

基于 DNA 计算的计算树逻辑模型检测方法研究进展

韩英杰 周清雷 朱维军
(郑州大学信息工程学院 郑州 450001)

摘 要 计算树逻辑(CTL)模型检测是保证系统正确性和可靠性的重要手段,但严峻的时空复杂性问题制约着 CTL 模型检测在工业界的应用。DNA 计算的大规模并行性和 DNA 分子巨大的存储密度为解决 CTL 模型检测的时空复杂性问题提供了新思路。文中介绍了基于 DNA 计算的 CTL 模型检测的背景,并概述了基于 DNA 计算的 CTL 模型检测方法的基本原理。从检测能力的提升、自治化程度的提升和相关问题的解决这 3 个方面综述了方法的研究进展。在方法检测能力的提升方面,分 3 个层次综述了研究进展,即从只能检测单个 CTL 基本公式到能够检测一般公式,从只能检测带未来时间算子的 CTL 公式到能够检测带过去时间算子的 CTL 公式,从只能检测 CTL 公式到能够检测线性时序逻辑、投影时序逻辑和区间时序逻辑公式,表明了方法的检测能力在公式数量和种类上均有大幅提升;在方法自治化程度的提升方面,综述了从基于无记忆过滤模型的人工操作的非自治方法到基于粘贴自动机的分子自治下的自治方法的研究进展,表明基于 DNA 计算的 CTL 模型检测方法已实现高度自治化;在相关问题的解决方面,阐述了提升 DNA 分子特异性杂交有效性预测的效率和构建 CTL 公式的 DNA 表示等的研究进展。最后,指出了基于 DNA 计算的 CTL 模型检测在研究新方法、构建专用的 DNA 计算模型和扩展应用领域等方面的研究趋势。

关键词 模型检测,计算树逻辑,DNA 分子,分子计算
中图法分类号 TP301,TP384 **文献标识码** A **DOI** 10.11896/jsjcx.181102091

Survey on DNA-computing Based Methods of Computation Tree Logic Model Checking

HAN Ying-jie ZHOU Qing-lei ZHU Wei-jun
(School of Information Engineering,Zhengzhou University,Zhengzhou 450001,China)

Abstract Computation tree logic (CTL) model checking is an important approach to ensuring the correctness and reliability of systems. However, the severe spatio-temporal complexity problems restrict the application of CTL model checking in industry. The large-scale parallelism of DNA computing and the huge storage density of DNA molecules provide new ideas for resolving the problems. The background and the principle of DNA-computing based methods of CTL model checking were introduced. The research progress was reviewed from three aspects: the improvement of power, the improvement of autonomy and the resolution of related problems. Firstly, the research progress of methods in terms of power was summarized from checking only one basic CTL-formula to general CTL-formulas, from CTL-formulas with only future operators to CTL-formulas with past-time operators, and from CTL-formulas to linear temporal logic, projection temporal logic and interval temporal logic formulas. Secondly, the research progress of methods in terms of autonomy from non-autonomous methods based on manual operations of memory-less filtering models to autonomous methods based on molecular autonomy of sticker automata was reviewed, showing that the methods are highly autonomous. At last, relevant problems in improving the predictive efficiency of specific hybridization of DNA molecules and constructing DNA molecules of CTL-formulas were described. In the end, corresponding research directions were discussed by concerning different methods, new models and new applications.

Keywords Model checking, Computation tree logic, DNA molecules, Molecular computing

1 引言

模型检测是一种能够提高系统安全性的形式化验证方法,在硬件验证^[1]、网络协议验证、安全协议验证^[2]、软件验

证^[3]和生物系统^[4]等领域被广泛应用。采用计算树逻辑(Computation Tree Logic, CTL)描述系统行为分支时序关系的模型检测方法被称为 CTL 模型检测。效率低和状态空间爆炸问题限制了其在工业界的应用。从目前的研究成果来

到稿日期:2018-11-14 返修日期:2019-03-22 本文受国家自然科学基金项目(61572444)资助。
韩英杰(1976—),女,硕士,讲师,CCF 会员,主要研究方向为形式化方法与模型检测、DNA 计算,E-mail:ieyjh@zhu.edu.cn;周清雷(1962—),男,博士,教授,CCF 高级会员,主要研究方向为模型检测、自动机理论和信息安全等,E-mail:ieqlzhou@zhu.edu.cn(通信作者);朱维军(1976—),男,博士,副教授,主要研究方向为 DNA 计算、模型检测和形式化方法等。

看,符号模型检测方法使用布尔函数表示模型的状态集合和状态转换关系,并采用二元决策图在计算机中存储布尔函数,很大程度地缓解了状态空间爆炸问题,扩大了模型检测处理问题的规模^[5]。其最大的瓶颈是存储和操作二元决策图时对内存的需求非常大,但硅芯片的存储极限导致电子计算机极难实现超大规模的存储容量。并行与分布式模型检测虽然提供了额外的计算硬件资源,但器件工艺技术与制造成本等问题是传统基于硅材料的微电子技术难以解决的,这些问题需要革命性的创新思维和技术手段来解决。其中,DNA 计算是一种备受关注的计算模式。

DNA 计算是以 DNA 分子与相关生物酶为基本材料、以某些生化反应为基础的计算模式^[6],具有运算速度快、计算所需能耗低等特点,在解决优化计算、信息安全、图与组合优化中的复杂困难问题时具有天然优势。DNA 分子的存储容量大和运算的大规模并行性等特点为 CTL 模型检测提供了新思路。2006 年,图灵奖得主 Emerson 教授利用 DNA 分子存储了“天文数量级”的状态空间图,并利用 DNA 分子自发的连锁反应实现了 CTL 公式 EFp (存在一条运行路径,在这条路径上原子命题 p 最终成立) 的模型检测,且在检测过程中不存在指数空间爆炸问题^[7]。他同时指出,基于 DNA 计算的模型检测方法在扩展自动验证领域方面具有极大的潜力。这是首次在 DNA 计算框架下的模型检测工作,证明了基于 DNA 计算的模型检测的科学可行性。但 Emerson 教授提出的方法不能检测其他 7 个 CTL 基本公式、描述性质的常用公式乃至一般公式,且在模型检测过程中使用了限制性内切酶,对生化反应环境的要求比较苛刻,在模型检测结果为“否”时不能提供反例。这些现象均表明基于 DNA 计算的 CTL 模型检测问题还远未解决,主要原因是缺少一套基于 DNA 计算的 CTL 模型检测方法。

基于 DNA 计算的 CTL 模型检测方法的主要研究内容是非自治的模型检测算法和纳米尺度下分子自组装的模型检测算法。从科学研究角度看,DNA 分子运算的大规模并行性有助于提升模型检测的效率,DNA 分子巨大的存储容量有助于存储和表示复杂系统的状态空间,为解决电子计算机无能为力的模型检测任务提供了可能。从应用角度看,CTL 模型检测有许多重量级的应用,例如,在 IEEE Futurebus+ 标准的缓存一致性协议中检测死锁等关键性错误^[8],在空中防撞系统中检测系统的有效性和安全性^[9]等。这些应用凸显了提升效率、解决状态空间爆炸问题、扩展 CTL 模型检测的应用领域的重要性。此外,因为使用 DNA 单链分别表示系统模型和 CTL 公式,CTL 模型检测有望在人体活细胞内实现,从而为人类重大基因突变疾病的分子诊断和治疗提供自治的、智能的、精确的手段,这也是传统模型检测无法实现的。

2 基本原理

2.1 CTL 模型检测

CTL 是一种基于命题逻辑的时序逻辑,用于表达系统模型的分支时序属性,由 Clarke 等^[10]提出;Queille 等^[11]将其用于模型检测。假设 p, q 是原子命题, AP 是原子命题集合, $p, q \in AP$, CTL 公式 φ 的定义如下:

$$\varphi ::= p \mid !\varphi \mid \varphi \vee \varphi \mid EX\varphi \mid E[\varphi U \varphi] \mid A[\varphi U \varphi]$$

其中,“!”是逻辑非运算符,“ $\vee$ ”是逻辑或运算符, E 是存在路径量词, A 是全称路径量词,时序运算符 X 表示“下一个”, U 表示“直到...”。

CTL 的基本公式有 8 个,分别是 $EFp, AFp, EGp, AGp, EXp, AXp, E[pUq]$ 和 $A[pUq]$ 。假设 M 是一个系统模型,则 CTL 基本公式的含义如下。

- 1) EFp : 在 M 中至少存在一条路径,在这条路径上 p 最终成立;
- 2) AFp : 在 M 的所有路径上均有 p 最终成立;
- 3) EGp : 在 M 中至少存在一条路径,在这条路径上 p 始终成立;
- 4) AGp : 在 M 的所有运行路径上 p 始终成立;
- 5) EXp : 在 M 中至少存在一条路径,在这条路径上从当前状态开始的下一个状态上 p 成立;
- 6) AXp : 在 M 的所有路径上,均有 p 在从当前状态开始的下一个状态上成立;
- 7) $E[pUq]$: 在 M 中至少存在一条路径,在这条路径上有 q 在未来某个状态成立,在此之前, p 一直成立;
- 8) $A[pUq]$: 在 M 的所有路径上都有 q 在未来某个状态成立,在此之前, p 一直成立。

CTL 模型检测的思想如下:给定一个表示有限状态并发系统的 Kripke 结构 $M=(S,R,L)$ (S 是一个有限、非空的状态集, $R \subseteq S \times S$ 是关系的集合, $L:S \rightarrow 2^{AP}$ 是状态的标签函数) 和一个描述系统性质的 CTL 公式 f , 找出 S 中满足 f 的所有状态的集合。如果所有初始状态都在这个集合中,则 M 满足 f ^[12]; 否则, M 不满足 f , 提供反例。

2.2 基于 DNA 计算的 CTL 模型检测方法的原理

基于 DNA 计算的 CTL 模型检测方法的原理如图 1 所示。采用 Kripke 结构建模待检测的系统,采用 DNA 单链编码系统模型的状态和变迁,遍历系统模型得到系统所有运行路径的 DNA 单链。采用 CTL 公式描述系统需满足的性质,采用 DNA 单链编码 CTL 公式。将系统模型运行路径的 DNA 单链与 CTL 公式的 DNA 单链混合杂交,根据杂交产生的双链 DNA 分子判断系统模型是否满足 CTL 公式描述的性质。在具体操作时,依据 DNA 计算模型及其提供的编码方案和生物操作来实现。

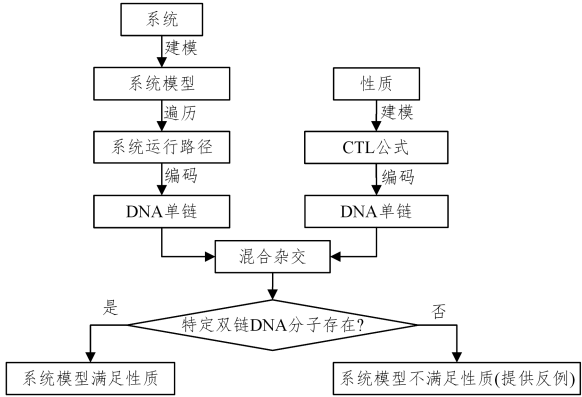


图 1 基于 DNA 计算的 CTL 模型检测方法的基本原理
Fig. 1 Principle of DNA-computing based methods of CTL model checking

3 研究进展

本节从 3 个方面综述基于 DNA 计算的 CTL 模型检测方法的研究进展。首先,在方法的检测能力提升方面,从 3 个层次综述基本公式模型检测的研究进展:从单个 CTL 基本公式到一般公式,从带未来时间算子的 CTL 公式到带过去时间算子的 CTL 公式,从 CTL 公式到线性时序逻辑(Linear Temporal Logic, LTL)、区间时序逻辑(Interval Temporal Logic, ITL)和投影时序逻辑(Projection Temporal Logic, PTL)。其次,综述了方法从非自治到自治的研究进展。最后,阐述提升 DNA 分子特异性杂交有效性的预测效率和构建 CTL 公式的 DNA 表示等相关问题的研究进展。

3.1 检测能力的提升

3.1.1 从计算树逻辑基本公式到一般公式

文献[13]提出了 CTL 基本公式 EXp 和 AXp 的模型检测的 DNA 计算方法。该方法的核心问题是将 CTL 模型检测问题规约为有限状态自动机接受的语言的问题。其主要思想是构造待检测系统的带标签的有限状态自动机模型,并用 DNA 单链编码其状态和变迁,遍历系统模型获得 DNA 单链表示的系统模型的所有运行路径;构造 CTL 公式的有限状态自动机,并进一步构造其粘贴自动机^[14]。粘贴自动机是一种 DNA 计算模型^[15],将系统模型运行路径的分子作为 CTL 公式的有限状态自动机的输入,判定其是否能够被 CTL 公式的有限状态自动机接受,即判断系统模型运行路径的 DNA 单链是否与 CTL 公式的有限状态自动机的粘贴自动机的 DNA 单链杂交,从而形成完全双链 DNA 分子。如果是,则系统模型满足 CTL 公式描述的性质;否则,系统模型不满足 CTL 公式描述的性质,不能被粘贴自动机接受的局部双链分子就是反例。文献[16]提出了基于无记忆过滤模型的 4 个基本公式即 AFp , EFp , AGp 和 EGp 的模型检测算法,并指出基于无记忆过滤模型的 CTL 模型检测算法能够在线性时间内完成模型检测任务,而在电子计算机上需要多项式的时间。可以看出,CTL 模型检测效率的提升得益于 DNA 分子强大的并行计算能力。文献[17]提出了基于粘贴自动机的 CTL 的 8 个基本公式的模型检测算法。因为 CTL 的一般公式由 8 个基本公式有机、递归地组合而成,所以 DNA 计算框架下的 CTL 模型检测问题已初步得到解决。文献[18]指出,在基于粘贴自动机的模型检测方法中,仅需对 DNA 编码方案做位数扩展即可实现对任意 LTL 公式的模型检测,这点对 CTL 模型检测同样适用。同时,这预示着基于 DNA 计算的 CTL 模型检测方法的完备性有望实现。

3.1.2 从带未来时间算子的公式到带过去时间算子的公式

文献[13,17]仅研究了基于 DNA 计算的带未来时间算子的 CTL 公式的模型检测方法。实际上,完整的时序逻辑应该是对称的,即既包含表示未来的时间算子,又包含表示过去的时间算子。用带过去算子的 CTL 公式描述性质具有自然、简单、易于理解的特点,既增强了 CTL 公式的表达能力^[19],又不增加模型检测问题的复杂性^[20],且在组合规范描述中有

着重要的作用^[21]。文献[22]提出了带过去时间算子的 CTL 基本公式 EOp , AOp , EHp 和 AHp 的模型检测算法。该算法的思想如下:建立系统的带标签的有限状态自动机模型,使用 DNA 单链编码系统状态和变迁,通过混合、杂交生成系统模型的运行路径的 DNA 单链;采用无记忆过滤模型的基本生物操作和 CTL 公式的语义对系统运行路径的 DNA 单链进行过滤,依据过滤结果判定模型检测的结果。该方法可在线性时间内完成带过去时间算子的 CTL 公式的模型检测,而经典算法的时间复杂度为指数阶。该方法不仅适用于检测带过去时间算子的 4 个 CTL 基本公式,还适用于检测其对偶的带未来时间算子的 4 个基本公式。带过去时间算子的 4 个 CTL 基本公式的含义如下:

- 1) EOp :在 M 中至少存在一条路径,在这条路径上 p 曾经成立;
- 2) AOp :在 M 的所有运行路径上均有 p 曾经成立;
- 3) EHp :在 M 中至少存在一条路径,在这条路径上 p 在过去始终成立;
- 4) AHp :在 M 的所有路径上 p 在过去始终成立。

实现带过去算子的 CTL 公式的模型检测有利于建立一套基于 DNA 计算的过去+未来 CTL 模型检测方法。

3.1.3 从 CTL 公式扩展到其他时序逻辑公式

文献[18]和文献[23]提出了 LTL 的 4 个基本公式和 5 个常用性质公式的模型检测的 DNA 计算方法。文献[24]提出了 ITL 和 PTL 的基本公式的模型检测方法,并在生物平台上通过仿真实验验证了该方法的可行性和有效性。在时序逻辑中,LTL 表达系统行为的线性时序关系,CTL 公式描述系统行为的分支时序关系,ITL 公式描述系统行为的简单区间关系,PTL 公式描述间隔以及它们对系统行为影响的一般关系。几种时序逻辑在表达能力上互补,能够更准确地描述系统的时序属性和系统行为的时序关系。在实际应用中,人体活细胞分裂过程中的 DNA 修复和错配被认为是引发癌症的原因之一^[15]。文献[25]构建了细胞内的“可计算基因模型”,其能够对肿瘤分子标志是否同时满足条件 1 和条件 2,满足条件 3 或者满足条件 4 等静态条件进行判断,但不能描述肿瘤分子标志随时间动态变化的性质和时序关系。CTL, LTL, ITL 和 PTL 等时序逻辑公式足以描述肿瘤分子标志随时间动态变化的性质和时序关系,诸如“某时刻满足条件 1,后来又满足条件 2”“曾经满足条件 3”“始终满足条件 4”等,在基因突变引发的疾病的自动诊断方面具有巨大的应用潜力。

表 1 列出了已有基于 DNA 计算的时序逻辑模型检测方法的检测能力。从仅能检测单个 CTL 基本公式到可以检测多种时序逻辑的多个基本公式,方法的检测能力不断地增强。在提升检测方法的检测能力的同时,必须考虑 DNA 生化反应对环境的要求。表 2 列出了已有方法对反应环境的要求。文献[7]由于在 DNA 生化实验中使用了限制性内切酶,对生化反应的温度、PH 值、离子浓度、反应时间等要求比较苛刻,求解的可靠性也受到了影响。文献[13,16-18,22-24]提出的方法因不使用限制性内切酶,降低了对生化反应环境的要求。

表 2 还列出了已有方法在提供反例方面的能力,文献[7]提出的方法在系统模型不满足性质时不能提供反例。文献[13, 16-18, 22-24]的方法将原子命题的 DNA 编码嵌入到系统模型状态的 DNA 编码中,因此能够提供反例。从不能提供反例到能够提供反例,该方法正在逐步完善。

表 1 基于 DNA 计算的 CTL 模型检测方法的检测能力对比

Table 1 Comparison of power among DNA-computing based methods of CTL model checking									
时序逻辑	公式	文献 [7] 方法	文献 [13] 方法	文献 [16] 方法	文献 [17] 方法	文献 [18] 方法	文献 [22] 方法	文献 [23] 方法	文献 [24] 方法
CTL	$AG\rho$	—	—	✓	✓	—	✓	—	✓
	$EG\rho$	—	—	✓	✓	—	✓	—	✓
	$AF\rho$	—	—	✓	✓	—	✓	—	✓
	$EF\rho$	✓	—	✓	✓	—	✓	—	✓
	$AX\rho$	—	✓	—	✓	—	—	—	✓
	$EX\rho$	—	✓	—	✓	—	—	—	✓
	$A[\rho Uq]$	—	—	—	✓	—	—	—	✓
	$E[\rho Uq]$	—	—	—	✓	—	—	—	✓
	$AH\rho$	—	—	—	—	—	✓	—	—
	$EH\rho$	—	—	—	—	—	✓	—	—
LTL	$AO\rho$	—	—	—	—	—	✓	—	—
	$EO\rho$	—	—	—	—	—	✓	—	—
	ρUq	—	—	—	—	✓	—	✓	—
	$F\rho$	—	—	—	—	✓	—	✓	—
	$G\rho$	—	—	—	—	✓	—	✓	—
	$X\rho$	—	✓	—	—	✓	—	—	—
	$G! \rho$	—	—	—	—	✓	—	—	—
ITL	$G(\rho \rightarrow Fq)$	—	—	—	—	✓	—	—	—
	$GF\rho \rightarrow GFq$	—	—	—	—	✓	—	—	—
	$FG\rho \rightarrow GFq$	—	—	—	—	✓	—	—	—
	$(\rho_1 Uq_1); (\rho_2 Uq_2)$	—	—	—	—	—	—	—	✓
	$(\rho Uq)^*$	—	—	—	—	—	—	—	✓
PTL	$((\rho_1 Uq_1), (\rho_2 Uq_2))$	—	—	—	—	—	—	—	✓
	$\text{prj}(\rho_3 \wedge Xq_3)$	—	—	—	—	—	—	—	✓

表 2 基于 DNA 计算的 CTL 模型检测方法对反应环境要求及提供反例能力的对比

Table 2 Comparison of requirements of reaction and capability of providing counter-examples among DNA-computing based methods of CTL model checking									
	文献 [7] 方法	文献 [13] 方法	文献 [16] 方法	文献 [17] 方法	文献 [18] 方法	文献 [22] 方法	文献 [23] 方法	文献 [24] 方法	
不用限制性内切酶	×	✓	✓	✓	✓	✓	✓	✓	✓
提供反例	×	✓	✓	✓	✓	✓	✓	✓	✓

3.2 自治化程度的提升

DNA 计算模型包括计算所使用的 DNA 分子的结构及可实施的生物操作,是实现 CTL 模型检测的关键。DNA 计算模型分为非自治模型和自治模型两大类^[15]。由此,基于 DNA 计算的 CTL 模型检测方法也分为非自治的方法和自治的方法。非自治的方法依托非自治的 DNA 计算模型实现,其原理如图 2 所示。基于人工操作的非自治的 DNA 计算模型有无记忆过滤模型^[15]、记忆过滤模型^[26]、表面模型^[27]、粘贴模型^[28]、剪接模型^[29]等,用于解决如哈密顿路径问题^[30]、SAT 问题^[31-32]、最大团问题^[33]、顶点覆盖问题^[34]、图着色问

题^[35]等复杂的 NP(Non-deterministic Polynomial)困难问题。文献[16, 22]提出了基于无记忆过滤模型的 CTL 模型检测的非自治的计算方法,但其需要人工操作的干预,易错且耗时。自治的方法依托自治的 DNA 计算模型实现,其原理如图 3 所示。自治的 DNA 模型具有分子尺度下自治和部分可编程的特点。DNA 自组装模型^[36]、发夹模型^[37]、有限状态自动机模型^[14, 38-39]、巨磁电阻^[40]和探针机模型^[41]等均属于自治的 DNA 模型。其应用领域包括求解复杂的计算问题^[42-43],构建 DNA 分子逻辑门^[44-47]和生物传感器^[48-49],设计纳米药物递送系统^[50]和纳米材料^[51-52]等。文献[13]、文献[17-18]、文献[23-24]分别提出了几种时序逻辑模型检测的自治方法。这类方法只须设计适当的 DNA 分子,设定分子自组装的物理、化学条件,混合 DNA 软、硬件。当分子被自组装为超分子结构时,即可从中读出表征模型检测结果的信息。从结构上看,自治的方法使用分子自组装的纳米技术取代了传统的人工设计的算法。从功能上看,自治的方法采用了完全不同的计算机制完成计算任务。从非自治的模型检测方法到自治的模型检测方法,方法的自治化程度不断提高。

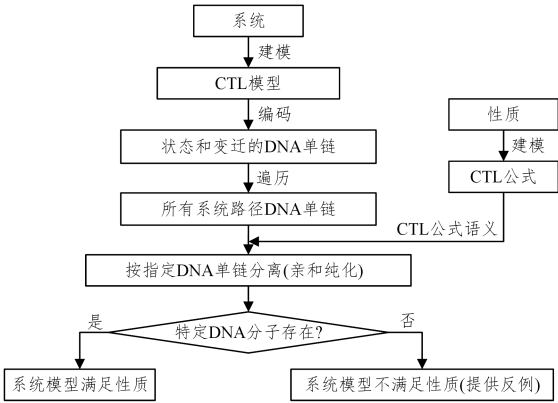


图 2 基于 DNA 计算模型的 CTL 模型检测非自治方法的原理
Fig. 2 Principle of non-autonomous DNA-computing based methods of CTL model checking

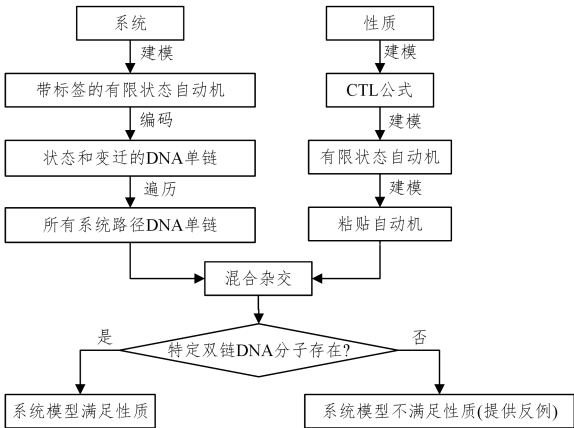


图 3 基于 DNA 计算模型的 CTL 模型检测自治方法的原理
Fig. 3 Principle of autonomous DNA-computing based methods of CTL model checking

3.3 其他进展

在研究基于 DNA 计算的 CTL 模型检测方法的过程中,

解决了 DNA 分子特异性杂交有效性预测效率低的问题,并构建了计算树逻辑公式的 DNA 表示。这对研究基于 DNA 计算的 CTL 模型检测方法具有十分重要的意义。

3.3.1 DNA 分子特异性杂交有效性预测效率的提升

研究者们基于生物仿真平台来验证基于 DNA 计算的 CTL 模型检测方法的可行性、正确性和有效性。NUPACK^[53]是分析与设计 DNA 分子的仿真工具。实验表明,NUPACK 对一组 $n(n \geq 7)$ 个 DNA 单链的杂交有效性的分析时间超过 24h。对于这个问题,文献[54]引入了机器学习方法来提高预测分子特异性杂交有效性的效率。采用提升树算法(Boosting Tree,BT)预测 DNA 分子特异性杂交有效性的具体步骤如下:

- 1) 构造训练样本集;
- 2) 采用 BT 算法学习训练样本集,生成预测模型;
- 3) 用第 2)步生成的预测模型对测试样本集进行预测,并统计特异性杂交的时间和有效性。

其中,构造训练样本集的过程如下:采用 NUPACK 和随机算法分别生成 m 个记录,每个记录包含 n 个 DNA 单链。采用 NUPACK 分析每一个记录中 n 个 DNA 单链杂交的有效性,记录分析所用的时间和完全杂交产生的双链 DNA 分子浓度,重复 $2m$ 次。将 DNA 单链能否特异性杂交作为一个二分类问题,当一组 DNA 单链杂交产生的双链 DNA 分子浓度大于或等于 98%时,认定其能够进行有效的特异性杂交,将其分类标签置为“1”;否则将其分类标签置为“0”。采用 NUPACK 设计的 DNA 单链的杂交有效性通常可达 98%以上,这些 DNA 单链构成了训练集的正例样本子集。采用随机算法设计的 DNA 单链的杂交有效性低于 90%,原因是没有考虑杂交的生物学约束。这些 DNA 单链构成了训练集的负例样本子集。

实验结果表明,BT 算法预测 DNA 分子杂交有效性在速度上比 NUPACK 快 9 万余倍,且准确率达到 94.2%,相比 NUPACK 仅下降不到 6%,在可接受的范围内。这也证明:机器学习算法能够对生物分子特异性杂交的有效性进行预测,且在效率上比 NUPACK 更有优势。

3.3.2 构建计算树逻辑公式的 DNA 表示

要在 DNA 计算框架下实现 CTL 模型检测,CTL 公式的 DNA 表示是前提和基础。构建基于粘贴自动机的 CTL 公式 f 的 DNA 表示的步骤如下:

- 1) 预处理 f ,通过等价变换去掉 f 中的逻辑运算符“与”“蕴含”“等价”,保留“或”,并使否定运算符仅作用于原子命题,由此得到预处理后的公式 f' 。
- 2) 构建 f' 的有限状态自动机 $M(f') = (\Sigma, S, T, s_0, F)$,其中, Σ 是原子命题公式集合, S 是状态集合, T 是变迁集合, s_0 是初始状态, F 是终止状态集合。
- 3) 构建 $M(f')$ 的粘贴自动机,并设计粘贴自动机的启动区 I_1 、终止区 I_2 和间隔区 $X_0, \dots, X_m (m = |S| + 1)$ 的 DNA 单链。对于原子命题及含有否定符号的原子命题,采用定长的 DNA 单链表示。对于状态转移函数 $T(s_i, p) = s_j (p \in \Sigma)$,

用形如 $3'\overline{X_{i+1}} \dots \overline{X_m} \delta(p) \overline{X_0} \dots \overline{X_j} 5'$ 的 DNA 单链表示,其中, $\overline{X}$ 表示 X 的沃森-克里克补链, $\delta(p)$ 表示原子命题 p 的 DNA 单链, $\overline{\delta(p)}$ 表示 $\delta(p)$ 的沃森-克里克补链。对初态 s_i ,采用形如 $3'\overline{I_1} \overline{X_0} \dots \overline{X_i} 5'$ 的 DNA 单链表示。对于每个终态 s_j ,采用形如 $3'\overline{X_{j+1}} \dots \overline{X_m} \overline{I_2} 5'$ 的 DNA 单链表示。

由此得到的 DNA 单链的集合就构成了 CTL 公式 f 的 DNA 表示。该方法可用于求得任意一个 CTL 公式的 DNA 表示。

结束语 CTL 模型检测是形式化验证技术的研究热点。效率和状态空间爆炸问题限制了 CTL 模型检测的应用。研究成果表明,基于 DNA 计算的 CTL 模型检测方法能够提升 CTL 模型检测的效率,缓解状态空间爆炸问题,扩展其应用领域。本文论述了基于 DNA 计算的 CTL 模型检测方法的研究进展。随着检测方法在检测的公式数量、时间维度、多种时序逻辑上的扩展,方法的检测能力不断提升。CTL 模型检测在自治的 DNA 计算模型上的实现,预示着方法的自治化程度不断增强。相关问题的解决则促进了方法的完善。

针对基于 DNA 计算的 CTL 模型检测,未来可从以下几方面展开相关研究工作。

- 1) 扩展基于 DNA 计算的 CTL 模型检测方法,用于检测更多表达能力更强的时序逻辑,如 CTL* 等,进一步提升了现有方法的检测能力。
- 2) 研究基于 DNA 计算的时序逻辑模型检测的新方法。这将有利于精确表达系统模型的性质,也有利于扩展模型检测的应用领域。
- 3) 研究专用于时序逻辑模型检测的 DNA 计算模型,最大限度地发挥 DNA 分子巨大的运算并行性和存储密度大的特点,解决传统电子计算机尚无法解决的模型检测问题。在 DNA 计算机研究中,所建模型的好坏直接影响编码的难易程度、生物操作或生化反应的设计、解空间的大小、计算时间、应用范围以及通用性等^[55]。建立快速的、功能强的、具有一定通用性的模型检测的 DNA 计算模型,应该受到关注。
- 4) 扩展基于 DNA 计算的 CTL 模型检测的应用领域,如研究人体活细胞内的模型检测,为基因突变引起的人类疾病提供自治的、智能的分子诊断和治疗方法。DNA 计算的软、硬件只有在满足人体活细胞所需的物理、化学、生物等条件时才能够嵌入到细胞内,这也是工作的难点。

参 考 文 献

[1] BARNAT J, BAUCH P, BRIM L, et al. Designing fast LTL model checking algorithms for many-core GPUs[J]. Journal of Parallel and Distributed Computing, 2012, 72(9): 1083-1097.

[2] CARBONE R. LTL model-checking for security protocols [J]. AI Communications, 2011, 24(3): 281-283.

[3] FU S, TAYSSIR T. Efficient CTL model-checking for pushdown systems [J]. Theoretical Computer Science, 2014, 549(3): 127-145.

[4] BRIM L, ČEŠKA M, ŠAFRÁNEK D. Model checking of biologi-

cal systems[M]// Formal Methods for Dynamical Systems. Berlin: Springer-Verlag, 2013: 63-112.

[5] HOU G, ZHOU K J, YONG J W, et al. Survey of state explosion problem in model checking [J]. Computer Science, 2013, 40(S1): 77-86. (in Chinese)

侯刚, 周宽久, 勇嘉伟, 等. 模型检测中状态爆炸问题研究综述 [J]. 计算机科学, 2013, 40(S1): 77-86.

[6] XU J. Forthcoming era of biological computer [J]. Bulletin of Chinese Academy of Sciences, 2014, 29(1): 42-54. (in Chinese)

许进. 生物计算机时代即将来临 [J]. 中国科学院院刊, 2014, 29(1): 42-54.

[7] EMERSON E A, HAGER K D, KONIECZKA J H. Molecular model checking [J]. International Journal of Foundations of Computer Science, 2006, 17(4): 733-741.

[8] CLARKE E M, GRUMBERG O, HIRAISHI H, et al. Verification of the futurebus+ cache coherence protocol [J]. Formal Methods in System Design, 1995, 6(2): 217-232.

[9] ANDERSON R J, BEAME P, BURNS S, et al. Model checking large software specifications [J]. IEEE Transactions on Software Engineering, 1998, 24(7): 498-520.

[10] CLARKE E M, EMERSON E A. Design and synthesis of synchronization skeletons using branching-time temporal logic[C]// The Workshop on Logic of Programs. Berlin: Springer, 1981, 131: 52-71.

[11] QUEILLE J P, SIFAKIS J. Specification and verification of concurrent systems in CESAR [C]// Colloquium on International Symposium on Programming. Berlin: Springer-Verlag, 1982: 337-351.

[12] CLARKE E M, GRUMBERG O, PELED D A. Model checking [M]. Massachusetts: The MIT Press, 1999: 35-36.

[13] HAN Y J, ZHU W J, JIAO L F, et al. DNA computing methods of LTL model checking on Xp [J]. Mini-Micro Systems, 2017, 38(3): 553-558. (in Chinese)

韩英杰, 朱维军, 焦林枫, 等. 线性时序逻辑公式 Xp 模型检测的 DNA 计算方法 [J]. 小型微型计算机系统, 2017, 38(3): 553-558.

[14] MARTÍNEZ-PÉREZ I M, ZIMMERMANN K H, IGNATOVA Z. An autonomous DNA model for finite state automata [J]. International Journal of Bioinformatics Research & Applications, 2009, 5(1): 81-96.

[15] ZIMMERMANN K H, ISRAEL M P, ZOYA I. DNA computing models [M]. New York: Springer, 2008: 127-128.

[16] ZHU W J, HAN Y J, ZHOU Q L. Performing CTL model checking via DNA computing [J/OL]. <https://doi.org/10.1007/s00500-018-3314-7>.

[17] ZHU W J, WANG Y F, ZHOU Q L, et al. Model checking computational tree logic using sticker automata [M]// Bio-Inspired Computing-Theories and Applications. Singapore: Springer, 2016: 12-20.

[18] ZHU W J, ZHOU Q L, ZHANG Q X. A LTL model checking approach based on DNA computing [J]. Chinese Journal of Computers, 2016, 39(12): 2578-2597. (in Chinese)

朱维军, 周清雷, 张钦宪. 基于 DNA 计算的线性时序逻辑模型检测方法 [J]. 计算机学报, 2016, 39(12): 2578-2597.

[19] KUPFERMAN O, PNUELI A. Once and for all [C]// Proceedings of Tenth Annual IEEE Symposium on Logic in Computer Science. New Jersey: IEEE Computer Society Press, 1995: 25-35.

[20] LICHTENSTEIN O, PNUELI A, ZUCK L. The glory of the past [M]// Logics of Programs. Berlin: Springer, 1985: 196-218.

[21] EMERSON E A. Temporal and modal logic [M]// Handbook of Theoretical Computer Science. Massachusetts: The MIT Press, 1990: 995-1072.

[22] HAN Y J, ZHOU Q L, JIAO L F, et al. Model checking for computation tree logic with past based on DNA computing [M]// Bio-Inspired Computing-Theories and Applications. Singapore: Springer, 2017: 131-147.

[23] ZHU W J, ZHOU Q L, LI Y L. LTL model checking based on DNA computing [J]. Acta Electronica Sinica, 2016, 44(6): 1265-1271. (in Chinese)

朱维军, 周清雷, 李永亮. 以 DNA 为载体的线性时序逻辑模型检测 [J]. 电子学报, 2016, 44(6): 1265-1271.

[24] ZHU W J, FENG C, WU H M. Model checking temporal logic formulas using sticker automata [J]. BioMed Research International, 2017, 2017(3): 1-33.

[25] MARTÍNEZ-PÉREZ I M, ZHANG G, IGNATOVA Z, et al. Computational genes: a tool for molecular diagnosis and therapy of aberrant mutational phenotype [J]. BMC Bioinformatics, 2007, 8(1): 365-365.

[26] ADLEMAN L M. On constructing a molecular computer [M]// Lipton R J, Baum E B, eds. DNA Based Computers. American Mathematical Society, 1996: 1-21.

[27] LIU Q, WANG L, FRUTOS A G, et al. DNA computing on surfaces [J]. Nature, 2000, 403(6766): 175-178.

[28] ROWEIS S, WINFREE E, BURGOYNE R, et al. A sticker-based model for DNA computation [J]. Journal of Computational Biology, 1998, 5(4): 615-629.

[29] HEAD T. Splicing Schemes and DNA [C]// Lindenmayer Systems: Impacts on theoretical computer science, computer graphics and developmental biology. Heidelberg: Springer, 1992: 371-383.

[30] ADLEMAN L M. Molecular computation of solutions to combinatorial problems [J]. Science, 1994, 266(5187): 1021-1023.

[31] LIPTON R J. DNA solution of hard computational problems [J]. Science, 1995, 268(5210): 542-545.

[32] BRAICH R S, CHELYAPOV N, JOHNSON C, et al. Solution of a 20-variable 3-SAT problem on a DNA computer [J]. Science, 2002, 296(5567): 499-502.

[33] OUYANG Q, KAPLAN P D, LIU S, et al. DNA solution of the maximal clique problem [J]. Science, 1997, 278(5337): 446-449.

[34] GAO L, XU J. DNA solution of vertex cover problem based on sticker model [J]. Chinese Journal of Electronics, 2002, 11(2): 280-284.

[35] LI K L,ZHOU X,XU J. An $O(2^n)$ Volume molecular solutions for the 3-colorable problem on DNA—based supercomputing [J]. Acta Electronica Sinica,2008,36(11):2096-2101. (in Chinese)
李肯立,周旭,许进. 图 3-着色问题的 $O(2^n)$ 链数 DNA 计算机算法[J]. 电子学报,2008,36(11):2096-2101.

[36] WINFREE E,LIU F,WENZLER L A,et al. Design and self-assembly of two-dimensional DNA crystals [J]. Nature,1998,394(6693):539-544.

[37] SAKAMOTO K,GOUZU H,KOMIYA K,et al. Molecular computation by DNA hairpin formation [J]. Science,2000,288(6469):1223-1226.

[38] BENENSON Y,PAZ-ELIZUR T,ADAR R,et al. Programmable and autonomous computing machine made of biomolecules [J]. Nature,2001,414(6862):430-439.

[39] KURAMOCHI J,SAKAKIBARA Y. Intensive in vitro experiments of implementing and executing finite automata in test tube[M]//DNA Computing. Berlin:Springer,2005:193-202.

[40] XIAO J H,XU J. The DNA computation model based on giant magnetoresistance for sat problem [J]. Chinese Journal of Computers,2013,36(4):829-835. (in Chinese)
肖建华,许进. 可满足性问题的巨磁电阻型 DNA 计算模型[J]. 计算机学报,2013,36(4):829-835.

[41] XU J. Probe machine [J]. IEEE Transactions on Neural Networks and Learning Systems,2016,27(7):1405-1416.

[42] LI K L,LUO X,WU F,et al. An algorithm in tile assembly model for maximum clique problem [J]. Journal of Computer Research and Development,2013,50(3):666-675. (in Chinese)
李肯立,罗兴,吴帆,等. 基于自组装模型的最大团问题 DNA 计算算法[J]. 计算机研究与发展,2013,50(3):666-675.

[43] ZHANG G,IGNATOVA Z. Biomolecular autonomous solution of the Hamiltonian path problem via hairpin formation [J]. International Journal of Bioinformatics Research & Applications,2005,1(4):389-398.

[44] YANG J,DONG C,DONG Y F,et al. Logic nanoparticle beacon triggered by the binding-induced effect of multiple inputs[J]. ACS Applied Materials & Interfaces,2014,6(16):14486-14492.

[45] YANG J,JIANG S X,LIU X R,et al. Aptamer-Binding Directed DNA Origami Pattern for Logic Gates[J]. ACS Applied Materials & Interfaces,2016,8(49):34054-34060.

[46] PAN L Q,WANG Z Y,LI Y F,et al. Nicking enzyme-controlled toehold regulation for DNA logic circuits[J]. Nanoscale,2017,9(46):18223-18228.

[47] YANG J,WU R F,LI Y F,et al. Entropy-driven DNA logic circuits regulated by DNzyme[J]. Nucleic Acids Research,2018,46(16):8532-8541.

[48] SHI X L,WANG Z Y,DENG C Y,et al. A novel bio-sensor based on DNA strand displacement [J]. Plos One,2014,9(10),e108856-e108856.

[49] LI X,SONG T,SHI X L,et al. A universal fast colorimetric method for DNA signal detection with DNA strand displacement and gold nanoparticles [J]. Journal of Nanomaterials,2015(1/2):1-9.

[50] LI X,LI H,SONG T,et al. Highly biocompatible drug-delivery systems based on DNA nanotechnology [J]. Journal of Biomedical Nanotechnology,2017,13(7):747-757.

[51] SHI X L,CHEN C Z,LI X,et al. Size controllable DNA nanoribbons assembled from three types of reusable brick single-strand DNA tiles [J]. Soft matter,2015,11(43):8484-8492.

[52] SHI X L,WU X X,SONG T,et al. Construction of DNA nanotubes with controllable diameters and patterns by using hierarchical DNA sub-tiles [J]. Nanoscale,2016,8(31):14785-14792.

[53] The NUPACK team. NUPACK [EB/OL]. <http://www.nupack.org>.

[54] ZHU W J,HAN Y J,WU H M,et al. Predicting the results of molecular specific hybridization using boosted tree algorithm[J/OL]. https://www.researchgate.net/publication/327760975_Predicting_the_results_of_molecular_specific_hybridization_using_boosted_tree_algorithm.

[55] XU J,TAN G J,FAN Y K,et al. DNA computer principle,advances and difficulties(IV): on the models of DNA computer [J]. Chinese Journal of Computers,2007,30(6):881-893. (in Chinese)
许进,谭钢军,范月科,等. DNA 计算机原理、进展及难点(IV): 论 DNA 计算机模型[J]. 计算机学报,2007,30(6):881-893.