Cyber-Physical Systems

Laura Nenzi

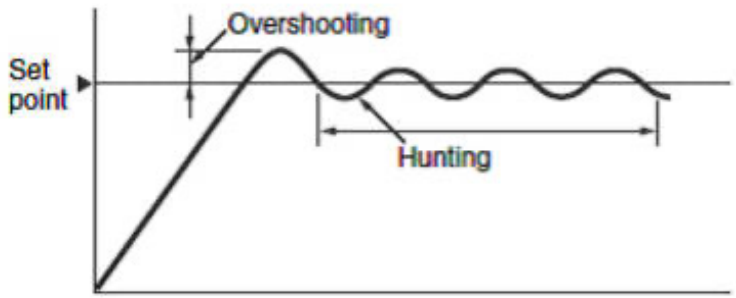
Università degli Studi di Trieste
I Semestre 2023

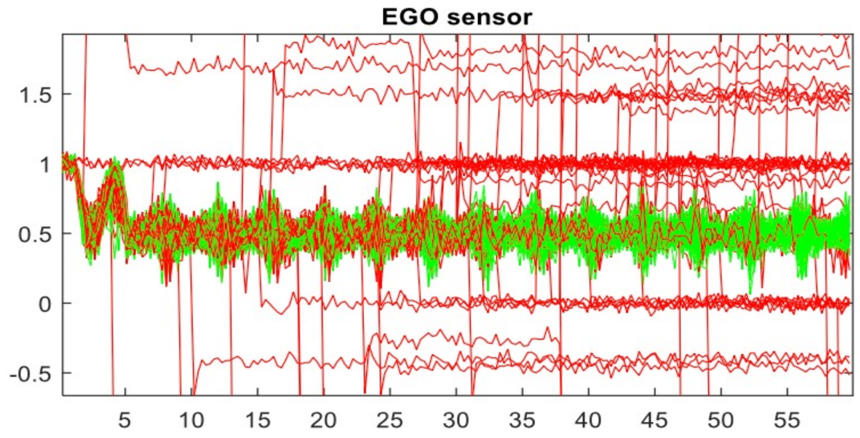
Lecture 21: STL learning

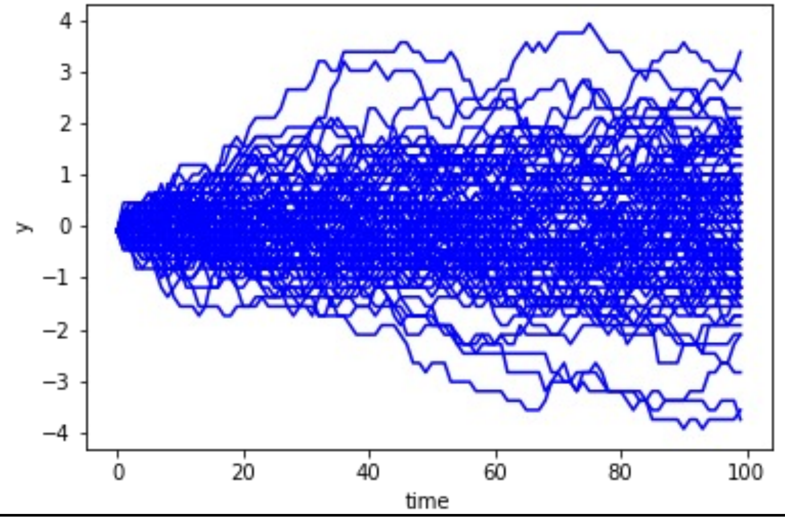
[Many Slides due to J. Deshmukh, S. Silvetti]

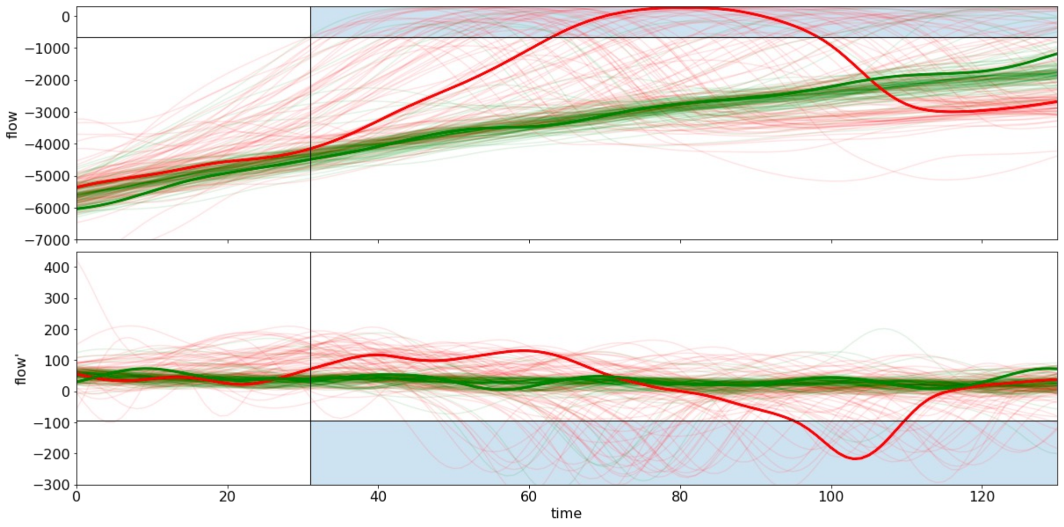
Context and Problem

Data are trajectories

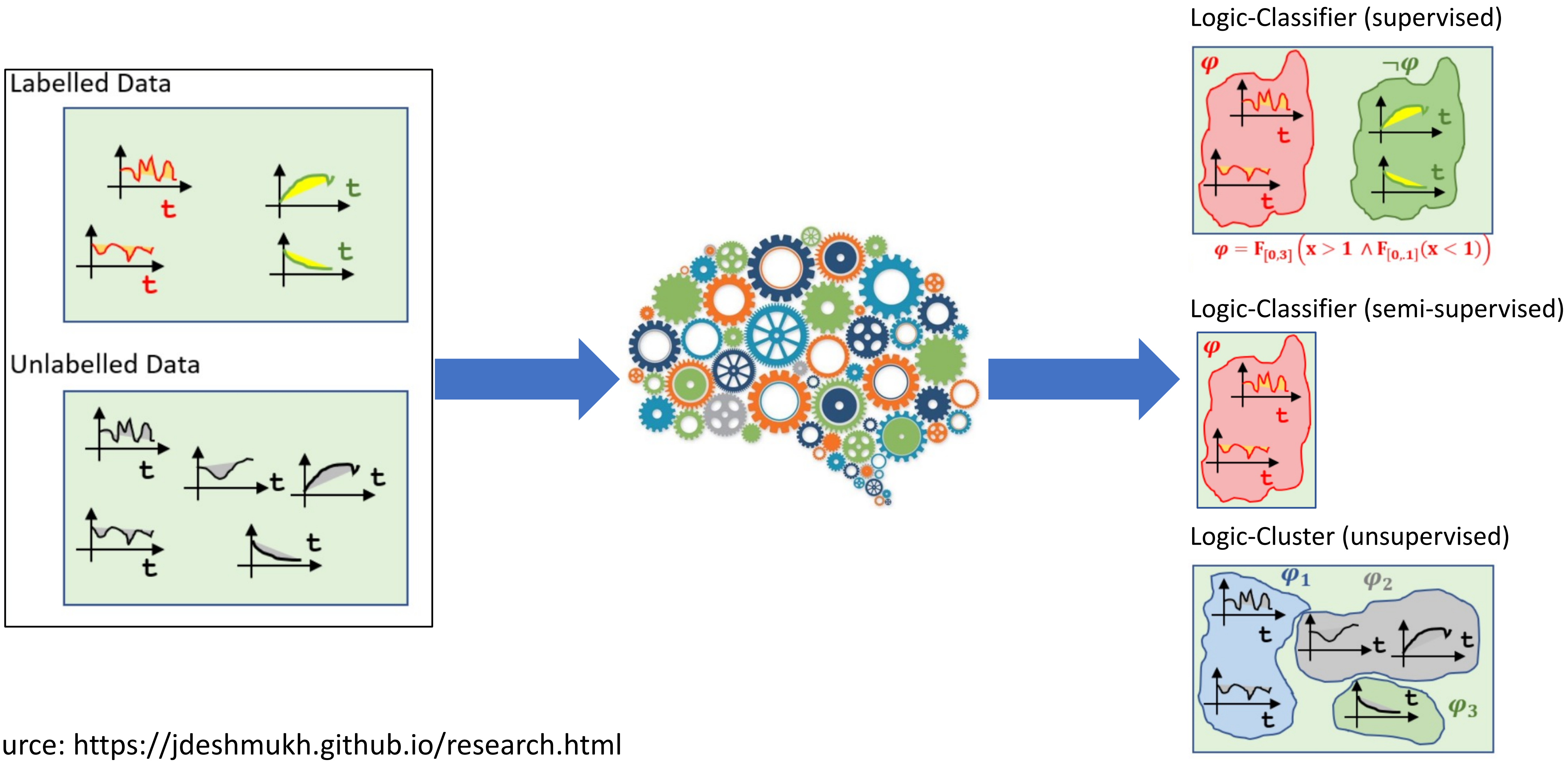








Temporal Logic requirement mining



Source: https://jdeshmukh.github.io/research.html

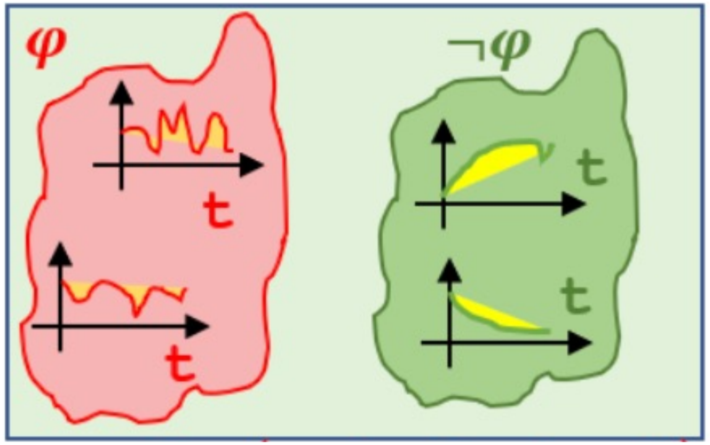
STL Classifier: Problem Statement

We want a way to search in the space of STL formulae considering training data X_{learn}

Supervised two-class classification problem

Training data set: two sets

- regular X^+_{learn}
- anomalous X^-_{learn}

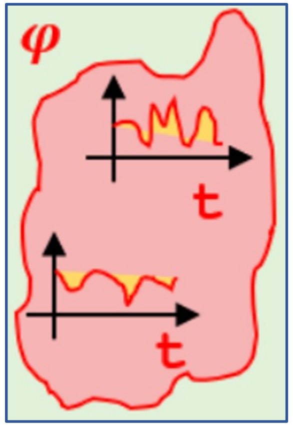


Find the best ϕ that better separates the two sets.

Semi-supervised one-class classification prob

Training data set: one set

- regular X^+_{learn}



Find the "tight" ϕ that is satisfied by the set

Parametric Signal Temporal Logic

Definition (PSTL syntax)

$$\phi := (x_i \bowtie \pi) \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \, \mathcal{U}_{[\tau_1, \tau_2]} \, \varphi_2$$

with $\bowtie \in \{>, \leq\}$

- π is **threshold** parameter
- τ_1, τ_2 are **temporal** parameters

- $\mathbb{K} = (\mathcal{T} \times \mathcal{C})$ be the **parameter space**
- $\boldsymbol{\theta} \in \mathbb{K}$ is a **parameter configuration**

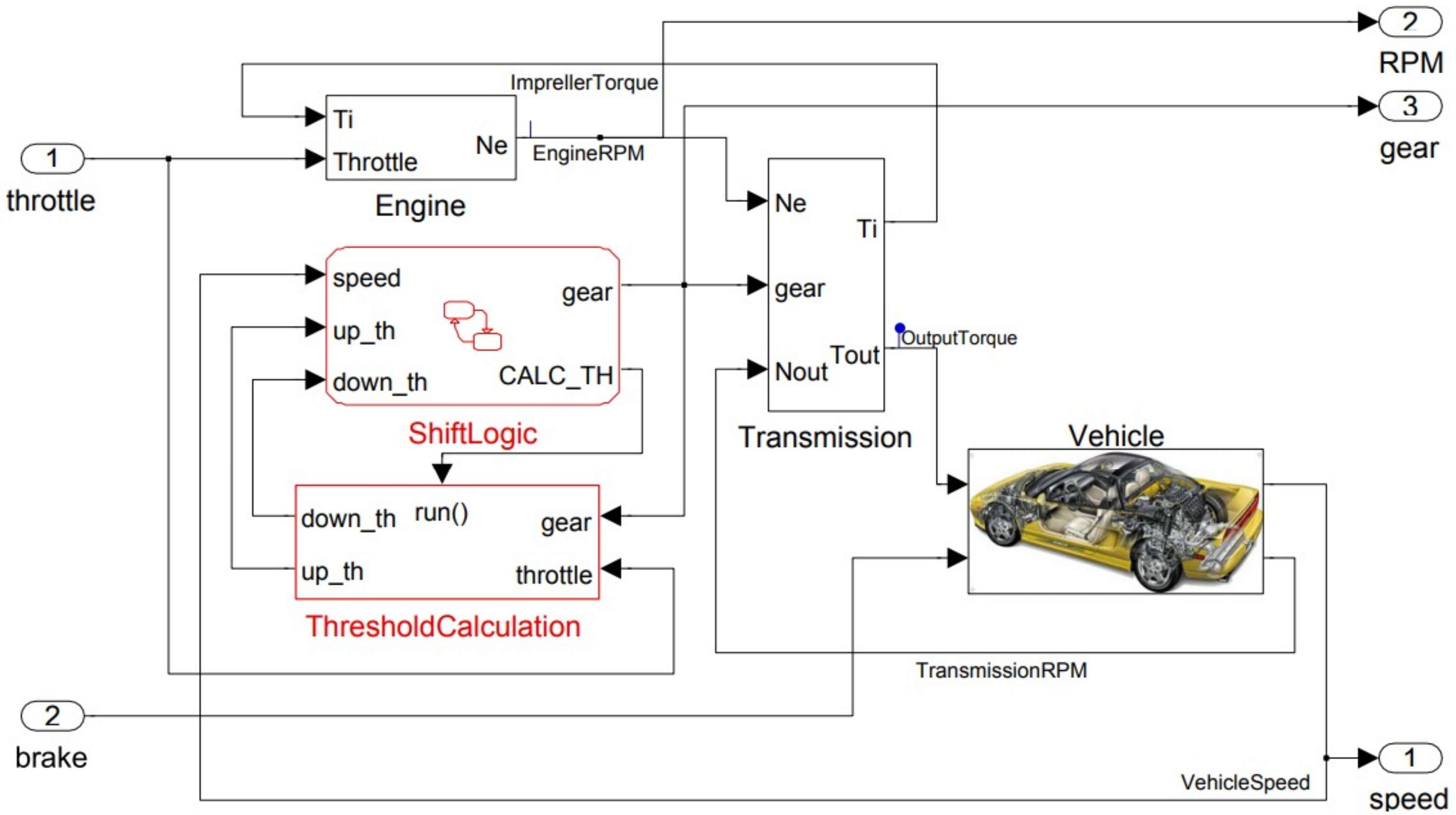
e.g., $\phi = \mathcal{F}_{[a,b]}(x_i > k), \boldsymbol{\theta} = (0, 2, 3.5)$ then $\phi_{\boldsymbol{\theta}} = \mathcal{F}_{[0,2]}(x_i > 3.5)$.

Specification Mining

- Specification Mining: Try to find values of parameters of a PSTL formula from a given model

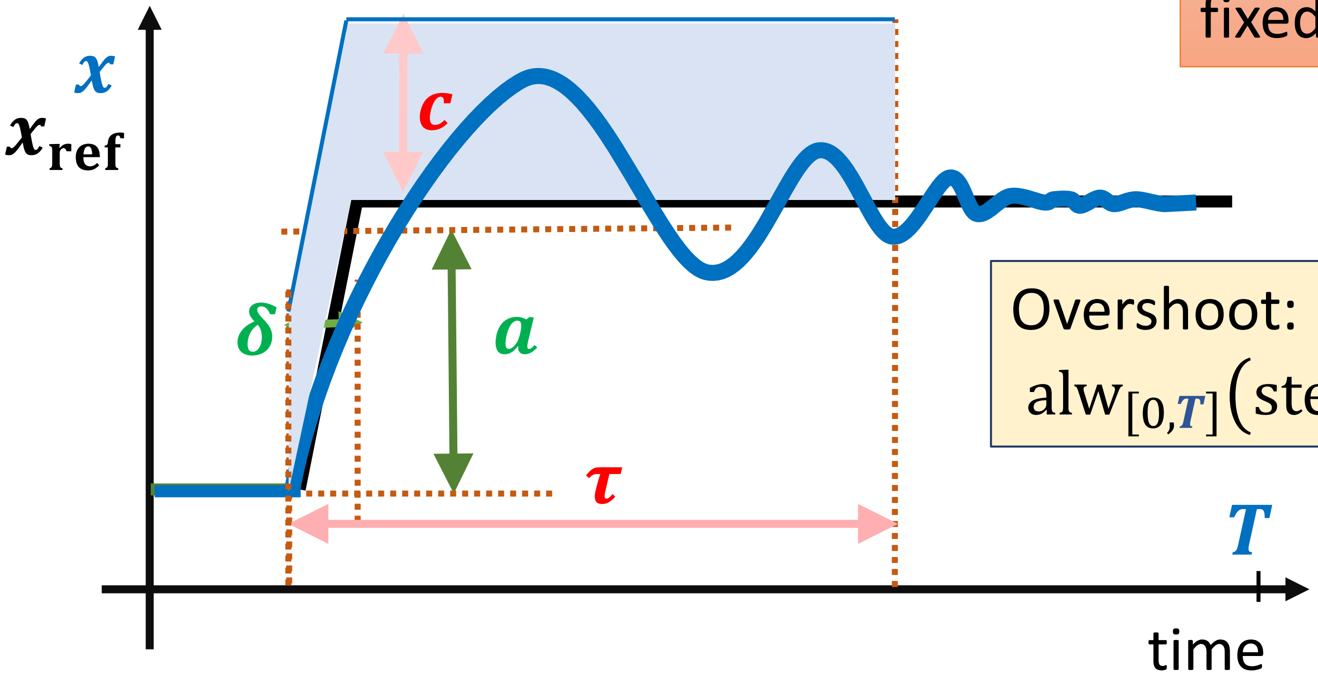
- Why?
  - Good to know “as-is” properties of the model
  - Finds worst-case behaviors of the model
  - Discriminates between regular and anomalous behaviours

Specification Mining



- What is the maximum speed that the vehicle can reach ?
- What is the minimum d well time in a given gear ?

Specification Templates using PSTL



In previous lecture, a, c, T, τ, δ were some fixed values, here they represent parameters

Step:

$$\text{step}(y) \coloneqq y(t+\delta) - y(t) > a$$

Overshoot:

$$\text{alw}_{[0,T]}\big(\text{step}(x_{\text{ref}}) \Rightarrow \text{alw}_{[0,\tau]}(x(t) - x_{\text{ref}}(t) < c)\big)$$

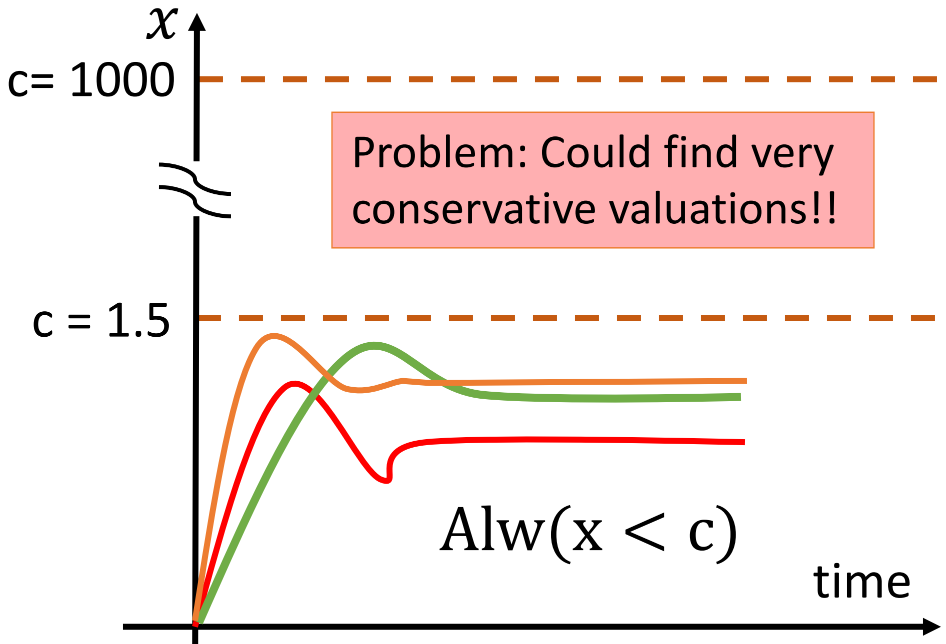
Parameter inference for PSTL

- Given:
  - PSTL formula $\varphi(\mathbf{p})$, $[\mathbf{p} = (p_1, p_2, \dots, p_m)]$
  - Traces $x_1, \dots, x_n$
- Find:
  - ~~Valuation~~ $\nu(\mathbf{p})$ such that: $\forall i : x_i \vDash \varphi(\nu(\mathbf{p}))$

    δ-tight valuation

  - and $\exists i: x_i \nvDash \varphi(\nu(\mathbf{p}) \pm \delta)$:

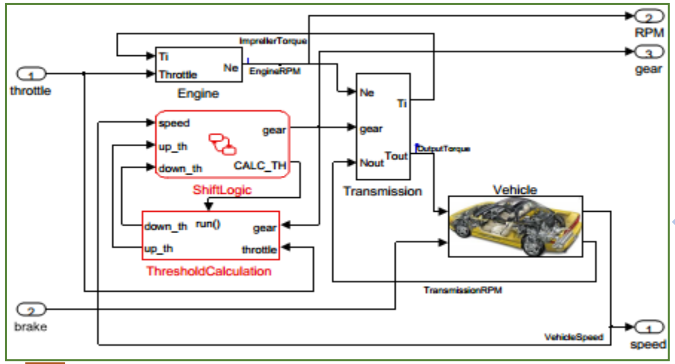
    i.e. small perturbation in $\nu(\mathbf{p})$ makes some trace not satisfy formula



formula sat for given valuation $\Rightarrow$ $\forall$ greater (or lesser) valuations sat

Finding δ-tight valuations hard in general, but **efficient** for **Monotonic PSTL**

Binary search on parameter space



Falsification:
∃ trace ⊭ Property?

1.

1.

YES

NO

Specificatio
n Mining

Secret Sauce:

- Infer parameters for a given PSTL formula from traces
- Falsify given STL formula

Parar

Find "Tightest" Answers

Settling Time is 6 ms
Overshoot is 5.5 KPa
Upper Bound on x is 5

Settling Time is **??**
Overshoot is **??**
Bounds on *x* are **??**

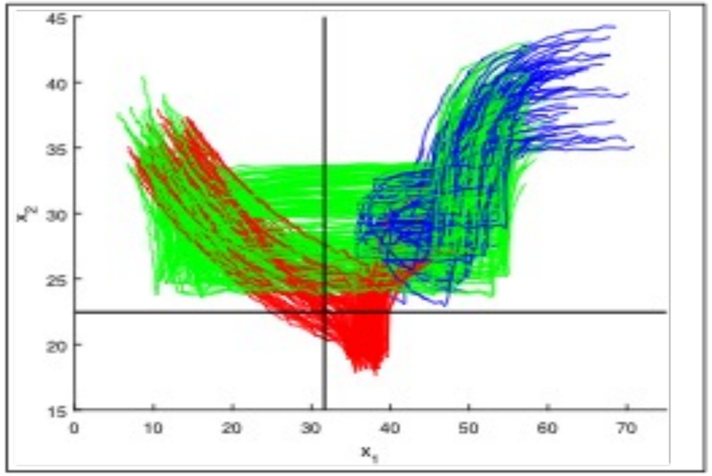
Settling Time is 6 ms
Overshoot is 5.5 KPa
Upper Bound on x is 5

Learning STL classifiers

Goal: Given sets of good and bad trajectories (or generative models), learn STL properties that can separate the two behaviours (a STL classifier)

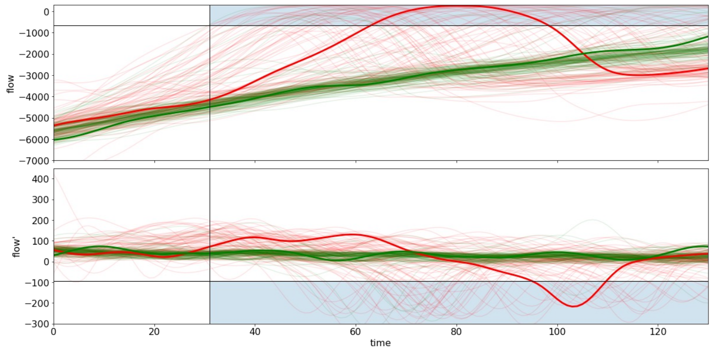
Idea: for a fixed template formula, learn optimally separating parameters by **Bayesian Optimisation**.

$\varphi = ((x_2 > 22.46)\,\mathcal{U}_{[49,287]}\,(x_1 \leq 31.65))$



Maritime surveillance

$\phi = \mathcal{F}_{[31,130]}((flow \geq -670) \vee (flow' \leq -94))$



Light entrainment of biological oscillator

Idea: explore formula structure by **genetic programming** on syntactic trees

Problem Formulation

A supervised two-class classification problem

Given a training set of D_p(good) and D_n(bad) find the best ϕ that better separates the two sets.

Discrimination Function

$$G(\phi) = \frac{\mathbb{E}(R_\phi|\vec{X}_p) - \mathbb{E}(R_\phi|\vec{X}_n)}{\sigma(R_\phi|\vec{X}_p) + \sigma(R_\phi|\vec{X}_n)}.$$

Observation: only statistical and noisy evaluations of $G(\phi)$

Goal: maximize $G(\phi)$

ROGE – RObustness GEnetic Algorithm

It is a bi-level optimization algorithm. A GEnetic algorithm to learn the structure and a Bayesian optimization algorithm to learn the parameters.

Require: $\mathcal{D}_p, \mathcal{D}_n, \mathbb{K}, Ne, Ng, \alpha, s$

1: $gen \leftarrow$ GENERATEINITIALFORMULAE(Ne, s)
2: $gen_\Theta \leftarrow$ LEARNINGPARAMETERS$(gen, G, \mathbb{K})$
3: **for** $i = 1 \ldots Ng$ **do**
4: $\quad subg_\Theta \leftarrow$ SAMPLE(gen_Θ, F)
5: $\quad newg \leftarrow$ EVOLVE$(subg_\Theta, \alpha)$
6: $\quad newg_\Theta \leftarrow$ LEARNINGPARAMETERS$(newg, G, \mathbb{K})$
7: $\quad gen_\Theta \leftarrow$ SAMPLE$(newg_\Theta \cup gen_\Theta, F)$
8: **end for**
9: **return** gen_Θ

$$\phi_{best} = \operatorname{argmax}_{\phi_\theta \in gen_\Theta}(G(\phi_\theta))$$

Learning the Parameters

<u>Problem</u>

Given a PSTL formula ϕ, a parameter space $\mathbb{K}$, find Θ^* that maximises the discrimination function $G(\phi_\Theta)$.

<u>Methodology</u>

1. Sample $\{(\theta_{(i)},y_{(i)}),\ i = 1,\dots,n\}$
2. Emulate (**GP Regression**): $G[R_\phi] \sim GP(\mu,k)$
3. Optimize the emulation via **GP-UCB algorithm**, new $\theta_{(n+1)}$

$\exists\delta$ s.t. $\mathbb{E}(R_{\phi_\Theta *}|\vec{X}_p) > \delta$ and $\mathbb{E}(R_{\phi_\Theta *}|\vec{X}_n) \leq \delta$

Translation. $(\vec{x} - \delta) \Rightarrow \mathbb{E}(R_{\phi_\Theta *}|\vec{X}_p) > 0$ and $\mathbb{E}(R_{\phi_\Theta *}|\vec{X}_n) \leq 0$

Learning the Structure

Problem

Given a set of PSTL formulas *gen*, find the best ϕ such that ϕ_Θ maximises the discrimination function $G(\phi_\Theta)$.

Methodology

1. GENERATEINITIALFORMULAE: $gen=\{\phi_1,\cdots,\phi_{N_e}\}$
2. SAMPLE$(gen_\Theta, F)=subg_\Theta$, $N_e/2$ formulae, $F(\phi)=G(\phi)-S(\phi)$
3. EVOLVE$(subg_\Theta, \alpha) = newg_\Theta$, based on two genetic operators, a **recombination** and a **mutation** operator.

Regularization

Formula size penalty $S(\phi)$ and complexity of initial population.

Learning the Structure

Problem

Given a set of PSTL formulas *gen*, find the best ϕ such that ϕ_Θ maximises the discrimination function $G(\phi_\Theta)$.

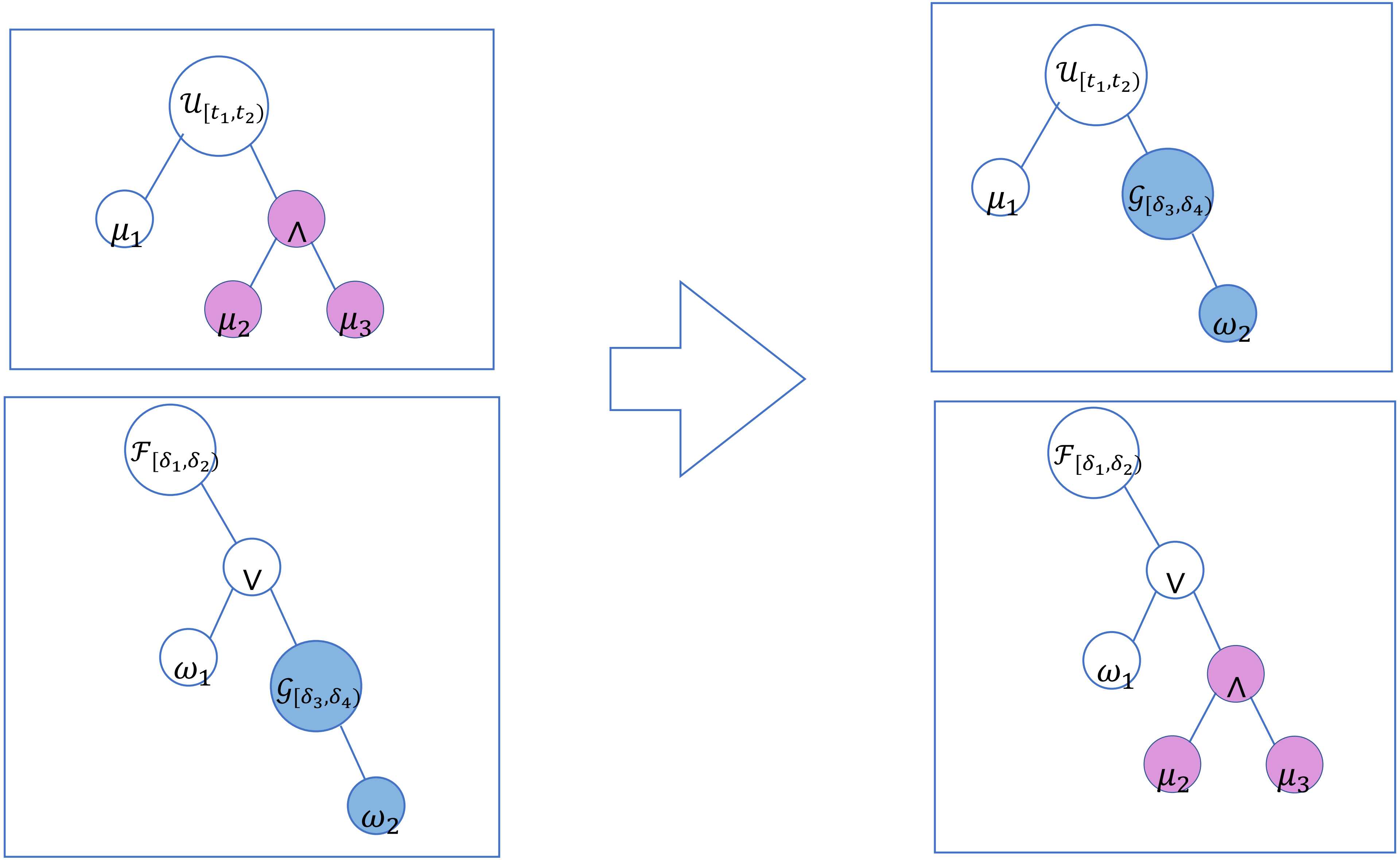
Methodology

1. GENERATEINITIALFORMULAE: $gen=\{\phi_1,\cdots,\phi_{N_e}\}$
2. SAMPLE$(gen_\Theta, F)=subg_\Theta$, $N_e/2$ formulae, $F(\phi)=G(\phi)-S(\phi)$
3. EVOLVE$(subg_\Theta, \alpha) = newg_\Theta$, based on two genetic operators, a **recombination** and a **mutation** operator.

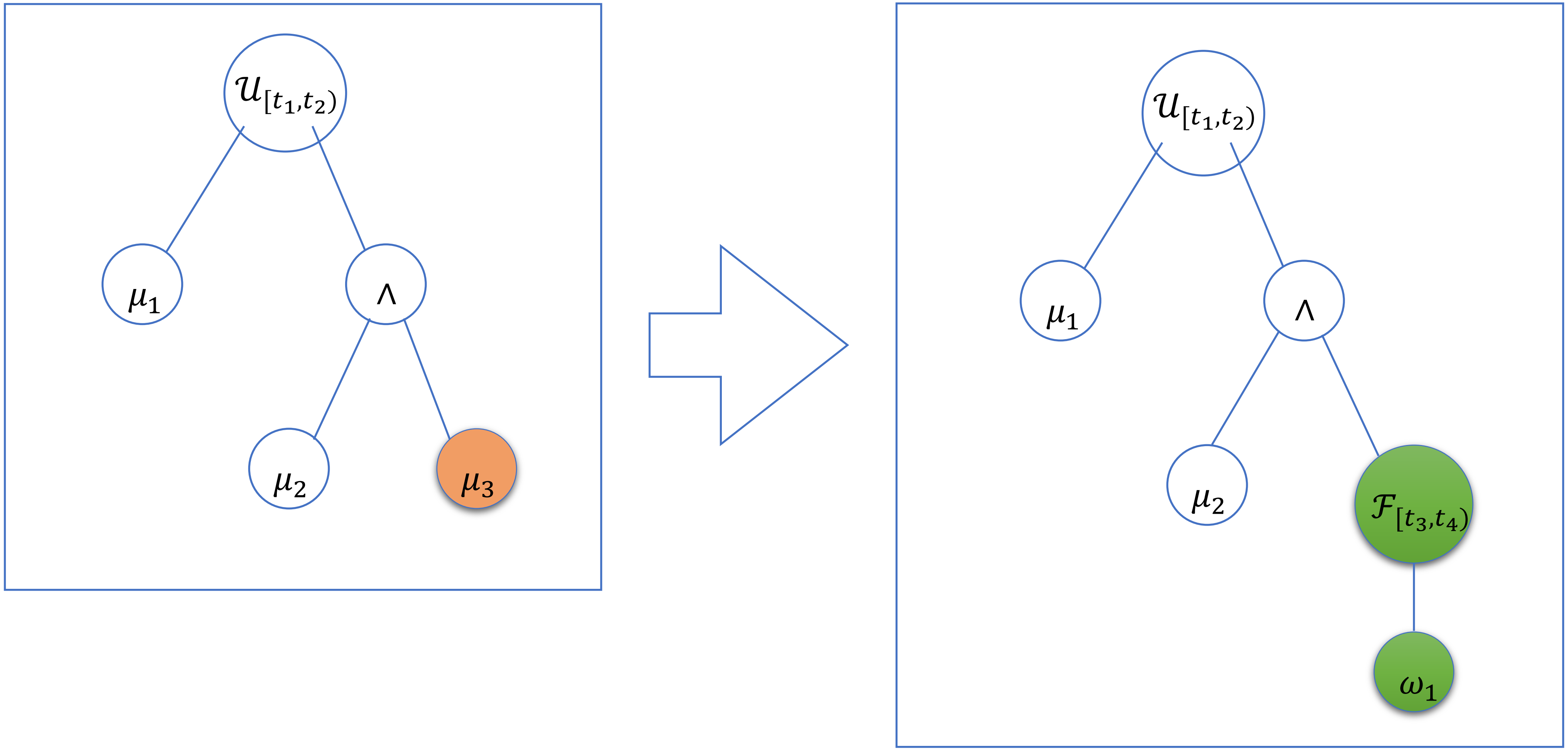
Regularization

Formula size penalty $S(\phi)$ and complexity of initial population.

Recombination operator

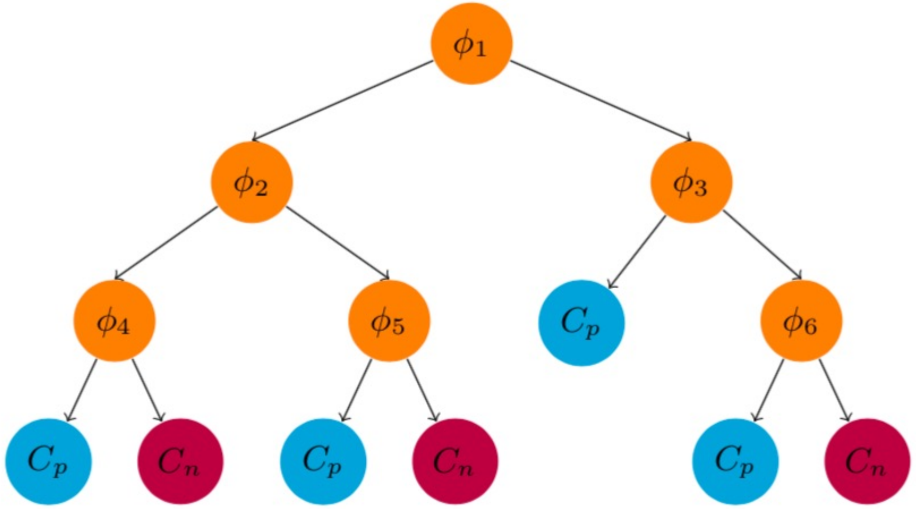


Mutation operator



A Decision Tree Approach to Data Classification

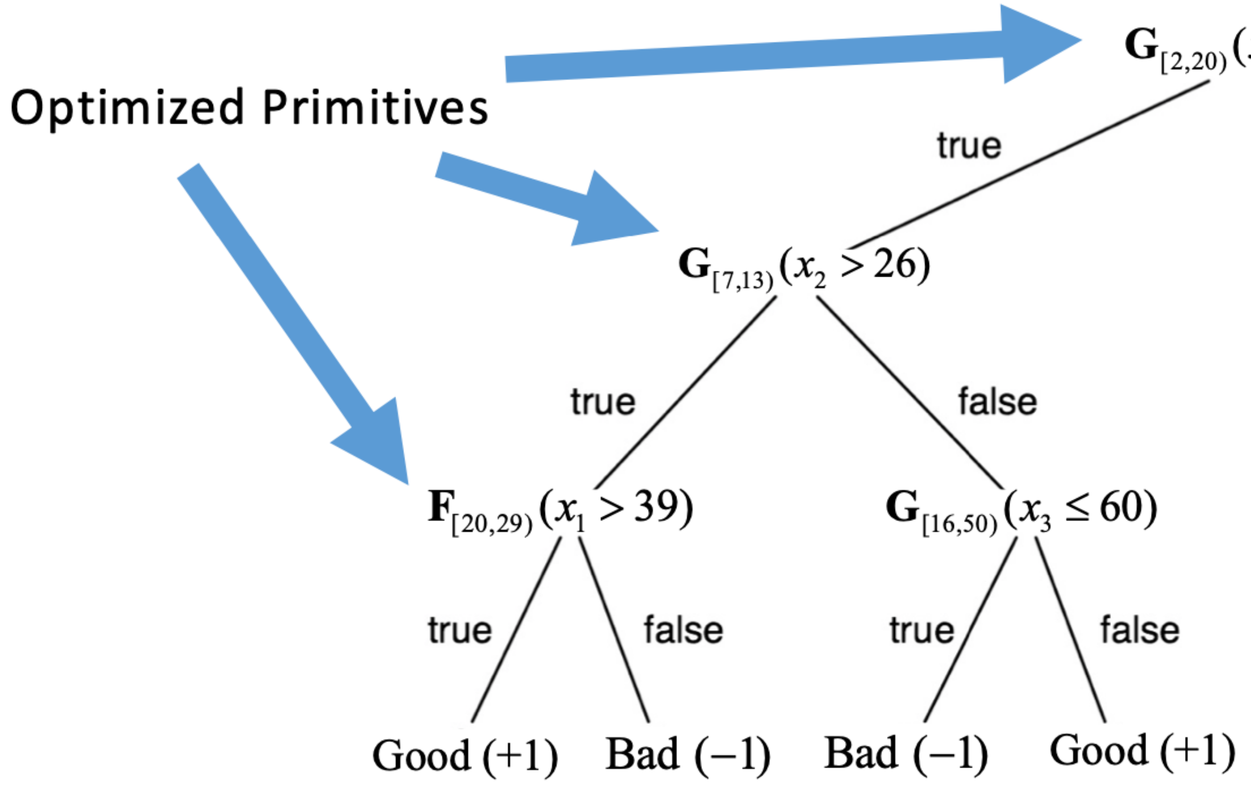
DTL4STL [2], that uses a decision tree algorithm for the structure of the formula and an heuristic impurity measure for parameter synthesis



$$\phi_{Tree} = (\phi_1 \wedge ((\phi_2 \wedge \phi_4) \vee (\neg\phi_2 \wedge \phi_5))) \vee (\neg\phi_1 \wedge (\phi_3 \vee (\neg\phi_3 \wedge \phi_6)))$$

[2] Bombara, G et all, A Decision Tree Approach to Data Classification Using Signal Temporal Logic. In: Proc. of HSCC, 2016.

A Decision Tree Approach to Data Classification



Construction Process

S = set of traces reaching this node

P = set of primitives

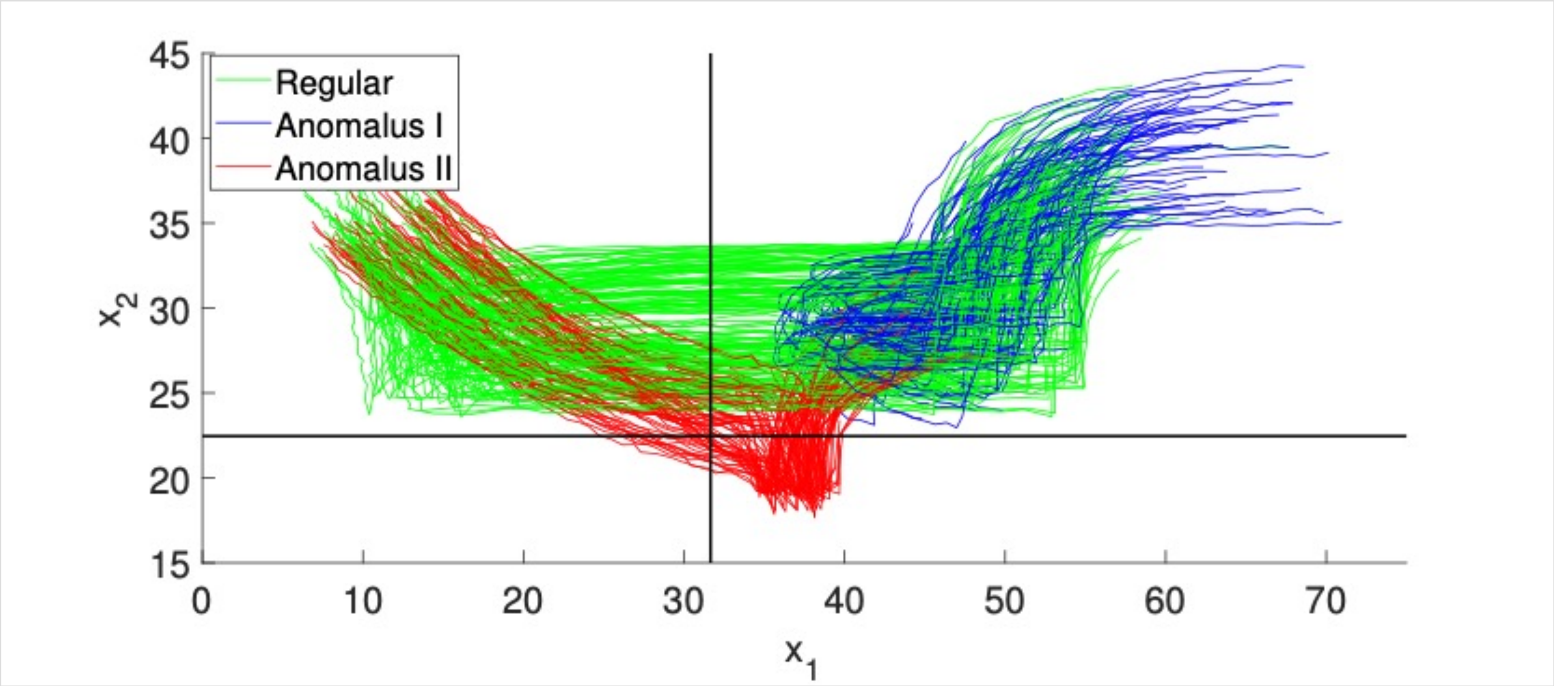
IR = Impurity Reduction Measure

$$\phi^* = \arg\max_{\psi \in \mathrm{P},\, \theta \in \Theta} IR(S, \psi(\theta))$$

[2] Bombara, G et all, A Decision Tree Approach to Data Classification Using Signal Temporal Logic. In: Proc. of HSCC, 2016.

Maritime Surveillance

Synthetic dataset of naval surveillance of 2-dimensional coordinates traces of vessels behaviours.



$$\phi_{ROGE} = ((x_2 > 22.46)\,\mathcal{U}_{[49,287]}\,(x_1 \leq 31.65))$$

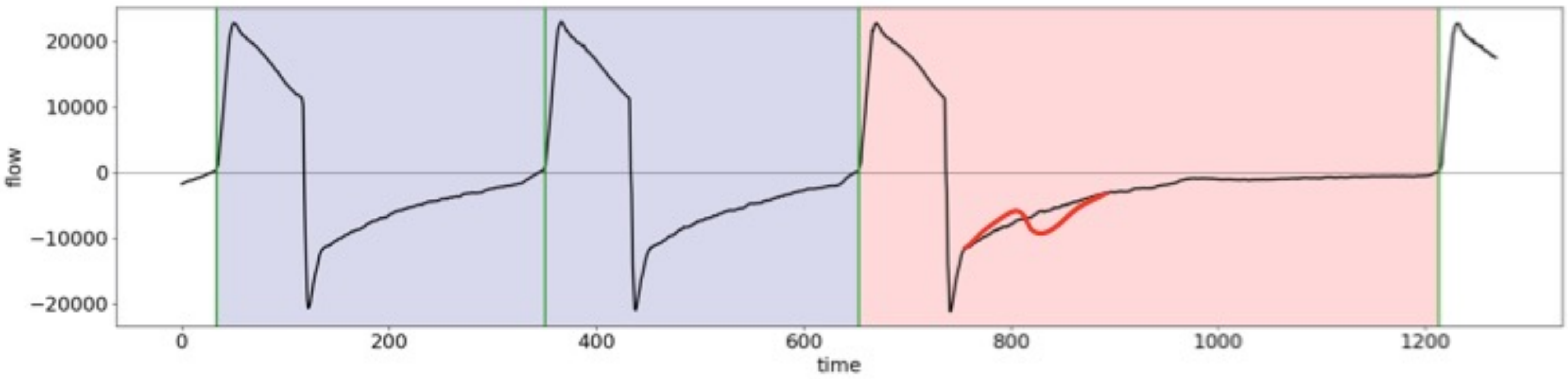
Maritime Surveillance

$$\phi_{ROGE} = ((x_2 > 22.46)\,\mathcal{U}_{[49,287]}\,(x_1 \leq 31.65))$$

$$\begin{aligned}\psi_{DTL4STL} = (((&\mathcal{G}_{[187,196)}x_1 < 19.8) \wedge (\mathcal{F}_{[55.3,298)}x_1 > 40.8)) \vee ((\mathcal{F}_{[187,196)}x_1 > 19.8) \wedge \\ &((\mathcal{G}_{[94.9,296)}x_2 < 32.2) \vee ((\mathcal{F}_{[94.9,296)}x_2 > 32.2) \wedge (((\mathcal{G}_{[50.2,274)}x_2 > 29.6) \wedge \\ &(\mathcal{G}_{[125,222)}x_1 < 47)) \vee ((\mathcal{F}_{[50.2,274)}x_2 < 29.6) \wedge (\mathcal{G}_{[206,233)}x_1 < 16.7))))))\end{aligned}$$

	ROGE	DTL4STL	DTL4STL$_p$
Mis. Rate	0	0.01 ± 0.013	0.007 ± 0.008
Comp. Time (sec.)	73 ± 18	144 ± 24	-

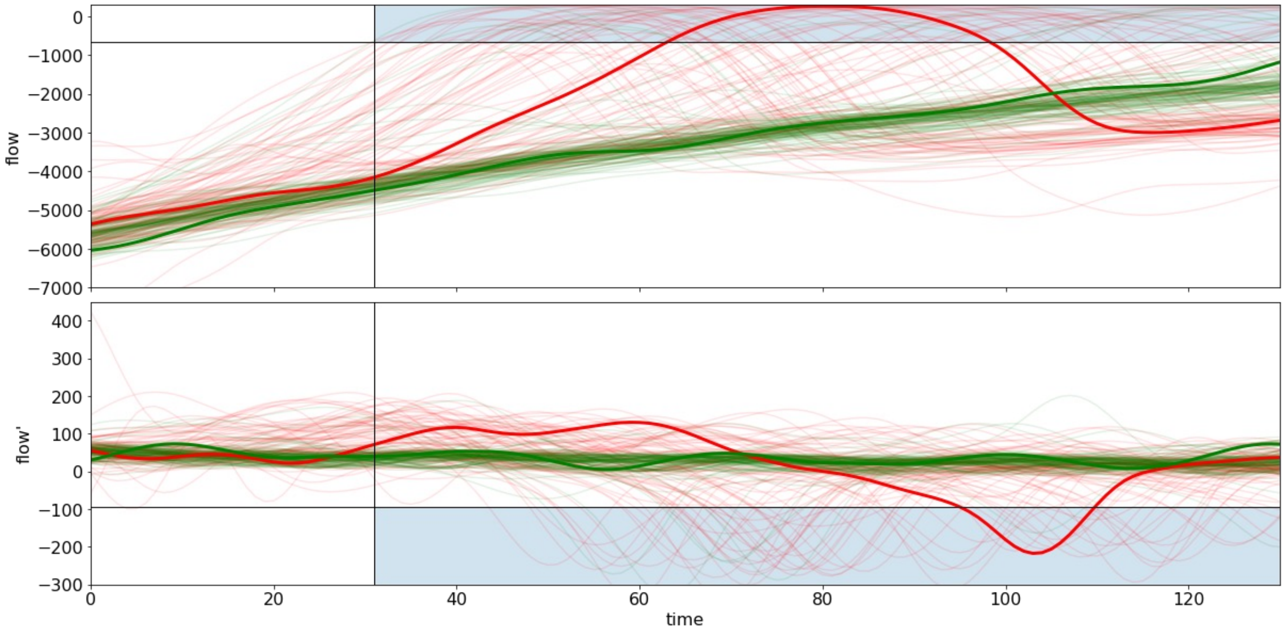
Ineffective Inspiratory Effort (IIE)



- Regular breaths
- Irregular breaths (IIE)

Ineffective Inspiratory Effort (IIE)

The dataset consists of 2-dim traces of flow and its derivative, flow’



	ROGE	DTL4STL
Mis. Rate	0.17 ± 0.01	0.23 ± 0.07
False. Pos	0.20 ± 0.02	0.23 ± 0.07
False. Neg	0.14 ± 0.02	0.20 ± 0.15
Comp. Time (sec.)	65 ± 7	201 ± 7

$$\phi = \mathcal{F}_{[31,130]}((flow \geq -670) \vee (flow' \leq -94))$$

Bibliography

Mining Requirements:

- Ezio Bartocci, Luca Bortolussi, Laura Nenzi, Guido Sanguinetti, System design of stochastic models using robustness of temporal properties. Theor. Comput. Sci. 587: 3-25 (2015)
- Jyo Deshmukh et al. Mining Requirements from Closed-loop Control Models (HSCC '13, IEEE Trans. On Computer Aided Design '15)
- Bartocci, E., Bortolussi, L., Sanguinetti, G.: Data-driven statistical learning of temporal logic properties, FORMATS, 2014
- Bufo, S., Bartocci, E., Sanguinetti, G., Borelli, M., Lucangelo, U., Bortolussi, L.i,Temporal logic based monitoring of assisted ventilation in intensive care patients, ISoLA, 2014.
- Nenzi L., Silvetti S., Bartocci E., Bortolussi L. (2018) *A Robust Genetic Algorithm for Learning Temporal Specifications from Data.* QEST 2018. LNCS, vol 11024. Springer, Cham.
- Bombara, G et all, A Decision Tree Approach to Data Classification Using Signal Temporal Logic. In: Proc. of HSCC, 2016.