Hyperproperties & Autonomy

Miroslav Pajic

Cyber-Physical Systems Lab – CPSL@Duke

Department of Electrical and Computer Engineering

Department of Computer Science

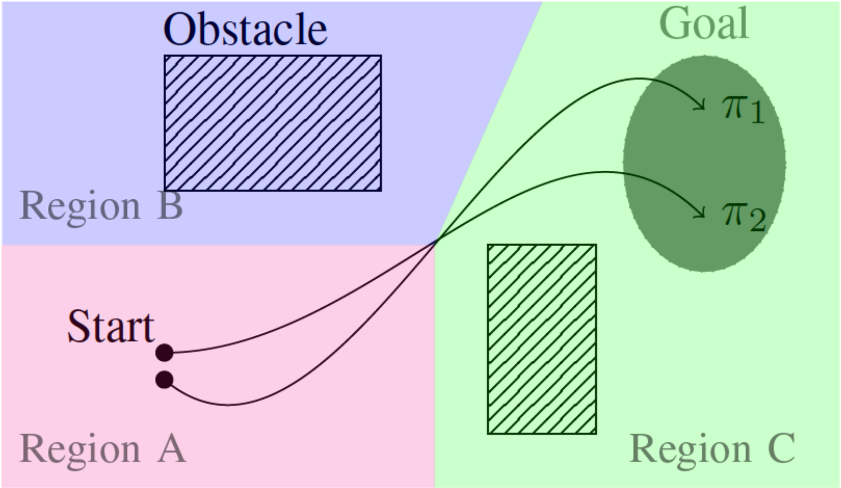
Duke University

Duke

PRATT SCHOOL of ENGINEERING

DUKE ROBOTICS

Hyperproperties for Motion Planning



Motion Planning with Privacy/Opacity

$$\exists \pi_1 \exists \pi_2 . (\pi_1 \text{ and } \pi_1 \text{ are different paths}) \\ \wedge (\pi_1 \text{ and } \pi_1 \text{ give identical observation}) \\ \wedge (\pi_1 \text{ and } \pi_1 \text{ reach goal}).$$

$$\exists \pi_1 \exists \pi_2 . (\mathrm{sec}(\pi_1) \neq \mathrm{sec}(\pi_2)) \wedge (\mathrm{obs}(\pi_1) = \mathrm{obs}(\pi_2))$$

Optimality of Synthesized Plans

$$\exists \pi . \Big((\pi \text{ reaches goal}) \wedge \\ (\forall \pi' . ((\pi' \text{ reaches goal}) \Rightarrow (\pi \text{ reaches goal})) \Big)$$

$$\exists \pi_1 \forall \pi_2 . \ (\mathbf{s_0}^{\pi_1} \wedge \mathbf{s_0}^{\pi_2}) \wedge (\Diamond_T (g^{\pi_2} \Rightarrow \Diamond_T g^{\pi_1}));$$

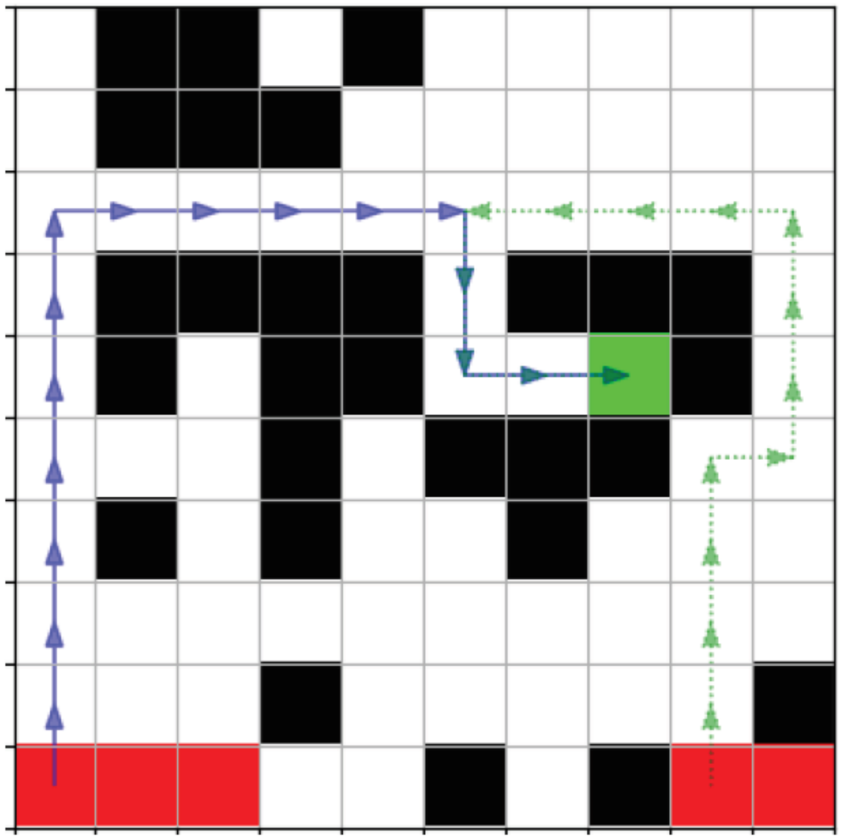
$$\exists \pi_1 \forall \pi_2 . \ (\mathbf{s_0}^{\pi_1} \wedge \mathbf{s_0}^{\pi_2}) \wedge (\Diamond_T (g^{\pi_1} \Rightarrow \Diamond_T g^{\pi_2}))$$

Robustness of Synthesized Plans

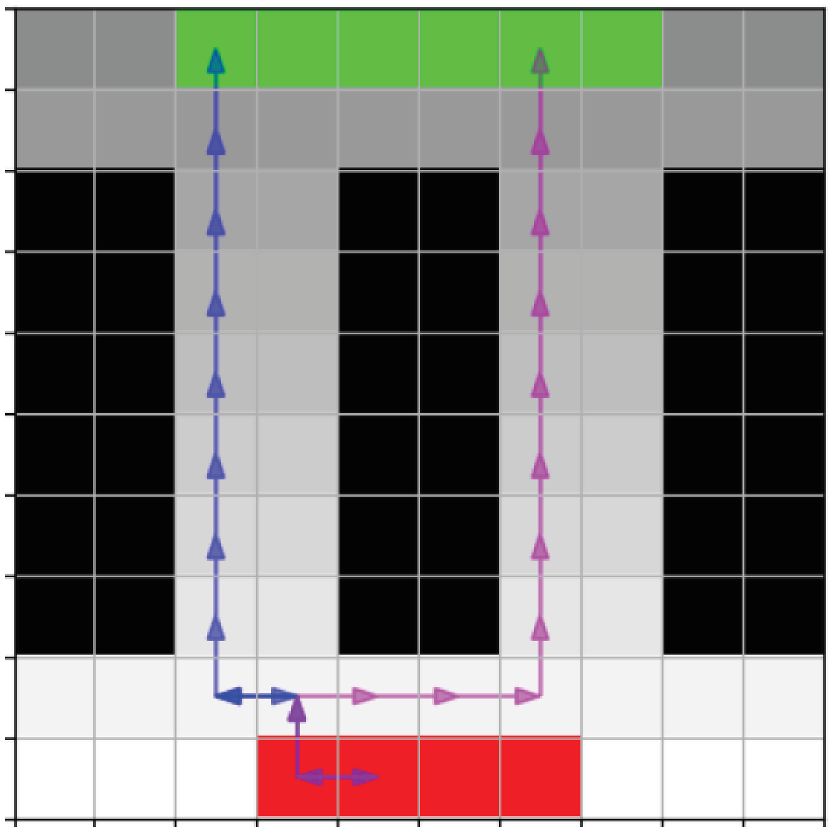
$$\exists \pi \forall \pi' . (\pi \text{ is derived by disturbing } \pi') \\ \wedge (\pi \text{ and } \pi' \text{ reach goal}).$$

$$\exists \pi_1 \forall \pi_2 . \ \mathrm{cls}_{\mathbf{s_0}}(\pi_1, \pi_2) \wedge \mathrm{cls}_{\mathtt{A}}(\pi_1, \pi_2) \Rightarrow (\varphi^{\pi_1} \wedge \varphi^{\pi_2})$$

Symbolic Synthesis from HyperLTL [ICRA'20*]



Shortest path

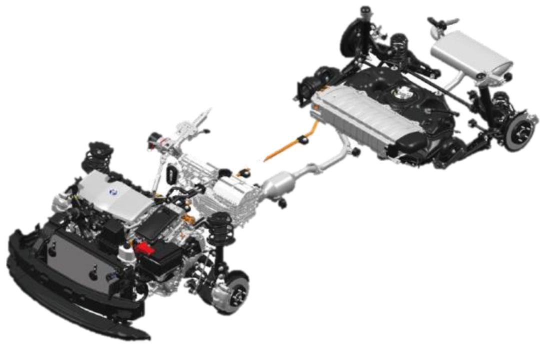


Opacity

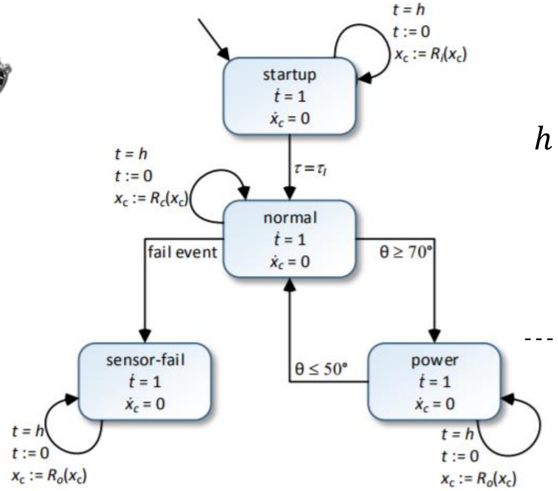
System *Sensitivity* to Modeling Errors

Duke PRATT SCHOOL of ENGINEERING

Toyota Powertrain Benchmark



Embedded Controller

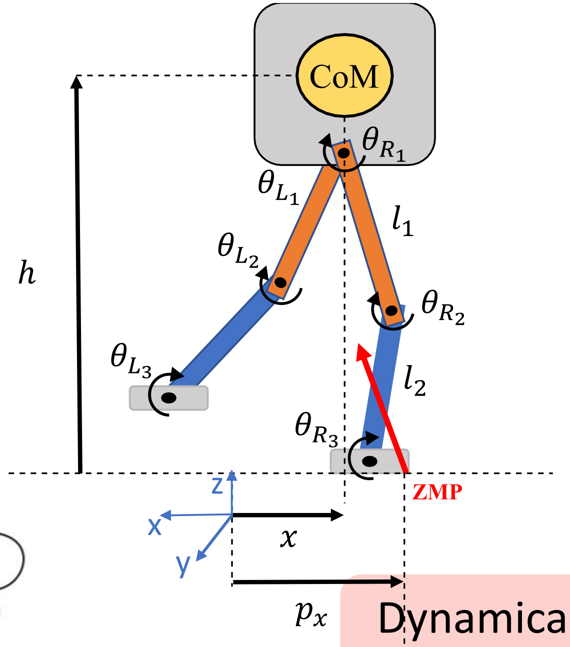


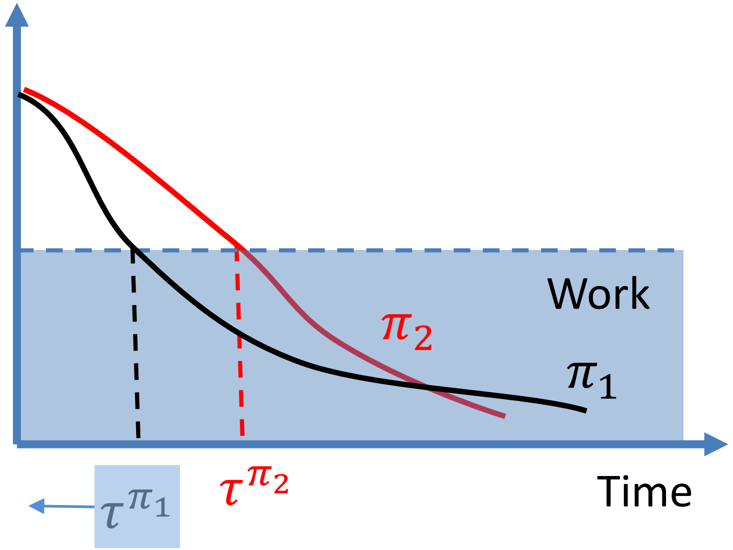
Combustion Process

State	Unit	Description
p	bar	Intake Manifold Pressure
λ_c	-	A/F Ratio in Cylinder
λ_m	-	Transfer Function Output
p_e	bar	Estimated Manifold Pressure
i	-	Integrator State, PI
$\dot{m}_{af}$	g/s	Inlet Air Mass Flow Rate
$\dot{m}_c$	g/s	Air Flow Rate to Cylinder
$\dot{m}_\phi$	g/s	Fuel Mass Aspirated into the Cylinder
$\dot{m}_\psi$	g/s	Fuel Mass Injected into Intake Manifold
θ_{in}	degrees	Throttle Angle Input
θ	degrees	Delay-Filtered Throttle Angle
$\hat{\theta}$	-	O/P of Throttle Polynomial
F_c	g/s	Command fuel
ω	rad/sec	Engine Speed
n	round/sec	Engine Speed ($\frac{\omega}{2\pi}$)

[Jin et.al, HSCC 14]

Walking Robot Benchmark With Reinforcement Learning





Dynamical response depends on system parameters

How does dynamical response change due to modeling errors or *wear-and-tear*?

- For example, start time change under probabilistic uncertainty?

Probabilistic hyperproperties*:* Sensitivity under probabilistic parameter change

$$\mathbf{Pr}_{\pi_1,\pi_2}(|\tau^{\pi_1}-\tau^{\pi_2}|\leq\delta)>1-\varepsilon$$

We need new logic to reason over ***multiple*** random paths!

HyperPSTL: Hyper Probabilistic Signal Temporal Logic

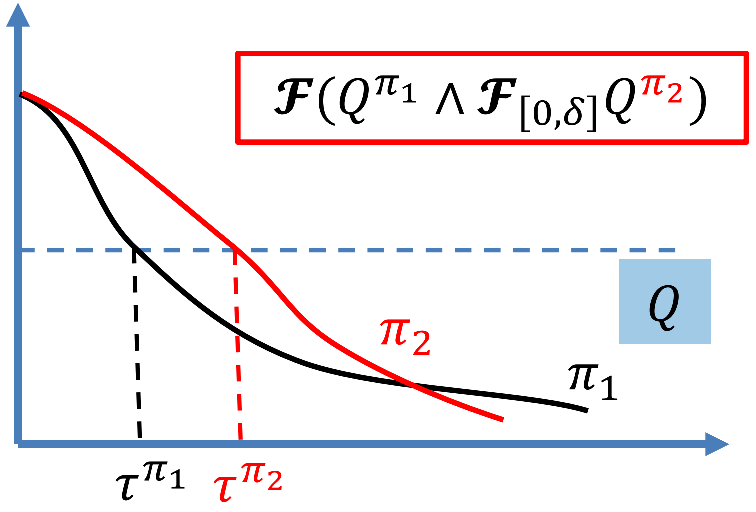
Duke PRATT SCHOOL of ENGINEERING

STL → Add reference to different paths → HyperSTL → Add probabilistic quantifications → HyperPSTL → Add probabilistic arithmetic → (full) HyperPSTL

HyperPSTL: $\varphi ::= \mathrm{a}^{\pi} \mid \varphi^{\pi} \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi\,\mathcal{U}_{[t_1,t_2]}\varphi \mid p \bowtie p$

$$p ::= \mathbb{P}^{\Pi}\varphi \mid \mathbb{P}^{\Pi}p \mid c$$

- a ∈ AP, and AP is the finite set of *atomic propositions*,
- $t_1 < t_2$ with $t_1, t_2 \in \mathbb{Q}_{\infty}$,
- π is a path variable, and Π is a set of path variables,
- $\mathbb{P}$ is the probability operator, $c \in [0,1]$
- $\bowtie \in \{<, >, =, \leq, \geq\}$,
- $\mathrm{fv}(\varphi) = \emptyset$



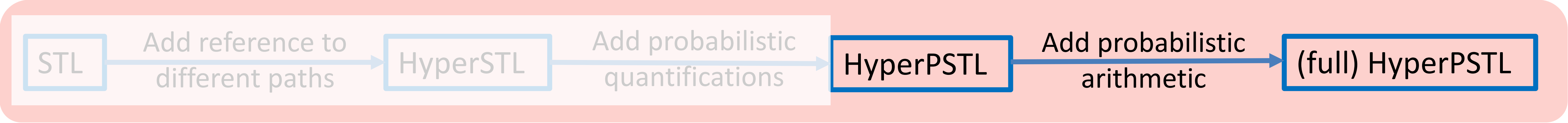
Probabilistic quantifications of multiple parallel paths

$$\mathbb{P}^{(\pi_1,\pi_2)}\left(\mathcal{F}\left(Q^{\pi_1} \wedge \mathcal{F}_{[0,\delta]}Q^{\pi_2}\right)\right) > p$$

Nested probabilistic path quantification

$$\mathbb{P}^{\pi_1}\left(\mathbb{P}^{\pi_2}\left(\mathcal{F}\left(Q^{\pi_1} \wedge \mathcal{F}_{[0,\delta]}Q^{\pi_2}\right)\right) > p_2\right) > p_1$$

HyperPSTL: Hyper Probabilistic Signal Temporal Logic



HyperPSTL: $\varphi ::= \mathrm{a}^{\pi} \mid \varphi^{\pi} \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi\,\mathcal{U}_{[t_1,t_2]}\varphi \mid p \bowtie p$

$$p ::= \mathbb{P}^{\Pi}\varphi \mid \mathbb{P}^{\Pi}p \mid f(p, \dots, p)$$

- a $\in$ AP, and AP is the finite set of *atomic propositions*,
- $t_1 < t_2$ with $t_1, t_2 \in \mathbb{Q}_{\infty}$,
- π is a path variable, and Π is a set of path v
- $\mathbb{P}$ is the probability operator,
- $\bowtie \in \{<, >, =, \leq, \geq\}$,
- $f: \mathbb{R}^n \to \mathbb{R}$ is a n-ary elementary function, constants are viewed as 0-ary functions,
- $\mathrm{fv}(\varphi) = \emptyset$

Kullback-Leibler divergence of two satisfaction probabilities $\mathbb{P}^{\pi_1}\varphi_1^{\pi_1}$ and $\mathbb{P}^{\pi_2}\varphi_2^{\pi_2}$:

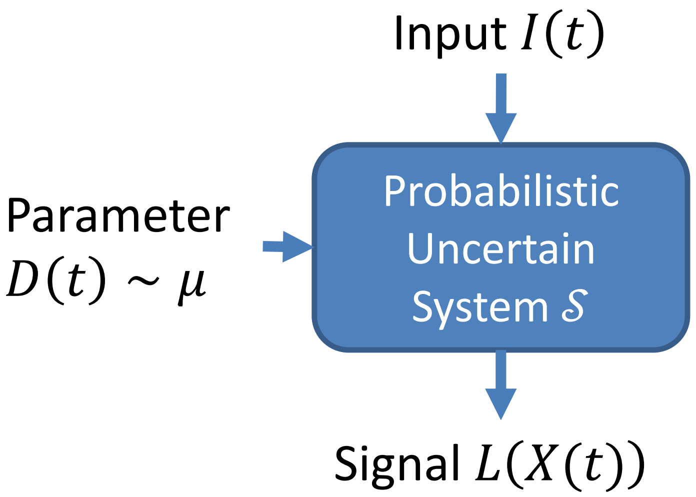
$$\mathbb{P}^{\pi_1}\varphi_1^{\pi_1}\log\left(\frac{\mathbb{P}^{\pi_1}\varphi_1^{\pi_1}}{\mathbb{P}^{\pi_2}\varphi_2^{\pi_2}}\right) + \left(1 - \mathbb{P}^{\pi_1}\varphi_1^{\pi_1}\right)\log\left(\frac{1 - \mathbb{P}^{\pi_1}\varphi_1^{\pi_1}}{1 - \mathbb{P}^{\pi_2}\varphi_2^{\pi_2}}\right) < c$$

Semantics: HyperPSTL on Probabilistic Uncertain System

Duke PRATT SCHOOL of ENGINEERING

Probabilistic uncertain system (PUS): $\mathcal{S} = (\mathcal{X}, I, D, \mu, \mathrm{AP}, L)$ where

- $\mathcal{X}$ is the **state space**, $X^{\mathrm{init}} = (x_1^{\mathrm{init}}, \ldots, x_l^{\mathrm{init}}) \in \mathcal{X}$ is an **initial state**
- **Parameter** $D(t) = (d_1(t), \ldots, d_n(t))$ for $t \in \mathbb{R}_{\geq 0}$ is drawn from probability distribution μ
- **Input** $I(t) = (i_1(t), \ldots, i_m(t))$ is an m-dimensional function of time t
- Given $I(t)$ and $D(t)$, the system generates a **path** $X: \mathbb{R}_{\geq 0} \to \mathcal{X}$ with $X(t) = (x_1(t), \ldots, x_l(t))$
- AP is a set of **atomic propositions**, $\mathrm{L}: \mathcal{X} \to 2^{\mathrm{AP}}$ is a **labeling function**
- a **path** of the system induces a **signal** $\sigma(t) = \mathrm{L}\big(X(t)\big): \mathbb{R}_{\geq 0} \to 2^{\mathrm{AP}}$.

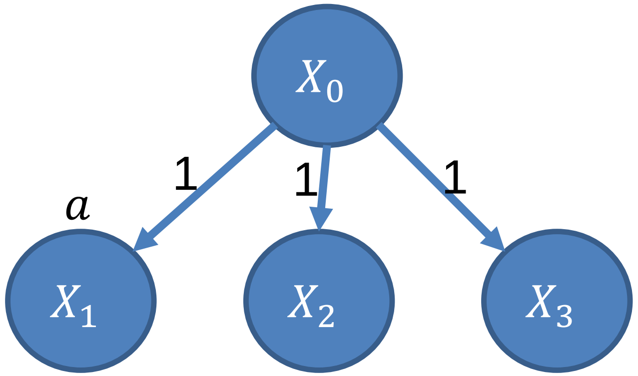


The PUS modeling allows capturing
- Hybrid I/O automata with probabilistic parameters (e.g., powertrain)
- continuous-time Markov chains (CTMCs) as in queueing networks

HyperPSTL: Expressiveness

Theorem: HyperPSTL *strictly subsumes* PSTL (its non-hyper fraction) on CTMCs.

Prof idea: find a CTMC and a property, such that this property only be expressed in HyperPSTL

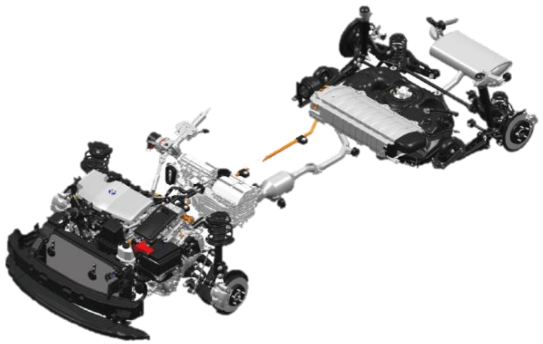


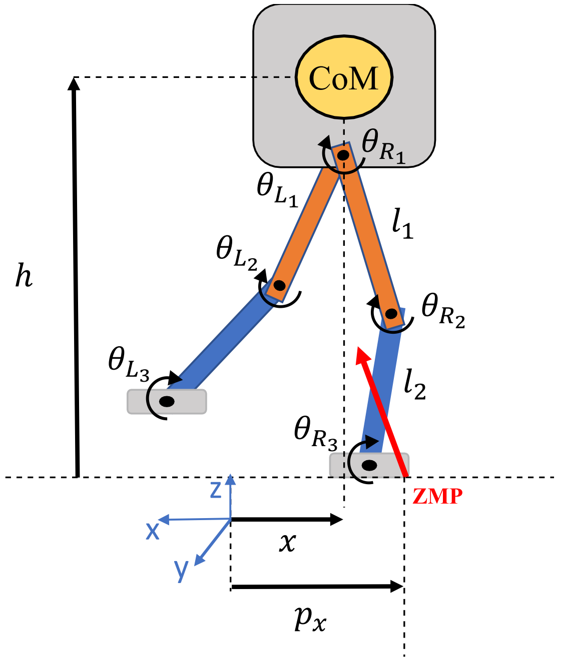
- CTMC has only 3 paths
- Satisfaction probability of any STL is $0, \frac{1}{3}, \frac{2}{3}, 1$, so $P(\varphi) = \frac{1}{9}$ is always false for any (φ)
- HyperPSTL $P^{(\pi_1,\pi_2)}\big(\boldsymbol{\mathcal{F}}(a^{\pi_1} \wedge a^{\pi_2})\big) = \frac{1}{9}$ is true

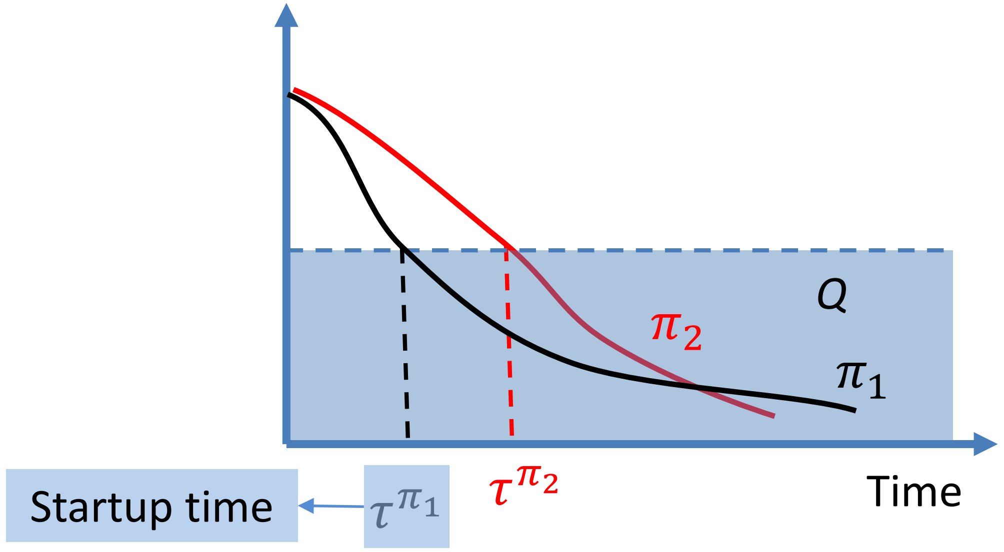
HyperPSTL in Action: Sensitivity to Modeling Errors

Duke PRATT SCHOOL of ENGINEERING

Toyota Powertrain Benchmark







Dynamical response depends on system parameters

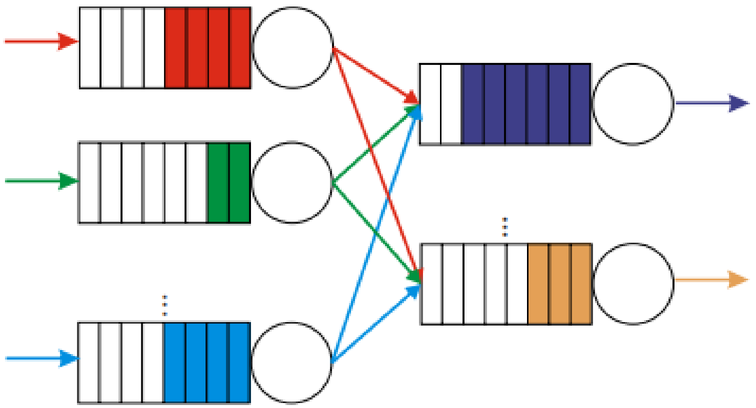
Design specification: Sensitivity of startup time

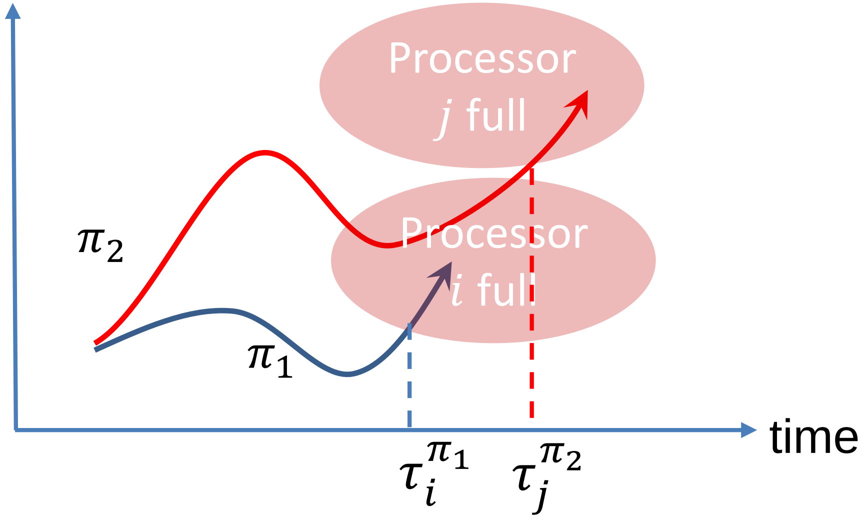
$$\mathbf{Pr}_{\pi_1,\pi_2}(|\tau^{\pi_1} - \tau^{\pi_2}| \leq \delta) > 1 - \varepsilon$$

$$\mathbb{P}^{(\pi_1,\pi_2)}\left(\begin{array}{c} (\neg Q^{\pi_1} \wedge \neg Q^{\pi_2}) \\ \boldsymbol{\mathcal{U}}\left(\left(Q^{\pi_1} \wedge \boldsymbol{\mathcal{F}}_{[0,\delta]} Q^{\pi_2}\right) \vee \left(Q^{\pi_2} \wedge \boldsymbol{\mathcal{F}}_{[0,\delta]} Q^{\pi_1}\right)\right) \end{array} \right) > 1 - \varepsilon$$

HyperPSTL in Action: Workload Fairness

Duke
PRATT SCHOOL of ENGINEERING





Design specification: Workload Fairness

$$\mathbf{Pr}_{\pi_1}\left(\left|\mathbf{Pr}_{\pi_2}\left(\tau_i^{\pi_1}-\tau_j^{\pi_2}>\mathrm{t}\right)-\mathbf{Pr}_{\pi_2}\left(\tau_i^{\pi_1}-\tau_j^{\pi_2}>\mathrm{t}\right)\right|<\delta\right)>1-\varepsilon$$

For any fixed $\tau_i^{\pi_1}$, the probability difference between $\boldsymbol{\tau_i^{\pi_1}-\tau_j^{\pi_2}>\mathbf{t}}$ and $\boldsymbol{\tau_i^{\pi_1}-\tau_j^{\pi_2}<-\mathbf{t}}$ should be less than δ

This should hold with probability more than $1-\varepsilon$ for π_1

$$\mathbb{P}^{\pi_1}(|\mathbb{P}^{\pi_2}((\neg Q_i^{\pi_1}\wedge\neg Q_j^{\pi_2})\mathcal{U}(Q_i^{\pi_1}\wedge\Diamond_{[\tau,\infty)}Q_j^{\pi_2}))-\mathbb{P}^{\pi_2}((\neg Q_i^{\pi_1}\wedge\neg Q_j^{\pi_2})\mathcal{U}(Q_j^{\pi_2}\wedge\Diamond_{[\tau,\infty)}Q_i^{\pi_1}))|\leq\delta)\geq 1-\varepsilon.$$

HyperPSTL in Action: Probabilistic Detectability

Captured independently of the type of used ***sound*** detector as probabilistic ***overshoot observability*** on system outputs, when input overshoot captures that an anomaly has occurred

- Let x be the input and y be the output. After a "step" event, the output signal should be different if the input (1) stays bounded or (2) overshoots.

$$\mathbb{P}^{\{\pi,\pi'\}}((\square(step^{\pi} \Rightarrow \square_I(x^{\pi} < c)) \wedge \diamond(step^{\pi'} \wedge \diamond_I(x^{\pi'} > c))) \Rightarrow (\diamond_I\, d\,(y^{\pi}, y^{\pi'}) > c')) > 1 - \varepsilon$$

SMC of HyperPSTL: Overview

Hyper features beyond existing methods for Statistical Model Checking (SMC)

- Probabilistic quantifications of ***multiple*** parallel paths (e.g., sensitivity) $\mathbb{P}^{(\pi_1,\pi_2)}\varphi^{(\pi_1,\pi_2)} < p$
- ***Nested*** probabilistic path quantification (e.g., fairness) $\mathbb{P}^{\pi_1}\left(\mathbb{P}^{\pi_2}\varphi^{(\pi_1,\pi_2)} < p_2\right) < p_1$
- ***Joint*** probabilities (e,g., KL-divergence) $(\mathbb{P}^{\Pi_1}\varphi_1, \mathbb{P}^{\Pi_2}\varphi_2) \in D$

$$\mathbb{P}^{\pi_1}\varphi_1^{\pi_1}\log\left(\frac{\mathbb{P}^{\pi_1}\varphi_1^{\pi_1}}{\mathbb{P}^{\pi_2}\varphi_2^{\pi_2}}\right) + (1 - \mathbb{P}^{\pi_1}\varphi_1^{\pi_1})\log\left(\frac{1 - \mathbb{P}^{\pi_1}\varphi_1^{\pi_1}}{1 - \mathbb{P}^{\pi_2}\varphi_2^{\pi_2}}\right) < c$$

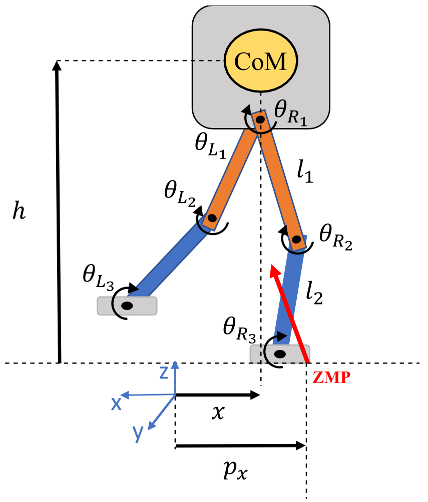
⇓

$$\left(\mathbb{P}^{\pi_1}\varphi_1^{\pi_1}, \mathbb{P}^{\pi_2}\varphi_2^{\pi_2}\right) \in D \text{ with } D = \left\{(x_1, x_2) \mid x_1\log\left(\frac{x_1}{x_2}\right) + (1 - x_1)\log\left(\frac{1-x_1}{1-x_2}\right) < c\right\}$$

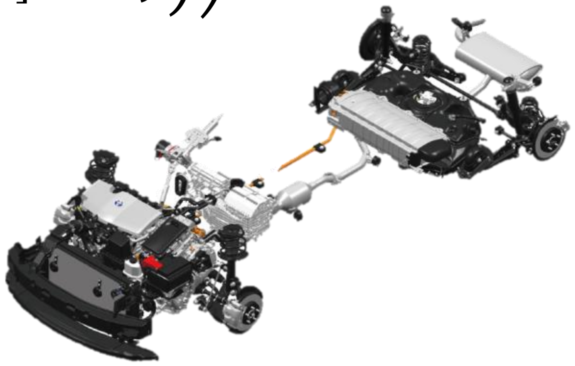
HyperSMC tool for SMC of HyperPSTL with desired **confidence/significance** level: https://gitlab.oit.duke.edu/cpsl/hypersmc

Sensitivity Verification of Real-World CPS

Duke PRATT SCHOOL of ENGINEERING



$$\mathbb{P}^{(\pi_1,\pi_2)}\left(\mathcal{U}\left(\left(Q^{\pi_1}\wedge\mathcal{F}_{[0,\delta]}Q^{\pi_2}\right)\vee\left(Q^{\pi_2}\wedge\mathcal{F}_{[0,\delta]}Q^{\pi_1}\right)\right)^{(\neg Q^{\pi_1}\wedge\neg Q^{\pi_2})}\right) > 1-\varepsilon$$



Walking Robot Benchmark With Reinforcement Learning Controller

δ	ε	α	Acc.	Sam.	Time (s)	Ans.
2.4	0.02	0.01	**1.00**	7.4e+01	3.0e-01	False
2.4	0.02	0.05	**0.99**	4.4e+01	1.4e-01	False
2.4	0.12	0.01	**1.00**	4.2e+01	1.2e-01	True
2.4	0.12	0.05	**1.00**	2.1e+01	7.0e-02	True
2.4	0.2	0.01	**1.00**	1.3e+01	4.0e-02	True
3.0	0.02	0.01	**1.00**	1.1e+01	2.4e-02	False
3.0	0.02	0.05	**1.00**	6.5e+00	1.1e-02	False
3.0	0.12	0.05	**0.98**	7.0e+01	4.3e-01	False
3.0	0.2	0.01	**1.00**	1.6e+02	5.5e-01	True
3.0	0.2	0.05	**0.98**	1.0e+02	2.9e-01	True

Toyota Powertrain Benchmark

δ	ε	α	Acc.	Sam.	Time (s)	Ans.
0.15	0.95	0.05	**1.00**	5.9e+01	8.1e+00	True
0.15	0.95	0.01	**1.00**	9.0e+01	1.3e+01	True
0.15	0.99	0.05	**0.99**	6.6e+01	9.1e+00	False
0.15	0.99	0.01	**1.00**	9.7e+01	1.4e+01	False
0.20	0.95	0.05	**0.98**	5.9e+01	8.1e+00	True
0.20	0.95	0.01	**1.00**	9.0e+01	1.2e+01	True
0.20	0.99	0.05	**1.00**	3.0e+02	4.2e+01	True
0.20	0.99	0.01	**0.99**	4.6e+02	1.8e+02	True

[EMSOFT'19, TACAS'20*]

- On ***continuous-time probabilistic*** models (e.g., powertrain, queueing network), how to capture ***properties between many paths*** (sensitivity, fairness, attack detectability)?

Hyper Probabilistic Signal Temporal Logics: ***HyperPSTL***

- How to reason about HyperPSTL on ***complex*** systems?

Statistical Model Checking (***SMC***) of HyperPSTL

- How does the SMC work in practice?

Evaluation on ***real-world CPS***

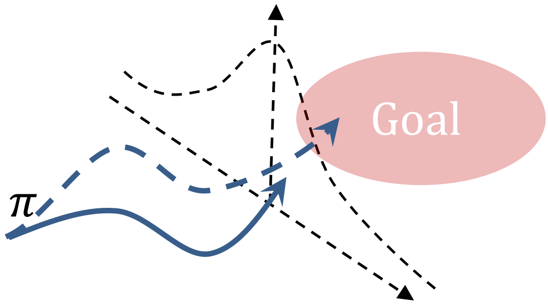
Current Work:

- Application to conformance testing
- Synthesis by reinforcement learning

Thank you

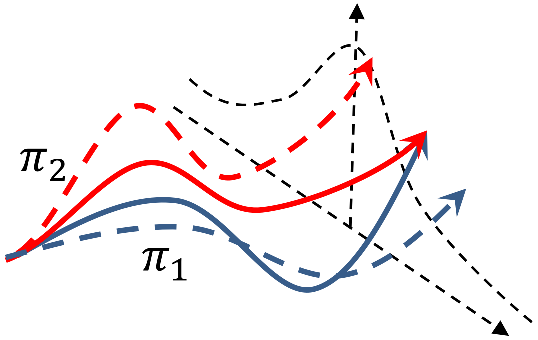
What is a probabilistic hyperproperty?

Probabilistic hyperproperty reasons over **multiple** random paths.



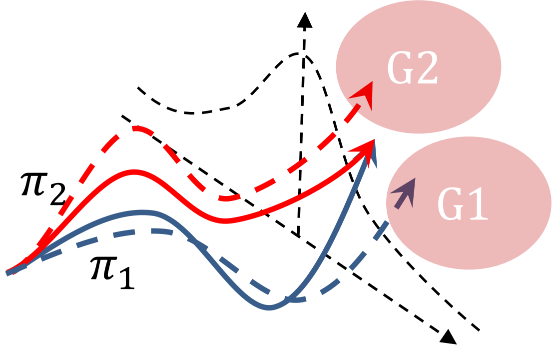
Probabilistic Property:

- Reachability
  $\mathbf{Pr}\big(\pi \vDash \boldsymbol{\mathcal{F}}(\text{Goal})\big) > 0.99$



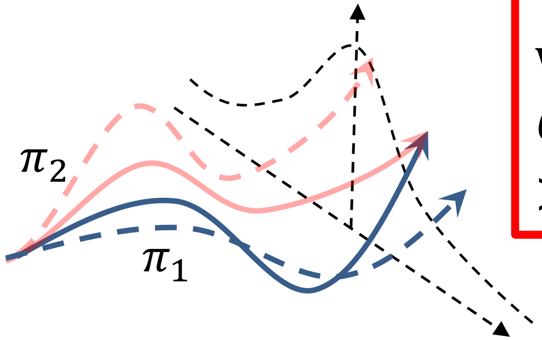
Probabilistic Hyperproperty:

- Two path meet
  $\mathbf{Pr}\big((\pi_1, \pi_2) \vDash \boldsymbol{\mathcal{F}}(\pi_1 = \pi_2)\big) > 0.99$



Probabilistic Hyperproperty:

- Compare satisfaction probabilities
  $\mathbf{Pr}\big(\pi \vDash \boldsymbol{\mathcal{F}}(\text{G1})\big) > \mathbf{Pr}\big(\pi \vDash \boldsymbol{\mathcal{F}}(\text{G2})\big)$



- One catchup another
  $\mathbf{Pr}(\pi_1 \vDash \boldsymbol{C}) > 0.5$

where

$\boldsymbol{C}: \mathbf{Pr}\big(\pi_2 \vDash \boldsymbol{\mathcal{F}}(\pi_1 = \pi_2)\big) > 0.99$