1. Introduction

2. Behavior discrimination

3. Methodology
   • Sampling-based behavior discrimination
   • Effective sampling schemes

4. Examples: biological systems
   • A model of the acute inflammatory response to infection
   • A model of collagen degradation
   • A model of the T-cell signalling pathway

5. Additional properties
Question: How can we map parameter/input space to a space of qualitative behaviors?

- What parameters yield oscillations versus steady state in a signalling model?
- What values of input lead to system failure?
- What control actions are feasible for a nonlinear control problem?
Behavior discrimination is the problem of identifying sets of parameters for which the system does (or does not) satisfy a qualitative property of interest.

Generally we'll use the term "feasible set" for the set where the property is satisfied.
Mathematical framework

- Consider an ODE system $S(\omega)$ with unknown parameters $\omega \in \Omega$:
  
  $$\dot{x} = \alpha(\omega, x)$$
  
  $$x(0) = x_0(\omega)$$

- A qualitative property of interest is expressed by a binary function on $\Omega$, where $G(\omega) = 1$ if the system satisfies the property, and $G(\omega) = -1$ otherwise.
In control theory: (Feasibility analysis) $G(\omega) = 1$ if certain constrained optimization problem is solvable.

In systems biology:
- $G(\omega) = 1$ if the given state leads to cell death.
- $G(\omega) = 1$ if the current patient state will lead to return to health.
Classical approach in systems biology/ control

Iterative subdivision of feasible/infeasible sets:

- Begin with a given hypercube (start with the entire parameter space).
- Determine if the hypercube is contained entirely in the feasible set \( F \) or the infeasible set \( I \).
- If so, then continue with any remaining hypercubes.
- If not, then subdivide along a direction that divides some points known to be in \( F \) from some known to be in \( I \).
- Continue until all hypercubes are in \( I \) or \( F \) (or are of insignificant size).
Figure: Summer et. al. (2011). The refinement process implies that the number of partitions increases exponentially with the dimension of state space.
Questions:

- How can we get a more efficient representation of the boundary (especially high-dimensional boundary)?
- How do we efficiently sample to determine the boundary?
- How do we estimate the uncertainty in our results?
A sampling-based behavior discrimination algorithm

1. Parameterizing the boundary
2. Data sampling
3. Probability distribution on the coefficient space
4. Probabilistic representation of the boundary
5. Boundary estimation Uncertainty in discrimination
Question: How can we represent the boundary efficiently?

Idea: Assume that there is some smooth function $g$ so that the feasible region is $g > 0$, infeasible is $g < 0$, boundary is $g = 0$.

Goal: Find an efficient approximation of $g$. 
Step 1: Parametrize the discrimination boundary

The discriminating boundary $\partial \Gamma$ is assumed to be the zero level set of a smooth function $\Gamma(\omega)$ that can be well approximated by polynomials

$$f(c, \omega) = \sum_{i=0}^{N} c_i \eta_i(\omega)$$

where $c_k$ is in a coefficient space $C$ that needs to be inferred.
Bayesian Inverse Problem

Figure 1: Uncertainty Quantification in Bayesian Inversion

Figure: Stuart (2014): Uncertainty Quantification in Bayesian Inversion

Vu Dinh, Ann E. Rundell and Gregory T. Buzzard
A Probabilistic Method for Efficient Behavior Classification
Step 2: Probabilistic representation of the boundary

1. A sampling scheme is used to select parameter values $\omega_i$ from $\Omega$, compute $G(\omega_i)$.

2. Based on this binary data, a probability distribution is generated on $C$

$$\pi_m(c) \propto \exp \left( - \sum_{i=0}^{m} (G(\omega_i) - \phi(c, \omega_i))^2 \right)$$

3. This distribution, when propagated to the space of all possible boundary curves, induces a probabilistic representation of the boundary.
Step 3: Boundary estimation and uncertainty in prediction

1. The expected predicted function is used to discriminate between behaviors.

   \[ \bar{\phi}_m(\omega) = E_{\pi_m}[\phi(c, \omega)] \]  \hspace{1cm} (3.1)

2. The uncertainty in the discrimination at a point \(\omega\) is represented by the variance in prediction

   \[ \text{Var}_{\pi_m}[\phi(c, \omega)] \]  \hspace{1cm} (3.2)
Uncertainty in discrimination

[Diagram with two plots]

Vu Dinh, Ann E. Rundell and Gregory T. Buzzard

A Probabilistic Method for Efficient Behavior Classification
Uncertainty-based sampling
Sampling schemes

1. Low-discrepancy sampling: \( \{ \omega_i \} \) has discrepancy approaches 0 when \( m \to \infty \).

\[
D_m(\{\omega_1, \ldots, \omega_m\}) = \sup_{B \subseteq \Omega} \left| \frac{\#B_m}{m} - \text{Vol}(B) \right| \to 0 \tag{3.3}
\]

2. Sequential sampling: data is collected sequentially. At step \( m \), the next data point \( \omega_{m+1} \) is taken at the point where the maximum of response variance with respect to \( \pi_m \) is achieved:

\[
\omega_{m+1} = \arg \max_{\omega \in \Omega} \text{Var}_{\pi_m}[\phi(c, \omega)]
\]
Low-discrepancy sampling gives asymptotically uniform coverage of space. This guarantees that we convergence to the correct result (assuming that the polynomials we choose can represent the boundary - the model is correct).

Sequential sampling focuses on the regions of greatest uncertainty - should give faster convergence, but we may miss some features.
Low-discrepancy sampling: rate of convergence

**Theorem**

Assume that the approximate model is correct, and \( \{ \omega_i \} \) has discrepancy \( D_m \) tending to 0 when \( m \to \infty \). Then

\[
\lim_{m \to \infty} \bar{\phi}_m(\omega) = G(\omega) \quad \forall \omega \in \Omega
\]

Moreover, for all \( \epsilon > 0 \), denote \( D^{-1}(\epsilon) = \sup \{m + 1 : D_m \geq \epsilon \} \), then for \( m = \Theta \left( D^{-1}(\epsilon) + N_{\frac{1}{\epsilon}} \log \frac{1}{\epsilon} \right) \), where \( N \) is the number of terms in the polynomial expansion, we have

\[
\int_{\Omega} |\bar{\phi}_m(\omega) - G(\omega)| d\omega \leq \epsilon
\]
Theorem

Assume that the approximate model is correct, the parameter space has finite cardinality, and data is collected sequentially.

\[ \omega_{m+1} = \arg \max_{\omega \in \Omega} \text{Var}_{\pi_m}[\phi(c, \omega)] \]

Then:

\[ \lim_{m \to \infty} \bar{\phi}_m(\omega) = G(\omega) \quad \forall \omega \in \Omega \]
A model of the acute inflammatory response to infection

- 4 equations, 22 parameters.
- State variables P, N, D, and C, correspond to the amounts of pathogen, pro-inflammatory mediators, tissue damage, and anti-inflammatory mediators, respectively.
- Death is defined as a sustained amount of tissue damage (D) above a specified threshold value.
A model of collagen degradation

- 12 equations, 22 parameters
- Models the loosening of the extra-cellular matrix, a crucial process in angiogenesis, the sprouting of new blood vessels as a reaction to signals that indicate the need for additional oxygen in certain tissues.
- Takes 20 minutes for one evaluation of the ODEs.
Discrimination: Sparse sampling

Initial concentration of TMP2 vs Initial concentration of MT1-MMP

Vu Dinh, Ann E. Rundell and Gregery T. Buzzard
A Probabilistic Method for Efficient Behavior Classification
Discrimination: Sequential sampling

(A)

Initial concentration of TIMP2 vs. Initial concentration of MT1-MMP
Uncertainty in discrimination

Vu Dinh, Ann E. Rundell and Gregery T. Buzzard

A Probabilistic Method for Efficient Behavior Classification
Uncertainty in discrimination

Left: Design points and predicted boundary with sequential sampling scheme. Right: Variance. Notice that the points with high variance lie around the true boundary, which explains why the data sampled on the figure on the left also tends to focus around the true boundary.
Coupled system of ODEs: 37 equations, 19 parameters.
Part of the adaptive immune system.
Identify where pERK stabilizes at a high level.
Use sensitivity analysis to identify most in sequential parameters.
Discrimination: Three dimensions
In the ideal case, currently misclassified points are added in the best possible order (requires knowledge not usually available in practice).
We consider a few more properties of this method.
Convergence

Graph showing prediction error vs. number of samples for Sparse sampling and Sequential sampling.
Boundary with disconnected component
Robustness

![Diagram showing some data points and curves with labels:]
- Sequential samples
- Boundary
- Sparse sampling
- Sequential sampling

---

Vu Dinh, Ann E. Rundell and Gregery T. Buzzard

A Probabilistic Method for Efficient Behavior Classification
Conclusions/Future works

- Efficient method for behavior discrimination in moderate dimension.
- Convergence results for low-discrepancy and sequential sampling
- Future works
  1. Non-parametric approach.
  2. Convergence rate of the sequential setting.
