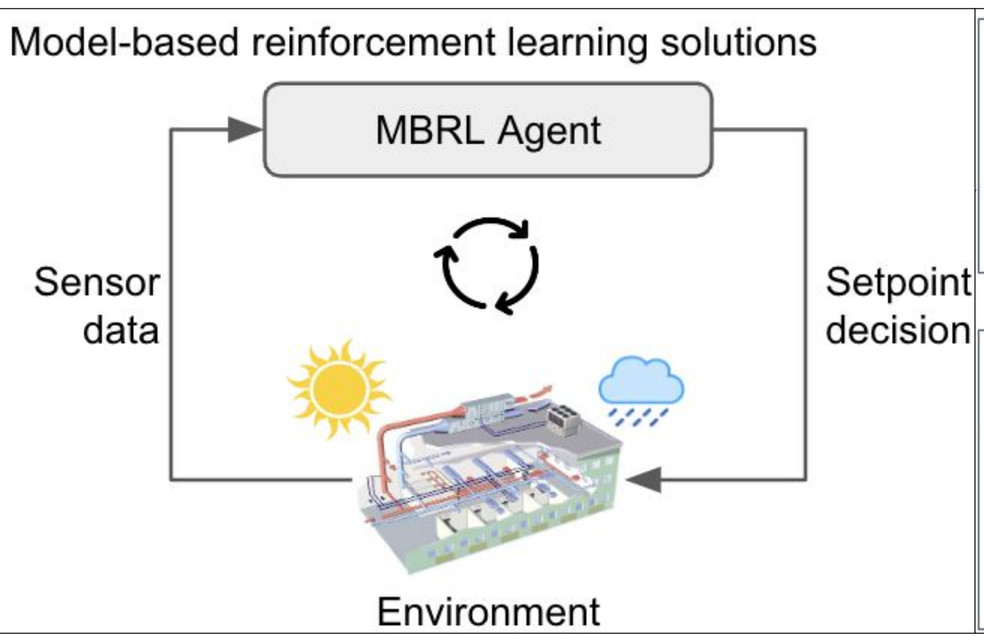
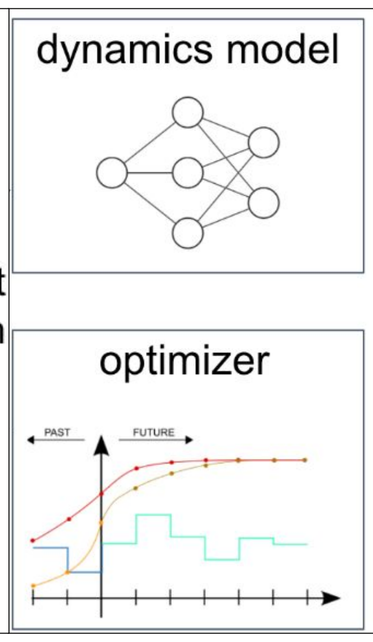
Go Beyond Black-box Policies:
Rethinking the Design of Learning Agent for Interpretable and Verifiable HVAC Control

Zhiyu An, Xianzhong Ding, Wan Du
University of California, Merced

1. Model-based Reinforcement Learning (MBRL) is used to improve energy efficiency of HVAC systems. Although they work in labs, they lack properties that allow us to verify their safety. This motivated us to rethink this method.



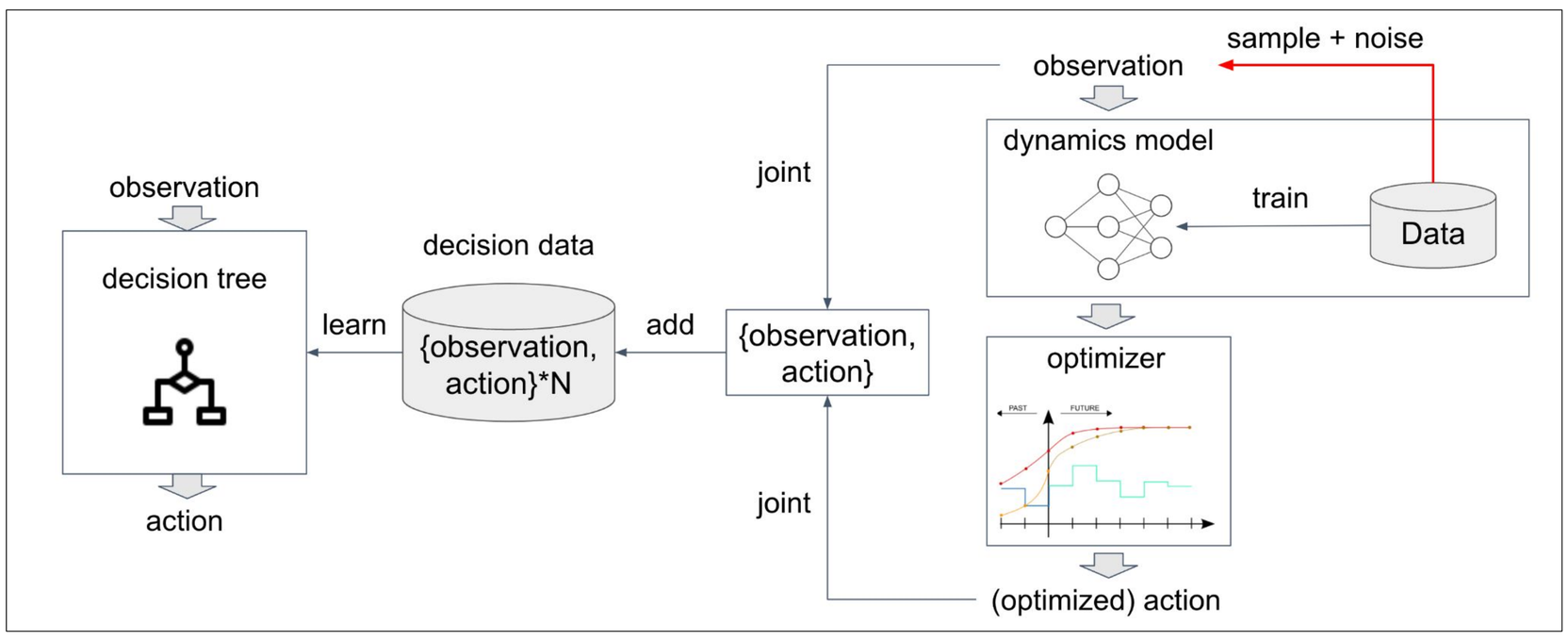


Not interpretable
(incomprehensible decision process)

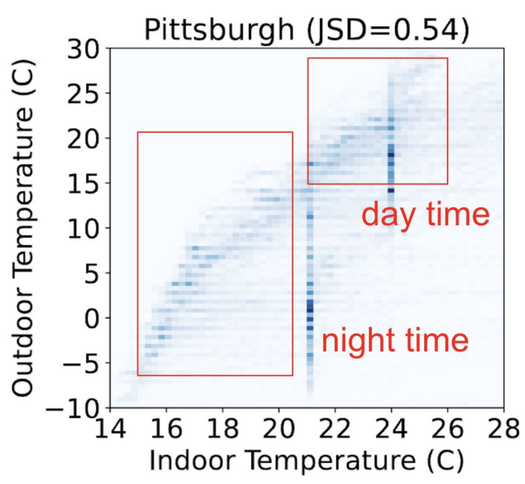
Stochastic
(chance of failure is always > 0)

Rethink the design of the agent!

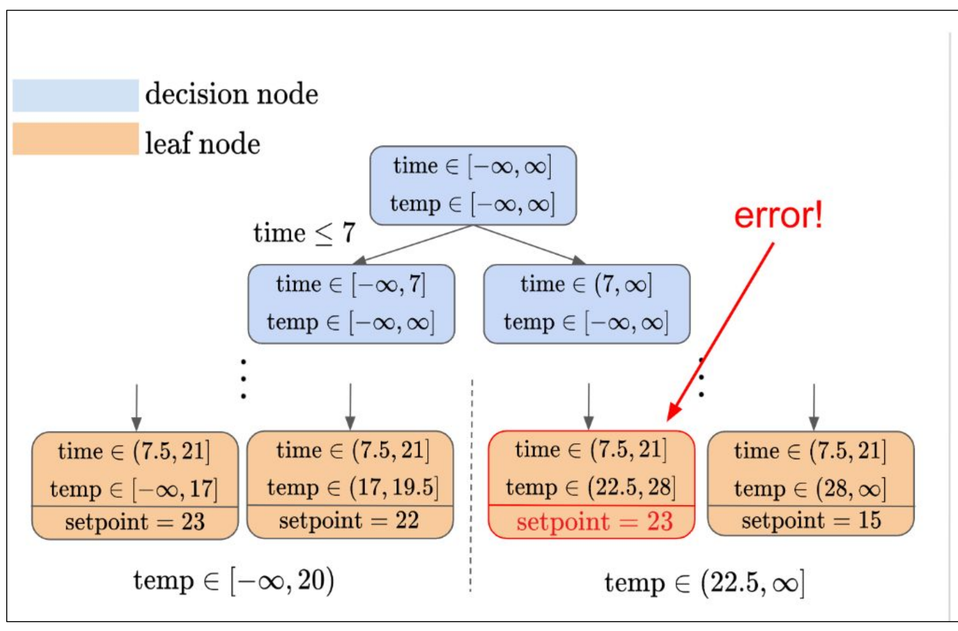
2. We propose to use decision tree as the control agent, which has the desired properties. We designed a process to learn the decision tree from the actions optimized by MBRL. We take advantage of the bias in data for efficient learning.

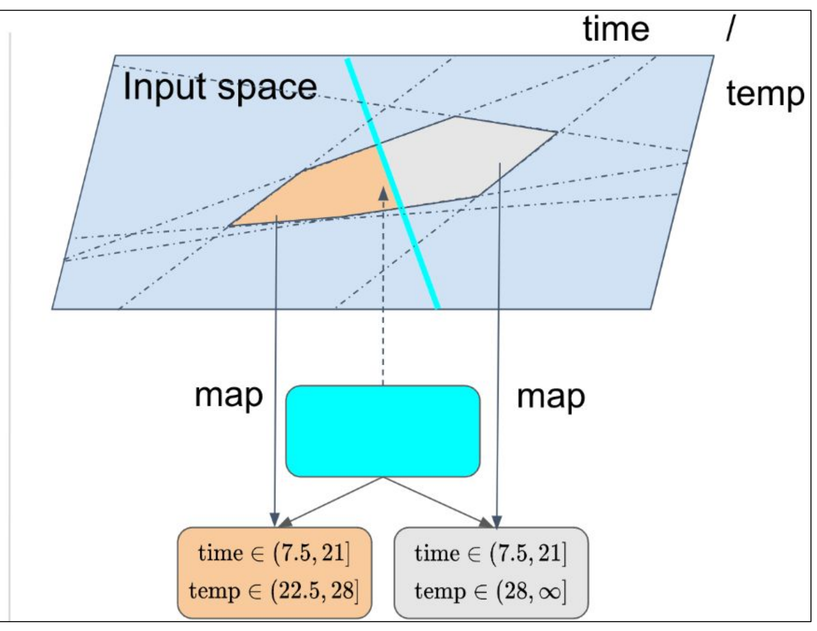


$$\widehat{p(x)} = X + \mathcal{N}\left(0, \text{noise\_level} \times \sqrt{\frac{\sum(x_i - \bar{x})^2}{|X|}}\right)$$



3. We designed a Decision Path Verification algorithm that finds the input set for every leaf node. This enables us to verify the input-output logical statements. This algorithm also allows us to interpret the tree by finding the in-out mapping.

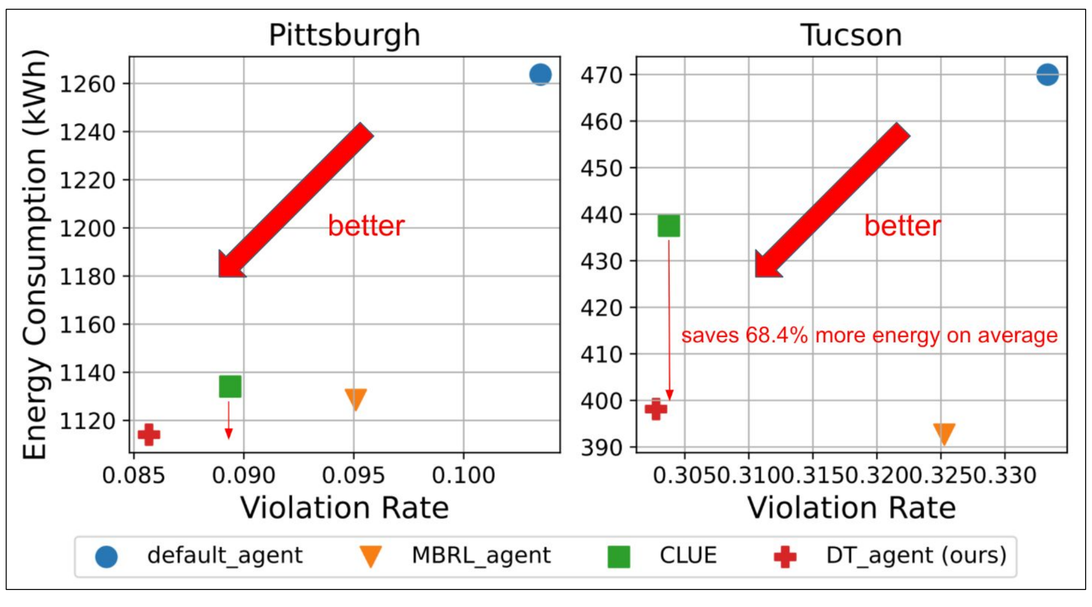




Algorithm 1: Decision path verification

for *each leaf node* T_i **do**
  $P = \{T_0, \cdots, T_i\} \leftarrow$ extract path from T_0 to T_i
  $C = \mathbb{R}^{|\mathcal{X}|}$ ▸ initialize state box boundaries
  for $T_j \in (P \setminus T_i)$ **do**
    $C \leftarrow C \bigcap$ (input box of the rules from T_j to T_{j+1})
  if $C \subseteq ((s_t > \overline{z}) \vee (s_t < \underline{z})) \bigcup \mathbb{R}^{|\mathcal{X}|-1})$ **then**
    check criteria compliance by Eq.4

4. We evaluated our method in EnergyPlus simulations using real weather profiles from two cities. Our method shows higher energy efficiency (increase energy savings by 68.4%), fast convergence, and fast inference(1124x less overhead).



	Pittsburgh	Tucson
Total No. of nodes	1199	3291
No. of leaf nodes (unique path)	599	1646
Safe probability estimated by crit. #1	94.6%	95.1%
No. of nodes corrected by crit. #2	0	0
No. of nodes corrected by crit. #3	0	88

	default [12]	MBRL [9]	CLUE [1]	DT (ours)
average (ms)	0.0	212.87	326.30	0.1888
std (ms)	0.0	266.89	102.30	0.4423

code available at https://github.com/ryeii/Veri_HVAC

DAC 2024, June 23-27, Moscone West, San Francisco