



基于神经符号融合推理的可验证的线性时态逻辑可满足性近似检测

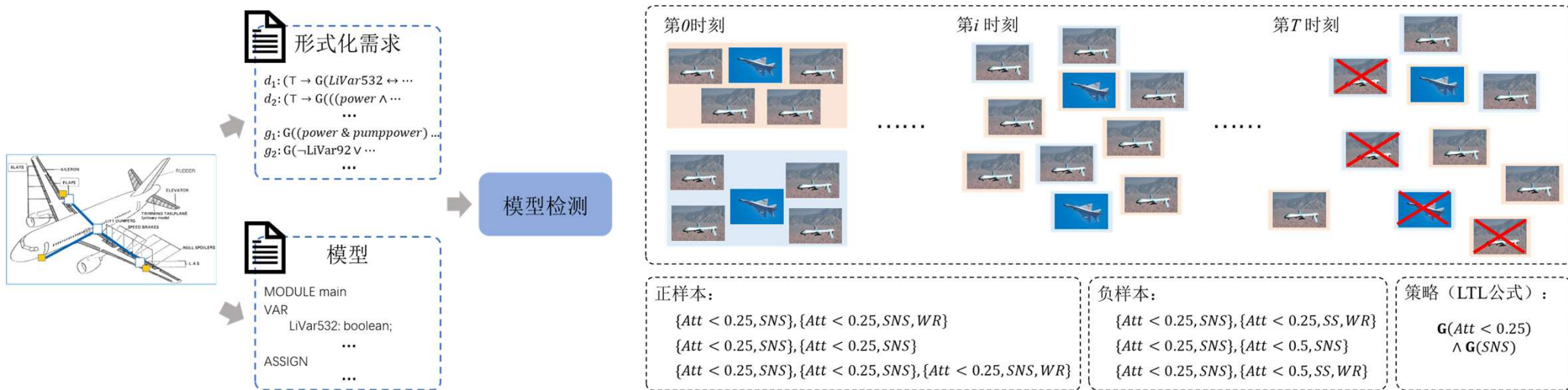
罗炜麟

中山大学

2024-7 贵阳、毕节

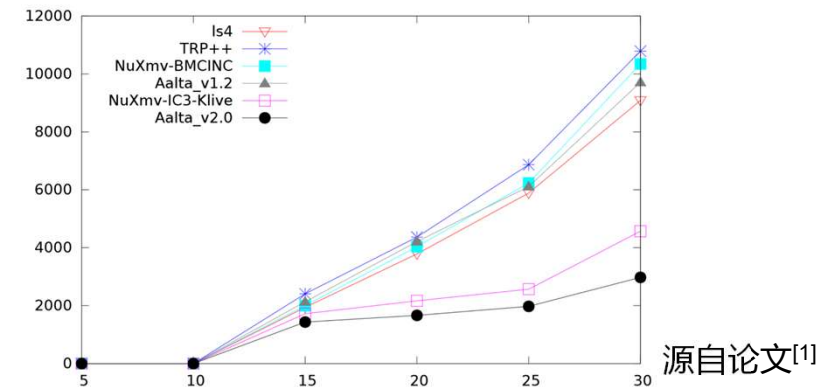
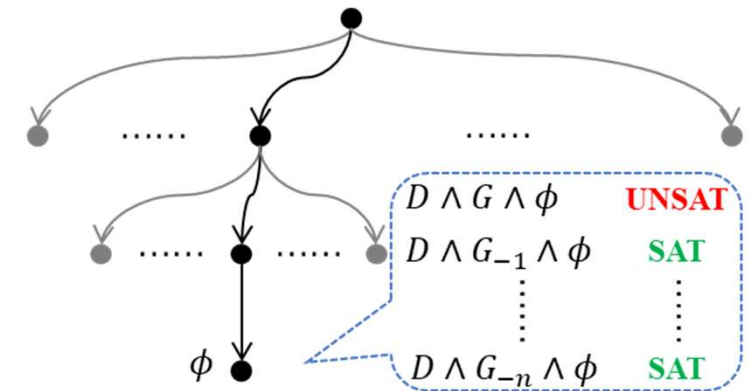
研究背景和动机

- 线性时态逻辑 (Linear Temporal Logic, LTL)
 - 时态算子: next (X)、until (U)、always (G)、eventually (F) 等
 - 例如, $G(req \rightarrow F grant)$
 - 应用: 形式化验证、规则学习



研究背景和动机

- LTL可满足性检测 (LTL Satisfiability Checking)
 - 检测给定的LTL公式是否是可满足的 (Satisfiable)
 - PSPACE-complete
- 可满足性检测繁重的下游任务
 - 例如, 需求修复 (Specification Repair) 和边界条件识别 (Boundary Condition Identification)



[1] Jianwen Li, Shufang Zhu, Geguang Pu, Lijun Zhang, Moshe Y. Vardi: SAT-based explicit LTL reasoning and its application to satisfiability checking. Formal Methods Syst. Des. 54(2): 164-190 (2019)



研究背景和动机

- 相关工作

- 基于逻辑的方法：
 - 基于制表式 (Tableau) 的方法^[1]
 - 基于布尔可满足性问题 (Boolean Satisfiability Problem, SAT) 的方法^[2,3]
 -
- 正确 (Sound) 且完备 (Complete)
- 依赖于精心设计的启发式搜索规则
- 设计良好的启发式搜索规则**困难且耗时!**

[1] Matteo Bertello, Nicola Gigante, Angelo Montanari, Mark Reynolds: Leviathan: A New LTL Satisfiability Checking Tool Based on a One-Pass Tree-Shaped Tableau. IJCAI 2016: 950-956

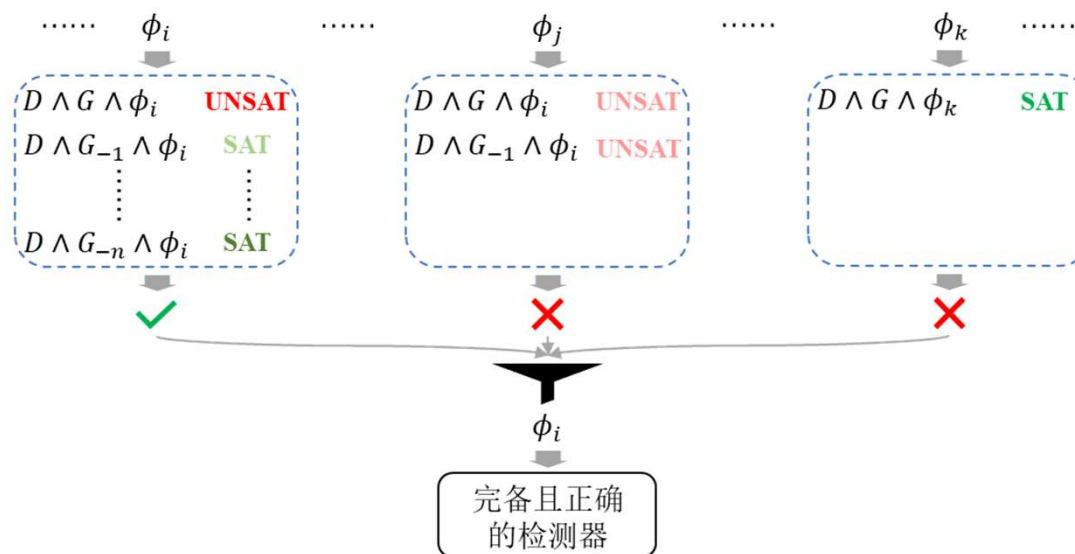
[2] Jianwen Li, Shufang Zhu, Geguang Pu, Lijun Zhang, Moshe Y. Vardi: SAT-based explicit LTL reasoning and its application to satisfiability checking. Formal Methods Syst. Des. 54(2): 164-190 (2019)

[3] Jianwen Li, Geguang Pu, Yueling Zhang, Moshe Y. Vardi, Kristin Y. Rozier: SAT-based explicit LTLf satisfiability checking. Artif. Intell. 289: 103369 (2020)

研究背景和动机

• 相关工作

- 缺乏一种**快速-精确**权衡的LTL可满足性**近似检测方法**
 - 多项式时间 (Polynomial Time)
 - 高精度地检测可满足性





研究背景和动机

- 总结
 - 针对特定领域的LTL公式存在高效的专用算法
 - 专用算法设计困难并且依然存在难以求解的公式
 - 面对可满足性检测繁重任务，缺乏一种快速-精确权衡的方法
- 我们的工作
 - 小模型是否可以进行可满足性**近似检测**？
 - **大模型**是否可以进行可满足性近似检测？
 - **可验证**的线性时态逻辑可满足性近似检测
- 核心挑战
 - 构建LTL在离散域的推理和深度神经网络在连续域的推理的联系



基于逻辑属性启发的端到端方法^[1,2,3]

——小模型是否可以进行可满足性近似检测？

[1] Weilin Luo, Hai Wan, Jianfeng Du, Xiaoda Li, Yuze Fu, Rongzhen Ye, Delong Zhang: Teaching LTLf Satisfiability Checking to Neural Networks. IJCAI 2022: 3292-3298

[2] Weilin Luo, Hai Wan, Delong Zhang, Jianfeng Du, Hengdi Su: Checking LTL Satisfiability via End-to-end Learning. ASE 2022: 21:1-21:13

[3] Weilin Luo, Yuhang Zheng, Rongzhen Ye, Hai Wan, Jianfeng Du, Pingjia Liang, Polong Chen: SAT-Verifiable LTL Satisfiability Checking via Graph Representation Learning. ASE 2023: 1761-1765



基于逻辑属性启发的端到端方法

- 设计具有LTL逻辑属性的神经网络
 - 局部相关
 - 置换不变
 - 序列性
- 分析现有常见神经网络是否可以对齐LTL逻辑属性
 - Transformer
 - Recursive Neural Network (TreeNN)
 - Graph Neural Network (GNN)

基于逻辑属性启发的端到端方法

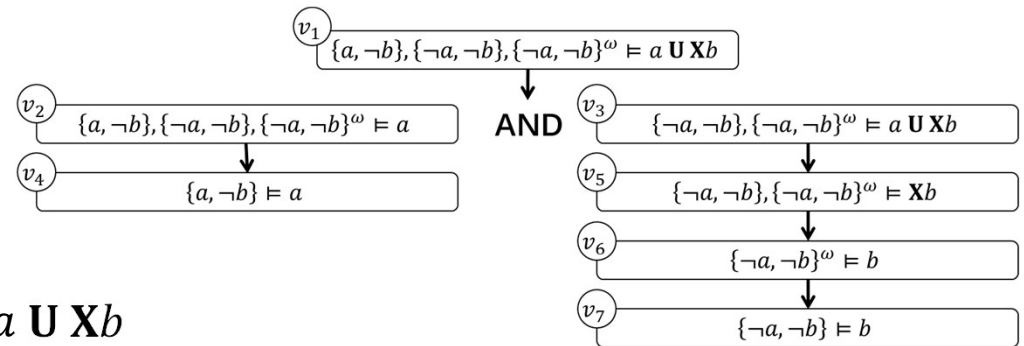
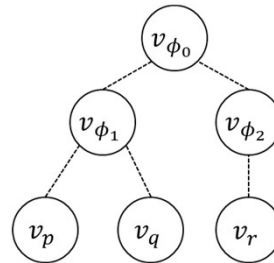
• 局部相关

• 语法

• 例如, $(p \text{ U } q) \wedge \text{X}r$

• 语义

• 例如, $\{a, \neg b\}, \{\neg a, \neg b\}, \{\neg a, \neg b\}^\omega \models a \text{ U } \text{X}b$



• 置换不变

• 子公式置换不变性

• 例如, $p \vee q \equiv q \vee p$

• 原子命题置换不变性

• 例如, $(p \text{ U } q) \wedge \text{G}\neg(r \wedge q)$ vs. $(p \text{ U } a) \wedge \text{G}\neg(r \wedge a)$

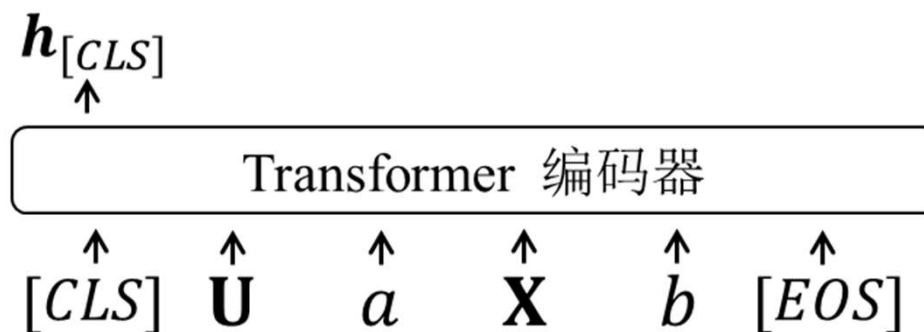
• 序列性

• 例如, $(p \text{ U } q) \wedge \text{G}\neg r$ vs. $(q \text{ U } r) \wedge \text{G}\neg r$



基于逻辑属性启发的端到端方法

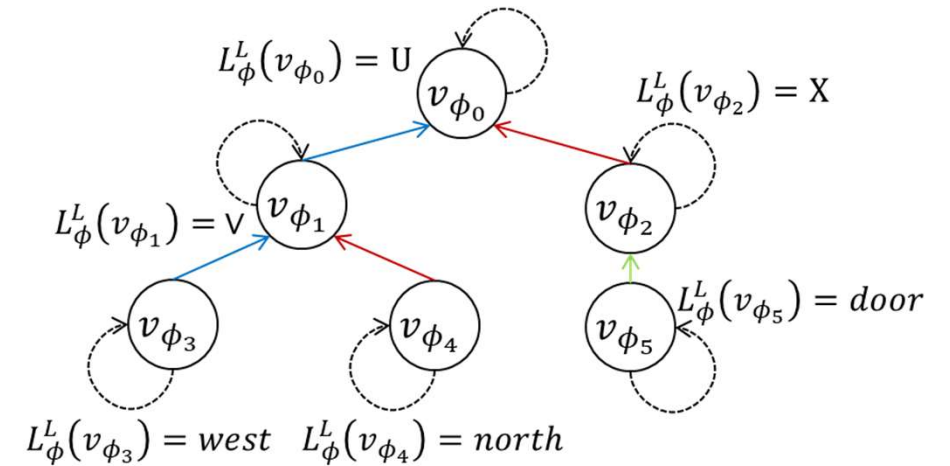
- 基于 Transformer 的嵌入模型
 - LTL公式可以视为词元 (token) 序列
 - 例如, $a \text{ U } Xb \rightarrow \text{U}aXb$
 - 通过训练 Transformer 模型预测可满足性
 - **序列性**: 位置编码
 - **局部相关**和**置换不变**: 多头自注意力机制 (Multi-head Self-attention) 有潜力学习



基于逻辑属性启发的端到端方法

- 基于 GNN 的嵌入模型
 - LTL公式的图表示
 - 抽象语法树
 - 自动机
 - 基于relational graph convolutional network (RGCN) 的嵌入方法^[1]
 - **局部相关**: 消息传递机制 (Message Passing)
 - **序列性**: \mathbf{W}_r

$$\mathbf{x}_v^{(t+1)} = \sigma \left(\sum_{r \in R_\phi^L} \sum_{u \in \mathcal{N}(v, r)} \frac{1}{|\mathcal{N}(v, r)|} \mathbf{W}_r \mathbf{x}_u^{(t)} \right)$$



$(north \vee west) \mathbf{U} \mathbf{X} door$ 基于抽象语法树的图表示

[1] Pashootan Vaezipoor, Andrew C. Li, Rodrigo Toro Icarte, Sheila A. McIlraith: LTL2Action: Generalizing LTL Instructions for Multi-Task RL. ICML 2021: 10497-10508

基于逻辑属性启发的端到端方法

- 基于 GNN 的嵌入模型

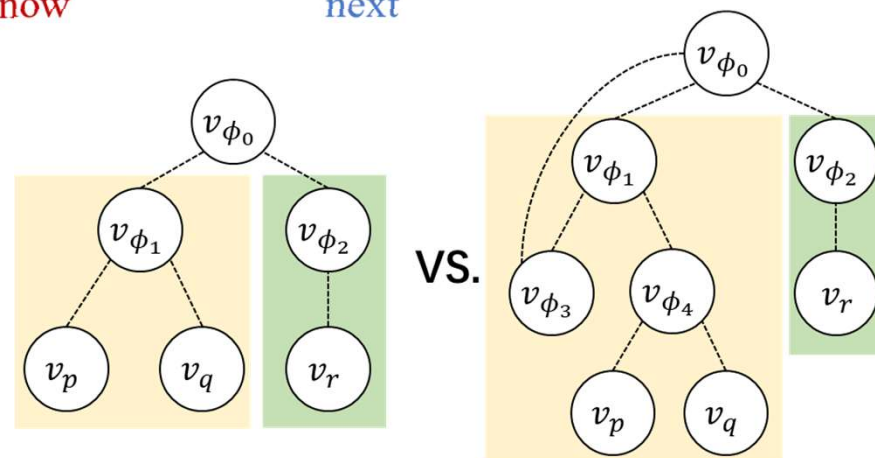
- 一步展开图 (One-step Unfolded Graph, OSUG)

$$\mathbf{X}\phi_i \equiv \underbrace{\top}_{\text{now}} \wedge \underbrace{\mathbf{X}\phi_i}_{\text{next}} \quad \phi_i \mathbf{U} \phi_j \equiv \underbrace{\phi_j}_{\text{now}} \vee (\underbrace{\phi_i}_{\text{now}} \wedge \underbrace{\mathbf{X}(\phi_i \mathbf{U} \phi_j)}_{\text{next}})$$

- **局部相关**：消息传递机制
- **序列性和置换不变性**：节点类型

$$\mathbf{v}_i^{(t)} = \mathbf{W}_1 \mathbf{v}_i^{(t-1)} + \mathbf{W}_2 \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{v}_j^{(t-1)}$$

- 更强的特征抽取能力

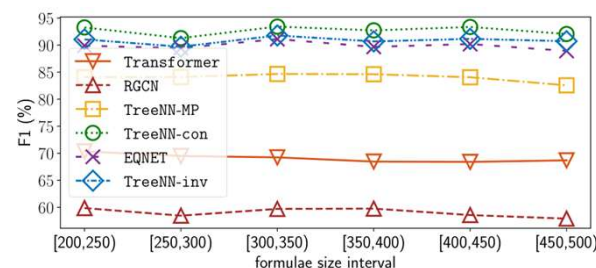


$(p \wedge q) \mathbf{U} \mathbf{X}r$ 或者 $(p \wedge q) \vee \mathbf{X}r$?

基于逻辑属性启发的端到端方法

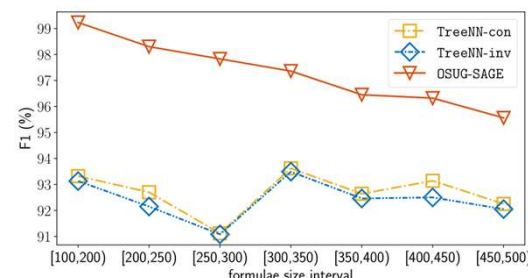
• 实验结果

approach	acc.	pre.	rec.	F1	time
Transformer	70.60	71.02	69.61	70.31	57.09
RGCN	65.42	71.06	52.01	60.06	3,642.99
TreeNN-MP	86.15	90.55	80.73	85.36	1,792.11
TreeNN-con	93.76	98.17	89.19	93.47	1,814.88
EQNET	90.73	94.87	86.13	90.29	485.08
TreeNN-inv	91.79	96.23	87.00	91.38	416.96



- TreeNN > Transformer > RGCN: 满足越多的逻辑属性越有利

approach	acc.	pre.	rec.	F1	time
random	50.00	50.00	50.00	50.00	-
TreeNN-con	93.61	97.70	89.32	93.32	5,371.56
TreeNN-inv	93.42	97.45	89.18	93.13	4,968.67
Transformer	-	-	-	-	-
OSUG-SAGE	98.48	99.58	98.89	99.23	68.11



- OSUG > TreeNN > Transformer : OSUG拥有更强的特征抽取能力



基于大模型迭代交互的迹生成方法^[1]

——大模型是否可以进行可满足性近似检测？

[1] Weilin Luo, Weiyuan Fang, Junming Qiu, Hai Wan, Yanan Liu, Rongzhen Ye: ITG: Trace Generation via Iterative Interaction between LLM Query and Trace Checking. NIER@ICSE 2024: 11-15



基于大模型迭代交互的迹生成方法

- LTL迹生成
 - 例如, 输入 $a \text{ U } \text{X}b$, 输出 $\{a, \neg b\}, \{\neg a, \neg b\}, \{\neg a, \neg b\}^\omega$
 - 验证LTL可满足性检测
- 神经网络是否可以进行端到端的生成?
 - 以**多项式时间**进行**可验证**的LTL可满足性识别





基于大模型迭代交互的迹生成方法

- 相关工作：基于Transformer的迹生成模型^[1]
 - 神经网络有能力进行迹生成
 - 有限原子命题
 - 公式规模受限
 - 训练代价巨大
- 大模型是否可以满足性近似检测？
 - 有能力进行逻辑推理
 - 有能力通过交互修复结果
 - 有限原子命题 -> 理解组合符号的能力
 - 巨大的训练代价 -> 少量样本提示学习

[1] Christopher Hahn, Frederik Schmitt, Jens U. Kreber, Markus Norman Rabe, Bernd Finkbeiner: Teaching Temporal Logics to Neural Networks. ICLR 2021

基于大模型迭代交互的迹生成方法

• ITG: 迭代交互方法

- 期望大模型能够利用丰富的 LTL 专业知识启发式地生成可满足迹
- 大模型可以定位在特定领域, 并通过提示引出该领域的专业知识
- 大模型推理+逻辑推理

LLM Reasoning

Generate a satisfiable trace from input LTL formula. The possible operators are $\&$, $|$, $!$, X , F , G , U . {... more ...}
LTL: $a \ U \ (Xb)$.



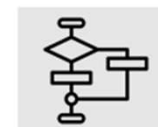
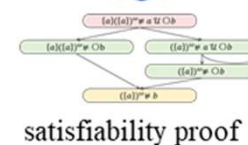
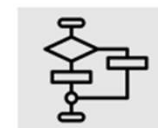
Trace: $[[a], [a]]$ FINISH

The trace you give does not satisfy the LTL formula. There is a proof to guide you to regenerate a satisfiable trace. {... more ...}



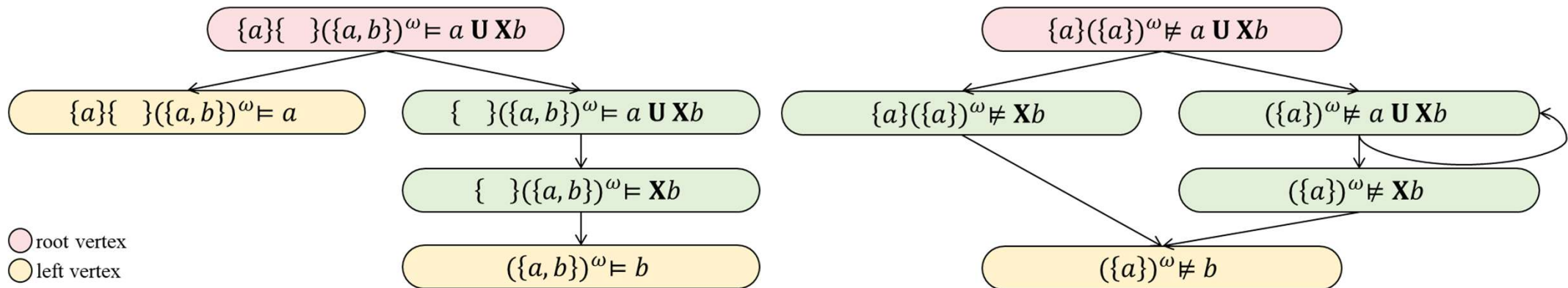
Repaired_Trace: $[[a], [a], [b]]$ FINISH

Logical Reasoning



基于大模型迭代交互的迹生成方法

- 如何**高效**反馈生成迹的**正确性**?
 - 迹和LTL公式的**可满足性证明**
 - 多项式时间
 - 例如, $\{a\}\{\ }(\{a, b\})^\omega \models a \text{ U } \text{X}b$ 和 $\{a\}(\{a\})^\omega \not\models a \text{ U } \text{X}b$





基于大模型迭代交互的迹生成方法

- 初始提示

- 输入（公式）+输出（迹）+迹生成任务描述+约束

1 Generate a satisfiable trace from input LTL formula. The possible atomic propositions
2 will be given. The possible operators are $\&$, $|$, $!$, X , F , G , U . Trace should be less than
3 10 states. Each state should not contain duplicated atomic propositions. The output trace
4 should be a list of states. Do not use ... in output. For example: LTL: $a \ U \ (X \ b)$ Trace:
5 $[[a],[],[b]]$ FINISH LTL: $F(a \ \& \ X \ b)$ Trace: $[[a],[b]]$ FINISH LTL: $G(a \ | \ ! \ b)$ Trace:
6 $[[],[a]]$ FINISH LTL: $a \ U \ (X \ b)$ Trace: $[[a],[a]]$ FINISH

- 修复提示

- 输入（公式，迹，证明）+输出（修复迹）+实例

1 The trace you give does not satisfy the LTL formula. There is a proof to guide you to
2 regenerate a satisfiable trace. For example: LTL: $a \ U \ (X \ b)$ Trace: $[[a],[a]]$ Proof: $\{[[a]]$
3 not satisfies b ; $[[a]]$ not satisfies $X \ b$; $[[a]]$ not satisfies $a \ U \ (X \ b)$; $[[a],[a]]$ not satisfies
4 $X \ b$ Repaired Trace: $[[a], [a], [b]]$ FINISH



基于大模型迭代交互的迹生成方法

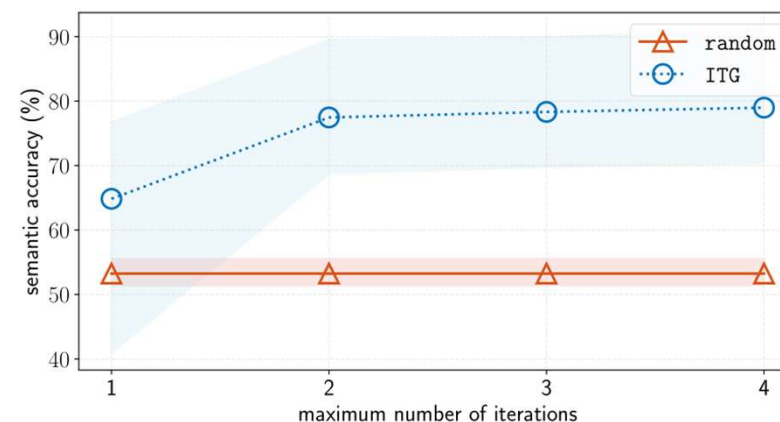
• 实验结果

- SOTA 神经网络的方法无法拟合大规模公式
- ITG可以进行有效的迹生成
- ITG具有跨公式长度的泛化能力
- 可满足性证明对于提高迹生成性能是有效的
- ITG优于其他常见的基于提示的方法

	[5, 20)	[20, 40)	[40, 60)	[60, 80)	[80, 100)
CoT-node	77.00	74.50	69.30	64.70	59.80
CoT-tree	80.20	74.20	69.50	65.40	62.30
CoT-SC	88.00	80.40	76.50	73.60	72.60
ITG	91.20	81.00	70.30	75.50	75.00

	[5, 20)	[20, 40)	[40, 60)	[60, 80)	[80, 100)
Transformer	93.30	76.21	58.66	58.32	56.35

	[5, 20)	[20, 40)	[40, 60)	[60, 80)	[80, 100)
random	54.20	52.10	53.10	53.10	54.00
Transformer	93.30	71.30	60.30	54.20	51.60
ITG-init	69.90	61.50	59.60	57.60	57.20
ITG	91.20	81.00	70.30	75.50	75.00





基于可微迹检测的可满足性检测^[1]

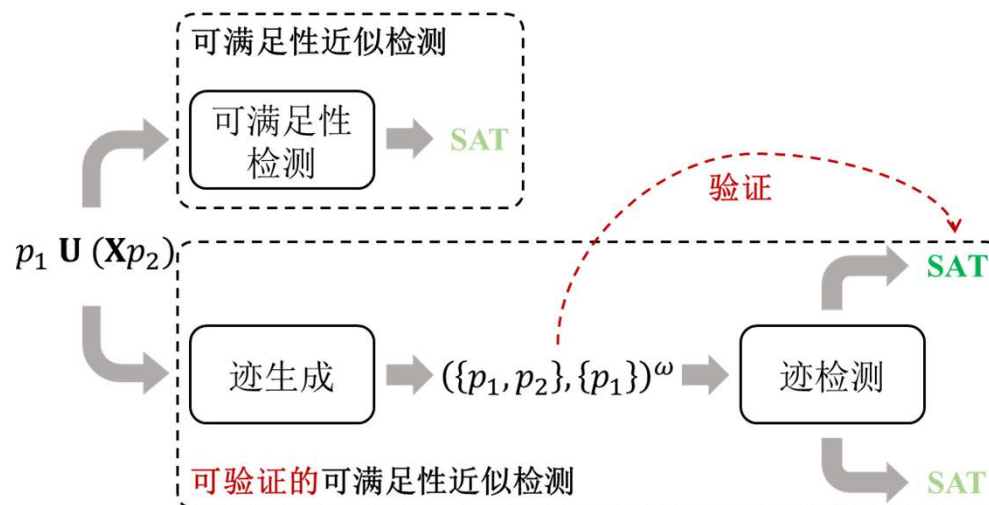
——可验证的线性时态逻辑可满足性近似检测

[1] Weilin Luo, Pingjia Liang, Junming Qiu, Polong Chen, Hai Wan, Jianfeng Du, Weiyuan Fang: Learning to SAT-verifiably Check LTL Satisfiability via Differentiable Trace Checking. ISSTA 2024



基于可微迹检测的可满足性检测

- 研究现状
 - 迹生成方法无法扩展到具有大规模原子命题的公式
 - 逐字符输出的生成模型效率低下
- 核心挑战
 - 连续域和离散域之间的差距
 - 多解问题
- 我们的方法
 - LTL 编码 \rightarrow 神经迹检测
 - 联合学习





基于可微迹检测的可满足性检测

- LTL 编码

- 语法树的先序遍历序列 $\langle v_1, \dots, v_{|\phi|} \rangle$, v_i 对应一个子公式
 - v_{i+1} 对应的子公式是 v_i 对应的子公式的**左子公式**
- 6元组 $\eta = (\eta_{right}, \eta_{atom}, \eta_{\neg}, \eta_{\wedge}, \eta_X, \eta_U)$
 - $(\eta_{right})_{i,j}$: v_i 对应的子公式的**右子公式**是否是 v_j 对应的子公式
 - $(\eta_{atom})_{i,j}$: v_i 对应的子公式是否是**原子命题** p_j
 - $(\eta_{op})_i$: v_i 对应的子公式是否是**op-公式**, 其中 $op \in \{\neg, \wedge, X, U\}$
- 例如, $\langle Up_1 Xp_2 \rangle$ 对应LTL编码是
 - $(\eta_U)_1 = 1$, $(\eta_{right})_{1,3} = 1$, $(\eta_{atom})_{2,1} = 1$, $(\eta_X)_3 = 1$, $(\eta_{atom})_{,2} = 1$
 - 其它为0



基于可微迹检测的可满足性检测

- 张量化迹

- 2元组(s, l)

- $(s)_{i,j}$: i 时刻 p_j 是否为真

- $(l)_i$: i 时刻是否是环开始

- 例如, $(\{p_1, p_2\}, \{p_1\})^\omega$ 的张量化迹是

$$s = \begin{bmatrix} 1.0 & 1.0 \\ 1.0 & 0.0 \end{bmatrix}, \quad l = \begin{bmatrix} 1.0 \\ 0.0 \end{bmatrix}$$

- NTCNet: LTL 编码参数化神经网络

- $(\theta_{right}, \theta_{atom}, \theta_{\neg}, \theta_{\wedge}, \theta_X, \theta_U) = (\eta_{right}, \eta_{atom}, \eta_{\neg}, \eta_{\wedge}, \eta_X, \eta_U)$



基于可微迹检测的可满足性检测

- NTCNet的推理：可微迹检测

- 输入：张量化迹，LTL编码
- 推理过程 ($NTCNet(*)$)

$$(p_i)_j = \sum_{k=1}^{|\mathbb{P}|} (\theta_{atom})_{j,k} (s)_{i,k}. \quad (3)$$

$$(r_i)_j = \sum_{k=j+2}^{|\pi|} (\theta_{right})_{j,k} (x_i)_k. \quad (4)$$

$$(x_i)_j = \sigma((p_i)_j + (\theta_{\neg})_j \sigma(1 - (x_i)_{j+1}) + (\theta_{\wedge})_j \sigma((x_i)_{j+1} + (r_i)_j - 1) + (\theta_{\circ})_j (x_{i+1})_{j+1} + (\theta_{\mathcal{U}})_j (u_i)_j^{(T)}), \quad (1)$$

$$(x_{|\pi|+1})_{j+1} = \sum_{k=1}^{|\pi|} (l)_k (x_k)_{j+1}, \quad (2) \quad (u_i)_j^{(t)} = \begin{cases} (r_i)_j, & t = 1, \\ \sigma(\sigma((x_i)_{j+1} + (u_{i+1})_j^{(t-1)} - 1) + (u_i)_j^{(1)}), & \text{otherwise}, \end{cases} \quad (5)$$

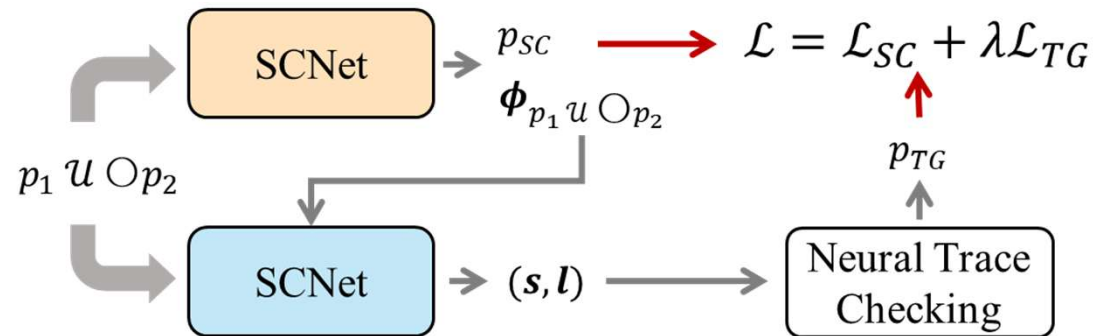
$$(u_{|\pi|+1})_j^{(t)} = \sum_{k=1}^{|\pi|} (l)_k (u_k)_j^{(t)}. \quad (6)$$

- **定理 1.** 令 ϕ 是LTL公式， η 是 ϕ 的LTL编码。对于任意的迹 π ，其张量化迹为 (s, l) ， $NTCNet((s, l), \eta) = 1$ 当且仅当 $\pi \models \phi$ 。

基于可微迹检测的可满足性检测

• SAT可验证的可满足性检测

训练



推理

Algorithm 2: SAT-verifiable LTL satisfiability checking

Input: An LTL formula ϕ .

Output: Whether ϕ is satisfiable or not and a satisfiable trace π of ϕ if it is satisfiable.

```

1  $(p_{SC}, \phi) \leftarrow \text{SCNet}(\phi)$ 
2  $(s, l) \leftarrow \text{TGNet}(\phi)$ 
3  $\pi \leftarrow \text{ROUND}((s, l))$ 
4  $st \leftarrow \text{TRACECHECK}(\pi, \phi)$ 
5 if  $st$  is SAT or  $p_{SC} > \beta_{SAT}$  then
6   | return SAT,  $\pi$ 
7 else
8   | return UNSAT
    
```

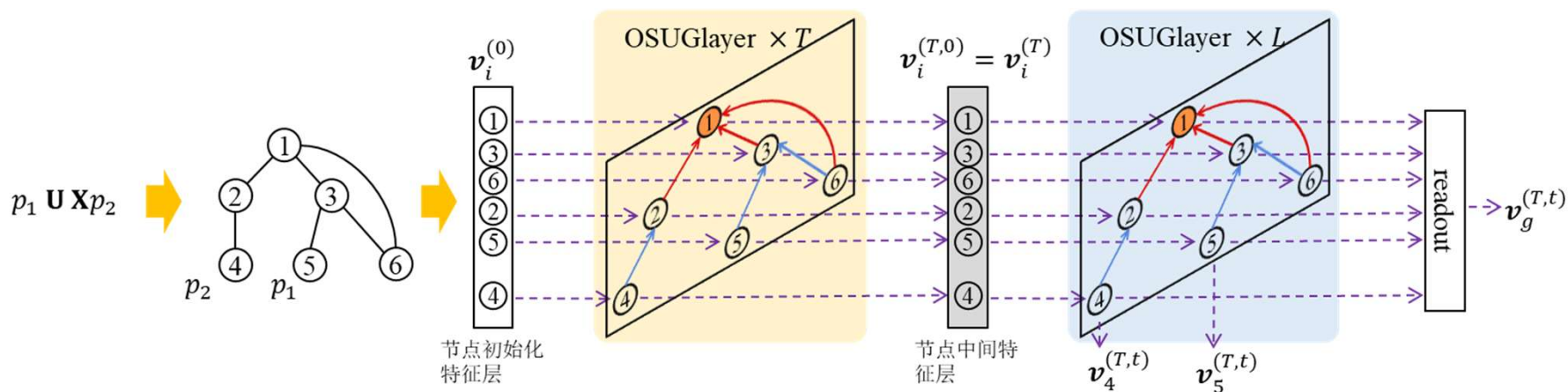
基于可微迹检测的可满足性检测

- SCNet: 端到端的可满足性检测
 - OSUGlayer: 基于OSUG图的表示学习
- TGNet: 端到端的迹生成
 - Encoder: 多层OSUGlayer
 - Decoder: 一层OSUGlayer解码一个时刻

$$(s)_{i,j} = (\text{softmax}(\text{ReLU}(W_9 \text{ReLU}(W_8 \mathbf{v}_j^{(T,i)} + b_8) + b_9)))_2 \quad (7)$$

$$l = \text{softmax}([x_1, \dots, x_{N_t}]),$$

$$x_i = W_{11} \text{ReLU}(W_{10} \mathbf{v}_g^{(T,i)} + b_{10}) + b_{11}, \quad (8)$$



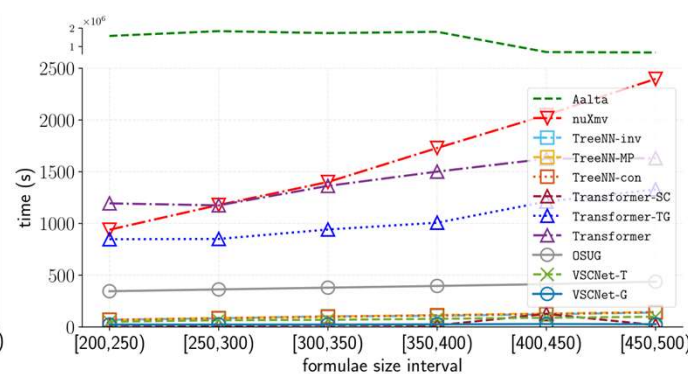
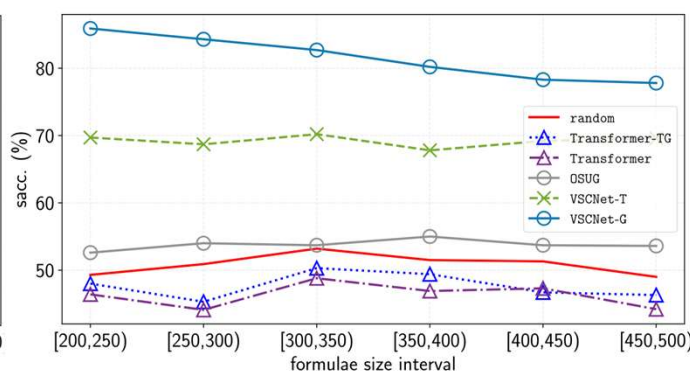
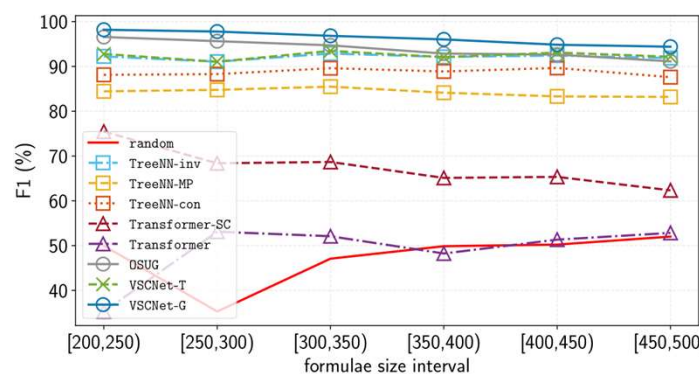


基于可微迹检测的可满足性检测

• 实验结果

- VSCNet拟合分布内数据的能力最强
- VSCNet具有一定跨公式长度泛化的能力
- 运行时间具有显著优势

approach	acc.	pre.	rec.	F1	sacc.	time
random	50.01	50.01	49.28	49.64	51.88	0
Aalta	100.00	100.00	100.00	100.00	100.00	12,784,302
nuXmv	100.00	100.00	100.00	100.00	100.00	6,848
TreeNN-inv	93.27	97.75	88.58	92.94	-	434
TreeNN-MP	86.54	90.74	81.40	85.82	-	455
TreeNN-con	89.38	95.96	82.22	88.56	-	459
Transformer-SC	72.56	74.24	69.10	71.58	-	104
Transformer-TG	-	-	-	-	47.67	5,484
Transformer	59.30	62.79	45.64	52.86	45.09	10,601
OSUG	98.47	99.15	97.79	98.46	55.22	2,816
VSCNet-T (our)	93.45	97.47	89.22	93.16	71.86	364
VSCNet-G (our)	99.15	99.47	98.83	99.15	91.01	156





总结

1. 提出了一种可验证的LTL可满足性近似检测方法
2. 探究了大模型在LTL推理任务上的潜力
3. 两种神经符号融合推理机制
 - 神经网络属性对齐逻辑属性
 - 研究神经网络在逻辑推理任务上的表达能力



Thank You



Backup

- Semantics of LTL

- $\pi_t \models p$ iff $p \in s_t, p \in \mathbb{P}$
- $\pi_t \models \neg\varphi$ iff $\pi_t \not\models \varphi$
- $\pi_t \models \varphi_1 \vee \varphi_2$ iff $\pi_t \models \varphi_1$ or $\pi_t \models \varphi_2$
- $\pi_t \models F\varphi$ iff $\pi_{t+1} \models \varphi$
- $\pi_t \models \varphi_1 U \varphi_2$ iff $\exists k \geq t, \pi_k \models \varphi_2$ and $\forall t \leq j < k, \pi_j \models \varphi_1$