HTN 规划在实际领域的应用, 学者们为了

消除这种障碍进行了不同的研究.

针对 HTN 方法不完备的情况, 有学者提出结合基于案例推理的技术弥补方法的缺失, 例如 Muñoz-Avila 等的工作; 也有学者提出在方法缺失时利用因果链推导^[45] 取代任务分解, 例如 Kambhampati 等^[46] 的工作. 针对无法给出 HTN 方法的情况, 学者们提出利用机器学习 (Machine learning) 的方法从以往的经验中学习 HTN 方法^[47], 包括方法结构的学习^[48] 和方法前提条件的学习^[49-52], 或者同时学习方法结构和前提条件^[53-54]. 然而, 领域知识缺失并不会引起任务分解不确定, 领域知识“冗余”——即一个复合任务存在多个可分解它的方法才是任务分解不确定的原因.

假设存在多个方法可以分解复合任务 ct , 记作方法子集 $M(ct)$. 一般地, 规划会从 $M(ct)$ 中任意选择一个方法 $m \in M(ct)$ 分解该任务, 如果 ct 分解后的任务网络无法被分解得到规划解, 规划会回溯到选择方法 m 的地方再从 $M(ct) - m$ 中任意选择一个方法分解 ct , 如此重复直到合成规划解或者返回失败. 合成规划解的任务分解路径会影响规划解所包含的动作, 而分解路径与选择的方法有关. 如果方法选择具有随机性, 动作的执行具有代价且状态有奖励, 则不同的任务分解路径对应的规划解具有不同的期望效用. 此时, 规划需要求解的不仅仅是一个动作序列, 而是具有最大期望效用的任务分解路径所对应的动作序列. 事实上, 领域建模者可以根据主观判断或以往经验对一个方法用于分解任务的可能性进行概率赋值. 基于此, 提出任务分解概率扩展表示任务分解不确定和最优分解路径扩展表示对规划解的要求.

扩展 8 (任务分解概率). 将规划领域 $D = \langle S, O, M, \gamma \rangle$ 扩展为 $D = \langle S, CT, O, M, \gamma, DP \rangle$, 其中 CT 是有限的复合任务集合; $DP: CT \times M \rightarrow [0, 1]$ 是任务分解的概率分布函数, $dp(m|ct)$ 表示用方法 m 分解任务 ct 的概率.

扩展 9 (最优分解路径). 假设 HTN 规划从初始状态 s_0 沿分解路径 $path = \langle ct_0, m_0, ct_1, m_1, \dots, pt_0, a_0, ct_i, m_i, \dots, pt_i, a_i, \dots, ct_k, m_k, \dots, pt_n, a_n \rangle$ 分解初始任务网络得到原子任务网络, 该路径可划分为复合任务分解路径 $CT_{path} = \langle ct_0, m_0, ct_1, m_1, \dots, ct_i, m_i, ct_k, m_k \rangle$ 和原子任务执行路径 $PT_{path} = \langle pt_0, a_0, pt_i, a_i, \dots, pt_n, a_n \rangle$. 如果动作有确定效果, PT_{path} 对应一条动作执行路径 $EX = \langle s_0, a_0, \dots, s_i, a_i, \dots, s_n, a_n \rangle$, 其效用值为 $V(EX) = \sum_{i=0}^n [R(s_i) - C(s_i, a_i)] \times P_{a_i}(s_{i+1}|s_i)$, 而分解路径 $path$ 的期望效用 $V(path) = \prod_{i=0}^k dp(m_i|ct_i)V(EX)$. 如果动作有不确定的效果, PT_{path} 对应一组动作执行路径 $\{EX_j = \langle s_0, a_0, \dots, s_i, a_i, \dots, s_n, a_n \rangle\}$, 则分解路径 $path$ 的期望效用 $V(path) = \prod_{i=0}^k dp(m_i|ct_i) \times \sum_j V(EX_j)$. 如果不存在其他分解路径 $path'$ 使得 $V(path) < V(path')$, 则称分解路径 $path$ 为最优分解路径.

3 不确定 HTN 规划研究现状

上文阐述了不确定 HTN 规划问题中的三类不确定, 并分析了现有方法针对不确定对原有 HTN 规划模型进行的扩展. 现有文献中也存在利用 HTN 规划技术求解不确定规划问题的研究工作, 表 2 给出了这些研究工作的情况, 包括求解问题类型、不确定表示方式和所作的扩展.

3.1 逻辑表示下状态完全可观、动作效果不确定 (P1) 的 HTN 规划

对于不确定规划问题 P1, Kuter 等^[55] 提出了 ND-SHOP2 (Non-deterministic simple hierarchical ordered planner 2) 规划器^[56], 其将前向链规划算法扩展为不确定前向链规划. 为了表示动作效果不确定, ND-SHOP2 对操作符的效果进行类似扩展 1 的修改. 另一方面, 不确定动作效果的存在, 使得

表 2 现有的不确定 HTN 规划研究

Table 2 Existing research on HTN planning under uncertainty

不确定 HTN 规划研究	不确定规划问题	不确定表示方式	扩展
ND-SHOP2、YoYo	P1	逻辑表示	扩展 1、2、3
Cond-SHOP2	P2	逻辑表示	扩展 4、5(2)、6、7
RTDP ^{SHOP2} 、LRTDP ^{SHOP2} 、 Fwd-VI ^{SHOP2}	P1	概率表示	扩展 1'、2、3'(1)
Hierarchical Factored POMDPs	P3	概率表示	扩展 1'、4'、5(1)、6'
C-SHOP、PC-SHOP	P3	概率表示	扩展 1'、4'、5(1)、6'、7'(2)
Probabilistic-HTN	P1+ 任务分解不确定	概率表示	扩展 1'、8、9

初始任务网络分解得到原子任务网络并不能保证一定达到目标状态, 为此 ND-SHOP2 对规划问题进行了类似扩展 2 的修改以显式地指定目标状态. ND-SHOP2 将规划解表示为扩展 3 中的形式, 并可以获得弱、强和强循环规划解. 在算法上, 前向链规划算法的特点是总是知道当前状态, 这方便了基于状态的剪枝技术的使用, 尤其是领域可裁剪规划器 (例如 SHOP2、TALplanner (Temporal action logics planner)^[57]、TLPlan (Temporal logics plan)^[58]) 中剪枝技术的使用. 当具有不确定效果的动作应用后, 存在多个可能的后继状态, 规划需要对每个可能的后继状态进行处理. 因此, 对前向链规划算法的不确定扩展是将动作应用后所有后继状态添加至未求解状态集, 然后任意选择一个状态进行处理并将该状态从未处理状态集中删除, 如此重复直到未求解状态集为空且达到目标状态. 实验结果表明, ND-SHOP2 相比其他一些非 HTN 规划器 (如 MBP (Model based planner)^[59]) 具有较高的效率. 然而, ND-SHOP2 的实用性有待加强, 表现为较低的可靠性和有待进一步提高的规划效率; 其也不具备求取复杂规划问题的能力.

针对上面提到的 ND-SHOP2 规划器的不足, Kuter 等^[60] 在 ND-SHOP2 的基础上, 结合基于 OBDD (Ordered binary decision diagram)^[61] 的状态集表示技术提出了 YoYo 规划器. YoYo 集成了 ND-SHOP2 中类似扩展 1 和扩展 2 的修改, 同样将规划解表示成扩展 3 中的形式, 也求取弱、强和强循环规划解. 不同的是, YoYo 利用有向 BDD 符号化表示状态集和状态转移关系, 并将该表示与 HTN 规划利用领域知识的高效搜索过程融合, 进一步提高了规划效率. 具体地, 状态集和状态转移关系被 OBDD 符号化表示后, 原来应用于单个状态的操作符变为应用于一个 OBDD 表示状态集合, 并依据 OBDD 运算规则计算操作符应用后的状态集合; 类似地, 应用于单个状态的任务分解变为一个状态集合下的任务分解. 依赖 OBDD 表示特有的运算规则, 基于状态集的推理过程比基于单个状态的推理过程高效很多. 实验结果也表明, YoYo 的规划效率高于只利用领域知识的 HTN 规划器 ND-SHOP2 和只利用 BDD 技术的非 HTN 规划器 MBP^[59], 在大规模问题上优势更加明显.

3.2 逻辑表示下状态部分可观、动作效果确定 (P2) 的 HTN 规划

对于不确定规划问题 P2, Kuter 等^[62] 提出了 Cond-SHOP2 规划器, 它是在研究部分可观条件下的前向链规划算法的基础上对 SHOP2 规划器扩展得到的. 在部分可观条件下, 为了表示信念状态及

其转移关系, Cond-SHOP2 对 SHOP2 进行了类似扩展 4 和扩展 6 的修改. 具体地, Cond-SHOP2 以状态变量形式表示状态, 一个状态对应所有状态变量的赋值集合, 则信念状态可以利用只知道部分状态变量赋值的部分状态表示; 而表示观察值函数的感知动作被用以对状态变量进行赋值, 并且区分感知动作与执行动作, 类似扩展 5-2). 此外, Cond-SHOP2 将规划解被表示为扩展 7 中的形式, 即由部分状态-动作序对组成. 在算法上, 为处理感知动作对状态变量的不同赋值, Cond-SHOP2 扩展前向链规划算法, 以对感知动作应用后产生的可能的部分状态进行分支处理. 具体地, Cond-SHOP2 在分解任务网络的同时, 利用执行动作转移部分状态, 并对感知动作产生的可能的部分状态进行分支处理, 直到部分状态集中所有部分状态均可到达目标状态为止. 最后的实验结果表明, 相比其他非 HTN 规划器 (如 PKS (Planning with knowledge and sensing)^[63-64]、MBP^[65]), Cond-SHOP2 能处理更多的问题且有较好的表现. 尽管 Cond-SHOP2 只能处理不确定问题 P2, 但结合 ND-SHOP2 中处理不确定动作效果的扩展, 也可以被用来处理不确定规划问题 P3, Kuter 等已经在进行相关研究.

3.3 概率表示下状态完全可观、动作效果不确定 (P1) 的 HTN 规划

对于不确定规划问题 P1, Kuter 等^[66] 将以领域知识作为控制策略的规划方法应用于 MDP 规划中, 通过修剪搜索空间, 提高规划效率. MDP 规划将问题 P1 对应的规划领域看作为一个随机系统, 利用扩展 1' 的形式表示具有随机性的动作效果, 利用扩展 2 的形式表示规划问题, 利用扩展 3'(1) 的形式定义规划解. 现有 MDP 规划算法的严重缺陷之一在于每次迭代时需要更新整个状态空间, 严重影响规划效率和可求解问题规模. 为此, Kuter 等将 SHOP2 的控制策略分别与 MDP 迭代算法 RTDP (Real-time dynamic programming)^[67]、LRTDP (Labelled RTDP)^[68] 和 Fwd-VI (Forward chaining version of value iteration)^[69] 结合, 相应地提出了 RTDP^{SHOP2}、LRTDP^{SHOP2} 和 Fwd-VI^{SHOP2} 规划器. 具体地, 规划器在前向链 MDP 规划算法框架下进行迭代, 每次迭代时利用 HTN 领域知识修剪可应用于当前状态的操作符集合, 这大大减少了每次迭代需要计算的状态数量, 进而有效提高了规划效率. Kuter 等也给出了搜索控制规则可应用于前向链 MDP 规划算法的条件, 并从理论和实验两方面说明搜索控制规则能够产生指数级增速.

3.4 概率表示下状态部分可观、动作效果不确定 (P3) 的 HTN 规划

对于不确定规划问题 P3, 现有的研究根据问题求解目标大体可以分为两类, 一类是求期望效用最大的规划解, 通常针对有奖励值和 (或) 代价值的规划问题; 另一类是求成功概率最大的规划解, 通常针对没有奖励值和 (或) 代价值的规划问题. 下面分别进行总结:

1) 求期望效用最大的规划解

对于不确定问题 P3, Müller 等^[70-71] 将 HTN 应用于 POMDP 求解, 通过利用领域知识可求解大规模 POMDP 问题. a) 对于一个部分可观的随机系统, 规划领域可以被刻画为 POMDP (Partially observable Markov decision processes) 模型, 与扩展 1' 一样, 利用带概率分布的动作效果表示具有随机性的状态转移关系. 此外, Müller 等讨论的动作不仅具有不确定效果, 还具有条件效果^[31], 即行动的效果依赖于执行时所处的状态, 但条件效果本身没有引入不确定. b) POMDP 模型中利用带概率分布的状态集合表示当前不确定的状态集合, 即形如扩展 4' 的概率信念状态. c) 系统部分可观时, 感知动作被用于获得观察值以推断当前的不确定状态, 并且随机环境下感知动作的观察值带有概率分布, 因而需要进行扩展 6' 所定义的观察值函数扩展, 以描述随机条件下信念状态的转移关系, 通常也采取类似扩展 5-1) 的形式表示感知动作.

除了上述扩展外, Müller 等的研究还包括以下几方面的内容: a) POMDP 规划问题通常定义形如扩展 7-1) 的信念状态规划解, 但也有研究利用有限状态控制器 FSC (Finite state controller) 表示 POMDP 问题的策略^[72]. FSC 由表示动作的节点集合和表示动作转移关系的转移函数组成, 转移函数指明了从一个动作节点转移到另一个动作节点的条件. 由于要枚举所有的观察值, FSC 的转移函数会变得很庞大而不方便使用. 作者在观察值可以分割的事实^[61] 基础上定义了逻辑有限状态控制器 LFSC (Logical finite state controller), 可以紧凑地表示 FSC. b) 为结合 HTN 规划的优势, 该研究引入抽象动作概念与 HTN 规划中复合任务相对应, 并利用 MFSC (Method FSC) 表示抽象动作的求精策略, 对应 HTN 规划中的任务分解过程. 在此基础上, 作者定义了 Hierarchical factored POMDP 问题, 并依据 FSC 执行的期望效用^[72] 定义了问题解, 为执行的期望效用最大的 FSC. c) 在层次推理中嵌入 A* 算法^[73] 和 UCT^[74] (Upper confidence bound applied to trees) 算法, 用于在 LFSC 空间中搜索规划解. 具体地, 规划利用 MFSC 对部分抽

象的 LFSC 求精, 直到 LFSC 变为不含抽象动作而只含执行动作的原子控制器. 在求精过程中运用算法和 UCT 算法选择 MFSC 可求精的控制器节点. d) 通过实验说明, 应用 UCT 算法能够更高效地求解大规模的 POMDP 规划问题.

2) 求成功概率最大的规划解

对于不确定问题 P3, Bouguerra 等扩展了经典 HTN 规划器 SHOP 提出了 C-SHOP 规划器^[75], 以获得成功概率大于设定值的规划解. 为处理动作效果不确定, C-SHOP 对操作符进行了形如扩展 1' 的修改, 以表示动作执行后带有概率分布的效果. 为了表示部分可观条件下带有概率分布的状态集合, C-SHOP 定义了形如扩展 4' 的信念状态; 为了描述信念状态之间的转移关系, C-SHOP 在操作符扩展中加入观察值以表示动作执行后可能获得的观察值, 类似扩展 5-1) 并采用类似扩展 6' 的方式推断当前信念状态. C-SHOP 同时定义了 COND 任务, 该类任务依据不同的观察值条件有不同的扩展分支. 需要指出, C-SHOP 存在与扩展 4' 不同的地方, 即不设置目标状态而以完全分解任务网络为规划终止条件. 另外, C-SHOP 的规划解被表示为决策树而非策略, 决策树^[33] 中的节点表示当前状态所对应的观察值, 而连接节点的弧表示动作. 因为动作执行后的观察值在规划中无法确定, 因而 C-SHOP 需要对所有可能的观察值进行分支处理.

此外, Bouguerra 等在 C-SHOP 的基础上进一步提出了 PC-SHOP^[76] 规划器. PC-SHOP 求解的问题与 C-SHOP 相同, 在规划算法上与 C-SHOP 也基本一致, 唯一不同的是在规划中利用动作效果和信念状态所带的概率信息进行简单概率推理, 在任务分解时以最大成功概率为启发式函数, 选择成功概率最大的分解分支, 以求得成功概率最大的规划解, 即扩展 7'-2).

3.5 概率表示下状态完全可观、动作效果不确定、任务分解不确定的 HTN 规划

针对利用 HTN 规划求解随机环境下的问题 P1 存在任务分解不确定的情况, Tang 等^[77] 提出利用概率层次任务网络求解此类问题. 具体地, 作者利用 MDP 模型对随机环境下的问题 P1 进行建模, 采用扩展 1' 的方式表示随机的动作效果, 同时引入扩展 3'-1) 中奖励值函数和代价值函数; 而为了表示任务分解的不确定还必须进行类似扩展 8 的修改. 由于 MDP 模型对应的状态转移图与 HTN 规划对应的任务分解树具有不同的结构, 为了将两者结合, 作者提出利用 Earley 图^[78] 表示不确定的任务分解产生的带有概率的层次任务网络, 再通过层次任务网络中的原子任务对应的动作与 MDP 模型融合, 以评

估 Earley 中路径的期望效用. 具体地, 基于任务分解与语法解析的关系^[79], 作者考虑修改 Earley 解析器^[80]来处理 HTN 规划的方法集 M 以产生 Earley 图, 利用 Earley 图的结点表示任务与方法的对应关系, 利用 Earley 图的连接表示任务间的分解关系, 连接带有的概率即任务分解时方法被选择的概率. 表示方法集的 Earley 图中包含任务的分解路径, 结合分解路径中动作对应的状态转移关系可以得到任务的分解-执行路径, 而根据路径上方法被选择的概率、动作转移的概率、动作执行的代价以及到达状态获得奖励值计算该分解-执行路径的期望效用. 为此, 作者提出 Earley 图遍历算法, 以搜索具有最大期望效用的分解路径作为规划问题的解, 类似扩展 9. 此外, 作者指出, 方法被选择的概率既可从专家的主观评估获得, 也可以从以往的分解统计数据中获得; 如果无法事先给定该概率, 可以先设定一组先验概率, 再通过学习得到后验概率来更新该概率, 此工作类似文献 [78] 中语法概率的学习.

3.6 小结

HTN 规划本质上是一种基于状态前向搜索的规划方法(简称前向规划), 而任务分解使 HTN 规划在处理不确定规划问题上相比非 HTN 的前向规划有明显的优势, 这主要体现在以下两个方面. 1) 任务分解具有动作选择能力. 非 HTN 的前向规划从当前状态下面前提条件被满足的动作集合中选择动作来构建达到目标状态的路径. 相比之下, HTN 规划只需要考虑那些沿着任务分解产生的原子任务对应的动作, 而不必考虑当前状态下可应用的动作集合中的所有动作, 减少了搜索需要访问的状态进而提高规划效率. 2) 任务分解具有感知选择能力. 针对状态部分可观的不确定规划问题, 非 HTN 的前向规划需要考虑当前条件下所有可执行的感知的观察值组合, 而不确定 HTN 规划可以根据任务分解的条件选择需要执行的感知, 不必考虑所有可执行感知的观察值组合, 减少了搜索需要访问的信念状态进而提高规划效率.

4 总结和未来研究方向

HTN 规划利用领域知识进行分层分解在求解确定性规划问题时获得了令人满意的效果, 这促使人们想要将这种方式应用到不确定规划问题的求解中, 因而出现了以一些基于 HTN 规划的不确定规划方法的研究.

现有文献针对不确定 HTN 规划的研究主要关注三类不确定, 即系统状态不确定、动作效果不确定和任务分解不确定, 然而并没有研究同时考虑三类不确定而只考虑其中的一类或两类不确定. 此外, 不

同的不确定表示, 逻辑表示或概率表示, 也会带来规划方法的不同. 现有文献的工作分别对 5 种情形下的不确定 HTN 规划进行了研究, 包括逻辑表示下状态完全客观、动作效果不确定的 HTN 规划, 逻辑表示下状态部分可观、动作效果确定的 HTN 规划, 概率表示下状态完全可观、动作效果不确定的 HTN 规划, 概率表示下状态部分可观、动作效果不确定的 HTN 规划以及概率表示下状态可观、动作效果不确定且考虑任务分解不确定的 HTN 规划.

尽管不确定 HTN 规划较早被人们关注, 但其研究仍然处于起步阶段, 还有一些尚未考虑的问题有待研究, 现有的研究也存在一些不足的地方需要继续研究.

1) 系统状态完全不可观的不确定 HTN 规划研究. 在航空航天、机器人控制和系统智能控制等领域, 规划智能体无法感知系统信息或感知代价太高^[81], 导致系统状态完全不可观. 求解此类规划问题需要产生一个动作序列可以“强制”规划智能体从初始信念状态到达目标信念状态, 即不论规划智能体处于哪种状态, 都能被强制进入一个给定的状态. 对于无法感知或感知代价很昂贵的领域, 这种强制是合适的; 而如果可以利用领域知识表达这种状态强制转移的规则, 那么 HTN 规划也将在这类不确定规划问题求解中发挥重要作用.

2) 带扩展目标的不确定 HTN 规划研究. 考虑到动作效果的不确定和可能的执行失败, 不确定规划问题对规划目标不仅有可达性要求, 还可能有其他方面的要求, 主要表现在以下两个方面: a) 对于考虑动作代价和状态奖励的不确定规划领域, 规划问题可能要求达到的规划目标具有最大的期望效用; b) 规划问题可能要求在达到规划目标的过程中避免或达到某些特定的中间状态, 还可能根据现实环境的严苛程度决定实现这种扩展的规划目标的强度要求是“必须”、“尽量”或其他强度. 扩展的规划目标在原本已经相当复杂的不确定规划问题上添加了新的难度, 因而研究带扩展目标的不确定 HTN 规划是一项富有挑战性的工作.

3) 考虑时间和资源不确定的 HTN 规划研究. 不确定 HTN 规划研究仅在网络服务组合 (Web service composition)^[82]、交互式故事系统 (Interactive storytelling system)^[83] 和软件开发管理^[84] 等现实领域有所应用, 但这些研究缺乏对时间和资源方面的不确定的考虑, 而这在实际应用中相当重要. 因此, 开展研究探索考虑时间和资源不确定的 HTN 规划具有重要的现实意义, 而确定性 HTN 规划中处理时间^[85-87] 和资源^[88] 的研究思路值得借鉴, 其他规划方法处理时间和资源不确定的机制^[89] 也值得借鉴.

4) 基于问题转换的不确定 HTN 规划研究. 现有的不确定 HTN 规划研究都是在扩展的规划模型上构建搜索空间并寻找规划解, 这种方式在面对现实世界的大规模问题时也会因为搜索空间爆炸而不适用. 近年来, FF-replan (Fast-forward replan) 规划器^[90] 在国际规划大赛中的高效表现, 使其将不确定规划问题转换为确定性规划问题并利用重规划合成规划解的方法成为研究的热点^[35, 38, 91-92], 其研究重点是转换算法和规划解合成. 如果不确定 HTN 规划问题可以被转换为确定性 HTN 规划问题, 并利用成熟的 HTN 规划器求解, 那么确定性 HTN 规划的高效性将使 HTN 规划在求解大规模不确定规划问题上有更好的表现.

5) 结合强化学习 (Reinforcement learning) 的不确定 HTN 规划研究. 任务分解是 HTN 规划过程中的重要步骤之一, 通过选择可行的方法来分解复合任务进而得到更小的子任务. 一般地, 领域知识中可能存在多个方法对应一个复合任务, 现有的 HTN 规划方法都是无差别地从这些可行的方法集合中选择一个方法分解任务, 最终得到的规划方案的质量可能令人不满意, 而因为选择一个方法后继续分解失败导致回溯也会增加规划消耗的时间. 如果能够借助强化学习的策略优化能力学习 HTN 方法的选择策略, 保证方法选择后的成功率以及对规划方案质量的提升, 那么将这些方法的选择策略应用到 HTN 规划中不仅可以提高规划方案的质量, 甚至可以提高规划效率.

6) 不确定领域的领域知识学习. 除了关注规划方法的研究外, 人们也很重视对领域知识学习的研究. 除了 Hogg 等^[93] 的工作, 现有文献对领域知识的学习都局限于确定的规划领域. Hogg 等的工作也只是对动作效果不确定情况下 HTN 方法的学习进行了研究, 并没有考虑状态不完全可观的情况. 然而, 状态不完全可观势必会带来 HTN 方法在前提条件和分解结构上的不同, 如何在这种状态不完全可观, 甚至动作效果不确定的情况下学习正确的 HTN 方法是一个值得研究的问题.

References

- Ghallab M, Nau D, Traverso P [Author], Jiang Yun-Fei, Yang Qiang, Ling Ying-Biao, Wu Li-Hua, Chen Ai-Xiang, Zhang Xue-Nong, Huang Wei [Translator]. *Automated Planning: Theory & Practice*. Beijing: Tsinghua University Press, 2008.
(Ghallab M, Nau D, Traverso P [著], 姜云飞, 杨强, 凌应标, 伍丽华, 陈嵩祥, 张学农, 黄巍 [译]. 自动规划: 理论与实践. 北京: 清华大学出版社, 2008.)
- Sacerdoti E D. The nonlinear nature of plans. In: Proceedings of the 4th International Joint Conference on Artificial Intelligence. San Francisco, CA, USA: Morgan Kaufmann, 1975. 206-214
- Yang Q. Formalizing planning knowledge for hierarchical planning. *Computational Intelligence*, 1990, **6**(1): 12-24
- Kambhampati S, Hendler J A. A validation-structure-based theory of plan modification and reuse. *Artificial Intelligence*, 1992, **55**(2-3): 193-258
- Erol K, Hendler J, Nau D S. UMCP: a sound and complete procedure for hierarchical task-network planning. In: Proceedings of the 2nd International Conference on Artificial Intelligence Planning Systems. Palo Alto, California: AAAI Press, 1994. 249-254
- Erol K, Nau D S, Subrahmanian V S. Complexity, decidability and undecidability results for domain-independent planning. *Artificial Intelligence*, 1995, **76**(1-2): 75-88
- Tate A. Generating project networks. In: Proceedings of the 5th International Joint Conference on Artificial Intelligence. San Francisco, CA, USA: Morgan Kaufmann, 1977. 888-893
- Wilkins D E. Domain-independent planning representation and plan generation. *Artificial Intelligence*, 1984, **22**(3): 269-301
- Wilkins D E. Can AI planners solve practical problems? *Computational Intelligence*, 1990, **6**(4): 232-246
- Currie K, Tate A. O-Plan: the open planning architecture. *Artificial Intelligence*, 1991, **52**(1): 49-86
- Tate A, Drabble B, Kirby R. O-Plan2: an open architecture for command, planning and control. *Intelligent Scheduling*. San Francisco, CA, USA: Morgan Kaufmann, 1994. 213-239
- Nau D, Cao Y, Lotem A, Munoz-Avila H. SHOP: simple hierarchical ordered planner. In: Proceedings of the 16th International Joint Conference on Artificial Intelligence. San Francisco, CA, USA: Morgan Kaufmann, 1999. 968-973
- Nau D, Au T C, Ilghami O, Kuter U, Murdock J W, Wu D, Yaman F. SHOP2: an HTN planning system. *Journal of Artificial Intelligence Research*, 2003, **20**: 379-404
- de la Asunción M, Castillo L, Fdez-Olivares J, García-Pérez O, González A, Palao F. SIADEX: an interactive knowledge-based planner for decision support in forest fire fighting. *AI Communications*, 2005, **18**(4): 257-268
- Georgievski I, Aiello M. HTN planning: overview, comparison, and beyond. *Artificial Intelligence*, 2015, **222**: 124-156
- Wilkins D E, desJardins M. A call for knowledge-based planning. *AI Magazine*, 2001, **22**(1): 99-115
- Nau D, Au T C, Ilghami O, Kuter U, Wu D, Yaman F, Munoz-Avila H, Murdock J W. Applications of SHOP and SHOP2. *IEEE Intelligent Systems*, 2005, **20**(2): 34-41
- Wilkins D E. *Practical Planning: Extending the Classical AI Planning Paradigm*. San Francisco, CA, USA: Morgan Kaufmann, 1988.
- Biundo S, Schattenberg B. From abstract crisis to concrete relief — a preliminary report on combining state abstraction and HTN planning. In: Proceedings of the 6th European Conference on Planning. Palo Alto, California: AAAI Press, 2001. 157-168
- Arentoft M M, Fuchs J J, Parrod Y, Gasquet A, Stader J, Stokes I, Vadon H. OPTIMUM-AIV: a planning and scheduling system for spacecraft AIV. *Future Generation Computer Systems*, 1992, **7**(4): 403-412

- 21 Estlin T A, Chien S A, Wang X M. An argument for a hybrid HTN/operator-based approach to planning. In: Proceedings of the 4th European Conference on Planning: Recent Advances in AI Planning. Heidelberg, Berlin: Springer, 1997. 182–194
- 22 Agosta J M. Formulation and implementation of an equipment configuration problem with the SIPE-2 generative planner. In: Proceedings of AAAI-95 Spring Symposium on Integrated Planning Applications. Menlo Park, California: AAAI Press, 1995. 1–10
- 23 Smith S J J, Hebbar K, Nau D S, Minis I. Integrating electrical and mechanical design and process planning. In: Proceedings of the IFIP TC5 WG5. 2 International Conference on Knowledge Intensive CAD. US: Springer, 1997. 269–288
- 24 Muñoz-Avila H, Aha D W, Nau D S, Weber R, Breslow L, Yamal F. SiN: integrating case-based reasoning with task decomposition. In: Proceedings of the 17th International Joint Conference on Artificial Intelligence-Volume 2. San Francisco, CA, USA: Morgan Kaufmann, 2001. 999–1004
- 25 Smith S J, Nau D, Throop T. Computer Bridge: a big win for AI planning. *AI Magazine*, 1998, **19**(2): 93–106
- 26 Morisset B, Ghallab M. Learning how to combine sensory-motor modalities for a robust behavior. *Advances in Plan-Based Control of Robotic Agents*. Heidelberg, Berlin: Springer, 2002. 157–178
- 27 Morisset B, Ghallab M. Synthesis of supervision policies for robust sensory-motor behaviors. In: Proceedings of the 7th International Conference on Intelligent Autonomous Systems. Amsterdam: IOS Press, 2002. 236–243
- 28 Sirin E, Parsia B, Wu D, Hendler J, Nau D. HTN planning for web service composition using SHOP2. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2004, **1**(4): 377–396
- 29 González-Ferrer A, Ten Teije A, Fdez-Olivares J, Milian K. Automated generation of patient-tailored electronic care pathways by translating computer-interpretable guidelines into hierarchical task networks. *Artificial Intelligence in Medicine*, 2013, **57**(2): 91–109
- 30 Wang Hong-Wei, Qi Chao. *Hierarchical Task Network Planning based Emergency Response Decision Making Theory and Method*. Beijing: Science Press, 2015.
(王红卫, 祁超. 基于层次任务网络规划的应急响应决策理论与方法. 北京: 科学出版社, 2015.)
- 31 Russell S, Norvig P [Author], Jiang Zhe, Jin Yi-Jiang, Zhang Min, Yang Lei, et al. [Translator]. *Artificial Intelligence: A Modern Approach* (Second Edition). Beijing: Posts & Telecom Press, 2010.
(Russell S, Norvig P [著], 姜哲, 金奕江, 张敏, 杨磊等 [译]. 人工智能: 一种现代的方法 (第二版). 北京: 人民邮电出版社, 2010.)
- 32 Geier T, Bercher P. On the decidability of HTN planning with task insertion. In: Proceedings of the 22nd International Joint Conference on Artificial Intelligence. Palo Alto, California: AAAI Press, 2011. 1955–1961
- 33 Alford R, Shivashankar V, Kuter U, Nau D. HTN problem spaces: structure, algorithms, termination. In: Proceedings of the 5th Annual Symposium on Combinatorial Search. Palo Alto, California: AAAI Press, 2012. 2–9
- 34 Ramírez M, Sardina S. Directed fixed-point regression-based planning for non-deterministic domains. In: Proceedings of the 24th International Conference on Automated Planning and Scheduling. Palo Alto, California: AAAI Press, 2014. 235–243
- 35 Alford R, Kuter U, Nau D, Goldman R P. Plan aggregation for strong cyclic planning in nondeterministic domains. *Artificial Intelligence*, 2014, **216**: 206–232
- 36 Keller T, Eyerich P. PROST: probabilistic planning based on UCT. In: Proceedings of the 22nd International Conference on Automated Planning and Scheduling. Palo Alto, California: AAAI Press, 2012. 119–127
- 37 Muise C, Belle V, McIlraith S A. Computing contingent plans via fully observable non-deterministic planning. In: Proceedings of the 28th AAAI Conference on Artificial Intelligence. Palo Alto, California: AAAI Press, 2014. 2322–2329
- 38 Maliah S, Brafman R, Karpas E, Shani G. Partially observable online contingent planning using landmark heuristics. In: Proceedings of the 24th International Conference on Automated Planning and Scheduling. Palo Alto, California: AAAI Press, 2014. 163–171
- 39 Zhang S Q, Stone P. CORPP: commonsense reasoning and probabilistic planning, as applied to dialog with a mobile robot. In: Proceedings of the 29th AAAI Conference on Artificial Intelligence. Austin, Texas, USA: AAAI Press, 2015. 1394–1400
- 40 Taig R, Brafman R I. A relevance-based compilation method for conformant probabilistic planning. In: Proceedings of the 28th AAAI Conference on Artificial Intelligence. Palo Alto, California: AAAI Press, 2014. 2374–2380
- 41 Nguyen K, Tran V, Son T C, Pontelli E. On computing conformant plans using classical planners: a generate-and-complete approach. In: Proceedings of the 22nd International Conference on Automated Planning and Scheduling. Palo Alto, California: AAAI Press, 2012. 190–198
- 42 Bresina J, Dearden R, Meuleau N, Ramakrishnan S, Smith D, Washington R. Planning under continuous time and resource uncertainty: a challenge for AI. In: Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence. San Francisco, CA, USA: Morgan Kaufmann, 2002. 77–84
- 43 Celorrio S J. Planning and Learning under Uncertainty [Ph. D. dissertation], Universidad Carlos III de Madrid, Spain, 2011
- 44 Cimatti A, Pistore M, Roveri M, Traverso P. Weak, strong, and strong cyclic planning via symbolic model checking. *Artificial Intelligence*, 2003, **147**(1–2): 35–84
- 45 McAllester D, Rosenblitt D. Systematic nonlinear planning. In: Proceedings of the 9th National conference on Artificial intelligence-Volume 2. Palo Alto, California: AAAI Press, 1991. 634–639
- 46 Kambhampati S, Mali A, Srivastava B. Hybrid planning for partially hierarchical domains. In: Proceedings of the 15th National Conference on Artificial Intelligence. Palo Alto, California: AAAI Press, 1998. 882–888
- 47 Kuter U. Learning and planning (intelligent systems). *Encyclopedia of Complexity and Systems Science*. New York: Springer, 2009. 5188–5206
- 48 Garland A, Ryall K, Rich C. Learning hierarchical task models by defining and refining examples. In: Proceedings of the 1st International Conference on Knowledge Capture. New York: ACM, 2001. 44–51
- 49 Ilghami O, Muñoz-Avila H, Nau D S, Aha D W. Learning approximate preconditions for methods in hierarchical plans. In: Proceedings of the 22nd International Conference on Machine Learning. New York: ACM, 2005. 337–344

- 50 Xu K, Muñoz-Avila H. A domain-independent system for case-based task decomposition without domain theories. In: Proceedings of the 20th National Conference on Artificial Intelligence. Palo Alto, California: AAAI Press, 2005. 234–239
- 51 Nejati N, Langley P, Konik T. Learning hierarchical task networks by observation. In: Proceedings of the 23rd International Conference on Machine Learning. New York: ACM, 2006. 665–672
- 52 Reddy C, Tadepalli P. Learning goal-decomposition rules using exercises. In: Proceedings of the 14th International Conference on Machine Learning. San Francisco: Morgan Kaufmann Publishers Inc., 1997. 278–286
- 53 Zhuo H H, Muñoz-Avila H, Yang Q. Learning hierarchical task network domains from partially observed plan traces. *Artificial Intelligence*, 2014, **212**: 134–157
- 54 Hogg C, Muñoz-Avila H, Kuter U. HTN-MAKER: learning HTNs with minimal additional knowledge engineering required. In: Proceedings of the 23th National Conference on Artificial Intelligence. Palo Alto, California: AAAI Press, 2008. 950–956
- 55 Kuter U, Nau D. Forward-chaining planning in nondeterministic domains. In: Proceedings of the 19th National Conference on Artificial Intelligence. Palo Alto, California: AAAI Press, 2004. 513–518
- 56 Kuter U. Planning under Uncertainty: Moving Forward [Ph.D. dissertation]. Maryland University College Park, USA, 2006
- 57 Kvarnström J, Doherty P. TALplanner: a temporal logic based forward chaining planner. *Annals of Mathematics and Artificial Intelligence*, 2000, **30**(1–4): 119–169
- 58 Bacchus F, Kabanza F. Using temporal logics to express search control knowledge for planning. *Artificial Intelligence*, 2000, **116**(1–2): 123–191
- 59 Bertoli P, Cimatti A, Pistore M, Roveri M, Traverso P. MBP: a model based planner. In: Proceedings of the IJCAI-2001 Workshop on Planning under Uncertainty and Incomplete Information. San Francisco, USA: Morgan Kaufmann, 2001. 473–478
- 60 Kuter U, Nau D, Pistore M, Traverso P. Task decomposition on abstract states, for planning under nondeterminism. *Artificial Intelligence*, 2009, **173**(5–6): 669–695
- 61 Bryant R E. Symbolic boolean manipulation with ordered binary-decision diagrams. *ACM Computing Surveys*, 1992, **24**(3): 293–318
- 62 Kuter U, Nau D, Reisner E, Goldman R. Conditionalization: adapting forward-chaining planners to partially observable environments. In: ICAPS-2007 Workshop on Planning and Execution for Real-World Systems. Providence, Rhode Island, USA, 2007.
- 63 Petrick R P A, Bacchus F. A knowledge-based approach to planning with incomplete information and sensing. In: Proceedings of the 6th International Conference on AI Planning and Scheduling. Palo Alto, California: AAAI Press, 2002. 212–222
- 64 Petrick R P A, Bacchus F. Extending the knowledge-based approach to planning with incomplete information and sensing. In: Proceedings of the 9th International Conference on Knowledge Representation and Reasoning. Palo Alto, California: AAAI Press, 2004. 613–622
- 65 Bertoli P, Cimatti A, Roveri M, Traverso P. Strong planning under partial observability. *Artificial Intelligence*, 2006, **170**(4–5): 337–384
- 66 Kuter U, Nau D. Using domain-configurable search control for probabilistic planning. In: Proceedings of 20th National Conference on Artificial Intelligence. Palo Alto, California: AAAI Press, 2005. 1169–1174
- 67 Bonet B, Geffner H. Planning with incomplete information as heuristic search in belief space. In: Proceedings of the 5th International Conference on AI Planning and Scheduling. Palo Alto, California: AAAI Press, 2000. 52–61
- 68 Bonet B, Geffner H. Labeled RTDP: improving the convergence of real-time dynamic programming. In: Proceedings of the 13th International Conference on Automated Planning and Scheduling. Trento, Italy: AAAI Press, 2003. 12–21
- 69 Bertsekas D P. *Dynamic Programming and Optimal Control*. Belmont, Massachusetts: Athena Scientific, 1995.
- 70 Müller F, Späth C, Geier T, Biundo S. Exploiting expert knowledge in factored POMDPs. In: Proceedings of the 20th European Conference on Artificial Intelligence. Amsterdam: IOS Press, 2012. 606–611
- 71 Müller F, Biundo S. HTN-style planning in relational POMDPs using first-order FSCs. *KI 2011: Advances in Artificial Intelligence*. Heidelberg, Berlin: Springer, 2011. 216–227
- 72 Hansen E A. Solving POMDPs by searching in policy space. In: Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence. San Francisco, CA, USA: Morgan Kaufmann, 1998. 211–219
- 73 Nilsson N J. *Principles of Artificial Intelligence*. San Francisco, CA, USA: Morgan Kaufmann, 2014.
- 74 Kocsis L, Szepesvári C. Bandit based Monte-Carlo planning. In: Proceedings of 17th European Conference on Machine Learning. Heidelberg, Berlin: Springer, 2006. 282–293
- 75 Bouguerra A, Karlsson L. Hierarchical task planning under uncertainty. In: Proceedings of the 3rd Italian Workshop on Planning and Scheduling. Perugia, Italy, 2004.
- 76 Bouguerra A, Karlsson L. PC-SHOP: a probabilistic-conditional hierarchical task planner. *Intelligenza Artificiale*, 2005, **2**(4): 44–50
- 77 Tang Y Q, Meneguzzi F, Sycara K, Parsons S. Planning over MDPs through probabilistic HTNs. In: AAAI-2011 Workshop on Generalized Planning. San Francisco, CA, USA: AAAI, 2011.
- 78 Stolcke A. An efficient probabilistic context-free parsing algorithm that computes prefix probabilities. *Computational Linguistics*, 1995, **21**(2): 165–201
- 79 Barrett A, Weld D S. Task-decomposition via plan parsing. In: Proceedings of the 12th National Conference on Artificial Intelligence. San Francisco, CA, USA: AAAI Press, 1994. 1117–1122
- 80 Earley J. An efficient context-free parsing algorithm. *Communications of the ACM*, 1970, **13**(2): 94–102
- 81 Zhao J, Sun J, Yin M. Recent advances in conformant planning. In: Proceedings of the 2007 IEEE International Conference on Robotics and Biomimetics (ROBIO). Sanya: IEEE, 2007. 2001–2006
- 82 Kuter U, Sirin E, Parsia B, Nau D, Hendler J. Information gathering during planning for web service composition. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2005, **3**(2–3): 183–205

- 83 da Silva F A G, Ciarlini A E M, Siqueira S W M. A planning algorithm for incorporating attempts and nondeterminism into interactive stories. In: Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence. Boca Raton, FL: IEEE, 2011. 96–101
- 84 Biundo S, Holzer R, Schattenberg B. Project planning under temporal uncertainty. In: Proceedings of Planning, Scheduling and Constraint Satisfaction from Theory to Practice. Amsterdam: IOS Press, 2005. 189–198
- 85 Yaman F, Nau D S. Timeline: an HTN planner that can reason about time. In: Proceedings of AIPS-02 Workshop on Planning for Temporal Domains. San Francisco, CA, USA: AAAI Press, 2002. 75–81
- 86 Castillo L A, Fdez-Olivares J, García-Pérez O, Palao F. Efficiently handling temporal knowledge in an HTN planner. In: Proceedings of 16th International Conference on Automated Planning and Scheduling. San Francisco, CA, USA: AAAI Press, 2006. 63–72
- 87 Tang P, Wang H W, Qi C, Wang J. Anytime heuristic search in temporal HTN planning for developing incident action plans. *AI Communications*, 2012, **25**(4): 321–342
- 88 Wang Z, Wang H W, Qi C, Wang J. A resource enhanced HTN planning approach for emergency decision-making. *Applied Intelligence*, 2013, **38**(2): 226–238
- 89 Beaudry É, Kabanza F, Michaud F. Planning with concurrency under resources and time uncertainty. In: Proceedings of the 19th European Conference on Artificial Intelligence. Amsterdam, The Netherlands: IOS Press, 2010. 217–222
- 90 Yoon S W, Fern A, Givan R. FF-Replan: a baseline for probabilistic planning. In: Proceedings of the 17th International Conference on Automated Planning and Scheduling. San Francisco, CA, USA: AAAI Press, 2007. 352–359
- 91 Palacios H, Albore A, Geffner H. Compiling contingent planning into classical planning: new translations and results. In: Proceedings of ICAPS-2014 Workshop on Models and Paradigms for Planning under Uncertainty: a Broad Perspective. San Francisco, CA, USA: AAAI Press, 2014. 27–34
- 92 Shani G, Brafman R I. Replanning in domains with partial information and sensing actions. In: Proceedings of the 22nd International Joint Conference on Artificial Intelligence. San Francisco, CA, USA: AAAI Press, 2011. 2021–2026
- 93 Hogg C, Kuter U, Muñoz-Avila H. Learning hierarchical task networks for nondeterministic planning domains. In: Proceedings of the 21st International Joint Conference on Artificial Intelligence. San Francisco, CA, USA: Morgan Kaufmann, 2009. 1708–1714



王红卫 华中科技大学自动化学院系统工程专业博士, 教授。主要研究方向为物流系统, 公共安全与应急管理, 工程管理。本文通信作者。

E-mail: hwwang@mail.hust.edu.cn
(**WANG Hong-Wei** Ph.D., professor in systems engineering at the School of Automation, Huazhong University of

Science and Technology. His research interest covers logistics systems, public security and emergency management, and modeling simulation of complex systems. Corresponding author of this paper.)



刘典 华中科技大学自动化学院博士研究生。2008 年获得华中科技大学控制科学与工程系学士学位。主要研究方向为 HTN 规划, 应急管理, 人工智能。

E-mail: liudianop@hust.edu.cn

(**LIU Dian** Ph.D. candidate at the School of Automation, Huazhong University of Science and Technology. He

received his bachelor degree from Huazhong University of Science and Technology in 2008. His research interest covers HTN planning, emergency management, and artificial intelligence.)



赵鹏 华中科技大学自动化学院博士研究生。2008 年获得华中科技大学控制科学与工程系学士学位。主要研究方向为 HTN 规划, 应急管理, 人工智能。

E-mail: pengzhao@hust.edu.cn

(**ZHAO Peng** Ph.D. candidate at the School of Automation, Huazhong University of Science and Technology. He

received his bachelor degree from Huazhong University of Science and Technology in 2008. His research interest covers HTN planning, emergency management, and artificial intelligence.)



祁超 华中科技大学自动化学院系统科学与工程系副教授, 博士。2006 年毕业于新加坡南洋理工大学。主要研究方向为自动规划, 调度与优化, 应急管理与决策。E-mail: qichao@hust.edu.cn

(**QI Chao** Ph.D., associate professor in the Department of Systems Science and Engineering, School of Au-

tomation, Huazhong University of Science and Technology, China. She graduated from Nanyang Technological University, Singapore in 2006. Her research interest covers automated planning, scheduling and optimization, emergency management, and decision making.)



陈曦 华中科技大学自动化学院系统科学与工程系副教授, 博士。2007 年毕业于华中科技大学。主要研究方向为复杂系统建模与仿真, 计算实验, 决策支持, 公共安全与应急管理。

E-mail: chenxi@hust.edu.cn

(**CHEN Xi** Ph.D., associate professor in the Department of Systems Science and Engineering, School of Automation, Huazhong

University of Science and Technology. He graduated from Huazhong University of Science and Technology in 2007. His research interest covers modeling simulation of complex systems, computational experiment, decision making support, public security, and emergency management.)