A Framework for Software Health Management Using Bayesian Statistics

Yuning He
NASA Ames Research Center
Johann Schumann
KBR

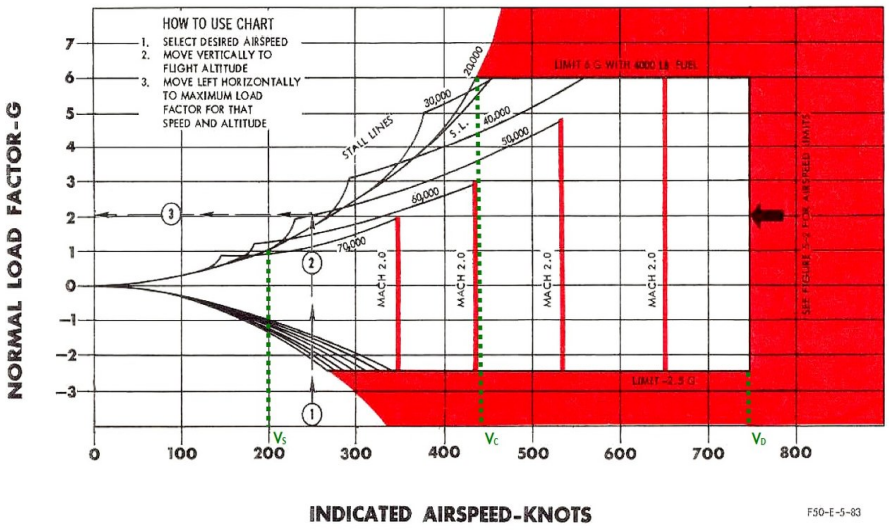
Introduction

- Large and complex software systems are becoming ubiquitous and increasingly critical
  - Cyberphysical systems (automotive, aerospace, medical)
  - Business systems
- Software systems often part of a large SW Ecology

How can we assure software health of such a software system?

A Complex System

can work safely only in certain regions of its state space



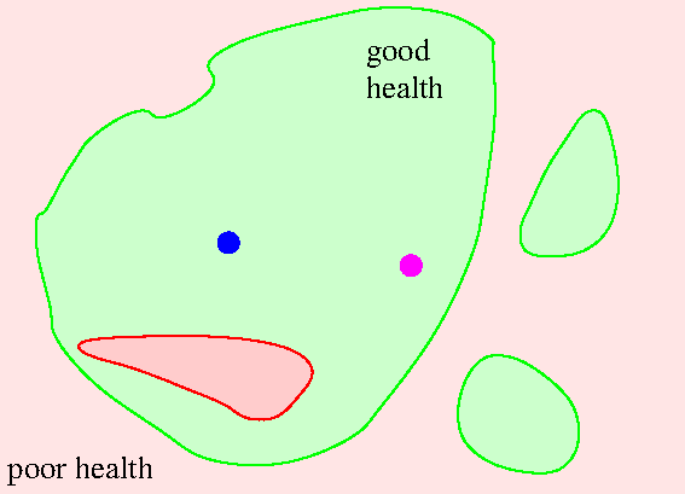
- Flight envelope of an aircraft [wikipedia]

- **Safety Boundaries**
- are in a high-dimensional space,
- can be nonlinear,
- can be hard to estimate,
- can change when the system changes or adapts

A Complex *Software* System

can work safely only in certain regions of its state space

- SW or SECO can be modeled as a continous system with a high-dimensional state space
  - CPU load, % available memory, % lost packages, ...
- Healthy regions are enclosed by boundaries

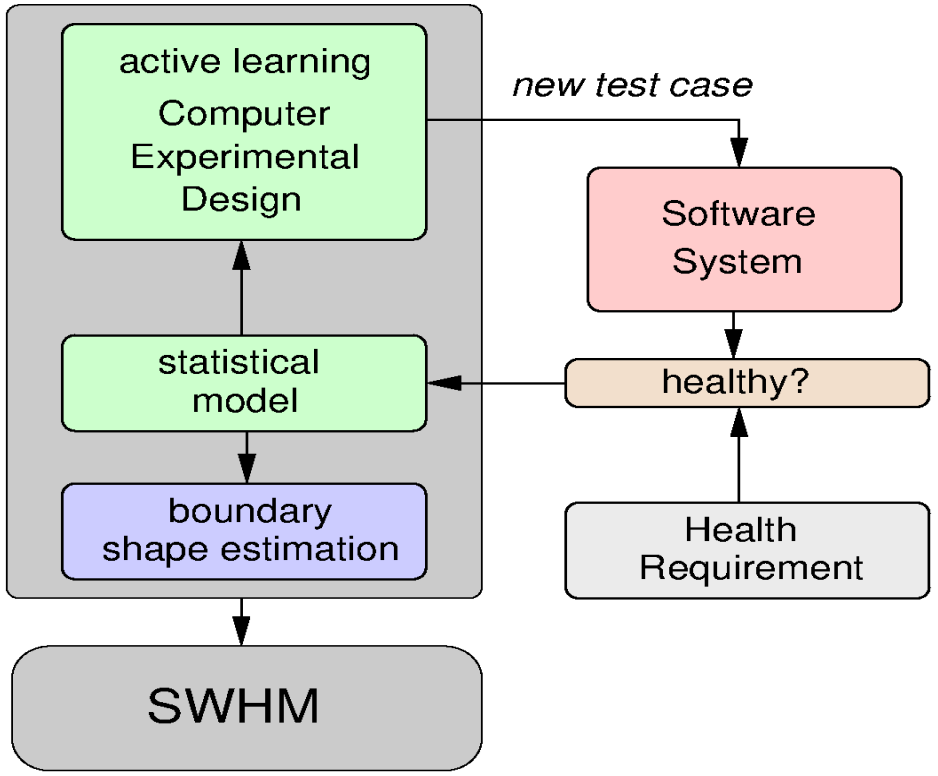


Knowledge about location and shapes of boundaries is important for

- Design
- Verification and Validation
- System Operations

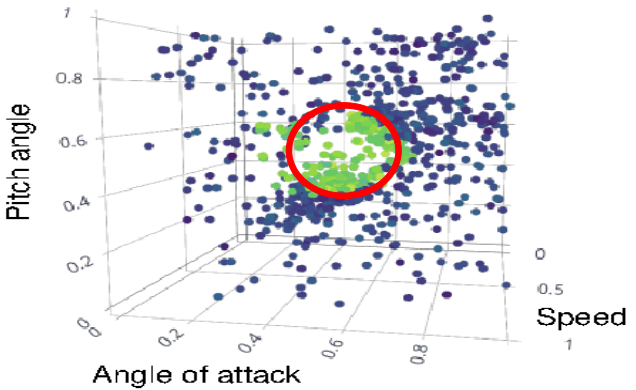
Testing Framework

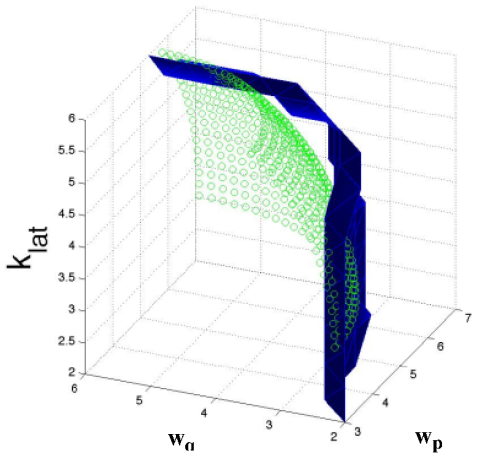
Hierarchical Bayesian statistical modeling with Active Learning in Computer Experiment Design

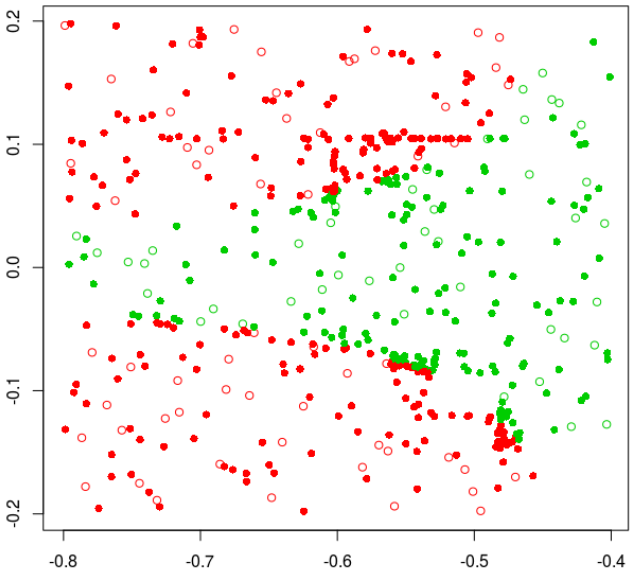


- Generate test cases to find regions where SW is not healthy as defined by Health Requirements
- Active learning selects new test cases close to the estimated boundaries for higher efficiency
- Geometric boundary shape estimation for feedback to designer and SWHM

Examples



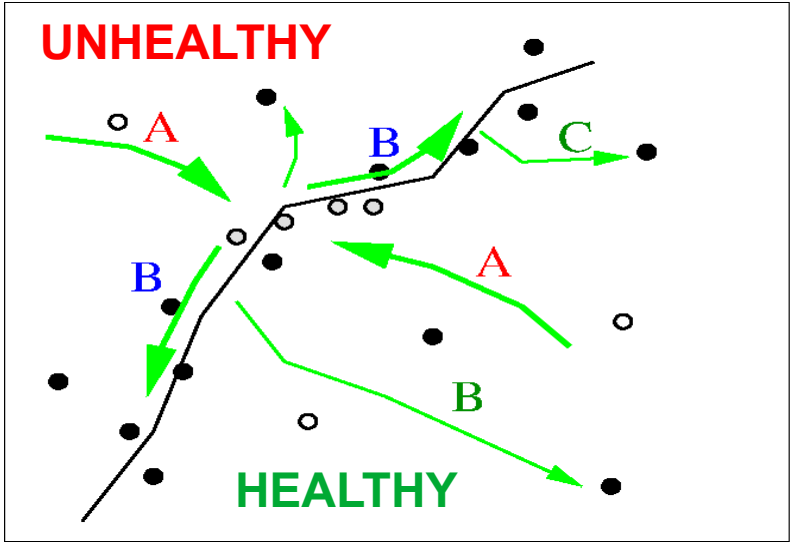




Projection of safety boundaries into 2D and 3D parameter space

Boundary-aware metric

$$E[I(x)] = -\int_{0.5-\alpha s(x)}^{0.5+\alpha s(x)} (y-\hat{y}(x))^2 \phi\left(\frac{y-\hat{y}(x)}{\sigma(x)}\right) dy$$
$$+2(\hat{y}-0.5)\sigma^2(x)\left[\phi\left(\frac{0.5-\hat{y}(x)}{\sigma(x)}+\alpha\right)-\phi\left(\frac{0.5-\hat{y}(x)}{\sigma(x)}-\alpha\right)\right]$$
$$+(\alpha^2\sigma^2(x)-(\hat{y}(x)-0.5)^2)\left[\Phi\left(\frac{0.5-\hat{y}(x)}{\sigma(x)}+\alpha\right)-\Phi\left(\frac{0.5-\hat{y}(x)}{\sigma(x)}-\alpha\right)\right.$$



- New test cases proposed that are

A) Close to the suspected boundary

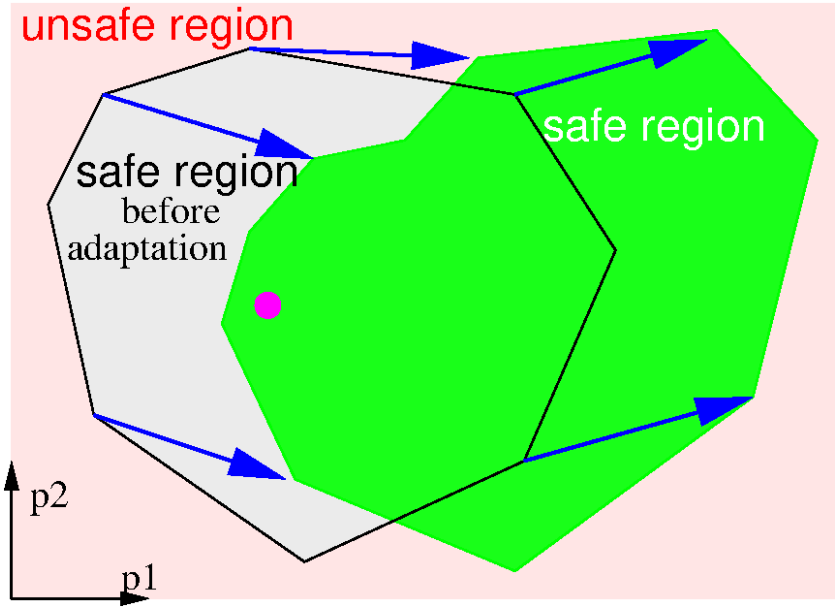
B) Along the boundary

C) Away from boundary in high-variance areas

Discussion I

- Our framework can find and characterize boundaries separating healthy from unhealthy regions
  - Complex system can be considered as a Black Box
  - Efficient exploration of high-dimensional space
  - Shape estimation for
    - effective feedback to designer
    - Basis for efficient SW health monitoring
  - Support for system V&V

Discussion and Future Work



- Health monitoring for changing SW systems
  - Hierarchical Bayesian statistical learning model for dynamic models
  - Active learning for low number of active test cases
  - Bayesian model to represent uncertainties and time-dependent information

Thank you!