*Seminar at National Institute of Standards and Technology*

# ANOMALY DETECTION AND FAILURE MITIGATION IN COMPLEX DYNAMICAL SYSTEMS

Asok Ray

**The Pennsylvania State University**
**University Park, PA 16802**
Email: axr2@psu.edu
Tel: (814) 865-6377

*Supported by*
Army Research Office
Grant No. DAAD19-01-1-0646

November 19, 2004

# Outline of the Presentation

## Anomaly Detection in Complex Dynamical Systems

- ❖ **Microstate Information Based on Macroscopic Observables**
  - **Thermodynamic Formalism of Multi-time-scale Nonlinearities**
  - **Symbolic Time Series Analysis of Macroscopic Observables**
  - **Pattern Discovery via Information-theoretic Analysis**
- ❖ **Real-time Experimental Validation on Laboratory Apparatuses**
  - **Active Electronic Circuits and Three-phase Electric Induction Motors**
  - **Multi-Degree-of-Freedom Mechanical Vibration and Chaotic Systems**
  - **Fatigue Damage Testing in Polycrystalline Alloys**

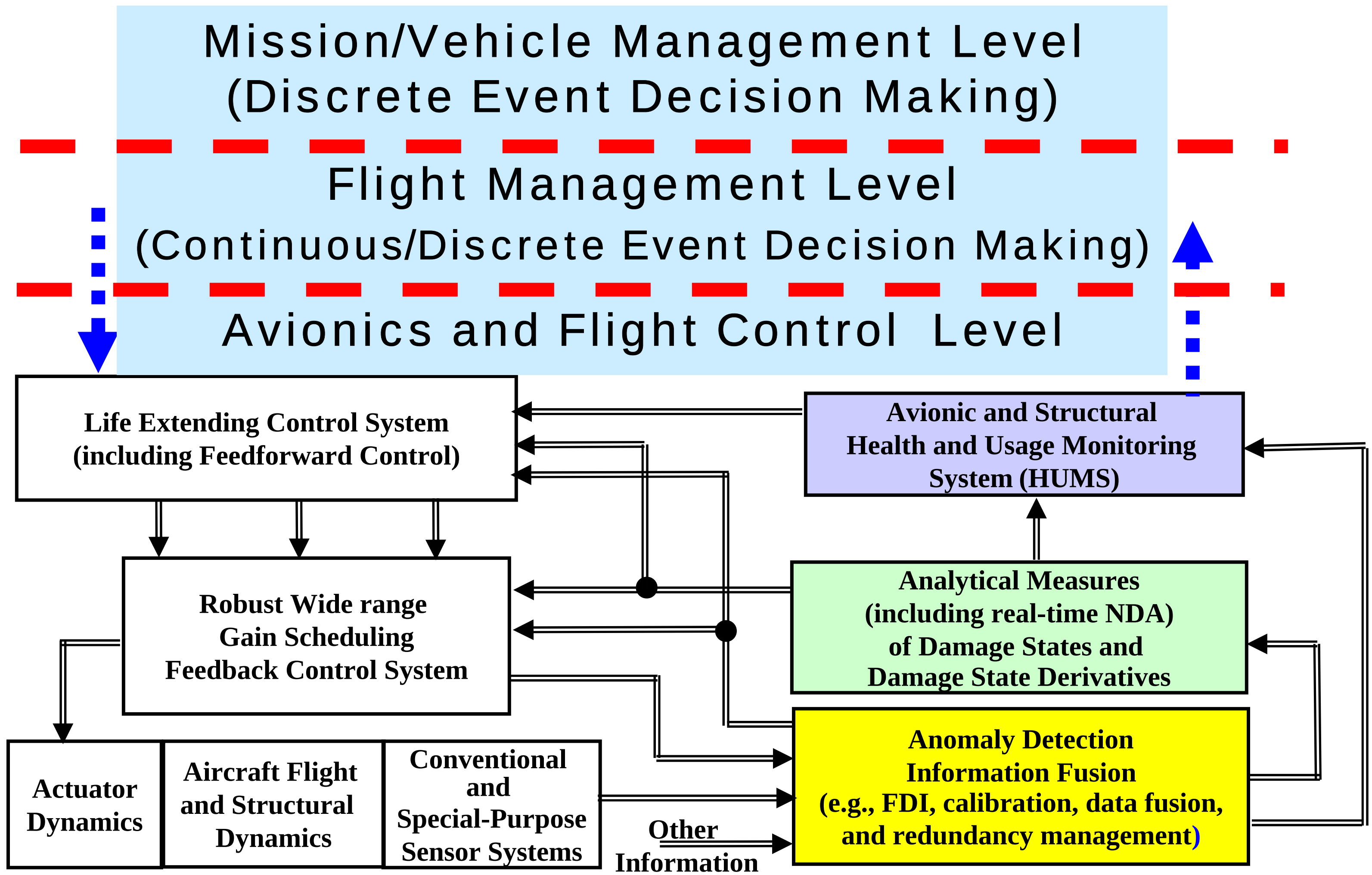
## Discrete Event Supervisory Control for Failure Mitigation

- ❖ Quantitative Measure for Language-based Decision and Control
- ❖ Real-time Identification of Language Measure Parameters
- ❖ Robust and Optimal Control in Language-theoretic Setting

## Future Collaborative Research in Complex Microstructures

- ❖ **Modeling and Control of Hidden Anomalies and their Propagation**
- ❖ Experimentation on Real-time Detection and Mitigation of Malignant Anomalies on a Hardware-in-the-loop Simulation Test Bed

# Intelligent Health Management and Failure Mitigating Control of Aerial Vehicle Systems



# Anomaly Detection and Classification:
# Symbolic Time Series Analysis

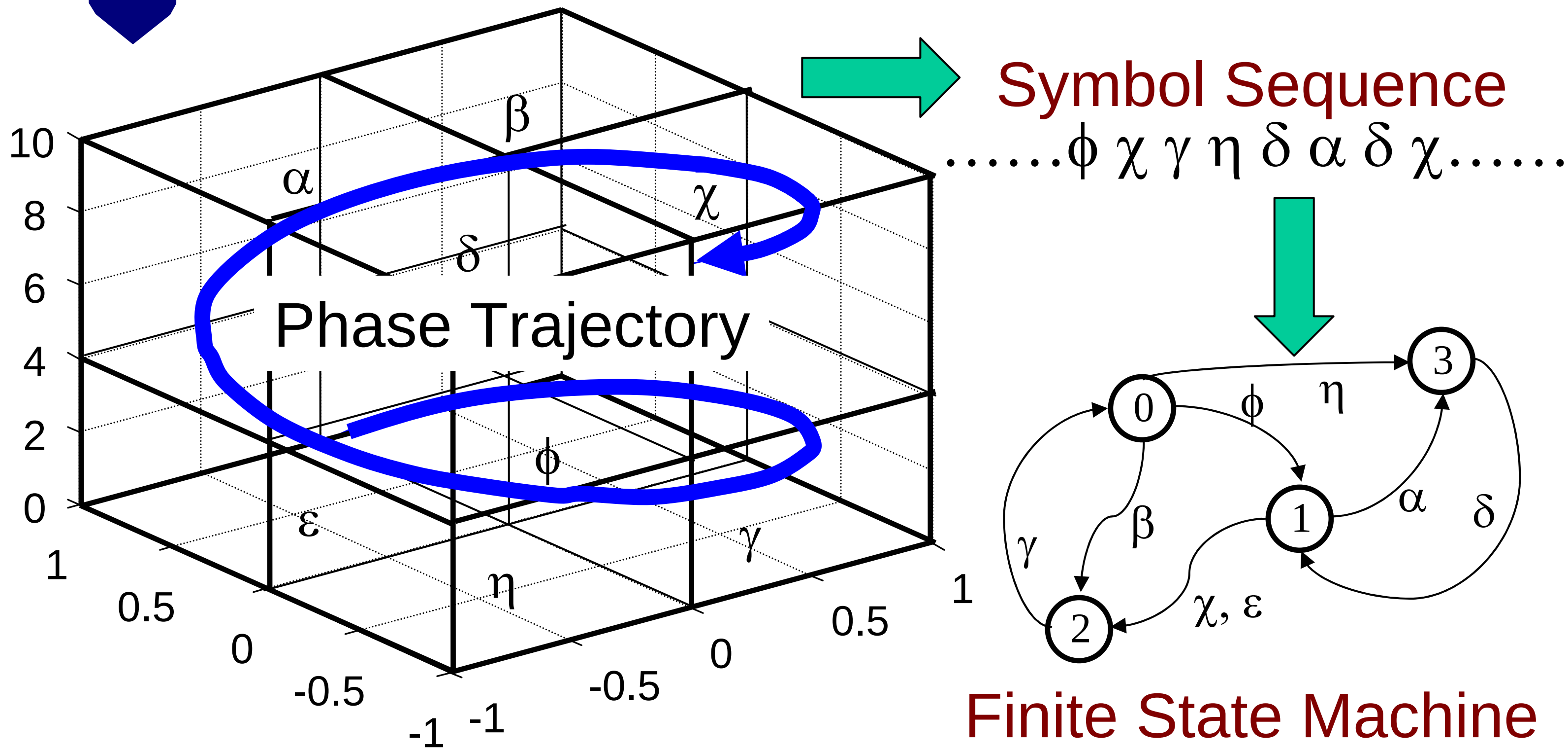
## Multi-Time-Scale Nonlinear Dynamics

- Slow Time Scale: Anomaly Propagation (Non-stationary Statistics)
- Fast Time Scale: Process Response (Stationary Statistics)

## Model-based Statistical Methods

- Modeling with Nonlinear Stochastic Differential Equations
  - **Ito Equation:** $d x_t(\zeta) = \varphi\left(x_t(\zeta), t\right) dt + \gamma_t\left(x_t(\zeta), t\right) d\beta_t(\zeta) \;\; \forall t \geq t_0$
  - **Fokker Planck Equation:** $\frac{\partial p(x,t|y,\tau)}{\partial t} = -\frac{\partial\left[p(x,t|y,\tau)\varphi(x,t)\right]}{\partial x} + \frac{1}{2}\frac{\partial^2\left[p(x,t|y,\tau)\gamma^2(x,t)\right]}{\partial x^2}$
  - **Uncertainties in Model Identification and Loss of Robustness**
- Statistical Mechanical Modeling (Canonical Ensemble Approach)
  - **Symbolic Time Series Analysis**
    - **Small perturbation stimulus**
    - **Self-excited oscillations**
  - **Thermodynamic Formalism and Information Theory**
  - **Hidden Markov Modeling (HMM) and Shift Spaces**

# Notion of Symbolic Dynamics




- Discretization of the Dynamical System in Space and Time
- Representation of Trajectories as Sequences of Symbols

# State Machine Construction

## D-Markov Machine

### State to state transitions from a symbol sequence

- **alphabet size = |A|** $\in \mathbb{N}$
- **window size = D** $\in \{0\} \bigcup \mathbb{N}$
- **kth word (state)=** $W^k = w_0^k w_1^k \ldots w_{D-1}^k$
- **kth word value =** $\sum_{i=0}^{D-1} (w_i^k) A^{D-1-i}$

$$w_0^k w_1^k \ldots\ldots\ldots\ldots w_{D-1}^k$$

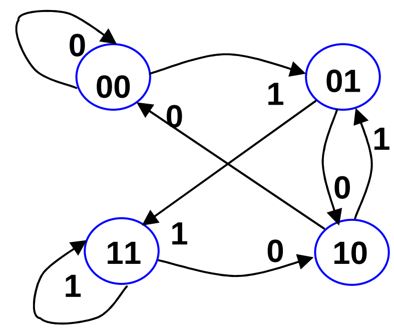
$$w_0^k \leftarrow w_0^{k+1} w_1^{k+1} \ldots\ldots\ldots w_{D-1}^{k+1} \leftarrow w_{D-1}^{k+1}$$

Out In

**Example**

1-D Ising (Spin-1/2) Model

Nearest Neighbor Interactions




$$\begin{bmatrix} p_{00} & 1-p_{00} & 0 & 0 \\ 0 & 0 & p_{01} & 1-p_{01} \\ p_{10} & 1-p_{10} & 0 & 0 \\ 0 & 0 & p_{11} & 1-p_{11} \end{bmatrix}$$

$|A|=2$; $D=2$; $A^D = 4$

- **(k+1)th word =** $W^{k+1} = w_0^{k+1} w_1^{k+1} \ldots w_{D-1}^{k+1}$
- **(k+1)th word value =** $\left(W^k - w_0^k A^{D-1}\right) A + w_{D-1}^{k+1}$

# State Space Construction via D-Markov Machine

- **Computationally efficient for anomaly measure**
- **Fixed depth D and alphabet size A**
- **Only the state transition probabilities to be determined based on symbol strings derived from time series data or wavelet-transformed data**
- **States represented by an equivalence class of strings whose last D strings are identical**

PENNSTATE 1855

# Anomaly Measure

## Based on the D-Markov Machine

### State Transition Matrix Construction

- **Banded structure**
- **Separation into irreducible subsystems**
- **Stationary state probability vector**
- **Information on the dynamical system characteristics**
  - **Chaotic motion, period doubling, and bifurcation**

### State Probability Vector

- **Reference Point: Nominal Condition $\mathbf{p}(\tau_o)$**
- **Epochs $\{\tau_k\}$ of Slow Time Scale $\{\mathbf{p}(\tau_k)\}$**

### Anomaly Measure at Slow-Time Epochs

$$\mathcal{M}(\tau_k; \tau_o) = \mathbf{d}(\mathbf{p}(\tau_k), \mathbf{p}(\tau_o))$$

# Comparison of Epsilon Machine and D-Markov Machine

## Epsilon Machine [Santa Fe Institute]

- A priori unknown machine structure
- Optimal prediction of the symbol process
- Maximization of mutual information
  (i.e., minimization of conditional entropy)

  I[X;Y] = H[X] – H[X|Y]
- Analogous to the class of ***Sofic Shifts*** in Shift Spaces

## D-Markov Machine

- A priori known machine structure
  (Fixed order fixed structure with given **|A|** and **D**)
- Excess states yielding redundant reducible matrices
  (Perron-Frobenius Theorem)
- Suboptimal prediction of the symbol process
- Analogous to the class of ***Finite-type Shifts*** in Shift Spaces

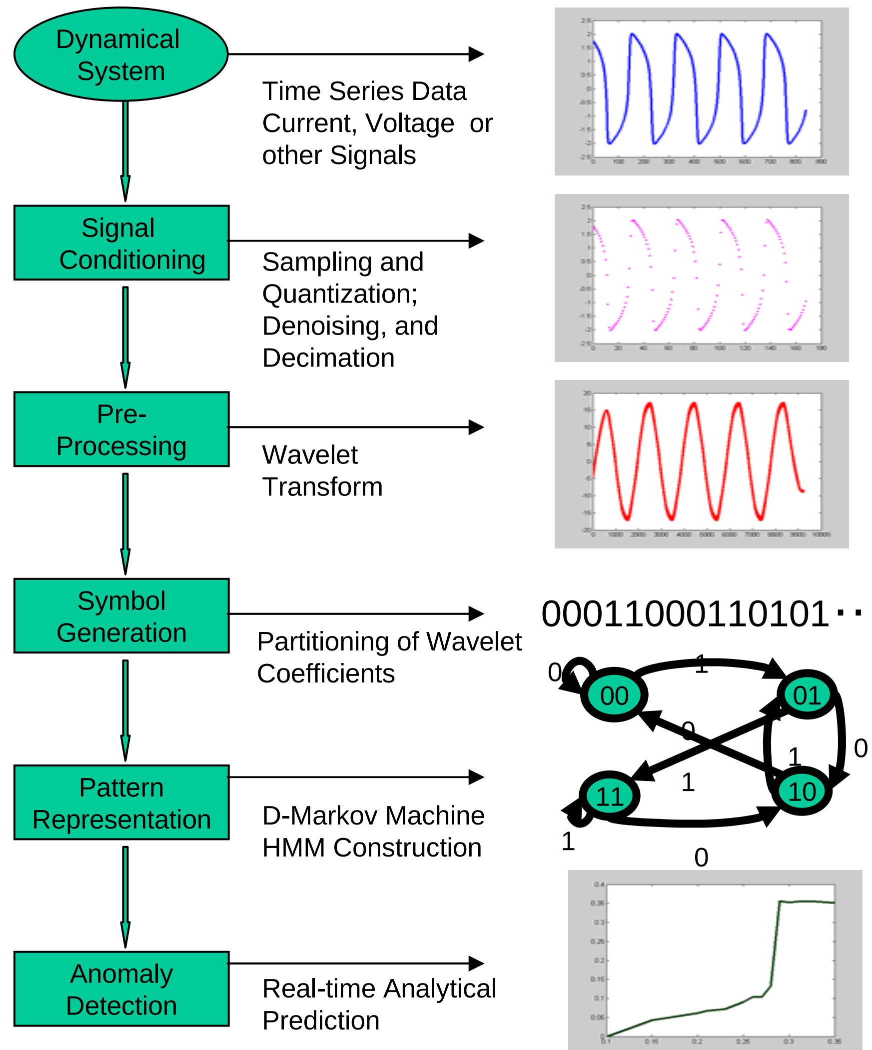
# Anomaly Detection Procedure

- **Forward (Analysis) Problem:**
  - **Characterization of system dynamical behavior**
    - **Parametric and non-parametric anomalies**
  - **Evolution of the grammar in the system dynamics**
    - **Representation of dynamical behavior as formal languages**
    - **Thermodynamic formalism of anomaly measure**
- **Inverse (Synthesis) Problem**
  - **Estimation of feasible ranges of anomalies**
    - **Fusion of information generated from responses under several stimuli chosen in the forward problem**

# Summary of Anomaly Detection Procedure

## Anomaly Detection and Classification

- Signal Conditioning and Decimation
  - Denoising
  - Embedding
- Symbol Sequence Generation
  - Phase space partitioning
  - Wavelet space partitioning
- Markov Modeling of Symbol Dynamics
  - Epsilon machine (sofic shift)
  - D-Markov machine (finite type shift)
- Thermodynamic Formalism of Generated Information



# Externally Stimulated Duffing Equation

## with a single slowly varying parametric anomaly

## Governing Equations:

$$\frac{d^2 y(t)}{dt^2} + \theta(t_s)\frac{dy(t)}{dt} + y(t) + \alpha\, y^3(t) = ACos(\omega t) \qquad t \in [t_o, \infty)$$

## Random Initial Conditions

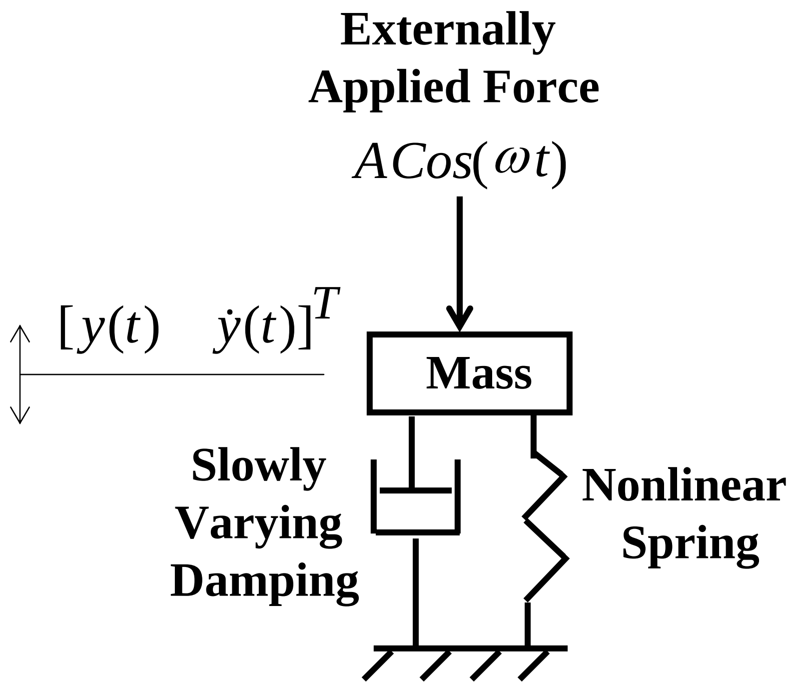
$$[y(t_o) \quad \dot{y}(t_o)]^T \in B_\delta(\underline{0})$$

$t$ *fast time* $t \in [t_o, \infty)$

$t_s$ *slow time*

*Parameters:*

$\alpha = 1; \delta = 0.01; A = 22; \omega = 5.0$

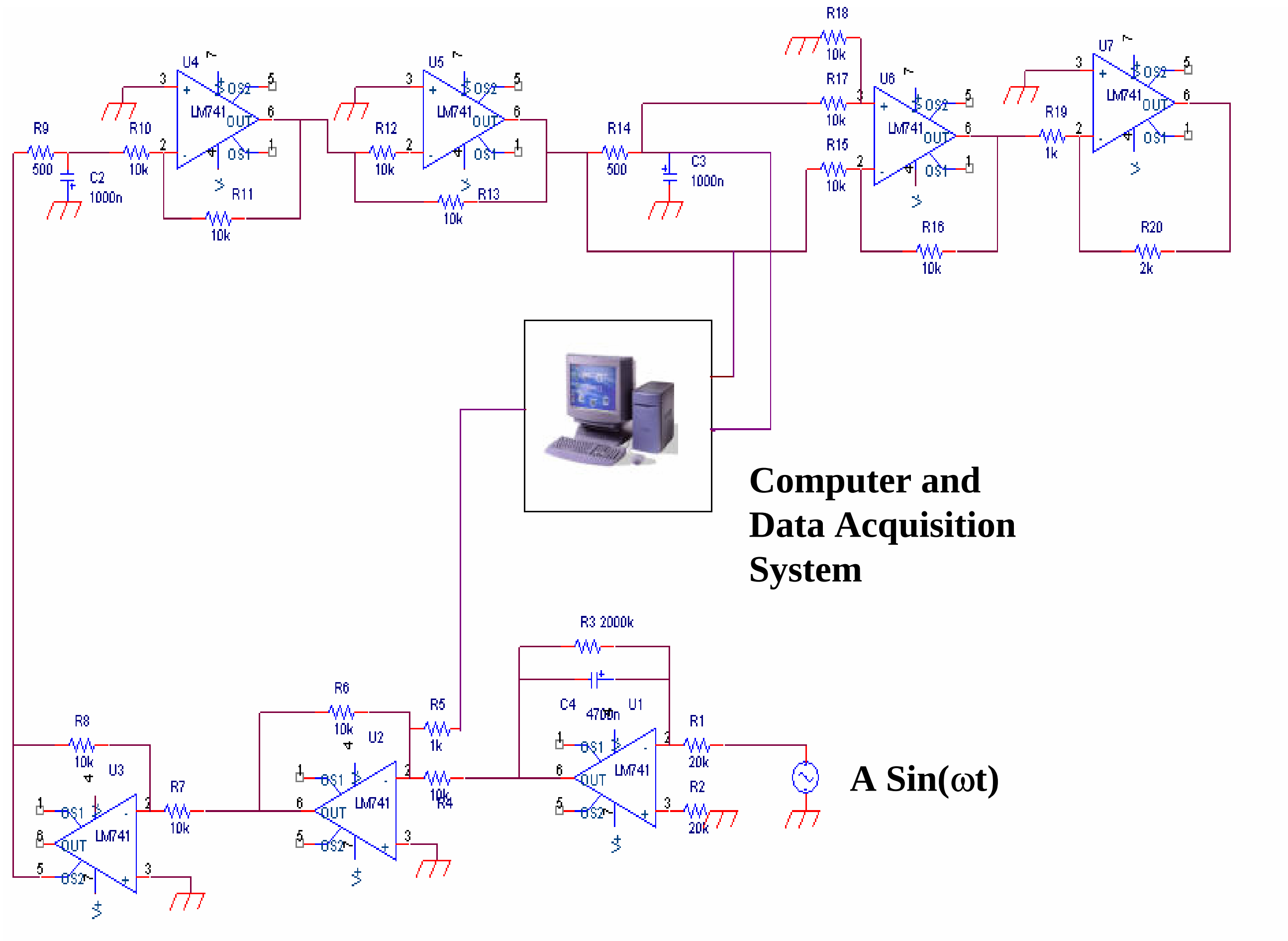


PENNSTATE 1855

# Electromechanical Systems Laboratory

## Anomaly Detection Apparatus for Hybrid Electronic Circuits

### *Externally Stimulated Duffing Equation*

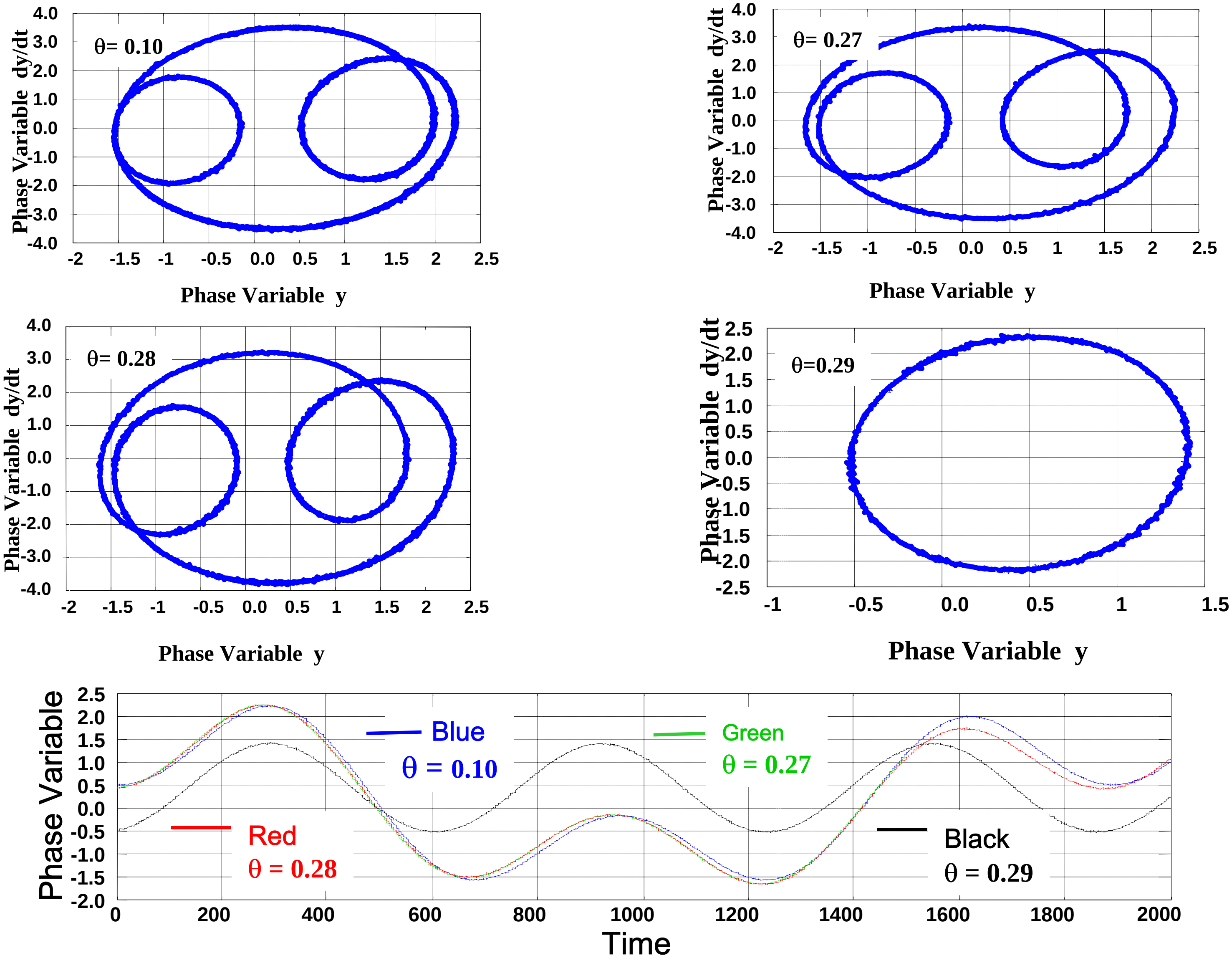


PENNSTATE

1855

# Electromechanical Systems Laboratory

## Phase-plane Plots under Nominal and Anomalous Conditions



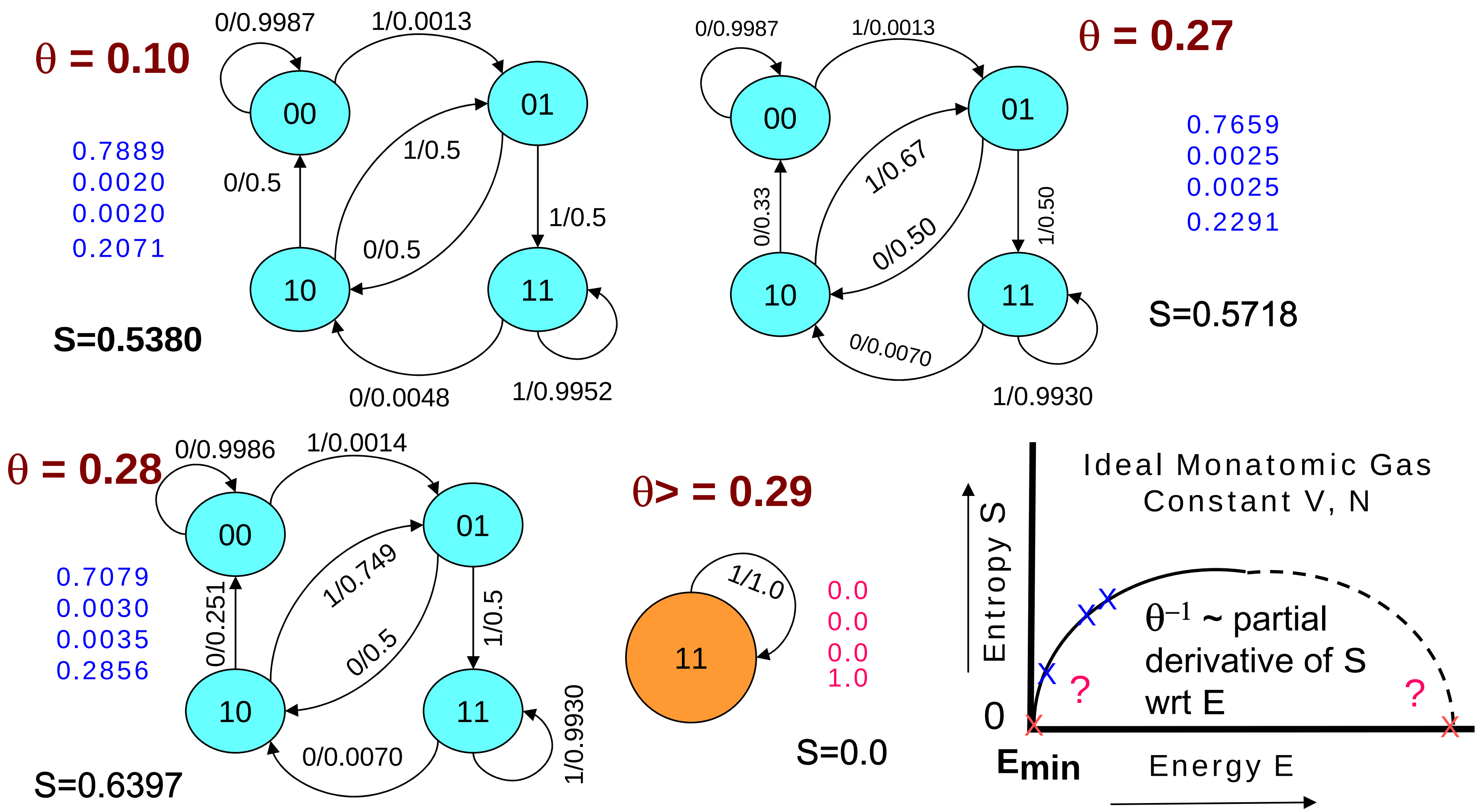
# Construction of Finite State Machines

PENNSTATE 1855

## Interpretation using

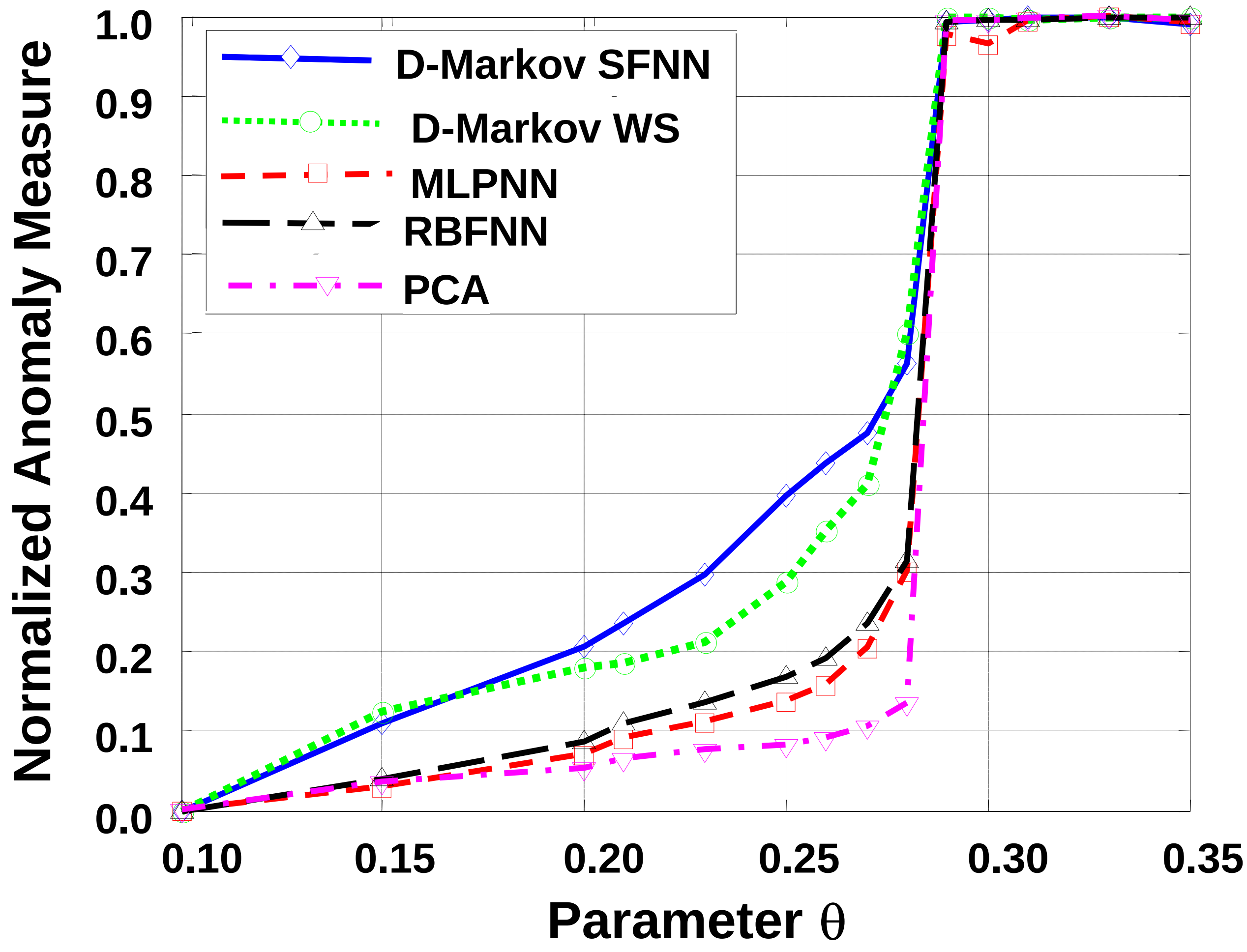
## Thermodynamic Formalism

$|A|=2; \;\; D=2; \;\; A^D = 4$



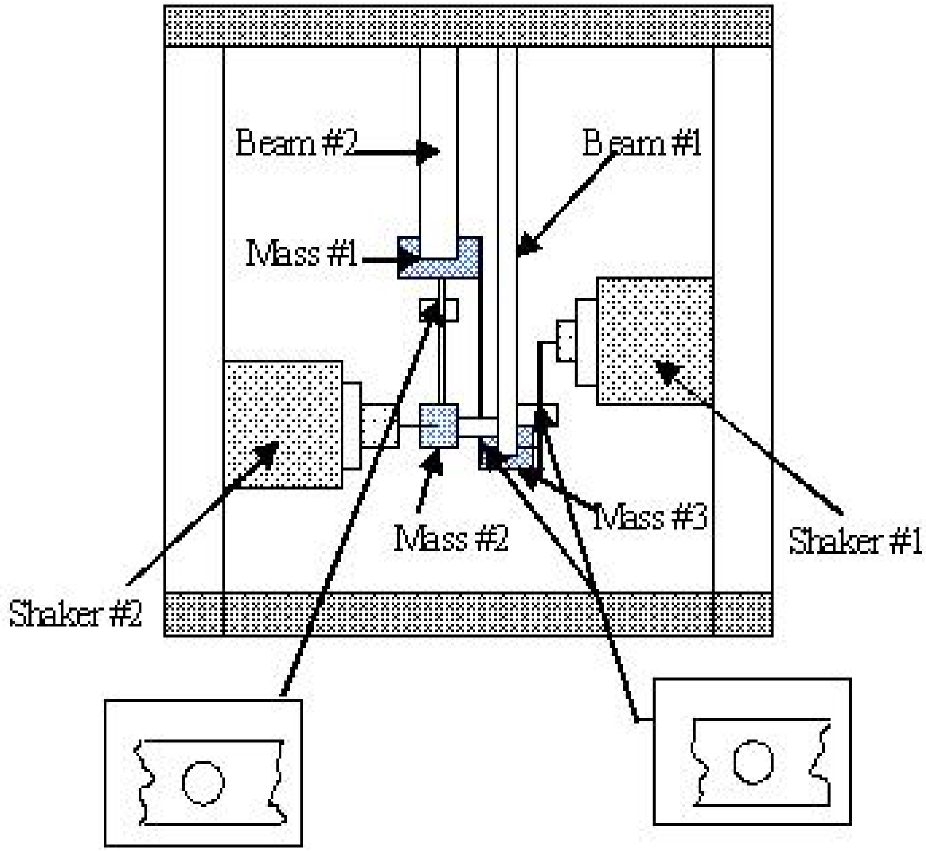
# Electronic Circuit Apparatus

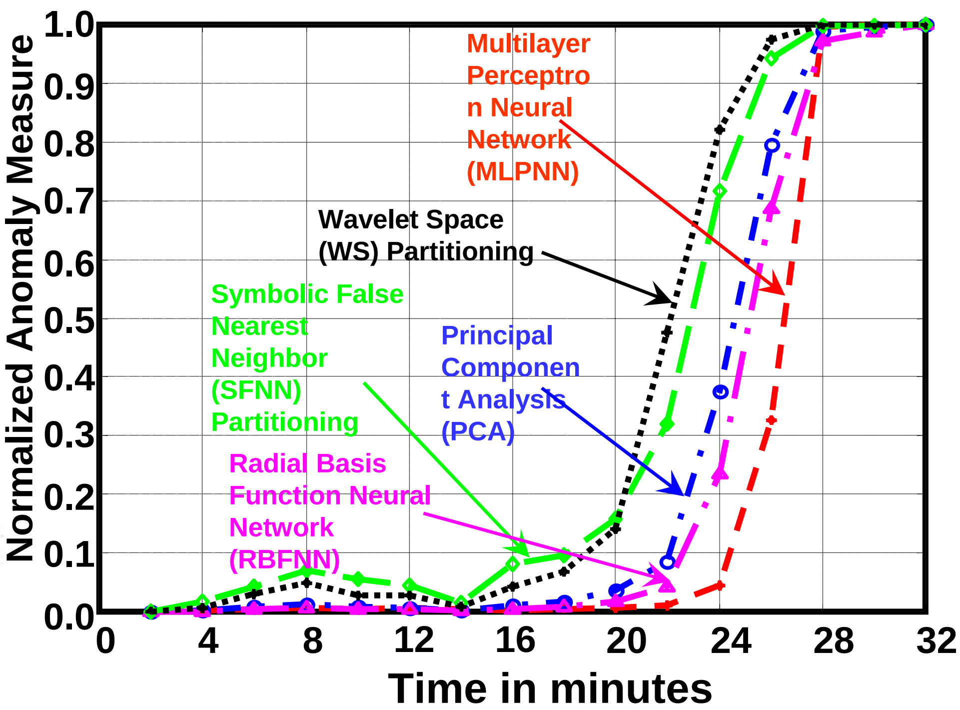
## Sensitivity of the Detection Algorithm to the Anomalous Parameter $\theta$



# Electromechanical Systems Laboratory

## Anomaly Detection Apparatus for Mechanical Vibration Systems



PENNSTATE Electromechanical Systems Laboratory

1855

# Fatigue Test Apparatus for Damage Sensing in Ductile Alloys

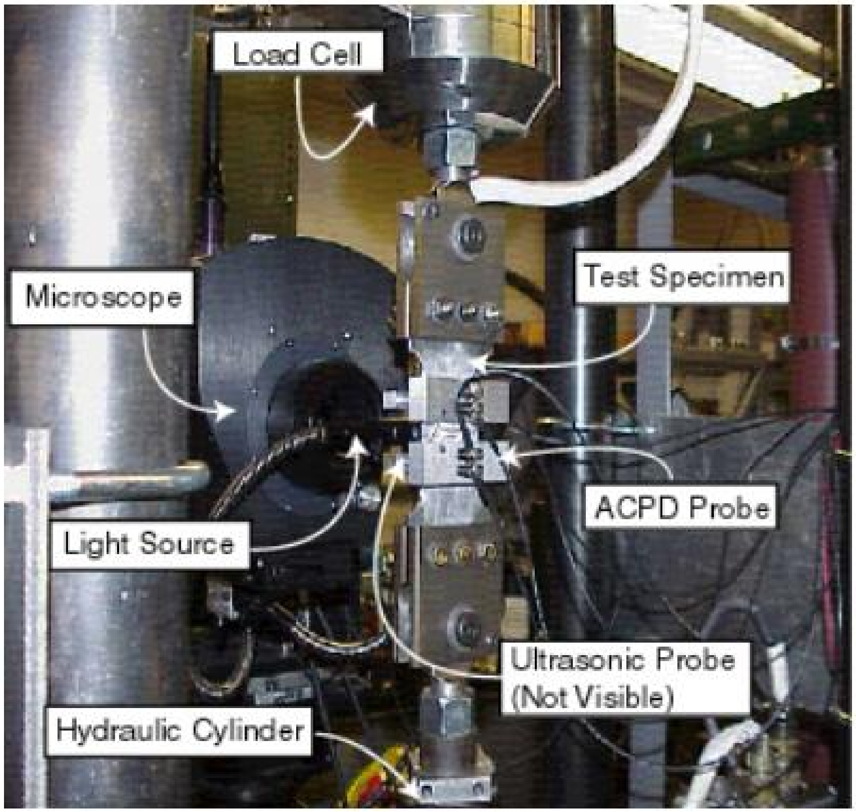
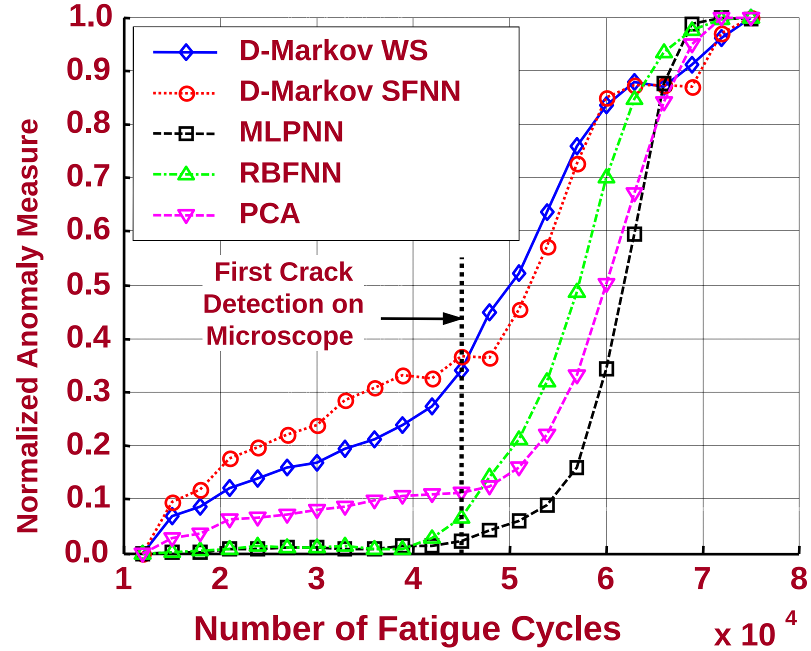



Figure A-1 Fatigue Damage Apparatus



# Anomaly Detection

## Symbolic Time Series Analysis

**A. Ray, "Symbolic Dynamic Analysis of Complex Systems for Anomaly Detection," *Signal Processing*, Vol. 84, No. 7, July 2004, pp. 1115-1130.**

## Advantages

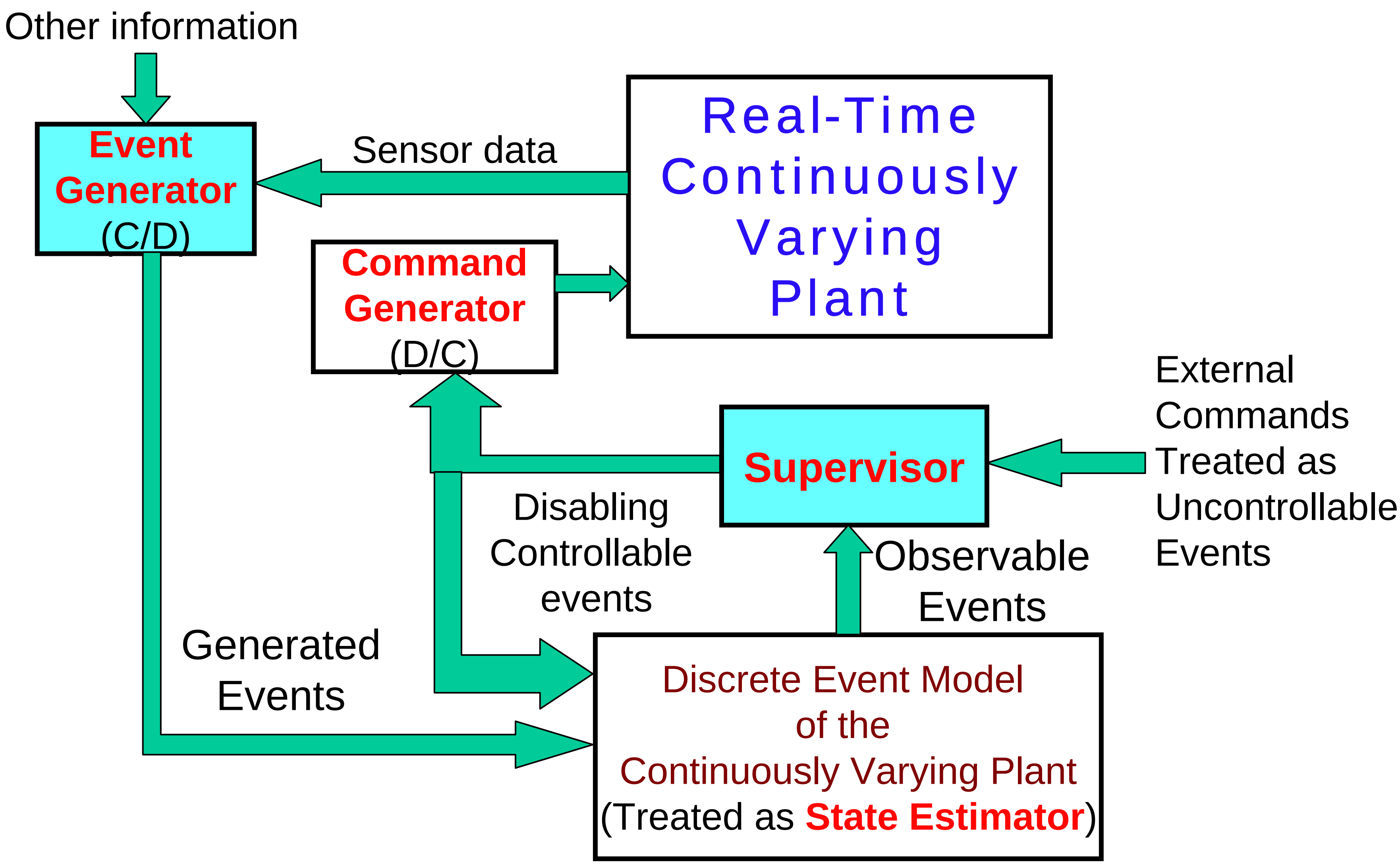
- Foundations on fundamental principles of physics and mathematics
- Quantitative measure as opposed to qualitative measure
- Robustness to measurement noise and spurious signal distortion
- Sensitivity to signal distortion due to nonlinearity and nonstationarity
- Adaptability to low-resolution sensing
- Applicability to real-time anomaly detection

## (Near-term) Disadvantages

- Requirement for advanced knowledge to understand the basics
- Need for much theoretical and experimental research
- Seemingly counter-intuitive to inadequately trained technical personnel

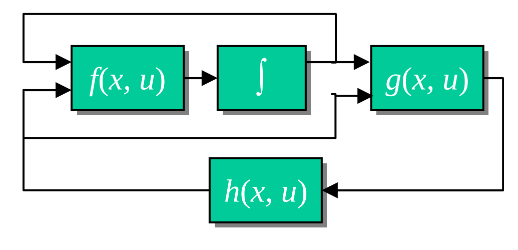
# Quantitative Measure for Discrete Event Supervisory Control

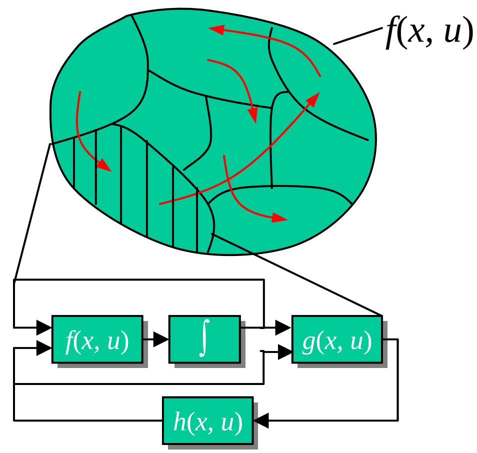
# Discrete Event Supervisory (DES) Control of Continuous Plants

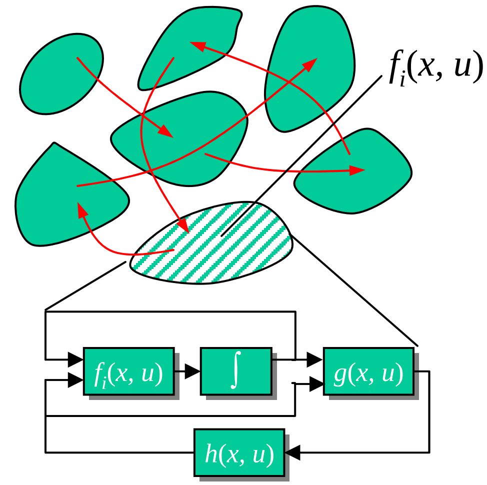


# Modeling of Discrete Event Supervisory Control Systems

PENNSTATE 1855

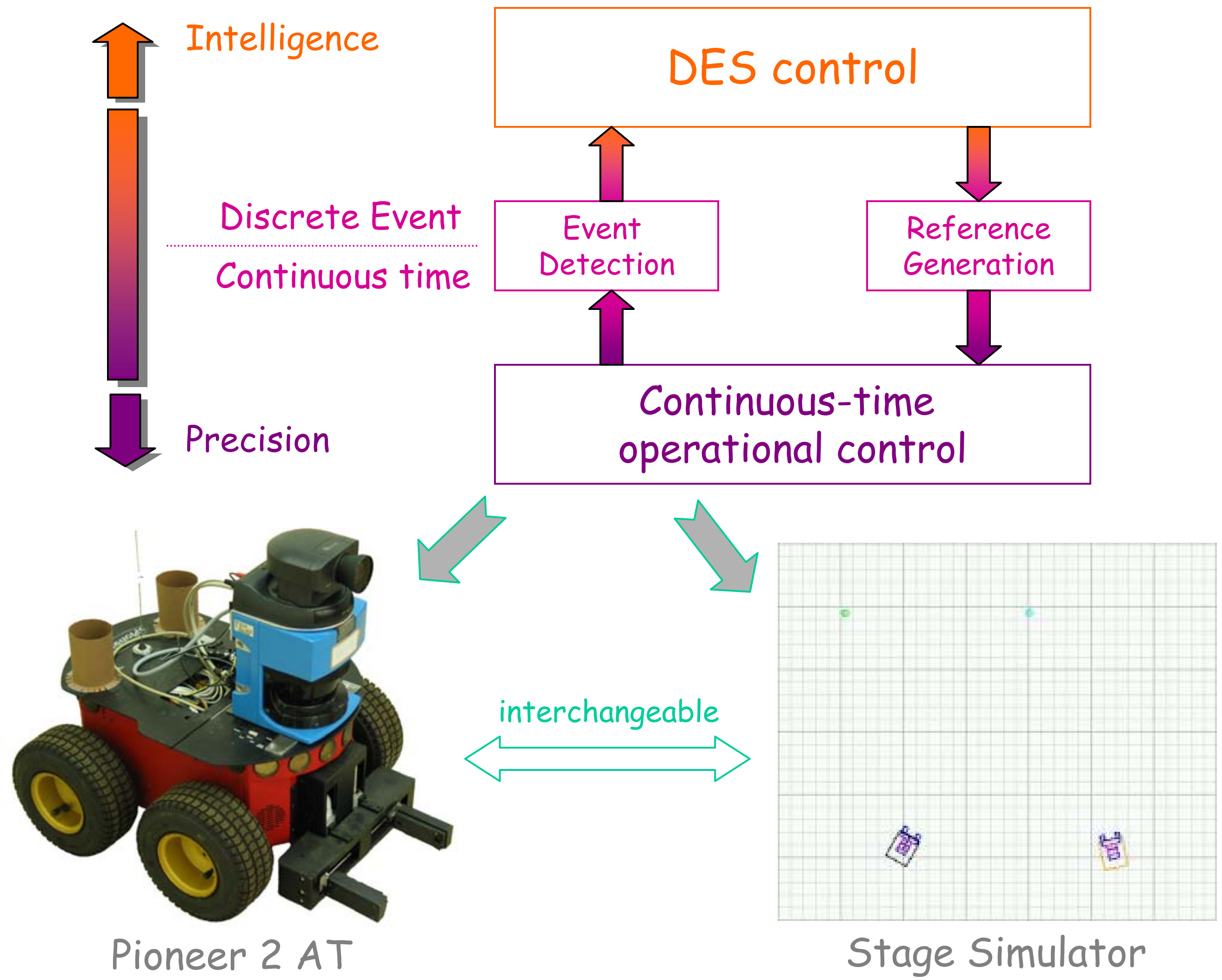



| | | |
|---|---|---|
| $\mathcal{C} = (\mathcal{X}, U, f)$<br>$\mathcal{X} \subseteq \mathbb{R}^n$ *is the state space*<br>$U \subseteq \mathbb{R}^m$ *is the system input,*<br>$f : \mathcal{X} \times U \to \mathbb{R}^n$ | $G = (Q, \Sigma, \delta, q_0, Q_m)$<br>$Q$ *is the discrete state space*<br>$\Sigma$ *is the set of events*<br>$\delta : Q \times \Sigma \to Q$<br>$q_0 \in Q$ *is the initial state* | $\mathcal{H} = (Q, X, \Sigma, \delta, \mathcal{X}_0, I, f)$<br>$X \subset \mathbb{R}^n$ *is a bounded continuous state space*<br>$\delta : (Q \times X) \times \Sigma \times (Q \times X) \to 2^X$<br>$\mathcal{X}_0 \subseteq \mathcal{X} = (Q \times X)$ *the set of initial states*<br>$f : Q \to (X \times \mathbb{R}^m \to \mathbb{R}^n)$<br>$I : Q \to 2^X$ *invariant conditions* |
| $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}, t), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0$<br>$\boldsymbol{y} = g(\boldsymbol{x}, \boldsymbol{u}, t)$ | $\phi(k+1) \in \delta(\phi(k), \sigma)$<br>$\phi(0) = q_0 \quad k \in \mathbb{N} \cup \{0\}$<br>$L(G) = \{s \in \Sigma^* \mid \delta(q_0, s) \in Q\}$ | |
| | Decoupling: continuous evolutions and discrete transitions | Coupling: continuous evolutions and discrete transitions |
| Continuously Varying Systems | Discrete Event Systems | Hybrid Systems |

# Behavior-based Robotic System
## DES Control Architecture



# Signed Real Measure of Regular Languages

## Salient Features

- **Quantitative Measure of $*$-Languages**
  - **Finite alphabet** $\Sigma \Longrightarrow \Sigma^*$ cardinality $N_0$
  - **Language Measure** $\mu: 2^{\Sigma^*} \longrightarrow R$
- **Language Metric $|\mu|$**
  - **Total variation of the signed real measure**
  - **(Real positive) distance between two languages**
- **Vector Space of Formal Languages**
  - **Infinite-dimensional space**
  - **Galois field GF(2)**
  - **Vector addition operator - Exclusive-OR**
- **Applications of the Language Measure for Failure Mitigation**
  - **Robust and optimal control of discrete-event systems**
  - **Anomaly quantification, classification, and mitigation**

# State-Based Partitioning

- The set $Q_m$ of marked states is partitioned as:

$$\mathbf{Q_m} = \mathbf{Q_m^+} \cup \mathbf{Q_m^-}; \ \mathbf{Q_m^+} \cap \mathbf{Q_m^-} = \varnothing$$

By using

- **Myhill-Nerode Theorem**
- **Hahn Decomposition Theorem**

where $\mathbf{Q_m^+}$ is the set of good marked states (positive measure)

$\mathbf{Q_m^-}$ is the set of bad marked states (negative measure)

# Main Result:

## Signed Real Measure of Regular Languages

A. Ray, V. V. Phoha and S. Phoha, **QUANTITATIVE MEASURE FOR DISCRETE EVENT SUPERVISORY CONTROL: Theory and Applications,** ***Springer*, New York, 2004.** ISBN 0-387-02108-6

**Lemma:** The regular expressions $L_i \equiv L(G_i),\ i \in I \equiv \{1,2,\cdots,n\}$ can be expressed by the following set of symbolic equations:

$$L_i = \sum_j \sigma_i^j L_j + \vartheta_i \quad \forall i \in I \qquad \text{where} \quad \vartheta_i = \begin{cases} \varepsilon & \text{if } q_i \in Q_m \\ \varnothing & \text{otherwise} \end{cases}$$

**Theorem:** The language measure of the regular expressions $L_i,\ i \in \{1,2,\cdots,n\}$ is given by the unique solution of the following set of algebraic equations: $\mu_i = \sum_j \pi_{ij} \mu_j + \chi_j \quad \forall i \in I$

In vector notation, the system $\mu = \Pi\mu + X$ has a unique solution:

$$\boxed{\mu = [I - \Pi]^{-1} X}$$

**Remark:** $[I - \Pi]^{-1}$ exists and is bounded above by $1/\|\Pi\|_\infty$

# Algorithm for the Unconstrained Optimal Control Law

$$\Pi^0 = \begin{bmatrix} \pi^0{}_{11} & \pi^0{}_{12} & \cdots & \pi^0{}_{1n} \\ \pi^0{}_{21} & \pi^0{}_{22} & \cdots & \pi^0{}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \pi^0{}_{n1} & \pi^0{}_{n2} & \cdots & \pi^0{}_{nn} \end{bmatrix} \qquad \overline{\mu}^0 = \begin{pmatrix} \mu_1{}^0 \\ \mu_2{}^0 \\ \vdots \\ \mu_n{}^0 \end{pmatrix} = \left(I - \Pi^0\right)^{-1} \overline{\chi}$$

***Given the information on the plant model***

$$G_i = (Q, \Sigma, \delta, q_i, Q_m)$$

***along with the state transition cost matrix*** $\Pi$ ***and characteristic vector*** $\overline{\chi}$***, the unconstrained optimal control maximizes the language measure by deleting some of the "bad" strings so that optimality of the supervised plant sublanguage is achieved***

# Supporting Theorems

PENNSTATE 1855

A. Ray, J. Fu and C.M. Lagoa, "**Optimal Supervisory Control of Finite State Automata**," *International Journal of Control*, Vol. 77, No. 12, August 2004, pp. 1083-1100.

Theorem #1 (*Monotonicity*) : Disabling the controllable events leading to states with negative (positive) performance does decrease (increase) supervised plant performance.

Theorem #2 (*Monotonicity*) : Enabling the controllable event(s) leading to states with non-negative performance does not decrease the performance for any state.

Theorem #3 (*Global Performance*): The controller at the termination of the algorithm is the global optimal controller in terms of supervised plant performance.

Theorem #4 (*Computational Complexity*): The optimal control law is solved in at most $n$ steps and each step requires a solution of $n$ linear algebraic equations where $n$ is the number of states of the plant model. Therefore, the computational complexity for synthesis of the optimal control algorithm of a polynomial order in $n$.

# Language-Measure-Theoretic Discrete-Event Supervisory Control

**A. Ray, V. V. Phoha and S. Phoha, QUANTITATIVE MEASURE FOR DISCRETE EVENT SUPERVISORY CONTROL: Theory and Applications,** ***Springer*****, New York, 2004.** ISBN 0-387-02108-6

## Advantages

- Foundations on principles of automata theory and functional analysis
- Quantitative measure as opposed to qualitative measure
- Robustness to measurement noise and spurious signal distortion
- Capability for emulation of human reasoning in a quantitative way
- Adaptability to low-resolution sensing
- Applicability to real-time decision-making at multiple time scales

## Disadvantages

- Potential source of instability under switching actions
- Need for much theoretical and experimental research
- Requirement for advanced knowledge to understand the basics

# Future Collaborative Research in Complex Microstructures

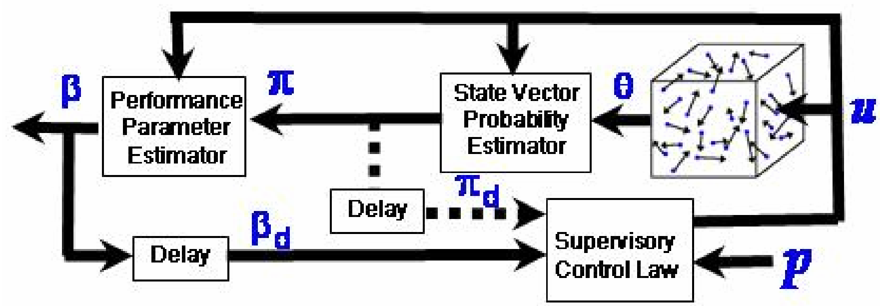
PENNSTATE

1855

# Modeling and Control of Hidden Anomalies and their Propagation

## Problem Definition

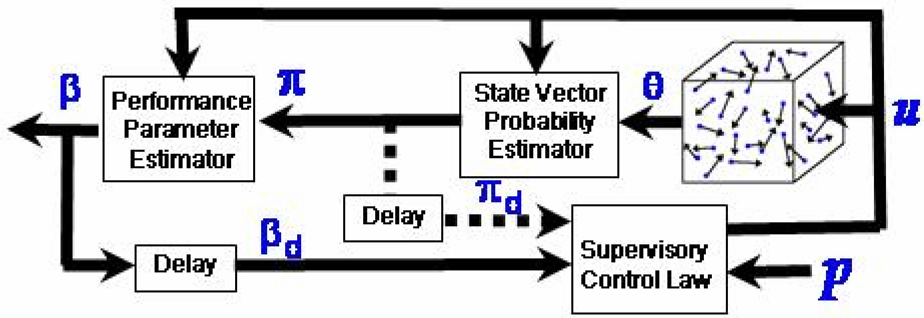
- Let the time series of (macroscopic) measurement(s) $\theta$ be available.
- The **first problem** is to estimate the unobservable performance parameter(s) $\beta$ (e.g., damage states and derivatives).
- The **second problem** is to control the microstates via manipulation of macroscopically controllable inputs $u$ to satisfy desired performance specifications $p$.




## Proposed Solution

- Construction of a canonical-ensemble model with the state probability vector $\pi(\theta, u)$ of the unobservable phenomena that are macroscopically controlled by inputs $u$ through usage of the time series data and a microstructural model, such as the OOF of NIST.
- Formulation of constitutive equations for the unobservable parameters $\beta(\pi, u)$ that are indicator(s) of the internal microstates and control laws $u(\beta_d, \pi_d, p)$ to satisfy desired performance specifications (e.g., remaining service life and reliability)

# Experimentation for Real-time Detection and Mitigation of Malignant Anomalies




- ❑ Experimental Validation of the Novel Constitutive Relations
  - ❖ **Special-purpose Fatigue Testing Machine at Penn State**
  - ❖ **Object Oriented Finite-element (OOF) Modeling Package at NIST**

- ❑ Experimental Validation of the Supervisory Control Concept
  - ❖ **Special-purpose Multi-degree-of-freedom Machine at Penn State**
  - ❖ **Control of the unobservable parameters, such as plastic zone size, that are indicator(s) of the internal microstates**
  - ❖ **Discrete Event Supervisory Control at the Upper level**
    - **Derived parameter(s) $\beta$ (e.g., damage states and their derivatives) to provide the input event sequence to the supervisory control module at the upper level**
    - **Supervisor command(s) to provide the control inputs $u$, such as shaft torque, to the test apparatus to satisfy the desired performance specifications $p$**

# Questions & Suggestions