Immune Systems as Complex Learning Machines

Terran Lane and Melanie Moses
Department of Computer Science
University of New Mexico

THE UNIVERSITY of NEW MEXICO
Computer Science

CETI
Center for Evolutionary and Theoretical Immunology
University of New Mexico • Los Alamos National Laboratory • Santa Fe Institute
NCRR

Abstract

We view immune systems as *learning machines* that extract information from their environments, encoding it in an internal representation that enables the machine to improve its performance (e.g., measured as host survival) over time.

The core resulting hypothesis is that such machines can operate in at least two performance regimes. In a "stationary environment", the machine faces an essentially fixed distribution of environmental stimuli, such as an individual that faces a predominantly fixed population of pathogens in its lifetime. This regime corresponds to "classical" learning machines, which display a characteristically-shaped asymptotic learning curve. Such learning machines are well understood in the applied and theoretical learning literature, and we intend to apply those tools and techniques to understand how certain features of adaptive immunity allow effective response to a wide variety of pathogens. For example, classical learning results predict that fixed-capacity learning machines are sufficient to learn a finite complexity, stationary concept. This may explain why a fixed-length antigen-binding variable region (essentially a fixed-length pattern recognizer) in each antibody and a fixed B-cell population size are sufficient to recognize almost any pathogen.

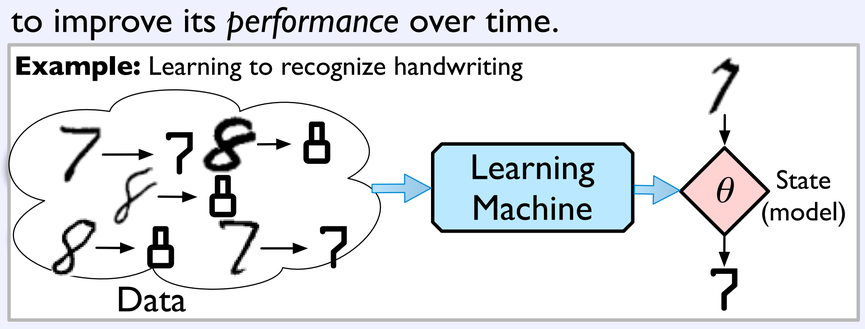
However, the adaptive immune system must also perform well in a second, "non-stationary," regime, in which the environment adapts in response to the learning machine. Here, both pathogen and host immunity are considered learning machines, engaged in a competitive game or "arms race". For example, both immune system and pathogen "learn" to respond to each other over evolutionary time via differential survival of host organisms and pathogens. The adaptive immune system also learns to recognize an evolving population of pathogens within an individual's lifespan through somatic hypermutation. This non-stationary regime is less well studied or understood from a learning theory perspective, however fixed-capacity learning machines do not appear to be adequate.

We propose to investigate this hypothesis both to understand the evolution of complexity in the immune system and to better understand the dynamics of learning in the nonstationary regime.

Machine Learning Background

A *learning machine* is a system that observes some environmental *data* and modifies its internal *state* so as to improve its *performance* over time.

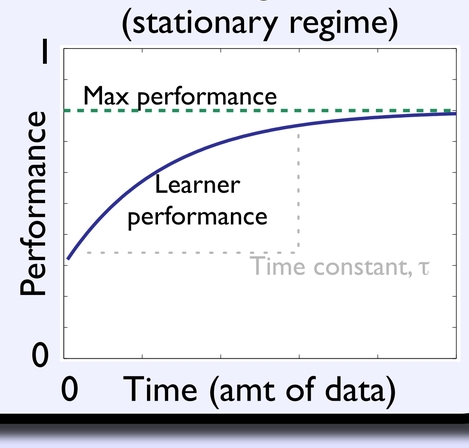
Example: Learning to recognize handwriting

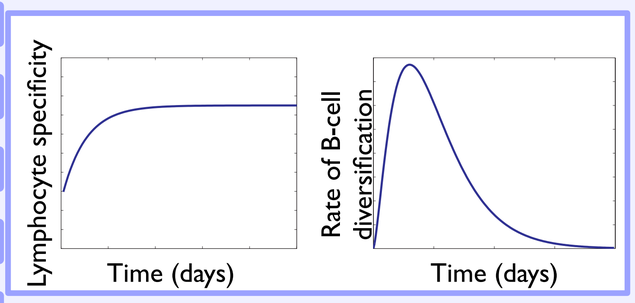
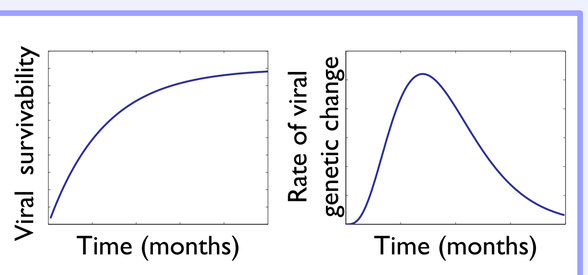
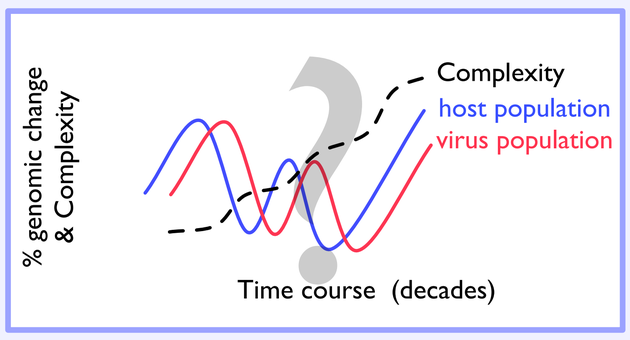


Two performance regimes:

- *Stationary*: distribution of data unchanging w.r.t. time
  - Asymptotic performance [4]
  - Finite information in environment and model
  - Fixed capacity (fixed model size) sufficient
- *Nonstationary*:
  - Data changing as model updated
  - Much less well understood
  - May require unbounded model

The canonical learning curve (stationary regime)



	$\tau_h \gg \tau_p$	$\tau_h \ll \tau_p$	$\tau_h \approx \tau_p$
Host/Pathogen	Individual mouse vs pathogens exposed to in its lifetime	Individual human vs HIV	Human population vs flu population or Bird population vs WNV population
Adaptive Element	**Host somatic hypermutation** generates effective B cell repertoire. Lymphocyte diversity declines with age [1] due to repeat exposure to similar pathogens	**Viral genome evolves** within each host to evade CD8 CTL [2].	**Hosts**: Somatic hypermutation, MHC evolution, cross reactive immunity **Viral genome:** evolves multiple serotypes, antigenic drift & shift [3].
Predicted Response			
Hypotheses/ Experiments	Measure Ab diversity with age in control mice exposed to a constant pathogen population and mice continuously exposed to new pathogens: Ab diversity asymptotes only in control mice	Measure change in viral genome diversity in response to drug treatment and CTL response: Expect viral titer to asymptote and diversity to peak & decline Question: do CTL continuously evolve in response to HIV?	Measure Viral genome & host MHC diversity in different ecosystems. Expect -Viral strain diversity increases with host diversity -Higher MHC evolution in species with high mortality (higher in corvids vs robins) -MHC evolution asymptotes only where viral diversity asymptotes

Stationary Regime | Nonstationary Regime

Discussion

- We attempt to predict how pathogens and immune responses co-evolve given different time scales of evolution for each.
- Viruses & adaptive immunity evolve over short (e.g. HIV & somatic hypermutation) and long (e.g. antigenic shift in influenza & MHC evolution) timescales.
- Effective immune response & therapies operate where $\tau_h \gg \tau_p$
- Hypothesis 1: Nonstationary regime drives complexity
  - E.g. arms race between MHC and viral strains increases complexity of both: mutual moving targets
- Hypothesis 2: Representation/learning capacity for stationary regime immune mechanisms can be bounded; learning capacity for nonstationary regime must be unbounded
  - E.g., Fixed-length antigen strings vs. unbounded genetic string length representation for MHC et al.
- Studying evolutionary arms races may guide ML algorithms for nonstationary learning domains

We Seek Suggestions for experimental systems to test stationary and nonstationary learning

References
[1] Messaoudi et al JEM 200(10) 2004
[2] Leslie et al Nature Med 10(3) 2004
[3] Fitch et al PNAS 94(15) 2007
[4] Kearns, M. J. and Vazirani, U.V., *An Introduction to Computational Learning Theory*. MIT Press, 1994.

Acknowledgements: Prof. Moses's work is supported by the NIH Grant P20RR018754 to the UNM Center for Evolutionary and Theoretical Immunology (CETI). Prof Lane's work is supported by NSF grant IIS-0705681.