TEXAS INSTITUTE FOR INTELLIGENT BIO-NANO MATERIALS AND STRUCTURES FOR AEROSPACE VEHICLES
A NASA University Research, Engineering and Technology Institute (URETI)

Dimensionality Effects on the Markov Property in Shape Memory Alloy Hysteretic Environment

Kenton Kirkpatrick
John Valasek

Aerospace Engineering Department
Texas A&M University

Compos Volatus

IEEE SMC Conference
13 October 2009
San Antonio, TX

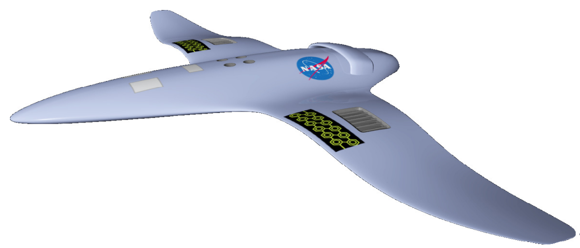
- Funded by the National Aeronautics and Space Administration
- Administered by the Texas Engineering Experiment Station
- A collaborative effort among: Prairie View A&M University | Rice University | Texas A&M University | | Texas Southern University | University of Houston | University of Texas - Arlington

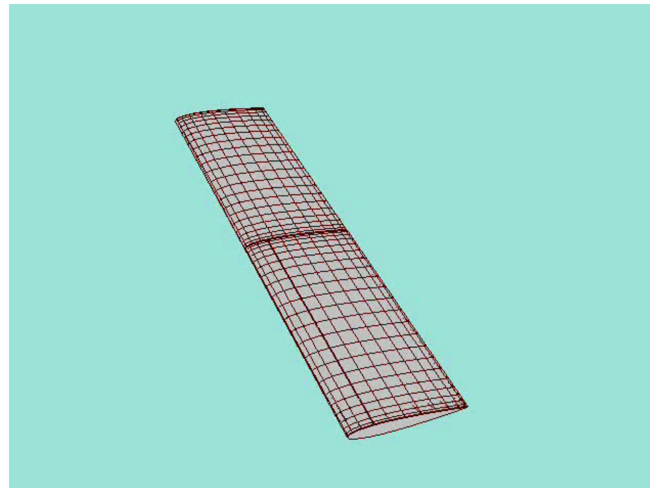
Basic Overview

- Motivation
- Reinforcement Learning
- The Markov Property
- Simulation
- Conclusions
- Challenges and Open Problems

Motivation

Morphing Aircraft





Why should it have morphing capabilities?

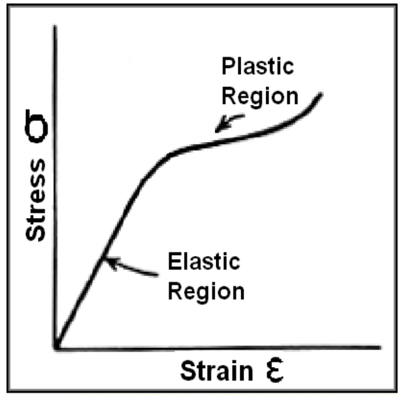
- Improved efficiency at multiple flight conditions
- Wider range of possible maneuvers

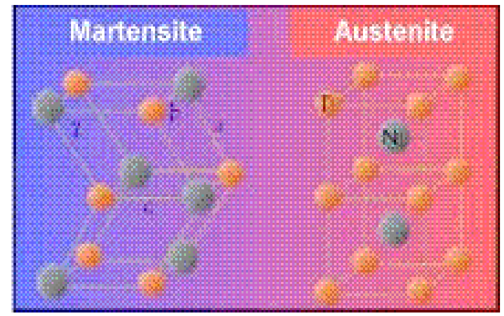
How would it physically achieve morphing?

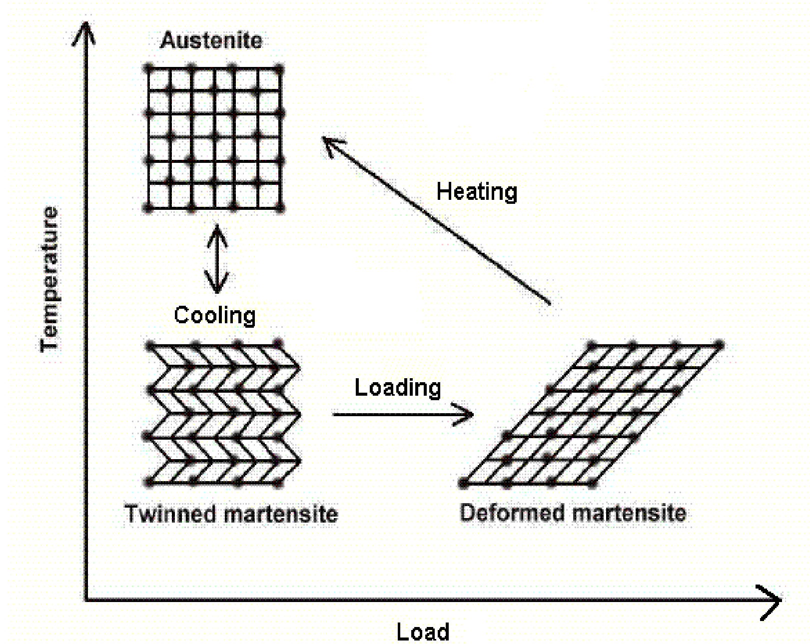
- SMA Actuators

Shape Memory Alloys (SMA)

- Metallic Alloys used as actuators
- Shape Memory Effect – SMA can fully recover from a plastically deformed shape change by the addition of heat
- Electricity can be used to induce a cycle of heating, cooling, and deformation in an SMA in order to execute a dynamic task





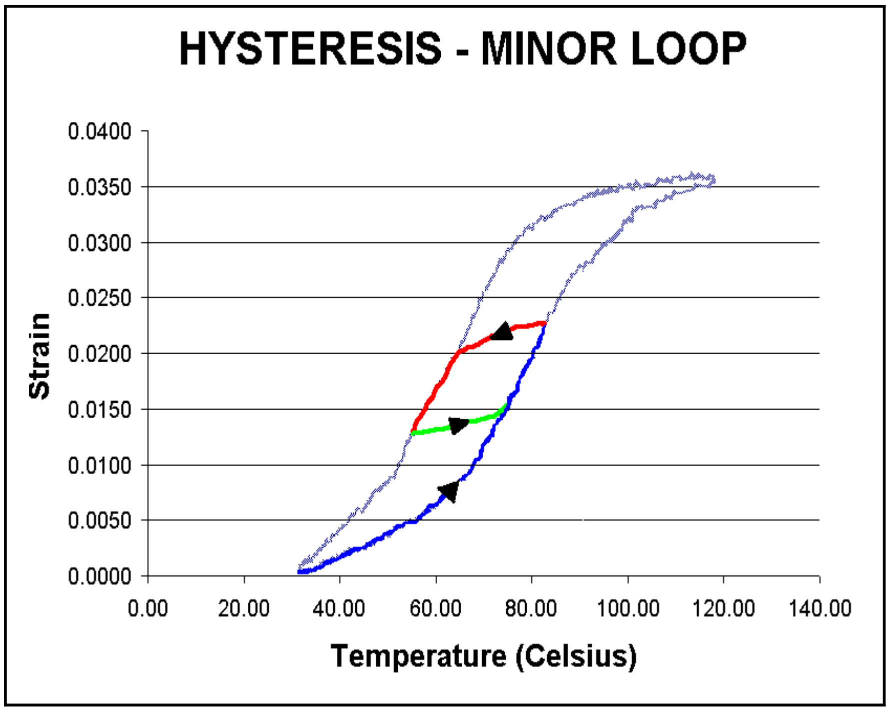


- PROBLEM:

  Efficiently characterizing and controlling SMA behavior

Why Is Characterizing SMAs Such A Challenge?

1. Uncertain Model Parameters
2. Temperature–Strain Relationship: **HYSTERESIS**
   - Behavior based on a <u>2-D FIELD</u>, not a 2-D PATH



Reinforcement Learning

Reinforcement Learning

- Does not require any prior knowledge.
  - Knowledge is based on experience and interaction with the environment, not on input-output data supplied by an external supervisor

- Achieves a specific **goal** by learning from interactions with the environment.
  - Considers **state** information (s)
  - Performs sequences of **actions**, (a), observing the consequences
  - Attempts to maximize **rewards** (r) over time
    - **These specify what is to be achieved,** ***not how to achieve it***
  - Constructs an **action value function** (Q)
    - Learns an optimal control **policy**

- Memory is contained in the action value function

RL Algorithm: Sarsa

$$Q(S,A,G) \leftarrow Q(S,A,G) + \alpha[R(S',A',G) + \gamma Q(S',A',G) - Q(S,A,G)]$$

3-D Control Policy Matrix

- States (S): Temperature and Strain Dependent
- Actions (A): Change Temperature (Voltage Application)
- Goal (G): Desired Strain
- α: Repetition Penalty
- γ: Future Policy Weight

Action Choice Method: ε-Greedy

- Explore or Exploit: Dependent upon ε (which varies with Episodes)

Why Not Q-Learning?

- On-policy v. Off-policy Learning

The Markov Property

The Markov Property

- In a general system, the probability of achieving a specific goal from a specific initial state is a function of both current and past information about states, actions, and rewards.

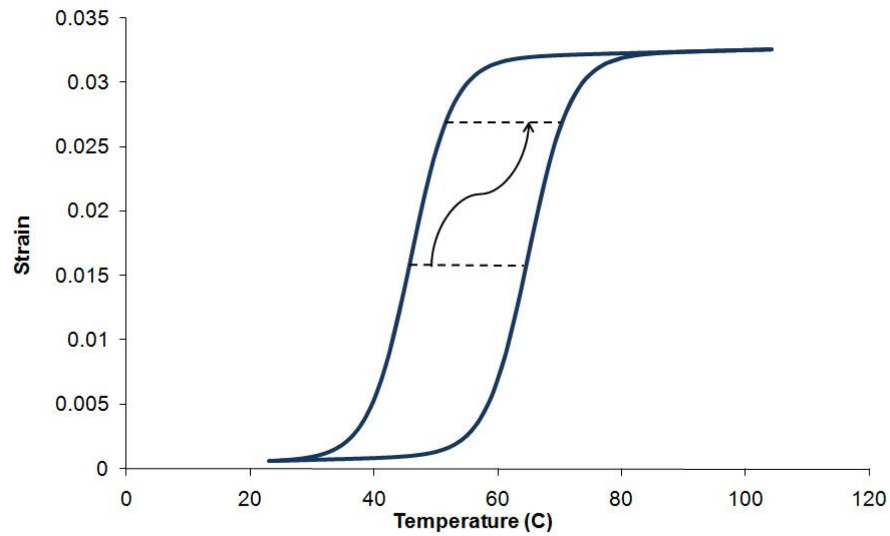
$$\Pr\left\{s_{t+1} = s', r_{t+1} = r \mid s_t, a_t, r_t, s_{t-1}, a_{t-1}, \ldots, r_1, s_0, a_0\right\}$$

- In a system with the Markov Property, the same probability distribution is obtained with only the current state and action information required.

$$\Pr\left\{s_{t+1} = s', r_{t+1} = r \mid s_t, a_t\right\}$$

The Markov Property

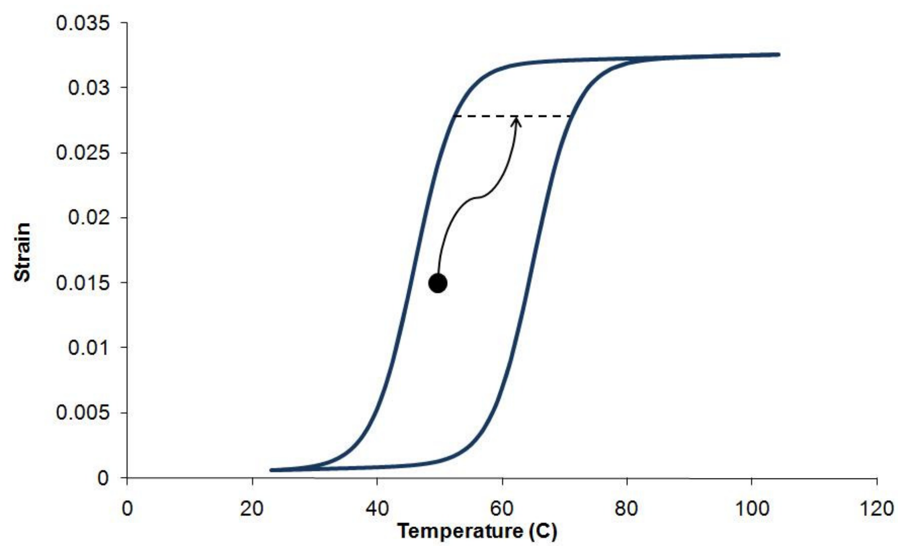
- Hysteresis is non-Markovian in nature because moving from one state to another in hysteresis space typically requires knowledge of state history.
- In the problem of learning to control the strain of an SMA wire during a thermally-induced transformation, this non-Markovian behavior is apparent.



The Markov Property

- For this research, we recognized that the past strain history is only needed so that we know the current point in temperature/strain space.

- By measuring temperature and including it in the current state information, the system becomes Markovian.



Simulation

Simulation Model

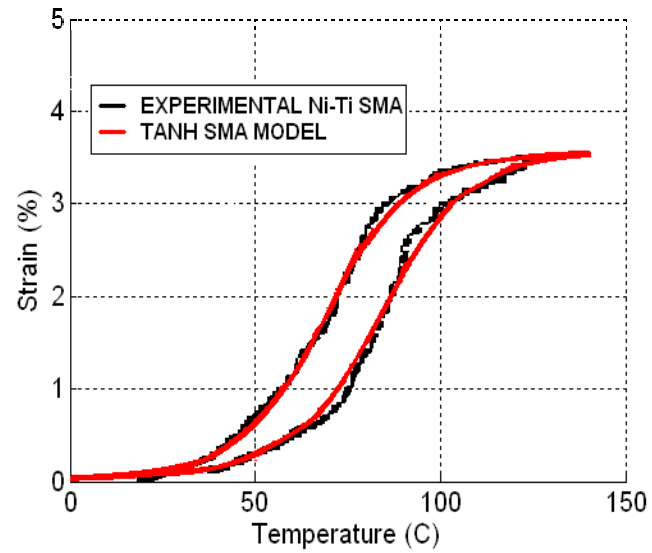
- **Temperature-Strain Relation:**
  - Hyperbolic Tangent Model

$$\varepsilon(i) = \frac{1}{2} h \tanh\left((T(i) - ctl)a\right) + s\left(T(i) - \frac{1}{2}ctl - \frac{1}{2}ctr\right) + \frac{1}{2}h + cs$$

- **Voltage-Temperature Relation:**

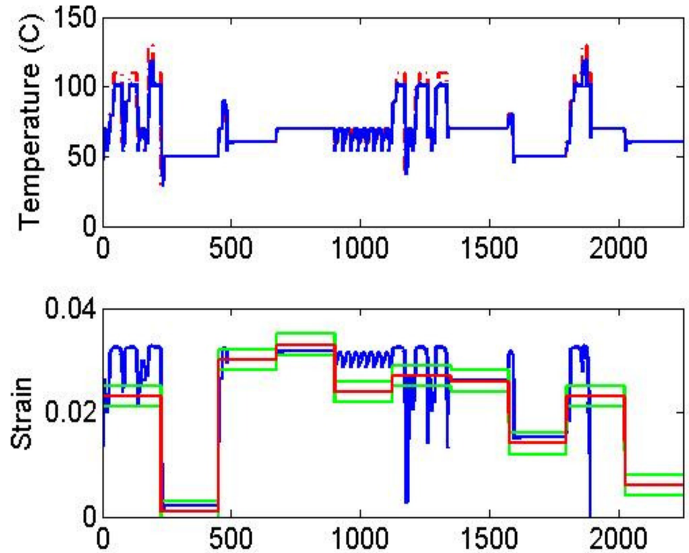
$$\frac{dT}{dt} = \frac{\frac{V^2}{R} - h\,\pi\,D\,L(T - T_\infty)}{m\,C_P}$$

- Learns input-output data (how to apply voltage to achieve a particular position state), not the constitutive model of an SMA

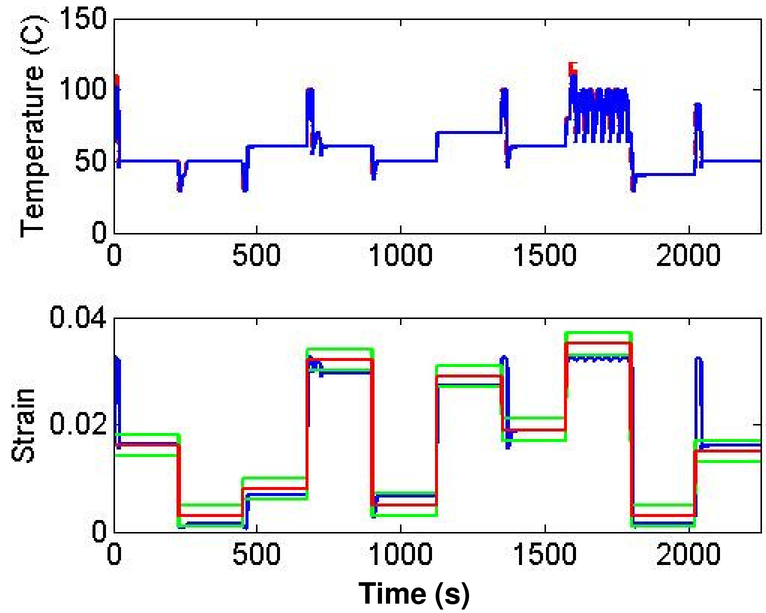


Simulation Results

Strain Learning Only



Strain and Temperature Learning



- **Goal States: Random Strain Goals**
- **Error range: ± 0.2% Strain**
- **Episodes: 50,000**

Conclusions

- Measuring temperature and including it in the information composing the current state allows the Sarsa algorithm to converge for the SMA learning problem
  - The SMA hysteresis is a non-Markovian environment when only strain is considered in state information
  - Measuring temperature alters the system and provides the Markov Property

- With the improved description of the state-space, Sarsa is able to converge to a near-optimal control policy in finite time using the Sarsa algorithm.

Challenges and Open Problems

- Open Problems
  - Expanding scope from wire (one spatial dimension displacement) to surface (two and three-dimensional displacements) would expand these methods to greater generalities.
  - Learning how to control arrays of 1-D wires needed to make 1-D SMA actuators useable for morphing.
  - Design and construct SMA-based actuator for implementation on a Morphing UAV.

- Challenges
  - Extension to higher dimensions produces much longer learning times due to spatial complexity.
  - The ability to heat 2-D and 3-D SMA objects is difficult to accomplish through electrical resistance.
  - May need separate actuator designs for each degree of freedom being morphed.

Acknowledgements

- NASA Headquarters as part of the University Research & Education Training Institute (URETI) program

Texas Institute for Intelligent
Bio-Nanomaterials and Structures for Aerospace
Vehicles

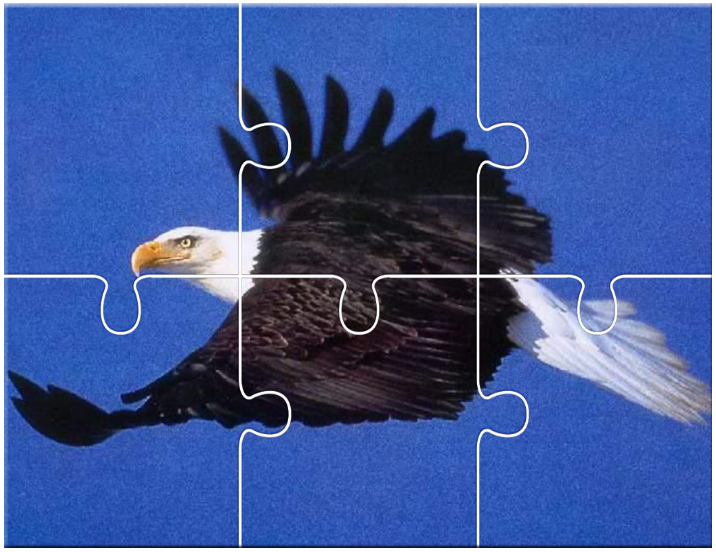
- National Science Foundation ***Graduate Research Fellowship Program***

- Texas A&M University ***Undergraduate Summer Research Grant*** (USRG)

- This work was sponsored (in part) by the Air Force Office of Scientific Research, USAF, under grant/contract number FA9550-08-1-0038. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

This support is gratefully acknowledged by the authors

Questions?

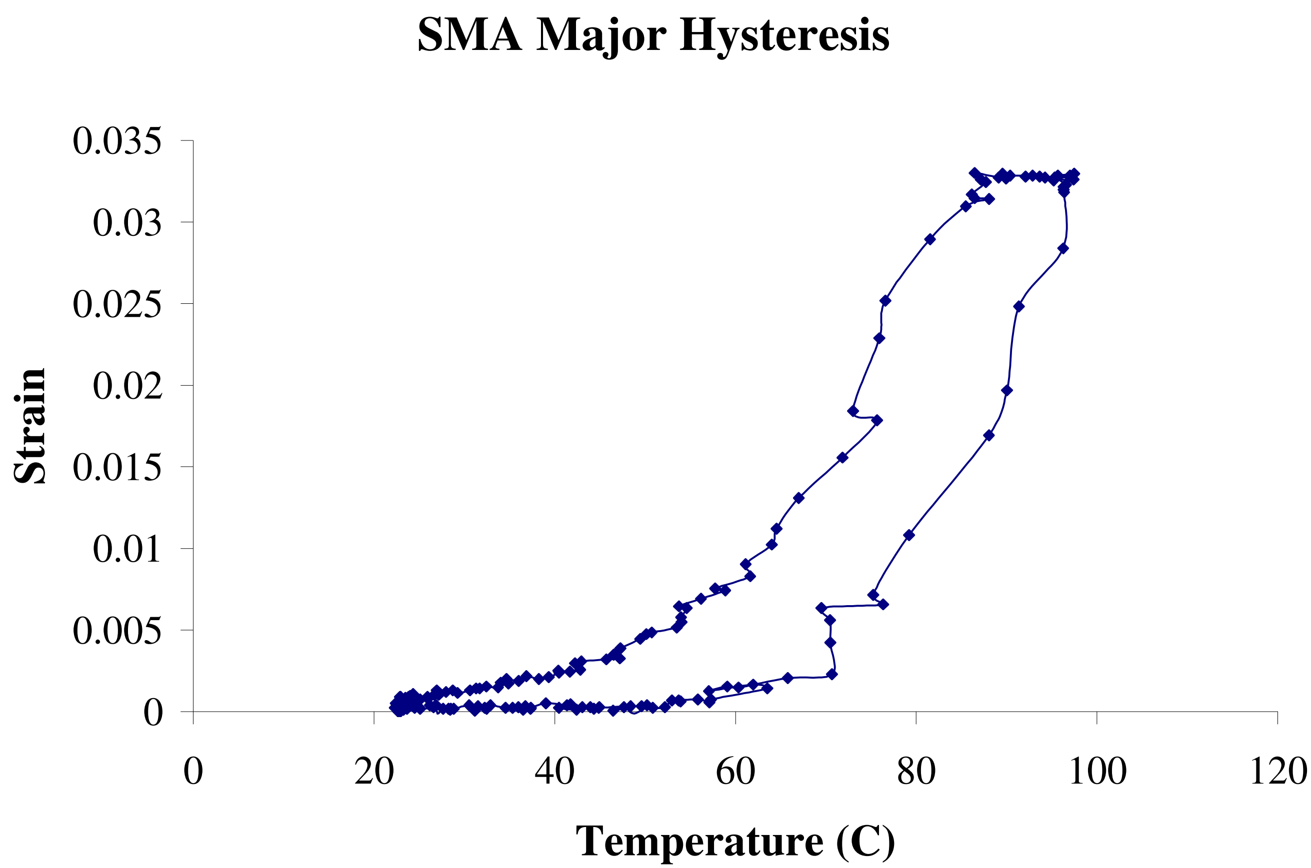


SMA Characterization Methodologies

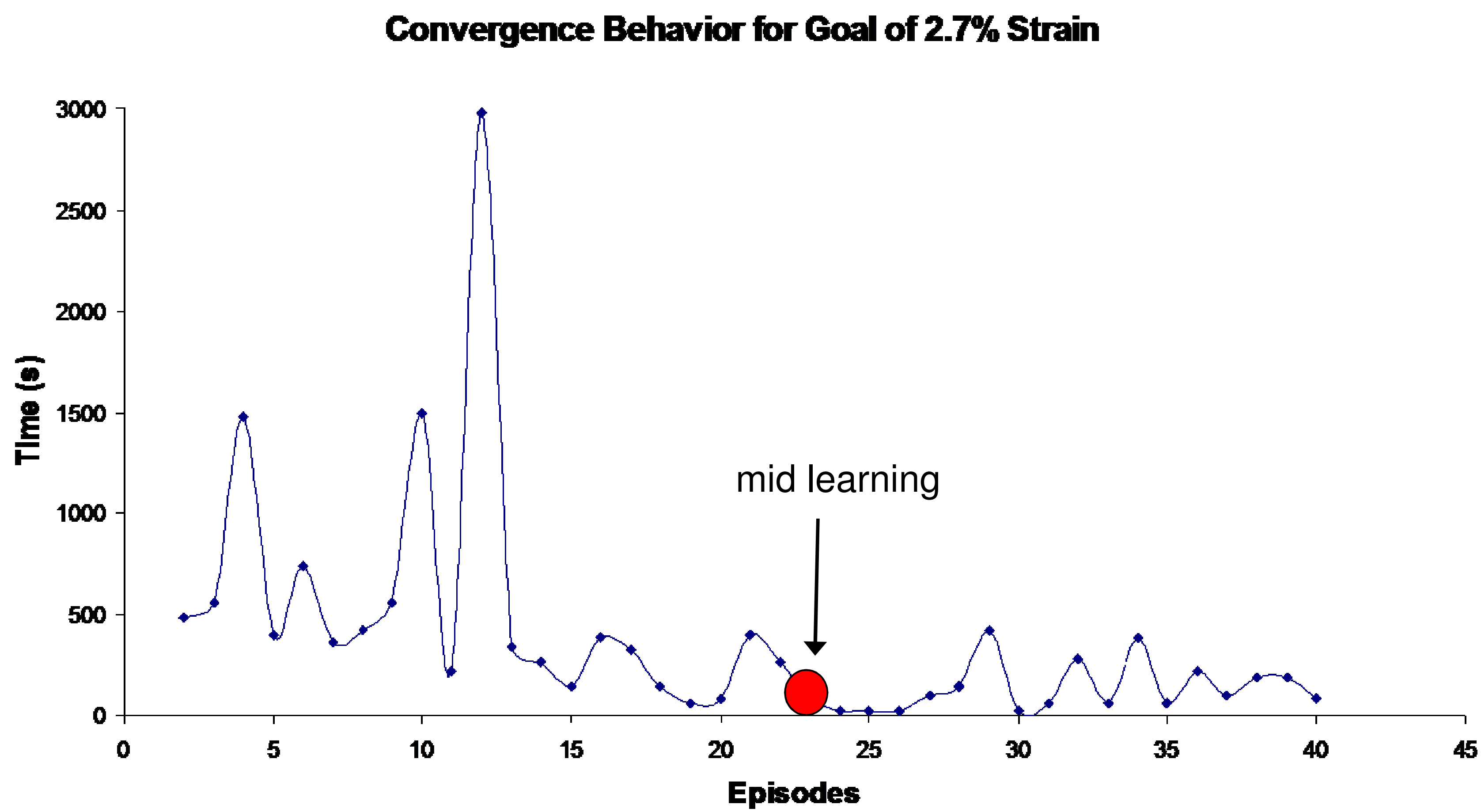
CURRENT METHOD: Experimentation and Testing to approximate constitutive model

- Mainly based on material property parameters OR system identification
  - Approximate Models:
    - Neglect hysteresis
    - Arrange antagonist SMAs to “cancel” the hysteretic effect
    - Modified plasticity model composed of averaged thermal effects
    - Disregard coupling of hysteretic and structural response
    - Other averaging techniques

Experimental Forced Characterization: Water



Experiment – Goals v. Time



Experiment – Mid Learning Refinement