Provably Correct Design of Observation for Fault Diagnosis and State Estimation under Privacy and Network Constraints

CISE/CNS/Computer Systems Research, Grant #1618369

PI: Agung Julius (Rensselaer Polytechnic Institute)

Background

➢ In complex cyber-physical systems, operation planning and control are performed using limited information, e.g., because of measurement limitation and privacy concerns.

➢ We seek to build a framework for **formal verification and synthesis** that the system operates correctly under such limitations.

➢ **Key question:** Can the right inference be made using the available information, e.g., for fault diagnosis and state estimation? (verification)

➢ **Key question:** What observation/information is needed for fault diagnosis and state estimation? How to synthesize the diagnoser/observer? (synthesis)

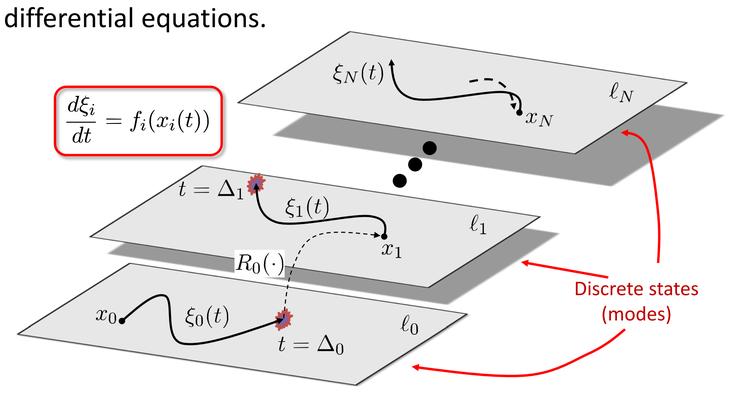
➢ **Key question:** How do limitations of communication network (transmission delay, limited bandwidth) affect the system?

➢ We propose a **model-based approach**, using hybrid systems to model the system dynamics.

Hybrid Systems

➢ Hybrid systems have both continuous and discrete dynamics.

➢ Can be considered as a product of automata and differential equations.



➢ Continuous states evolve according to differential equations (dependent on the discrete state).

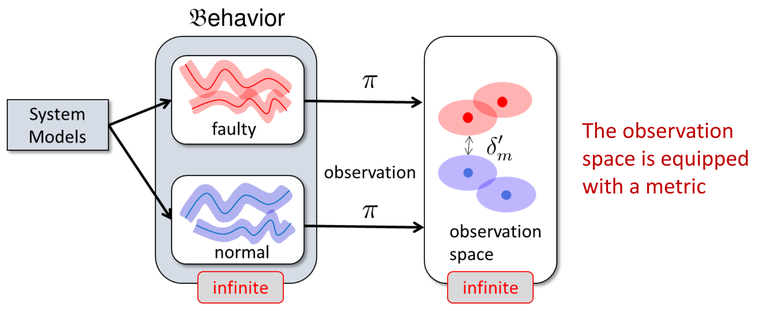
➢ Discrete state is updated at transition event times (guarded by conditions on the continuous state). The continuous state can be reset upon transition event.

➢ Ideal for modeling cyber-physical systems. Special cases include timed automata, finite transition systems, dynamical systems.

➢ Execution trajectories of a hybrid system include the evolution of the continuous states, discrete states, and the transition events.

Observation Map

➢ **Behavior:** collection of all execution trajectories of a hybrid system (could be varying by init state or parameters).



➢ Examples of observation map: hiding some of the variables, projection to the sequence of (timed) events, defining logical predicates on the variables.

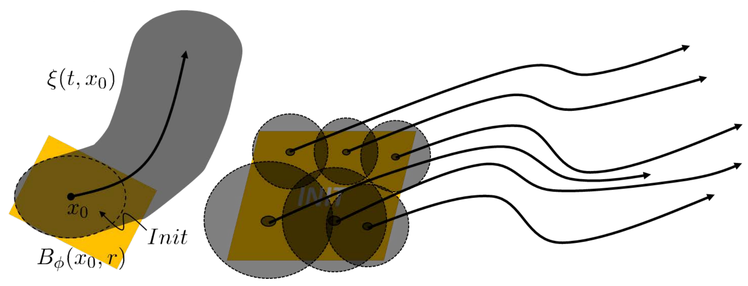
➢ **Fault diagnosis:** are the observations of normal and faulty behaviors separated (with sufficient distance)?

➢ **State estimation:** can be system state be identified using the observed information?

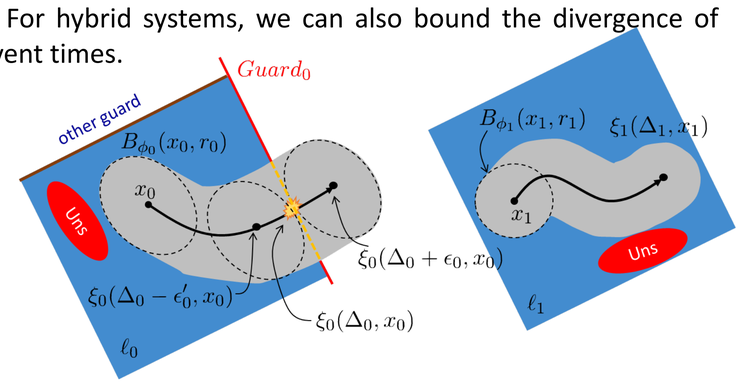
➢ **Privacy preservation:** can we guarantee that two behaviors are indistinguishable under the observation map?

Trajectory Robustness

➢ With trajectory robustness, we can formally bound the variation of the execution trajectories as we vary the initial state or parameters.



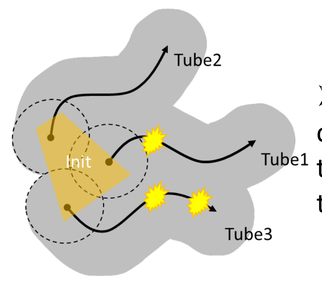
➢ For hybrid systems, we can also bound the divergence of event times.



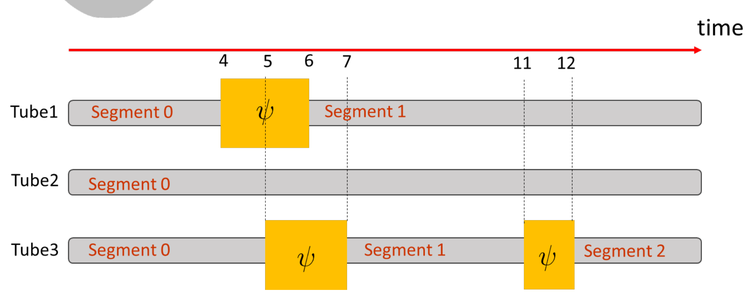
➢ **Key enabling result:** Compact behavior can be approximated by **finitely many trajectories**.

Timed Automata as Observer

➢ Suppose that we can only observe event times. Fault diagnosis can be done using the framework of timed automata.



➢ For simplicity, suppose we can cover the behavior with 3 robust tubes. Each tube has its own event timing (with uncertainty)



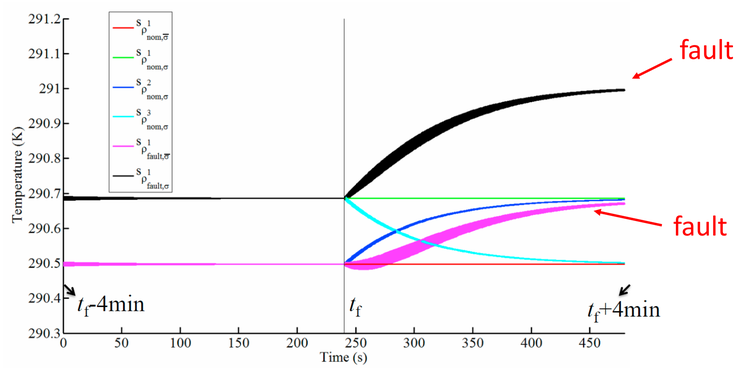
➢ Depending on **when the event ψ is observed**, we can deduce in which tube the state is. For example, no event until t=7 means the state is in Tube 2.

Privacy Preserving Fault Detection

➢ We consider a family of ODE models for the temperature and humidity in a room under 4 conditions: normal and occupied (NO), normal and empty (NE), faulty and occupied (FO), and faulty and empty (FE). Note: faulty means a window is open.

➢ Given continuous measurement of temperature and humidity, we can diagnose fault (i.e., if the window is open). That is, we can distinguish between (NE or NO) and (FE and FO).

➢ However, what if we do not want to divulge whether the room is occupied. That is, we do not want to distinguish between (FE or NE) with (FO or NO).

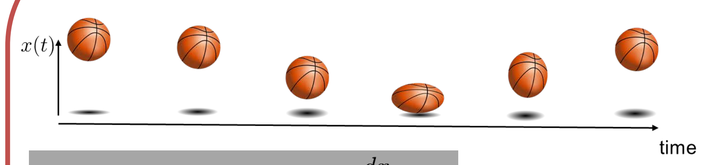


➢ A temporal logic formula is defined as a fault monitor

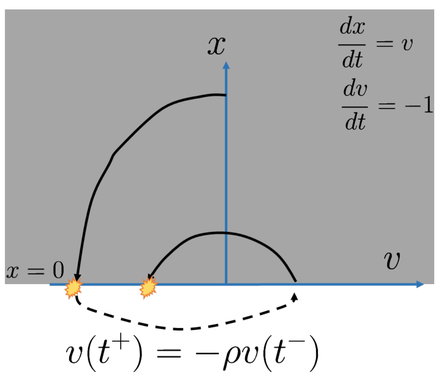
$$\phi(\alpha^*) = \left(\Box_{[160,180)} T > 290.75\right) \vee \left(\left(\Diamond_{[0,20)} T < 290.516\right) \wedge \left(\Diamond_{[39,65)} T > 290.525 \wedge \left(\Diamond_{[122,161)} T < 290.625\right)\right)$$

➢ This formula is robustly satisfied by all faulty trajectories, and robustly violated by all non-faulty trajectories.

Conceptual Example



➢ Bouncing basketball is a simple example of hybrid systems. The events occur when the ball bounces off the floor, with reset in the velocity state. The entire execution trajectories would consist of the *x(t)* and *v(t)* trajectories and the "bounce" event times.



➢ An observation map, e.g., extracts only the "bounce" event times and discard the rest.

➢ **State observation:** can we reconstruct the state of the system based on the "bounce" timing (assuming the model is known)? **Answer:** Yes (after two bounces).

➢ **Fault diagnosis:** can we deduce whether the ball is properly inflated (ρ=0.8) or under-inflated (ρ=0.1) from the "bounce" timing? **Answer:** Yes (after three bounces).

➢ **Privacy preservation:** Suppose that we do not want to reveal whether the ball was dropped from $x(0) \in [10, 11]$ or $x(0) \in [11, 12]$. Using the "bounce" timing would violate this privacy constraint. However, if instead of the "bounce" times we report a logical predicate of the inter-event time, "short" if the interval is less than 1 unit, and "long" otherwise, we would respect the privacy constraint, while still being able to detect underinflated ball.

Bibliography

- **A.A. Julius**, Trajectory-based theory for hybrid systems, in *Mathematical Control Theory I: Nonlinear and Hybrid Control Systems,* Springer, 2015.
- Y. Deng, A. D'Innocenzo, **A.A. Julius**, Trajectory-based observer for hybrid automata fault diagnosis, in *Proc. IEEE Conf. Decision and Control*, pp. 942-947, Osaka, Japan, 2015.
- Y. Deng, A. D'Innocenzo, M.D. DiBenedetto, S. Di Gennaro, **A.A. Julius**, Verification of hybrid automata diagnosability with measurement uncertainty, *IEEE Trans. Automatic Control*, vol. 61(4), pp. 982-993, 2016.
- Z. Xu, S. Saha, **A.A. Julius**, Provably correct design of observations for fault detection with privacy preservation, submitted to *IEEE Conf. Decision and Control 2017*.

Rensselaer

CISE/CNS/CSR