System Design as a Mechanism for Generalization

Sponsor: OUSD(R&E) | CCDC

Mr. Tyler Cody

11th Annual SERC Sponsor Research Review and
7th Annual SERC Doctoral Students Forum
November 18-19, 2019
FHI 360 CONFERENCE CENTER
1825 Connecticut Avenue NW, 8th Floor
Washington, DC 20009

www.sercuarc.org

Executive Summary

- I am studying how system design can be used as a mechanism for the generalization of learning systems.
- I am taking a systems approach to machine learning by:
  1. Using systems theory as a mathematical superstructure for learning, specifically theory on **input-output systems**
  2. Extending the **boundary of the system** not just around the algorithm and data, but to the system within which the learner operates
  3. Eliciting trade-offs wherein **stakeholder values** can select an operating point, instead of presupposing values and metrics
- We show how systems theory connects to learning theory, and how systems trade-offs can be elicited in applied machine learning

Contents

1. Systems Theory
   - Introduction
   - Connecting Systems Theory to Learning Theory
   - How System Design Emerges as a Mechanism for Generalization
2. Real-World Example: Eliciting Trade-Offs in Design and Generalization
   - Case study in design of rebuild procedure for machinery
3. Mathematical Framework for Understanding System and Algorithm Design for Generalization
4. Closing Remarks

What is systems theory?

- *"… there exists models, principals, and laws that apply to generalized systems or their subclasses, irrespective of their particular kind, the nature of their component elements, and the relationships of "forces" between them. It seems legitimate to ask for a theory, not of systems of a more or less special kind, but of universal principals applying to systems in general."*

- Ludwig von Bertalanffy, General System Theory (1968)

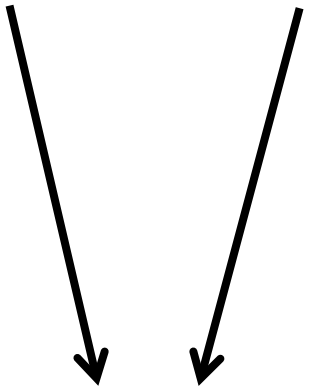
What is systems theory?

Descriptive General Systems Theory
Ludwig von Bertalanffy
School of Thought

- *Anatol Rapoport*
- *Kenneth Boulding*

Cybernetics
Norbert Weiner
School of Thought

- *Ross Ashby*
- *Herbert Simon*



Mathematical General Systems Theory

- M.D. Mesarovic (Abstract Systems Theory)
- Wayne Wymore (Model-Based Systems Engineering)

Levels of Abstraction

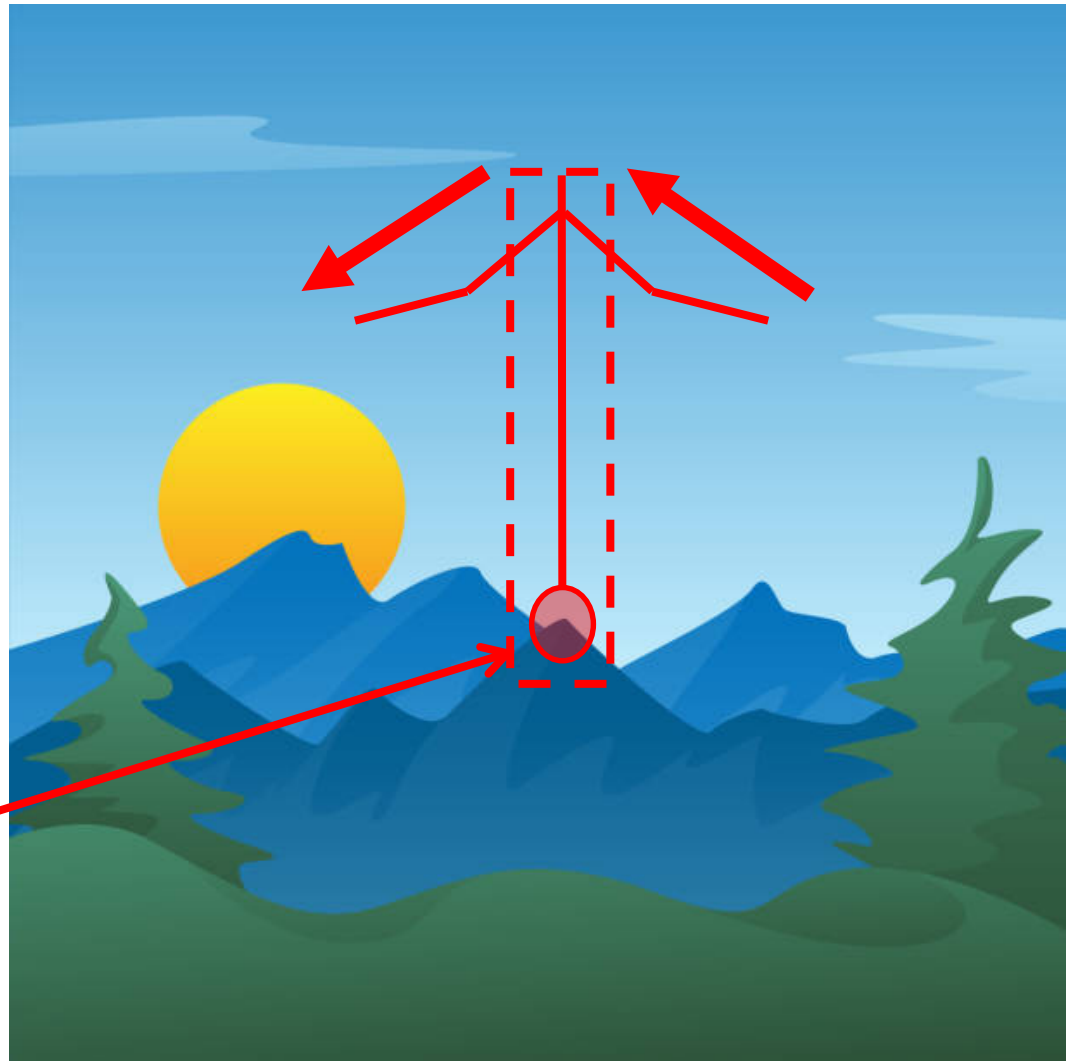
Systems Theory

Learning Theory

Machine Learning

Transfer Learning

Biology



Mesarovician Systems Theory

- **Definition**. *System*.

A general system is a relation on non-empty (abstract) sets,

$$S \subset \times \{V_i : i \in I\}$$

where $\times$ is the Cartesian product, I is the index set, and V_i are the component sets.

- **Definition**. *Input-Output System*.

An input-output system is a general system where the component sets can be partitioned into an input object and output object,

$$X = \times \{V_i : i \in I_x\} \text{ and } Y = \times \{V_i : i \in I_y\}$$

where $I_x \cup I_y = I$. Thus,

$$S \subset X \times Y$$

Learning as a Mesarovician Input-Output System

- **Definition**. *Learning System*.

A learning system is an input-output system,

$$S: X \to Y$$

with a sample D and a learning algorithm $A: D \to f^{\theta}$, where $f^{\theta}: X \to Y$ is a parameterized mapping.

- **Definition**. *An Empirical Risk Minimization Learning System*.

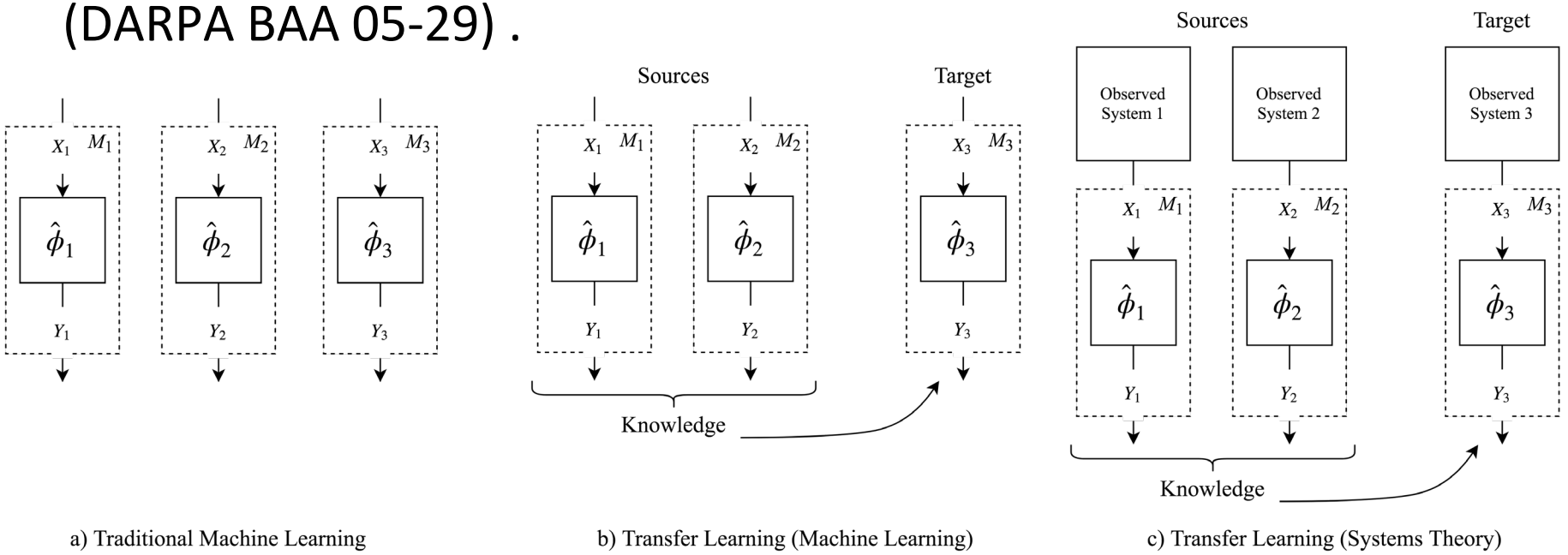
An empirical risk minimization learning system is a learning system where D is an i.i.d. sample of l input-output observations,

$A: D \to f^{\min_{\theta} R_{emp}(\theta)}$, where, $R_{emp}(\theta) = \frac{1}{l}\sum_{i=1}^{l} L(y_i, f^{\theta}(x_i))$

and L is a loss function.

Systems Theoretic Perspective on Transfer Learning

- By treating learning as an input-output system, learning problems can be straightforwardly embedded into their systems context
- *Transfer learning* is "the ability of a system to recognize and apply knowledge and skills learned in previous tasks to novel tasks" (DARPA BAA 05-29) .

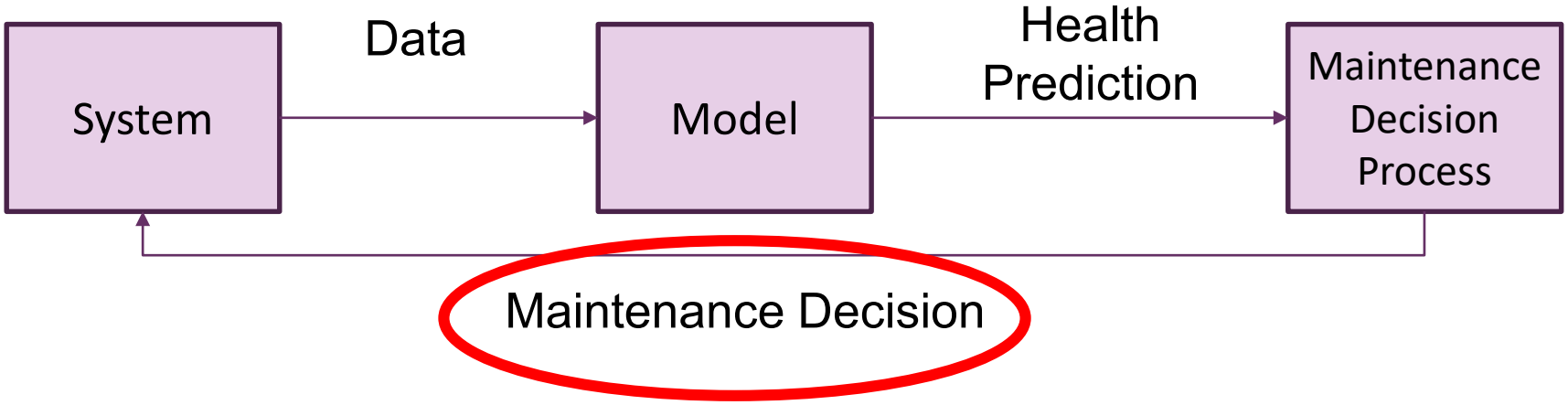


a) Traditional Machine Learning b) Transfer Learning (Machine Learning) c) Transfer Learning (Systems Theory)

Cody, Adams, & Beling. "A Systems Theoretic Approach to Transfer Learning." *IEEE SYSCON* 2019.

Actuator Health Monitoring

- Learning algorithms are used to predict current and future health states of actuators
- Actuators have similar underlying physics, but physical and functional differences exist between actuators and over time
- For example, **actuators change between rebuilds**.



Actuator Health Monitoring

- The transfer learning problem…

How do we transfer knowledge between actuators to make learning easier/feasible while accounting for individual differences?

- The machine learning approach…

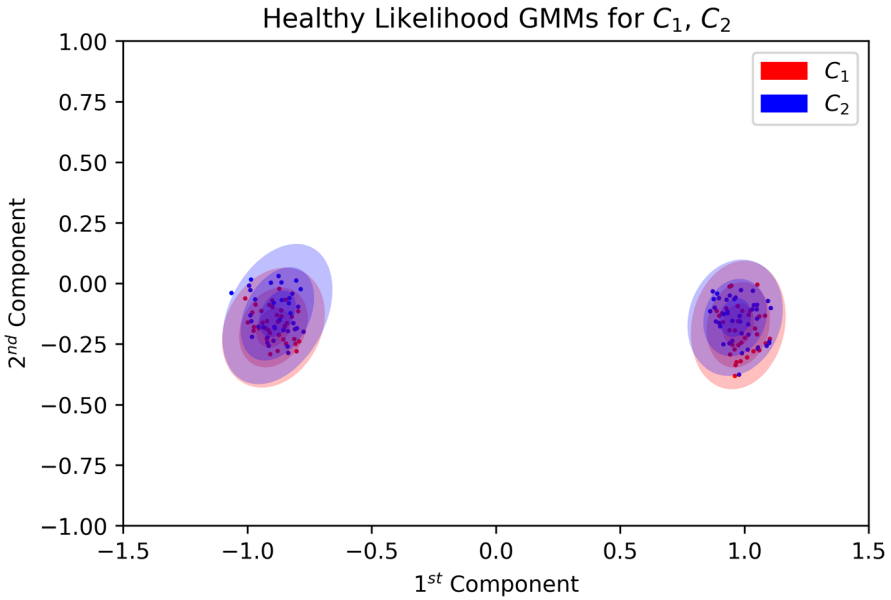
Statistically relate source and target actuators by using selection, weighing, and transformation of data, features, and parameters.

- **The systems design approach**…

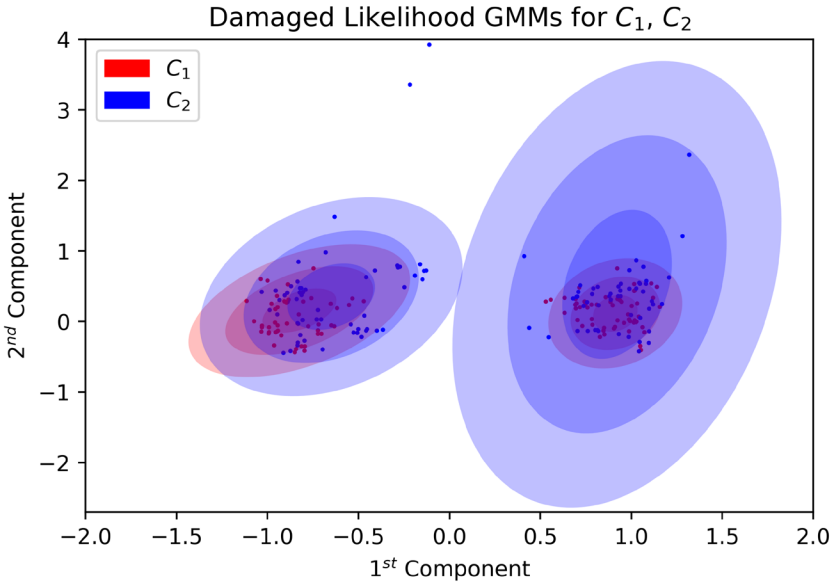
*Use systems knowledge to **understand and design the cyber-physical nature of the distributional change** between actuators.*

Probabilistic Models of Behavior

- We use probabilistic models to understand the distributional change associated with a rebuild process



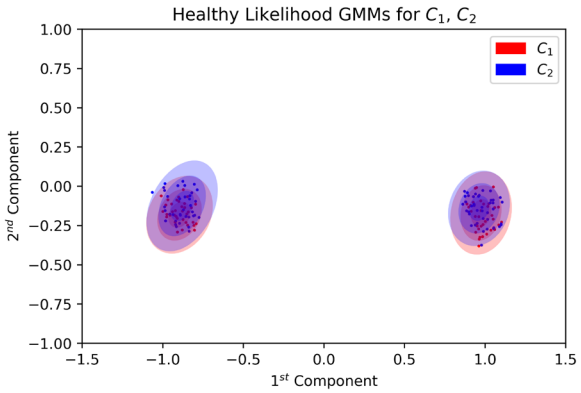
$$p(X = x|Y = healthy)$$

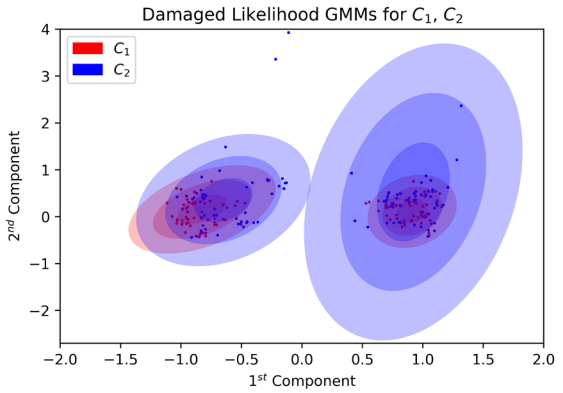


$$p(X = x|Y = damaged)$$

Probabilistic Models of Behavior

$$p(X = x|Y = healthy) \qquad p(X = x|Y = damaged)$$





- Using these, and a prior probability $P(Y = y)$, we can construct:

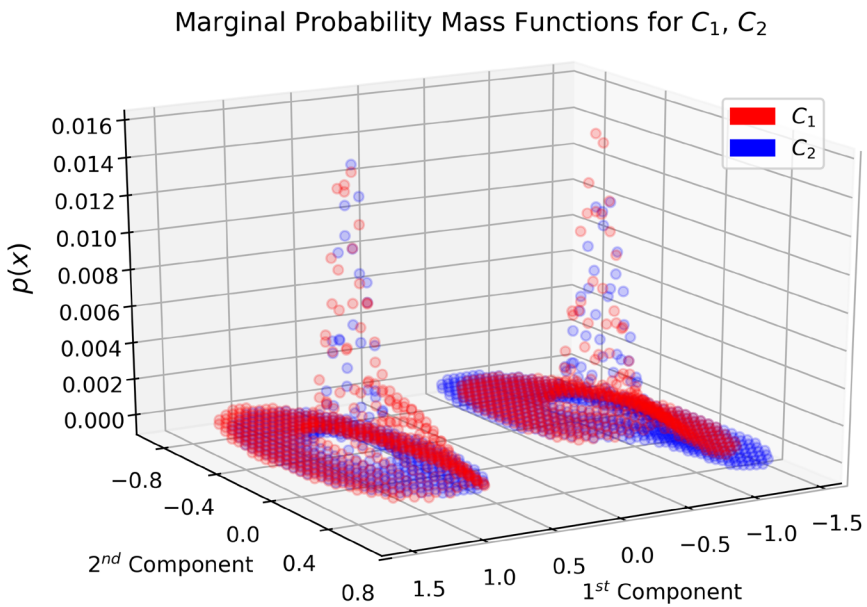
$$p(X = x) = \sum_{y} P(Y = y)\, p(X = x|Y = y)$$

- And using Bayes Theorem:

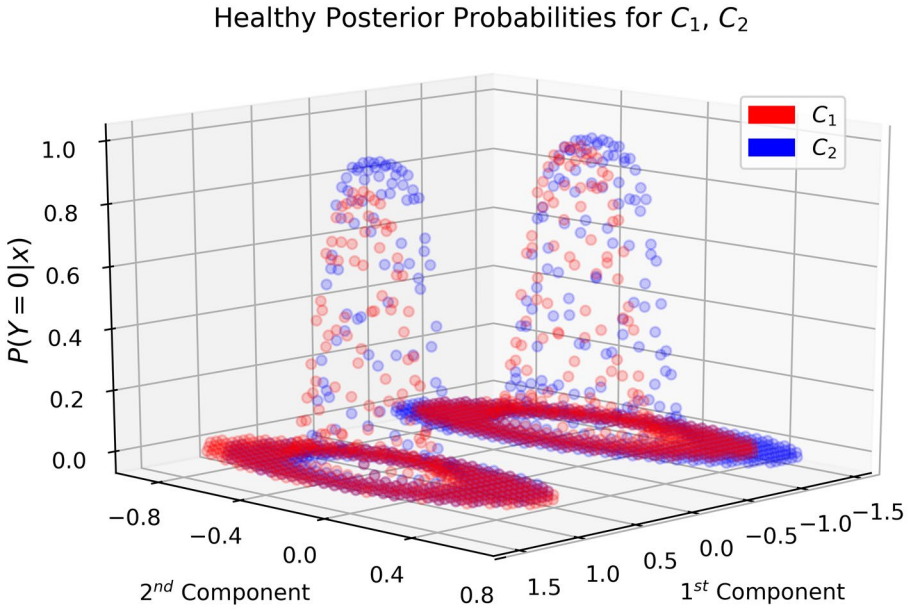
$$P(\mathrm{Y} = \mathrm{y}|\mathrm{X} = \mathrm{x}) = \frac{p(X = x|Y = y)P(Y = y)}{p(X = x)}$$

Characterizing Transfer Distance/Transferability

$p(X = x)$



$p(Y = healthy|X = x)$

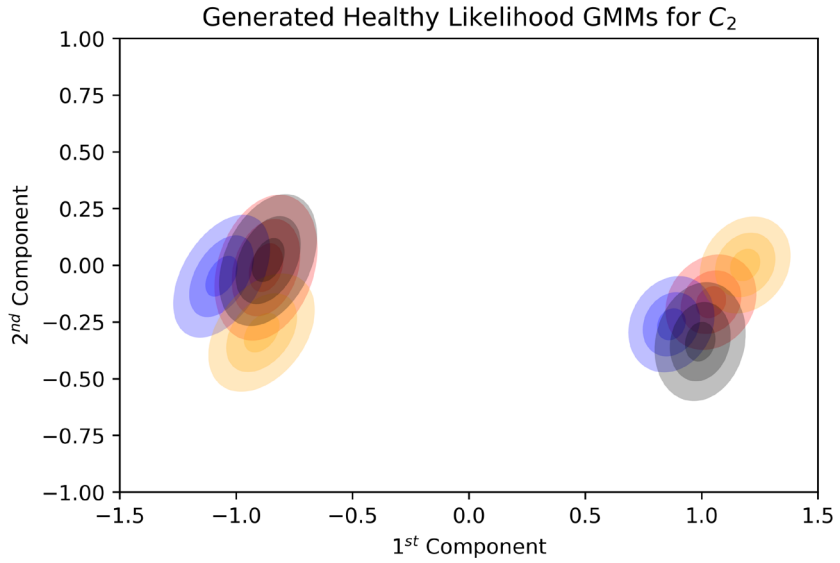


Distribution	Hellinger Distance
$p(X = x\|Y = healthy)$	0.23
$p(X = x\|Y = damaged)$	0.55
$p(X = x)$	0.41
$p(Y = healthy\|X = x)$	0.27

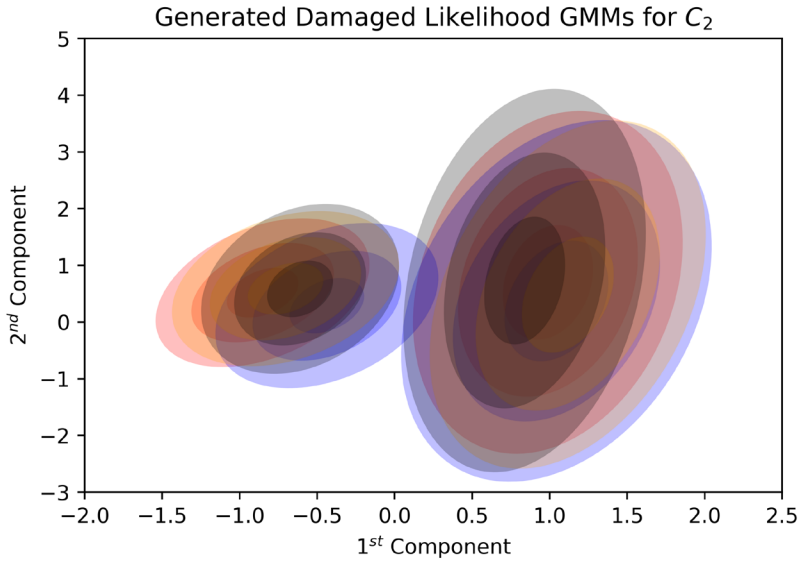
Trade-Offs: Transfer Distance and Rebuild Procedure

- This characterization specifies a family of distributions

$$p(X = x|Y = healthy)$$



$$p(X = x|Y = damaged)$$



- Design of the rebuild procedure influences this family, determining the potential non-stationarities the learning system may face, and thus, the generalization problem faced by the learning system.

Towards a Mathematical Framework

Recall, a learning algorithm A is a map,

$$A: D \rightarrow f^{\theta}\ , f^{\theta}: X \rightarrow Y.$$

Given an evaluation function, $v: f^{\theta} \rightarrow \mathbb{R}$, and a real threshold ϵ, we are interested in identifying the neighborhood,

$$N = \{P(X,Y) | v\left(f^{\theta}\right) \geq \epsilon\}$$

System behavior is captured by the random process,

$$R(X,Y) = \{P_t(X,Y) | t = 1, \ldots, T\}$$

Under this model of learning, system design influences generalization through influence over $R(X,Y)$.

In other words, generalization is both a systems and algorithm design problem.

Synopsis

We showed that:

1. Systems theory is a superstructure for learning theory that can be used to mathematically connect learning algorithms to the systems within which they operate

2. Under this framework, system design emerges as an important component in the generalization of learning algorithms over system lifecycles

Closing Thoughts

- Mathematical systems theory offers a research path towards systems engineering principles for learning algorithms
- This sort of research is dependent on access to systems and testbeds; data sets alone are insufficient
- General, systems-independent constraints on the use of learning algorithms are important, but limited; properties of learning systems are tied to both systems and learning theoretic properties

Acknowledgements

This material is based upon work supported by the Naval Sea Systems Command under Contract No. N00024-17-C-4008. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Naval Sea Systems Command.

Thank you!
Questions?

Tyler Cody – tmc4dk@virginia.edu

Peter Beling – pb3a@virginia.edu