Case Study using Posto

Lucky M. Kispotta

luckymk.mcs2024@cmi.ac.in

Chennai Mathematical Institute

2025-10-28

Table of Contents

Content

Introduction

Given an autonomous system which evolves in discrete times.
Devise a statistical method to monitor an autonomous system independent of it's nature (Linear / Non-linear).

Introduction

Given an autonomous system which evolves in discrete times.
Devise a statistical method to monitor an autonomous system independent of it's nature (Linear / Non-linear).

This method could be used to argue the safety of a self-driving car system is $> c$. Here c is the confidence.

System I/O execution model

Definition

The system I/O model is defined as :

$$f_{\text{sys}} : 2^{\mathbb{R}^n} \times \mathbb{R} \times \bigcup_{i \in [t-1]} o_i \rightarrow 2^{\mathbb{R}^n}$$

$$f_{\text{sys}}(\theta_0, t, [O]_{t-1}) = \theta_t$$

where, $\theta_0, \theta_1 \subset \mathbb{R}^n$, $\forall_{i \in [t-1]} o_i \subset \mathbb{R}^n$

Intuitively, f defines a transition function which maps the initial state to the next "t"th step w.r.t to some environment inputs.

Log

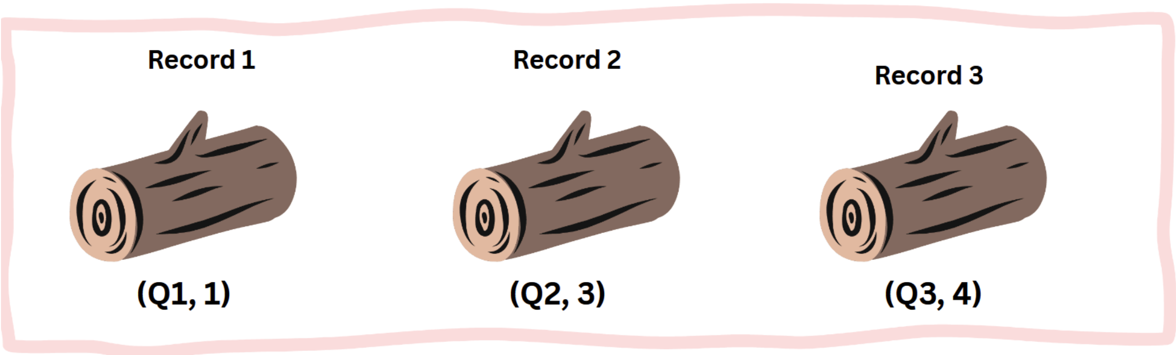
Definition

Given a system I/O execution model a finite size log of the system is defined as follows :

$$l = \left\{ \left(\hat{\theta}, t \right) \mid \theta_t \subseteq \hat{\theta}, t \leq H \right\}$$

where,

$$t, H \in R$$



Log

2) Visualization of Log

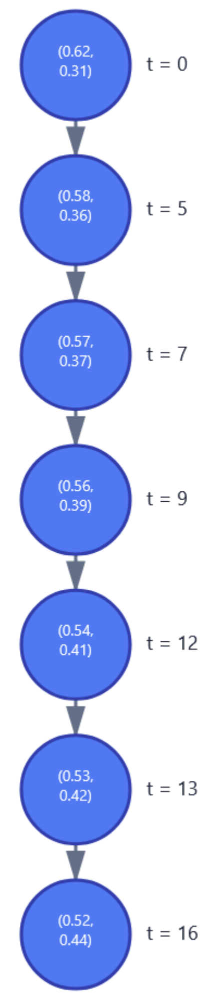


Figure 4: Given $T = 20$ and Porbablilty $\log = 20$
a) The visualization of bounds of states for uncertain log.
b) The visualization of log where state has a single value.

Trajectory

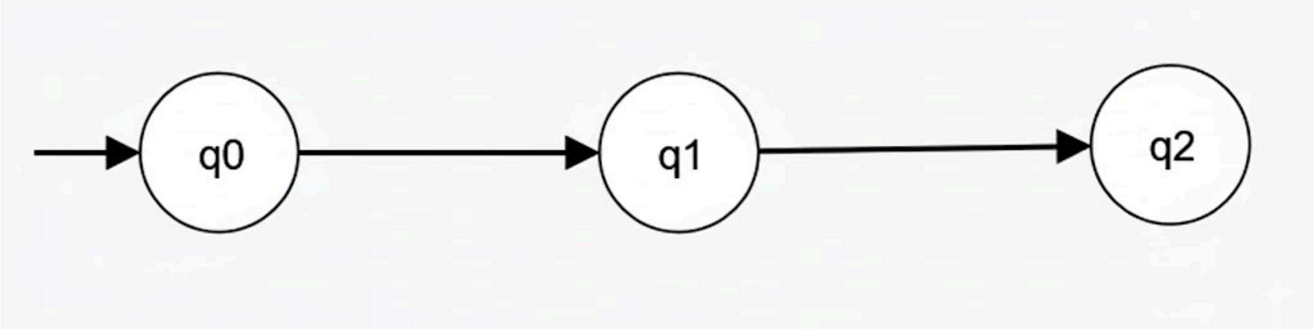
Definition

A trajectory τ of the system is an ordered sequence of states given as follows: $\tau = \{x_0, x_1, \cdots, x_H\}$ where $\forall t \in [0, H]$ and $f_{sys}(x_0, t, [O]_{t-1}) = x_t$.

Trajectory

Definition

A trajectory τ of the system is an ordered sequence of states given as follows: $\tau = \{x_0, x_1, \cdots, x_H\}$ where $\forall t \in [0, H]$ and $f_{sys}(x_0, t, [O]_{t-1}) = x_t$.



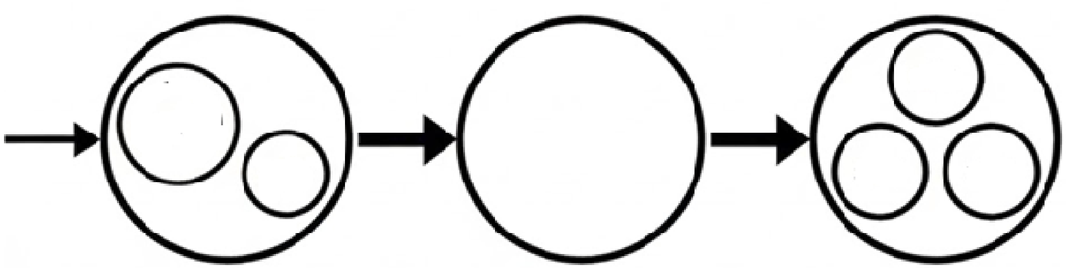
Here, each $q_i \in \mathbb{R}^2$. Since we are trying to model a system which represents the movement of an object in a space (here 2d).

Valid trajectories

Definition

A trajectory $\tau = \{x_0, x_1, \cdots, x_H\}$ is said to be valid with respect to a given log $l = \left\{\left(\hat{\theta}_t, t\right) \mid \theta_t \subseteq \hat{\theta}_t, t \leq H\right\}$ if $\forall_{\left(\widehat{\theta}_t, t\right) \in l} x_t \in \hat{\theta}_t$.

Intuitively,



Random Trajectory

Definition

Let a trajectory τ be randomly chosen from the set of all valid trajectories τ_{val} (w.r.t. to and environmental inputs $[O]_H$).

This is randomly drawn according to the distribution D, and formally expressed as $\tau = \text{Sample}\big(f_{sys}(\cdot), l, [O]_H, D\big)$

Visualization of Random Trajectory 1)

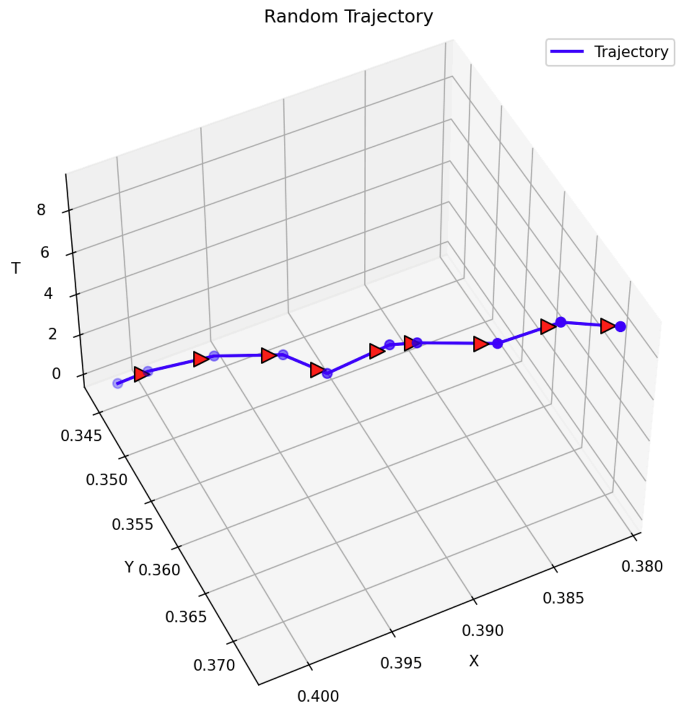


Figure 8: $3D$ trajectory plot.

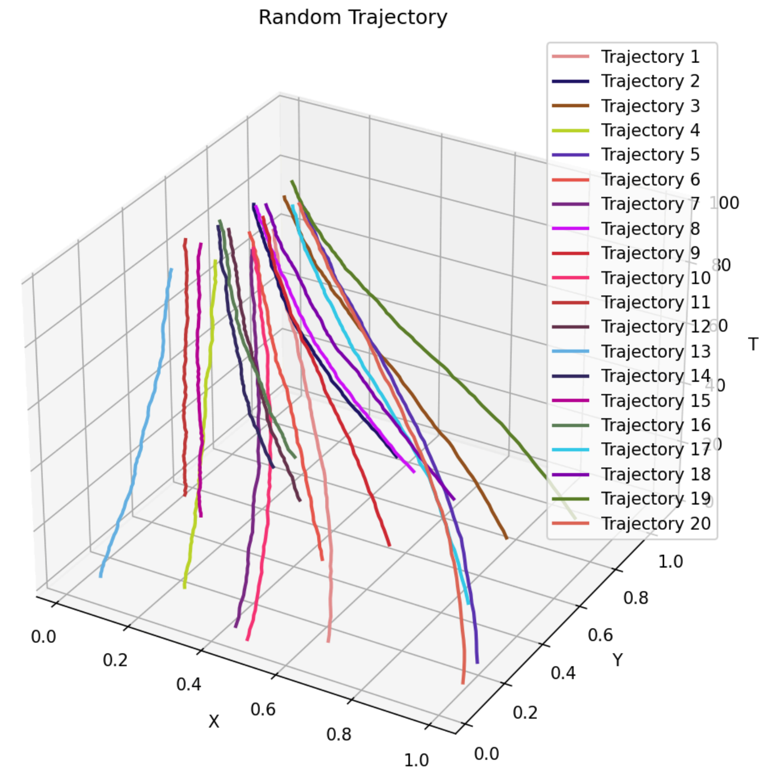


Figure 9: Random Trajectories.

Problem Statement

Now we formally define the Problem statement in hand.

Problem Statement

Now we formally define the Problem statement in hand.

Problem Statement

Given,

1. The system I/O model that is f_{sys}.
2. An uncertain log l.
3. Environment inputs $[O]_H$
4. The probabilistic distribution D 5. An unsafe set of trajectories $\mathcal{U}$.
6. A confidence parameter $c \in (0, 1)$ **desired**.

The problem is to perform monitoring to ensure safety of the system with confidence c as defined by Jeffery Bayes Factor based hypothesis testing.

Question

I was given a Non-Linear System to monitor.

$$x_{i+1} = x_i - dt(y_i + 1.5x_i + 1.5x_i^2)$$
$$y_{i+1} = y_i + dt(3x_i^2 - y_i)$$

Content

Overview

Overview

- Let K be the number of trajectories that need to be checked.
- The goal of this method is to correctly devise a value for K such that it "guarantees" that the system will work correctly with confidence $> c$.
- To enable hypothesis testing : Formulate two hypothesis.
- First one (H_0) represents the undesired result and the second one (H_1) represents the desired result.
- For each sample or trajectory check if it supports hypothesis H_0 or H_1.
- If any sample is in favour of H_0 (safe with probability $< c$) [includes unsafe case] return **False**.
- else return **True**.

Hypothesis Testing

- Let Null hypothesis be $H_0 : \Pr\left[f_{sys}(.), l, \mathcal{D}, \mathcal{U}\right] < c$

Hypothesis Testing

- Let Null hypothesis be $H_0 : \Pr\left[f_{sys}(.), l, \mathcal{D}, \mathcal{U}\right] < c$
- This is the probability that a random trajectory is safe with probability less than c.

Hypothesis Testing

- Let Null hypothesis be $H_0 : \Pr\left[f_{sys}(.), l, \mathcal{D}, \mathcal{U}\right] < c$
- This is the probability that a random trajectory is safe with probability less than c.
- Let alternative hypothesis be $H_1 : r\left[f_{sys}(.), l, \mathcal{D}, \mathcal{U}\right] \geq c$

Hypothesis Testing

- Let Null hypothesis be $H_0 : \Pr\left[f_{sys}(.), l, \mathcal{D}, \mathcal{U}\right] < c$
- This is the probability that a random trajectory is safe with probability less than c.
- Let alternative hypothesis be $H_1 : r\left[f_{sys}(.), l, \mathcal{D}, \mathcal{U}\right] \geq c$
- This is the probability that a random trajectory is safe with probability $\geq c$.

Hypothesis Testing

- Let Null hypothesis be $H_0 : \Pr\left[f_{sys}(.), l, \mathcal{D}, \mathcal{U}\right] < c$
- This is the probability that a random trajectory is safe with probability less than c.
- Let alternative hypothesis be $H_1 : r\left[f_{sys}(.), l, \mathcal{D}, \mathcal{U}\right] \geq c$
- This is the probability that a random trajectory is safe with probability $\geq c$.
- We want the hypothesis testing to conclude that that the H_1 is true.

Derivation of K

The probability that set of trajectories X is safe with probability m is m^K.

Derivation of K

The probability that set of trajectories X is safe with probability m is m^K.

Therefore the probability that the set of trajectories is safe given probability of a random trajectory is safe $< c$ is :

$$\Pr[\forall \tau \in X : \tau \cap = \emptyset \mid H_0] = \int_0^c q^K dq = \frac{c^{K+1}}{K+1}$$

Derivation of K

The probability that set of trajectories X is safe with probability m is m^K.

Therefore the probability that the set of trajectories is safe given probability of a random trajectory is safe $< c$ is :

$$\Pr[\forall \tau \in X : \tau \cap = \emptyset \mid H_0] = \int_0^c q^K dq = \frac{c^{K+1}}{K+1}$$

Therefore the probability that the set of trajectories is safe given probability of a random trajectory is safe $\geq c$ is :

$$\Pr[\forall \tau \in X : \tau \cap = \emptyset \mid H_1] = \int_c^1 q^K dq = \frac{1 - c^{K+1}}{K+1}$$

Bayes Factor

In any hypothesis testing problem the Bayes Factor measures how much likely is the data under H_1 than H_0.

Bayes Factor

In any hypothesis testing problem the Bayes Factor measures how much likely is the data under H_1 than H_0.

Bayes Factor

Bayes Factor is formally defined as the ratio :

$$\frac{\Pr[D \mid H_1]}{\Pr[D \mid H_0]}$$

Here, D are K *safe trajectories.*

Bayes Factor

Interpretation of Bayes Factor

If Bayes Factor is > 1 implies the data is more in favour of H_1.
$B = 10$ implies *"the observed data is 10 times more likely under H_1 than H_0"*.

The paper uses a "hardcoded" *Bayes Factor* as a threshold to accept the hypothesis H_1. Hence,

$$\frac{1 - c^{K+1}}{c^{K+1}} > B \Longleftrightarrow K > -\log_c(B + 1)$$

Intuition : is to find a K such that if K trajectories are safe then Bayes Factor of observed data$> B$.

Content

Flowchart

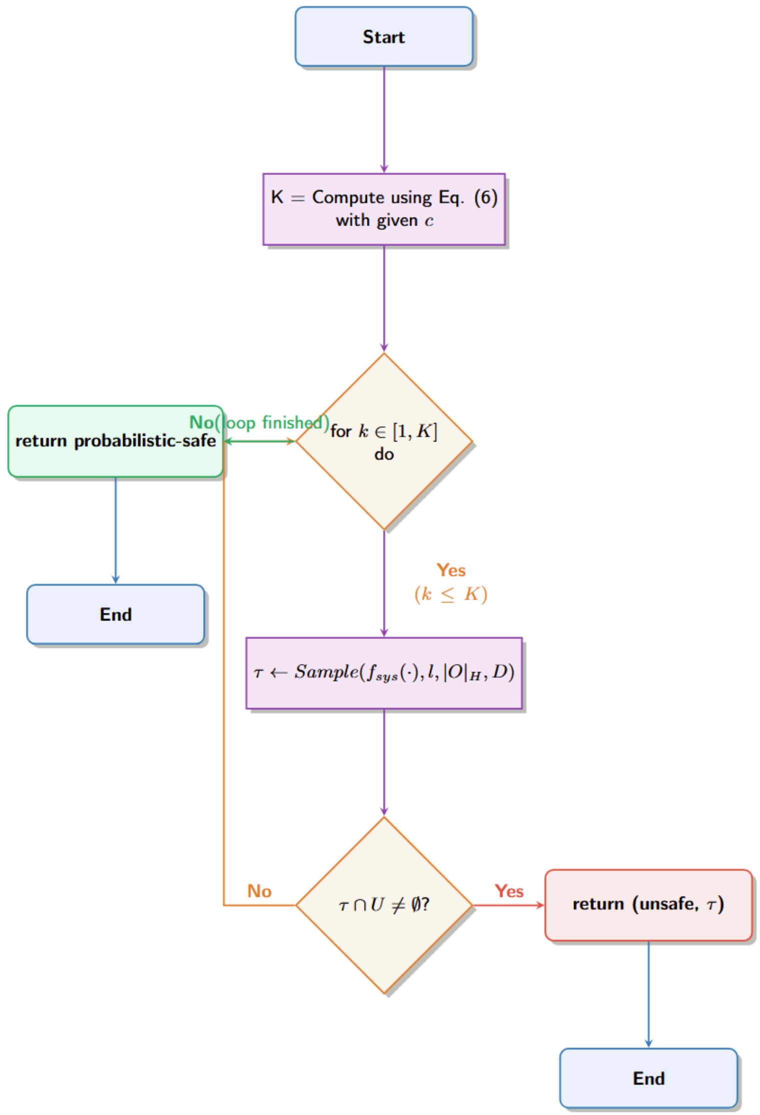


Figure 10: Flowchart of the algorithm

Content

Few Results (Logging Probability as variable)

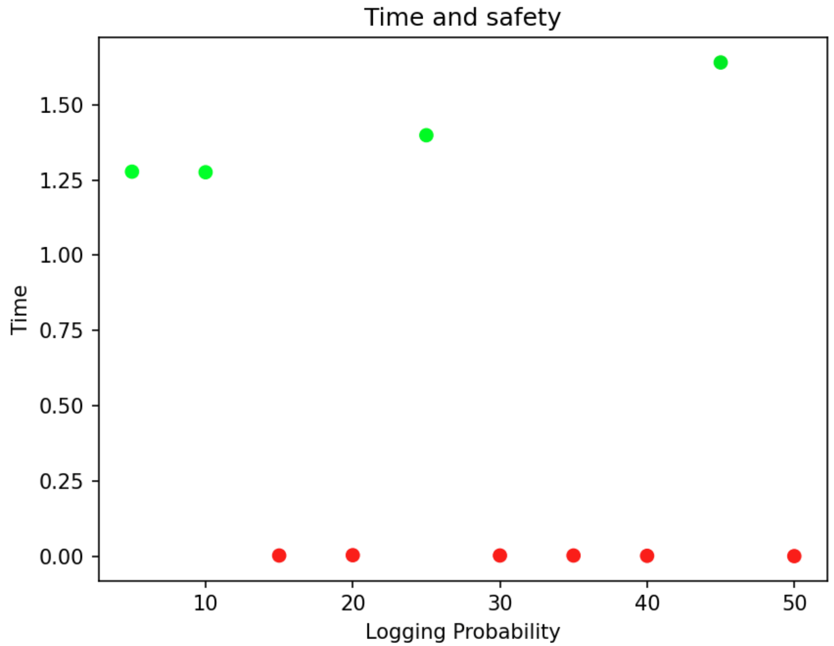


Figure 11: The values of constraints are : *unsafe = 0.7, op = 'ge', state = 1 (y)*

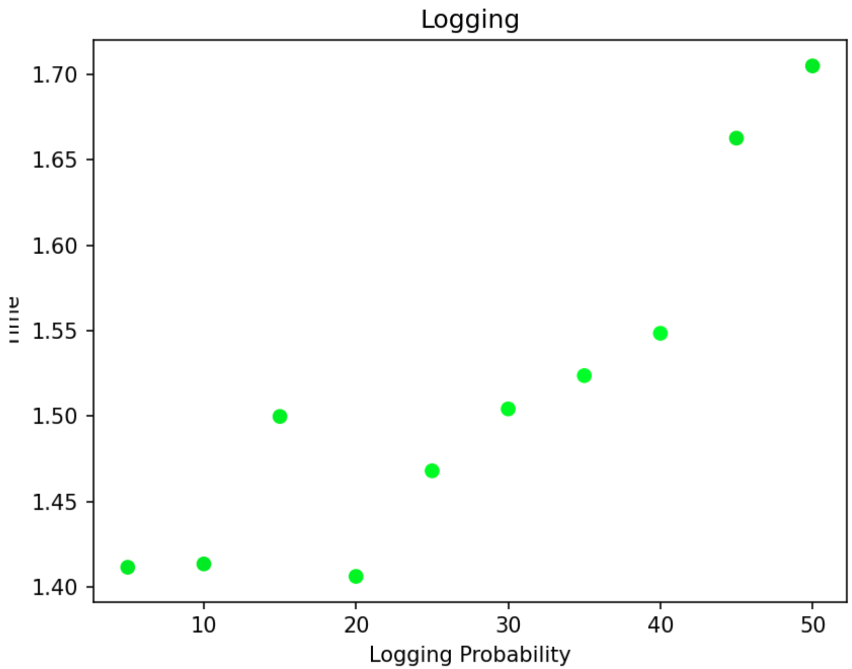


Figure 12: As the Logging Probability increases the time increases. The color of the points depend on the ratio $\frac{|\text{valid trajectories}|}{|\text{total trajectories}|}$. Green implies safe and red implies not safe. The values of constraints are : *unsafe = −0.1, op = 'le', state = 0 (x)*

More Results (Logging Probability as variable)

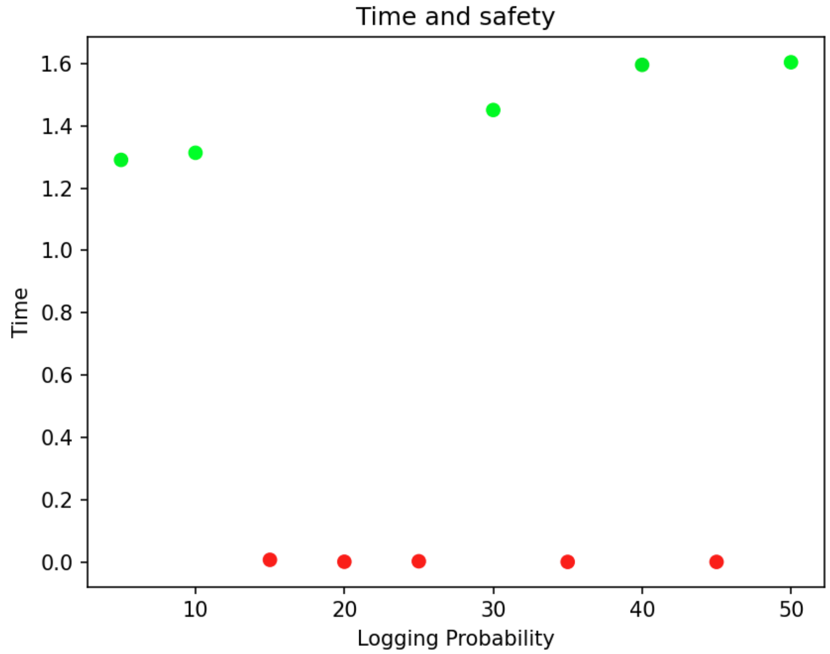


Figure 13: The values of constraints are : *unsafe = 0, op = 'le', state = 0 (x)*

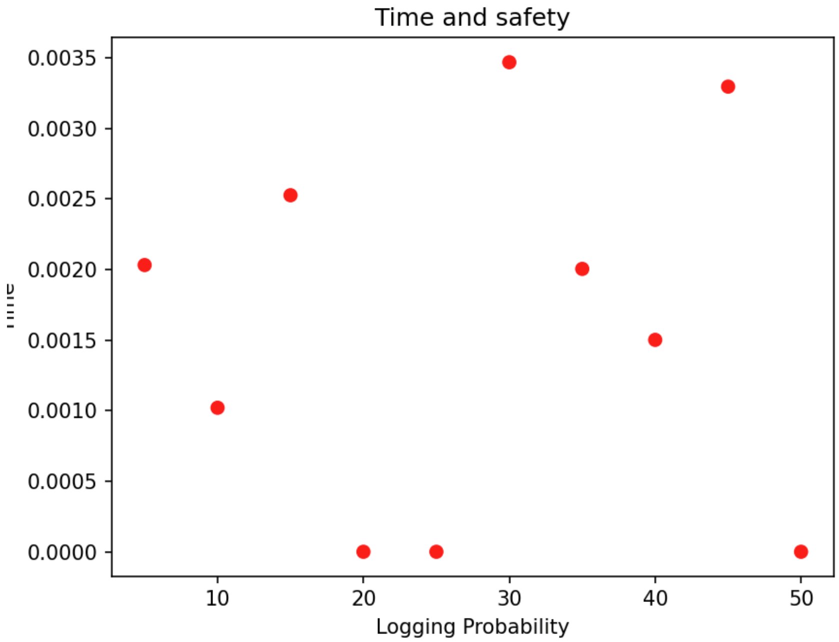


Figure 14: This is a very strict safety condition. The values of constraints are : *unsafe = 0.1, op = 'ge', state = 0 (x)*

Few Results (*Confidence* as variable)

An observation to make is that the algorithm immediately produces counter examples.

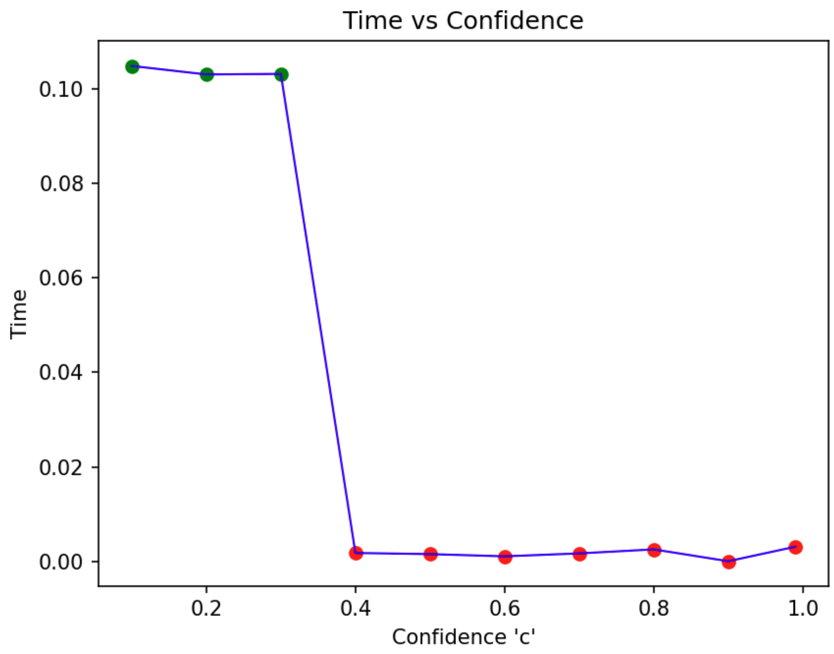


Figure 15: The values of constraints are : *unsafe = 0.7, op = 'ge', state = 1 (y)*

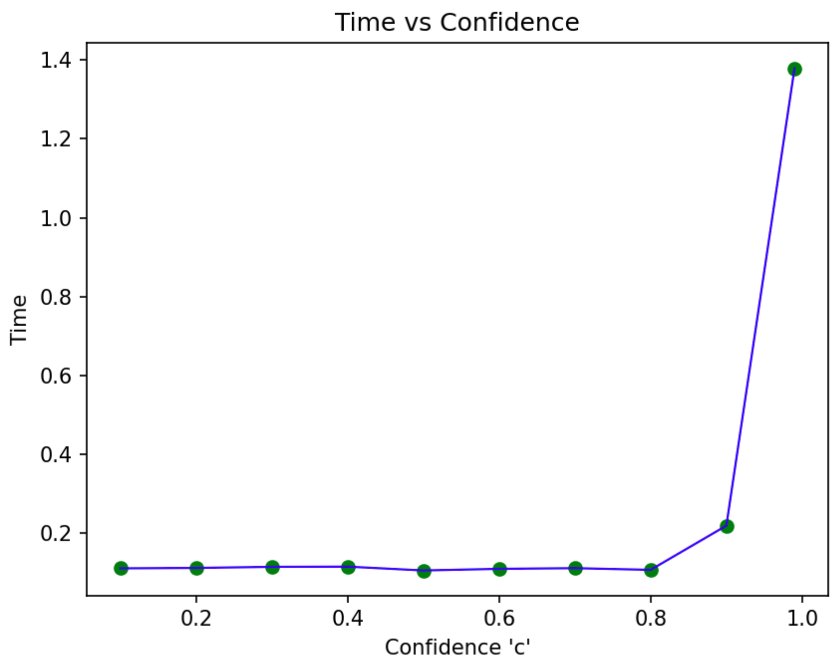


Figure 16: The values of constraints are : *unsafe = −0.1, op = 'le', state = 0 (x)*

More Results (*Confidence* as variable)

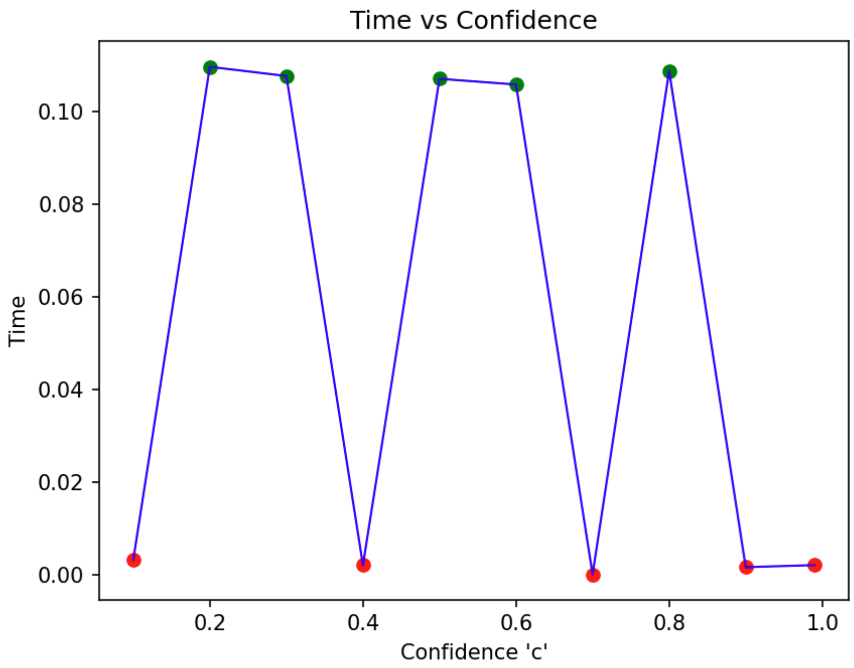


Figure 17: The values of constraints are : *unsafe = 0, op = 'le', state = 0 (x)*

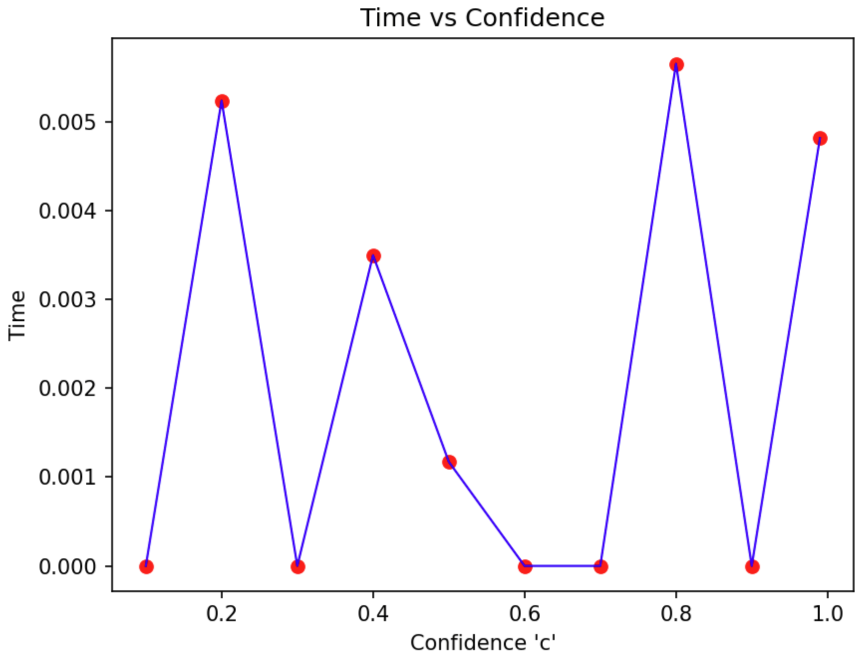


Figure 18: This is a very strict safety condition. The values of constraints are : *unsafe = 0.1, op = 'ge', state = 0 (x)*

Thanks for Listening.

Bye Bye