Parameter Invariant Monitoring for Signal Temporal Logic

Nima Roohi, Ramneet Kaur, James Weimer, Oleg Sokolsky, Insup Lee

Manasvi Saxena

Motivations

Challenges in monitoring Real Time Systems -

- Partially Observable States.
- Partially Observable Traces.
- High Computation Cost.

Motivations

Challenges in monitoring Real Time Systems -

- Partially Observable States.
- Partially Observable Traces.
- High Computation Cost.

Extends Parameter Invariant (PAIN) tests to -

- STL to support continuous systems
- Efficiently monitor STL online.

Monitoring a Diabetic Patient

STL formula, "If patient has a meal (M), then (s)he either recieved bolus t_1 units ago, or will recieve one in t_2 units"

$$\square_{[t_1,\infty)}\left((M>c_1)\to \lozenge_{(-t_1,t_2)}(B>c_2)\right)$$

.

Monitoring a Diabetic Patient

STL formula, "If patient has a meal (M), then (s)he either recieved bolus t_1 units ago, or will recieve one in t_2 units"

$$\square_{[t_1,\infty)}\left((M>c_1)\to \lozenge_{(-t_1,t_2)}(B>c_2)\right)$$

- State variables not directly Observable.
- Output variables affected by sensor and environment noise.
- Models are parametric. Estimating (nuissance) parameters is not feasible.

Monitoring a Diabetic Patient

STL formula, "If patient has a meal (M), then (s)he either recieved bolus t_1 units ago, or will recieve one in t_2 units"

 B (Bolus) is observable, but M (Meal) is not. Furthermore, B can be affected by noise.

Monitoring a Diabetic Patient

STL formula, "If patient has a meal (M), then (s)he either recieved bolus t_1 units ago, or will recieve one in t_2 units"

- B (Bolus) is observable, but M (Meal) is not. Furthermore, B can be affected by noise.
- Avoid relying on exact models of the system.
- Rely on probabilistic beliefs about system state at a given point in time.

Monitoring a Diabetic Patient

STL formula, "If patient has a meal (M), then (s)he either recieved bolus t_1 units ago, or will recieve one in t_2 units"

- B (Bolus) is observable, but M (Meal) is not. Furthermore, B can be affected by noise.
- Avoid relying on exact models of the system.
- Rely on probabilistic beliefs about system state at a given point in time.
- Furthermore, use maximally invariant statistics to derive probabilistic beliefs

Signal Temporal Logic

Signals

Given a finite set of variables Z, signal $f: \mathbb{R}_{\geq 0} \to \mathbb{R}^{Z}$ is a mapping from a point in *time* to a *valuation* of Z.

Signal Temporal Logic

Signals

Given a finite set of variables Z, signal $f : \mathbb{R}_{\geq 0} \to \mathbb{R}^{Z}$ is a mapping from a point in *time* to a *valuation* of Z.

Formulas over Signals

- An formula $\theta: \mathbb{R}^{\mathsf{Z}} \to \mathbb{R}$ maps a valuation to a real value.
- Function $\theta \circ f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ maps points in time to values defined by θ .
- At time $t \in \mathbb{R}_{\geq 0}$, signal f is true w.r.t. θ whenever $\theta(f(t)) > 0$. Conversely, f is false w.r.t θ w.r.t. θ whenever $\theta(f(t)) < 0$.
- Whenever, $\theta(f(t)) = 0$, value of f is *unknown* on θ at time t.
- $\theta(f(t))$, is the *robustness degree*. Larger *robustness degree* values signify greater *belief* about signal's adherence to θ .

Signal Temporal Logic

Syntax

- $\varphi ::= \top \mid \bot \mid \Theta \mid \neg \Theta \mid \varphi \land \varphi \mid \varphi \ \mathcal{U}_{\mathcal{I}_{>0}} \ \varphi \mid \varphi \mathcal{R}_{\mathcal{I}_{>0}} \varphi$
- $\Diamond_{\mathcal{I}}\varphi \equiv \top \mathcal{U}_{\mathcal{I}} \varphi$
- $\Box_{\mathcal{I}}\varphi \equiv \bot \mathcal{R}_{\mathcal{I}}\varphi$
- Law of Excluded Middle not assumed. Robustness 0 means signal neither satisfies nor not satisfies a formula. It is possible for $\varphi \vee \neg \varphi$ to not be true.
- Law of non contradiction still holds. It is not possible for both φ and $\neg \varphi$ to be true.

Continuous Semantics

Given signal f and a point in time $r : \mathbb{R}_{\geq 0}$, Continuous time STL semantics are given inductively as -

$$f, r \models_{\text{csr}} \top := \infty \qquad f, r \models_{\text{csr}} \bot := -\infty$$

$$f, r \models_{\text{csr}} \theta := \theta(f(r)) \qquad f, r \models_{\text{csr}} \neg \theta := -\theta(f(r))$$

$$f, r \models_{\text{csr}} \phi \lor \psi := f, r \models_{\text{csr}} \phi \sqcup f, r \models_{\text{csr}} \psi$$

$$f, r \models_{\text{csr}} \phi \land \psi := f, r \models_{\text{csr}} \phi \sqcap f, r \models_{\text{csr}} \psi$$



Continuous Semantics

Given signal f and point in time r: $\mathbb{R}_{\geq 0}$, Continuous time STL semantics are given inductively as -

$$egin{aligned} f, r & artriangleq_{ ext{cut}} arphi \ \mathcal{U}_{\mathcal{I}} \ \psi \ := \ igsqcup_{t:r+\mathcal{I}} \left(f, t & artriangle_{ ext{cut}} \ \psi \cap igcap_{r \leq t' < t} \left(f, t' & artriangle_{ ext{cut}} \ arphi
ight)
ight) \ f, r & artriangle_{ ext{cut}} \ arphi \ \mathcal{R}_{\mathcal{I}} \ \psi \ := \ igcap_{t:r+\mathcal{I}} \left(f, t & artriangle_{ ext{cut}} \ \psi \sqcup igcup_{r \leq t' < t} \left(f, t' & artriangle_{ ext{cut}} \ arphi
ight)
ight) \end{aligned}$$

Discrete Semantics

Given signal f and sampling function $\tau: \mathbb{N} \to \mathbb{R}_{\geq 0}$, discrete time semantics for $g := f \circ \tau: \mathbb{N} \to \mathbb{R}^{\mathbb{Z}}$ at step $n: \mathbb{N}$ are given inductively as -

$$egin{align*} g, n & artriangledows & = & \infty & g, n & artriangledows & = & -\infty \ g, n & artriangledows & = & \theta(g(n)) & g, n & artriangledows & = & -\theta(g(n)) \ g, n & artriangledows & = & 0 & 0 & 0 & 0 \ g, n & artriangledows & = & 0 & 0 \ g, n & artriangledows & = & 0 \ g, n & artriangledo$$

Discrete Semantics

Given signal f and sampling function $\tau: \mathbb{N} \to \mathbb{R}_{\geq 0}$, discrete time semantics for $g := f \circ \tau: \mathbb{N} \to \mathbb{R}^{\mathsf{Z}}$ at step $n: \mathbb{N}$ are given inductively as -

$$egin{aligned} egin{aligned} egi$$

Strengthening STL Formulas

For STL formula φ , and $\delta: \mathbb{R}_{\geq 0}$, φ^{δ} (strengthening) is defined as -

•
$$\perp^{\delta} := \perp \quad \top^{\delta} := \top \quad p^{\delta} := p \quad (\neg p)^{\delta} := \neg p$$

•
$$(\varphi \lor \psi)^{\delta} := \varphi^{\delta} \lor \psi^{\delta} \quad (\varphi \land \psi)^{\delta} := \varphi^{\delta} \land \psi^{\delta}$$

•
$$(\varphi \ \mathcal{U}_{\mathcal{I}} \ \psi)^{\delta} := \varphi^{\delta} \ \mathcal{U}_{(\underline{\mathcal{I}} + \delta, \overline{\mathcal{I}} - \delta)} \ \psi^{\delta}$$

•
$$(\varphi \ \mathcal{R}_{\mathcal{I}} \ \psi)^{\delta} := \varphi^{\delta} \ \mathcal{R}_{((\underline{\mathcal{I}} - \delta)^+, \overline{\mathcal{I}} + \delta)} \ \psi^{\delta}$$

Intuitvely, strengthening slightly shortens (lengthens) intervals for $\mathcal{U}(\mathcal{R})$. Along with conditions about the signal and sampling function, strengthening allows using discrete-time STL to reason about continuous-time STL.

Discrete to Continuous Time STL

Given signal f, formula φ , strictly increasing sampling function τ , where $\Delta \tau = \sqcup_{n:\mathbb{N}} (\tau(n+1) - \tau(n))$, and $\delta : \mathbb{R}_{\geq 0}$, where

- $\exists \lambda : \mathbb{R}_+$ such that for $\theta : \Theta$, $\theta \circ f$ is λ -Lipschitz continuous.
- δ is sufficiently large $\lambda \Delta \tau < \delta$.
- Sampling is sufficient $\Delta \tau < \frac{1}{3} \min_{\mathcal{I}: \mathcal{I}_{\varphi}} (\overline{\mathcal{I}} \underline{\mathcal{I}})$, where \mathcal{I}_{φ} are intervals in temporal operators of φ .
- Sampling started early $\tau(0) < \Delta \tau$

Discrete to Continuous Time STL

Given signal f, formula φ , strictly increasing sampling function τ , where $\Delta \tau = \sqcup_{n:\mathbb{N}} (\tau(n+1) - \tau(n))$, and $\delta : \mathbb{R}_{\geq 0}$, where

- $\exists \lambda : \mathbb{R}_+$ such that for $\theta : \Theta$, $\theta \circ f$ is λ -Lipschitz continuous.
- δ is sufficiently large $\lambda \Delta \tau < \delta$.
- Sampling is sufficient $\Delta \tau < \frac{1}{3} \min_{\mathcal{I}: \mathcal{I}_{\varphi}} (\overline{\mathcal{I}} \underline{\mathcal{I}})$, where \mathcal{I}_{φ} are intervals in temporal operators of φ .
- ullet Sampling started early $au(0) < \Delta au$
- guarantees the following hold -
 - $f \circ \tau \models_{\scriptscriptstyle{\mathrm{DSC}}} \varphi^{\Delta \tau} \implies f \models_{\scriptscriptstyle{\mathrm{CNT}}} \varphi > 0.$
 - $f \circ \tau \models_{\scriptscriptstyle{\mathrm{DSC}}} \neg \varphi^{\Delta \tau} \implies f \models_{\scriptscriptstyle{\mathrm{CNT}}} \varphi < 0.$

Instead of proving $f \models_{\text{\tiny CNT}} \varphi > 0$, prove $f, \tau \models_{\text{\tiny DNC}} \varphi^{\Delta \tau} > \delta$.



Partially Observable States

Say we have test functions Θ , signal f and sampling time τ . For θ : Θ , $\theta \circ f$ cannot be directly observed.

- $\mathcal{O}(\theta, f, t) = F\mu + \theta^{+}(f(t))G_{0}\rho_{0} + \theta^{-}(f(t))G_{1}\rho_{1} + \sigma n$
- Y is a set of *observable* variables.
- μ : dom $(\mu) \to \mathbb{R}$, ρ_0 : dom $(\rho_i) \to \mathbb{R}$ for $i \in \{0,1\}$ are unknown nuissance vectors.
- $F : \mathbb{R}^{Y \times \text{dom}(\mu)}$, $G_i : \mathbb{R}^{Y \times \text{dom}(\rho_i)}$ for $i \in \{0,1\}$ are known signal matrices.
- σ is unkown noise multiplier, n is random noise.

Partially Observable States

- It is assumed that for each test function $\theta, \theta' \in \Theta$, $\theta \neq \theta' \implies Y_{\theta} \cup Y_{\theta'} = \emptyset$. In other words, different test functions use disjoint observable variables.
- $\mathcal{O}(f,t)$ random variable corresponding to the observation of signal f at time t is *uniquely determined* by $\mathcal{O}(\theta,f,t)$ for each $\theta \in \Theta$.

Invariance to Transformations

- Given Ω , the sampling space of a probabily-space and $\mathcal{K} \subseteq \Omega \to \Omega$ a group of transformations.
- Let $\eta: \Omega \to \Omega'$ be a statistic.
- ullet η is said to be *invariant* to group of transformations ${\cal K}$ iff

$$\forall \omega : \Omega, k : \mathcal{K}.\eta(\omega) = \eta(k(\omega))$$

ullet η is said to be $extit{maximally invariant}$ to group of transformations $\mathcal K$ iff

$$\forall \omega, \omega' : \Omega.\eta(\omega) = \eta(\omega') \implies (\exists k : \mathcal{K}.\omega = k(\omega'))$$

Parameter Invariant Robustness Estimation of Atomic Propositions

Suppose we have

- $\Omega_{f, au}:\mathbb{N} \to \mathbb{R}^X$, be the space of all observations at different sample time
- A sequence $x : \Omega_{f,\tau}$ test function $\theta \in \Theta$, error and indifference bounds $\alpha', \delta' : \mathbb{R}^+$.

Parameter Invariant Robustness Estimation of Atomic Propositions

Suppose we have

- $\Omega_{f, au}:\mathbb{N} \to \mathbb{R}^X$, be the space of all observations at different sample time
- A sequence $x : \Omega_{f,\tau}$ test function $\theta \in \Theta$, error and indifference bounds $\alpha', \delta' : \mathbb{R}^+$.

One can find a algorithm $\mathcal{A}(x,\tau,\theta,\alpha',\delta')$ such that -

- A is *terminating* and returns value $r : \mathbb{R}$.
- $\mathbb{P}(\mid (\theta \circ f \circ \tau)(0) r \mid > \delta') < \alpha'$
- The *Probability* that r (i.e. the *robustness degree estimate* returned by \mathcal{A}) is more than δ' away from the real value of $f \circ \tau \models_{\text{\tiny Isc}} \theta$ is less than α' .

Parameter Invariant Robustness Estimation of Atomic Propositions

- If *statistic* used in A is *Maximally Invariant*, then r, is invariant to nuissance parameters.
- A change in the noise in sensor measurements will not affect r.
- Use $\mathcal A$ to obtain estimates that are (maximally) invariant to nuissance parameters.

- Using A, we get *robustness estimates* for atomic proposition $\theta \in \Theta$.
- Furthermore, if the estimate in \mathcal{A} is derived using a maximally invariant test statistic, the estimate is invariant to nuissance parameters (like noise).
- For example, robustness estimate will be invariant to sensor noise.

- For given θ , say Y_{θ} it the set of observable variables of θ .
- Let $\mathcal{K}_{\theta,f,\tau(n)} \subseteq \mathbb{R}^{Y_{\theta}} \to \mathbb{R}^{Y_{\theta}}$ be group of transformations affecting θ .
- Let $\mathcal{K}_{f,\tau(n)} \subseteq \mathbb{R}^Y \to \mathbb{R}^Y$ be group of transformations for f at step n.
- For $v \in \mathbb{R}^Y$, $\theta \in \Theta$, function $k \in \mathcal{K}_{f,\tau(n)}$ maps $v \upharpoonright^{Y_{\theta}}$ to some $k_{\theta}(v \upharpoonright^{Y_{\theta}})$ where $k_{\theta} \in \mathcal{K}_{\theta,f,\tau(n)}$.
- $\mathcal{K}_{f,\tau(n)}$ is completely determined by $\mathcal{K}_{\theta,f,\tau(n)}$ since $\theta \neq \theta' \implies Y_{\theta} \cap Y_{\theta'} = \emptyset$

- Let $\eta_{\theta}: \mathbb{R}^{Y_{\theta}} \to \mathbb{R}$ be statistics that produces *estimates* for $\theta \in \Theta$.
- Define $\eta: \mathbb{R}^Y \to \mathbb{R}$ such that for $v \in \mathbb{R}^Y$, $\eta(v) = \eta_{\theta}(v \upharpoonright^{Y_{\theta}})$ for $\theta \in \Theta$.
- Suppose η_{θ} is invariant to $\mathcal{K}_{\theta,f,\tau(n)}$ for each $\theta \in \Theta$.

- Let $\eta_{\theta}: \mathbb{R}^{Y_{\theta}} \to \mathbb{R}$ be statistics that produces *estimates* for $\theta \in \Theta$.
- Define $\eta: \mathbb{R}^Y \to \mathbb{R}$ such that for $v \in \mathbb{R}^Y$, $\eta(v) = \eta_{\theta}(v \upharpoonright^{Y_{\theta}})$ for $\theta \in \Theta$.
- Suppose η_{θ} is invariant to $\mathcal{K}_{\theta,f,\tau(n)}$ for each $\theta \in \Theta$.
- Formally, for $v : \mathbb{R}^Y$ and $k_{\theta} \in \mathcal{K}_{\theta,f,\tau(n)}$, $\eta_{\theta}(v | Y_{\theta}) = \eta_{\theta}(k_{\theta}(v | Y_{\theta}))$
- For $k \in \mathcal{K}_{f,\tau(n)}$, $k(v) = k_{\theta}(v)^{Y_{\theta}}$.
- $\eta(k(v)) = \eta_{\theta}(k(v)) = \eta_{\theta}(k_{\theta}(v)^{\mathbb{R}^{Y_{\theta}}}))$, for each $\theta \in \Theta$.
- We get $\eta(k(v)) = \eta_{\theta}(k_{\theta}(v)^{Y_{\theta}}) = \eta_{\theta}(v)^{Y_{\theta}}$ for each $\theta \in \Theta$.
- ullet Thus, η is invariant to $\mathcal{K}_{f, au(n)}$

- We proved η is invariant to K_{f,τ(n)}. Moreover, if η_θ is maximally invariant, to K_{f,τ(n)}, then η is also maximally invariant.
 Thus, using maximally invariant estimates for atomic propositions, we
- I hus, using maximally invariant estimates for atomic propositions, we get a maximally invariant estimate for the entire signal at a point in time.
- Recall $\mathcal A$ gives us an *estimate* for $\theta \in \Theta$. If $\mathcal A$ uses a *maximally invariant* statistic, we get a *maximally invariant* estimate for each atomic proposition in an STL formula.

Assumptions

Given a signal f, sampling-time function τ , STL formula φ , such that τ is strictly increasing. $g = f \circ \tau$ is discretized signal

- ullet ${\cal M}$ is an algorithm for monitoring ${\it g}$ against arphi and -
 - needs a complete mapping of all atomic propositions in φ to robustness degrees.
 - takes input a pair (v, t), where v is a complete mapping of atomic propositions to robustness and t is time.
 - return interval [a, b] the inf and sup respectively of all possible robustness degrees if the observed sequence was arbitrarily extended.
- Main Algorithm (following slides) join \mathcal{A} , \mathcal{M} to have a Parameter Invariant Monitor for STL.

- (Inputs) signal f, formula φ Lipschitz constant λ , indifference region δ and an error parameter α .
- (Output) accept, reject or unknown
- Has 3 functions -
 - MAIN Initializes the algorithm, runs main loop, and checks termination conditions.
 - **ESTIMATE** Uses $\mathcal A$ to estimate values for each atomic proposition of φ .
 - OBSERVE Auxilliary function used by estimate to obtain estimates for atomic propositions.

Main Function

```
x \leftarrow \text{nil}
n \leftarrow 0
function MAIN
      \mathcal{M}_1 \leftarrow \text{initialize } \mathcal{M} \text{ for } \varphi^{\Delta \tau}
      \mathcal{M}_2 \leftarrow \text{Initialize } \mathcal{M} \text{ for } \neg (\varphi)^{\Delta \tau}
       [a_1,b_1] \leftarrow (-\infty,\infty)
       [a_2,b_2] \leftarrow (-\infty,\infty)
      while a_1 \leq \delta \wedge a_2 \leq \delta \wedge (\delta < b_1 \wedge \delta < b_2) do
              (t, v) \leftarrow \text{ESTIMATE}
              [a_1,b_1] \leftarrow \mathcal{M}_1(v,t)
              [a_2,b_2] \leftarrow \mathcal{M}_2(v,t)
              n \leftarrow n + 1
if a_1 > \delta then return accept
if a_2 > \delta then return reject
return unknown
```

▷ Sequence of observations▷ Current Step

Estimation Function

ESTIMATE uses \mathcal{A} to calculate estimates for each atomic proposition θ .

```
function ESTIMATE
     v: R^{\Theta}
     for all \theta \in \Theta do
          a \leftarrow \mathcal{A}(\mathsf{OBSERVE}, \theta, \alpha', \delta')
          if a - \delta' > 0 then
                v(\theta) \leftarrow a - \delta'
          else if a + \delta' < 0 then
                v(\theta) \leftarrow a + \delta'
          else
                v(\theta) \leftarrow 0
     // Recall x is global list containing observations
     (t, \_) \leftarrow pop(x)
     return (t, v)
```

Observation Function

```
function OBSERVE (n : \mathbb{N})

while |x| \le n do

v : \mathbb{R}^Y

t \leftarrow current time

for all \theta \in \Theta do

v \upharpoonright^{Y_{\theta}} \leftarrow Sensor reading for \theta

x \leftarrow x + (t, v)

(t, v) \leftarrow \mathbf{n}^{\text{th}} element of x

return (t, v) \upharpoonright^{Y_{\theta}}
```

 OBSERVE allows ESTIMATE to observe multiple sensor readings at a given point in time.

Algorithm Execution Flow

- MAIN initializes \mathcal{M} for $\varphi^{\Delta \tau}$ and its negation.
- \mathcal{M} is stateful. After initialization, every call to \mathcal{M} with observation-time pair (v,t) makes \mathcal{M} record pair.
- ESTIMATE returns robustness degrees for each atomic proposition, using sensor readings.
- If A uses a maximally invariant statistic to calculate robustness estimate, then estimate is invariant to noise.

Monitoring a Diabetic Patient

STL formula, "If patient has a meal (M), then (s)he either recieved bolus t_1 units ago, or will recieve one in t_2 units"

$$\square_{[t_1,\infty)}\left((M>c_1)\to \lozenge_{(-t_1,t_2)}(B>c_2)\right)$$

Monitoring a Diabetic Patient

STL formula, "If patient has a meal (M), then (s)he either recieved bolus t_1 units ago, or will recieve one in t_2 units"

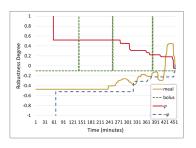
$$\square_{[t_1,\infty)}\left((M>c_1)\to \lozenge_{(-t_1,t_2)}\left(B>c_2\right)\right)$$

. Formula above not strictly STL (uses past time). Equivalently-

$$\square_{[t_1,\infty)}\left(\lozenge_{(0,t_1+t_2)}\left(B>c_2\right)\vee\square_{[t_1,t_1+t_2)}\left(\left(M>c_1\right)\to\lozenge_{(0,t_2)}\left(\left(B>c_2\right)\right)\right)\right)$$

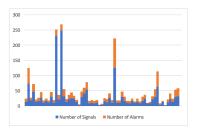
"At all points $r : \mathbb{R}_{\geq 0}$, either patient recieved Bolus or if meal is taken, then Bolus should be taken no later than t_2 units"

Monitoring a Diabetic Patient



- $t_1 = t_2 = 30 (minutes)$
- Plot shows b_1 for φ and a_2 for $\neg \varphi$
- Meal occurs at t=421, but no bolus detected in $391 \le t \le 421$. Thus, when t slightly above 451 formula is rejected, as $a_2 > \delta$.
- Noisy B (bolus) signal, unobservable M (meal) variable

Monitoring a Diabetic Patient



- Data for 61 patients, each monitored for different time lengths.
- Plot shows alarms vs number of observations for each patient.
- Evidence technique scales well for real world applications

Conclusions

- Introduce Parameteric Invariant Monitoring for Signal Temporal Logic
- Using *robustness estimates* for single atomic propositions, derive *robustness estimates* for entire formula.
- If robustness estimates for atomic propositions is maximally invariant then test procedure for entire formula is also maximally invariant to nuissance paramters.
- Demonstrate real-world applications on monitoring an STL formula with patients having T1-Diabetes.