

密级:_____



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

中国科学院大学学位论文 L^AT_EX 模板

作者姓名:_____刘坚_____

指导教师:_____蒋颖 研究员_____

_____中国科学院软件研究所_____

学位类别:_____工学博士_____

学科专业:_____计算机软件与理论_____

培养单位:_____中国科学院软件研究所_____

2018 年 02 月

L^AT_EX Thesis Template
of
The University of Chinese Academy of Sciences

by
Jian Liu

A thesis submitted to
The University of Chinese Academy of Sciences
in partial fulfillment of the requirements
for the degree of
Doctor of Computer Software and Theory

Institute of Software, Chinese Academy of Sciences

February, 2018

学位论文独创性声明

本人郑重声明：我所呈交的学位论文是本人在导师指导下进行的研究工作及所取得的研究成果。尽我所知，除了文中已经标注引用的内容外，本论文中不含其他个人或集体已经发表或撰写过的研究成果。对本文的研究做出贡献的个人和集体，均已在文中作了明确的说明或致谢。本人知道本声明的法律结果由自己承担。

作者签名：_____ 日期：_____

关于学位论文使用授权的说明

本人完全了解中国科学院软件研究所有关保留、使用学位论文的规定，即：中国科学院软件研究所有权保留送交论文的复印件，允许论文被查阅和借阅；可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

（保密的论文在解密后应遵守此规定）

作者签名：_____ 导师签名：_____ 日期：_____

摘 要

本文是中国科学院大学学位论文模板 ucasthesis 的使用说明文档。主要内容为介绍 L^AT_EX 文档类 ucasthesis 的用法，以及如何使用 L^AT_EX 快速高效地撰写学位论文。

关键词：中国科学院大学，学位论文，L^AT_EX 模板

Abstract

This paper is a help documentation for the \LaTeX class ucasthesis, which is a thesis template for the University of Chinese Academy of Sciences. The main content is about how to use the ucasthesis, as well as how to write thesis efficiently by using \LaTeX .

Keywords: University of Chinese Academy of Sciences (UCAS), Thesis, \LaTeX Template

目 录

摘 要	vii
Abstract	ix
目 录	xi
图形列表	xiii
表格列表	xv
符号列表	xvii
第一章 引言	1
第二章 定理证明与模型检测的结合	3
2.1 CTL_P	3
2.2 SCTL	6
2.3 SCTLProV	14
2.4 案例分析与实验结果	21
参考文献	37
发表学术论文	39
简历	41
致 谢	43

图形列表

2.1	无人车可能的所在位置	6
2.2	SCTL(\mathcal{M})	8
2.3	SCTLProV.	15
2.4	CPT 的重写规则.....	16
2.5	A scratch of process A and process B.	21
2.6	The input file “mutual.model”.	22
2.7	Visualization of the proof tree and the Kripke model in the illustrative example.	25
2.8	A simple solution of the mutual exclusion problem.....	25
2.9	The input file “mutual_solution.model” indicating the solution.	26
2.10	SCA zones, where right and left are relative to the pilot facing the runway, i.e., opposite from the reader point of view ^[1]	27
2.11	Properties $P_{01}, P_{02}, \dots, P_{12}$ to be checked in benchmark #1, #2, and #3.	30
2.12	Average verification time in benchmark #1.	31
2.13	Average memory usage in benchmark #1.	31
2.14	Average verification time in benchmark #1 and #2.	32
2.15	Average memory usage in benchmark #1 and #2.	32
2.16	Average verification time in SCTLProV vs. SCTLProV _R	33
2.17	Properties to be verified in benchmark #4.	34

表格列表

2.1	SCTLProV vs. SCTLProV _R	32
2.2	Solvable cases in Verds, NuSMV, NuXMV, and SCTLProV.	35
2.3	Cases where SCTLProV both runs faster and uses less memory.	35
2.4	Test cases where SCTLProV uses less time and less memory, respectively.	36

符号列表

Characters

Symbol	Description	Unit
R	the gas constant	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
C_v	specific heat capacity at constant volume	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
C_p	specific heat capacity at constant pressure	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
E	specific total energy	$\text{m}^2 \cdot \text{s}^{-2}$
e	specific internal energy	$\text{m}^2 \cdot \text{s}^{-2}$
h_T	specific total enthalpy	$\text{m}^2 \cdot \text{s}^{-2}$
h	specific enthalpy	$\text{m}^2 \cdot \text{s}^{-2}$
k	thermal conductivity	$\text{kg} \cdot \text{m} \cdot \text{s}^{-3} \cdot \text{K}^{-1}$
T	temperature	K
t	time	s
p	thermodynamic pressure	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
\hat{p}	hydrostatic pressure	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
\mathbf{f}_b	body force	$\text{kg} \cdot \text{m}^{-2} \cdot \text{s}^{-2}$
S	boundary surface	m^2
V	volume	m^3
\mathbf{V}	velocity vector	$\text{m} \cdot \text{s}^{-1}$
u	x component of velocity	$\text{m} \cdot \text{s}^{-1}$
v	y component of velocity	$\text{m} \cdot \text{s}^{-1}$
w	z component of velocity	$\text{m} \cdot \text{s}^{-1}$
c	speed of sound	$\text{m} \cdot \text{s}^{-1}$
\mathbf{r}	position vector	m
\mathbf{n}	unit normal vector	1
$\hat{\mathbf{t}}$	unit tangent vector	1
$\tilde{\mathbf{t}}$	unit bitangent vector	1
C_R	coefficient of restitution	1
Re	Reynolds number	1
Pr	Prandtl number	1

Ma	Mach number	1
α	thermal diffusivity	$\text{m}^2 \cdot \text{s}^{-1}$
μ	dynamic viscosity	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-1}$
ν	kinematic viscosity	$\text{m}^2 \cdot \text{s}^{-1}$
γ	heat capacity ratio	1
ρ	density	$\text{kg} \cdot \text{m}^{-3}$
σ_{ij}	stress tensor	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
S_{ij}	deviatoric stress tensor	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
τ_{ij}	viscous stress tensor	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
δ_{ij}	Kronecker tensor	1
I_{ij}	identity tensor	1

Operators

Symbol	Description
Δ	difference
∇	gradient operator
δ^\pm	upwind-biased interpolation scheme

Abbreviations

Acronym	Description
ANFO	Ammonium Nitrate Fuel Oil
CFD	Computational Fluid Dynamics
CFL	Courant-Friedrichs-Lewy
CJ	Chapman-Jouguet
EOS	Equation of State
JWL	Jones-Wilkins-Lee
TVD	Total Variation Diminishing
WENO	Weighted Essentially Non-oscillatory
ZND	Zel'dovich-von Neumann-Doering

第一章 引言

第二章 定理证明与模型检测的结合

本章首先介绍的逻辑系统 CTL_P , CTL_P 是计算树逻辑 CTL 的一个扩展; 然后介绍的是针对 CTL_P 的一个证明系统 SCTL; 之后介绍的是对证明系统 SCTL 的一个实现 SCTLProV, 最后介绍的是案例分析, 以及相关实验结果。

2.1 CTL_P

我们用逻辑 $CTL_P(\mathcal{M})$ 来刻画要验证的系统 \mathcal{M} 的性质, 其中 \mathcal{M} 通常指的是一个 Kripke 结构, 其定义如下。

定义 2.1.1 (Kripke 结构). 一个 Kripke 结构 $\mathcal{M} = (\mathcal{S}, \longrightarrow, \mathcal{P})$ 包含如下三个部分:

1. \mathcal{S} 是一个有穷的状态集合;
2. $\longrightarrow \subseteq \mathcal{S} \times \mathcal{S}$ 是一个二元关系; 对于每一个状态 $s \in \mathcal{S}$, 至少存在一个 $s' \in \mathcal{S}$ 使得 $s \longrightarrow s'$;
3. \mathcal{P} 是一个有穷的关系符号的集合; 对于每个关系符号 $P \in \mathcal{P}$, 都存在自然数 n 使得 $P \in \mathcal{S}^n$ 。

对于一个状态 $s \in \mathcal{S}$, 我们将 s 的所有的下一个状态的集合定义为

$$\text{Next}(s) = \{s' \mid s \longrightarrow s'\}.$$

一个路径是一个有穷或无穷的状态序列, 通常形式为 s_0, \dots, s_n 或者 s_0, s_1, \dots , 其中, 对于任意自然数 i , 如果 s_i 不是该序列的最后一个元素, 那么就有 $s_{i+1} \in \text{Next}(s_i)$ 。

我们称 T 是一棵路径树当且仅当对于 T 上的所有由 s 标记的非叶子节点, 该节点的所有后继节点正好由 $\text{Next}(s)$ 中的所有元素一一标记。一棵路径树上的所有节点既可以是有穷个也可以是无穷个。

语法。 一个 Kripke 结构 \mathcal{M} 的性质由 $CTL_P(\mathcal{M})$ 公式表示:

定义 2.1.2. 对于一个给定的 Kripke 模型 $\mathcal{M} = (\mathcal{S}, \longrightarrow, \mathcal{P})$, $CTL_P(\mathcal{M})$ 公式的语法定义如下:

$$\phi := \left\{ \begin{array}{l} \top \mid \perp \mid P(t_1, \dots, t_n) \mid \neg P(t_1, \dots, t_n) \mid \phi \wedge \phi \mid \phi \vee \phi \mid \\ AX_x(\phi)(t) \mid EX_x(\phi)(t) \mid AF_x(\phi)(t) \mid EG_x(\phi)(t) \mid \\ AR_{x,y}(\phi_1, \phi_2)(t) \mid EU_{x,y}(\phi_1, \phi_2)(t) \end{array} \right.$$

其中, x 与 y 为变量, 取值范围为 S , 而 t_1, \dots, t_n 既可以是代表状态的常量, 也可以是取值范围为 S 的变量。

在定义 2.1.2 中, 我们用模态词来绑定公式中的变量。比如, 模态词 AX , EX , AF 以及 EG 在公式 ϕ 中绑定了变量 x ; 而模态词 AR 和 EU 则在公式 ϕ_1 和 ϕ_2 中分别绑定了变量 x 和 y . 变量的替换则写为 $(t/x)\phi$, 表示将公式 ϕ 中所有自由出现的变量 x 都替换为 t 。

不失一般性地来说, 我们假定所有的否定符号都出现在原子命题上; 而且有如下缩写:

- $\phi_1 \Rightarrow \phi_2 \equiv \neg\phi_1 \vee \phi_2$,
- $EF_x(\phi)(t) \equiv EU_{z,x}(\top, \phi)(t)$,
- $ER_{x,y}(\phi_1, \phi_2)(t) \equiv EU_{y,z}(\phi_2, ((z/x)\phi_1 \wedge (z/y)\phi_2))(t) \vee EG_y(\phi_2)(t)$, 其中变量 z 既不在 ϕ_1 , 也不在 ϕ_2 中出现,
- $AG_x(\phi)(t) \equiv \neg(EF_x(\neg\phi)(t))$,
- $AU_{x,y}(\phi_1, \phi_2)(t) \equiv \neg(ER_{x,y}(\neg\phi_1, \neg\phi_2)(t))$.

我们称模态词 AF , EF , AU , 以及 EU 为归纳模态词; 模态词 AR , ER , AG , 以及 EG 为余归纳模态词。

语义。 相应地, 对于一个给定的 Kripke 模型 \mathcal{M} , $\text{CTL}_P(\mathcal{M})$ 的语义定义如下:

- $\mathcal{M} \models P(s_1, \dots, s_n)$: 如果 $\langle s_1, \dots, s_n \rangle \in P$, 而且 P 是一个 \mathcal{M} 上的 n 元关系;
- $\mathcal{M} \models \neg P(s_1, \dots, s_n)$: 如果 $\langle s_1, \dots, s_n \rangle \notin P$, 而且 P 是一个 \mathcal{M} 上的 n 元关系;
- $\mathcal{M} \models \top$ 永远成立;
- $\mathcal{M} \models \perp$ 永远不成立;
- $\mathcal{M} \models \phi_1 \wedge \phi_2$: 如果 $\mathcal{M} \models \phi_1$ 和 $\mathcal{M} \models \phi_2$ 同时成立;
- $\mathcal{M} \models \phi_1 \vee \phi_2$: 如果 $\mathcal{M} \models \phi_1$ 成立, 或者 $\mathcal{M} \models \phi_2$ 成立;
- $\mathcal{M} \models AX_x(\phi_1)(s)$: 如果对于每个状态 $s' \in \text{Next}(s)$, 都有 $\mathcal{M} \models (s'/x)\phi_1$ 成立;
- $\mathcal{M} \models EX_x(\phi_1)(s)$: 如果存在一个状态 $s' \in \text{Next}(s)$, 使得 $\mathcal{M} \models (s'/x)\phi_1$ 成立;

- $\mathcal{M} \models AF_x(\phi_1)(s)$: 如果存在一个有无穷个节点的树 T , 而且 T 的根节点是 s , 那么对于 T 的任何一个非叶子节点 s' , s' 的子节点为 $\text{Next}(s')$, 对于 T 的任何一个叶子节点 s' , $\vdash (s'/x)\phi_1$ 成立;
- $\mathcal{M} \models EG_x(\phi_1)(s)$: 如果存在 \mathcal{M} 上的一个无穷路径 s_0, s_1, \dots (其中 $s_0 = s$), 那么对于任意的自然数 i , 都有 $\mathcal{M} \models (s_i/x)\phi_1$ 成立;
- $\mathcal{M} \models AR_{x,y}(\phi_1, \phi_2)(s)$: 如果存在一棵路径树 T , T 的根节点由 s 标记, 对于任意节点 $s' \in T$ 都有 $\mathcal{M} \models (s'/y)\phi_2$ 成立, 而且对于任意的叶子节点 $s'' \in T$ 都有 $\mathcal{M} \models (s''/x)\phi_1$ 成立;
- $\mathcal{M} \models EU_{x,y}(\phi_1, \phi_2)(s)$: 如果存在一个无穷路径 s_0, s_1, \dots (其中 $s_0 = s$) 和一个自然数 j , $\mathcal{M} \models (s_j/y)\phi_2$ 成立, 而且对于任意的自然数 $i < j$ 都有 $\mathcal{M} \models (s_i/x)\phi_1$ 成立。

CTL vs. CTL_P。在计算树逻辑 (CTL)^[2,3] 的语法中, 原子公式通常用命题符号来表示, 而命题符号在计算树逻辑的语义中通常解释为一个 Kripke 结构上的状态集合。在逻辑系统 CTL_P 中, 相比于计算树逻辑, 我们通过引入多元谓词来增加逻辑系统中公式的表达能力。CTL_P 相比于 CTL 的表达能力的提升可由如下的例子表示出来:

例子 2.1.1. 本例子受多机器人路径规划系统^[4,5]启发。在原例子中, 多机器人路径规划系统的规范可以写成 CTL 公式: 在一个多个区块的地图上, 每从初始位置出发的机器人都能到达指定的最终位置, 而且在行进的同时, 每个机器人都会避免经过某些位置。

在本例子中, 除了 CTL 所能表示的时序性质之外, 我们考虑一种“空间”性质, 即表示状态之间的关系。

假定有一个无人车正在一个星球表面行走, 这个星球的表面已经被分成了有无穷个小的区域。无人车一次能从一个区域行走到另一个区域, 那么我们将无人车的位置看作成一个状态, 无人车所有的可能的所在位置则可看作为状态空间, 而且无人车从一个位置到另一个位置的移动规律则可看作成迁移关系。无人车的设计需要满足一个基本的性质, 即无人车不能永远在一个很小的范围内移动。准确地说, 对于给定的距离 σ , 在任意状态 s , 随着无人车的移动会到达状态 s' , 使得 s 和 s' 的位置之间的距离大于 σ 。该性质可以由公式 $AG_x(AF_y(D_\sigma(x, y))(x))(s_0)$ 来刻画, 其中 s_0 是初始状态, 即无人车的降落点; 原子公式 $D_\sigma(x, y)$ 则刻画了一种空间性质, 即状态 x 和 y 的位置的距离大于 σ 。

例子2.1.1中的性质可以很容易由 CTL_P 中的公式进行刻画, 然而很难用传统的时序逻辑的公式进行表示。原因是在传统的时序逻辑的语法中通常没有表述一

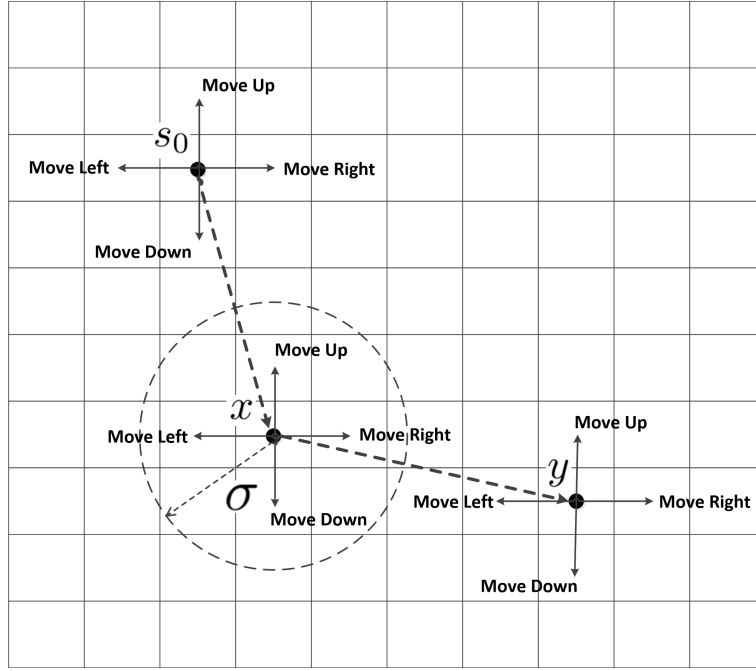


图 2.1: 无人车可能的所在位置

个特定的状态或者多个状态之间的关系的机制，即使在语义中，传统的时序逻辑通常只考虑当前的状态，而无法考虑多个状态之间的关系。

2.2 SCTL

在本节，我们针对逻辑 $\text{CTL}_P(\mathcal{M})$ 给出一个证明系统 $\text{SCTL}(\mathcal{M})$ (Sequent-calculus-like proof system for CTL_P)。在通常意义下的证明系统中，一个公式是可证的当且仅当该公式在所有的模型中都成立，而在 $\text{SCTL}(\mathcal{M})$ 中，一个公式是可证的当且仅当该公式在模型 \mathcal{M} 中是可证的。

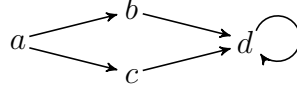
首先，让我们考虑一个 $\text{CTL}_P(\mathcal{M})$ 公式 $AF_x(P(x))(s)$ 。该公式在模型 \mathcal{M} 中成立当且仅当存在一个路径树 T ， T 的根节点由 s 标记，而且 T 上的每个叶子节点都满足 P 。

然后，我们考虑一个具有嵌套模态词的 $\text{CTL}_P(\mathcal{M})$ 公式 $AF_x(AF_y(P(x,y))(x))(s)$ 。如果试图说明该公式在模型 \mathcal{M} 中是成立的，那么就需要找到一个路径树 T ，使得 T 的根节点由 s 标记，而且对于 T 中的所有叶子节点 a ， $AF_y(P(a,y))(a)$ 是成立的。为了说明 $AF_y(P(a,y))(a)$ 是成立的，则需要又找到一棵路径树 T' 使得 T' 的根节点由 a 标记，而且 T' 上的所有叶子节点 b 都满足 $P(a,b)$ 。我们可以用以下的两个规则来刻画当前的嵌套的路径树。

$$\frac{\vdash (s/x)\phi}{\vdash AF_x(\phi)(s)} \text{ AF-R}_1$$

$$\frac{\vdash AF_x(\phi)(s_1) \quad \dots \quad \vdash AF_x(\phi)(s_n)}{\vdash AF_x(\phi)(s)} \text{AF-R}_2 \quad \{s_1, \dots, s_n\} = \text{Next}(s)$$

例子 2.2.1. 假设一个模型有如下图所示的迁移规则,



和一个原子谓词 $P = \{b, c\}$, 那么公式 $AF_x(P(x))(a)$ 的一个证明如下。

$$\frac{\frac{\overline{\vdash P(b)}}{\vdash AF_x(P(x))(b)} \text{AF-R}_1 \quad \frac{\overline{\vdash P(c)}}{\vdash AF_x(P(x))(c)} \text{AF-R}_1}{\vdash AF_x(P(x))(a)} \text{AF-R}_2$$

在此证明树中, 除了 AF-R_1 和 AF-R_2 , 我们还应用了如下规则。

$$\frac{}{\vdash P(s_1, \dots, s_n)} \text{atom-R} \quad \langle s_1, \dots, s_n \rangle \in P$$

例子 2.2.2. 假设另一个模型, 该模型的迁移规则与例子2.2.1中相同, 除此之外还有原子谓词 $Q = \{(b, d), (c, d)\}$ 。公式 $AF_x(AF_y(Q(x, y))(x))(a)$ 的证明如下。

$$\frac{\frac{\frac{\overline{Q(b, d)}}{\vdash AF_y(Q(b, y))(d)} \text{AF-R}_1}{\vdash AF_y(Q(b, y))(b)} \text{AF-R}_2 \quad \frac{\frac{\overline{Q(c, d)}}{\vdash AF_y(Q(c, y))(d)} \text{AF-R}_1}{\vdash AF_y(Q(c, y))(c)} \text{AF-R}_2}{\frac{\vdash AF_x(AF_y(Q(x, y))(x))(b)}{\vdash AF_x(AF_y(Q(x, y))(x))(a)} \text{AF-R}_1 \quad \frac{\vdash AF_x(AF_y(Q(x, y))(x))(c)}{\vdash AF_x(AF_y(Q(x, y))(x))(a)} \text{AF-R}_2} \text{AF-R}_2$$

在 SCTL 中, 每个相继式都有 $\Gamma \vdash \phi$ 形式, 其中 Γ 是一个可能为空的 SCTL 公式集合, ϕ 是一个 SCTL 公式。不同于通常的相继式演算,

So, as all sequents have the form $\vdash \phi$, the left rules and the axiom rule can be dropped as well. In other words, unlike the usual sequent calculus and like Hilbert systems, SCTL is tailored for deduction, not for hypothetical deduction.

As the left-hand side of sequents is not used to record hypotheses, we will use it to record a different kind of information, that occur in the case of co-inductive modalities, such as the modality EG .

Indeed, the case of the co-inductive formula, for example $EG_x(P(x))(s)$, is more complex than that of the inductive one, such as $AF_x(P(x))(s)$. To justify its validity, one needs to provide an infinite sequence starting from s , and each state

in the infinite sequence verifies P . However, as the model is finite, we can always restrict to regular sequences and use a finite representation of such sequences. This leads us to introduce a rule, called EG-merge, that permits to prove a sequent of the form $\vdash EG_x(P(x))(s)$, provided such a sequent already occurs lower in the proof. To make this rule local, we re-introduce hypotheses Γ to record part of the history of the proof. The sequent have therefore the form $\Gamma \vdash \phi$, with a non empty Γ in this particular case only, and the EG-merge rule is then just an instance of the axiom rule, that must be re-introduced in this particular case only.

SCTL(\mathcal{M}) 的证明规则如图2.2所示。

$\frac{}{\vdash P(s_1, \dots, s_n)} \text{atom-R}_{\langle s_1, \dots, s_n \rangle \in P}$		$\frac{}{\vdash \neg P(s_1, \dots, s_n)} \neg\text{-R}_{\langle s_1, \dots, s_n \rangle \notin P}$	
$\frac{}{\vdash \top} \top\text{-R}$	$\frac{\vdash \phi_1 \quad \vdash \phi_2}{\vdash \phi_1 \wedge \phi_2} \wedge\text{-R}$	$\frac{\vdash \phi_1}{\vdash \phi_1 \vee \phi_2} \vee\text{-R}_1$	$\frac{\vdash \phi_2}{\vdash \phi_1 \vee \phi_2} \vee\text{-R}_2$
$\frac{\vdash (s'/x)\phi}{\vdash EX_x(\phi)(s)} \text{EX-R}_{s' \in \text{Next}(s)}$		$\frac{\vdash (s_1/x)\phi \quad \dots \quad \vdash (s_n/x)\phi}{\vdash AX_x(\phi)(s)} \text{AX-R}_{\{s_1, \dots, s_n\} = \text{Next}(s)}$	
$\frac{\vdash (s/x)\phi}{\vdash AF_x(\phi)(s)} \text{AF-R}_1$		$\frac{\vdash AF_x(\phi)(s_1) \quad \dots \quad \vdash AF_x(\phi)(s_n)}{\vdash AF_x(\phi)(s)} \text{AF-R}_2_{\{s_1, \dots, s_n\} = \text{Next}(s)}$	
$\frac{\vdash (s/x)\phi \quad \Gamma, EG_x(\phi)(s) \vdash EG_x(\phi)(s')}{\Gamma \vdash EG_x(\phi)(s)} \text{EG-R}_{s' \in \text{Next}(s)}$		$\frac{}{\Gamma \vdash EG_x(\phi)(s)} \text{EG-merge}_{EG_x(\phi)(s) \in \Gamma}$	
$\frac{\vdash (s/y)\phi_2 \quad \Gamma' \vdash AR_{x,y}(\phi_1, \phi_2)(s_1) \quad \dots \quad \Gamma' \vdash AR_{x,y}(\phi_1, \phi_2)(s_n)}{\Gamma \vdash AR_{x,y}(\phi_1, \phi_2)(s)} \text{AR-R}_1_{\{s_1, \dots, s_n\} = \text{Next}(s), \Gamma' = \Gamma, AR_{x,y}(\phi_1, \phi_2)(s)}$			
$\frac{\vdash (s/x)\phi_1 \quad \vdash (s/y)\phi_2}{\Gamma \vdash AR_{x,y}(\phi_1, \phi_2)(s)} \text{AR-R}_2$		$\frac{}{\Gamma \vdash AR_{x,y}(\phi_1, \phi_2)(s)} \text{AR-merge}_{AR_{x,y}(\phi_1, \phi_2)(s) \in \Gamma}$	
$\frac{\vdash (s/y)\phi_2}{\vdash EU_{x,y}(\phi_1, \phi_2)(s)} \text{EU-R}_1$		$\frac{\vdash (s/x)\phi_1 \quad \vdash EU_{x,y}(\phi_1, \phi_2)(s')}{\vdash EU_{x,y}(\phi_1, \phi_2)(s)} \text{EU-R}_2_{s' \in \text{Next}(s)}$	

图 2.2: SCTL(\mathcal{M})

有效性与完备性。 命题2.2.1和命题2.2.2的作用是将有穷结构转换为无穷结构，这两个命题被用来证明 SCTL 的有效性；命题2.2.3和命题2.2.4的作用是将无穷结构转换到有穷结构，这两个命题被用来证明 SCTL 的完备性。

命题 2.2.1 (有穷状态序列到无穷状态序列). 给定一个有穷的状态序列 s_0, \dots, s_n , 其中对于任意 $0 \leq i \leq n-1$ 都有 $s_i \rightarrow s_{i+1}$, 而且存在 $0 \leq p \leq n-1$ 使得 $s_n = s_p$ 。那么, 一定存在一个无穷的状态序列 s'_0, s'_1, \dots 使得 $s_0 = s'_0$, 而且对于任意 $i \geq 0$ 都有 $s'_i \rightarrow s'_{i+1}$, 同时此无穷状态序列中的每个状态都在 s_0, \dots, s_n 中。

证明. 本命题所述无穷序列为: $s_0, \dots, s_{p-1}, s_p, \dots, s_{n-1}, s_p, \dots$, 其中 $s_0 = s'_0$. \square

命题 2.2.2 (有穷路径树到无穷路径树). 设 Φ 为一个状态集合, T 为一个有穷的路径树, T 的每个叶子节点都由某个状态 s 来标记, 其中, $s \in \Phi$; 或者存在从 T 的根结点到当前叶子节点的分支上的一个节点, 使得该节点同样由 s 所标记。那么, 一定存在一棵可能无穷的路径树 T' , 而且 T' 的所有叶子节点都由 Φ 中的某个状态标记, 同时用来标记 T' 节点的状态都用来标记 T 的节点。

证明. 令 T' 的根结点为 T 的根结点, 而且对于 T 的每个节点的标记 s 来说, 如果 $s \in \Phi$, 那么 s 标记 T' 的叶子节点; 否则, s 的后继节点分别由 $\text{Next}(s)$ 中的每个元素标记。显然, 标记 T' 中节点的状态都标记 T 中的节点。 \square

命题 2.2.3 (无穷状态序列到有穷状态序列). 给定一个无穷状态序列 s_0, s_1, \dots , 其中对于任意 $i \geq 0$ 都有 $s_i \rightarrow s_{i+1}$ 。那么, 一定存在一个有穷的状态序列 s'_0, \dots, s'_n , 对于任意 $0 \leq i \leq n-1$, 都存在一个 $0 \leq p \leq n-1$, 使得 $s'_i = s'_p$, 而且 s'_0, \dots, s'_n 中的所有状态都在 s_0, s_1, \dots 中出现。

证明. 由于 Kripke 模型的状态集是有穷的, 因此在状态序列 s_0, s_1, \dots 一定存在 $p, n \geq 0$, 使得 $s_p = s_n$ 。本命题所述有穷状态序列即为 s_0, \dots, s_n 。 \square

命题 2.2.4 (可能无穷的路径树到有穷路径树). 设 Φ 为一个状态集合; T 为一个可能无穷的路径树, 其中 T 的所有叶子节点都由 Φ 中的某个状态所标记。那么, 一定存在一个有穷的路径树 T' , 使得对于 T' 的每个叶子节点的标记 s , $s \in \Phi$, 或者存在从 T' 的根结点到该叶子节点的分支上的一个节点, 该节点同样由 s 标记。

证明. 由于 Kripke 模型的状态集是有穷的, 因此对于 T 的每个无穷分支, 都存在 $0 \leq p < n$, 使得 $s_p = s_n$ 。将 T 的每个这样的无穷分支在 s_n 处截断, 所得到的路径树即为 T' 。显然, 由于 T' 具有有穷个分支, 同时 T' 的每个分支都是有穷的, 因此 T' 也是有穷的。 \square

定理 2.2.1 (有效性). 设 \mathcal{M} 为一个 Kripke 模型, ϕ 为一个 $\text{CTL}_P(\mathcal{M})$ 闭公式。如果相继式 $\vdash \phi$ 具有一个证明, 则 $\mathcal{M} \models \phi$ 成立。

证明. 假设相继式 $\vdash \phi$ 具有证明 π , 以下对证明 π 的结构做归纳:

- 如果 π 的最后一条规则为 atom-R, 那么 $\vdash \phi$ 具有 $\vdash P(s_1, \dots, s_n)$ 形式, 因此 $\mathcal{M} \models P(s_1, \dots, s_n)$ 。
- 如果 π 的最后一条规则为 \neg -R, 那么 $\vdash \phi$ 具有 $\vdash \neg P(s_1, \dots, s_n)$ 形式, 因此 $\mathcal{M} \models \neg P(s_1, \dots, s_n)$ 。
- 如果 π 的最后一条规则为 \top -R, 那么 $\vdash \phi$ 具有 $\vdash \top$ 形式, 因此 $\mathcal{M} \models \top$ 。
- 如果 π 的最后一条规则为 \wedge -R, 那么 $\vdash \phi$ 具有 $\vdash \phi_1 \wedge \phi_2$ 形式。根据归纳假设, $\mathcal{M} \models \phi_1$ 与 $\mathcal{M} \models \phi_2$ 均成立, 因此 $\mathcal{M} \models \phi_1 \wedge \phi_2$ 。
- 如果 π 的最后一条规则为 \vee -R, 那么 $\vdash \phi$ 具有 $\vdash \phi_1 \vee \phi_2$ 形式。根据归纳假设, $\mathcal{M} \models \phi_1$ 成立或 $\mathcal{M} \models \phi_2$ 成立, 因此 $\mathcal{M} \models \phi_1 \vee \phi_2$ 。
- 如果 π 的最后一条规则为 AX-R, 那么 $\vdash \phi$ 具有 $\vdash AX_x(\phi_1)(s)$ 形式。根据归纳假设, 对于任意 $s' \in \text{Next}(s)$, 都有 $\mathcal{M} \models (s'/x)\phi_1$ 成立, 因此 $\mathcal{M} \models AX_x(\phi_1)(s)$ 。
- 如果 π 的最后一条规则为 EX-R, 那么 $\vdash \phi$ 具有 $\vdash EX_x(\phi_1)(s)$ 形式。根据归纳假设, 存在 $s' \in \text{Next}(s)$, 使得 $\mathcal{M} \models (s'/x)\phi_1$ 成立, 因此 $\mathcal{M} \models EX_x(\phi_1)(s)$ 。
- 如果 π 的最后一条规则为 AF-R₁ 或 AF-R₂, 那么 $\vdash \phi$ 具有 $\vdash AF_x(\phi_1)(s)$ 形式。根据证明 π , 我们利用归纳的方式构造一棵路径树 $|\pi|$ 。构造方式如下:
 - 如果 π 的最后一条规则为 AF-R₁, 而且 ρ 为 $\vdash (s/x)\phi_1$ 的证明, 则路径树 $|\pi|$ 只包含一个节点 s ;
 - 如果 π 的最后一条规则为 AF-R₂, 而且 π_1, \dots, π_n 分别为 $\vdash AF_x(\phi_1)(s_1), \dots, \vdash AF_x(\phi_n)(s_n)$ 的证明, 其中 $\{s_1, \dots, s_n\} = \text{Next}(s)$, 那么令 $|\pi|$ 等于 $s(|\pi_1|, \dots, |\pi_n|)$ 。

路径树 $|\pi|$ 的根结点为 s , 而且对于 $|\pi|$ 的每个叶子节点 s' 来说, $\vdash (s'/x)\phi_1$ 都有一个比 π 小的证明。根据归纳假设, 对于 $|\pi|$ 的每个叶子节点 s' 来说, 都有 $\mathcal{M} \models (s'/x)\phi_1$ 成立, 因此, $\mathcal{M} \models AF_x(\phi_1)(s)$ 成立。

- 如果 π 的最后一条规则为 EG-R, 则 $\vdash \phi$ 具有 $\vdash EG_x(\phi_1)(s)$ 形式。根据证明 π , 我们归纳构造一个状态序列 $|\pi|$ 。构造方式如下:
 - 如果 π 的最后一条规则为 EG-merge, 那么 $|\pi|$ 只包含一个单独的状态 s ;
 - 如果 π 的最后一条规则为 EG-R, 而且 ρ 和 π_1 分别为 $\vdash (s/x)\phi_1$ 和 $\Gamma, EG_x(\phi_1)(s) \vdash EG_x(\phi_1)(s')$ 的证明, 其中 $s' \in \text{Next}(s)$, 那么令 $|\pi|$ 等于 $s|\pi_1|$ 。

对于状态序列 $|\pi| = s_0, \dots, s_n$, $s_0 = s$; 对于任意 $0 \leq i \leq n-1$, $s_i \longrightarrow s_{i+1}$; 对于任意 $0 \leq i \leq n$, $\vdash (s_i/x)\phi_1$ 都有一个比 π 小的证明; 而且存在 $p < n$ 使得 $s_n = s_p$ 。根据归纳假设, 对于任意 $i \geq 0$, 都有 $\mathcal{M} \models (s_i/x)\phi_1$ 成立。由命题2.2.1可知, 存在一个无穷的状态序列 s'_0, s'_1, \dots , 其中对于任意 $i \geq 0$ 都有 $s'_i \longrightarrow s'_{i+1}$, 同时 $\mathcal{M} \models (s'_i/x)\phi_1$ 成立。因此, $\mathcal{M} \models EG_x(\phi_1)(s)$ 成立。

- 如果 π 的最后一条规则为 AR-R₁ 或 AR-R₂, 那么 $\vdash \phi$ 具有 $\vdash AR_x(\phi_1, \phi_2)(s)$ 形式。根据 π , 我们归纳构造一个有穷的路径树 $|\pi|$ 。构造方式如下:
 - 如果 π 的最后一条规则为 AR-R₁, 而且 ρ_1 和 ρ_2 分别为 $\vdash (s/x)\phi_1$ 和 $\vdash (s/x)\phi_2$ 的证明, 那么 $|\pi|$ 只包含一个节点 s ;
 - 如果 π 的最后一条规则为 AR-merge, 那么 $|\pi|$ 只包含一个节点 s ;
 - 如果 π 的最后一条规则为 AR-R₂, 而且 $\rho, \pi_1, \dots, \pi_n$ 分别为 $\vdash (s/y)\phi_2$, $\Gamma, AR_{x,y}(\phi_1, \phi_2)(s) \vdash AR_{x,y}(\phi_1, \phi_2)(s_1), \dots, \Gamma, AR_{x,y}(\phi_1, \phi_2)(s) \vdash AR_{x,y}(\phi_1, \phi_2)(s_n)$ 的证明, 其中 $\{s_1, \dots, s_n\} = \text{Next}(s)$, 那么令 $|\pi|$ 等于 $s(|\pi_1|, \dots, |\pi_n|)$ 。

路径树 $|\pi|$ 以 s 为根结点, 而且对于 $|\pi|$ 的每个节点 s' 来说, $\vdash (s'/y)\phi_2$ 都有一个比 π 小的证明; 对于 $|\pi|$ 的任意叶子节点 s' 来说, $\vdash (s'/x)\phi_1$ 有一个比 π 小的证明, 或者在从 $|\pi|$ 的根结点到当前叶子节点的分支上存在一个节点, 使得 s' 标记此节点。根据归纳假设, 对于 $|\pi|$ 的任意节点 s' , $\vdash (s'/y)\phi_2$ 成立, 而且对于 $|\pi|$ 的任意叶子节点 s' , $\vdash (s'/x)\phi_1$ 成立, 或者在从 $|\pi|$ 的根结点到当前叶子节点的分支上存在一个节点, 使得 s' 标记此节点。根据命题2.2.2, 存在一个可能无穷的路径树 T' , 使得对于 T' 的每个节点 s' , 都有 $\vdash (s'/y)\phi_2$ 成立, 而且对于 T' 的每个叶子节点 s' , 都有 $\vdash (s'/x)\phi_1$ 成立。因此, $\vdash AR_{x,y}(\phi_1, \phi_2)(s)$ 成立。

- 如果 π 的最后一条规则为 EU-R₁ 或 EU-R₂, 那么 $\vdash \phi$ 具有 $\vdash EU_{x,y}(\phi_1, \phi_2)(s)$ 形式。根据 π , 我们归纳构造一个有穷状态序列 $|\pi|$ 。构造过程如下:
 - 如果 π 的最后一条规则为 EU-R₁, 那么 $|\pi|$ 只包含一个状态 s ;
 - 如果 π 的最后一条规则为 EU-R₂, 而且 ρ 和 π_1 分别为 $\vdash (s/x)\phi_1$ 和 $\vdash EU_{x,y}(\phi_1, \phi_2)(s')$ 的证明, 那么令 $|\pi|$ 等于 $s|\pi_1|$ 。

在状态序列 $|\pi| = s_0, \dots, s_n$ 中, $s_0 = s$; 对于任意 $0 \leq i \leq n-1$, $s_i \longrightarrow s_{i+1}$; 对于任意 $0 \leq i \leq n-1$, $\vdash (s_i/x)\phi_1$ 有一个比 π 小的证明; 而且 $\vdash (s_n/y)\phi_2$ 有一个比 π 小的证明。根据归纳假设, 对任意 $0 \leq i \leq n-1$, $\vdash (s_i/x)\phi_1$ 和 $\vdash (s_n/y)\phi_2$ 均成立。因此, $\vdash EU_{x,y}(\phi_1, \phi_2)(s)$ 成立。

- π 的最后一条规则不能为 merge 规则。

□

定理 2.2.2 (完备性). 设 ϕ 是一个 $\text{CTL}_P(\mathcal{M})$ 闭公式。如果 $\mathcal{M} \models \phi$, 则 $\vdash \phi$ 在 $\text{SCTL}(\mathcal{M})$ 中是可证的。

证明. 对 ϕ 的结构作归纳:

- 如果 $\phi = P(s_1, \dots, s_n)$, 那么由 $\mathcal{M} \models P(s_1, \dots, s_n)$ 可知, $\vdash P(s_1, \dots, s_n)$ 是可证的。
- 如果 $\phi = \neg P(s_1, \dots, s_n)$, 那么由 $\mathcal{M} \models \neg P(s_1, \dots, s_n)$ 可知, $\vdash \neg P(s_1, \dots, s_n)$ 是可证的。
- 如果 $\phi = \top$, 那么显然 $\vdash \top$ 是可证的。
- 如果 $\phi = \perp$, 那么显然 $\vdash \perp$ 是不可证的。
- 如果 $\phi = \phi_1 \wedge \phi_2$, 那么由于 $\mathcal{M} \models \phi_1 \wedge \phi_2$, 因此 $\mathcal{M} \models \phi_1$ 和 $\mathcal{M} \models \phi_2$ 均成立。根据归纳假设, $\vdash \phi_1$ 和 $\vdash \phi_2$ 均可证。因此, $\vdash \phi_1 \wedge \phi_2$ 是可证的。
- 如果 $\phi = \phi_1 \vee \phi_2$, 那么由于 $\mathcal{M} \models \phi_1 \vee \phi_2$, 因此 $\mathcal{M} \models \phi_1$ 或 $\mathcal{M} \models \phi_2$ 成立。根据归纳假设, $\vdash \phi_1$ 或 $\vdash \phi_2$ 是可证的。因此, $\vdash \phi_1 \vee \phi_2$ 是可证的。
- 如果 $\phi = AX_x(\phi_1)(s)$, 那么由于 $\mathcal{M} \models AX_x(\phi_1)(s)$, 因此对于任意 $s' \in \text{Next}(s)$, 都有 $\mathcal{M} \models (s'/x)\phi_1$ 成立。根据归纳假设, 对于任意 $s' \in \text{Next}(s)$, $\vdash (s'/x)\phi_1$ 都是可证的。因此, $\vdash AX_x(\phi_1)(s)$ 是可证的。
- 如果 $\phi = EX_x(\phi_1)(s)$, 那么由于 $\mathcal{M} \models EX_x(\phi_1)(s)$, 因此存在 $s' \in \text{Next}(s)$ 使得 $\mathcal{M} \models (s'/x)\phi_1$ 成立。根据归纳假设, $\vdash (s'/x)\phi_1$ 是可证的, 因此 $\vdash EX_x(\phi_1)(s)$ 是可证的。
- 如果 $\phi = AF_x(\phi_1)(s)$, 那么由于 $\mathcal{M} \models AF_x(\phi_1)(s)$, 因此存在一棵有穷的路径树 T , 并且 T 以 s 为根结点; 对于 T 的每个非叶子节点 s' , s' 的后继节点分别由 $\text{Next}(s)$ 中的元素所标记; 对于 T 的每个叶子节点 s' , 都有 $\mathcal{M} \models (s'/x)\phi_1$ 成立。根据归纳假设, $\vdash (s'/x)\phi_1$ 是可证的。然后, 对于 T 的每个以子树 T' (设 T' 的根结点为 s'), 我们归纳构造 $\vdash AF_x(\phi_1)(s')$ 的一个证明 $|T'|$ 。构造过程如下:
 - 如果 T' 只包含一个节点 s' , 那么 $|T'|$ 的最后一条规则为 AF-R_1 , 同时 $|T'|$ 中包含 $\vdash (s'/x)\phi_1$ 的证明;

- 如果 $T' = s'(T_1, \dots, T_n)$, 那么 $|T'|$ 的最后一条规则为 AF-R₂, 同时 $|T'_1|, \dots, |T'_n|$ 分别为 $\vdash AF_x(\phi_1)(s_1), \dots, \vdash AF_x(\phi_1)(s_n)$ 的证明, 其中 $\{s_1, \dots, s_n\} = \text{Next}(s)$ 。

因此, $|T|$ 是 $\vdash AF_x(\phi_1)(s)$ 的一个证明。

- 如果 $\phi = EG_x(\phi_1)(s)$, 那么由于 $\mathcal{M} \models EG_x(\phi_1)(s)$, 因此存在一个状态序列 s_0, \dots, s_n 使得 $s_0 = s$, 而且对于任意 $0 \leq i \leq n$ 都有 $\mathcal{M} \models (s_i/x)\phi_1$ 成立。根据归纳假设, $\vdash (s_i/x)\phi_1$ 是可证的。根据命题2.2.3, 存在一个有穷的状态序列 $T = s_0, \dots, s_n$ 使得对任意 $0 \leq i \leq n-1$, $s_i \rightarrow s_{i+1}$, 同时 $\vdash (s_i/x)\phi_1$ 是可证的, 而且存在 $p < n$ 使得 $s_n = s_p$ 。对于 T 的每个后缀 s_i, \dots, s_n , 我们归纳构造 $|s_i, \dots, s_n|$ 为 $EG_x(\phi_1)(s_0), \dots, EG_x(\phi_1)(s_{i-1}) \vdash EG_x(\phi_1)(s_i)$ 的证明。构造方式如下:

- $|s_n|$ 的最后一条规则为 EG-merge;
- 如果 $i \leq n-1$, 根据归纳假设, 由于 $\vdash (s_i/x)\phi_1$ 是可证的, 而且 $|s_{i+1}, \dots, s_n|$ 是 $EG_x(\phi_1)(s_0), \dots, EG_x(\phi_1)(s_i) \vdash EG_x(\phi_1)(s_{i+1})$ 的一个证明。因此, $|s_i, \dots, s_n|$ 是 $EG_x(\phi_1)(s_0), \dots, EG_x(\phi_1)(s_{i-1}) \vdash EG_x(\phi_1)(s_i)$ 的一个证明, 而且最后一条规则为 EG-R。

因此, $|s_0, \dots, s_n|$ 是 $\vdash EG_x(\phi_1)(s)$ 的一个证明。

- 如果 $\phi = AR_{x,y}(\phi_1, \phi_2)(s)$, 那么由于 $\mathcal{M} \models AR_{x,y}(\phi_1, \phi_2)(s)$, 因此存在一棵以 s 为根节点的可能无穷的路径树, 对于该路径树的每个节点 s' , 都有 $\mathcal{M} \models (s'/x)\phi_2$; 对于该路径树的每个叶子节点 s' , 都有 $\mathcal{M} \models (s'/y)\phi_1$ 。根据归纳假设, 对于该路径树的每个节点 s' , $\vdash (s'/y)\phi_2$ 是可证的, 而且对于该路径树的每个叶子节点 s' , $\vdash (s'/y)\phi_1$ 是可证的。由命题2.2.4可知, 存在一棵有穷的路径树 T , 对于该路径树的每个节点 s' , $\vdash (s'/y)\phi_2$ 是可证的; 对于该路径树的每个叶子节点 s' , $\vdash (s'/y)\phi_1$ 是可证的, 或者 s' 为从 T 的根节点到该叶子节点分支上的节点。然后, 对于 T 的每个子树 T' , 我们归纳构造 $AR_{x,y}(\phi_1, \phi_2)(s_1), \dots, AR_{x,y}(\phi_1, \phi_2)(s_m) \vdash AR_{x,y}(\phi_1, \phi_2)(s')$ 的一个证明 $|T'|$, 其中 s' 为 T' 的根节点, 而且 s_1, \dots, s_m 为从 T 的根节点到 T' 的根节点的分支。构造方式如下:

- 如果 T' 只包含一个单独的节点 s' , 同时 $\vdash (s'/x)\phi_1$ 是可证的, 那么根据归纳假设, $\vdash (s'/x)\phi_1$ 和 $\vdash (s'/y)\phi_2$ 皆可证, 而且 $|T'|$ 的最后一条规则为 AR-R₁;
- 如果 T' 只包含一个单独的节点, 同时 s' 包含在 s_1, \dots, s_m 中, 那么 $|T'|$ 的最后一条规则为 AR-merge;

- 如果 $T' = s'(T_1, \dots, T_n)$, 那么根据归纳假设, $|T_1|, \dots, |T_n|$ 分别为

$$\begin{aligned} & AR_{x,y}(\phi_1, \phi_2)(s_1), \dots, AR_{x,y}(\phi_1, \phi_2)(s_m), \\ & AR_{x,y}(\phi_1, \phi_2)(s') \vdash AR_{x,y}(\phi_1, \phi_2)(s'_1) \\ & \dots \\ & AR_{x,y}(\phi_1, \phi_2)(s_1), \dots, AR_{x,y}(\phi_1, \phi_2)(s_m), \\ & AR_{x,y}(\phi_1, \phi_2)(s') \vdash AR_{x,y}(\phi_1, \phi_2)(s'_n) \end{aligned}$$

的证明, 同时 $|T'|$ 的最后一条规则为 AR-R₂, 其中 $s'_1, \dots, s'_n = \text{Next}(s')$ 。

因此, $|T|$ 是 $\vdash AR_{x,y}(\phi_1, \phi_2)(s)$ 的一个证明。

- 如果 $\phi = EU_{x,y}(\phi_1, \phi_2)(s)$, 那么由于 $\mathcal{M} \models EU_{x,y}(\phi_1, \phi_2)(s)$, 因此存在一个有穷的状态序列 $T = s_0, \dots, s_n$ 使得 $\mathcal{M} \models (s_n/y)\phi_2$ 成立, 而且对于任意 $0 \leq i \leq n-1$, $\mathcal{M} \models (s_i/x)\phi_1$ 成立。根据归纳假设, $\vdash (s_n/y)\phi_2$ 是可证的, 而且对于任意 $0 \leq i \leq n-1$, $\vdash (s_i/x)\phi_1$ 是可证的。然后, 对于 T 的每个后缀 s_i, \dots, s_n , 我们归纳构造 $|s_i, \dots, s_n|$ 为 $\vdash EU_{x,y}(\phi_1, \phi_2)(s_i)$ 的证明。构造方式如下:

- $|s_n|$ 的最后一条规则为 EU-R₁;
- 如果 $i \leq n-1$, 那么根据归纳假设, 由于 $|s_{i+1}, \dots, s_n|$ 是 $\vdash EU_{x,y}(\phi_1, \phi_2)(s_{i+1})$ 的证明, 而且 $\vdash (s_i/x)\phi_1$ 是可证的, 因此, $|s_i, \dots, s_n|$ 是 $\vdash EU_{x,y}(\phi_1, \phi_2)(s_i)$ 的证明。

因此, $|s_0, \dots, s_n|$ 是 $\vdash EU_{x,y}(\phi_1, \phi_2)(s)$ 的一个证明。

□

2.3 SCTLProV

本节介绍 SCTL 的一个实现—SCTLProV (图2.3) 以及该工具与其他模型检测工具的对比。SCTLProV 的工作方式如下: 首先, SCTLProV 读入一个输入文件, 并将该输入文件解析到一个 Kripke 模型以及若干个 SCTL 公式; 然后, 对于每个公式, SCTLProV 搜索该公式的证明, 如果该公式可证, 则并输出该证明 (或者只输出 True), 如果该公式不可正, 则输出该公式的非的证明 (或者只输出 False)。

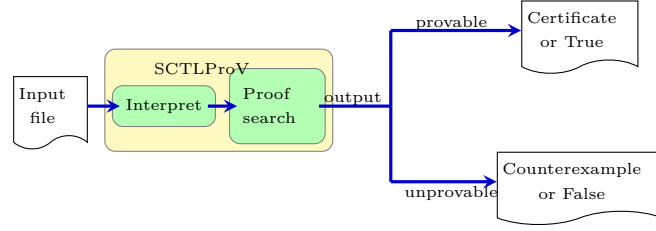


图 2.3: SCTLProV.

2.3.1 证明搜索

SCTLProV 的证明搜索方法如下：首先，对于要证明的相继式，以及 SCTL 规则将该公式的所有的前提给定一个序；然后，依次对这些前提进行证明搜索。我们将以上证明搜索方法定义成一系列对于连续传递树（定义2.3.1）的重写规则。下面我们介绍连续传递树的概念。

2.3.1.1 连续传递树

在连续传递树中，连续是一个基本的概念。在计算机程序设计语言理论^[6,7]中，连续是计算机程序将要执行的部分的显示表示。

定义 2.3.1 (连续传递树). 一个连续传递树（Continuation Passing Tree, 简称为 CPT）指的是一棵同时满足以下条件的二叉树：

- 每个叶子节点被 t 或 f 标记，其中 t 和 f 是不同的两个符号；
- 每个非叶子节点都被一个 SCTL 相继式标记。

对于 CPT 的每个非叶子节点来说，它的左子树称之为该节点的 t -连续；它的右子树称之为该节点的 f -连续。对于一个 CPT c 来说，若 c 的根节点为 $\Gamma \vdash \phi$ ，以及 c 的 t -连续和 f -连续分别为 c_1 和 c_2 ，那么我们将 c 记作 $\text{cpt}(\Gamma \vdash \phi, c_1, c_2)$ ，或者可以表示为如下形式：

$$\begin{array}{c} \Gamma \vdash \phi \\ \wedge \\ c_1 \quad c_2 \end{array}$$

在 SCTLProV 的证明搜索算法可总结为：对于给定的 SCTL 相继式 $\vdash \phi$ ，我们构造一个连续传递树 $c = \text{cpt}(\vdash \phi, t, f)$ ，然后根据图2.4所示的重写规则将 c 重写到 t 或 f 。如果 c 最终重写到 t ，那么 $\vdash \phi$ 是可证的；如果 c 最终重写到 f ，那么 $\vdash \phi$ 是不可证的。

在 CPT 的重写规则中，对一个 CPT c 的一步重写只需判断 c 的根节点，而与 c 的子表达式无关。例如，根据重写规则， $\text{CPT}_{\text{cpt}(\vdash \phi_1 \wedge \phi_2, t, f)}$ 重写到 $\text{cpt}(\vdash$

$\text{cpt}(\vdash \top, c_1, c_2) \rightsquigarrow c_1$	$\text{cpt}(\vdash \perp, c_1, c_2) \rightsquigarrow c_2$
$\text{cpt}(\vdash P(s_1, \dots, s_n), c_1, c_2) \rightsquigarrow c_1$	$[\langle s_1, \dots, s_n \rangle \in P]$
$\text{cpt}(\vdash P(s_1, \dots, s_n), c_1, c_2) \rightsquigarrow c_2$	$[\langle s_1, \dots, s_n \rangle \notin P]$
$\text{cpt}(\vdash \neg P(s_1, \dots, s_n), c_1, c_2) \rightsquigarrow c_2$	$[\langle s_1, \dots, s_n \rangle \in P]$
$\text{cpt}(\vdash \neg P(s_1, \dots, s_n), c_1, c_2) \rightsquigarrow c_1$	$[\langle s_1, \dots, s_n \rangle \notin P]$
$\text{cpt}(\vdash \phi_1 \wedge \phi_2, c_1, c_2) \rightsquigarrow \text{cpt}(\vdash \phi_1, \text{cpt}(\vdash \phi_2, c_1, c_2), c_2)$	
$\text{cpt}(\vdash \phi_1 \vee \phi_2, c_1, c_2) \rightsquigarrow \text{cpt}(\vdash \phi_1, c_1, \text{cpt}(\vdash \phi_2, c_1, c_2))$	
$\text{cpt}(\vdash AX_x(\phi)(s), c_1, c_2) \rightsquigarrow \text{cpt}(\vdash (s_1/x)\phi, \text{cpt}(\vdash (s_2/x)\phi, \text{cpt}(\dots \text{cpt}(\vdash (s_n/x)\phi, c_1, c_2), \dots, c_2), c_2))$	$[\{s_1, \dots, s_n\} = \text{Next}(s)]$
$\text{cpt}(\vdash EX_x(\phi)(s), c_1, c_2) \rightsquigarrow \text{cpt}(\vdash (s_1/x)\phi, c_1, \text{cpt}(\vdash (s_2/x)\phi, c_1, \text{cpt}(\dots \text{cpt}(\vdash (s_n/x)\phi, c_1, c_2) \dots)))$	$[\{s_1, \dots, s_n\} = \text{Next}(s)]$
$\text{cpt}(\Gamma \vdash AF_x(\phi)(s), c_1, c_2) \rightsquigarrow c_2$	$[AF_x(\phi)(s) \in \Gamma]$
$\text{cpt}(\Gamma \vdash AF_x(\phi)(s), c_1, c_2) \rightsquigarrow$ $\text{cpt}(\vdash (s/x)\phi, c_1, \text{cpt}(\Gamma' \vdash AF_x(\phi)(s_1), \text{cpt}(\dots \text{cpt}(\Gamma' \vdash AF_x(\phi)(s_n), c_1, c_2) \dots, c_2), c_2))$	$[\{s_1, \dots, s_n\} = \text{Next}(s), AF_x(\phi)(s) \notin \Gamma, \text{ and } \Gamma' = \Gamma, AF_x(\phi)(s)]$
$\text{cpt}(\Gamma \vdash EG_x(\phi)(s), c_1, c_2) \rightsquigarrow c_1$	$[EG_x(\phi)(s) \in \Gamma]$
$\text{cpt}(\Gamma \vdash EG_x(\phi)(s), c_1, c_2) \rightsquigarrow$ $\text{cpt}(\vdash (s/x)\phi, \text{cpt}(\Gamma' \vdash EG_x(\phi)(s_1), c_1, \text{cpt}(\dots \text{cpt}(\Gamma' \vdash EG_x(\phi)(s_n), c_1, c_2) \dots), c_2))$	$[\{s_1, \dots, s_n\} = \text{Next}(s), EG_x(\phi)(s) \notin \Gamma, \text{ and } \Gamma' = \Gamma, EG_x(\phi)(s)]$
$\text{cpt}(\Gamma \vdash AR_{x,y}(\phi_1, \phi_2)(s), c_1, c_2) \rightsquigarrow c_1$	$[(AR_{x,y}(\phi_1, \phi_2)(s) \in \Gamma)]$
$\text{cpt}(\Gamma \vdash AR_{x,y}(\phi_1, \phi_2)(s), c_1, c_2) \rightsquigarrow$ $\text{cpt}(\vdash (s/y)\phi_2, \text{cpt}(\vdash (s/x)\phi_1, c_1, \text{cpt}(\Gamma' \vdash AR_{x,y}(\phi_1, \phi_2)(s_1), \text{cpt}(\dots \text{cpt}(\Gamma' \vdash AR_{x,y}(\phi_1, \phi_2)(s_n), c_1, c_2) \dots, c_2), c_2))$	$[\{s_1, \dots, s_n\} = \text{Next}(s), AR_{x,y}(\phi_1, \phi_2)(s) \notin \Gamma, \text{ and } \Gamma' = \Gamma, AR_{x,y}(\phi_1, \phi_2)(s)]$
$\text{cpt}(\Gamma \vdash EU_{x,y}(\phi_1, \phi_2)(s), c_1, c_2) \rightsquigarrow c_2$	$[EU_{x,y}(\phi_1, \phi_2)(s) \in \Gamma]$
$\text{cpt}(\Gamma \vdash EU_{x,y}(\phi_1, \phi_2)(s), c_1, c_2) \rightsquigarrow$ $\text{cpt}(\vdash (s/y)\phi_2, c_1, \text{cpt}(\vdash (s/x)\phi_1, \text{cpt}(\Gamma' \vdash EU_{x,y}(\phi_1, \phi_2)(s_1), c_1, \text{cpt}(\dots \text{cpt}(\Gamma' \vdash EU_{x,y}(\phi_1, \phi_2)(s_n), c_1, c_2) \dots), c_2))$	$[\{s_1, \dots, s_n\} = \text{Next}(s), EU_{x,y}(\phi_1, \phi_2)(s) \notin \Gamma, \text{ and } \Gamma' = \Gamma, EU_{x,y}(\phi_1, \phi_2)(s)]$

图 2.4: CPT 的重写规则.

$\phi_1, \text{cpt}(\vdash \phi_2, \mathbf{t}, \mathbf{f}), \mathbf{f})$, 此步重写意味着: 如果搜索 $\vdash \phi_1$ 的证明成功, 则继续搜索 $\vdash \phi_2$ 的证明; 如果搜索 $\vdash \phi_1$ 的证明失败, 则直接判定 $\vdash \phi_1 \wedge \phi_2$ 不可证。接下来, 根据 ϕ_1 的结构, 继续对 $\text{cpt}(\vdash \phi_1, \text{cpt}(\vdash \phi_2, \mathbf{t}, \mathbf{f}), \mathbf{f})$ 进行重写。

2.3.2 证明搜索的可终止性

在证明利用图2.3.1表示的重写规则, 使得 SCTLProV 的证明搜索是可终止的之前, 我们需要引入以下定义和命题。

定义 2.3.2 (字典路径序 (lexicographic path ordering)^[8,9]). 设 \succeq 是函数符号集合 F 的一个拟序 (quasi-ordering), 其中 F 的每个符号的元数 (arity) 是固定不变的。集合 $T(F)$ (由 F 生成的项的集合) 上的字典路径序 \succeq_{lpo} 的归纳定义如下:

$s = f(s_1, \dots, s_m) \succeq_{\text{lpo}} g(t_1, \dots, t_n) = t$ 当且仅当以下至少一条断言成立:

- 存在 $i \in \{1, \dots, m\}$, 使得 $s_i \succeq_{\text{lpo}} t$ 成立。
- 对于任意 $j \in \{1, \dots, n\}$, $f \succ g$ 和 $s \succ_{\text{lpo}} t_j$ 同时成立。
- 对于任意 $j \in \{2, \dots, n\}$, $f = g$, $(s_1, \dots, s_m) \succeq'_{\text{lpo}} (t_1, \dots, t_n)$ 和 $s \succ_{\text{lpo}} t_j$ 都成立, 其中 \succeq'_{lpo} 是由 \succeq_{lpo} 产生的字典序。

命题 2.3.1 (字典路径序的良基性). 如果 \succeq 是函数符号集合 F 的一个拟序 (quasi-ordering), 其中 F 的每个符号的元数 (arity) 是固定不变的, 那么根据集合 $T(F)$ (由 F 生成的项的集合) 上的字典路径序 \succeq_{lpo} 是良基的当且仅当 \succeq 是良基的。

证明. 证明由 Dershowitz 提出, 参考[8]. □

定义 2.3.3 (相继式的权重). 假设一个 Kripke 模型的状态集的基数为 n ; $\Gamma \vdash \phi$ 是一个 $\text{SCTL}(\mathcal{M})$ 相继式; $|\phi|$ 是公式 ϕ 的大小; $|\Gamma|$ 是 Γ 的基数。相继式 $\Gamma \vdash \phi$ 的权重为

$$w(\Gamma \vdash \phi) = \langle |\phi|, (n - |\Gamma|) \rangle$$

命题 2.3.2 (可终止性). 假设 \mathcal{M} 是一个 Kripke 模型, ϕ 是一个 $\text{CTL}_P(\mathcal{M})$ 闭公式, 那么 $\text{cpt}(\vdash \phi, \mathbf{t}, \mathbf{f})$ 能在有限步之内重写到 \mathbf{t} 或 \mathbf{f} 。

证明. 令 $F = \{\mathbf{t}, \mathbf{f}, \text{cpt}\} \cup \text{Seq}$, 其中 Seq 是在 $\text{cpt}(\vdash \phi, \mathbf{t}, \mathbf{f})$ 的重写步骤中所出现的相继式的集合; cpt 的元数是 3, F 中其他符号的元数是 0。 F 上的拟序 \succeq ($\forall f, g \in F$, $f \succ g$ 是指 “ $f \succeq g$ 同时 $f \neq g$ ”) 定义如下:

- $\text{cpt} \succ \mathbf{t}$;
- $\text{cpt} \succ \mathbf{f}$;
- 对于每个相继式 $\Gamma \vdash \phi$ 都有 $\Gamma \vdash \phi \succ \text{cpt}$;
- $\Gamma \vdash \phi \succ \Gamma' \vdash \phi'$ 当且仅当 $w(\Gamma \vdash \phi) > w(\Gamma' \vdash \phi')$, 其中 $>$ 是自然数对上的字典序。

令 \succeq_{lpo} 为由 \succeq 生成的关于 CPT 的字典序。显然, \succeq 是良基的, 因此根据命题 2.3.1 可知, \succeq_{lpo} 也是良基的。

若要证明重写系统是可终止的, 只需证明对于每一步重写 $c \rightsquigarrow c'$, 都有 $c \succ_{\text{lpo}} c'$ 。下面我们针对重写规则逐条进行分析:

假设 c 是 $\text{cpt}(\Gamma \vdash \phi, c_1, c_2)$ 形式的。

- 如果 $\phi = \top, \perp, P(s_1, \dots, s_m)$ 或 $\neg P(s_1, \dots, s_m)$, 那么由于 c_1 和 c_2 是 $\text{cpt}(\Gamma \vdash \phi, c_1, c_2)$ 的子项, 因此, $\text{cpt}(\Gamma \vdash \phi, c_1, c_2) \succ_{\text{lpo}} c_1$ 以及 $\text{cpt}(\Gamma \vdash \phi, c_1, c_2) \succ_{\text{lpo}} c_2$ 。
- 如果 $\phi = \phi_1 \wedge \phi_2$, 那么由 \succ_{lpo} 的定义可知, 由于 $\vdash \phi_1 \wedge \phi_2 \succ \vdash \phi_1$, $\text{cpt}(\vdash \phi_1 \wedge \phi_2, c_1, c_2) \succ_{\text{lpo}} \text{cpt}(\vdash \phi_2, c_1, c_2)$ 以及 $\text{cpt}(\vdash \phi_1 \wedge \phi_2, c_1, c_2) \succ_{\text{lpo}} c_2$, 因此 $\text{cpt}(\vdash \phi_1 \wedge \phi_2, c_1, c_2) \succ_{\text{lpo}} \text{cpt}(\vdash \phi_1, \text{cpt}(\vdash \phi_2, c_1, c_2), c_2)$;
- 如果 $\phi = \phi_1 \vee \phi_2$, 那么由 \succ_{lpo} 的定义可知, 由于 $\vdash \phi_1 \vee \phi_2 \succ \vdash \phi_1$, $\text{cpt}(\vdash \phi_1 \vee \phi_2, c_1, c_2) \succ_{\text{lpo}} c_1$ 以及 $\text{cpt}(\vdash \phi_1 \vee \phi_2, c_1, c_2) \succ_{\text{lpo}} \text{cpt}(\vdash \phi_2, c_1, c_2)$, 因此 $\text{cpt}(\vdash \phi_1 \vee \phi_2, c_1, c_2) \succ_{\text{lpo}} \text{cpt}(\vdash \phi_1, c_1, \text{cpt}(\vdash \phi_2, c_1, c_2))$;
- 如果 $\phi = AX_x(\phi_1)(s)$, 那么根据 \succ_{lpo} 的定义可知, 由于 $\Gamma \vdash AX_x(\phi_1)(s) \succ \vdash (s_i/x)\phi_1$ 以及 $\text{cpt}(\Gamma \vdash AX_x(\phi_1)(s), c_1, c_2) \succ_{\text{lpo}} \text{cpt}(\vdash (s_i/x)\phi_1, \text{cpt}(\dots \text{cpt}(\vdash (s_n/x)\phi_1, c_1, c_2) \dots, c_2), c_2)$, 因此 $\text{cpt}(\Gamma \vdash AX_x(\phi_1)(s), c_1, c_2) \succ_{\text{lpo}} \text{cpt}(\vdash (s_1/x)\phi_1, \text{cpt}(\dots \text{cpt}(\vdash (s_n/x)\phi_1, c_1, c_2), \dots, c_2), c_2)$, 其中 $\text{Next}(s) = \{s_1, \dots, s_n\}$, 而且 $i \in \{1, \dots, n\}$;
- 对于 EX 情况的分析与 AX 类似;
- 如果 $\phi = EG_x(\phi_1)(s)$, 那么
 - 当 $EG_x(\phi_1)(s) \in \Gamma$ 时, 此时与第一种情况类似: $c \succ_{\text{lpo}} c'$;
 - 当 $EG_x(\phi_1)(s) \notin \Gamma$ 时, 根据 \succ_{lpo} 的定义可知, 由于 $\Gamma \vdash EG_x(\phi_1)(s) \succ \vdash (s/x)\phi_1$ 以及 $\forall i \in \{1, \dots, n\}, \Gamma \vdash EG_x(\phi_1)(s) \succ \Gamma' \vdash EG_x(\phi_1)(s_i)$, 其中 $\text{Next}(s) = \{s_1, \dots, s_n\}$ 以及 $\Gamma' = \Gamma \cup \{EG_x(\phi_1)(s)\}$;
- 对于 AF 情况的分析与 EG 类似;
- 如果 $\phi = AR_{x,y}(\phi_1, \phi_2)(s)$, 那么
 - 当 $AR_{x,y}(\phi_1, \phi_2)(s) \in \Gamma$ 时, 此时与第一种情况类似: $c \succ_{\text{lpo}} c'$;
 - 当 $AR_{x,y}(\phi_1, \phi_2)(s) \notin \Gamma$ 时, 由 \succ_{lpo} 的定义可知, 由于 $\Gamma \vdash AR_{x,y}(\phi_1, \phi_2)(s) \succ \vdash (s/y)\phi_2$, $\Gamma \vdash AR_{x,y}(\phi_1, \phi_2)(s) \succ \vdash (s/x)\phi_1$ 以及 $\forall i \in \{1, \dots, n\}, \Gamma \vdash AR_{x,y}(\phi_1, \phi_2)(s) \succ \Gamma' \vdash AR_{x,y}(\phi_1, \phi_2)(s_i)$, 因此 $c \succ_{\text{lpo}} c'$, 其中 $\text{Next}(s) = \{s_1, \dots, s_n\}$ 以及 $\Gamma' = \Gamma \cup \{AR_{x,y}(\phi_1, \phi_2)(s)\}$;
- 对于 EU 情况的分析与 AR 类似。

□

2.3.3 证明搜索算法的正确性

证明搜索算法的正确性可由以下命题来表示。

命题 2.3.3 (证明搜索算法的正确性). 对于给定闭公式 ϕ , $\text{cpt}(\vdash \phi, t, f) \rightsquigarrow^* t$ 当且仅当 $\vdash \phi$ 是可证的。

证明. Induction on the structure of ϕ . The details are presented in ??.

□

2.3.4 证明搜索算法的优化

2.3.5 其他 CTL_P 模型检测方法的对比

在这里, 我们讨论 SCTLProV 的证明搜索算法与其他 CTL_P 模型检测方法的对比。

2.3.5.1 基于 BDD 的符号模型检测

当 Kripke 模型中绝大多数状态变量是布尔类型 (比如在硬件模型检测问题中) 的时候, BDD 的应用可以用来减少模型检测算法的空间占用。迄今为止, 最好的基于 BDD 的符号模型检测工具是 NuSMV^[10,11] 以及 NuSMV 的扩展 NuXMV^[12]。下面我们举例说明基于 BDD 的符号模型检测方法的原理: 假设存在一个以 s_0 为初始状态, T 为迁移规则的 Kripke 模型 \mathcal{M} 。若要验证 $\mathcal{M}, s_0 \models EF\phi$, 基于 BDD 的符号模型检测工具 (例如 NuSMV) 会首先计算一个最小不动点 $\text{lfp} = \mu Y.(\phi \vee EXY)$, 然后, 若 $s_0 \in \text{lfp}$, 则 $\mathcal{M}, s_0 \models EF\phi$ 成立, 反之则不成立。计算不动点的过程中会不断地对迁移规则 T 进行展开直到得到不动点, 其中 s_0 不可达的状态也可能被计算在不动点之内。

与基于 BDD 的符号模型检测工具不同, SCTLProV 在验证过程中没有必要计算不动点: 迁移规则 T 是动态展开, 即展开 T 直到可以判定公式是否可证为止。SCTLProV 的验证过程只访问初始状态 s_0 可达的状态, 因此验证过程会节省空间的占用。不止如此, 当 Kripke 模型的状态变量绝大多数为布尔类型的时候, SCTLProV 可以用 BDD 来记录访问过的状态, 并以此来进一步节省空间占用; 反之, 当 Kripke 模型中包含多个非布尔类型的状态变量的时候, SCTLProV 可以选择直接记录 (通常用哈希表) 访问过的状态。不同于符号模型检测中将 Kripke 结构和要证明的性质都编码到 BDD 的做法, SCTLProV 在 Kripke 结构上直接做状态搜索, BDD 只被用来记录搜索过的状态。

2.3.5.2 动态 (On-the-fly) 模型检测

在验证时序逻辑公式的正确性时, 利用动态模型检测的方法可以避免访问 Kripke 模型的整个状态空间, 而只是访问由初始状态可达的状态集合。传统的 CTL 动态模型检测方法^[13,14] 通常是基于递归的: 即子公式的验证以及迁移规则的展开都是递归进行的。基于递归的 CTL 动态模型检测通常会涉及到大量的栈操作, 尤其是在验证大型系统的过程中。验证算法通常会在栈操作上浪费大量的时间。

与传统的 CTL 动态模型检测方法不同, 在 SCTLProV 中, 公式和迁移规则均按需展开, 而且验证算法基于连续 (连续传递风格) 而不是递归。基于连续的算法只需占用常数大小的栈空间^[6,7,15], 而在递归算法中, 栈空间的占用大小与递归深度成正比。

由于目前没有完整的基于传统的动态模型检测算法的工具存在, 因此, 为了将其与 SCTLProV 的证明搜索算法进行对比, 我们开发了一个递归版本的 SCTLProV, 即 SCTLProV_R¹。不同于 SCTLProV, SCTLProV_R 利用基于递归的算法来证明子公式并搜索状态空间。SCTLProV 与 SCTLProV_R 的实验结果对比见第2.4节。

2.3.5.3 限界模型检测

在传统的限界模型检测工具中, 若要证明一个时序逻辑公式, 首先需要人为规定 (或程序给定) 一个限界, 并将 Kripke 模型的迁移规则在限界之内展开, 然后判断该公式在迁移规则的有限步展开之内满足与否。若在当前的限界之内可以判断公式满足与否, 则算法终止; 否则, 继续扩大限界。举例来说, 若要判断 $\mathcal{M}, s_0 \models_{k+1} EF\phi$ 满足与否, 则要先在限界 $k+1$ 之内展开公式与迁移规则^[16]:

$$[\mathcal{M}, EF\phi]_{k+1} := \bigwedge_{i=0}^{k-1} T(s_i, s_{i+1}) \wedge \bigvee_{j=0}^k \phi(s_j)$$

与限界模型检测工具不同, SCTLProV 不需要限界展开迁移规则, 而是在展开证明公式的同时按需展开迁移规则。例如在证明 $\vdash EF_x(\phi)(s_0)$ 的时候, 公式和迁移规则的展开方式如下:

$$\text{unfold}(S, \vdash EF_x(\phi)(s_i)) := \phi(s_i) \vee ((s_i \notin S) \wedge T(s_i, s_{i+1}) \wedge \text{unfold}(S \cup \{s_i\}, \vdash EF_x(\phi)(s_{i+1})))$$

其中, S 是在证明过程中已经访问过的状态集合。

¹https://github.com/terminatorlxj/SCTLProV_R

2.4 案例分析与实验结果

To illustrate the feasibility and the efficiency of SCTLProV, we first use an example (Subsection 2.4.1) to show an application of SCTLProV, and then evaluate several benchmarks (benchmark #1, #2 and #3 in Subsection 2.4.3, benchmark #4 in Subsection 2.4.4, and benchmark #5 in Subsection 2.4.5) to show the efficiency of SCTLProV, and compare the experimental results with five other verification tools: the Resolution-based theorem prover iProver Modulo^[17], the QBF-based bounded model checker Verds version 1.49, the BDD-based unbounded model checker NuSMV version 2.6.0 and its extension NuXMV version 1.0.0, and the formal verification toolbox CADP. All examples and benchmarks are tested on a Linux platform with 3.0 GB memory and a 2.93GHz \times 4 CPU, and the time limit is 20 minutes.

All benchmarks used in this paper are available online².

2.4.1 An illustrative example

例子 2.4.1 (A Mutual Exclusion Problem^[18]). This example is a mutual exclusion algorithm of two concurrent processes (process *A* and process *B*). Mutual Exclusion means that at any time, the number of processes in the critical section is at most one. A scratch of the algorithm is depicted in Figure ??, where *flag* indicates whether there exist a process running at the moment, and *mutex* (has the initial value 0) is the number of processes that have entered the critical section. A violation of Mutual Exclusion means that in some state of the program, more than one process have entered the critical section, i.e., the value of *mutex* is 2.

/* Process A */	/* Process B */
1: while(flag);/*wait*/	1: while(flag);/*wait*/
2: flag = true;	2: flag = true;
3: mutex ++;	3: mutex ++;
/*critical section*/	/*critical section*/
4: mutex --;	4: mutex --;
5: flag = false;	5: flag = false;

图 2.5: A scratch of process A and process B.

²https://github.com/terminatorlxj/ctl_benchmarks

```

Model mutual()
{
  Var {
    flag : Bool; mutex : (0 .. 2); a : (1 .. 5); b : (1 .. 5);
  }
  Init {
    flag := false; mutex := 0; a := 1; b := 1;
  }
  Transition {
    a = 1 && flag = false : {a := 2;};
    a = 2 : {a := 3; flag := true;};
    a = 3 : {a := 4; mutex := mutex + 1;}; /*A has entered the critical section*/
    a = 4 : {a := 5; mutex := mutex - 1;}; /*A has left the critical section*/
    a = 5 : {flag := 0;};
    b = 1 && flag = false : {b := 2;};
    b = 2 : {b := 3; flag := true;};
    b = 3 : {b := 4; mutex := mutex + 1;}; /*B has entered the critical section*/
    b = 4 : {b := 5; mutex := mutex - 1;}; /*B has left the critical section*/
    b = 5 : {flag := 0;};
    /*If none of the conditions above are satisfied, then the current state goes to itself.*/
    (a = 1 || b = 1) && flag = true: {}
  }
  Atomic {
    bug(s) := s(mutex = 2);
  }
  Spec {
    find_bug := EU(x, y, TRUE, bug(y), ini);
  }
}

```

图 2.6: The input file “mutual.model”.

In the input file (Figure 2.6), two additional variables a and b are used to indicate the program counters of the two processes, respectively, and ini stands for the initial state. The property to be checked is that whether both processes are in the critical section at the the same time. We check this property in SCTLProV

using the following command.

```
sctl -output output.out mutual.model
```

The result is as follows, which indicates that there is a bug in the mutual exclusion problem, i.e., the mutual exclusion property is violated.

```
verifying on the model mutual...
find_bug: EU(x,y, TRUE, bug(y), ini)
find_bug is true.
```

The proof tree of the property is output to the file “output.out”, which is listed as follows. In the output of the proof tree, each node is shown as the form

$$id : seqt [id_1, \dots, id_n]$$

where *seqt* is the sequent, *id* the identity number of the sequent, and id_1, \dots, id_n the identity numbers of the hypothesis of the sequent.

```
0: |- EU(x,y,TRUE,bug(y),{flag:=false;mutex:=0;a:=1;b:=1}) [4, 1]
4: {flag:=false;mutex:=0;a:=1;b:=1}
|- EU(x,y,TRUE,bug(y),{flag:=false;mutex:=0;a:=2;b:=1}) [7, 5]
1: |- TRUE []
7: {flag:=false;mutex:=0;a:=1;b:=1}
{flag:=false;mutex:=0;a:=2;b:=1}
|- EU(x,y,TRUE,bug(y),{flag:=false;mutex:=0;a:=2;b:=2}) [23, 20]
5: |- TRUE []
23:{flag:=false;mutex:=0;a:=1;b:=1}
{flag:=false;mutex:=0;a:=2;b:=1}
{flag:=false;mutex:=0;a:=2;b:=2}
|- EU(x,y,TRUE,bug(y),{flag:=true;mutex:=0;a:=3;b:=2}) [27, 24]
20: |- TRUE []
27:{flag:=false;mutex:=0;a:=1;b:=1}
{flag:=false;mutex:=0;a:=2;b:=1}
{flag:=false;mutex:=0;a:=2;b:=2}
{flag:=true;mutex:=0;a:=3;b:=2}
|- EU(x,y,TRUE,bug(y),{flag:=true;mutex:=1;a:=4;b:=2}) [31, 28]
24: |- TRUE []
31:{flag:=false;mutex:=0;a:=1;b:=1}
```

```

{flag:=false;mutex:=0;a:=2;b:=1}
{flag:=false;mutex:=0;a:=2;b:=2}
{flag:=true;mutex:=0;a:=3;b:=2}
{flag:=true;mutex:=1;a:=4;b:=2}
|- EU(x,y,TRUE,bug(y),{flag:=true;mutex:=1;a:=4;b:=3}) [35, 32]
28: |- TRUE []
35:{flag:=false;mutex:=0;a:=1;b:=1}
{flag:=false;mutex:=0;a:=2;b:=1}
{flag:=false;mutex:=0;a:=2;b:=2}
{flag:=true;mutex:=0;a:=3;b:=2}
{flag:=true;mutex:=1;a:=4;b:=2}
{flag:=true;mutex:=1;a:=4;b:=3}
|- EU(x,y,TRUE,bug(y),{flag:=true;mutex:=2;a:=4;b:=4}) [37]
32: |- TRUE []
37: |- bug({flag:=true;mutex:=2;a:=4;b:=4}) []

```

According to the output above, we can find that after process A has entered the critical section, process B can also enter the critical section.

The output can be visualized by a visualization tool VMDV³ (Visualization for Modeling, Demonstration, and Verification), where both the proof tree and the Kripke model can be visualized in 3D space, as depicted in Figure 2.7.

Note that the violation of the mutual exclusion property can be avoided if the the algorithm of the two processes is modified^[18] as in Figure 2.8. The Kripke model built for this algorithm is depicted in Figure 2.9. Variable x and y are used to indicate whether process A and B are running, respectively; $turn$ is the variable indicating it is whose turn to enter the critical section.

When applying this solution, the mutual exclusion exclusion property will not be violated, which is indicated by the following result.

```

verifying on the model mutual...
find_bug: EU(x, y, TRUE, bug(y), ini)
find_bug is false.

```

³<https://github.com/terminatorlxj/VMDV>

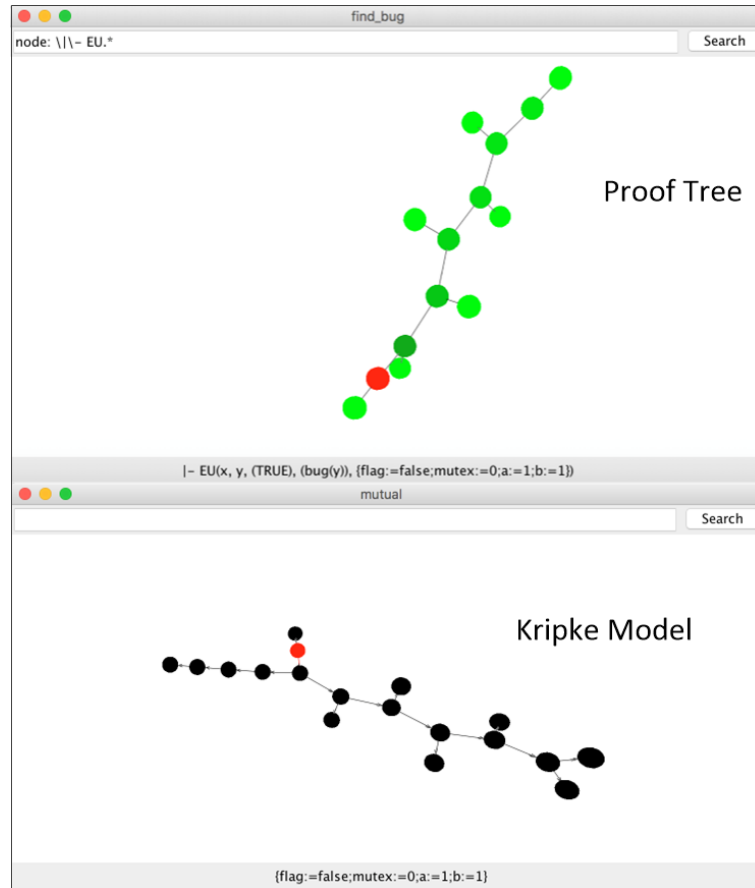


图 2.7: Visualization of the proof tree and the Kripke model in the illustrative example.

<code>/* Process A */</code>	<code>/* Process B */</code>
<code>1: x = true;</code>	<code>1: y = true;</code>
<code>2: turn = 1;</code>	<code>2: turn = 2;</code>
<code>3: while(y&&turn!=2); /*wait*/</code>	<code>3: while(x&&turn!=1); /*wait*/</code>
<code>4: mutex ++;</code>	<code>4: mutex ++;</code>
<code>/*critical section*/</code>	<code>/*critical section*/</code>
<code>5: mutex --;</code>	<code>5: mutex --;</code>
<code>6: x = false;</code>	<code>6: y = false;</code>

图 2.8: A simple solution of the mutual exclusion problem.

```

Model mutual()
{
  Var {
    x:Bool; y:Bool; mutex:(0 .. 2); turn:(1 .. 2); a:(1 .. 6); b:(1 .. 6);
  }
  Init {
    x := false; y := false; mutex := 0; turn := 1; a := 1; b := 1;
  }
  Transition {
    a = 1 : {a := 2; x := true;};
    a = 2 : {a := 3; turn := 1;};
    a = 3 && (y = false || turn = 2): {a := 4;};
    a = 4 : {a := 5; mutex := mutex + 1;}; /*A has entered the critical section*/
    a = 5 : {a := 6; mutex := mutex - 1;}; /*A has left the critical section*/
    a = 6 : {x := false;};
    b = 1 : {b := 2; y := true;};
    b = 2 : {b := 3; turn := 2;};
    b = 3 && (x = false || turn = 1): {b := 4;};
    b = 4 : {b := 5; mutex := mutex + 1;}; /*B has entered the critical section*/
    b = 5 : {b := 6; mutex := mutex - 1;}; /*B has left the critical section*/
    b = 6 : {y := false;};
    /*If none of the conditions above are satisfied, then the current state goes to itself.*/
    (a != 3 && (y = true && turn = 1)) || (b != 3 && (x = true && turn = 2)) : {};
  }
  Atomic {
    bug(s) := s(mutex = 2);
  }
  Spec {
    find_bug := EU(x, y, TRUE, bug(y), ini);
  }
}

```

图 2.9: The input file “mutual_solution.model” indicating the solution.

2.4.2 An application to the analysis of Air traffic control protocols

As an application to an engineering problem, we present a concept of operations for the Small Aircraft Transportation System (SATS)^[1,19] in SCTLProV⁴.

⁴<https://github.com/terminatorlxj/SATS-model>

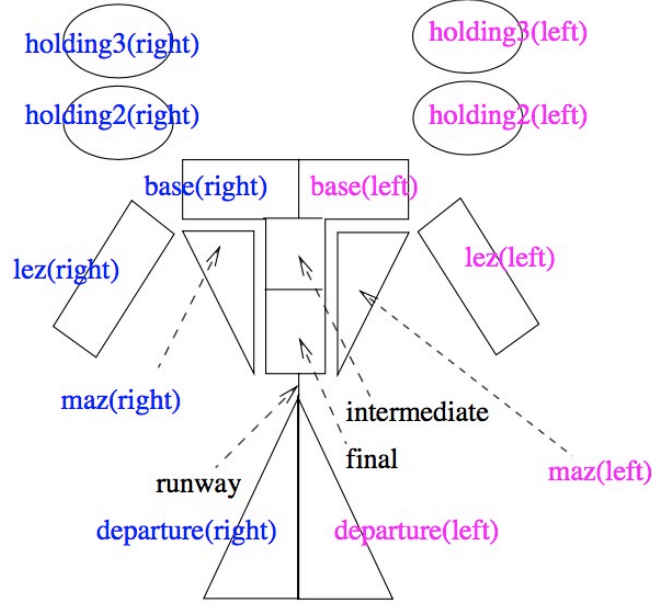


图 2.10: SCA zones, where right and left are relative to the pilot facing the runway, i.e., opposite from the reader point of view^[1].

In this concept of operation, the airspace volume surrounding an airport facility, called the self controlled area, is divided into 15 zones (Figure 2.10). For instance, the zone holding3(right) is a holding pattern at 3000 feet on the right of the self controlled area. Each zone contains a list of aircraft and 24 transition rules specify different SATS-procedures. For instance, the rule Vertical Entry (right) specifies the vertical entry of an aircraft in the zone holding3(right).

The model is non-deterministic, that is, for a given state, several transitions are possible and all must be considered. As there are no a priori bounds on the number of aircraft in each zone, the number of states in the model is potentially infinite. However, the number of states that are reachable from the initial state is finite: an enumeration of the model shows that there are 54221 such states (and around 3000 in the simplified model where departure operations are not considered).

There are eight properties of the model that we want to verify with SCTLProV, for instance that the SATS concept does not allow more than four simultaneous landing operations and none of the 15 zones contains too many aircraft (each zone is assigned a maximum number of aircraft and the actual number of aircraft is never higher than this number). The safety property is thus conjunction of these eight properties.

The verification problem is to check that this property holds on every reachable

state from the initial state (the state where there are no aircraft on each zone of the self controlled area), so the formula to be checked is $AG_x(\phi)(e)$ where ϕ is the conjunction of the eight properties and e is the initial state.

This is a typical model checking problem, but this problem is known to be cumbersome for traditional model checkers^[1] because:

- Each state of the model is represented by a complex data structure. For instance, a number of state variables are represented by lists of aircraft with unbounded length.
- The transition rules of the model are complex algorithms. For instance, some transitions rules involve recursive operations on lists of aircraft.
- The properties to be verified in the model are also represented by complex algorithms. For instance, some of the properties are inductively defined over lists of aircraft.

However, this example fits well in SCTLProV that provides a more expressive input language than most traditional model checkers. Indeed, SCTLProV provides both readable notations for the definition of data structures such as records or lists with unbounded length, and arbitrary algorithms for the definitions of transition rules and of properties. So we have been able to check in SCTLProV that the safety property holds on the model, and the verification was executed in less than 30 seconds on the same machine as which the benchmarks are evaluated.

2.4.3 Randomly generated programs

We consider three benchmarks in this part. The original description of benchmark #1^[20] is restated here. Based on benchmark #1, we extend the number of variables to tens, hundreds, and even thousands in benchmark #2 and benchmark #3. The randomness of the test cases in three benchmarks makes it rather fair for different CTL model checking approaches, and helps us recognize the strengths and weaknesses of each tool.

2.4.3.1 Benchmark #1

Benchmark #1 chosen in this subsection is originally introduced by Zhang^[20] in the evaluation of model checkers Verds and NuSMV. Later, Ji^[21] also uses this benchmark in the evaluation of the theorem prover iProver Modulo and the model

checker Verds. This benchmark consists of 2880 randomly generated test cases where two types of random Boolean programs are considered—Concurrent Processes and Concurrent Sequential Processes. In programs with Concurrent Processes, the parameters of the first set of random Boolean programs are as follows.

a : number of processes
b : number of all variables
c : number of shared variables
d : number of local variables in a process

The shared variables are initially set to a random value in $\{0, 1\}$, and the local variables are initially set to 0. For each process, the shared variables and the local variables are assigned the negation of a variable randomly chosen from these variables. We test different sizes of the programs with 3 processes ($a = 3$), and let b vary over the set of values $\{12, 24, 36\}$, then set $c = b/2, d = c/a$. Each of the 24 properties is tested on 20 test cases for each value of b .

In programs with Concurrent Sequential Processes, in addition to a, b, c, d specified above, the parameters of the second set of random Boolean programs are as follows.

t : number of transitions in a process
p : number of parallel assignments in each transition

For each concurrent sequential process, besides the b Boolean variables, there is a local variable representing program locations, with c possible values. The shared variables are initially set to a random value in $\{0, 1\}$, and the local variables are initially set to 0. For each transition of a process, p pairs of shared variables and local variables are randomly chosen among the shared variables and the local variables, such that the first element of such a pair is assigned the negation of the second element of the pair. Transitions are numbered from 0 to $t - 1$, and are executed consecutively, and when the end of the sequence of the transitions is reached, it loops back to the execution of the transition numbered 0. For this type of programs, we test different sizes of the programs with 2 processes ($a = 2$), and let b vary in the set of values $\{12, 16, 20\}$, and then set $c = b/2, d = c/a, t = c$, and $p = 4$. Similarly, each property is tested on 20 test cases for each value of b .

Twenty-four properties are to be checked in this benchmark: properties P_{01} to P_{12} are depicted in Figure 2.11, and P_{13} to P_{24} are simply the variations of P_{01} to P_{12} by replacing \wedge and \vee by \vee and \wedge , respectively.

P_{01}	$AG(\bigvee_{i=1}^c v_i)$
P_{02}	$AF(\bigvee_{i=1}^c v_i)$
P_{03}	$AG(v_1 \Rightarrow AF(v_2 \wedge \bigvee_{i=3}^c v_i))$
P_{04}	$AG(v_1 \Rightarrow EF(v_2 \wedge \bigvee_{i=3}^c v_i))$
P_{05}	$EG(v_1 \Rightarrow AF(v_2 \wedge \bigvee_{i=3}^c v_i))$
P_{06}	$EG(v_1 \Rightarrow EF(v_2 \wedge \bigvee_{i=3}^c v_i))$
P_{07}	$AU(v_1, AU(v_2, \bigvee_{i=3}^c v_i))$
P_{08}	$AU(v_1, EU(v_2, \bigvee_{i=3}^c v_i))$
P_{09}	$AU(v_1, AR(v_2, \bigvee_{i=3}^c v_i))$
P_{10}	$AU(v_1, ER(v_2, \bigvee_{i=3}^c v_i))$
P_{11}	$AR(AXv_1, AXAU(v_2, \bigvee_{i=3}^c v_i))$
P_{12}	$AR(EXv_1, EXEU(v_2, \bigvee_{i=3}^c v_i))$

图 2.11: Properties $P_{01}, P_{02}, \dots, P_{12}$ to be checked in benchmark #1, #2, and #3.

2.4.3.2 Benchmark #2 and #3

In benchmark #2, we increase the number of state variables in benchmark #1 to 48, 60, or 72 for Concurrent Processes, and 24, 28, or 32 for Concurrent Sequential Processes. The 2880 test cases are also randomly generated. The properties to be checked are the same as in benchmark #1.

In benchmark #3, we increase the number of state variables in benchmark #1 to 252, 504 and 1008 for both Concurrent Processes and Concurrent Sequential Processes, and check the same properties as benchmark #1 and #2.

2.4.3.3 Experimental data

The experimental results are shown below, and the detailed data is in ??.

Experimental data for benchmark #1. For 2880 test cases in this benchmark, iProver Modulo can solve 1816 (63.1%) cases, Verds can solve 2230 (77.4%) cases, SCTLProV can solve 2862 (99.4%) cases, and both NuSMV and NuXMV can solve all (100%) test cases. The numbers of test cases where SCTLProV runs faster are 2823 (98.2%) comparing with iProver Modulo, 2858 (99.2%) comparing with Verds, 2741 (95.2%) comparing with NuSMV, and 2763 (95.9%) comparing with NuXMV. According to Figure 2.12 and Figure 2.13, SCTLProV uses less time and space than the other four tools.

Experimental data for benchmark #2. For 2880 test cases in this benchmark, iProver Modulo can solve 1602 (55.6%) cases, Verds can solve 1874 (65.1%) cases, NuSMV can solve 728 (25.3%) cases, NuXMV can solve 736 (25.6%) cases, and SCTLProV can solve 2597 (90.2%) cases. The numbers of test cases where SCTLProV runs faster are 2597 (90.2%) comparing with iProver Modulo, 2594 (90.1%) comparing with Verds, and 2588 (89.9%) comparing both with NuSMV and NuXMV. According to Figure 2.14 and Figure 2.15, SCTLProV uses less time and space than the other four tools.

Experimental data for benchmark #3. For 2880 test cases in this benchmark, iProver Modulo can solve 1146 (39.8%) cases, Verds can solve 352 (12.2%) cases, SCTLProV can solve 1844 (64.0%) cases, while neither NuSMV nor NuXMV can solve any case.

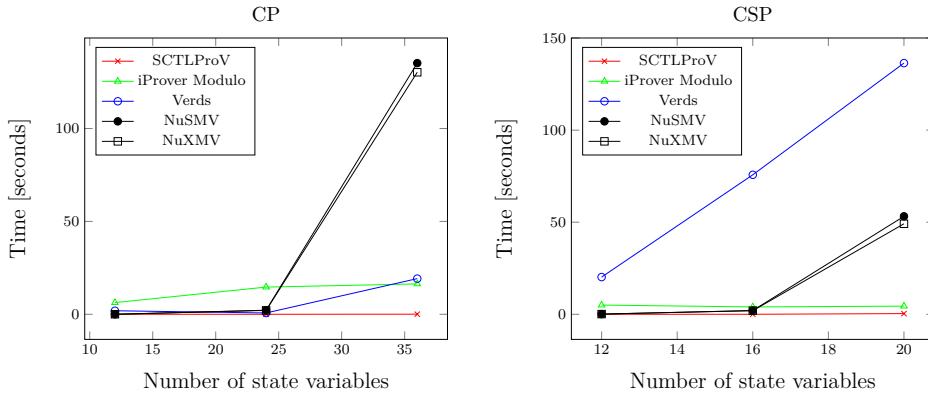


图 2.12: Average verification time in benchmark #1.

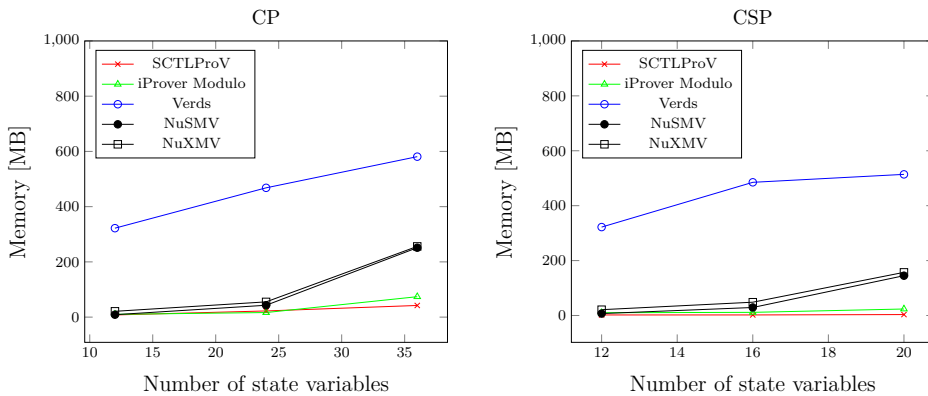


图 2.13: Average memory usage in benchmark #1.

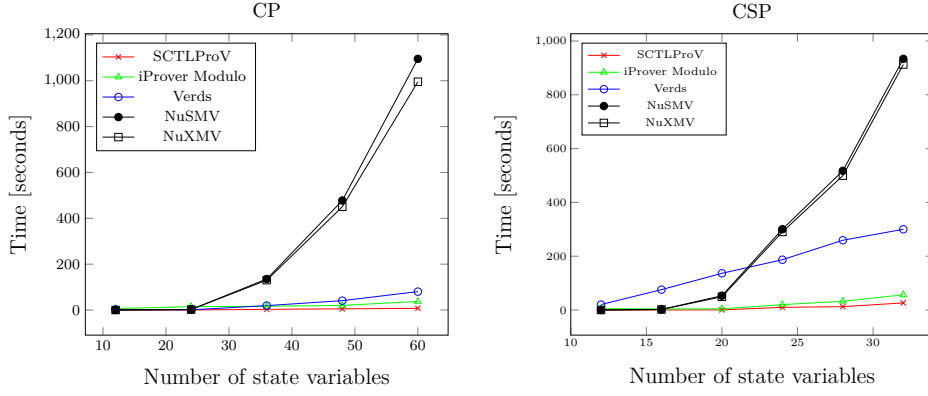


图 2.14: Average verification time in benchmark #1 and #2.

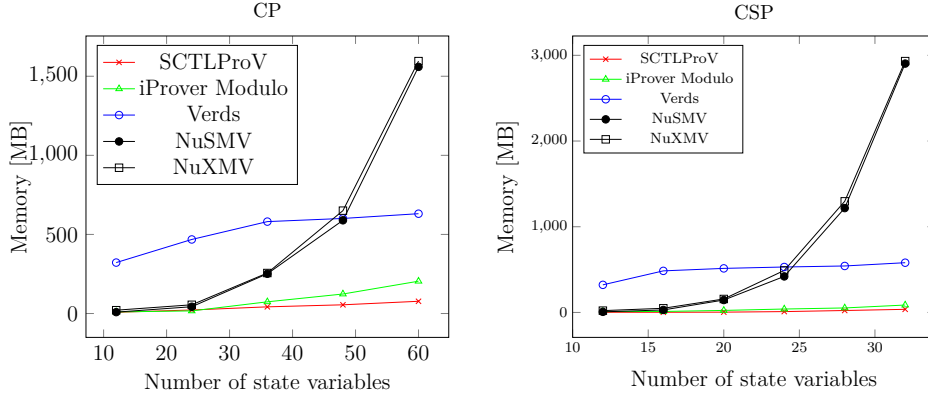


图 2.15: Average memory usage in benchmark #1 and #2.

2.4.3.4 Continuation vs. recursion.

To show the importance of using continuation-passing style, we have implemented a recursive version of our tool and compared the time efficiency. In benchmark #1, #2, and #3, SCTLProV solves about 10% more test cases than SCTLProV_R, and it outperforms SCTLProV_R in almost all solvable cases (Table 2.1). SCTLProV_R is more sensitive to the number of variables than SCTLProV (Figure 2.16).

Bench	SCTLProV solvable	SCTLProV _R solvable	$t(\text{SCTLProV}) < t(\text{SCTLProV}_R)$
#1	2862(99.4%)	2682(93.1%)	2598(90.2%)
#2	2597(90.2%)	2306(80.1%)	2406(83.5%)
#3	1849(64.2%)	1520(52.8%)	1735(60.2%)

表 2.1: SCTLProV vs. SCTLProV_R

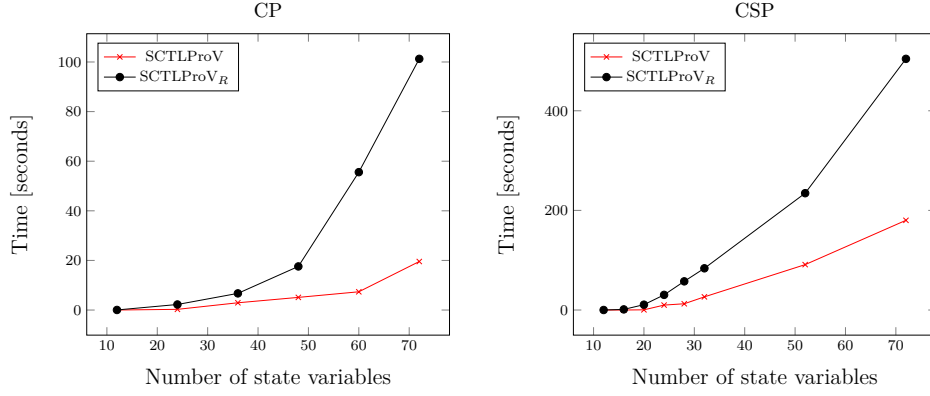


图 2.16: Average verification time in SCTLProV vs. SCTLProV_R.

In the comparison of average verification time of SCTLProV and SCTLProV_R, we extend the number of variables in Concurrent Sequential Processes to 52 and 72, on the basis of benchmark #2.

2.4.4 Programs with fairness constraints

In this part, we evaluate benchmark #4, which models mutual exclusion algorithms and ring algorithms⁵. Then, we compare the evaluation results of SCTLProV, Verds, NuSMV, and NuXMV, and we do not consider iProver Modulo because iProver Modulo cannot handle CTL properties with fairness constraints^[21].

Mutual exclusion and ring algorithms. This benchmark consists of two sets of concurrent programs: the mutual exclusion algorithms and the ring algorithms. Both kinds of algorithms consist of a set of concurrent processes running in parallel.

In the mutual exclusion algorithms, the scheduling of processes is simple: for all i between 0 and $n - 2$, process $i + 1$ performs a transition after process i , and process 0 performs a transition after process $n - 1$. Each formula in the algorithms needs to be verified under the fairness constraint that each process does not starve, i.e., no process waits infinitely long.

Each process in the mutual exclusion algorithms has three internal states: non-critical, trying, and critical. The number of processes vary from 6 to 51. There are five properties specified by CTL formulae are to be verified in mutual exclusion algorithms, as in Table 2.17. In these formulae, non_i (try_i , cri_i) indicates that process p_i has internal state noncritical (trying, critical). Note that because of the scheduling

⁵http://lcs.ios.ac.cn/~zwh/verds/verds_code/bp12.rar

Prop	Mutual Exclusion Algorithms
P_1	$EF(cri_0 \wedge cri_1)$
P_2	$AG(try_0 \Rightarrow AF(cri_0))$
P_3	$AG(try_1 \Rightarrow AF(cri_1))$
P_4	$AG(cri_0 \Rightarrow Acri_0U(\neg cri_0 \wedge A\neg cri_0Ucri_1))$
P_5	$AG(cri_1 \Rightarrow Acri_1U(\neg cri_1 \wedge A\neg cri_1Ucri_0))$
Prop	Ring Algorithms
P_1	$AGAFout_0 \wedge AGAF\neg out_0$
P_2	$AGEFout_0 \wedge AGEF\neg out_0$
P_3	$EGAFout_0 \wedge EGAF\neg out_0$
P_4	$EGEFout_0 \wedge EGEF\neg out_0$

图 2.17: Properties to be verified in benchmark #4.

algorithm, processes 0 and 1 are not symmetric, as exemplified by the difference in performance between the properties P_4 and P_5 .

Each process in the ring algorithms consists of 5 Boolean internal variables indicating the internal state, and a Boolean variable indicating the output. Each process receives a Boolean value as the input during its running time. For a ring algorithm with processes p_0, p_1, \dots, p_n , the internal state of p_i depends on the output of process p_{i-1} , and the output of p_{i-1} depends on its internal state, where $1 \leq i \leq n$. The internal state of p_0 depends on the output of process p_n , and the output of p_n depends on the internal state of its own. The number of processes vary from 3 to 10. There are four properties specified by CTL formulae are to be verified in ring algorithms, as in Figure 2.17. In these formulae, out_i indicates that the output of process p_i is Boolean value *true*.

The experimental results (Table 2.2 and Table 2.3) show that SCTLProV solves more test cases than Verds, NuSMV, and NuXMV. At the same time, SCTLProV is more time and space efficiency in more than 75 percent of the test cases than the other three tools.

The detailed experimental data is shown in ??.

Programs	Verds	NuSMV	NuXMV	SCTLProV
mutual exclusion	136 (59.1%)	50 (21.7%)	50 (21.7%)	191 (83.0%)
ring	16 (50.0%)	21 (65.6%)	21 (65.6%)	20 (62.5%)
Sum	152(58.0%)	71(27.1%)	71(27.1%)	211 (80.5%)

表 2.2: Solvable cases in Verds, NuSMV, NuXMV, and SCTLProV.

Programs	Verds	NuSMV	NuXMV
mutual exclusion	187 (81.3%)	191 (83.0%)	191 (83.0%)
ring	13 (40.6%)	20 (62.5%)	20 (62.5%)
Sum	200(76.3%)	211(80.5%)	211(80.5%)

表 2.3: Cases where SCTLProV both runs faster and uses less memory.

2.4.5 The VLTS benchmark

In this part, we evaluate benchmark #5, also called the VLTS (Very Large Transition Systems) benchmark⁶, which was originally proposed as a part of the CADP⁷ (Construction and Analysis of Distributed Processes) toolbox^[22]. As a formal verification toolbox, CADP focuses on action-based models, for instance, LTSs and Markov Chains. As pointed out by the authors of CADP: this benchmark has been obtained from the modeling of various communication protocols and concurrent systems, many of which corresponds to real life and industrial systems.

There are 40 test cases in benchmark #5, where deadlocks and livelocks are to be detected for each test case. All test cases in benchmark #5 are encoded as BCG (Binary-Coded Graphs) format, which is a binary format designed for encoding large state spaces^[22]. When verifying test cases in this benchmark in SCTLProV, each BCG file was parsed into a Kripke model, and perform the proving procedure on the Kripke model. We compare the experimental results of detecting deadlocks and livelocks both in SCTLProV and CADP.

The experimental result is depicted in Table 2.4: when detecting deadlocks, SCTLProV uses less time than CADP in 31 (77.5%) test cases, and uses less memory than CADP in 35 (87.5%) test cases; when detecting livelocks, SCTLProV uses less time than CADP in 22 (55%) test cases, and uses less memory than CADP in 27 (67.5%) test cases.

⁶<http://cadp.inria.fr/resources/vlts/>

⁷<http://cadp.inria.fr/>

For brevity's sake, more details about the experiment on benchmark #5 are explained in ??.

Cases	$t(\text{SCTLProV}) < t(\text{CADP})$	$m(\text{SCTLProV}) < m(\text{CADP})$
deadlock	31 (77.5%)	35 (87.5%)
livelock	22 (55%)	27 (67.5%)

表 2.4: Test cases where SCTLProV uses less time and less memory, respectively.

2.4.6 Discussion of the experimental results

In the evaluation of benchmark #1 to benchmark #4 in this paper, the performances of NuSMV, NuXMV, Verds, iProver Modulo, and SCTLProV in the comparisons are affected by two factors: the number of state variables, and the type of the property to be checked. The performances of NuSMV and NuXMV are mainly affected by the number of state variables, while the performances of iProver Modulo, Verds, and SCTLProV are mainly affected by the type of the property to be checked. When the number of state variables is rather small (such as test cases in benchmark #1), NuSMV and NuXMV solves more test cases than iProver Modulo, Verds and SCTLProV, but when the number of state variables becomes larger (such as test cases in benchmark #2 and #3), they performs worse than the other three tools. When checking properties where nearly all states must be searched (such as *AG* properties), NuSMV and NuXMV usually perform better than iProver Modulo, Verds and SCTLProV. However, for most properties, iProver Modulo, Verds and SCTLProV usually search much less states than NuSMV and NuXMV to check them, and are more time and space efficiency. Thus, iProver Modulo, Verds and SCTLProV scale up better than NuSMV and NuXMV when checking these properties. Moreover, SCTLProV scales up better than both iProver Modulo and Verds, and outperforms these two tools in most solvable cases.

In the evaluation of benchmark #5, the performances of SCTLProV and CADP are affected by both the size of the state space, and the properties to be verified. The verification processes in SCTLProV and CADP tend to use less time and memory in test cases which have smaller state spaces. For test cases where there exists a deadlock or livelock, the verification processes tend to use less time and memory than those where there are no deadlock or livelock. Moreover, SCTLProV outperforms CADP in more than half of the test cases.

参考文献

- [1] MUÑOZ C A, DOWEK G, CARREÑO V. Modeling and verification of an air traffic concept of operations[C]//Proceedings of the ACM/SIGSOFT International Symposium on Software Testing and Analysis, ISSTA 2004. Boston, Massachusetts, USA: ACM, USA, 2004: 175–182.
- [2] EMERSON E A, CLARKE E M. Using branching time temporal logic to synthesize synchronization skeletons[J]. Sci. Comput. Program., 1982, 2(3): 241–266.
- [3] EMERSON E A, HALPERN J Y. Decision procedures and expressiveness in the temporal logic of branching time[J]. J. Comput. Syst. Sci., 1985, 30(1): 1–24.
- [4] CRAIG J J. Introduction to robotics - mechanics and control (2. ed.)[M]. USA: Prentice Hall, 1989.
- [5] PARTOVI A, LIN H. Assume-guarantee cooperative satisfaction of multi-agent systems[C]//Proceedings of American Control Conference, ACC 2014. USA: IEEE, USA, 2014: 2053–2058.
- [6] APPEL A W. Compiling with continuations (corr. version)[M]. UK: Cambridge University Press, 2006.
- [7] SESTOFT P. Undergraduate topics in computer science: volume 50 programming language concepts[M]. Switzerland: Springer International Publishing, 2012.
- [8] DERSHOWITZ N. Termination of rewriting[J/OL]. J. Symb. Comput., 1987, 3(1/2): 69–116. [https://doi.org/10.1016/S0747-7171\(87\)80022-6](https://doi.org/10.1016/S0747-7171(87)80022-6).
- [9] SARNAT J, SCHÜRMANN C. Lexicographic path induction[C]//Proceedings of Typed Lambda Calculi and Applications, 9th International Conference, TLCA 2009. Brasilia, Brazil: Springer, Berlin, 2009: 279–293.
- [10] MCMILLAN K L. Symbolic model checking[M]. USA: Springer, 1993.
- [11] CIMATTI A, CLARKE E M, GIUNCHIGLIA F, et al. Nusmv: A new symbolic model verifier[C]//Proceedings of CAV'99. Trento, Italy: Springer-Verlag, Berlin, 1999: 495–499.
- [12] CAVADA R, CIMATTI A, DORIGATTI M, et al. The nuxmv symbolic model checker [C]//Proceedings of Computer Aided Verification - 26th International Conference, CAV 2014. Vienna, Austria: Springer International Publishing, Switzerland, 2014: 334–342.
- [13] VERGAUWEN B, LEWI J. A linear local model checking algorithm for CTL[C]//Proceedings of CONCUR '93, 4th International Conference on Concurrency Theory. Hildesheim, Germany: Springer-Verlag, Berlin, 1993: 447–461.
- [14] BHAT G, CLEAVELAND R, GRUMBERG O. Efficient on-the-fly model checking

- for $ctl^*[C]$ //Proceedings of LICS'95. San Diego, California, USA: IEEE Computer Society, USA, 1995: 388–397.
- [15] REYNOLDS J C. The discoveries of continuations[J]. Lisp and Symbolic Computation, 1993, 6(3-4): 233–248.
- [16] BIERE A, CIMATTI A, CLARKE E, et al. Symbolic model checking without BDDs [C]//CLEAVELAND W R. LNCS: volume 1579 Proceedings of TACAS'99. Amsterdam, the Netherlands: Springer, USA, 1999: 193–207.
- [17] BUREL G. Experimenting with deduction modulo[C]//SOFRONIE-STOKKERMANS V, BJØRNER N. Proceedings of CADE 2011. Wroclaw, Poland: Springer-Verlag, Berlin, 2011: 162–176.
- [18] PETERSON G L. Myths about the mutual exclusion problem[J]. Inf. Process. Lett., 1981, 12(3): 115–116.
- [19] NASA/TM-2004-213006. Abstract model of sats concept of operations: Initial results and recommendations[M]. USA: NASA, 2004.
- [20] ZHANG W. QBF Encoding of Temporal Properties and QBF-based Verification[C]//Proceedings of IJCAR 2014. Vienna: Springer-Verlag, Berlin, 2014: 224–239.
- [21] JI K. CTL Model Checking in Deduction Modulo[C]//Proceedings of Automated Deduction - CADE-25. Berlin: Springer International Publishing, Switzerland, 2015: 295–310.
- [22] GARAVEL H, LANG F, MATEESCU R, et al. CADP 2011: a toolbox for the construction and analysis of distributed processes[J/OL]. STTT, 2013, 15(2): 89–107. <https://doi.org/10.1007/s10009-012-0244-z>.

发表学术论文

学术论文

- [1] Jian Liu, Ying Jiang, Yanyun Chen. VMDV: A 3D Visualization Tool for Modeling, Demonstration, and Verification. TASE 2017. accepted.
- [2] Ying Jiang, Jian Liu, Gilles Dowek, Kailiang Ji. SCTL: Towards Combining Model Checking and Proof Checking. The Computer Journal. submitted.

项目资助情况

中法合作项目 LOCALI (项目编号 NSFC 61161130530 和 ANR 11 IS02 002 01)。

简历

基本情况

刘坚，男，山东省茌平县人，1989 年出生，中国科学院软件研究所博士研究生。

教育背景

- 2011 年 9 月至今：中国科学院软件研究所，计算机软件与理论，硕博连读
- 2007 年 9 月至 2011 年 7 月：山东农业大学，信息与计算科学，本科

联系方式

通讯地址：北京市海淀区中关村南四街 4 号，中国科学院软件研究所，5 号楼
3 层计算机科学国家重点实验室

邮编：100090

E-mail: dreammaker2010@yeah.net

致 谢

值此论文完成之际，谨在此向多年来给予我关心和帮助的老师、学长、同学、朋友和家人表示衷心的感谢！