软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2018,29(6):1566−1581 [doi: 10.13328/j.cnki.jos.005465]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

向量加法系统验证问题研究综述*

张文博[1], 龙 环[2]

[1](上海交通大学 软件学院,上海 200240)

[2](上海交通大学 计算机科学与工程系,上海 200240)

通讯作者:龙环, E-mail: longhuan@sjtu.edu.cn

摘 要: Petri 网是形式化验证领域最重要的模型之一,具有重要的理论和应用价值.从验证算法分析的角度,Petri 网可以被等价地抽象为向量加法系统.在对向量加法模型的研究中,人们又发展了一些重要的扩展模型.对近些年来国内外学者在向量加法系统验证领域取得的成果进行了系统总结.首先给出了向量加法系统及几个关键验证问题的形式化定义,并重点总结了一般向量加法系统模型上可达性问题的最新研究进展和关键技术;接着总结了当限定模型的维度为固定值时相关研究进展,重点给出了 2 维情况的核心定理;随后介绍了几个重要扩展模型,并总结了这些模型上验证问题研究的最新进展.在每一部分,都对未来研究方向及可能面临的挑战进行了展望.

关键词: Petri 网;向量加法系统;可达性;形式化验证;算法复杂性

中图法分类号: TP311

中文引用格式: 张文博,龙环.向量加法系统验证问题研究综述.软件学报,2018,29(6):1566−1581. http://www.jos.org.cn/1000-9825/5465.htm

英文引用格式: Zhang WB, Long H. State-of-the-Art survey of the verification of vector addition systems. Ruan Jian Xue Bao/ Journal of Software, 2018,29(6):1566−1581 (in Chinese). http://www.jos.org.cn/1000-9825/5465.htm

State-of-the-Art Survey on Verification of Vector Addition Systems

ZHANG Wen-Bo[1], LONG Huan[2]

[1](School of Software, Shanghai Jiaotong University, Shanghai 200240, China)

[2](Departmet of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200240, China)

Abstract: Petri nets is a fundamental model in the area of formal verification. It is popular in both theoretical study and application. For the analysis of algorithmic properties of Petri nets, they are often equivalently viewed as vector addition systems. This survey gives a comprehensive review of the recent achievements in this area. First, formal definitions of the vector addition systems and their key verification problems are provided with emphasis on the discussion about reachability problem, including the latest results and the main proof techniques. Then the development on the case where the dimension is a constant number rather than a variable is summarized along with some key theorems which are fundamental to the current complexity results. Furthermore, as some important variants of vector addition systems have been proposed in recent years, a brief introduction is given to the motivation and definitions of some of the most representative ones, and the latest results on verification relating to these models. In addition, possible future work are highlighted at the end of each section.

Key words: Petri nets; vector addition systems; reachability; formal verification; complexity

自 1962 年 Petri 网[1]的概念提出以来,经过半个多世纪的发展,其已成为描述与分析并发系统最重要的形式

* 基金项目: 国家自然科学基金(61472239, 61772336, 61572318)
Foundation item: National Natural Science Foundation of China (61472239, 61772336, 61572318)
本文由形式化方法的理论基础专题特约编辑傅育熙教授、李国强副教授、田聪教授推荐.
收稿时间: 2017-07-01; 修改时间: 2017-09-01; 采用时间: 2017-11-06; jos 在线出版时间: 2017-12-28
CNKI 网络优先出版: 2017-12-29 13:19:12, http://kns.cnki.net/kcms/detail/11.2560.TP.20171229.1318.005.html

化工具之一.除在并发程序语言形式化验证中的重要应用外,Petri 网模型还被广泛应用于生物、化学、金融、网络、安全等不同领域的建模和分析[2−9].其中,Ball 等人定义了并行库系统(parameterized library system,简称 PLS),在此基础上实现了多线程程序的模型检测工具[2],并证明了 PLS 系统上的一些验证问题与 Petri 网上的验证问题等价;Aalst 用 Petri 网对工作流管理系统进行建模,用 Petri 网理论来验证工作流过程的正确性[3];Heiner 等人[4]用 Petri 网对生化反应过程建模,实现了对生化系统中短时间内的行为进行有效的分析等.国内的研究者也对类似问题进行了深入研究,特别是将 Petri 网理论广泛应用于对网络及安全等领域的验证.代表性的工作包括:北京交通大学 Lei 教授等人用 Petri 网对无线网络系统建模,以实现对无线网络的性能的有效分析[5];清华大学林闯教授等人在 Petri 网的基础上定义了私有 Petri 网,对恶意软件造成的私有信息泄露行为进行分析[6];以及同济大学 Yu教授等人用 Petri 网对电子商务的支付过程建模,验证电子商务支付系统的安全性[7]等.

与 Petri 网等价的数学模型向量加法系统(vector addition system,简称 VAS)具有数学描述上的简洁性,且现实中并发系统的状态和迁移往往都可以用向量描述,故 VAS 本身也具有重要的理论和应用价值.VAS 模型验证的核心问题之一是可达性(reachability)问题,即对给定的 VAS 模型,判断从一个初始格局出发能否到达一个指定的目标格局.令人遗憾的是:虽然关于 VAS 已经有了大量的理论和应用研究,也有了一些高效的实现工具,但对可达性问题的复杂性,至今仍没有给出令人满意的回答.另一方面,VAS 是极为简洁的系统,在实际研究中,人们根据不同的应用背景提出了一些重要的扩展模型(如下推向量加法系统 PVAS、交互向量加法系统 AVASS、分枝向量加法系统 BVASS、带数据的 Petri 网模型 PND(Petri nets with data)等).但对这些模型的相关验证问题的研究都处于初始阶段.

本文第 1 节形式化定义向量加法系统及一些重要的验证问题.第 2 节为本文重点,总结一般 VAS 上可达性问题的主要结论和核心研究技术.第 3 节陈述当维度固定时 VAS 上可达性问题的主要结论和技术.第 4 节总结几个重要 VAS 扩展模型上的验证问题的结论.第 5 节给出现有主要验证结论一览表.本文覆盖了 VAS 研究领域的最新进展、关键技术和开放问题.

1 向量加法系统基础

1.1 VAS的形式化定义

向量加法系统(vector addition system,简称 VAS),可以表示为二元组 $V=\langle d,A\rangle$.$d\in\mathbb{N}$表示 V 的维度,$A\subseteq\mathbb{Z}^d$ 表示迁移规则(transition rule)集合.VAS 中的一次迁移 $u\xrightarrow{a}_V v$ 满足 $v=u+a$,其中,$u,v\in\mathbb{N}^d$,$a\in A$.VAS 的一次从向量 u_0 到 u_n 的运行(run)是一个有限长度的迁移序列 $u_0\xrightarrow{a_1}_V u_1...u_{n-1}\xrightarrow{a_n}_V u_n$,也可以写作 $u_0\xrightarrow{a_1\cdots a_n}_V u_n$.用符号 $\sigma=a_1\ldots a_n$ 表示迁移规则序列.

带状态的向量加法系统(vector addition system with states,简称 VASS)可以表示为三元组 $V=\langle Q,d,T\rangle$.Q 是有限的状态集合,$d\in\mathbb{N}$表示 V 的维度,$T\subseteq Q\times\mathbb{Z}^d\times Q$ 表示迁移规则集合.VASS 中,格局(configuration)指的是集合 $Confs_V=Q\times\mathbb{N}^d$ 中的元素.VASS 中的一次迁移 $(q,u)\xrightarrow{t}_V(q',v)$ 是满足 $v=u+a$ 的三元组,其中,$u,v\in\mathbb{N}^d$,$t=(q,a,q')\in T$.从格局 c_0 到格局 c_n 的一次运行是一个有限长度的迁移序列 $c_0\xrightarrow{t_1}_V c_1...c_{n-1}\xrightarrow{t_n}_V c_n$,也可以写作:

$$c_0\xrightarrow{t_1\cdots t_n}_V c_n.$$

针对上面两个模型,Hopcroft 和 Pansiot 在 1979 年的工作[10]中证明了 n 维 VASS 可以用 n+3 维 VAS 来表达,因此两者的可达性问题等价.

1.2 验证问题的形式化定义

本节具体给出 VAS 模型上几个著名验证问题的形式化定义.

可达性(reachability)问题.

输入:VAS V (VASS V),初始和结束向量 u,u'(格局 c,c');

问题:是否存在一次从 u 到 u'(从 c 到 c')的运行.

可覆盖(coverability)问题.

输入:VAS V,初始和结束向量 u,u';

问题:是否存在向量 $u''\geqslant u'$(u''的每一维都大于等于 u'的相应维),且存在从 u 到 u''的运行.

有界性(boundedness)问题.

输入:VAS V,初始向量 u;

问题:u 可达的向量集合是否有限.

VAS 的可达性关系(reachability relation)用$\rightarrow^*$表示,对于两个向量 $u,u'\in\mathbb{N}^d$,如果 u 到 u'可达,那么$(u,u')\in\rightarrow^*$.值得一提的是,VAS 模型验证的核心问题是可达性问题.实际应用中,很多具体的验证问题都可以归约到可达性问题.

2 VAS 上的验证问题的复杂性

2.1 主要结论

VAS 模型上的覆盖性和有界性都已经有了确定的结论.最早是 Lipton 和 Rackoff[11,12]给出的 EXPSPACE-完备的结果.Rosier 对他们的证明做了细化,得到了更精确的,几乎完全匹配的上下界[13].

Lipton 在 1976 年证明了可达性问题有 EXPSPACE 的算法下界[11].可判定结果则用了更长的时间:Sacerdote 与 Tenney 在 1977 年给出了关于此问题的部分结论[14];Mayr 在 1981 年完善了原结果,并给出了可判定性的完整证明,但该证明非常复杂[15];Kosaraju 在 1982 年对证明进行了简化[16];1992 年,Lambert 在前人工作的基础上给出了清晰的可判定性证明[17].他们工作中的主要证明技术被称为 KLMST 分解.遗憾的是,基于 KLMST 分解的算法没有给出复杂性上界.近年,关于 VAS 模型可达性结论的主要进步是 Leroux 等人自 2009 年开始的一系列工作:Leroux[18]利用 KLMST 分解给出了基于前向递归不变量的较简洁的可达性判定算法;基于此思想,首次给出了不基于 KLMST 的判定算法[19,20].这两个证明为 VAS 可达性问题建立了非常重要的几何直观,并使得研究算法复杂性成为可能.新进展是 Leroux 等人在 LICS 2015 会议上给出的基于理想(ideal)的分解算法[21],并首次给出了可达性问题的上界结论:F_{ω^3}.但这是一个超原始递归的界,且与 Lipton 给出的 EXPSPACE-难下界之间存在巨大差距.可以看出:一般 VAS 模型上的可达性及相关问题在历经数十年的研究后,依然保持着其神秘性.

2.2 VAS主要研究技术及其特点

VAS 上的可达性验证是领域内最受关注也最重要的问题.经过近 40 年的研究,研究者已发展了一套独立、有效的研究技术.下面从验证算法的上界和下界两方面来介绍现有主要研究成果,重点是对关键证明技术的总结和分析.

2.2.1 上界技术

迄今为止,VAS 可达性验证的技术可以分为两大类:2011 年之前的技术都属于 KLMST 分解算法(或对该算法的精化)以及 2011 年 Leroux 提出的基于递归不变量的算法.下面重点陈述此类算法并对其局限和未来可能的工作做分析.

- KLMST 分解算法

前文说过一般的 VAS 和 VASS 的可达性问题等价,本节将在 VAS 的定义下介绍 KLMST 分解算法.KLMST 分解算法会维护一个标记证据图序列(marked witness graph sequence)集,对其中不满足完美(perfect)条件(等价于 Kosaraju 文中的θ条件[16])的标记证据图作分解.如果分解结束时集合不为空,则原可达性问题可达;否则不可达.Leroux 在 2015 年给出了 KLMST 算法的上界 F_{ω^3},本节将给出这个复杂性类的定义.该分解算法用了 Karp-Miller 覆盖树[22],其构造算法有 F_ω的下界[23],由此得出 KLMST 算法的下界.这与目前的上界 F_{ω^3} 之间也有差距.

首先引入大于所有自然数的元素ω,并定义 $N_\omega=\mathbb{N}\cup\{\omega\}$.对于所有的 $z\in\mathbb{Z}$,$\omega+z=z+\omega=\omega$.在集合 $F\subseteq\{1,2,\ldots,d\}$ 上定义映射 $\pi_F:N_\omega^d\rightarrow N_\omega^d$,$\pi_F$在集合 F 的维度上保持不变,将 F 之外的维度映射到ω.证明中用到的核心概念是

证据图(witness graph).这是一个有限的强连通有向图 $G=(S,E)$,其中:点集为 $S\subseteq \mathbb{N}_\omega^d$;边集为 $E\subseteq S\times A\times S$,任意一边 $(u,a,v)\in E$ 满足 $u+a=v$.对于 $e\in E,\Delta(e)=a$.易见:证据图中的所有点,在某一维度上是否为 ω 是一致的.记取值为 ω 的分量集合为 I,记 $\overline{I}=\{1,2,...,d\}\setminus I$.

KLMST 分解算法使用的重要工具是标记证据图(marked witness graph) $M=(G,c^{in},s^{in},c^{out},s^{out})$,其中:$G$ 是证据图;s^{in},s^{out} 是 G 中的两个点,代表图的输入节点(input vertices)和输出节点(output vertices); $c^{in},c^{out}\in\mathbb{N}_\omega^d$ 表示输入限制(input constraint)和输出限制(output constraint).对于 $i\in\{1,2,\ldots,d\}$,如果 $s^{in}(i)\neq\omega$,那么 $c^{in}(i)=s^{in}(i)$,输出节点和输出限制也有同样的关系.记 c^{in},c^{out} 中取值为 ω 的分量集合为 I^{in},I^{out}:

$$\overline{I^{in}}=\{1,2,...,d\}\setminus I^{in},\overline{I^{out}}=\{1,2,...,d\}\setminus I^{out}.$$

Ω_M 表示所有的运行 $u\xrightarrow{a_1\ldots a_n}v$ 的集合,其中, $c^{in}=\pi_{\overline{I^{in}}}(u),c^{out}=\pi_{\overline{I^{out}}}(v)$,$\pi_T(a_1)\ldots\pi_T(a_n)$是 M 中的一条路径.

标记证据图序列(marked witness graph sequence)是标记证据图和迁移规则交叉组成的序列:$\xi=M_0,a_1,M_1,\ldots,a_k,M_k$,其中, $M_j=(G_j,c_j^{in},s_j^{in},c_j^{out},s_j^{out})$,Ω_ξ表示所有运行 $u_0\xrightarrow{\sigma_0}v_0\xrightarrow{a_1}u_1\xrightarrow{\sigma_1}v_1\ldots\xrightarrow{a_k}u_k\xrightarrow{\sigma_k}v_k$ 的集合,其中, $u_j\xrightarrow{\sigma_j}v_j\in\Omega_{M_j}$.定义 L_ξ为多元组$(u_0,\psi_0,v_0,\ldots,u_k,\psi_k,v_k)$的集合,其中,$\psi_i:E_i\to\mathbb{N}$满足 $v_i=u_i+\sum_{e\in E_i}\psi_i(e)\Delta(e)$.

标记证据图序列ξ如果满足以下 3 个条件,则称ξ满足完美条件.

(1) 可泵性(pumpability):对一个标记证明图 M,Ω_M 中存在运行使得在 I 中的维度都能达到无穷大,则称 M 是向前可泵的(forward pumpable).类似地,可以定义向后可泵的(backward pumpable).如果ξ中的所有标记证据图都是向前、向后可泵的,那么称ξ满足可泵性;

(2) 边无界性(edge unboundedness):对于ξ中的每一个标记序列图 M_j,Ψ_j 表示 L_ξ中所有 ψ_j 的集合.对于所有 $e\in E_j$,如果 $\sup\Psi_j(e)=\omega$,那么称ξ满足边无界性;

(3) 输入/输出限制无界性(input/output constraint unboundedness):对于ξ中的每一个标记序列图 M_j,U_j,V_j 表示 L_ξ中所有 u_j,v_j 的集合,如果 $\sup U_j=c_j^{in},\sup V_j=c_j^{out}$,那么称$\xi$满足输入/输出限制无界性.

KLMST 分解算法将会构建标记证据图序列集合的序列 $\Xi_0,\Xi_1,\Xi_2,\ldots$ 初始的 $\Xi_0=\{\xi_0\},\xi_0=M_0=(G_0,u,(\omega,\ldots,\omega),u',(\omega,\ldots,\omega)),G_0=(\{s\},\{s\}\times A\times\{s\}),s=(\omega,\ldots,\omega)$.算法的每一步,若当前集合$\Xi_i$ 为空,则算法回答“不可达”;若 Ξ_i 不为空,则会检测Ξ_i 中的每一个标记证据图序列,选出其中一个不满足完美条件的标记证据图序列ξ,将其分解成有限的标记证据图序列集合 $dec(\xi),\Xi_{i+1}=(\Xi_i\setminus\{\xi\})\cup dec(\xi)$.如果所有标记证据图序列都满足完美条件,则算法终止,回答“可达”.对满足完美条件的标记证据图序列ξ最重要的结论是:

引理 2.1[21]. 标记证据图序列ξ是否满足完美条件,是指数空间内可判定的.

另一方面,对不满足完美条件的标记证据图序列ξ,将对ξ作分解.若ξ中存在一个标记证明图 M 是不可泵的,假设它在第 $i\in I$ 维有上界,那么分解后的标记证据图将具体化第 i 维的数值,$I=I\cup\{i\}$;若ξ不满足边无界,其中,e 出现的次数有上界,那么会在标记证据图中去掉 e 这条边,得到一个由 e 连接的标记证据图序列;若ξ不满足输入限制无界,假如第 $i\in I^{in}$ 维不满足,那么将输入限制替换为有限多种可能的数值,$I^{in}=I^{in}\cup\{i\}$,ξ不满足输出限制无界.为方便理解,我们构造了具体例子介绍分解算法.

例:对 $V=\langle 3,A\rangle,A=\{a_1=(2,-1,0),a_2=(0,-1,1),a_3=(-1,1,-1),a_4=(2,1,-1)\},u=(0,0,3),u'=(1,2,0)$,问 u 到 u'是否可达.

初始的$\xi_0=M_0$ 如图 1 所示.

记 $p=(p_1,p_2,p_3,p_4)\in\mathbb{N}^4$ 表示每条规则使用的次数.解线性方程组 $\sum_{i\in\{1,2,3,4\}}a_ip_i=u'-u,p=(1,0,1,0)+n(0,3,2,1)$,$n\in\mathbb{N}$.规则 a_1 只会出现 1 次,不满足边无界条件.算法将ξ_0中的 a_1 边去掉,将其分解成多个由 a_1 连接的标记证据图.得到标记证据图序列 $\xi_1=M_0',a_1,M_1'$.如图 2 所示.

$$M_0'=(G_0',(0,0,3),(\omega,\omega,\omega),(\omega,\omega,\omega),(\omega,\omega,\omega)),M_1'=(G_1',(\omega,\omega,\omega),(\omega,\omega,\omega),(1,2,0),(\omega,\omega,\omega)).$$

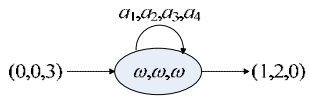


Fig.1 Initial marked witness graph sequence $\xi_0=M_0$

图 1 初始的标记证据图序列$\xi_0=M_0$

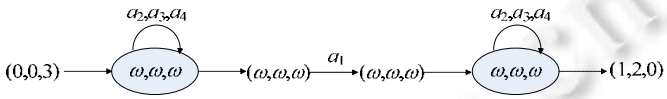


Fig.2 Next marked witness graph sequence $\xi_1 = M_0', a_1, M_1'$

图 2 标记证据图序列 $\xi_1 = M_0', a_1, M_1'$

接着构造从M_0', M_1'的输入限制出发的 Karp-Miller 树以及从输出限制出发的反向的 Karp-Miller 树,如图 3 所示.可以看出: M_0'不是向前可泵的, M_1'不是向后可泵的.ξ_1 不满足可泵条件.算法将 G_0' 中不可泵的维度具体化,因为不可泵的维度上取值是有限的,得到的依然是一个有限的图.ξ_1 可以分解为集合Ξ_2,图 4 是集合Ξ_2 中的一个元素,记作ξ_2: $\xi_2 = M_0'', a_1, M_1'', M_0'' = (G_0'', (0,0,3), (\omega,0,3), (\omega,1,2), (\omega,1,2)), M_1'' = (G_1'', (\omega,0,2), (\omega,0,2), (\omega,2,0), (1,2,0))$.

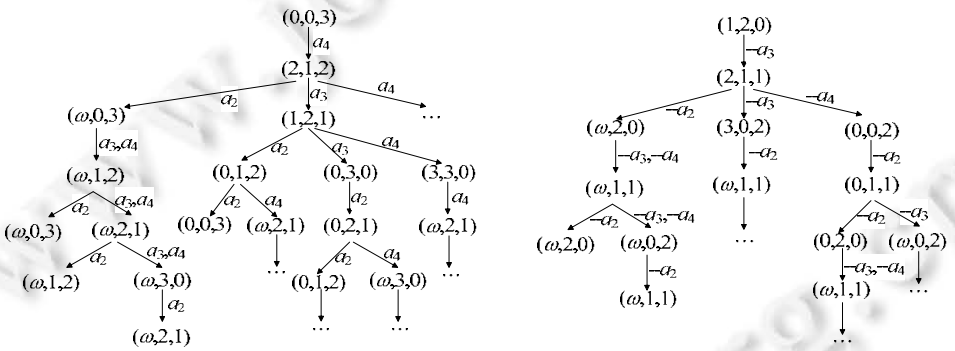


Fig.3 Karp-Miller tree from (0,0,3) and Karp-Miller tree constructed with inverse transitions from (1,2,0)

图 3 从(0,0,3)出发的 Karp-Miller 树和从(1,2,0)出发的反向的 Karp-Miller 树

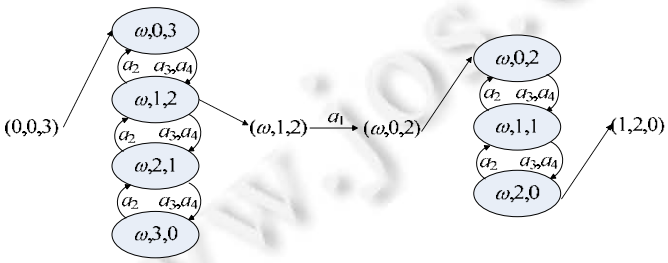


Fig.4 Final marked witness graph sequence $\xi_2 = M_0'', a_1, M_1''$

图 4 一个最终的标记证据图序列 $\xi_2 = M_0'', a_1, M_1''$

这里的ξ_2 是满足完美条件的标记证据图序列,算法不会再对其分解.算法结束时Ξ_i 至少包含ξ_2,因此 u 到 u' 可达.

为了说明上述计算过程会终止,为标记证据图序列ξ定义一个秩函数(ranking function)r.这个秩函数会将ξ映射到三元组的多重集合,包含每一个标记向量图的($|I|$,$|E|$,$|I^{in}|+|I^{out}|$).例如 $r(\xi_0)=\{(3,4,0)\}$,$r(\xi_1)=\{(3,3,3),(3,3,3)\}$,$r(\xi_2)=\{(1,9,1),(1,6,1)\}$.在 Dershowitz 定义的多重集合的序下[24],对于所有的$\xi \in dec(\xi)$,有 $r(\xi)>r(\xi')$.因为

这个在多重集合上的序是一个良序(well order),由良序的性质,所有秩函数严格递减的ξ序列都是有限的.又因为每次对ξ分解得到的集合 $dec(\xi)$也是有限集,由柯尼格引理,KLMST 分解算法一定终止.

下面对 KLMST 分解算法作复杂度分析,这里非常重要的是引入了一类所谓的非初等(non-elementary)复杂性类,它的定义依赖于集合论中的序数理论.

对于序列$\xi_0,\xi_1,\xi_2,\ldots$满足$\xi_{n+1}\in dec(\xi_n)$,有 $r(\xi_0)>r(\xi_1)>r(\xi_2)>\ldots$.分解过程中,这个序列长度的上界记做 L.序数$\alpha<\varepsilon_0(\varepsilon_0=\omega^{\omega^{\omega^{\cdots}}})$有唯一的康托范式$\alpha=\omega^{\alpha_1}c_1+\ldots+\omega^{\alpha_n}c_n$,其中,$c_1,\ldots,c_n\in\mathbb{N}$,序数$\alpha_1>\alpha_2>\ldots>\alpha_n$.也可以等价地写作 $\alpha=\omega^{\alpha_1}+\ldots+\omega^{\alpha_m}$,序数$\alpha_1\geqslant\alpha_2\geqslant\ldots\geqslant\alpha_m$.对于两个序数 $\alpha=\omega^{\alpha_1}+\ldots+\omega^{\alpha_n}$ 和 $\beta=\omega^{\beta_1}+\ldots+\omega^{\beta_m}$,$\alpha\oplus\beta=\omega^{\gamma_1}+\ldots+\omega^{\gamma_{n+m}}$,其中,$\gamma_1\geqslant\gamma_2\geqslant\ldots\geqslant\gamma_{m+n}$是$\alpha_i$和$\beta_i$的重排序.

我们可以把秩函数 r 的取值等价地看做ω^{ω^3}以下的序数.对于标记证据图 M,序数$\beta_M=\omega^2|I|+\omega|E|+(|I^{in}|+|I^{out}|)$.对于标记证据图序列$\xi=M_0,a_1,M_1,\ldots,a_k,M_k,\beta_\xi=\bigoplus_{1\leqslant j\leqslant k}\omega^{\beta_{M_j}}$.例如在前面的例子中,

$$\beta_{\xi_0}=\omega^{\omega^2\cdot3+\omega\cdot4},\beta_{\xi_1}=\omega^{\omega^2\cdot3+\omega\cdot3+3}\cdot2,\beta_{\xi_2}=\omega^{\omega^2+\omega\cdot9+1}+\omega^{\omega^2+\omega\cdot6+1}.$$

给定$\alpha=\omega^{\alpha_1}c_1+\ldots+\omega^{\alpha_n}c_n$,定义$N(\alpha)=\max_{1\leqslant j\leqslant n}\{c_i,N(\alpha_i)\}$.$N(\xi_0)=4,N(\xi_1)=3,N(\xi_2)=9$.注意,控制函数(control function) $g:\mathbb{N}\to\mathbb{N}$是严格增函数.$g_\alpha$是一个超限递归定义的函数:$g_0(n)=0,g_{\alpha+1}(n)=1+g_\alpha(g(n)),g_\lambda(n)=g_{\lambda(n)}(n)$.其中,$\lambda$是极限序数.$\lambda(n)$的超限递归定义是:$(\gamma+\omega^{\beta+1})(n)=\gamma+\omega^\beta(n+1),(\gamma+\omega^{\lambda'})(n)=\gamma+\omega^{\lambda(n)}$.

例:$\omega(n)=n+1,\omega^{\omega^3}(n)=\omega^{\omega^3(n)}=\omega^{\omega^2(n+1)}$;

$g_k(n)=k,g_\omega(n)=g_{n+1}(n)=n+1,g_{\omega+1}(n)=1+g_\omega(g(n))=2+g(n)$.

为了研究复杂性,还需要定义控制函数$H(n)=n+1,H_{\omega^2}(n)=H_{\omega(n+1)}(n)=(2^{n+1}-1)(n+1),H_{\omega^3}(n)$是一个 Non-elementary 函数,$H_{\omega^\omega}(n)$是非原始递归函数.

定义$g^k(n)=g(g^{k-1}(n))$表示函数 g 迭代 k 次,$g^\alpha(n)=g^{g^\alpha(n)}(n)$.最后可以定义复杂性类:

$$\mathcal{F}_{<\alpha}=\bigcup_{\beta<\omega^\alpha}FSPACE(H^\beta(n)),F_{h,\alpha}=\bigcup_{p\in\mathcal{F}_{<\alpha}}SPACE(h^{\omega^\alpha}(p(n))),F_\alpha=F_{H,\alpha}.$$

在 KLMST 分解中,假设$\xi=M_0,a_1,M_1,\ldots,a_k,M_k,\xi'\in dec(\xi)$.

可以看出,$N(\beta_\xi)$不超过$\|\xi\|=\max_{0\leqslant j\leqslant k}(2d,k,|E_j|,|I_j^{in}|+|I_j^{out}|,|I_j|)$.我们将会构造控制函数 g,使得$\|\xi\|\leqslant g(\|\xi\|)$.在$\xi$不满足边无界或输入/输出限制无界条件的情况下,$\|\xi'\|$的大小相比$\|\xi\|$有一个不超过指数的放大;而在$\xi$不满足可泵条件时,因为用到了 Karp-Miller 树,有$\|\xi'\|\leqslant H^{\omega^{d+1}}\|\xi\|^4$ [25],我们可以构造控制函数 $g(x)=H^{\omega^{d+1}}(p(x))$,其中,$p(x)$是给定的多项式函数.这个控制函数在 VAS 的维度 d 固定时是在$\mathcal{F}_{<\omega}$中的原始递归函数,当维度 d 作为输入的一部分时,是在$\mathcal{F}_{<\omega+1}$中的非原始递归函数.

综上,在 KLMST 分解中构造的标记证据图序列的大小有上界$g^L(n)$,其中,$L=g_{\omega^{\omega^3}}(n)$.因此,KLMST 分解算法用到的空间不超过$g^{\omega^{\omega^3}}(n)$.因此,VAS 可达性问题属于复杂性类$F_{g,\omega^3}$.因为$F_{g,\omega^3}=F_{\omega^3}$ [26],所以有如下结论.

定理 2.2[21]. VAS 的可达性问题上界不超过F_{ω^3}.

研究展望:关于 KLMST 分解的最新研究是 2015 年 Leroux 定义了理想的概念[21],给出了对 KLMST 分解的另一个直观解释:KLMST 分解是对路径集合的理想的分解.这可能对 VAS 的其他扩展模型的验证问题有启发作用.此外,KLMST 分解的算法目前已知的上界F_{ω^3}和下界F_ω有一个差距,与 VAS 可达性问题本身的下界 EXPSPACE-hard 之间有一个更大的差距,这些都值得进一步研究.

- 基于递归不变量(inductive invariant)的算法

基于递归不变量(inductive invariant)算法.这个算法有两个半可判定过程:第 1 个不断枚举所有的动作序列以证明可达,第 2 个枚举所有的 Presburger 公式证明不可达.Leroux 证明了如果向量 u 到向量 u'不可达,那么存在一个可以用 Presburger 公式表示的向量集合,向量 u 在这个集合中,而向量 u'不在这个集合中[16].并且这个集

合相对于可达性关系是一个前向递归不变量,即,所有集合中的向量的下一步依然在这个集合中.因此,这个Presburger集合就是不可达的一个证明(witness).目前,这个方法只有可判定的结果,没有算法复杂性上界.

如果集合 $S\subseteq\mathbb{Z}^d$ 可以写成如下形式: $L(C,P)=\left\{x\mid \exists c\in C,\exists \alpha_1,...,\alpha_k\in\mathbb{N},\exists p_1,...,p_k\in P,x=c+\sum_{i=1}^{k}\alpha_i p_i\right\}$,其中,$C,P\subseteq\mathbb{Z}^d$,则称 S 是半线性集(semi linear set).Hopcroft 等人证明了 5 维以下的 VAS 的可达集是可被有效计算的半线性集[10].集合 $S\subseteq\mathbb{Z}^d$ 如果可以被 Presburger 算术 $FO(\mathbb{Z},+,\leqslant,0,1)$中的公式表示,那么称 S 是 Presburger 集合.Ginsburg 等人证明了,集合 S 是 Presburger 集合当且仅当 S 是半线性集[27].

最近,Leroux 对一般情况下 VAS 的可达集给出了更加精确的描述.2011 年,Leroux 定义了 Lambert 集合和 Petri 集合的概念[19],证明了在一般情况下,VAS 的可达集是 Lambert 集合,可达性关系的集合是 Petri 集合.其中,Petri 集合一定是 Lambert 集合;Lambert 集合一定也是 Presburger 集合.下面给出这些集合的定义.

集合 $P\subseteq\mathbb{Z}^d$ 满足 $0\in P,P+P\subseteq P$,则称 P 为周期集合(periodic set).对于周期集合 P,如果 $\mathbb{Q}_{\geqslant 0}P$ 在 $FO(\mathbb{Q},+,\leqslant,0,1)$ 上可定义,则称 P 是多胞周期集合(polytope periodic set).如果集合 $L\subseteq\mathbb{Z}^d$ 是有限个集合的并,其中每个集合可以表示为 $b+P,b\in\mathbb{Z}^d,P\subseteq\mathbb{Z}^d$ 是多胞周期集合,则称 L 是 Lambert 集合.如果集合 $X\subseteq\mathbb{Z}^d$,满足对于任意 Presburger 集合 $S\subseteq\mathbb{Z}^d,S\cap X$ 是 Lambert 集合,则称 X 是 Petri 集合.

例:周期集合 $P_1=\{(m,n)\in\mathbb{N}^2\mid m\leqslant\sqrt{2}n\}$ 不是多胞周期集合.周期集合 $P_2=\{(0,0)\}\cup\{(2^n,1)\mid n\in\mathbb{N}\}\cup\{(m,n)\mid m\geqslant 1,b\geqslant 2,m,n\in\mathbb{N}\}$. $\mathbb{Q}_{\geqslant 0}P_2=\{(0,0)\}\cup\mathbb{Q}^2_{>0}$ 在 $FO(\mathbb{Q},+,\leqslant,0,1)$上可定义,$P_2$ 是多胞周期集合.因此,P_2 是 Lambert 集合.同时,P_2 不是 Petri 集合,因为对于 Presburger 集合 $\mathbb{N}\times\{1\},P_2\cap(\mathbb{N}\times\{1\})=\{(2^n,1)\mid n\in\mathbb{N}\}$ 不是 Lambert 集合.

对给定自反传递的关系 $R\subseteq\mathbb{Z}^d\times\mathbb{Z}^d$,集合 $X,Y\subseteq\mathbb{Z}^d$.定义:

$$post_R(X)=\bigcup_{x\in X}\{y\in\mathbb{Z}^d\mid(x,y)\in R\},\ pre_R(Y)=\bigcup_{y\in Y}\{x\in\mathbb{Z}^d\mid(x,y)\in R\}.$$

若 $post_R(X)\subseteq X$,则称 X 是相对于关系 R 的前向递归不变量(forward inductive invariant);若 $pre_R(Y)\subseteq Y$,则称 Y 是相对于关系 R 的后向递归不变量(backward inductive invariant).定义 R 的区分(separators) $(X,Y):X,Y\subseteq\mathbb{Z}^d$ 并且都是 Presburger 集合,满足 $R^*\cap(X\times Y)$是空集.如果关系 R 本身是 Petri 集合,则称 R 是 Petri 关系.

Leroux 证明了下面的重要结论:

引理 2.3[19]. VAS 的可达性关系是一个 Petri 关系.

进一步得到关于 Petri 关系的一个主要结论是:

引理 2.4[19]. 若 $R^*\subseteq\mathbb{Z}^d\times\mathbb{Z}^d$ 是一个 Petri 关系,$X',Y'\subseteq\mathbb{Z}^d$ 是两个 Presburger 集合,$R^*\cap(X'\times Y')=\varnothing$,那么存在一个区分$(X,Y)$,其中,$X$ 是前向递归不变量,Y 是后向递归不变量,满足 $X'\subseteq X,Y'\subseteq Y$.

如果 u 到 u'不可达,$\mathbb{Z}^d$ 就可以划分为前向递归不变量 X 和后向递归不变量 Y,X,Y 都是 Presburger 集合,其中,$u\in X,u'\in Y$.因此有一个判定 VAS 不可达的半可判定过程:通过不断枚举 Presburger 公式表示的向量集合,判断其是否包含 u 而不包含 u'.

研究展望:注意到,基于前向递归不变量的算法不依赖于 KLMST 分解算法.不过,该算法的复杂性目前仍然是公开的.这个复杂性依赖于最短的可达路径长度和最短的能够表示一个可以区分起点和终点的前向递归不变量的 Presburger 公式长度.值得一提的是:虽然 VAS 可能会有一个有限的 Ackermann 大小的可达集合,但该算法中表示前向递归不变量的 Presburger 集合是可达集合的一个超集,因此该算法还是可能有比 F_ω更好的下界.

总结:目前,VAS 可达性的上界算法主要有 KLMST 分解算法和基于递归不变量的算法.KLMST 分解算法试图去分析可达路径,最终得到的标记证据图序列可以看出路径的结构.该算法会有 F_ω的下界和 F_{ω^3} 的上界.基于递归不变量的算法更加直观,同时枚举所有的路径以证明可达,和所有的前向递归不变量以证明不可达.但该算法的复杂性分析似乎比 KLMST 分解算法更难.

2.2.2 下界技术

Lipton 在 1976 年给出了 VAS 可达性问题的 EXPSPACE 的算法下界.Lipton 首先提出了并行程序(parallel

programs)的模型,证明了并行程序上的可接受问题(the acceptance problem)可以多项式时间归约到 VAS 的可达性问题.然后,递归构造了可以模拟 3 个有界的计数器的并行程序(这里的计数器可以做加法、减法、测 0 操作),计数器的界是这个并行程序的双指数大小.而用 3 个双指数大小的计数器可以来模拟任意需要指数空间的图灵机,因此可以得到 VAS 的可达性问题 EXPSPACE-hard 的下界.

其中,Lipton 用到的直观概念是并行程序(parallel programs).并行程序是三元组 $P=\langle F,d,\boldsymbol{x}\rangle$,其中,$F$ 是并行的流程图(flowcharts)的集合,d 是维数,$\boldsymbol{x}=(x_1,\ldots,x_d)\in\mathbb{N}^d$ 是所有流程图共同维护的 d 维向量.每个流程图是有限的有向图,包含 4 种节点:起始节点、接受节点、猜测节点、赋值节点.如图 5 所示.

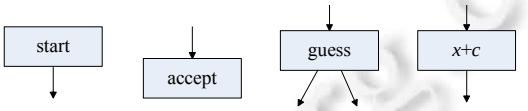


Fig.5 Four kinds of nodes in a parallel program

图 5 流程图的 4 种节点

并行程序的格局(configuration)可以表示为 $c=\langle p_1,\ldots,p_m,x_1,\ldots,x_d\rangle,m=|F|$.其中,$p_i$ 表示第 i 个流程图当前节点的编号.并行程序的初始格局 C_0 中,p_i 是第 i 个流程图的起始节点,$x_i=0$.

并行程序每一步运行 $\langle p_1,\ldots,p_m,x_1,\ldots,x_d\rangle\rightarrow_P\langle p_1',\ldots,p_m',x_1',\ldots,x_d'\rangle$ 需要满足以下中的一个条件.

(1) 存在 $i\in\{1,2,\ldots,m\}$,p_i 是起始节点,p_i' 是 p_i 唯一的后继节点.对于所有的 $j\neq i, p_j'=p_j$;对于所有的 j,$x_j'=x_j$;

(2) 存在 $i\in\{1,2,\ldots,m\}$,p_i 是猜测节点,p_i' 是 p_i 两个后继节点中的一个.对于所有的 $j\neq i, p_j'=p_j$;对于所有的 $j, x_j'=x_j$;

(3) 存在 $i\in\{1,2,\ldots,m\}$,p_i 是赋值节点,p_i' 是 p_i 唯一的后继节点.对于所有的 $j\neq i, p_j'=p_j$;对于所有的 j,$x_j'=x_j+c_j\geqslant 0$.

用 $\rightarrow_P^*$ 表示 $\rightarrow_P$ 的传递闭包.并行程序的可接受问题(acceptance problem)问是否存在一个格局 $C=\langle p_1,\ldots,p_m,x_1,\ldots,x_d\rangle$,使得 $C_0\rightarrow_P^* C$,其中,存在一个 i,使得 p_i 是接受节点.可以证明,并行程序上的可接受问题和 VAS 的可达性问题满足如下关系:

引理 2.5[11]. 并行程序上的可接受问题可以多项式时间归约到 VAS 的可达性问题.

我们可以通过研究并行程序上可接受问题的下界来回答 VAS 可达性问题的下界.注意,并行程序可以模拟计数器的加法和测零操作.加法操作的模拟比较直观;模拟测零操作的思路是用两维向量 $x,x'\in\mathbb{N}$,令它们始终保持 $x+x'=A_k=2^{2^k}$ 的关系.这样,判断 x 是否为 0,等价于判断 x' 是否能够等于 A_k.Lipton 递归地构造一个并行程序中的程序块 α_{x_i},程序块 α_{x_i} 可以在 $x_i+x_i'=A_i$ 的条件下测试 x_i 是否为 0:如果 x_i 为 0,则该程序块将走 YES 出口,并互换 x_i,x_i' 的值;如果 x_i 不为 0,则该程序块将走 NO 出口.

- i=1 时,程序块如图 6 所示.

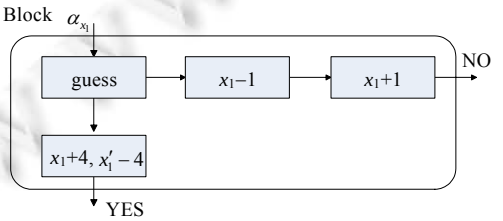


Fig.6 Block α_{x_1}

图 6 程序块 α_{x_1}

- i>1 时,设置辅助变量 $u_i+u_i'=A_{i-1},v_i+v_i'=A_{i-1},b_i+c_i=0$.$u_i,v_i$ 分别控制两层循环,赋值节点 $x_i+1,x_i'-1$ 将

会经过 $A_{i-1}^2 = A_i$ 次.程序块 $\alpha_{v_i'}, \alpha_{u_i'}$ 的作用类似于 $\alpha_{x_{i-1}}$,但为保证构造出的并行程序大小在多项式范围内,加入了新的流程图,如图 7(b),用来调用程序块 $\alpha_{x_{i-1}}$.b_i, c_i 在程序块中起到了调用和返回的作用.

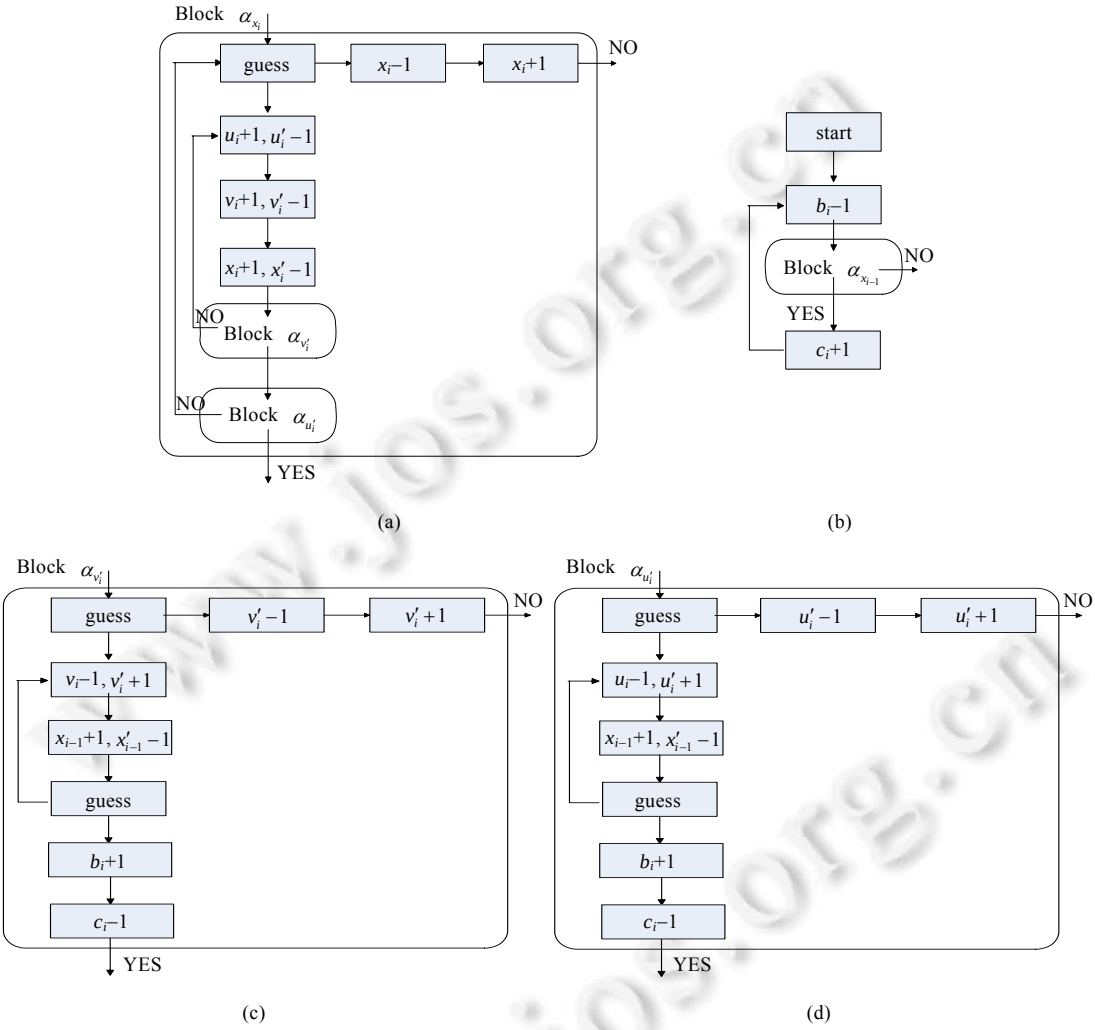


Fig.7 Block α_{x_k}

图 7 程序块 α_{x_k}

将变量 $x_i'(u_{i+1}', v_{i+1}')$ 赋值为 A_i,只需要把图程序块赋值节点 $x_i+1, x_i'-1$ 替换成 $x_i'+1$ 即可.而模拟计数器的测 0 操作只需执行两次程序块,第 2 个程序块中需要将其中的 x_k, x_k' 互换,作用是将 x_k, x_k' 的值再交换回来.第 2 个程序块中不会走到 NO 出口(如图 8 所示).

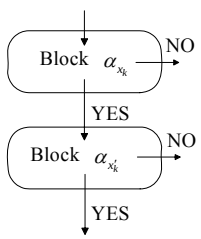


Fig.8 Test zero in a parallel program

图 8 在并行程序上实现测 0

因此,并行程序上的可接受问题是 EXPSPACE-hard.

由引理 2.5,VAS 可达性问题的下界也是 EXPSPACE-hard.

定理 2.6[11]. VAS 可达性问题的下界是 EXPSPACE-hard.

研究展望:在关于下界的研究方面,已有技术的核心是 VAS 系统可以模拟双指数大小的计数器,由此其可达性问题的下界是 EXPSPACE-hard.类似的技术可以推广到下推向量加法系统.下推向量加法系统在向量加法系统的基础上增加了一个栈,Lazic 证明了它可以模拟界是 $2\uparrow\uparrow n$ 的计数器[28], $2\uparrow\uparrow n=2^{2^{\cdot^{\cdot^{2}}}}\Big\}n$个2.因此,下推向量加法系统上的可达性问题是 Non-elementary 的.特别值得一提的是,可以看出,这类下界证明的共同点都是巧妙的“计数”.我们认为,这类方法可以被推广到其他模型的研究中.

3 固定维度时验证问题的复杂性

3.1 主要结论

在前面的讨论中,向量的维度 d 是输入的一部分.而在实际应用中,问题的维度往往存在固定的上界.Hopcroft 和 Pansiot 在 1979 年的工作[10]中指出:当 VAS 的维数在 5 之内时,其可达集是一个可被有效计算的半线性集.由于 n 维 VASS 可以用 n+3 维 VAS 表示,这就等价于得出了 2 维之内 VASS 可达性问题、等价性问题(equivalence)、包含问题(containment)等的可判定结论.他们同时用反例说明了该技术无法被应用到 3 维及更高维度的 VASS 的研究.

关于低维度 VASS 的研究结果主要包括:Hasse 与 Kreutzer 等人[29]在 2009 年证明了 1-VASS(即 1 维 VASS,等价于 4 维 VAS)的可达性问题在一进制编码下是 NL-完备的,而在二进制编码下是 NP-完备的.Haase 在其博士论文[30]中证明了同一模型在输入是二进制编码时的有界性问题和可覆盖问题都是 NP-完备的.而对 2-VASS,Howell 等人[31]对 Hopcroft 的算法进行了分析,指出原始算法的复杂度是非确定双指数时间(2-NEXPTIME),并进一步将算法上界改进到了确定双指数时间(2-EXPTIME).而关于二进制编码下 2-VASS 上可达性问题的最终结果则是在最近几年才得出:Fearnley 等人[32]给出了 PSPACE-hard 的证明;紧接着,Blondin 等人[33]在 2015 年找到了 PSPACE 的判定算法,从而证明了该问题实际上是 PSPACE-完备的.他们工作的一个额外结论是证明了对固定维度($d\geqslant 2$ 维)d-VASS,有界性和可覆盖性都是 PSPACE-完备的.值得一提的是,Blondin 等人的工作,本质上是对 Leroux 等人[34]在 2004 年有关 2-VASS 具有平坦性(flatness)证明的精化.当限制为输入是一进制编码时,2-VASS 可达性问题在 2016 年被证明是 NL-完备的[35].到目前为止,在可达性问题上没有关于 3-VASS 或更高维度模型的结果:既没有更准确的下界结论,也没有任何非平凡(即优于 F_{ω^3})的判定算法.下面总结其中代表性的研究技术,并指出未来可能的工作方向.

3.2 固定维度VASS主要研究技术及其特点

目前,固定维度 VASS 的复杂性结果集中在 2 维及以下,即:当向量维度大于 2 时,仅有平凡结论(即 2 维问题的下界和维度无限制时的上界).因此,这里集中讨论 2 维及 2 维以下模型相关的研究技术.

3.2.1 上界技术(算法)

3.2.1.1 二维 VASS 的平坦化(flatness)

已有 d 维($d\leqslant 2$)VASS 的算法都依赖于一个重要的结论:此维度限制下的 VASS 是可平坦化的.对给定的 VASS $V=(Q,T)$,其中,Q 代表状态集合,T 代表迁移规则集合,V 的输入规模被定义为$|V|=|Q|+|T|\cdot d\cdot\log_2|T|$.用$\alpha_0,\beta_1,\alpha_1,\ldots,\beta_k,\alpha_k$代表 V 上一组不含循环、可执行的迁移规则序列,则一组形如 $\rho=\alpha_0\beta_1^*\alpha_1\ldots\beta_k^*\alpha_k$ 的串,被称为一条线性路径策略(linear path scheme).典型的线性路径策略如图 9 所示.

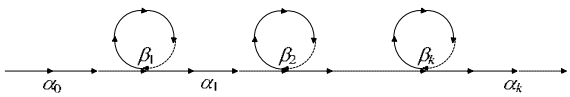


Fig.9 Linear path scheme

图 9 线性路径策略

有限条线性路径策略的并集被称为半线性路径方案(semi-linear path scheme).若 V 的可达关系集合可以表示为一个半线性路径方案,则我们称 V 是具有平坦性的(相应称对应的可达关系集合是平坦的).直观上讲,平坦性意味着对应的可达路径可以被表示为一组只依次含有若干简单环的路径(即,不存在环的嵌套),而具有这样结构的路径所对应的可达关系是有效半线性的.对给定的二元关系 $R_V \subseteq Q\times Q$, R_V^* 为 R_V 的自反传递闭包. Leroux 等人的主要结论[34]是:当 2 维 VASS 的初始向量和目标向量中两个分量的值都大于某个可计算的常数时,对应的二元关系是平坦的.这个性质称做终极平坦性(ultimately flat). R_V^* 的其他子集能相对简单地证明具有平坦性.最终, R_V^* 集合是有效半线性集.但 Leroux 等人的工作中并没有给出对应方案的具体取值范围,故而无法直接得出更有效的复杂性上界.

直到 2015 年,Blondin 等人[33]对 Leroux 的研究进行了量化,给出了上述常数 c 的上界$(|Q|+T)^{O(1)}$.并证明了如果初始格局到目标格局可达,则在对应的平坦性的关系中,半线性集合将由一组有效的、指数规模的线性路径策略(即 $|\rho| \overset{def}{=} |\alpha_0\beta_1...\beta_k\alpha_k|$ 为输入的指数函数)组成,特别地,其中简单环个数(即 k)相对输入是多项式大小,绕每条简单环重复次数的上界相对输入是指数大小.综合上述结果,算法可猜测(利用非确定性)路径上的中间状态,并利用多项式空间的计数器控制搜索的上界,即:对一组可达格局对,算法必然能在多项式空间内停止并给出一个正面的回答.再根据 Savitch 定理 PSPACE=NPSPACE,就完成了关于 2-VASS 可达性具有(确定)多项式空间算法的证明.这是目前关于固定维度 VASS 最好的结果.

Blondin 工作中一个相对独立的结果是,证明了非确定有限自动机所识别的路径的 Parikh 像(Parikh image,用于表达路径中每个字符出现的次数的函数)存在一个多项式规模的使用线性路径策略的表达方式.基于这个结论,很容易就能证明:当将原始可达性验证问题的条件放宽到整数意义下时,任意 d 维 VASS 上的可达性等价于存在一组有限的、多项式规模的线性路径策略.当然,这与各维向量值都应限制为自然数下的可达性还有距离.为此,他们将问题分成了 3 种子类型.

(1) 初始格局和目标格局的状态相同,且始末状态两个计数器的值都大于常数 c,但对中间运行情况不做限制;

(2) 初始格局到目标格局的整条路径上,两个计数器的值都大于常数 c;

(3) 初始格局到目标格局的整条路径上,至少有一个计数器的值不超过常数 c.

以上 3 种情况分别对应于图 10 中的 3 种不同路径类型.

Blondin 等人证明了这 3 种情况下的 VASS 的可达关系都有指数规模的刻画.具体而言,类型(3)可规约到维度为 1 的情况从而化简.类型(2)可以用类型(1)的结论推导得出.对最关键的类型(1),与 $d\leqslant 2$ 这个限制直接相关的核心技术引理是:

引理 3.1[33]. 令 $b\in Z^2, P\subseteq Z^2$ 且 $b\in P$,令 Z 是 2 维空间中的一个象限,则有 $L(b;P)\cap Z=\bigcup_{i\in I} L(c_i;P_i)$,且对每个 $i\in I$,如下结论成立:

- $|P_i|\leqslant 2$;
- $P_i\subseteq (P\cup L(b;P))\cap Z$;且
- 存在 $e\leqslant P^{O(1)}$,满足 $\{c_i\}\cup(P_i\cap L(b;P))\subseteq b+cone_{[0,e]}(P)$.

即:在 2 维情况下,对满足引理前提的线性集 $L(b;P)$,考虑其与任意象限的交集,结果都可以表示为一组简单的半线性路径的并;或者更具体地说,由一组周期(period,即 P_i)的基数为 2,且周期和基(base,即 c_i)来自一个简单集合的线性路径组成.值得注意的是,这个引理对 $d\geqslant 3$ 的情况均不成立.

最后,通过证明任意 2 维 VASS 的可达路径都可分解为多项式规模段类型为类型(1)、类型(2)、类型(3)的路径后,就完成了 2 维情况下问题上界的证明.

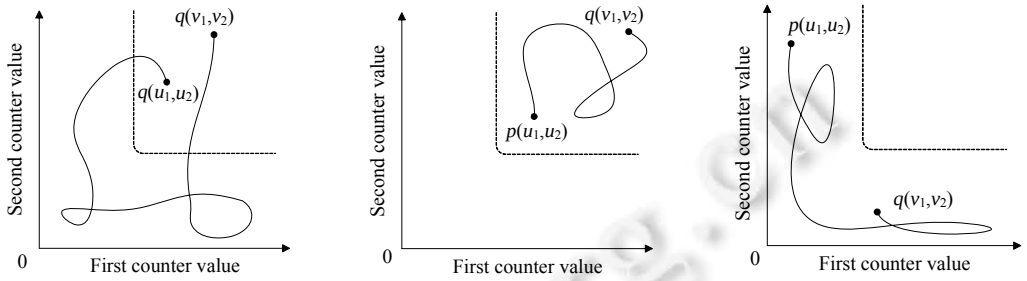


Fig.10 Three kinds of reachable paths

图 10 3 种类型的可达路径

研究展望:当维数固定时的研究,是近年关于 VASS 验证研究的核心.特别值得注意的是:除了上面的引理 3.1 及其相关推论外,Blondin 在文献[33]中的其他结论都可以被直接应用或推广到更高维的情况.因此,对更高维情况的研究,重点和难点是对 3 维及以上的情况找到类似引理 3.1 的降维结论.

3.2.2 下界技术

下界来自对有界单计数器自动机(bounded one-counter automata)模型上可达性问题的规约.

有界单计数器自动机(可以表示为三元组 $V=(Q,T,b)$,其中,(Q,T)就是 1-VASS,而 $b\in N$ 是用二进制表达的计数器值的上界.令 $B=[0,b]$,对给定的初始格局 $p(u)$和目标格局 $q(v)$,$u,v\in B$.与一般的可达性问题不同的是:可达性关注是否存在从 $p(u)$到 $q(v)$的可行路径,且路径上每一个中间格局对应的计数器的值都不超过 b.即问关系 $p(u)\to_B^* q(v)$ 是否成立.

Fearnley 等人[32]在 2013 年证明了有界单计数器自动机上的可达性问题是 PSPACE-complete.

对任意有界单计数器自动机 $V=(Q,T,b)$,可构造其可达性问题到 2-VASS 可达性问题的规约.规约的核心是利用界限参数 b:构造 2-VASS $V'\overset{def}{=}(Q,\{h(t):t\in T\})$,其中,$h$ 函数的具体定义是 $h(p,z,q)\overset{def}{=}(p,(z,-z),q)$.注意:因为原模型 V 中 $u,v\in B$,所以 V'中两个向量对应的计数器的值都不会小于 0.此时,如下的当且仅当关系显然成立:

模型 V 中 $p(u)\xrightarrow{\pi}_B q(v)$ 当且仅当模型 V'中 $p(u,b-u)\xrightarrow{h(\pi)}_{N^2} q(v,b-v)$.

研究展望:注意,这部分所得到的下界结论可以平凡地延伸到维度 $d\geqslant 3$ 的情况.但因为目前对 3 维的情况了解甚少,可优先考虑相关算法的研究,增加对模型的认知之后,再考虑合适的下界规约.

4 重要扩展模型

随着应用领域研究的推进,研究者们逐渐意识到现有 VASS 模型的局限并定义和发展了若干重要的扩展模型.鉴于篇幅限制,这里只列举领域内近年来备受关注的几个重要模型,并总结关于这些模型的验证问题的最新结论.

4.1 PVAS

下推向量加法系统(pushdown vector addition system,简称 PVAS)是在 VASS 模型之上增加了栈及相应的入栈(push)和出栈(pop)操作.

PVAS 上一条典型的迁移规则为:$(p,u,X)\xrightarrow{v,op}(q,u+v,X')$,其中,$(p,u)\xrightarrow{v}(q,u+v)$ 的含义与 VASS 完全一致,$X\xrightarrow{op}X'$ 用于记录栈的变化(op 为栈操作),PVAS 为研究同时具有递归与整数变量的程序语言提供了一个简洁的数学模型[36].

PVAS 模型是 VASS 的非平凡扩展.Leroux 等人证明了 PVAS 终止性和有界性都可判定[37],但没有进一步的

上界结论;目前仍不知道 PVAS 上可覆盖性和可达性问题是否可判定,仅仅知道它们都是 Tower-hard[28],复杂性远高于 VASS 上的对应验证问题.

在维度固定时,仅当维度为 1 时有一些进展.Leroux 等人证明了 1-PVAS 可覆盖性是 NP-hard,上界是 EXPSPACE[38,39],而有界性问题是 NP-hard,相应的上界是 EXPTIME[39,40].

关于 PVAS 非常特殊的一点是,n-PVAS 的可达性,可以规约到 n+1-PVAS 的可覆盖性,即,PVAS 模型下的可达性问题和可覆盖性问题难度本质上一样.这也从另一个角度说明了 PVAS 的特殊性.

4.2 AVASS

Alternating VASS 最早是由 Lincoln 等人在研究命题线性逻辑(propositional linear logic)的可判定性时被提出来的.AVASS 是一个四元组 $A=Q,d,T_u,T_f$.其中:Q 是有限的状态集合;$d\in N$ 代表维度;$T_u\subseteq Q\times Z^d\times Q$ 是一元迁移规则,如 $q\xrightarrow{v}q_1$;而 $T_f\subseteq Q^3$ 是状态分叉(fork)规则,如 $q\to q_1\wedge q_2$.作为例子,当一元迁移规则作用在格局(q,u)上时,下一格局是$(q_1,u+v)$;而当分叉规则作用在同一格局上时,下一格局是(q_1,u)或(q_2,u).显然,VASS 是 AVASS 的一个特殊子类,即 $T_f=\varnothing$的情况.

关于 AVASS 的主要结论是可达性不可判定[41].Courtois 等人证明了:若将求解(格局)可达性削弱为状态可达性,则问题难度是双指数完备的[42].当 AVASS 模型中向量维度固定时,状态可达性问题(及非终止性问题)难度降低为 EXPTIME 完备的.值得一提的是:AVASS 状态可达性问题的研究可以有效地规约到 BPP、VASS 与有限状态系统的模拟问题上,从而得到了新的下界结论.这也成为研究 AVASS 的重要原因之一.

4.3 BVASS

VASS 的计算过程可以理解为一个线性过程.Branching VASS[43]的计算则是一棵计算树:从叶子节点出发到根节点的过程,每个非叶子节点的向量值被定义为:该节点的儿子节点所对应的向量值之和,再加上一个规则向量.得到的模型可有效刻画计算语言学、逻辑、乃至 XML 等中的一些核心问题,从而引起研究者的广泛重视.

Lazic[44]和 Demri[45]分别研究了 BVASS 的可达性问题以及可覆盖性、有界性问题,证明了前者是 2-EXPSPACE-难的,而后两个问题是 2-EXPTIME-完备的.Lazic 等人[46,47]进一步证明了一般 BVASS 可达性问题的下界甚至是 Tower-hard,也就是 Non-elementary,即,任何验证算法不具有现实可行性.但非常有趣的是:与对维度不做限制时的高复杂性相比,2016 年,Göller 等人证明了一维 BVASS 上的可达性、可覆盖以及有界性问题都是多项式时间完备的[48].这些结果都表明,BVASS 并非 VASS 的平凡扩展.

4.4 PND模型

PND 模型是在 Petri 网中引入数据与数据操作的概念而为相关建模提供了便利,PND 模型中,每个 place 里不再是存储同样类型的 token,而是转为存储数据,故而迁移规则相应地变为消耗和产生数据.关于 PND,比较集中的工作是 Lazic[49],Haddad[50],Rosa-Velardo[51]等人近年给出的一系列结论,其中最重要的结论之一是证明了有序数据子模型上的可覆盖性问题是 F_{ω^3} -完备的.PND 研究的最新成果是关于无序数据模型(unordered data Petri nets,简称 UDPN)上可覆盖性问题的研究:2016 年,Hofman 等人利用 Karp-Miller 树的构造,得出了一个超 Ackermannian(hyper-Ackermannian)的算法上界[52],而此问题目前已知的最好下界是 2017 年 Lazic 等给出的 Ackermannian-hard[53].截至目前,PND 模型相关问题的上下界之间仍有很大的差距,这也是未来一个可能的研究方向.

研究展望:本节集中介绍了几个典型的 VAS 扩展模型,并对这些模型上的验证问题做了较为简要的总结.这几个模型都具有理论或应用上的重要价值.从已有结论可以看出,许多关于这些模型的验证问题都丞待解决.我们认为,其中最值得关注的是:1 维和 2 维 PVAS 可覆盖问题的上、下界、任意固定维度 AVASS 的验证问题的复杂性、2 维 BVASS 的可达性问题的复杂性以及 UDPN 可达性问题的判定性.

5 本文总结

本文总结了加法向量系统验证领域的若干核心问题、最新进展和关键技术;特别地,指出了本领域中若干重要开放问题及部分结论的相互联系;并对一些扩展模型和验证问题间的相互关系做了相应探讨,提出了未来的研究方向.我们认为:向量加法系统是一个简洁而强大的数学模型,无论从理论还是应用上而言,对向量加法系统及其扩展模型上验证问题的深入研究都具有极重要的价值.最后,作为对加法向量系统验证领域研究现状的小结,我们将本领域最新结论总结在表 1 中(注意,这里用#表示开放问题).

Table 1 State-of-the-Art results on VASS and its extensions

表 1 VASS 及其扩展模型上最新结论总结

模型&问题 \ 维度		参数	1 维	2 维	d 维($d\geq 3$)
VASS	可达性	[EXPSPACE-难, F_{ω^3}]	NP-完备	PSPACE-完备	#
PVAS	可覆盖性	[F_3,#]	[NP-难,EXPSPACE]	#	#
	有界性	[F_3,#]	[NP-难,EXPTIME]		
	可达性	[F_3,#]	[NP-难,#]		
AVASS	可达性	不可判定	#	#	#
BVASS	可覆盖性	2-EXPTIME-完备	P-完备	#	#
	有界性	2-EXPTIME-完备			
	可达性	[Non-elementary,#]			
UDPN	可覆盖性	[F_ω,$F_{\omega 2}$]	#	#	#
	有界性	[F_3,Decidable]			
	可达性	[F_ω,#]			

References:

[1] Petri CA. Kommunikation mit Automaten. Bonn: Institut fur Instrumentelle Mathematik, Schriften des IIM Nr. 2, 1962.

[2] Ball T, Chaki S, Rajamani S. Parameterized verification of multithreaded software libraries. In: Proc. of the Tools and Algorithms for the Construction and Analysis of Systems. LNCS 2031, Springer-Verlag, 2001. 158−173. [doi: 10.1007/3-540-45319-9_12]

[3] van der Aalst W. The application of Petri nets to workflow management. Journal of Circuits, Systems, and Computers, 1998,8(1): 21−66. [doi: 10.1142/S0218126698000043]

[4] Heiner M, Gilbert D, Donaldson R. Petri nets for systems and synthetic biology. In: Proc. of the Formal Methods for Computational Systems Biology. 2008. 215−264. [doi: 10.1007/978-3-540-68894-5_7]

[5] Lei L, Lin C, Zhong ZD. Stochastic Petri nets for wireless networks. In: Springer Briefs in Electrical and Computer Engineering, Springer-Verlag, 2015. 1−101. [doi: 10.1007/978-3-319-16883-8]

[6] Fan LJ, Wang YZ, Li JY, Cheng XQ, Lin C. Privacy Petri net and privacy leak software. Journal of Computer Science and Technology, 2015,30(6):1318−1343. [doi: 10.1007/s11390-015-1601-7]

[7] Yu WY, Yan CG, Ding ZJ, Jiang CJ, Zhou MC. Modeling and validating E-commerce business process based on Petri nets. IEEE Trans. on Systems, Man, and Cybernetics: Systems, 2014,44(3):327−341. [doi: 10.1109/TSMC.2013.2248358]

[8] 林闯.随机 Petri 网和系统性能评价.北京:清华大学出版社,2005.

[9] Jiang YX, Lin C, Qu Y, Yin H. Research on Model-Checking Based on Petri Nets. Ruan Jian Xue Bao/Journal of Software, 2004,15(9):1265−1276 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/15/1265.htm

[10] Hopcroft J, Pansiot JJ. On the reachability problem for 5-dimensional vector addition systems. Theoretical Computer Science, 1979, 8(2):135−159. [doi: 10.1016/0304-3975(79)90041-0]

[11] Lipton RJ. The reachability problem is exponential-space-hard. Technical Report, 62, Department of Computer Science, Yale University, 1976.

[12] Rackoff C. The covering and boundedness problems for vector addition systems. Theoretical Computer Science, 1978,6:223−231. [doi: 10.1016/0304-3975(78)90036-1]

[13] Rosier L, Yen HC. A multiparameter analysis of the boundedness problem for vector addition systems. Journal of Computer and System Sciences, 1986,32:105−135. [doi: 10.1016/0022-0000(86)90006-1]

[14] Sacerdote GS, Tenney RL. The decidability of the reachability problem for vector addition systems (preliminary version). In: Proc. of the 9th Annual ACM Symp. on Theory of Computing. ACM Press, 1977. 61−76. [doi: 10.1145/800105.803396]

[15] Mayr EW. An algorithm for the general Petri net reachability problem. In: Proc. of the STOC'81. ACM Press, 1981. 238−246. [doi: 10.1145/800076.802477]

[16] Kosaraju SR. Decidability of reachability in vector addition systems. In: Proc. of the STOC'82. ACM Press, 1982. 267−281. [doi: 10.1145/800070.802201]

[17] Lambert JL. A structure to decide reachability in Petri nets. Theoretical Computer Science, 1992,99(1):79−104. [doi: 10.1016/0304-3975(92)90173-D]

[18] Leroux J. The general vector addition system reachability problem by presburger inductive invariants. In: Proc. of the 24th IEEE Symp. on Logic in Computer Science (LICS 2009). 2009. [doi: 10.1109/LICS.2009.10]

[19] Leroux J. Vector addition system reachability problem (a short self-contained proof). In: Proc. of the Principles of Programming Languages. Austin: ACM Press, 2011. 307−316. [doi: 10.1145/1926385.1926421]

[20] Leroux J. Vector addition systems reachability problem (a simpler solution). In: Voronkov A, ed. Proc. of the Alan Turing Centenary Conf. (Turing-100). 2012.

[21] Leroux J, Schmitz S. Demystifying reachability in vector addition systems. In: Proc. of the 30th IEEE Symp. on Logic in Computer Science (LICS 2015). 2015. 56−67. [doi: 10.1109/LICS.2015.16]

[22] Karp RM, Miller RE. Parallel program schemata. Journal of Computer and System Sciences, 1969,3(2):147−195. [doi: 10:1016/S0022-0000(69) 80011-5] [doi: 10.1016/S0022-0000(69)80011-5]

[23] Muller H. The reachability problem for VAS. In: Proc. of the Advances in Petri Nets 1984. NCS 188. Springer-Verlag, 1985. 376−391.

[24] Dershowitz N, Manna Z. Proving termination with multiset orderings. Communications of the ACM, 1979,22(8):465−476. [doi: 10:1145/359138: 359142] [doi: 10.1145/359138.359142]

[25] Figueira D, Figueira S, Schmitz S, Schnoebelen P. Ackermannian and primitive-recursive bounds with Dickson's Lemma. In: Proc. of the LICS 2011. IEEE Press, 2011. [doi: 10:1109/LICS:2011:39]

[26] Schmitz S. Complexity hierarchies beyond Elementary. ACM Trans. on Computation Theory, 2015. http://arxiv:org/abs/1312:5686 [doi: 10.1145/2858784]

[27] Ginsburg S, Spanier EH. Semigroups, presburger formulas and languages. Pacific Journal of Mathematics, 1966,16(2):285−296. [doi: 10.2140/pjm.1966.16.285]

[28] Lazic R. The reachability problem for vector addition systems with a stack is not elementary. 2013. https://arxiv.org/pdf/1310.1767.pdf

[29] Haase C, Kreutzer S, Ouaknine J, Worrell J. Reachability in succinct and parametric one-counter automata. In: Proc. of the Concurrency Theory (CONCUR 2009). 2009. 369−383. [doi: 10.1007/978-3-642-04081-8_25]

[30] Haase C. On the complexity of model checking counter automata [Ph.D. Thesis]. University of Oxford, 2012.

[31] Howell RR, Rosier LE, Huynh DT, Yen HC. Some complexity bounds for problems concerning finite and 2-dimensional vector addition systems with states. Theoretical Computer Science, 1986,46(3):107−140. [doi: 10.1016/0304-3975(86)90026-5]

[32] Fearnley J, Jurdzinski M. Reachability in two-clock timed automata is PSPACE-complete. In: Proc. of the 40th Int'l Colloquium on Automata, Languages, and Programming (ICALP), Vol.2. 2013. 212−223. [doi: 10.1007/978-3-642-39212-2_21]

[33] Blondin M, Finkel A, Göller S, Haase C, McKenzie P. Reachability in two-dimensional vector addition systems with states is PSPACE-complete. In: Proc. of the 30th Annual ACM/IEEE Symp. on Logic in Computer Science (LICS 2015). 2015. 32−43. [doi: 10.1109/LICS.2015.14]

[34] Leroux J, Sutre G. On flatness for 2-dimensional vector addition systems with states. In: Proc. of the CONCUR 2004—15th Int'l Conf. on Concurrency Theory. 2004. 402−416. [doi: 10.1007/978-3-540-28644-8_26]

[35] Englert M, Lazic P, Totzke P. Reachability in two-dimensional unary vector addition systems with states is NL-complete. In: Proc. of the LICS 2016. 2016. 477−484. [doi: 10.1145/2933575.2933577]

[36] Ganty P, Majumdar R. Algorithmic verification of asynchronous programs. ACM Trans. on Programming Languages and Systems, 2012,34(1):6:1−6:48. [doi: 10.1145/2160910.2160915]

[37] Leroux J, Praveen M, Sutre G. Hyper-Ackermannian bounds for pushdown vector addition systems. In: Proc. of the Joint Meeting of the 23rd EACSL Annual Conf. on Computer Science Logic (CSL) and the 29th Annual ACM/IEEE Symp. on Logic in Computer Science (LICS) (CSL-LICS 2014). 2014. [doi: 10.1145/2603088.2603146]

[38] Leroux J, Sutre G, Totzke P. On the coverability problem for pushdown vector addition systems in one dimension. In: Proc. of the ICALP, Vol.2. 2015. 324−336. [doi: 10.1007/978-3-662-47666-6_26]

[39] Finkel A, Leroux J. Recent and simple algorithms for Petri nets. Software and System Modeling, 2015,14(2):719−725. [doi: 10.1007/s10270-014-0426-0]

[40] Leroux J, Sutre G, Totzke P. On boundedness problems for pushdown vector addition systems. In: Proc. of the RP. 2015. 101−113. [doi: 10.1007/978-3-319-24537-9_10]

[41] Lincoln P, Mitchell J, Scedrov A, Shankar N. Decision problems for propositional linear logic. Annals of Pure and Applied Logic, 1992,56(1-3):239−311. [doi: 10.1016/0168-0072(92)90075-B]

[42] Courtois JB, Schmitz S. Alternating vector addition systems with states. In: Proc. of the MFCS, Vol.1. 2014. 220−231. [doi: 10.1007/978-3-662-44522-8_19]

[43] Verma KN, Goubault-Larrecq J. Karp–Miller trees for a branching extension of VASS, discr. Mathematics & Theoretical Computer Science, 2005,7:217−230.

[44] Lazic R. The reachability problem for branching vector addition systems requires doubly-exponential space. Information Processing Letters, 2010,110(17):740−745. [doi: 10.1016/j.ipl.2010.06.008]

[45] Demri S, Jurdzinski M, Lachish O, Lazic R. The covering and boundedness problems for branching vector addition systems. Journal of Computer and System Sciences, 2013,79(1):23−38. [doi: 10.1016/j.jcss.2012.04.002]

[46] Lazic R, Schmitz S. Non-Elementary complexities for branching VASS, MELL, and extensions. In: Proc. of the CSL/LICS. 2014. [doi: 10.1145/2603088.2603129]

[47] Lazic R, Schmitz S. Non-Elementary complexities for branching VASS, MELL, and extensions. ACM Trans. on Computational Logic, 2015,16(3):1−30. [doi: 10.1145/2733375]

[48] Göller S, Haase C, Lazic R, Totzke P. A polynomial-time algorithm for reachability in branching VASS in dimension one. 2016, 105:1−13. https://arxiv.org/abs/1602.05547

[49] Lazic R, Newcomb T, Ouaknine J, Roscoe A, Worrell J. Nets with tokens which carry data. Fund Information, 2008,88(3):251−274. [doi: 10.1007/978-3-540-73094-1_19]

[50] Haddad S, Schmitz S, Schnoebelen P. The ordinal recursive complexity of timed-arc Petri nets, data nets, and other enriched nets. In: Proc. of the LICS. IEEE Press, 2012. 355−364. [doi: 10.1109/LICS.2012.46]

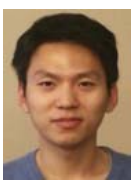
[51] Rosa-Velardo F. Ordinal recursive complexity of unordered data nets. Technical Report, TR-4-14, Departamento de Sistemas Informaticosy Computacion, Universidad Complutense de Madrid, 2014.

[52] Hofman P, Lasota S, Lazić R, Leroux J, Schmitz S, Totzke P. Coverability trees for Petri nets with unordered data. In: Proc. of the 19th Int'l Conf. on Foundations of Software Science and Computation Structures (FoSSaCS). 2016. [doi: 10.1007/978-3-662-49630-5_26]

[53] Lazic R, Totzke P. What makes Petri nets harder to verify: Stack or data? In: Proc. of the Concurrency, Security, and Puzzles 2017. 2017. 144−161. [doi: 10.1007/978-3-319-51046-0_8]

附中文参考文献:

[9] 蒋屹新,林闯,曲扬,尹浩.基于 Petri 网的模型检测研究.软件学报,2004,15(9):1265−1276. http://www.jos.org.cn/1000-9825/15/1265.htm



张文博(1992－),男,河南洛阳人,学士,主要研究领域为理论计算机科学.

龙环(1980－),女,博士,副教授,博士生导师,主要研究领域为理论计算机科学.