

基于缩减信念状态的 Conformant 规划方法*

魏唯^{1,2}, 欧阳彤^{1,2}, 吕帅^{1,2}

¹(吉林大学 计算机科学与技术学院, 吉林 长春 130012)

²(符号计算与知识工程教育部重点实验室(吉林大学), 吉林 长春 130012)

通讯作者: 欧阳彤, E-mail: ouyd@jlu.edu.cn

摘要: Conformant 规划问题通常转化为信念状态空间的搜索问题来求解. 提出了通过降低信念状态的不确定性来提高规划求解效率的方法. 首先给出缩减信念状态的增强爬山算法, 在此基础上, 提出了基于缩减信念状态的 Conformant 规划方法, 设计了 CFF-Lite 规划系统. 该规划器的求解过程包括两次增强爬山过程, 分别用于缩减信念状态和搜索目标. 首先对初始信念状态作最大程度的缩减, 提高启发函数的准确性, 然后从缩减后的信念状态开始执行启发式搜索. 实验结果表明, CFF-Lite 规划系统通过快速缩减信念状态降低了问题的求解难度, 在大多数问题上, 求解效率和规划解质量与 Conformant-FF 相比, 都有显著的提高.

关键词: Conformant 规划问题; 信念状态; 增强爬山; 启发式搜索

中图法分类号: TP18 **文献标识码:** A

中文引用格式: 魏唯, 欧阳彤, 吕帅. 基于缩减信念状态的 Conformant 规划方法. 软件学报, 2013, 24(7): 1557-1570. <http://www.jos.org.cn/1000-9825/4289.htm>

英文引用格式: Wei W, Ouyang DT, Lü S. Conformant planning based on reducing belief states. Ruan Jian Xue Bao/Journal of Software, 2013, 24(7): 1557-1570 (in Chinese). <http://www.jos.org.cn/1000-9825/4289.htm>

Conformant Planning Based on Reducing Belief States

WEI Wei^{1,2}, OUYANG Dan-Tong^{1,2}, LÜ Shuai^{1,2}

¹(College of Computer Science and Technology, Jilin University, Changchun 130012, China)

²(Key Laboratory of Symbolic Computation and Knowledge Engineering, Ministry of Education (Jilin University), Changchun 130012, China)

Corresponding author: OUYANG Dan-Tong, E-mail: ouyd@jlu.edu.cn

Abstract: Conformant planning is usually transformed into a search problem in the space of belief states. In this paper, a method which can improve efficiency of planning by reducing the nondeterministic degree of belief states is proposed. An enforced hill-climbing algorithm for reducing belief states is presented first. Then, the method of Conformant planning based on reducing belief states is proposed. A planner named CFF-Lite implements this idea and is designed. The planner includes two phases of enforced hill-climbing which are used to reduce belief states and search the goal respectively. Before the search phase, the initial belief state is reduced furthest to an intermediate state which is much more deterministic. Next, the precision of heuristic information is improved and the heuristic search phase is performed. Experimental results show that the CFF-Lite planner can decrease the difficulty of Conformant planning problems by reducing belief states and with most of the test problems this method outperforms Conformant-FF in both planning efficiency and planning quality.

Key words: Conformant planning; belief state; enforced hill-climbing; heuristic search

* 基金项目: 国家自然科学基金(61133011, 60973089, 61003101, 61170092, 61272208); 国家教育部博士点专项基金(20100061110031); 吉林省科技发展计划(20101501, 20100185, 201101039); 浙江师范大学计算机软件与理论省级重中之重学科开放基金(ZSDZZZXK12); 浙江省自然科学基金(Y1100191)

收稿时间: 2011-07-20; 修改时间: 2012-05-18; 定稿时间: 2012-07-23

近十几年来,智能规划研究领域取得了巨大的突破,在求解规模与求解效率上都有数量级的提高^[1,2].为了提高智能规划系统的性能,研究人员提出了多种规划求解方法,例如基于问题转换的规划方法^[3,4]、基于启发式搜索的规划方法^[5-7]等,国内很多研究者对智能规划领域的研究也提出了不同的求解方法与策略^[8-10].经典的智能规划问题建立在强约束的基础上,对现实世界问题的限制过强,阻碍了智能规划在实际中的应用.在经典规划的基础上,不确定性规划通过扩展规划问题域定义、修改操作规则、设置规划运行的环境参数和添加辅助信息等方法发展起来,具有更广泛的实际应用前景.Conformant 规划是一类不确定性规划,能够更好地刻画智能体对现实世界认知的不完备,要求在初始状态不确定或动作效果不确定并且假设智能体对周围环境无任何观测能力的情况下求出规划解.规划解的形式是线性的动作序列,确保在任意可能的初始状态上执行这一动作序列都能够到达目标状态^[11].

Conformant 规划问题通常转化为信念状态空间上的搜索问题,设计相应的启发式策略以提高求解性能.搜索空间中的元素不是具体的世界状态,而是表示当前所有可能世界状态的信念状态.Bonet 与 Geffner 最早提出基于信念状态空间搜索的规划方法,由两人共同设计的规划系统 GPT 中显式地表示信念状态,列举所有可能的世界状态,使得系统性能严重受到问题规模的限制^[12].此后的一系列规划方法通常采用 BDDs 结构表示信念状态,存储空间上有一定的压缩,但 BDDs 结构的规模以及复杂的构建与操作步骤仍然限制了规划系统的求解能力^[13,14].Hoffman 与 Brafman 设计的 Conformant Fast Forward 规划系统(简称 Conformant-FF)采用了新颖的隐式信念状态表示方法,信念状态用描述动作执行序列的 CNF 公式来表示,牺牲存储空间换取计算信念状态的时间效率.Conformant-FF 规划系统建立在著名的启发式搜索规划系统 Fast Forward(简称 FF)基础之上,利用了 FF 系统中的增强爬山算法、放松规划解启发式和帮助动作剪枝策略等核心框架,在国际规划比赛上表现突出,取得了优异的成绩^[15].

然而,Conformant-FF 系统在一些测试域上存在爬山算法效率不高甚至失败的情况,此时,调用最好优先搜索算法重新求解,使得整体性能有所下降.造成 Conformant-FF 规划系统性能不一致的原因在于信念状态中不确定信息所占的比重较大,对信念状态进行代价估值时,由不确定信息上的蕴含关系给出的帮助动作在一些问题上存在错误引导的情况,使得信念状态的不确定性无法下降,导致爬山算法始终无法快速搜索到目标状态,规划解的质量通常较差.针对这一局限性,本文提出降低信念状态不确定性的规划方法,求解过程包含两次增强爬山过程,分别用于降低信念状态的不确定性和搜索目标状态.首先对初始信念状态进行缩减,利用增强爬山算法将初始信念状态转换为完全确定的世界状态或不确定性更低的信念状态;然后,再从这个中间状态开始搜寻目标.我们在 Conformant-FF 规划系统的基础上实现了该方法,设计了高效的 CFF-Lite 规划系统,实验结果表明,CFF-Lite 规划系统通过快速缩减信念状态的不确定性,降低了问题的求解难度,在大多数问题上,求解效率和求解质量与 Conformant-FF 相比,都有显著的提高.

1 Conformant 规划的相关概念

本文考虑初始状态带有不确定性的 Conformant 规划问题,允许动作带有条件效果,对动作效果不确定性的处理方法可以通过计算信念状态转移时考虑不确定效果进行简单的扩展.

1.1 Conformant 规划问题

定义 1^[15]. Conformant 规划问题是一个三元组 (A, I, G) ,其中, A 为所有动作的集合; I 为 CNF 公式,表示的是初始状态需要满足的条件; G 为目标条件.

Conformant 规划问题不指定任何具体的世界状态为初始状态,而是通过一个 CNF 公式 I 给出初始状态需要满足的条件,任意一个满足公式 I 的世界状态均是可能的初始状态.用 2^I 表示所有可能的初始状态的集合,构成初始信念状态.动作 $a \in A$ 以 $(pre(a), E(a))$ 的形式给出, $pre(a)$ 表示动作 a 的前提条件, $E(a)$ 为动作 a 所有条件效果的集合;条件效果 $e \in E(a)$ 的形式为 $(con(e), add(e), del(e))$,其中, $con(e)$ 为条件效果 e 的发生条件, $add(e)$ 和 $del(e)$ 分别为条件效果 e 的添加效果和删除效果.集合 A 中的动作均不指定任何执行代价,即所有动作默认为统一的单元执行代价.

定义 2^[15]. 对于 Conformant 规划问题 (A, I, G) ,称 $Act = \langle a_1, \dots, a_n \rangle$ 为动作集合 A 上的动作序列,其中, $a_i \in A$, $i \in \{1, \dots, n\}$,记为 $Act \in A^*$.

定义 3^[15]. 对于信念状态 S 以及动作序列 $Act \in A^*$,如果 Act 在每一个世界状态 $w \in S$ 上都可执行,则称动作序列 Act 在信念状态 S 上可执行.

Conformant 规划问题的解是一个在初始信念状态上可执行的动作序列,从任意可能的初始状态 $w_0 \in 2^I$ 开始执行这一动作序列,都能够实现目标条件.由于 Conformant 规划解的形式是线性的动作序列,当动作不涉及执行代价时,将规划解的长度(动作序列中所包含的动作个数)作为规划解质量的衡量标准.

1.2 隐式信念状态表示

Conformant-FF 规划系统在信念状态空间上执行前向搜索,求解给定的规划任务,采用隐式信念状态表示方法.对于一个动作序列 Act 和一个命题 P ,判断在任意一个可能的初始状态上执行动作序列 Act 之后命题 P 是否一定成立,由此获得动作序列 Act 所到达的信念状态,相关定义如下:

定义 4^[15]. 对于 Conformant 规划问题 (A, I, G) 、信念状态 S 以及到达该信念状态的动作序列 $Act \in A^*$,如果命题 P 在任意可能的初始状态 $w_0 \in 2^I$ 上执行 Act 后所得世界状态上都成立,那么命题 P 是 S 上的 Known 命题;如果命题 P 在从任意可能的初始状态 $w_0 \in 2^I$ 上执行 Act 后所得世界状态上都不成立,那么命题 P 是 S 上的 Negatively Known 命题.

信念状态 S 上的 Known 命题表示该命题在每一个世界状态 $w \in S$ 上都成立,是 S 上确定为真的命题,记为 $Known(S)$;Negatively Known 命题表示该命题在每一个世界状态 $w \in S$ 上都不成立,是 S 上确定为假的命题,记为 $Negatively_Known(S)$.可见,Known 命题与 Negatively Known 命题与初始状态的不确定性无关,无论从哪个初始状态出发,执行 Act 后 $Known(S)$ 与 $Negatively_Known(S)$ 集合上的命题在信念状态 S 上的取值都确定为真和为假.对于一个命题 P ,如果 $P \notin Known(S)$ 且 $P \notin Negatively_Known(S)$,那么 P 为信念状态 S 上的 Unknown 命题.

定义 5^[15]. 对于 Conformant 规划问题 (A, I, G) 、信念状态 S 以及到达该信念状态的动作序列 $Act \in A^*$,如果命题 P 在某些可能的初始状态 $w_0 \in 2^I$ 上执行 Act 后所得世界状态上成立,但在其余初始状态 $w'_0 \in 2^I$ 上执行 Act 后所得世界状态上不成立,那么命题 P 是 S 上的 Unknown 命题.

与 Known 命题和 Negatively Known 命题不同,信念状态 S 上的 Unknown 命题与具体从哪一个初始状态开始有关.根据初始状态的不同,命题在当前时刻既可能为真也可能为假.因此从信念状态的角度上看,命题为真和为假这两种真值均属于信念状态 S ,记为 $Unknown(S)$.

由以上命题取值的定义可知,在 Conformant 规划问题中,命题在信念状态上的真值包括 True, False 和 Unknown 这 3 种,分别对应 Known, Negatively Known 和 Unknown 命题,这些命题的取值可以通过对表示初始信念状态和动作序列的 CNF 公式进行推理得到.这类 CNF 公式 $\Phi(Act)$ 包括初始状态公理、 Act 中各个动作的添加效果公理、删除效果公理和框架公理,公式 $\Phi(Act)$ 刻画了在初始信念状态上执行动作序列 Act 的语义.

定义 6^[15]. 对于 Conformant 规划问题 (A, I, G) 、信念状态 S 以及到达该信念状态的长度为 n 的动作序列 Act ,命题 P 是信念状态 S 上的 Known 命题当且仅当 $\Phi(Act) \Rightarrow P(n)$,其中, $P(n)$ 表示命题 P 在第 n 个时间步的真值.

Conformant-FF 规划系统采用基于 SAT 的求解方法计算信念状态上所有命题的真值,计算过程如下:

- 对于动作序列 Act 以及每一个命题 P 调用一次 SAT 求解器,判定公式 $\Phi(Act) \wedge \neg P(n)$ 的可满足性,当判定结果为不可满足时,表明 $\Phi(Act) \Rightarrow P(n)$ 是恒成立的,那么命题 P 在所得信念状态上确定为真,是 Known 命题;
- 同理,调用 SAT 求解器对公式 $\Phi(Act) \wedge P(n)$ 的可满足性进行判定以计算 Negatively Known 命题集合.两者都不成立的命题即为 Unknown 命题.

由此,Conformant-FF 系统采用表示初始状态公式与动作序列的 CNF 公式来隐式地表示信念状态.对于一个给定的规划问题 (A, I, G) ,任意可能的初始状态均由空的动作序列 $Act = ()$ 到达,因此,表示 Act 执行语义的 CNF 公式 $\Phi(Act)$ 只包含描述初始条件的公式 I .调用上述 SAT 判定过程即可得到初始信念状态上的所有 Known 命题、Negatively Known 命题以及 Unknown 命题,本文以下用 S_I 标记由 SAT 方法计算得到的初始信念状态.

2 基于缩减信念状态的 Conformant 规划算法

在利用启发式搜索方法求解 Conformant 规划问题时,启发信息需要对当前信念状态进行评价,对当前所有可能的世界状态给出一个统一的代价估值.为了保证足够的信息量,启发函数的设计需要充分考虑信念状态中的所有不确定性信息,以确保所得启发信息适用于当前所有可能的世界状态.Conformant-FF 规划系统采用问题的放松规划解作为启发信息,在放松规划图中额外加入了 Unknown 命题层.由于 Unknown 命题可以包含两种真值,并不确定为真或为假,因此,应用在 Unknown 命题上的动作给出了相邻时间步之间 Unknown 命题的蕴含关系.即执行动作时,前一层上某一个 Unknown 命题所取的真值将影响下一层上某些 Unknown 命题的取值.扩展放松规划图时,同时添加 Unknown 命题上的蕴含关系,根据这些约束关系推导哪些 Unknown 命题通过执行动作能够确定为 Known 命题;提取放松规划解时,同时回溯放松规划图并判断这些蕴含关系对于目标的实现,得到的放松规划解中包含确定的命题层上给出的动作和 Unknown 命题层上导致蕴含关系的动作.

很多 Conformant 规划问题在搜索过程中,信念状态始终具有较强的不确定性,启发信息主要来源于信念状态中不确定的部分,大部分帮助动作由 Unknown 命题上扩展出的蕴含关系给出.而 Conformant-FF 在处理 Unknown 命题之间的约束关系时作了进一步的放松处理,当一个动作的条件效果中包含多个 Unknown 的发生条件时,只任选其中一个 Unknown 条件,插入这个条件命题及其添加效果之间的蕴含关系,忽略其余的 Unknown 条件.由此,启发式评估过程得到了简化.但是,随着放松规划图的扩展,逐渐会有更多的信息丢失,启发信息无法准确估计出当前任意可能的世界状态与目标状态之间的代价估值.不可靠的启发信息和帮助动作将会影响求解效率与规划解质量,甚至可能导致增强爬山算法搜索目标失败.实际上,在很多 Conformant 规划问题中,当信念状态的不确定性过强时很难给出合适的启发信息,不适宜立即执行启发式搜索过程.降低信念状态的不确定性有助于降低问题的求解难度,提高启发信息的准确性.

2.1 缩减信念状态的增强爬山算法

本文第 1 节给出了信念状态的隐式表示方法,其中,Unknown 命题刻画了信念状态的不确定性.当 Unknown 命题数目越多时,信念状态的不确定性越强,所代表的可能世界状态数目越多;Unknown 命题个数越少时,所代表的可能世界状态数目越少;当 Unknown 命题数目为 0 时,所有的命题都有一个确定的真值,信念状态为普通的世界状态.研究发现,很多 Conformant 规划问题中的信念状态可以通过执行规划动作减少信念状态中 Unknown 命题数,Unknown 命题数目降低时,Known 与 Negatively Known 命题个数增加,信念状态中确定信息的部分变大,不确定性有所下降.当 Known 命题包含所有目标条件时,搜索过程成功到达目标状态.

根据增强爬山算法的高效性,本节提出利用增强爬山过程缩减隐式信念状态的方法,下面首先给出缩减信念状态的相关定义.

定义 7. 对于信念状态 S 以及 $Unknown(S)$,称 $num_U(S)=|Unknown(S)|$ 为信念状态 S 的 Unknown 命题数.

定义 8. 对于信念状态 S 以及在 S 上可执行的动作序列 Act ,如果在 S 上执行 Act 后所得信念状态 S' 满足条件: $num_U(S') < num_U(S)$,则称动作序列 Act 缩减了信念状态 S .

将信念状态 S 上对 Unknown 命题数目的度量 $num_U(S)$ 作为评价函数,从初始信念状态开始,爬山过程中反复搜索更优的信念状态.考虑在缩减信念状态的过程中到达目标状态的情况,在每一次迭代过程中加入目标条件的判断,当在增强爬山过程中搜索到 Unknown 命题数目更小的信念状态时,判断当前信念状态是否已实现了所有目标条件.如果缩减过程始终没有到达某一目标状态且 Unknown 命题数无法进一步缩减,那么算法到达一个不可缩减的中间状态,记为 $Inter_S$.

定义 9. 对于信念状态 S ,如果不存在任何动作序列 $Act \in A^*$ 使得执行 Act 后所得信念状态 S' 满足 $num_U(S') < num_U(S)$,则称信念状态 S 是不可缩减的.

将初始信念状态逐步缩减为中间状态 $Inter_S$ 的增强爬山过程如下所示:

算法 1. $Ehc_to_Inter_S()$.

1. 初始化: S 置为初始信念状态 S_I ,计算 $num_U(S)$;

2. 反复执行步骤 2.1~步骤 2.5,直到 $num_U(S)=0$:
 - 2.1. 以 S 为初始节点执行宽度优先搜索,寻找满足 $num_U(S')<num_U(S)$ 的信念状态 S' ;
 - 2.2. 如果不存在满足条件的信念状态,则转步骤 3;
 - 2.3. 找到满足条件的信念状态 S' ,把将 S 缩减为 S' 的动作序列追加到 Act 之后;
 - 2.4. 更新当前信念状态,令 $S=S'$;
 - 2.5. 若目标条件 $G \subseteq Known(S), Found_plan=True$,则转步骤 3;
3. 令 $Inter_S=S$,返回.

算法 1 给出了缩减隐式信念状态的增强爬山算法,从初始信念状态开始执行,利用 $Found_plan$ 记录在缩减信念状态的增强爬山过程中是否已找到规划解.如果到达目标状态,搜索过程结束,不再继续执行缩减信念状态的过程;当前信念状态 S 没有到达任何目标状态时,从 S 开始执行宽度优先搜索,寻找满足 $num_U(S')<num_U(S)$ 的信念状态 S' .当搜索过程无法找到满足条件的信念状态时,表明当前信念状态无法进一步缩减,否则,增强爬山过程找到第 1 个 Unknown 命题的数目更小的信念状态 S' ,此时认为搜索过程中从信念状态 S 到达 S' 的路径是有效的,将这部分路径上的所有动作追加到当前所求得的动作序列之后.根据增强爬山算法的思想,更新当前信念状态 S 为 S' ,清除先前迭代过程的搜索信息,从信念状态 S' 开始重新搜索 Unknown 命题数目更小的信念状态.反复执行这一过程,直到增强爬山过程实现目标条件或到达一个不可缩减的中间状态 $Inter_S$.至此,缩减信念状态的增强爬山过程结束.此后,规划算法从这个中间状态开始继续求解规划任务,这就是基于缩减信念状态的 Conformant 规划方法的基本思想.算法 1 给出的 $Ehc_to_Inter_S()$ 是规划算法中用于缩减信念状态的子过程,所用到的 $Found_plan$ 和 Act 等标识是从全局求解过程中接收的全局变量,详细说明将在第 2.2 节中给出.

下面通过一个简单的积木世界问题描述算法 1 缩减信念状态的过程.

例 1:Block 域规划问题描述的是一类按照指定的目标形态排列积木块的问题,域中给出的可用动作如下:

- $(move-b-to-b \ ?bm \ ?bf \ ?bt)$: $pre: (clear \ ?bm), (clear \ ?bt), (on \ ?bm \ ?bf)$;
 $e1: con: \emptyset; add: (on \ ?bm \ ?bt), (clear \ ?bf); del: (clear \ ?bt), (on \ ?bm \ ?bf)$;
- $(move-b-to-t \ ?bm \ ?bf)$: $pre: \emptyset$;
 $e1: con: (on \ ?bm \ ?bf); add: (on-table \ ?bm), (clear \ ?bf); del: (on \ ?bm \ ?bf)$;
- $(move-t-to-b \ ?bm \ ?bt)$: $pre: (clear \ ?bm), (clear \ ?bt), (on-table \ ?bm)$;
 $e1: con: \emptyset; add: (on \ ?bm \ ?bt); del: (clear \ ?bt), (on-table \ ?bm)$.

所有动作均包含一个条件效果,其中,动作 $(move-b-to-b \ ?bm \ ?bf \ ?bt)$ 表示将积木块 $?bm$ 从积木块 $?bf$ 上移到积木块 $?bt$ 上, $(move-b-to-t \ ?bm \ ?bf)$ 表示将积木块 $?bm$ 从积木块 $?bf$ 上移到桌子上, $(move-t-to-b \ ?bm \ ?bt)$ 表示将积木块 $?bm$ 从桌子上移到积木块 $?bt$ 上.

图 1 给出的问题实例包括 $B1, B2, B3$ 这 3 个积木块和一个桌子 $Table$,具体的问题描述如下:

- 初始状态信念:
 $I = \{(on-table \ B2), (on-table \ B3), (clear \ B1), (on \ B1 \ B2) \vee (on \ B1 \ B3), \neg (on \ B1 \ B2) \vee \neg (on \ B1 \ B3), (on \ B1 \ B2) \vee (clear \ B2), \neg (on \ B1 \ B2) \vee \neg (clear \ B2), (on \ B1 \ B3) \vee (clear \ B3), \neg (on \ B1 \ B3) \vee \neg (clear \ B3)\}$.
- 目标条件: $G = \{(on \ B3 \ B2), (on \ B2 \ B1), (on-table \ B1)\}$.



Fig.1 A planning problem from the Block domain

图 1 Block 域规划问题

公式 I 即为初始状态各个积木块必须满足的条件,给出了初始时确定成立的事实以及不确定信息之间的约束关系,其中, B_2 和 B_3 的位置是确定的, B_1 可能位于 B_2 或 B_3 上,并且约束两者只能有一个成立.可以看出,无论 B_1 的初始位置如何,求解上述规划问题首先要将 B_1 的位置确定,把 B_1 移到桌子上,然后再进行 B_2 和 B_3 的移动.下面通过例 1 描述算法 1 的具体实施过程,对于各个信念状态只列出了 Known 命题与 Unknown 命题,其余未出现的命题均为 Negatively Known 命题,由于 $Negatively_Known(S)$ 数目过多,这里不再一一列举.

通过 SAT 判定方法对初始条件公式给出的约束关系进行推理,得到初始信念状态为

$$S_i: Known(S_i) = \{(on-table B_2), (on-table B_3), (clear B_1)\};$$

$$Unknown(S_i) = \{(on B_1 B_2), (on B_1 B_3), (clear B_2), (clear B_3)\}.$$

算法 1 只关注 Unknown 命题的个数,得到 $num_U(S_i) = 4$.从初始信念状态开始执行增强爬山过程,初始时 $Act = \langle \rangle$.由定义 3 可知,信念状态 S 上的可执行动作为所有前提包含于 $Known(S_i)$ 的动作,扩展初始信念状态时可执行动作为 $(move-b-to-t B_1 B_2)$ 和 $(move-b-to-t B_1 B_3)$.假设算法先评估执行 $(move-b-to-t B_1 B_2)$ 后得到的状态 S' ,调用 SAT 求解器计算动作效果导致的状态转移,得到信念状态 S' 为

$$S': Known(S') = \{(on-table B_2), (on-table B_3), (clear B_1), (clear B_2)\};$$

$$Unknown(S') = \{(on B_1 B_3), (clear B_3), (on-table B_1)\}.$$

执行这个动作将 B_1 位于 B_2 上的情况排除, $num_U(S') = 3$, Unknown 命题数降低.此时, B_1 确定不在 B_2 上,但可能位于 B_3 或桌子上.由此得知, $(move-b-to-t B_1 B_2)$ 动作缩减了初始信念状态,将其追加到 Act 之后.从缩减后的信念状态 S' 开始重复这一过程,提取出动作 $(move-b-to-t B_1 B_3)$,排除 B_1 位于 B_2 上的情况,进一步将信念状态的 Unknown 命题数降为 0.此时,得到不可缩减的中间状态 $Inter_S$.

$$Inter_S: Known(Inter_S) = \{(on-table B_1), (clear B_1), (on-table B_2), (clear B_2), (on-table B_3), (clear B_3)\};$$

$$Unknown(Inter_S) = \emptyset.$$

在中间状态 $Inter_S$ 上, B_1, B_2, B_3 全部位于桌子上,这是一个确定的世界状态,缩减过程得到下面的动作序列:

$$Act = \langle (move-b-to-t B_1 B_2), (move-b-to-t B_1 B_3) \rangle.$$

缩减信念状态的增强爬山算法充分利用了隐式信念状态表示方法的优势,具有如下优点:1) 能够直接获取当前信念状态上的 Unknown 命题数,不需要增加额外的计算量;2) 增强爬山过程中采用宽度优先搜索寻找更优的节点,能够保证满足条件的信念状态距离当前信念状态最近;3) 借助于 SAT 求解的高效性,能够快速给出逐步缩减信念状态的动作序列;4) 如果在缩减信念状态的过程中就已经到达目标,那么算法 1 能够直接给出规划解,避免了多次计算启发式估值的过程,从而减少了计算量.

2.2 CFF-Lite 规划系统

本文在 Conformant-FF 规划系统的框架之上实现了基于缩减信念状态的 Conformant 规划方法,设计了相应的 CFF-Lite 规划系统,基本结构如图 2 所示.对于给定的 Conformant 规划问题,首先执行第 2.1 节给出的增强爬山过程,对初始信念状态 S_i 进行缩减,根据 $Found_plan$ 标识判断缩减过程是否到达了目标状态.当缩减过程没有完成规划任务的求解时,得到中间状态 $Inter_S$,根据 Unknown 命题数判断 $Inter_S$ 不确定的程度,采用不同的规划方法.当中间状态 $Inter_S$ 为确定的世界状态时(即 $num_U(Inter_S) = 0$ 成立),剩余的规划任务转变为经典规划任务,以 $Inter_S$ 为初始状态,调用经典规划方法进行求解(本文选用 FF 规划系统);当中间状态 $Inter_S$ 为信念状态时,剩余的规划任务仍为 Conformant 规划任务,以 $Inter_S$ 为初始信念状态,调用 Conformant_FF 规划系统求解.

下面具体给出基于缩减信念状态的 Conformant 规划方法的求解过程:

算法 2. $CFF-Lite(A, I, G)$.

输入: Conformant 规划问题 (A, I, G) ;

输出: 规划解 Act .

全局变量: $Found_plan, Act, Inter_S$.

1. 初始化全局变量: $Found_plan=False;Act=();Inter_S$ 置为空;
2. 对公式 I 调用 SAT 求解器计算初始信念状态 S_i ;
3. 缩减信念状态:执行增强爬山过程 $Ehc_to_Inter_S()$;
4. 若 $Found_plan=True$,则已到达目标状态,算法结束,输出规划解 Act ;
5. 若 $Found_plan=False$,得到剩余规划问题 $(A,Inter_S,G)$,判断中间状态 $Inter_S$ 中是否包含 Unknown 命题;
 - 5.1. 当 $num_U(Inter_S)=0$ 时,调用 FF 规划系统求解剩余规划任务 $(A,Inter_S,G)$;
 - 5.2. 当 $num_U(Inter_S)\neq 0$ 时,调用 Conformant_FF 规划系统求解剩余规划任务 $(A,Inter_S,G)$;
6. 若 $Found_plan=True$,返回规划解 Act ;否则,失败退出.

算法 2 的输入为 Conformant 规划问题 (A,I,G) ,输出为规划解 Act ,需要使用的全局变量包括用于存储规划解的动作序列 Act 、缩减初始信念状态所得中间状态 $Inter_S$ 和记录是否成功求解的变量 $Found_plan$.首先对全局变量进行初始化,并计算初始信念状态 S_i ,以便获取初始时 Unknown 命题的个数,用于缩减信念状态的子过程.如果执行 $Ehc_to_Inter_S()$ 后 $Found_plan$ 为真,缩减过程所求得的动作序列 Act 即为有效的规划解,算法成功结束;否则,从 $Inter_S$ 和当前动作序列 Act 开始继续求解,所求得的规划动作追加在 Act 之后,构成最终的规划解.

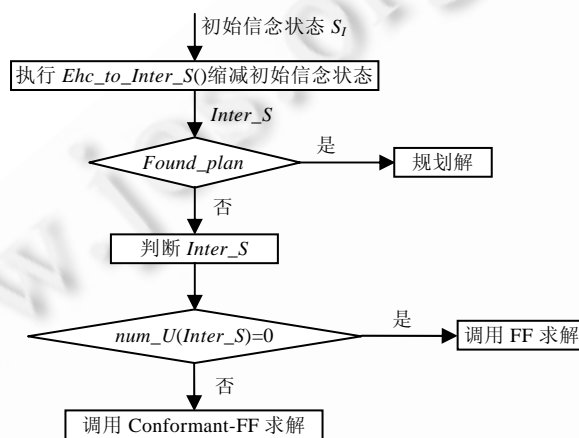


Fig.2 The basic framework of the CFF-Lite planner

图 2 CFF-Lite 规划系统的基本结构

对于例 1 给出的规划问题,图 3 为算法 2 求解的主要步骤,执行第 2.1 节中缩减信念状态的过程后,得到世界状态 $Inter_S=\{(on-table\ B1),(clear\ B1),(on-table\ B2),(clear\ B2),(on-table\ B3),(clear\ B3)\}$.然后,从这一状态继续完成求解,利用 FF 规划系统得到最终规划解为

$$Act=\langle (move-b-to-t\ B1\ B2),(move-b-to-t\ B1\ B3),(move-t-to-b\ B2\ B1),(move-t-to-b\ B3\ B2)\rangle.$$

当然,Conformant-FF 规划系统搜索时也必将经过这个中间状态 $Inter_S$,两者的主要区别在于所使用的启发信息不同,算法 2 将不确定信息的处理过程单独提取出来,先由 Unknown 命题数来引导爬山过程,然后再去完成目标条件的实现;而 Conformant-FF 求解时将直接计算初始信念状态的启发式估值,根据各个访问状态的启发式估值反复爬山.对于上述规划问题,算法 2 的求解过程只对搜索空间的两个世界状态计算启发式估值,Conformant-FF 完成求解任务需要经过 4 次启发式评估过程,其中包括两次对信念状态计算启发式估值.

由上述算法给出的求解过程可见,对于给定的 Conformant 规划任务,CFF-Lite 规划系统不直接搜索目标状态,而是先尽量降低信念状态的不确定性,再执行目标状态的搜索过程,利用启发信息策略引导搜索过程.信念状态的不确定性降低后,启发信息中确定性信息的比重增加,使得启发信息估算目标距离的准确性有所提高,避免发生错误引导的情况.当中间状态降低为普通的世界状态时,剩余搜索空间中所有的状态节点均为确定的世

界状态,利用 FF 放松规划图启发函数进行代价估值,搜索过程不包含任何不确定信息;当中间状态仍为信念状态时,剩余任务的搜索空间中的元素仍包含信念状态,但在信念状态缩减后启发式评估过程中 Unknown 命题上蕴含关系的数量有所降低,减少了对 Unknown 命题之间约束关系的推理.此时,启发信息和帮助动作更多地来自于信念状态中 Known 命题部分,能够更准确地引导搜索过程.

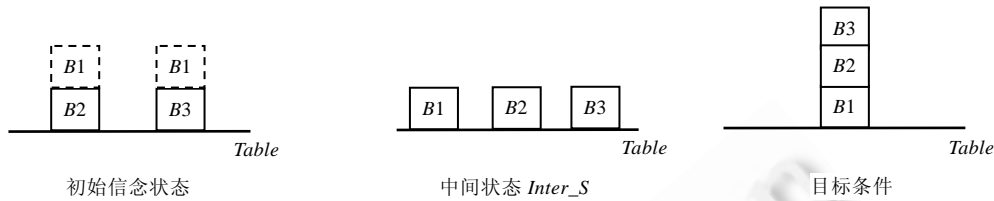


Fig.3 The main steps of Alrorthm 2

图3 算法2的主要求解步骤

关于搜索算法的计算复杂度,利用启发信息引导搜索过程是克服状态空间组合爆炸问题的有效策略,但在很多情况下,计算启发式信息是比较耗时的过程.因此,启发式评估的计算代价与准确度直接影响了搜索的整体效率.Conformant-FF 规划系统采用了基于放松规划解的启发式设计思想,计算放松规划解的过程更为复杂,需要充分考虑初始状态的不确定性对于提取放松规划解的影响.如果希望得到尽可能多的信息,保留完整的初始条件公式,直接调用 SAT 求解器来计算放松规划解,那么不能够保证启发式估值的准确性,并且最坏情况下,SAT 判定过程可能是指数级的;如果对初始条件进行适当的放松,启发函数可以转化为 2-CNF 推理问题,放松规划任务更容易求解,计算复杂度在最坏情况下是多项式时间的,但信息量将大量流失.Conformant-FF 实际的评估过程在计算代价与信息量之间加以折中,获取启发信息的复杂度介于上述两者之间.本文提出缩减信念状态方法的目的就是克服 Conformant-FF 难以快速、准确地计算启发式估值的问题,规划求解过程先对 Conformant 问题进行简化,尽量避免不确定信息对启发式评估过程的影响.CFF-Lite 的第 1 次爬山过程逐渐降低信念状态的规模,可以将 Unknown 命题数看做是引导搜索过程的启发信息.虽然 Unknown 命题之间存在约束关系,不能单独凭借其个数来断定信念状态中所包含的状态数,但 Unknown 命题数仍然能够合理地度量信念状态之间不确定程度的比较,启发信息的可信度较高.缩减信念状态后,如果得到了确定的世界状态,那么,第 2 次爬山过程只需要进行 FF 启发式估值,FF 系统给出的放松规划解是近似最优的启发式,且能够在多项式时间内求解.由此可见,与 Conformant-FF 相比,CFF-Lite 的两部分爬山过程都采用了准确度更高且计算复杂度更低的启发函数.此外,搜索算法的计算量还可以通过搜索空间的规模来度量,在计算放松规划解启发式时,给出的帮助动作能够有效地修剪搜索空间.由于 Unknown 命题数可以直接获取,CFF-Lite 第 1 次爬山时的计算量主要是展开搜索空间,缩减信念状态的过程在扩展状态时没有任何可用的修剪策略,需要生成所有可执行动作产生的后继状态,当分支数较大时,访问的状态数增多,对搜索算法的性能有一定的影响.然而,观察放松规划解启发式的评估过程可以发现,信念状态上的可执行动作中通常包含处理 Unknown 命题的条件效果,扩展信念状态时,这一类条件效果所在的动作分支能够帮助爬山算法快速获得 Unknown 命题数更少的后继,当初始信念状态包含的不确定信息较多时,缩减信念状态的爬山过程需要经过多次迭代才能够完全消除信念状态的 Unknown 命题数.因此,初始信念状态的 Unknown 命题数给出了爬山过程迭代次数的上限.状态空间的规模对搜索效率的影响与具体给定的规划问题密切相关,设计由领域专门知识给出的剪枝策略能够帮助压缩搜索空间,这也是我们未来的研究工作中将要解决的问题.

3 实验结果与分析

本文以 Conformant-FF 规划系统为基础,实现了基于缩减信念状态的 Conformant 规划系统 CFF-Lite,采用具有代表性的国际通用 Conformant 规划域:Cube, Ring, Bomb 和 Coins 进行测试,主要从求解效率和规划质量两个方面比较 Conformant-FF 和 CFF-Lite 规划系统.在以下各表中, T 表示求解的总时间(单位:秒); S 为求解过程所

访问的状态数,CFF-Lite 访问的状态包括缩减信念状态过程中评估 Unknown 命题数的信念状态以及第 1 节完成规划求解的过程中计算启发式估值的状态,Conformant-FF 访问的状态是求解过程中所有进行启发式评估的状态; L 为规划解的长度,用于评价规划解的质量;“-”表示求解失败(时间限定为 1 800s).实验测试的机器配置如下:操作系统 Ubuntu 8.0,处理器 1.90GHz,内存 1.0GB.

Conformant-FF 规划系统中设置了 3 种不同的启发函数 h^l, h^{inc} 和 $h^{\phi(Act)}$,主要区别在于对蕴含图的推理部分^[15]: h^l 利用初始信念状态公式作蕴含图的推理; h^{inc} 与 h^l 类似,但推理时用简单的集合包含关系取代逻辑蕴含关系; $h^{\phi(Act)}$ 利用当前访问的信念状态公式作蕴含图的推理.多数情况下, h^l 与 h^{inc} 显示的求解效率与规划质量基本一致,因此本节主要对 CFF-Lite 规划系统与 Conformant-FF 规划系统采用启发函数 h^l 和 $h^{\phi(Act)}$ 时的测试结果进行比较.

3.1 Cube域

Cube 是一个经典的 Conformant 规划域,描述一个边长为 n 的三维立方体.初始状态时,智能体可能位于立方体中任意一个位置,每一维坐标轴上包括向上和向下两个操作,用于改变当前坐标值.根据目标状态的位置将测试问题分为 Cube-center 和 Cube-corner,分别表示目标状态的坐标位于立方体的中心和立方体的一个角落.

对于规模为 n 的立方体,由于初始时 X 轴、 Y 轴、 Z 轴坐标均可位于 $\{1, \dots, n\}$ 上任意一点,每个坐标轴都有 n 个可能的取值,因此,初始信念状态中所有描述坐标轴取值的命题都为 Unknown 命题,个数为 $3n$.实验发现:CFF-Lite 求解此类问题时,首先反复执行向上或向下动作,将各个坐标轴上可能取值的个数逐渐降低为 $n-1, n-2, \dots, 2, 1$.例如:在 X 轴上如果执行 $n-1$ 次向上动作, X 轴上的坐标值将确定为 n ;如果执行 $n-1$ 次向下动作, X 轴上的坐标值将确定为 0.通过这样的缩减过程,CFF-Lite 得到立方体边缘上一个确定的中间状态,使问题得到简化.CFF-Lite 求解 Cube 域规划问题的优势是缩减信念状态的爬山过程,每执行一次动作都可能得到 Unknown 命题数更小的信念状态,能够迅速降低问题的不确定性,只需要执行 $3(n-1)$ 个动作即可消除所有 Unknown 命题.与 Conformant-FF 规划系统相比,CFF-Lite 没有对任何信念状态进行启发式评估,不需要利用启发信息探索信念空间,缩减过程结束后利用 FF 启发式引导搜索,搜索空间中的元素都是确定的世界状态,启发式估值的计算量小,准确性较高.

表 1 为不同规模的 Cube-center 测试问题上的测试结果,Cube-center- n 表示边长为 n 的立方体问题.对比表 1 中的测试结果可知,Conformant-FF 系统对于 Cube-center 问题的求解能力很差.在采用 h^l 启发函数时,只能对规模较小的问题成功求解,问题规模增加时,搜索空间迅速膨胀,启发信息无法正确引导搜索过程,不能在规定时间内求出规划解; $h^{\phi(Act)}$ 启发函数的求解效率优于 h^l ,但是,随着问题规模的增大,求解效率明显下降,启发信息仍然没有起到很好的引导作用,搜索过程访问状态数目过多,最终求得的规划解质量较差.CFF-Lite 显示的求解能力明显优于 Conformant-FF,由于目标条件给出的是一个完整的世界状态,因此,CFF-Lite 首先对初始信念状态进行缩减,得到确定的中间状态后再考虑目标条件的实现,搜索方向更明确,所访问状态的数目较少;从时间效率上看,CFF-Lite 系统的求解效率不受问题规模的限制,始终能够快速求出规划解;从规划质量上看,CFF-Lite 系统求得的规划解长度远远小于 Conformant-FF 系统,不包含任何冗余动作,规划解质量更高.

Table 1 The results of Cube-center problems from the Cube domain

表 1 Cube 域中 Cube-center 问题实验结果

Problem	CFF-Lite	CFF $h^{\phi(Act)}$	CFF h^l
	T/S/L	T/S/L	T/S/L
Cube-center-3	0.00/17/9	0.03/61/9	0.02/93/15
Cube-center-5	0.02/32/18	0.40/262/30	13.50/2211/45
Cube-center-7	0.07/47/27	5.49/825/55	-
Cube-center-9	0.17/62/36	77.13/2052/97	-
Cube-center-11	0.40/77/45	1104.74/4913/147	-
Cube-center-13	0.93/92/54	-	-
Cube-center-15	1.41/107/63	-	-

表 2 为 Cube-corner 问题上的测试结果.从表中测试结果可知,Conformant-FF 系统采用 h^l 启发函数的效果

略优于 $h^{\phi(Ac)}$ 启发函数,求解速度更快,规划解长度更短.CFF-Lite 系统求得的规划解长度与 Conformant-FF 系统采用 h^l 启发函数求得的规划解长度一致,优于 $h^{\phi(Ac)}$ 启发函数.在执行缩减信念状态的增强爬山过程时,扩展状态考虑的是所有可用动作的分支,因此,CFF-Lite 求解过程访问的状态数目更多,但时间效率却普遍优于 Conformant-FF 系统.这是由于 CFF-Lite 借助于高效的 SAT 求解技术完成状态转移的计算,并且访问信念状态时能够直接获取 Unknown 命题数,整体上评估信念状态的计算代价更小.

Table 2 The results of Cube-corner problems from the Cube domain

表 2 Cube 域中 Cube-corner 问题实验结果

Problems	CFF-Lite	CFF $h^{\phi(Ac)}$	CFF h^l
	T/S/L	T/S/L	T/S/L
Cube-corner-3	0.00/14/6	0.01/7/7	0.01/6/6
Cube-corner-5	0.02/26/12	0.07/13/13	0.03/12/12
Cube-corner-7	0.06/38/18	0.24/19/19	0.12/18/18
Cube-corner-9	0.18/50/24	1.16/25/25	0.35/24/24
Cube-corner-11	0.46/62/30	2.65/31/31	0.79/30/30
Cube-corner-13	0.94/74/36	6.60/37/37	1.45/36/36
Cube-corner-15	1.56/86/42	16.19/43/43	2.10/42/42

3.2 Ring域

Ring 域也是一个常见的 Conformant 规划域, n 个连通的房间构成一个环形,执行 Move 动作在相邻房间之间来回移动,每个房间内有一个窗户,有“开着”、“关着”和“锁着”这 3 种状态,可以通过 Close 和 Lock 操作来改变窗户的状态.该域具体的每个问题实例在初始状态时智能体可能位于 n 个房间中的任意一个,每个房间内窗户的状态也是不确定的,可能处于 3 种状态中的任意一种;目标条件是每个房间的窗户都被锁上.表 3 给出了 Ring 域中不同测试问题上的测试结果,Ring- n 表示包含 n 个房间的规划问题.

Table 3 The results of the Ring domain——1

表 3 Ring 域实验结果——1

Problem	CFF-Lite	CFF $h^{\phi(Ac)}$	CFF h^l
	T/S/L	T/S/L	T/S/L
Ring-2	0.00/19/6	0.01/22/12	0.01/18/7
Ring-3	0.18/204/10	0.14/41/18	0.13/36/15
Ring-4	2.62/1490/14	-	-
Ring-5	69.09/10842/18	-	-

由于初始位置和各个房间内窗户的状态是不确定的,即使移动到相邻的房间,仍然无法进一步确定所处的位置和各个窗户的状态,因此,初始信念状态附近区域的各个信念状态始终保持较高的不确定性.实际上,该域上规划任务的目标是使各个窗户都达到“锁着”的状态,求解时并不需要确认当前的具体位置,正确的规划解只需要遍历 n 个房间,对每个窗户分别执行一次“关闭”动作和一次“锁上”动作即可实现目标.

从表 3 给出的测试结果可见,Conformant-FF 系统并不适宜求解 Ring 域上的规划问题,求解这类问题时,直接执行增强爬山过程通常会失败,然后再采用最好优先搜索的方法求解,只能对房间数目较小的问题成功求解;并且规划解质量不高,房间数目增加时,启发信息无法正确地引导搜索过程,系统丧失了求解能力,无法在有效时间内求得规划解.Conformant-FF 爬山失败的主要原因是启发信息不能挖掘问题的特殊结构,无法控制正确的搜索方向,爬山过程始终在较小的区域内徘徊,最终耗尽搜索空间,导致失败.

CFF-Lite 的处理方法是判断动作对信念状态上 Unknown 命题的改变,在执行第 2 节中算法 1 给出的缩减过程时,如果始终沿着一个方向依次遍历所有房间,并对每个窗户执行一次“关闭”动作,得到的信念状态上 Unknown 命题数则会有所减少,使得每一个窗户至少处于“关着”的状态;再沿同一方向循环一圈,对每个窗户执行一次“锁上”动作,所有窗户将达到目标条件要求的“锁着”的状态.此时,虽然智能体所处的位置仍不确定,信念状态无法再进一步缩减,但当前信念状态上的 Known 命题已经实现了所有目标条件.与 Conformant-FF 计算启发式估值过程相比,CFF-Lite 利用 Unknown 命题个数度量信念状态,不需要额外的计算量,能够在较短的时间内

访问大量的状态.从表3的实验结果上看,CFF-Lite 系统求解速度快,并且规划解的质量更好.由于 Ring 域的特殊结构, Conformant-FF 系统搜索过程中重复信念状态检测的环节成为了限制系统性能的主要因素,取消重复信念状态检测时,求解效率大幅度提高.对比结果见表4,Conformant-FF 系统的求解效率显然优于 CFF-Lite 系统,对各个问题均能成功求解,但规划解质量很差,启发函数没有起到很好的引导作用,导致规划解中包含很多冗余动作;而 CFF-Lite 系统始终能够快速求出质量更高的规划解.

Table 4 The results of the Ring domain——2

表4 Ring 域实验结果——2

Problem	CFF-Lite	CFF h^l -R
	T/S/L	T/S/L
Ring-2	0.00/19/6	0.00/18/7
Ring-3	0.18/204/10	0.04/50/20
Ring-4	2.62/1490/14	0.29/88/39
Ring-5	69.09/10842/18	1.79/139/64

3.3 Bomb域

Bomb 域是一个具有代表性的 Conformant 规划域,域中物体包括炸弹和厕所,炸弹分为武装状态和解除武装状态,厕所分为阻塞状态和非阻塞状态.炸弹和厕所的状态可以通过 Dunk 动作和 Flush 动作加以更改.不同测试问题下,炸弹和厕所的数目不同,Bomb- $b-t$ 表示炸弹数目为 b ,厕所数目为 t 的规划问题,每个测试问题在初始状态时所有炸弹的状态是不确定的,目标条件为所有炸弹处于解除武装状态.求解 Bomb 域规划问题的基本思路是,逐步消除各个炸弹的不确定性,当厕所数目较小时,需要交替执行 Dunk 动作和 Flush 动作来处理所有炸弹.表5为 Bomb 域中不同测试问题上的测试结果.

Table 5 The results of the Bomb domain

表5 Bomb 域实验结果

Problems	CFF-Lite	CFF $h^{(Act)}$	CFF h^l
	T/S/L	T/S/L	T/S/L
Bomb-5-1	0.00/11/9	0.00/19/9	0.01/19/9
Bomb-10-1	0.01/21/19	0.03/64/19	0.03/64/19
Bomb-20-1	0.00/41/39	0.53/229/39	0.25/229/39
Bomb-50-1	0.45/101/99	10.98/1324/99	5.07/1324/99
Bomb-100-1	5.39/201/199	-	129.70/5149/199
Bomb-5-5	0.01/21/5	0.01/5/5	0.00/5/5
Bomb-10-5	0.07/71/15	0.00/30/15	0.02/30/15
Bomb-20-5	0.41/171/35	0.32/155/35	0.17/155/35
Bomb-50-5	4.77/471/95	16.98/1130/95	4.67/1130/95
Bomb-100-5	65.49/971/195	-	120.99/4755/195
Bomb-5-10	0.03/21/5	0.00/5/5	0.00/5/5
Bomb-10-10	0.15/66/10	0.01/10/10	0.00/10/10
Bomb-20-10	0.90/266/30	0.04/85/30	0.05/85/30
Bomb-50-10	9.93/866/90	12.98/910/90	3.14/910/90
Bomb-100-10	136.02/1866/190	-	109.65/4285/190

实验过程中发现,当厕所数 $b=1$ 时,每执行一次 Duck 动作之后必须执行一次 Flush 动作,将厕所变为非阻塞状态.厕所数目增加时,需要协调炸弹数 b 与厕所数 t .当 $b \leq t$ 时,厕所数目充足,规划解只包含 b 个 Duck 动作,不需要执行任何 Flush 动作;当 $b > t$ 时,执行 t 次 Duck 动作后,所有厕所都处于阻塞状态,处理剩余炸弹时需要交替执行 Duck 与 Flush 动作,反复提供可用的厕所才能解除所有炸弹.Conformant-FF 启发信息对于各个信念状态与目标接近程度的衡量并不准确,有些情况下解除一个炸弹时,启发式估值没有改变,当执行一段动作序列之后爬山算法才能够找到估值更小的信念状态,因此搜索过程中评估的状态较多.当 CFF-Lite 利用 Unknown 命题数引导增强爬山过程时,每解除一个炸弹,Unknown 命题的个数就有所减少,爬山过程能够敏锐地察觉到信念状态的变化,用于缩减信念状态的爬山速度较快.由于求解 Bomb 域问题的目标就是将初始时状态不确定的炸弹确认为解除武装的状态,因此,当所有炸弹的不确定性都被消除时,目标条件得以实现,问题得到求解.

分析表 5 给出的测试结果可知,在厕所数目为 1 的测试问题上,CFF-Lite 系统的求解效率明显优于 Conformant-FF 系统,搜索过程访问的状态数目更少,两个系统规划解长度相同,均求出了最优规划解;炸弹数逐渐增大时,CFF-Lite 系统的求解效率无明显变化,但 Conformant-FF 系统的求解效率有所下降,采用 $h^{(Act)}$ 启发函数时效率下降的幅度更大;当炸弹数目增大到 100 时,Conformant-FF 系统已无法在规定时间内求出规划解.当厕所数目增加时,在一些问题上,CFF-Lite 系统的优势不明显,求解效率不如 Conformant-FF 系统的 h' 启发函数.导致 CFF-Lite 系统性能有所下降的原因是,缩减信念状态的增强爬山算法采用宽度优先搜索寻找不确定性更低的信念状态,随着物体数目的增加,搜索空间快速膨胀,分支数增多,在一定程度上限制了求解速度.但整体上,CFF-Lite 系统优于 Conformant-FF 系统的 $h^{(Act)}$ 启发函数,当炸弹数目逐渐增加时, $h^{(Act)}$ 启发函数求解效率迅速下降,甚至丧失了求解能力,而 CFF-Lite 系统仍然能够快速求出最优规划解.

3.4 Coins域

Coins 域是 IPC5 上采用的一个较新的 Conformant 规划域,域中物体包括电梯、楼层、楼层上的不同位置和硬币,域描述给出的动作包括电梯在楼层之间上下移动的动作、智能体进出电梯的动作、在楼层不同位置之间来回移动的动作以及收集硬币的动作.对于具体的规划问题,初始状态时,每个电梯所停留的楼层以及每个硬币在楼层上的具体位置都是不确定的,目标条件是智能体收集到问题给出的所有硬币.Coins 域具体给出的测试问题分为 4 组,各组问题之间在硬币数、楼层数与楼层上的位置数有所不同,每组各个问题之间初始状态时硬币在各个楼层与位置上的分布不同.表 6 给出该域中不同问题的测试结果.

Table 6 The results of the Coins domain

表 6 Coins 域实验结果

Problem	CFF-Lite	CFF $h^{(Act)}$	CFF h'
	T/S/L	T/S/L	T/S/L
Coins-01	0.01/42/10	0.01/15/12	0.00/15/12
Coins-02	0.00/42/10	0.01/15/12	0.00/15/12
Coins-03	0.01/47/11	0.01/18/13	0.01/18/13
Coins-04	0.01/42/10	0.01/15/12	0.00/15/12
Coins-05	0.01/47/11	0.01/18/13	0.01/18/13
Coins-06	0.11/140/29	0.07/50/31	0.05/50/31
Coins-07	0.08/133/27	0.30/111/34	0.25/111/34
Coins-08	0.08/134/27	0.07/41/28	0.06/41/28
Coins-09	0.07/119/27	0.03/29/26	0.03/29/26
Coins-10	0.09/126/27	0.23/106/38	0.17/106/38
Coins-11	1.01/353/74	4.10/140/78	1.76/140/78
Coins-12	1.18/311/67	3.28/129/72	1.46/129/72
Coins-13	1.17/327/69	15.50/360/95	5.73/360/95
Coins-14	1.41/349/74	3.20/116/76	1.32/116/76
Coins-15	1.06/317/67	15.50/325/89	6.56/325/89
Coins-16	20.93/13783/87	145.92/1727/145	64.32/1727/145
Coins-17	108.04/33551/95	5.29/206/94	2.41/206/94
Coins-18	43.13/21651/91	25.33/552/118	10.65/552/118
Coins-19	57.20/19947/87	71.56/1070/128	29.73/1070/128
Coins-20	13.78/7169/86	74.38/1235/143	31.51/1235/143

求解此类硬币拾捡问题的主要流程首先需要将电梯的位置确定下来并移动到智能体所在的位置,然后智能体乘坐电梯到达各个可能存在硬币的位置,按照目标条件的要求收集硬币.CFF-Lite 从初始信念状态开始,检测到执行电梯移动的动作能够将电梯的位置确定,信念状态上 Unknown 命题数变小,电梯的不确定信息被消除之后,智能体到达各个楼层上,收集不同位置上的硬币,捡起动作能够令智能体收集到硬币,将硬币的状态确定,进一步减少信念状态上的 Unknown 命题数.智能体收集到目标条件要求的所有硬币后,Unknown 命题数减为 0, CFF-Lite 系统求得给定问题的规划解.

由表 6 给出的结果不难看出,CFF-Lite 系统求解 Coins 域问题时显示了较强的优越性,在大多数问题上,CFF-Lite 系统相比 Conformant-FF 系统的求解效率和规划解质量都有不同程度的提高,个别问题上,CFF-Lite 系统的求解速度不如 Conformant-FF 系统;但 Conformant-FF 系统求得的规划解质量通常较差,而 CFF-Lite 系统仍

能很快求出质量更高的规划解,在大多数问题上,CFF-Lite 系统求得的规划解质量更好。

综合以上测试结果可以看出,Conformant-FF 系统对于求解带有几何结构的不确定规划域没有表现出更多的优势,由于问题结构的特殊性,在一些情况下很难给出通用的启发函数,导致求解效率低或规划质量不高。相比之下,CFF-Lite 系统更适于求解此类规划问题,通过降低信念状态的不确定性加速求解过程,多数情况下,CFF-Lite 能够得到质量更好的规划解。实际上,很多 Conformant 规划任务的目标都是使初始时状态不确定的物体到达目标条件指定的确定状态,CFF-Lite 规划系统的处理方法顺应 Conformant 规划问题的要求,对于 Coins 域等贴近实际应用的规划问题也能够显示出较好的求解能力。当问题给定的物体数目较多时,求解效率受到影响,但仍然能够得到高质量的规划解。此外,由于 CFF-Lite 系统在第 1 次增强爬山过程中大量调用 SAT 判定过程,求解效率与动作序列对应的 CNF 公式的求解难度有关。Conformant-FF 系统利用启发信息对搜索空间进行压缩,虽然求得的规划解质量不如 CFF-Lite 系统,但受问题给定的物体数目的影响较小。由此表明:CFF-Lite 系统在求解能力上弥补了 Conformant-FF 系统的不足。从以上测试结果可以看出,除 Bomb 域上 CFF-Lite 与 Conformant-FF 系统都能够求出最优规划解之外,在其余测试域上,CFF-Lite 的规划质量相比 Conformant-FF 系统都有不同程度的提高。

4 结 语

Conformant-FF 规划系统利用信念状态空间启发式搜索方法求解 Conformant 规划问题,一直受到研究者的关注。采用新颖的隐式信念状态表示方法,借助于 SAT 求解方法计算信念状态,能够快速地完成状态转移的计算过程。研究发现,信念状态始终保持较强的不确定性对启发信息与帮助动作的准确性会产生一定的干扰,从而限制了 Conformant-FF 系统的求解能力。

针对这一问题,本文提出了隐式信念状态表示形式下新的问题求解思想,通过增强爬山过程缩减信念状态,逐渐降低信念状态中不确定的信息,从而减小可能世界状态集合的规模,使得缩减后的规划任务更易于求解。实验测试验证了基于缩减信念状态的 Conformant 规划方法的求解能力。分析可知,缩减信念状态方法的搜索空间规模与具体问题有关,不确定性对于算法性能的制约是目前很多规划方法求解 Conformant 问题的难点。在将来的工作中,我们将继续研究如何处理不确定动作的方法以及部分可观察环境下的条件规划等不确定规划问题,例如不确定规划域上领域知识的提取策略、利用领域专门知识控制搜索空间的规模等方法。

致谢 感谢欧阳丹彤教授在智能规划与自动推理的研究过程中给予的悉心指导。

References:

- [1] Bryce D, Cushing W, Kambhampati S. State agnostic planning graphs: Deterministic, non-deterministic, and probabilistic planning. *Artificial Intelligence*, 2011,175(3-4):848-889. [doi: 10.1016/j.artint.2010.12.002]
- [2] Gerevini AE, Haslum P, Long D, Saetti A, Dimopoulos Y. Deterministic planning in the 5th Int'l planning competition: PDDL3 and experimental evaluation of the planners. *Artificial Intelligence*, 2009,173(5-6):619-668. [doi: 10.1016/j.artint.2008.10.012]
- [3] Vidal V, Geffner H. Branching and pruning: An optimal temporal POCL planner based on constraint programming. *Artificial Intelligence*, 2006,170(3):298-335. [doi: 10.1016/j.artint.2005.08.004]
- [4] Lü S, Liu L, Shi L, Li Y. Artificial intelligence planning based on automated reasoning techniques. *Ruan Jian Xue Bao/Journal of Software*, 2009,20(5):1226-1240 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3468.htm> [doi: 10.3724/SP.J.1001.2009.03468]
- [5] Meuleau N, Benazera E, Brafman RI, Hansen EA, Mausam. A heuristic search approach to planning with continuous resources in stochastic domains. *Journal of Artificial Intelligence Research*, 2009,34:27-59. [doi: 10.1613/jair.2529]
- [6] de la Rosa T, Jiménez S, Fuentetaja R, Borrajo D. Scaling up heuristic planning with relational decision trees. *Journal of Artificial Intelligence Research*, 2011,40:767-813. [doi: 10.1613/jair.3231]
- [7] Hoffmann J, Nebel B. The FF planning system: Fast plan generation through heuristic search. *Journal of Artificial Intelligent Research*, 2001,14:253-302. [doi: 10.1613/jair.855]

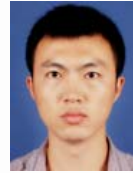
- [8] Wu XJ, Jiang YF, Ling YB. Research and development of StepByStep AI planner. Ruan Jian Xue Bao/Journal of Software, 2008,19(9):2243–2264 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/19/2243.htm> [doi: 10.3724/SP.J.1001.2008.02243]
- [9] Cai DB, Yin MH, Gu WX, Sun JG, Liu KC. Fast forward planning system based on delayed partly reasoning. Chinese Journal of Computers, 2008,31(5):793–802 (in Chinese with English abstract).
- [10] Lai Y, Ouyang DT, Cai DB, Lü S. Model counting and planning using extension rule. Journal of Computer Research and Development, 2009,46(3):459–469 (in Chinese with English abstract).
- [11] Bonet B. Conformant plans and beyond: Principles and complexity. Artificial Intelligence, 2010,174(3-4):245–269. [doi: 10.1016/j.artint.2009.11.001]
- [12] Bonet B, Geffner H. Planning with incomplete information as heuristic search in belief space. In: Proc. of the 5th Int'l Conf. on Artificial Intelligence Planning Systems. Menlo Park: AAAI Press, 2000. 53–61. [doi: 10.1.1.42.9524]
- [13] Cimatti A, Roveri M. Conformant planning via symbolic model checking. Journal of Artificial Intelligence Research, 2000,13: 305–338. [doi: 10.1613/jair.774]
- [14] Cimatti A, Roveri M, Bertoli P. Conformant planning via symbolic model checking and heuristic search. Artificial Intelligence, 2004,159(1-2):127–206. [doi: 10.1016/j.artint.2004.05.003]
- [15] Hoffmann J, Brafman RI. Conformant planning via heuristic forward search: A new approach. Artificial Intelligence, 2006,170(6): 507–541. [doi: 10.1016/j.artint.2006.01.003]

附中文参考文献:

- [4] 吕帅,刘磊,石莲,李莹.基于自动推理技术的智能规划方法.软件学报,2009,20(5):1226–1240. <http://www.jos.org.cn/1000-9825/3468.htm> [doi: 10.3724/SP.J.1001.2009.03468]
- [8] 吴向军,姜云飞,凌应标.智能规划器 StepByStep 的研究与开发.软件学报,2008,19(9):2243–2264. <http://www.jos.org.cn/1000-9825/19/2243.htm> [doi: 10.3724/SP.J.1001.2008.02243]
- [9] 蔡敦波,殷明浩,谷文祥,孙吉贵,刘科成.基于延迟部分推理的快速前向规划系统.计算机学报,2008,31(5):793–802.
- [10] 赖永,欧阳丹彤,蔡敦波,吕帅.基于扩展规则的模型计数与智能规划方法.计算机研究与发展,2009,46(3):459–469.



魏唯(1984—),女,吉林省吉林市人,博士生,主要研究领域为智能规划,自动推理.
E-mail: wei_wei10@mails.jlu.edu.cn



吕帅(1981—),男,博士,讲师,CCF 高级会员,主要研究领域为智能规划,自动推理.
E-mail: lus@jlu.edu.cn



欧阳丹彤(1968—),女,博士,教授,博士生导师,CCF 高级会员,主要研究领域为基于模型的诊断,智能规划,自动推理.
E-mail: ouyd@jlu.edu.cn