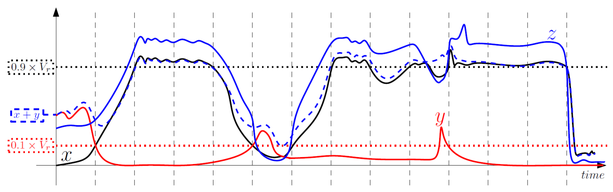
Flexible Mining of Causal Relations from Hybrid System Traces

António Bruto da Costa[1], Goran Frehse[2], Pallab Dasgupta[1]

[1]Indian Institute of Technology Kharagpur; [2]ENSTA ParisTech France

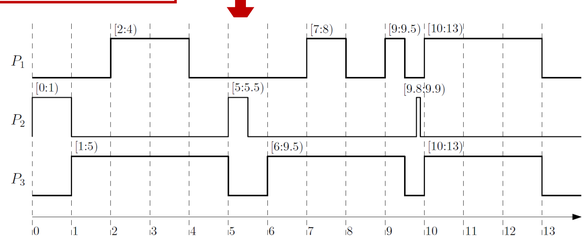
Acknowledging Partial Support from Intel Labs, Intel Corporation

Booleanizing Traces → **Build Decision Tree *for Prefix Sequence*** ↔ **Mine New Predicates** → **Translate Tree to Prefix Sequences**



Real Valued Time-Series

Lists of Truth Intervals

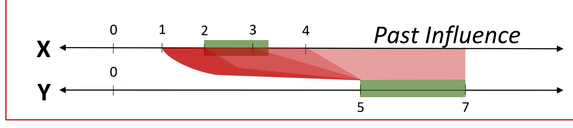


Predicates

P_1: $x \geq 0.9 * V_r$
P_2: $y \geq 0.1 * V_r$
P_3: $z \geq x + y$

Pseudo-Targets

Resolution of K up to a sequence expression length of N.
K=1; N=4



Unified Entropy: Pseudo-Target State Overlap

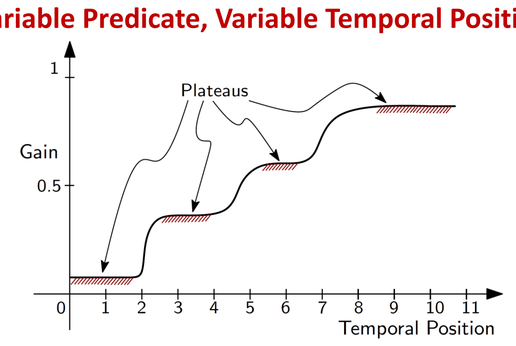
Node Entropy =
Entropy(True Target) + Entropy(False Target)
– ***Overlap Entropy***

Unified Gain: Branch Time-Points Overlap

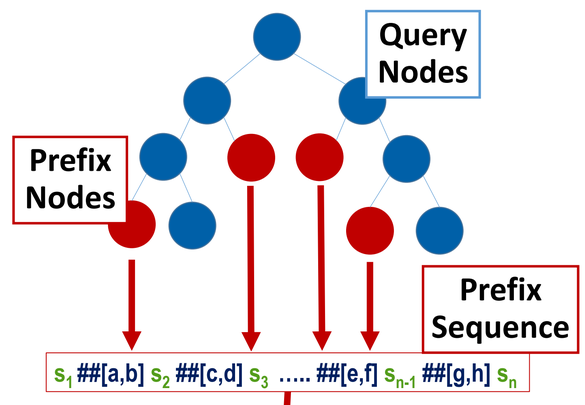
Decision Gain =
Node Entropy + α_{left}(Entropy$_{left}$) + α_{right}(Entropy$_{right}$)

Decision: ***Choose the Predicate and its position in the prefix sequence, maximizing the Unified Gain.***

Variations in Gain:
Variable Predicate, Variable Temporal Position



- For a fixed predicate, Gain increases *monotonically* with increasing temporal distances from the target.
- Gain of choosing a predicate x ⋈ c varies *unpredictably* with variations in constant c ∈ ℝ.
- We use Simulated Annealing to learn new predicates.



Learned Explanation

s_1 ##[a,b] s_2 ##[c,d] s_3 ….. ##[e,f] s_{n-1} ##[g,h] s_n => Target

INTRODUCTION

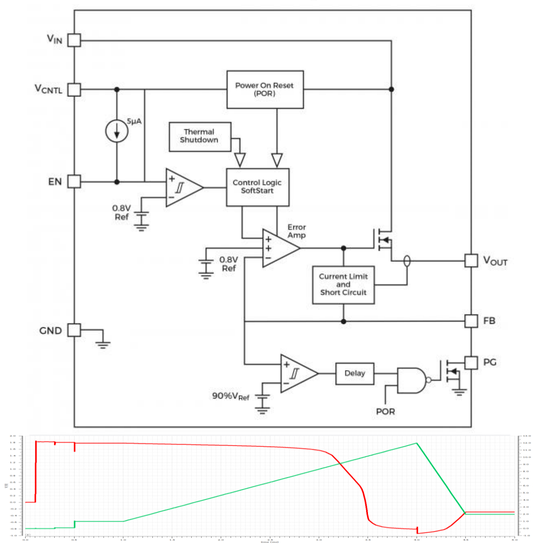
- ❖ **Mining Temporal Properties from Time-series data.**
  - Filter *real* properties from incidental ones
- ❖ **Property Mining: Complex for Continuous / Hybrid Systems.**
  - Degrees of influence on a consequent - varies continuously - dense time.
- ❖ **Parameterized method to learn likely causes of an well defined event.**
  - Interval arithmetic
  - Flexible learning - Over Time Scales and Predicates over system variables.

Interval Arithmetic & Decision Trees can be used to *Learn* Temporal Sequences of Events that Explain Observations in Time-Series data.

WHY?

- ❖ **Root Cause Analysis**
  - Infeasible: Manually examine sizeable time-series & identify common causes.
- ❖ **Anomaly Detection**
  - Learned patterns: Detect rare and suspicious deviations.
- ❖ **Explaining Classifications**
  - Derive succinct, human – interpretable explanations for events and their absence.

Mixed Signal Circuit Properties



Low Dropout Regulator

Functional block diagram of the AP7176B 3A ultra-low dropout LDO, Diodes Inc.

Circuit Properties Learned (5x10^5 time points)

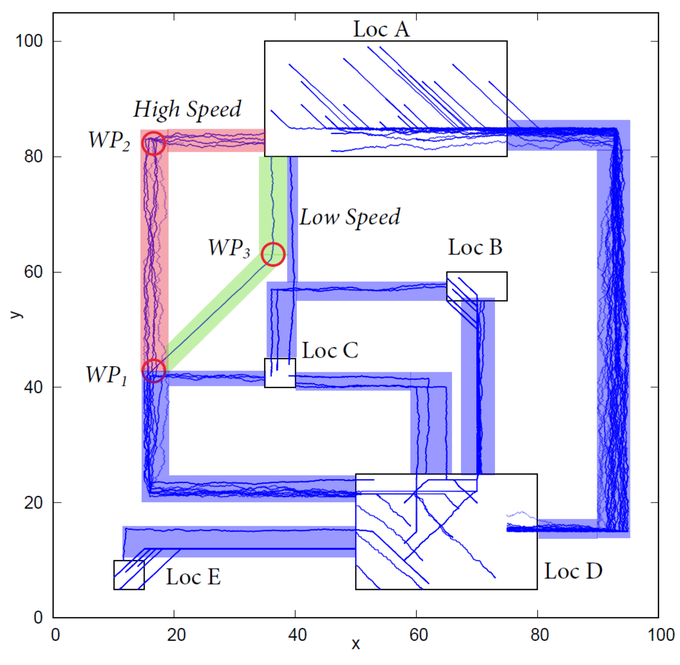
*The regulator has a **rise time** of 4us*

`!InShrtCkt && VLowerBand |=> ##[ 4.06us : 4.09us ] VUpperBand`

*The regulator enters a **short circuit state** when the voltage is under 1.6V and the current is above 8.5V.*

`!VAbove && ShrtCktEvent |=>##[0:20us] InShrtCkt`

Traffic (Vehicular/Network/Other)



Traffic Patterns (2.5x10^5 time points)

Movement** from Loc A to Loc D with **timing

`LocA |=> ##[ 1h24m50s : 2h3m ] LocD`

*From LocD, **passing through waypoint-WP2** within 1h to 6h, it is possible to reach LocA within the following 1h.*

`LocD ##[ 1h4m : 6h11m22s ] WP2 |=> ##[ 0h : 1h ] LocA`

Applications And Opportunities

- ❖ **Cyber Physical Systems**
  - Transactional relationships: Patterns of interaction between control and plant. Root cause analysis and anomaly detection
- ❖ **Circuit Design and Testing**
  - Aid for formal specification writing.
  - Design guidance, Root Cause, Timing Analysis.
  - Circuit Behavior abstracted and used in an Assume/Guarantee reasoning framework.
- ❖ **Process Control**
  - Root cause analysis, Process optimization.
- ❖ **Business and Finance**
  - Learning from transactional patterns.
- ❖ **Music and the Arts**
  - Understanding how artists compose, the structure of musical movements, frequently used patterns.
- ❖ **Behavioral Sciences**
  - Pattern extraction, understanding and explaining behaviours.
- ❖ **Medicine**
  - Root cause analysis.