Safety verification of decision-tree policies in continuous time

NeurIPS 2023 spotlight


Christian Schilling


Anna Lukina


Emir Demirović


Kim Guldstrand Larsen

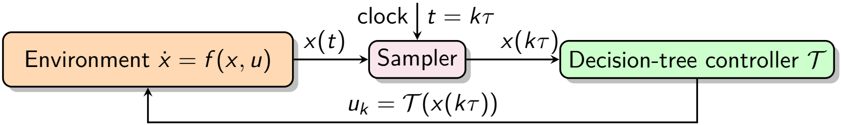
AALBORG UNIVERSITY

TUDelft

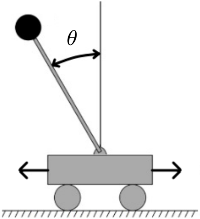
Decision-tree control systems

A **decision-tree control system** (DTCS) is a triple $(f, \mathcal{T}, \tau)$ with

- f: continuous-time environment $\dot{x} = f(x, u) : \mathbb{R}^{n+m} \to \mathbb{R}^n$
- $\mathcal{T}$: decision-tree policy $\mathcal{T} : \mathbb{R}^n \to U$ where $U \subseteq \mathbb{R}^m$
- τ: control period $\tau \in \mathbb{R}^+$



Example: Cart/pole system

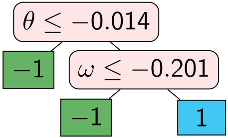


$$\dot{p} = v \qquad \phi = \frac{9.8\sin(\theta) - \cos(\theta)\psi}{2/3 + 5/11\cos(\theta)^2}$$

$$\dot{\theta} = \omega$$

$$\dot{\omega} = \phi \qquad \psi = \frac{10\mathbf{u} + 0.05\omega^2\sin(\theta)}{1.1}$$

$$\dot{v} = \psi - \frac{1}{22}\phi\cos(\theta)$$

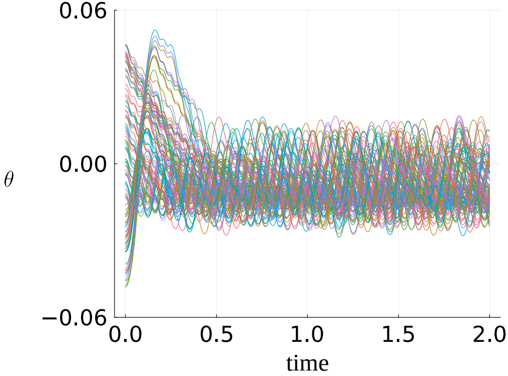


$\tau = 0.02$

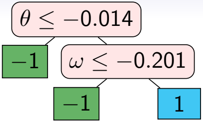
Problem statement

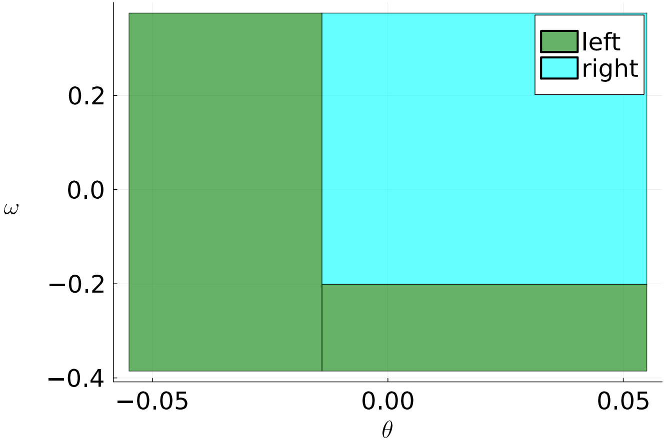
Given a DTCS $(f, \mathcal{T}, \tau)$, a set of initial states $\mathcal{X}_0 \subseteq \mathbb{R}^n$, and an iteration bound $k_{\max}$, compute the **set of reachable states**

Example: $p_0 = v_0 = 0$; $\theta_0, \omega_0 \in [-0.05, 0.05]$ and $k_{\max} = 100$



Reachability example

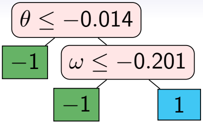


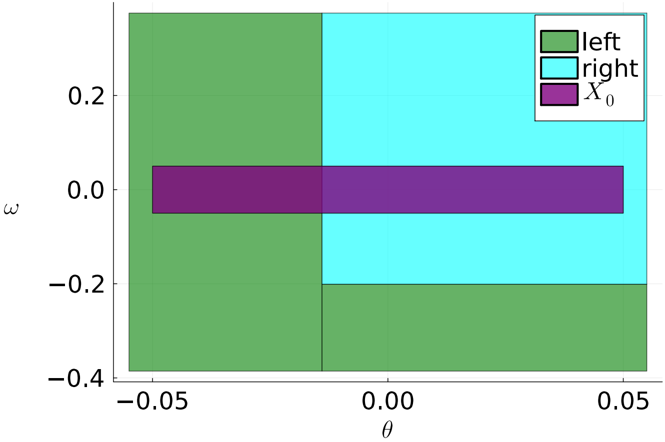


Reachability example

$\mathcal{X}_0$: $p_0 = v_0 = 0$; $\theta_0, \omega_0 \in [-0.05, 0.05]$

τ: 0.02

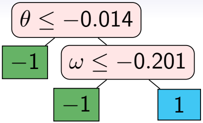


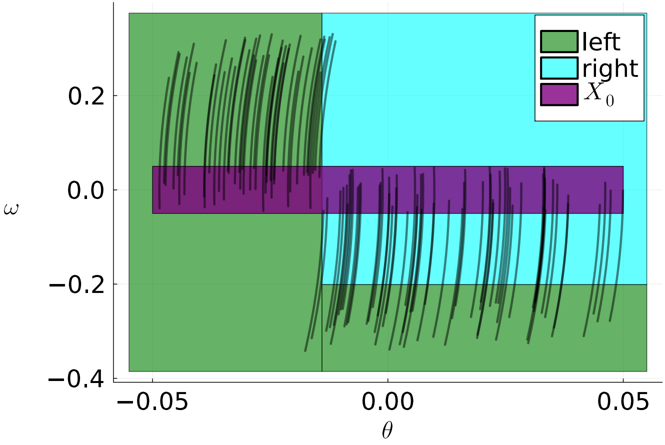


Reachability example

$\mathcal{X}_0$: $p_0 = v_0 = 0$; $\theta_0, \omega_0 \in [-0.05, 0.05]$

τ: 0.02

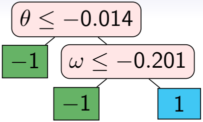


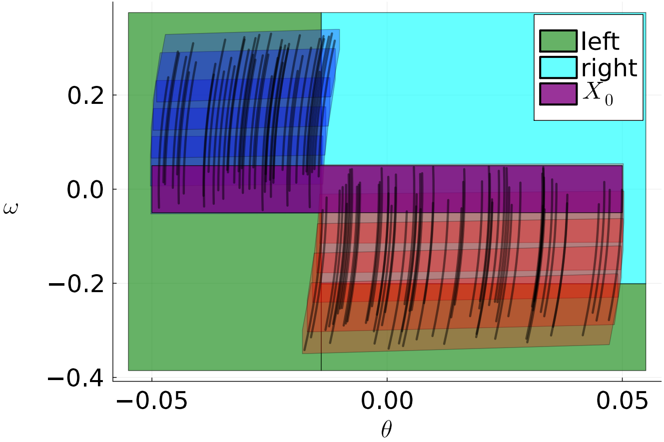


Reachability example

$\mathcal{X}_0$: $p_0 = v_0 = 0$; $\theta_0, \omega_0 \in [-0.05, 0.05]$

τ: 0.02

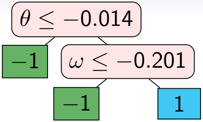


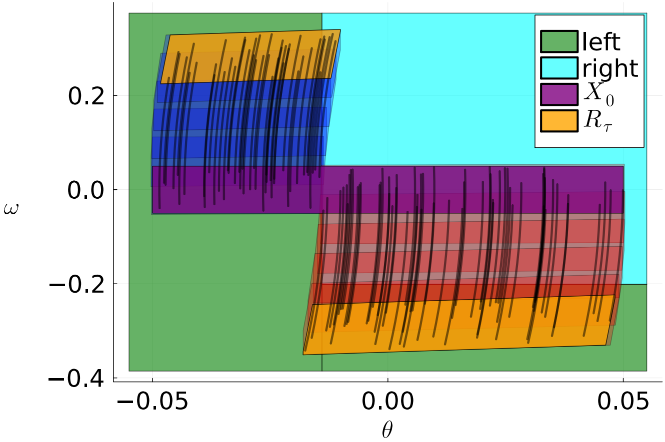


Reachability example

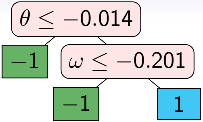
$\mathcal{X}_0$: $p_0 = v_0 = 0$; $\theta_0, \omega_0 \in [-0.05, 0.05]$

τ: 0.02



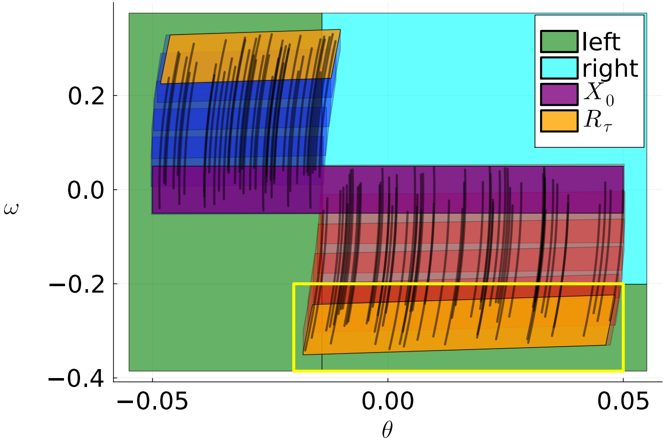


Reachability example



$\mathcal{X}_0$: $p_0 = v_0 = 0$; $\theta_0, \omega_0 \in [-0.05, 0.05]$

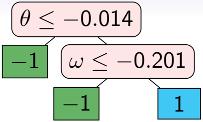
τ: 0.02

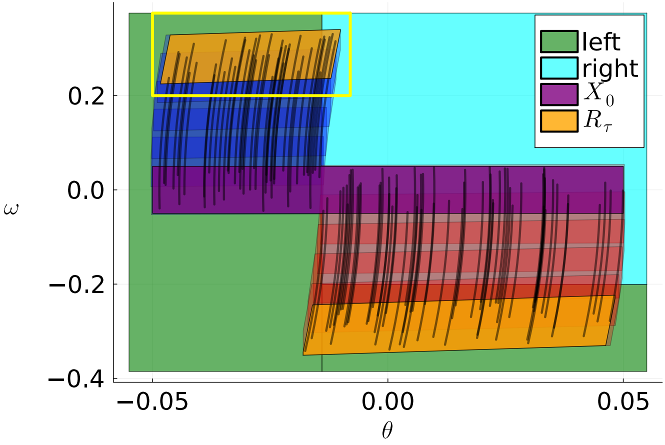


Reachability example

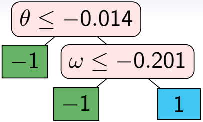
$\mathcal{X}_0$: $p_0 = v_0 = 0;\ \theta_0, \omega_0 \in [-0.05, 0.05]$

τ: 0.02



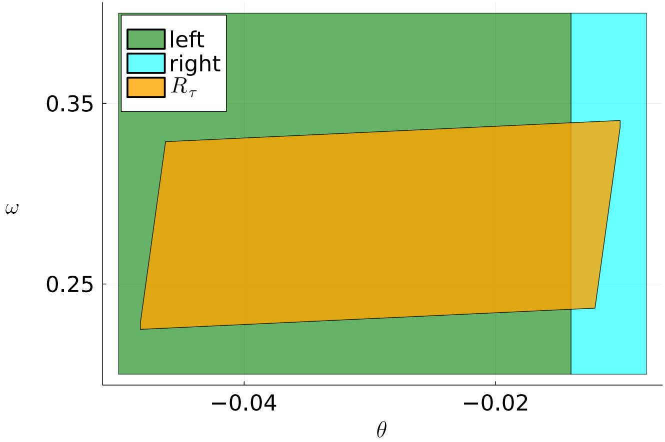


Reachability example



$\mathcal{X}_0$: $p_0 = v_0 = 0;\ \theta_0, \omega_0 \in [-0.05, 0.05]$

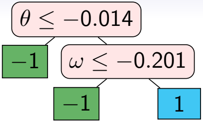
τ: 0.02

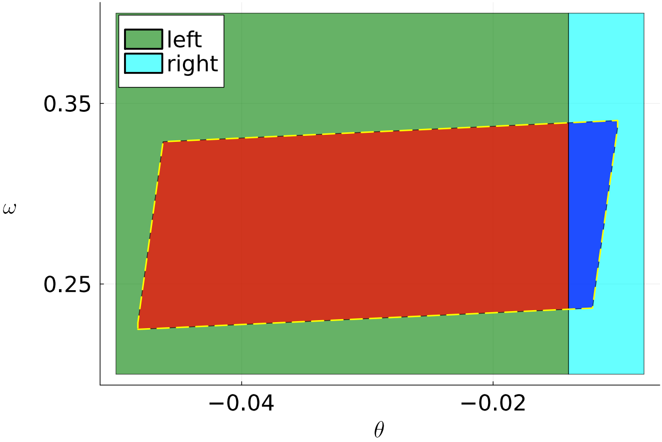


Reachability example

$\mathcal{X}_0$: $p_0 = v_0 = 0$; $\theta_0, \omega_0 \in [-0.05, 0.05]$

τ: 0.02

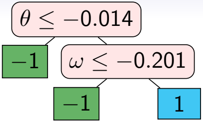


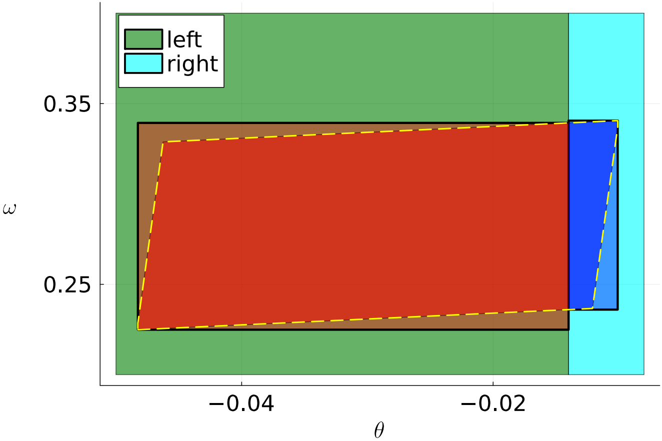


Reachability example

$\mathcal{X}_0$: $p_0 = v_0 = 0$; $\theta_0, \omega_0 \in [-0.05, 0.05]$

τ: 0.02





Reach-avoid problem

- By computing the **reachable states** we can analyze **reach-avoid specifications**: "avoid A and reach R at time t"

$$\Box^{\leq t} \neg A \ \wedge \ \Diamond^{t} R$$

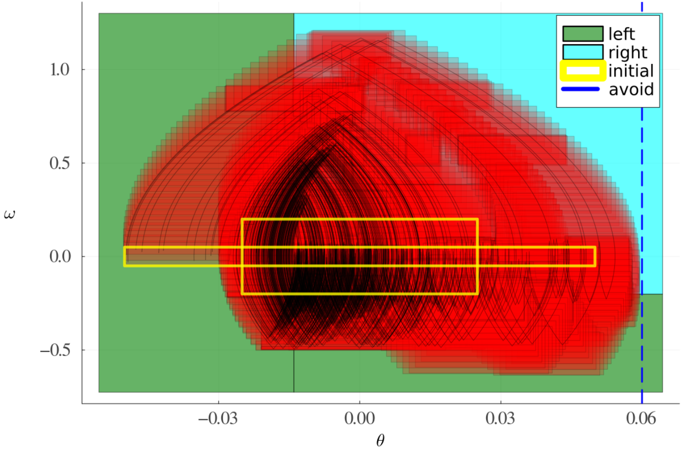
- Deciding whether such specifications hold is **undecidable** already for nonlinear environments $f(x, u)$
- If $f(x, u) = u$, we call f **state independent**

Theorem

The **reach-avoid problem** for **state-independent** DTCS is

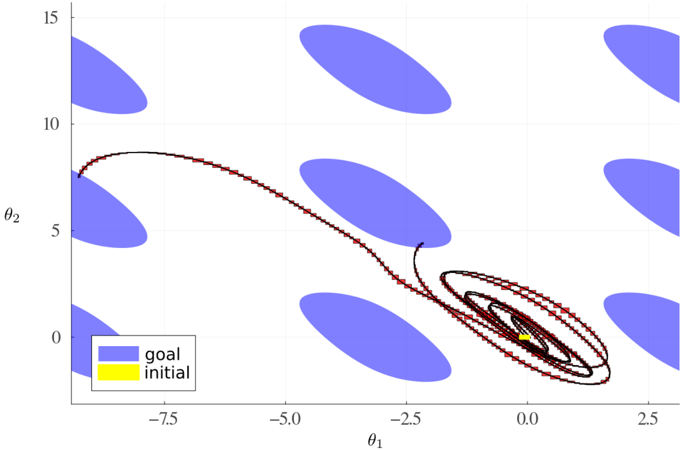
1. **undecidable** for unbounded time
2. **PSPACE-complete** for bounded time

Experimental evaluation



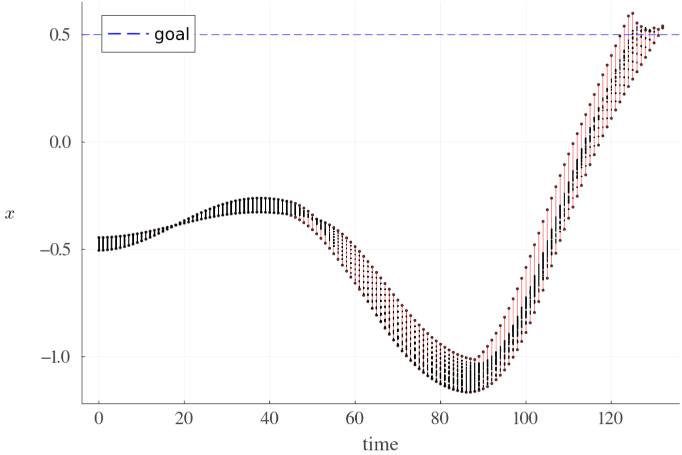
Stabilization of cart/pole policy

Experimental evaluation



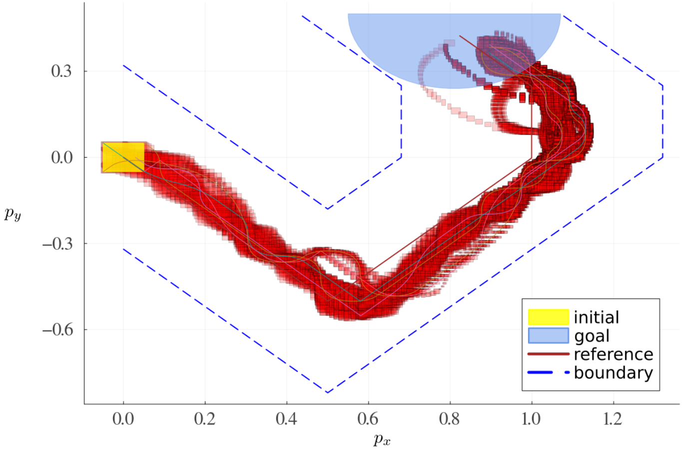
Acrobot policy reaching swing-up goal

Experimental evaluation



Car policy reaching top of the mountain (discrete-time setting)

Experimental evaluation



Quadrotor policy following reference trajectory to goal region

Contributions

- **Parametric reachability algorithm** with sufficient conditions for soundness and relative completeness
- **Instantiated algorithm** based on Taylor models and axis-aligned decisions (“$x \leq c$”), exploiting problem structure
- **Complexity proof**
- **Public implementation** and experimental evaluation
  https://neurips.cc/virtual/2023/poster/70218