Statistics and the Verification Validation & Testing of Adaptive Systems

Roman D. Fresnedo
M&CT, Phantom Works
The Boeing Company

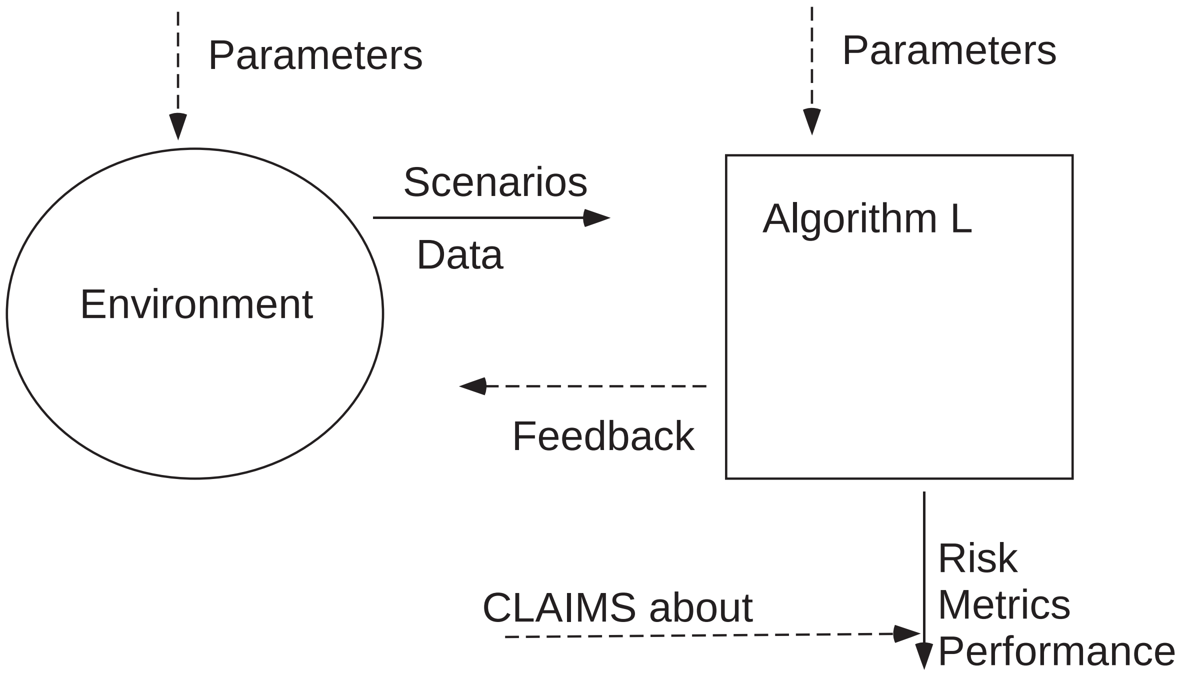
Preliminaries

- Few *old* reflections on *quick* readings and discussions on controllers and other adaptive systems
- General in nature
- Intention: formulate problems, suggest statistical connections, promote dialogue and future interaction

Overview

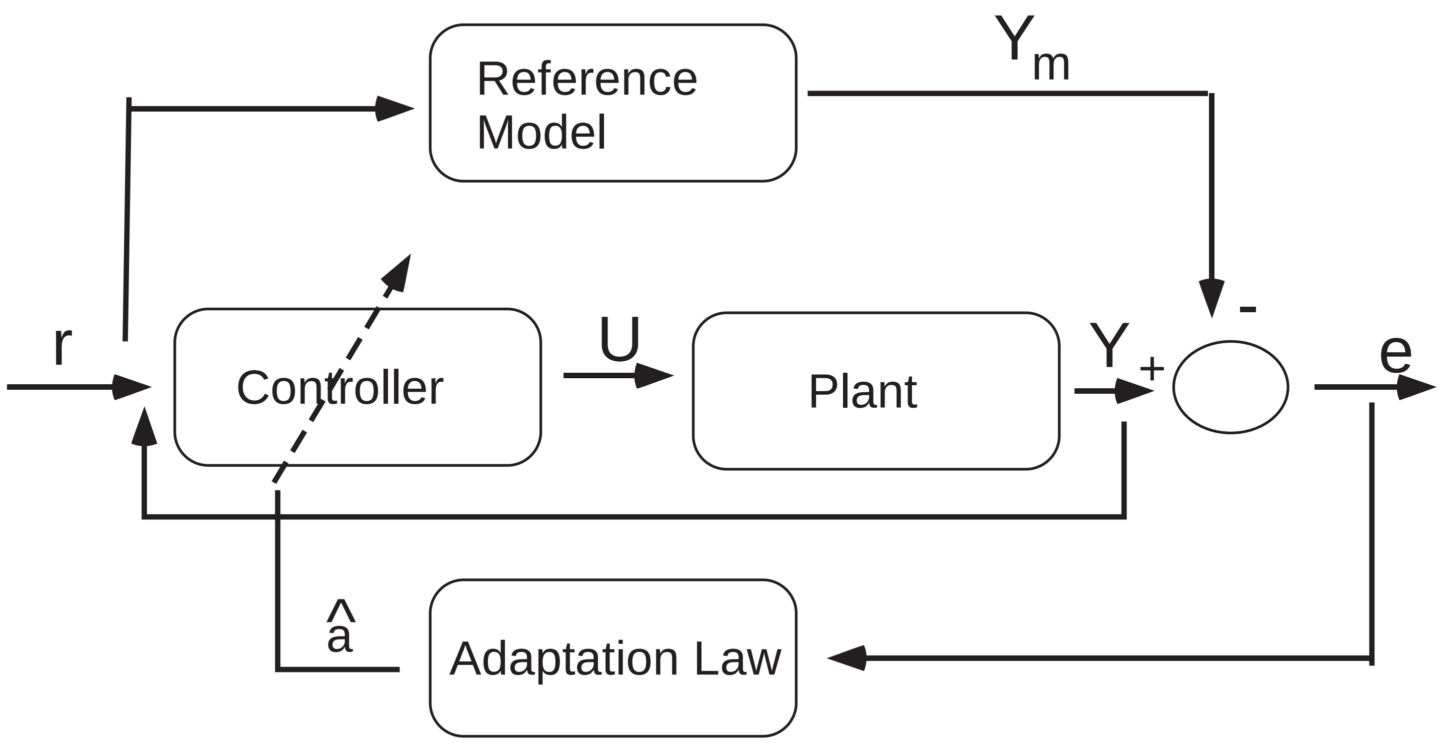
- Adaptive systems
- A couple of examples (Ahmed, Slotine)
- What *are* the problems? *Claims* to be "validated"?
- Which are *statistical* problems?
- Experimental and inferential methods:
  - DOE: where do we test, simulate?
  - How do we infer properties in untested points?
  - How do we calibrate against real systems?
  - Bayesian methods:
    - Estimation, calibration, model selection and uncertainty, extrapolation....

Adaptive System



Clarity about: Definition of Environment, Parameters, Performance Metrics and Claims

MRAC Controller



Where proof ends and induction begins

Identify an *adaptive system* with the evolution of its “state”.
Consider

A “scoring function” that evaluates its performance

Stability, settling, raising times, deviations from 0….

An *environment* (e.g flight cond.) where it is evaluated.

Either we can prove “adaptability” everywhere, or we can prove it at some selected points and *inductively* assert it elsewhere

$A = \{\alpha_t \in \mathbf{A} : t \in T\}$: Adaptive system immersed in one of

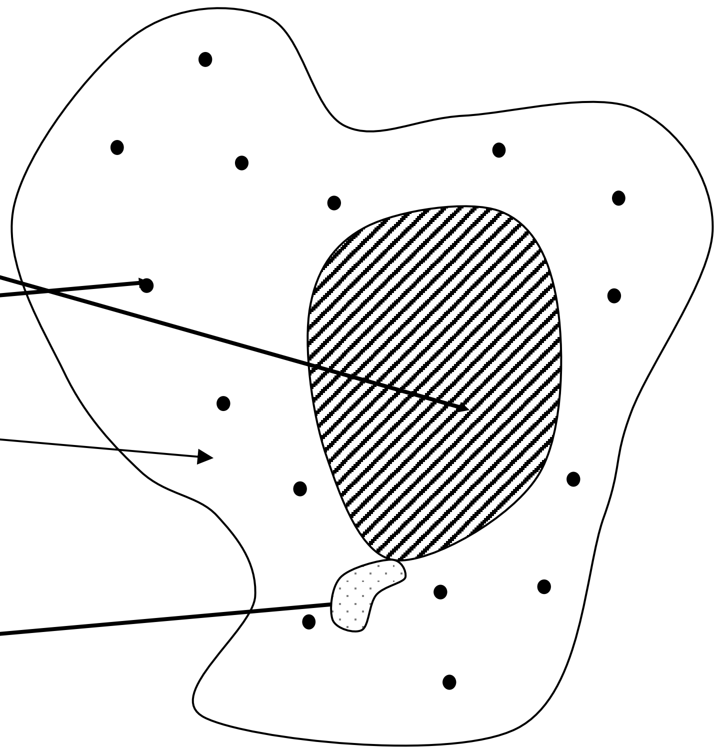
$\{E_t^\gamma : t \in T, \gamma \in \Gamma\}$: a set of environments

adaptability is judged by a “scoring function” $\theta : A \to \Theta \supseteq \Theta_0$

A is adaptive with respect to $(\{E_\gamma\}, \Theta_0)$ if $\gamma \in \Gamma \Rightarrow \theta_\gamma \in \Theta_0$.

Induction (probability)

Large Parameter/Function Space



Significant inductive work

Known with mathematical certainty (e.g., stability proofs)

Experimental points

Must be inferred in the rest of the space

Essentially we are interested in C.I.'s on the performance measure, i.e., what values might it assume *here*?

These are *experimental design* and *inferential* problems.

Example: (Ahmed NN2000)

Pseudo linear model; measurable "operating point" ϑ_t:

$$\dot{x} = A(\vartheta_t)x + b(\vartheta_t)u \tag{4.1}$$

Reference model with command r:

$$\dot{x}_m = A_m x_m + B_m r \tag{4.2}$$

Choose plant input u with gains $k(\vartheta_t)$ (unknown nonlinear functions) so as to drive plant to r=0 (regulator):

$$u = k^T(\vartheta_t)x + k_{n+1}(\vartheta_t)r \tag{4.3}$$

Approximate: $k_i(\vartheta_t) \approx \sum_{j=1}^{\gamma} \theta_{ij}\Psi_{ij}(\vartheta_t)$ Choose A_m for stability ... then an adaptation law and *stable* (Lyapunov) closed loop evolution can be obtained:

$$\dot{x} = (A_m + b(\vartheta_t)\Psi^T\tilde{\Theta})x \; ; \;\; \dot{\tilde{\theta}}_i = -\Gamma_i(h^T x)\Psi x_i \tag{4.4}$$

Illustrative results (Ahmed)

- Simulations: plant selection bias?
  - Plant models (a couple)
  - A few “trials + errors” to choose A (i.e. Q for stab.), h …
  - What if we change the plant model (or Q) by a bit? by a lot?

- 25 to 200+ differential equations for parameter evolution
  - Expensive simulations: need DOE
  - Isn’t something in learning theory/practice that would cut the complexity?

Sanner+Slotine (IEEE 1992)

adaptive control for the following dynamical system:

$$x^{(n)} + f(x, \dot{x}, \ldots, x^{(n-1)}) = bu \tag{4.12}$$

$f(.)$ unknown function, u =control input, b =gain.

Objective: find u so plant $\mathbf{x} = (x, \dot{x}, \ldots, x^{(n-1)})$ follows a desired trajectory $\mathbf{x}_d = (x_d, \dot{x}_d, \ldots, x_d^{(n-1)})$ in such a way that the error remains bounded and $\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{x}_d \to 0$.

For basis $\{g_i(x, \xi_i)\}$ a function estimation method can be found so that, for some set A_d the estimate $\hat{f} = \sum_i \hat{c}_i g_i(x, \xi_i)$ satisfies:

$$\| \hat{f}(x) - f(x) \| < \epsilon_f \quad x \in A_d \tag{4.13}$$

Choose a sufficiently fine grid I_σ in A_d and radial basis functions $g_i = g_\sigma(x, \xi_i)$ centered at grid points ξ_i with constant variance σ^2; A_d, the grid size, and variance are estimated from a-priori known spectral (smoothness) properties of $f(.)$.

Comments and results (S+S)

- Solution of learning problem: state clearly prior information (e.g. spectral smoothness), push the math as far as you can and exploit it to decide approx. method
- ….will bridge the setting up of inference problems (e.g. measures on space of plant models)
- Interplay between theory, controller machinery (e.g. sliding controllers), and "ad-hockeries" is complex
  - Outsider would have difficulty in setting up a DOE
- As in Ahmed, very few experimental illustrations
- Paper if full of intuitive and often rigorous insights on the approximations depending on the dynamics (NN nodes being activated etc.)
- …which probably can be examined more systematically

More questions

- Which class of universal “approximator”?
- How many basis functions?
- For a given basis choice, find the set F of f(.) (plants) for which the approximation is adequate
- For f(.) in F, find the distribution of the “scoring function”

Design of Computer Experiments

DOE + Gaussian processes for the analysis and prediction of complex simulation experiments

- Expensive code of multiparameter models
  - Aerodynamics PDEs

- Run code for selected *design points (DOE)* and interpolate (Gaussian process)

- Successfully used in a wider variety of optimization problems

Model Calibration: Connection to Reality

- Fit the parameters of a *computer model* that represents a real model, given measurements on the latter
  - Intended to make predictions
  - Computer code/real process are deterministic or not
  - Code is expensive to run, yet cheaper than direct measurements
  - Many parameters
- Sources of uncertainty include:
  - Parameter uncertainty
  - Model inadequacy
  - Residual variability (stochastic or not)
  - Observation errors
  - Code uncertainty

- Gaussian processes have been proposed in *other* contexts

Suggestions

- Statistics (probability) is not necessarily related to “noise” or “randomness”; it is about inference in the presence of uncertainty
  – Our uncertainty included: inference about untested cases or about model-reality relationship
- A step into fruitful interaction: clarity about
  – The function space (e.g. plant models) and nature of non-linearities: what do we know, do not know?
    - Example: Sanner-Slotine more principled than Ahmed
  – Alternatives to space of basis functions (NNs)
  – What are the *claims?* (scoring function)
    - Beyond Lyapunov stability, what else can we say, mathematically, experimentally?

Suggestions

- Make a list of function and parameter spaces, describe what is known and design experiments to run models
- Construct a probability on relevant function spaces

- DOE principles are known
  - …although for much simpler models (e.g ANOVA)
  - Need to adapt to function spaces (dynamical systems')

- Gaussian-like processes could be used to interpolate experimental results *and* calibrate models