Robust Bayesian Estimators for Transition Parameters in Probabilistic Model Checking

Xingyu Zhao

[Sing-You]

26th Mar. 2020

ASSURING AUTONOMY INTERNATIONAL PROGRAMME

HERIOT WATT UNIVERSITY

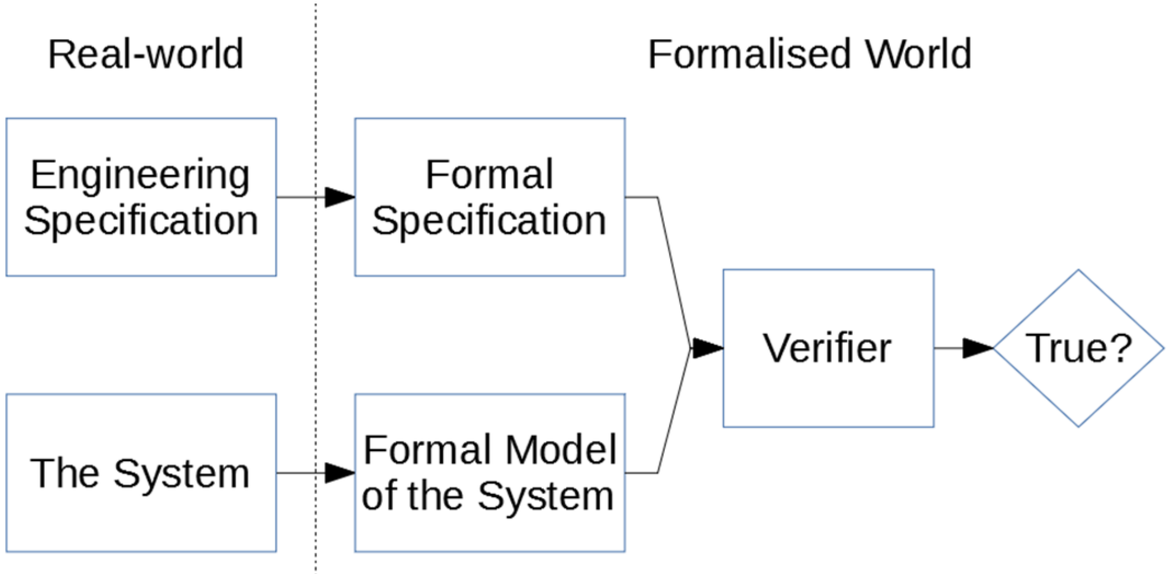
ORCA HUB
Offshore Robotics for Certification of Assets
Remote Safety and Integrity

Agenda (20+10)

- Background
  - Motivations – how to learn transitions parameters in PMC effectively.
- Solutions – Some runtime estimators
  - Preliminaries – MLE, KAMI (ICSE’09), COVE (ICPE’14)
  - CBI (Conservative Bayesian Inference) -- very rare events
  - IPSP (Imprecise Probability with Sets of Priors) – regularly observed events
- Extensions and future ideas
  - CTMC, one-off events.
  - Change-point detection – IPSP bound width indicates prior-data conflicts
  - Anything else?

Background

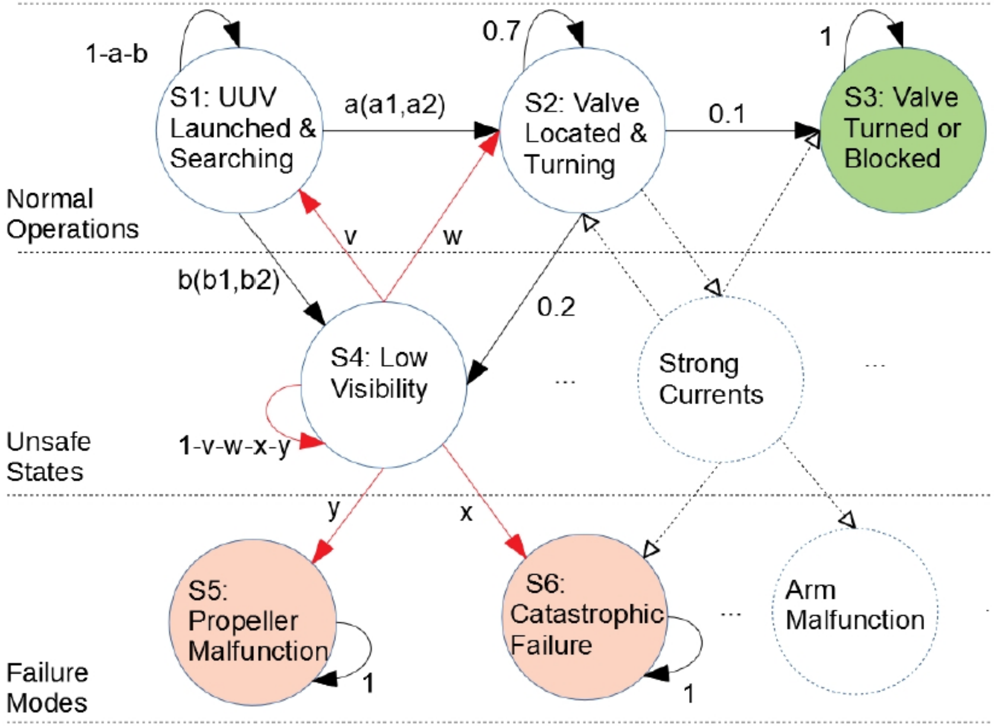
- What and Why: Probabilistic Model Checking (PMC)?
  - Probabilistic formal models capture inevitable uncertainties
    - Stochastic environments/Random H/W components failures
    - Black-box nature of DNNs.
  - Quantitative properties are of more interest
    - Chance of a successful mission
    - Chance of seeing a catastrophic failure
    - Expected time/energy costs



Background – Motivations

- One inherent problem for PMC (or any formal verification):
  - Assuming the formal model accurately represents the real-world.
    - Fairly easy to argue for simple systems, but hard for e.g., robots under uncertainties

- What' the structure of the Markov model?
- What're the transition probabilities?
  - Bayesian estimators at runtime



The MLE estimator

- Say r.v. *X* – the unknown transition probability from state *i* to *j*.
  - Observe *k* transitions from i to j in total *n* outgoing transitions from *i*.
  - A sequence of i.i.d. Bernoulli trails, given the probability x.
    - Flip a coin *n* times and see *k* heads.

$$L(x; k\&n) = x^k(1-x)^{n-k}$$

$$\hat{X}_{MLE} = \frac{k}{n}$$

The KAMI estimator

- A Bayesian estimator
  - Conjugacy – Beta prior and Binomial likelihood

$$f(x|k\&n) = \frac{L(x;k\&n)f(x)}{\int_0^1 L(x;k\&n)f(x)dx}$$

$$f(x) \sim Beta(a,b)$$

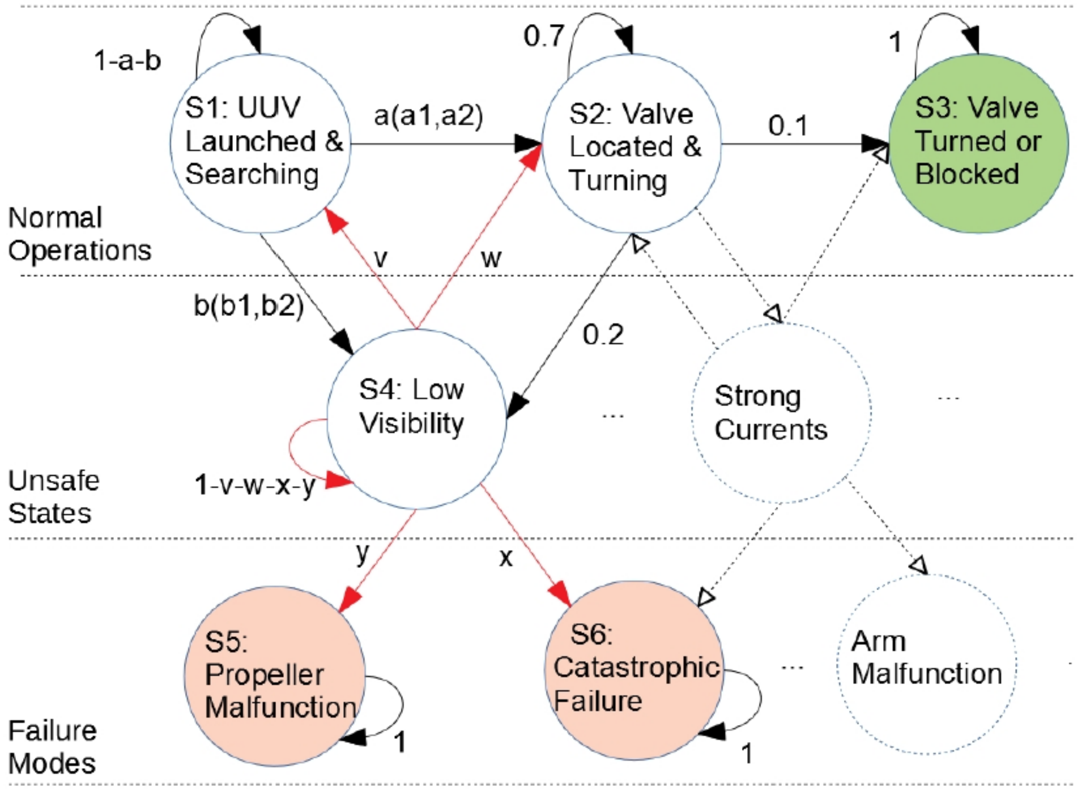
$$f(x|k\&n) \sim Beta(a+k, b+n-k)$$

$$E[X|k\&n] = \frac{a+k}{a+b+n}$$

- COVE
  - Add aging factors of each data point

$$E(f(x)|k\&n) = \frac{a+\sum_1^m w_i k_i}{a+b+\sum_1^m w_i n_i}, \qquad \sum k_i = k, \sum n_i = n$$

The CBI estimator



- X -- very rare events
  - E.g. catastrophic failures
  - 10^-6? 10^-8?
  - MLE? – 0
  - KAMI (COVE)?
    - Why it has to be Beta? Can you justify any complete prior distribution?
      - implicitly assumptions introduce optimistic bias.
      - Sensitive to the choice of priors

The gist of CBI

$$E(X|k\&n) = \frac{\int_0^1 xL(x;k\&n)f(x)dx}{\int_0^1 L(x;k\&n)f(x)dx} \quad (1)$$

$$Pr(x < 10^{-4}) = 0.9 \quad (2)$$

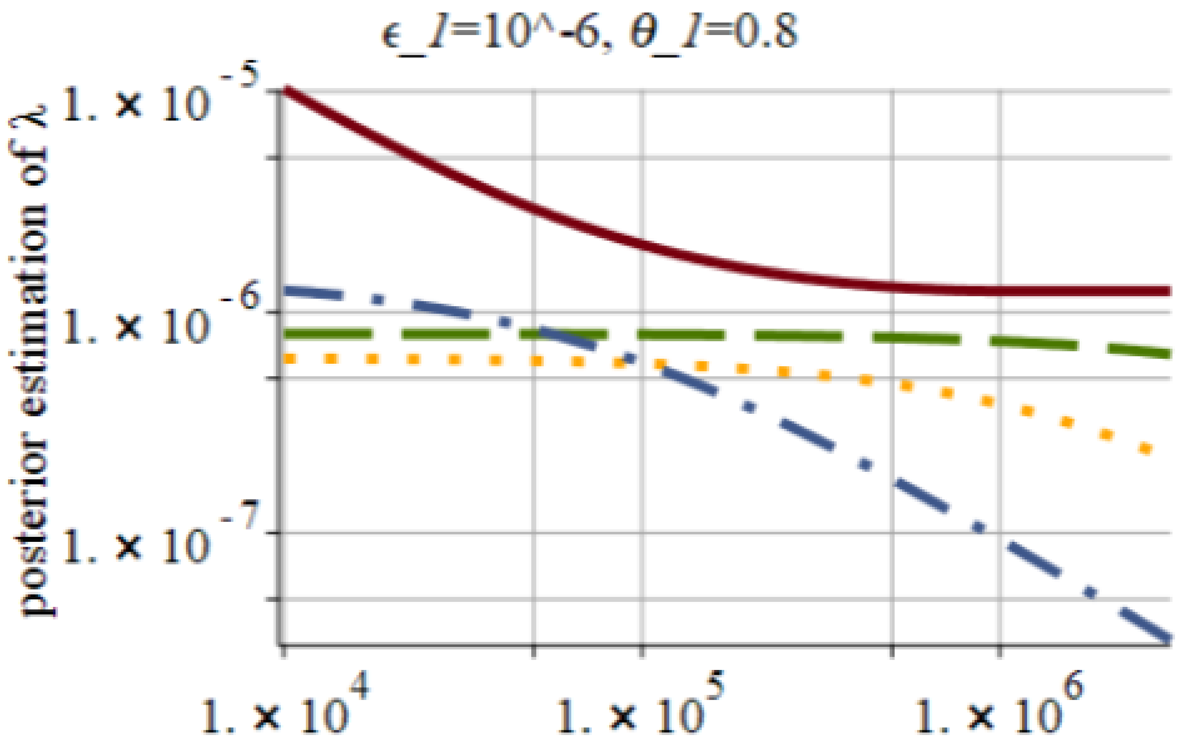
- What about limited and partial prior knowledge?
  - Much easier to justify/obtain..
    - IEC61508 SIL4 implies a confidence bound
  - Partial in the sense of
    - A infinite number of priors satisfying (2).
- CBI: To maximise (1) (i.e. being conservative), subject to (2), what is the corresponding *f(x)*?
  - Generalised to other objective functions and partial prior knowledge.

– $\mathbb{E}[pfd] \leq m$: the prior mean pfd cannot be worse than a stated value;
– $Pr(pfd \leq \epsilon) = \theta$: a prior confidence bound on pfd;
– $Pr(pfd = 0) = \theta$: a prior confidence in the perfection of the system;
– $\mathbb{E}[(1 - pfd)^n] \geq \gamma$: prior confidence in the reliability of passing n tests.

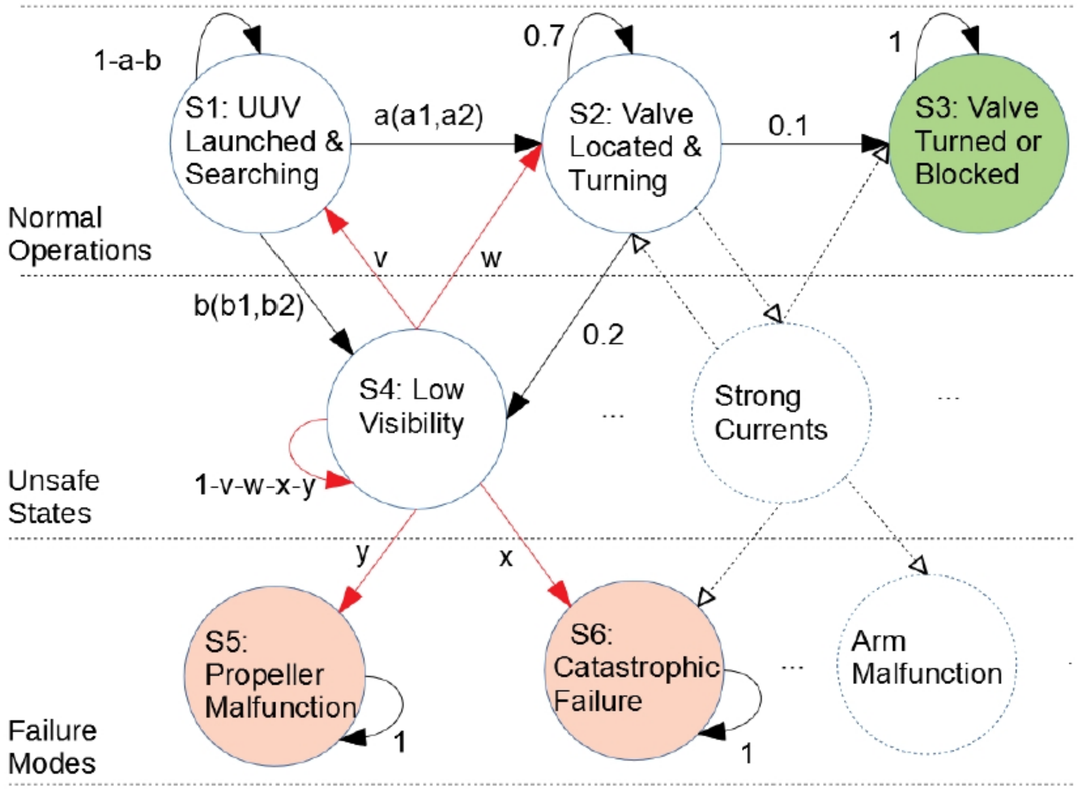
SIL	PFD	PFD (power)
1	0.1–0.01	$10^{-1} – 10^{-2}$
2	0.01–0.001	$10^{-2} – 10^{-3}$
3	0.001–0.0001	$10^{-3} – 10^{-4}$
4	0.0001–0.00001	$10^{-4} – 10^{-5}$

CBI VS KAMI

- CBI (red solid line)
  - Worst-case prior
- KAMI (3 dotted lines)
  - 3 arbitrary prior distributions satisfying the partial prior knowledge.
- CBI is guaranteed to be conservative.



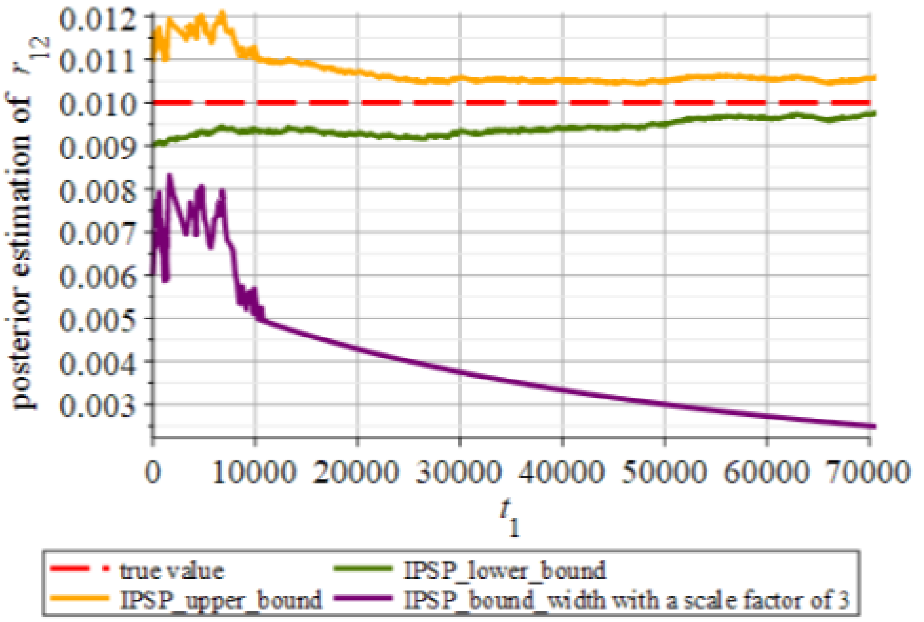
The IPSP estimator



- W – regularly observed transitions
  - MLE, KAMI is ok…
- But, can we do better?
  - Point estimates without est. err.
  - Est. err. propagates and compound.
- Assessors might be reluctant to express a single prior distribution
  - Their vague and imperfect prior knowledge
  - More practical and flexible way to express priors?
  - A community of imprecise probability gives us some solution..

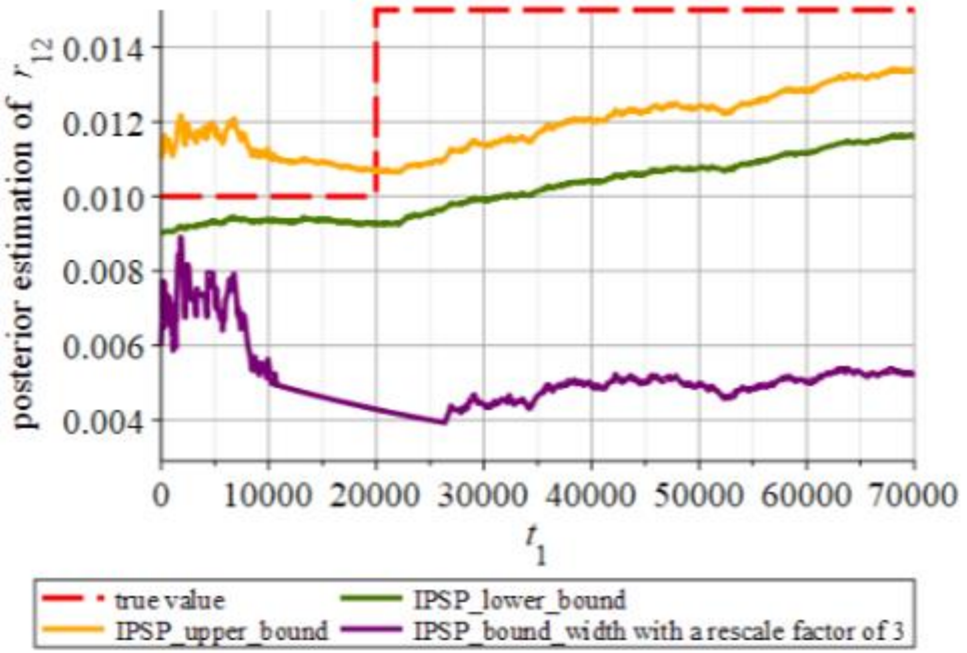
The gist of IPSP

- A set of Beta distribution as priors – *a_u, a_l, b_u, b_l*
  - Instead of a single Beta(a, b).
  - A set of posteriors, in theory.
  - The maximum and minimum of the set can be determined by
    - What you observed and your parameters of the set of Beta.
    - Closed-form expressions
- The width of the bounds
  - Measures the est. err.
  - When data confirms priors
    - Confidence in est. increases, width reduces.



Extension & future ideas

- CBI, IPSP for CTMC
- From cataphoric failure to one-off events
  - E.g., probability of finishing a difficult task
  - Minimisation means being conservative
- Change-point detection
  - ``prior-data conflict''
  - Triggers the change-point detector.
- Anything else?



THANK YOU

- xingyu.zhao@hw.ac.uk