October 2010

Investigating Human-Algorithm Interactions in Supervisory Control Settings

Missy Cummings

missyc@mit.edu |Humans and Automation Laboratory| http://halab.mit.edu

Some Background

Human Supervisory Control

Human Operator
(Supervisor)

Controls

Displays

Computer
(Vehicle)

Actuators

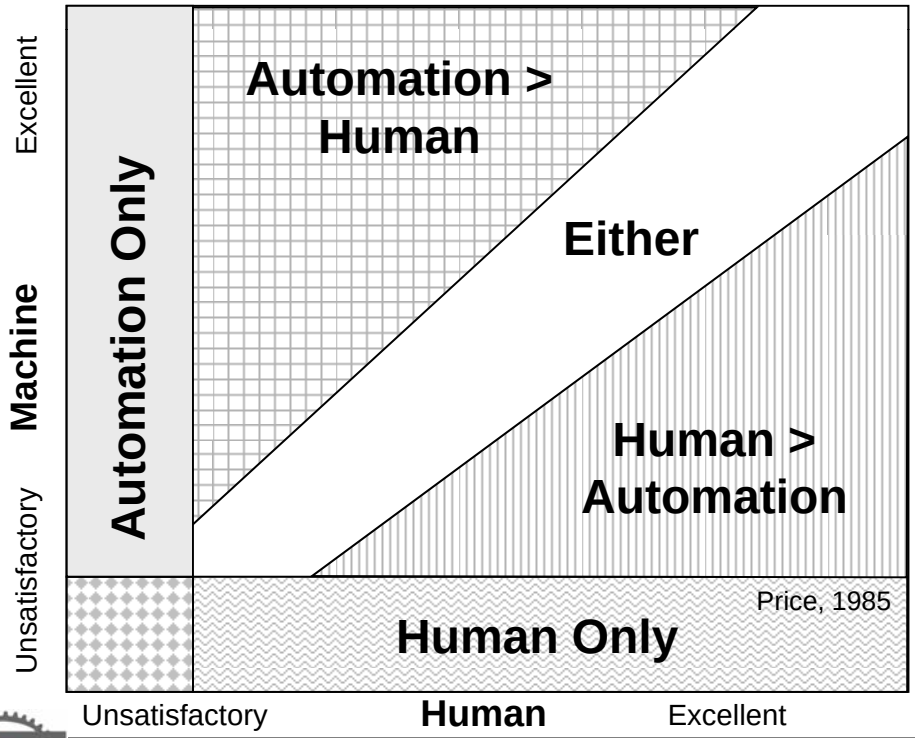
Sensors

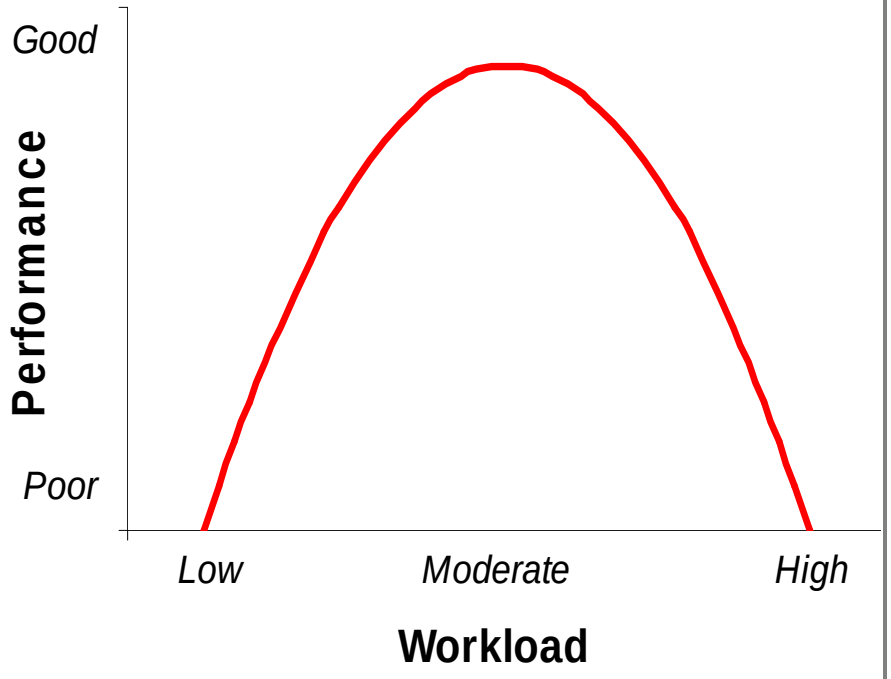
Task

- Complex, time-pressured, high risk domains
- Systems require embedded autonomy with human supervision
- Individuals & teams

The need for human-automation collaboration

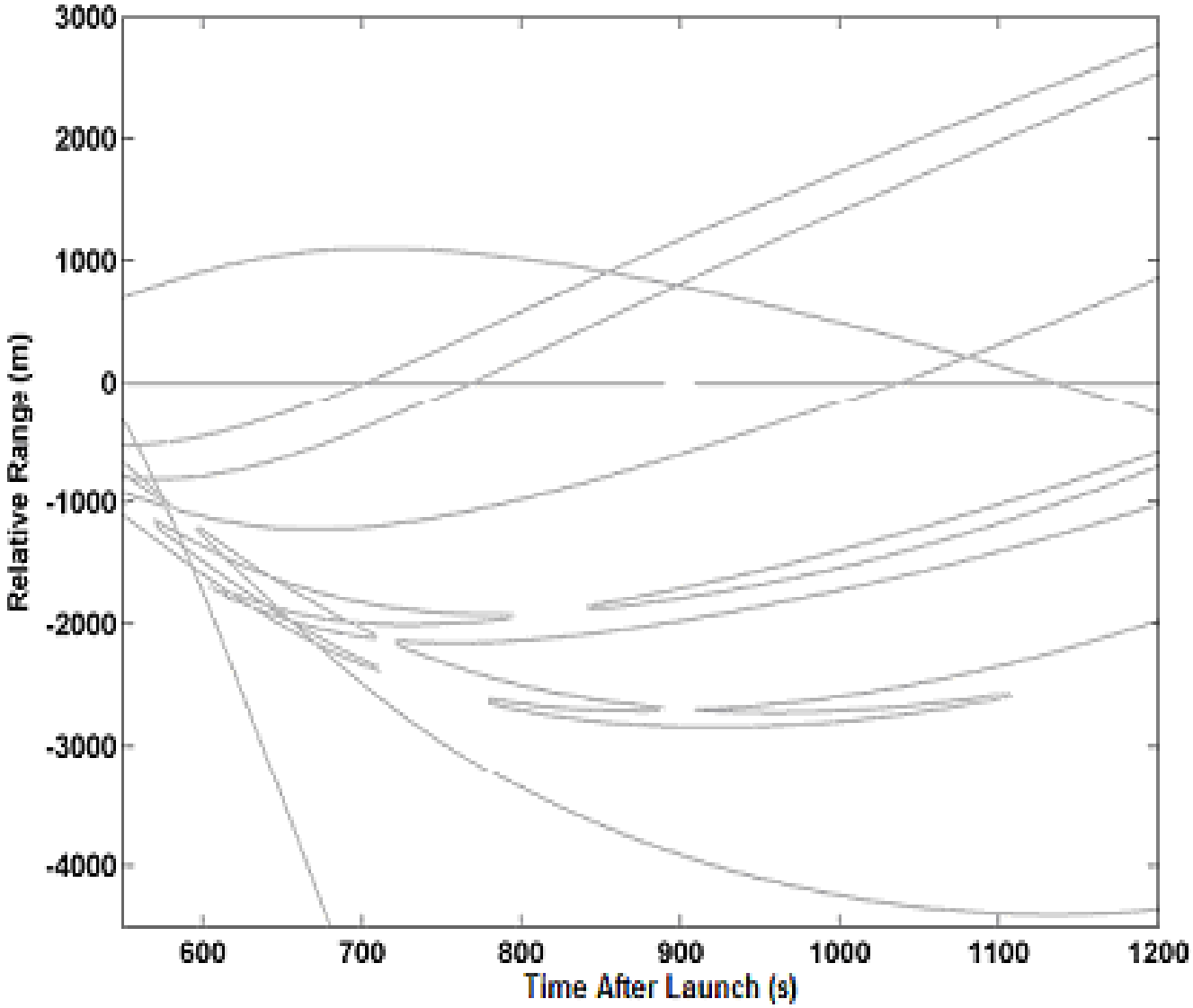
- Growing system complexity and mission requirements mean that mutually exclusive role allocation between humans and automation will not be sufficient
- Humans and automation both have strengths & limitations
  - Brittleness vs. attention inefficiencies





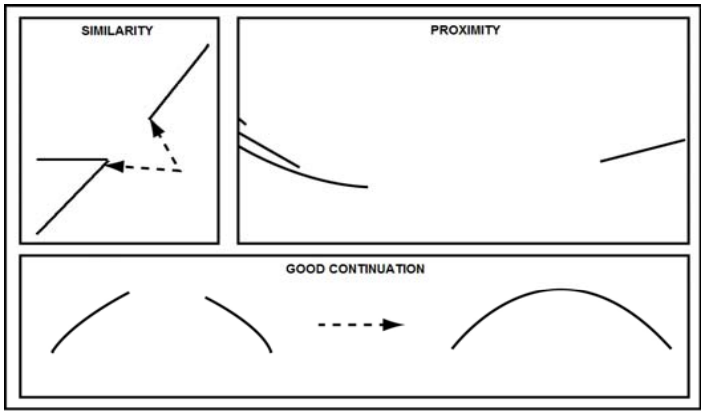
Humans helping algorithms “see”

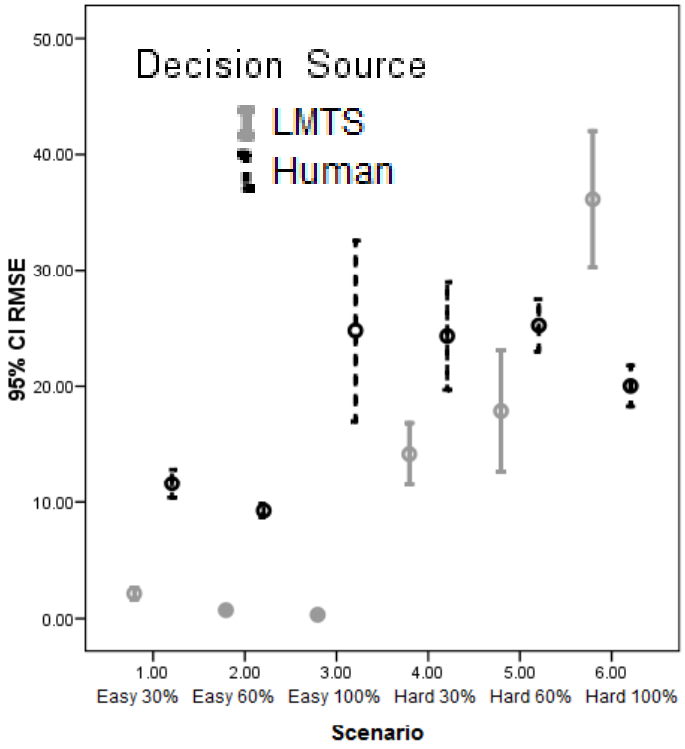
Human-Automation Collaboration in Track Smoothing



The power of perceptual-based reasoning

- Two experiments with data gap and crossing angle factor levels
  - 30/60/90, shallow/steep
- 17% & 26% people did as well or better than algorithm
  - Algorithm did better than all participants only 42% & 28% of the time.
- Interpolation vs. extrapolation
- Clear function allocation criteria

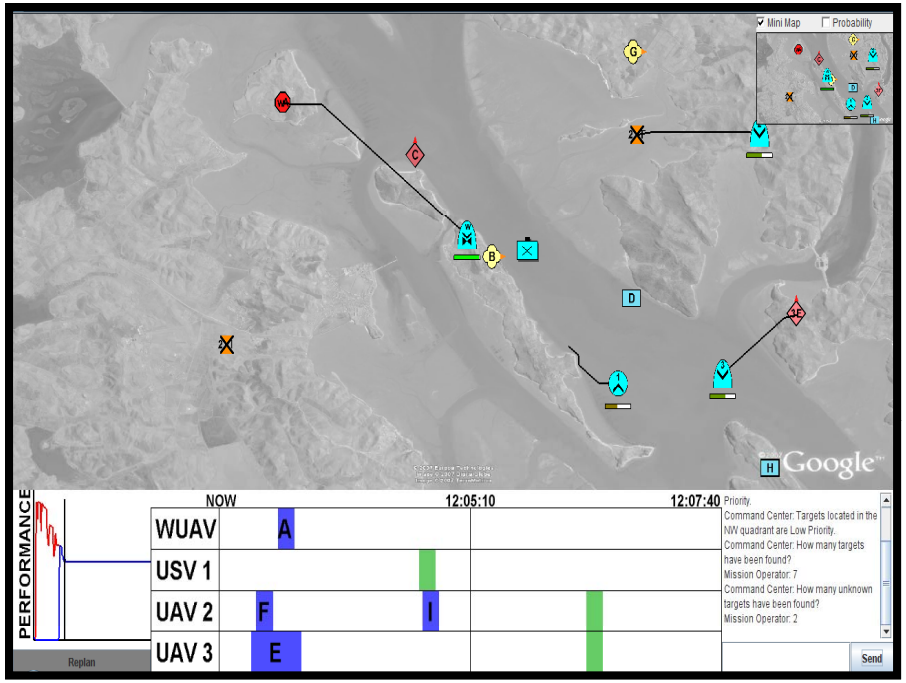




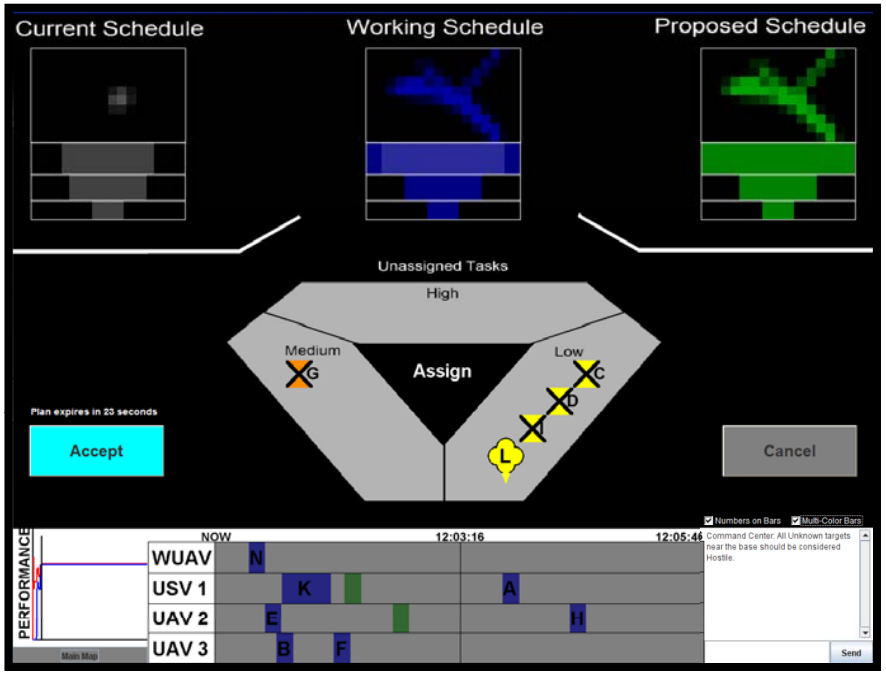
Algorithms helping humans by reducing workload, humans helping algorithms optimize

Single Operator-Multiple Unmanned Vehicle Control

Map Display



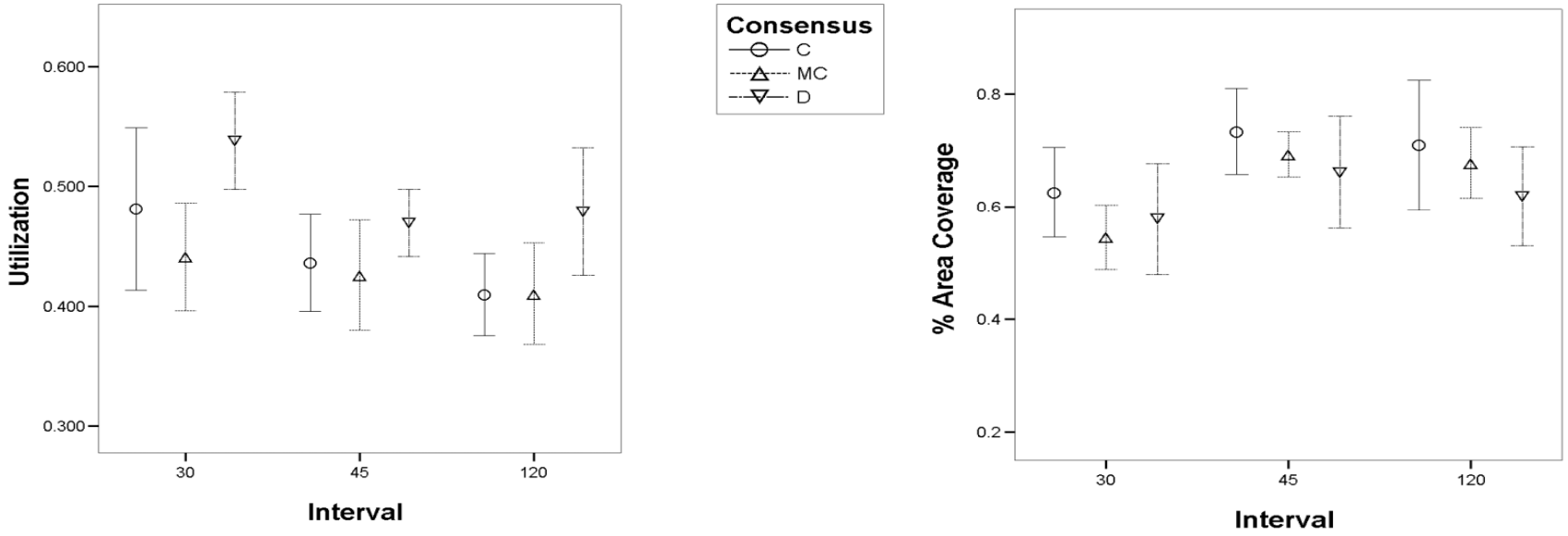
Schedule Comparison Tool



- Task-based control vs. vehicle-based control
  - MILP planners (ACL lab/Aurora Flight Sciences) can replan ~ every 10 secs
- Humans are needed for plan approval
- How often should automation have human in the loop?
  - 30, 45, 120s replanning interval over 10 min experiment w/ 30 subjects

The Role of Consensus

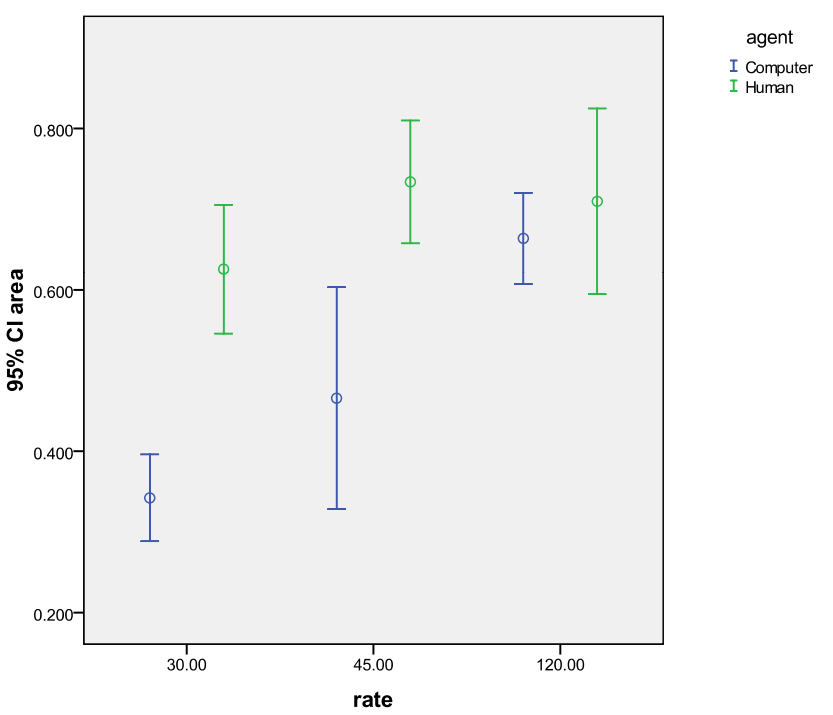
	30s Replan Interval	45s Replan Interval	120s Replan Interval
Rate Dissenters	32 ± 6.9s	37 ± 13.6s	38 ± 11.7s
Mixed Rate Consenters	35 ± 6.5s	36 ± 8.8s	65 ± 26.1s
Rate Consenters	33 ± 6.3s	46 ± 6.3s	**79 ± 18.7s**

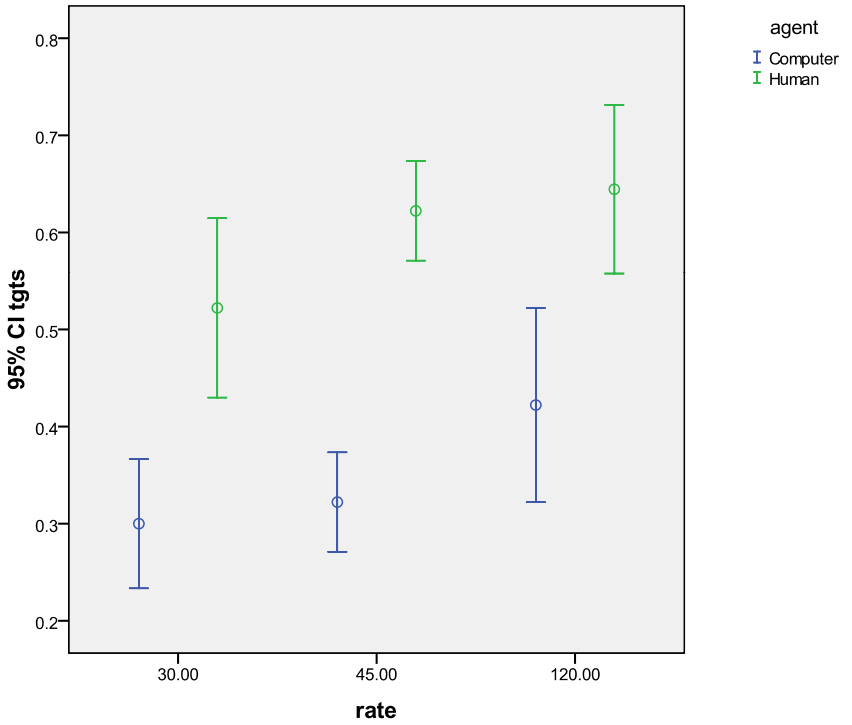


Gaming experience mattered…

Human-algorithm assistance

- What if we have a perfectly compliant human?



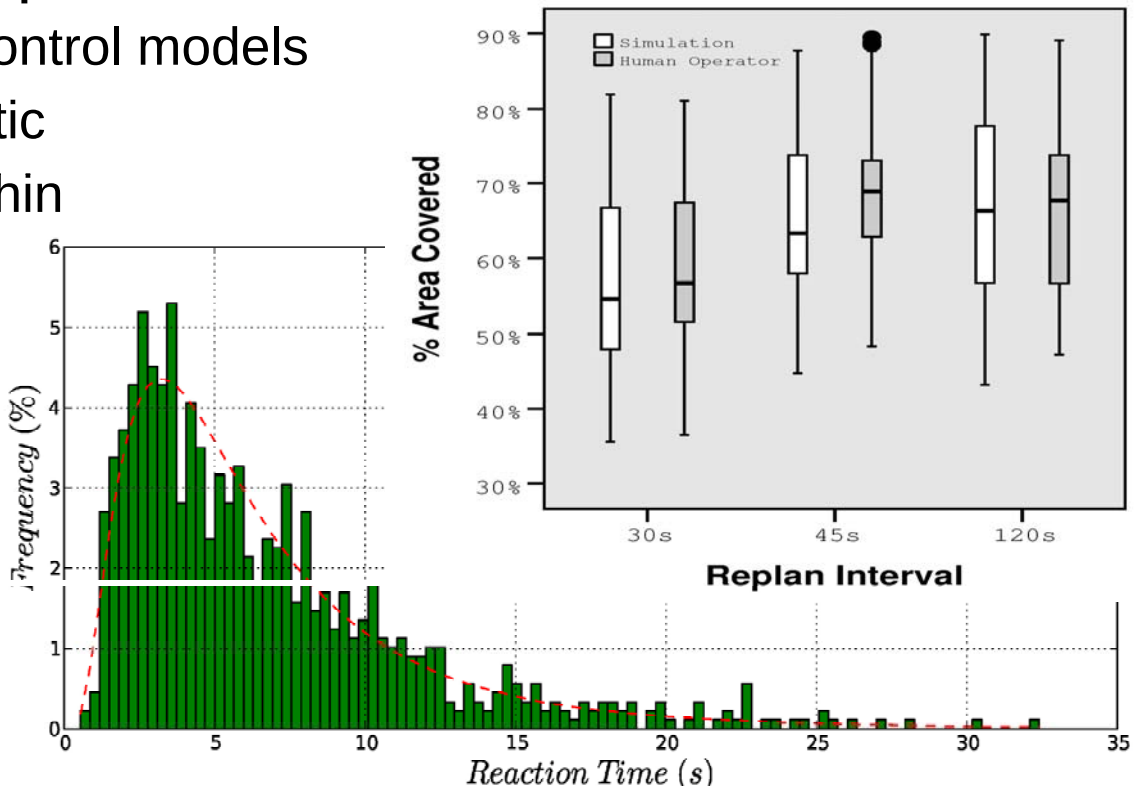


Human search guidance was critical

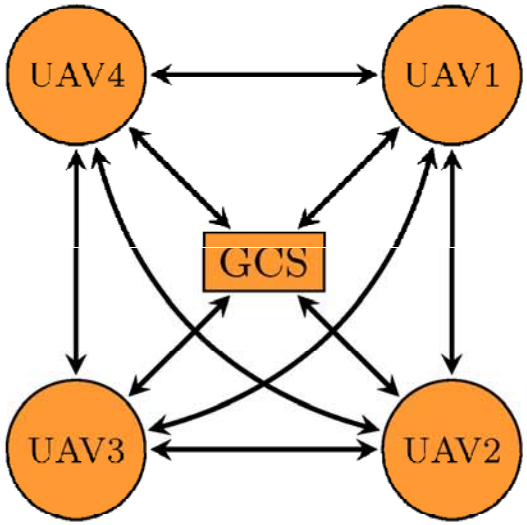
Modeling the operator in the decentralized UV control world to investigate network performance

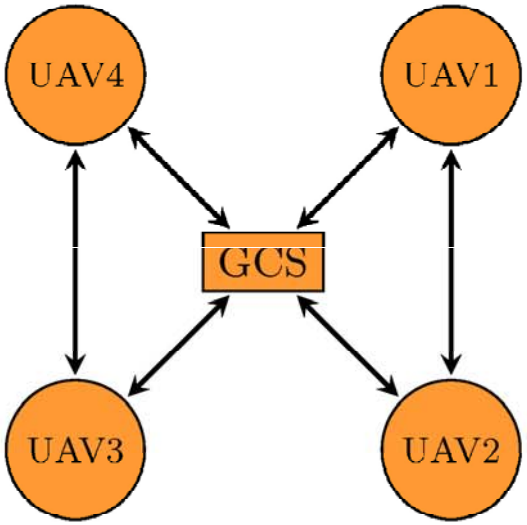
Operator Modeling

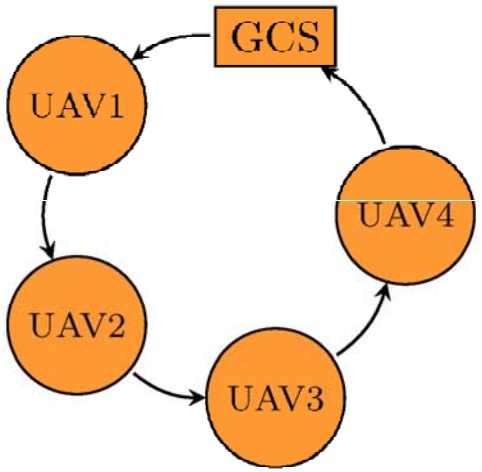
- Critical for investigation of ANY supervisory control system metric performance
  - Expensive and time consuming
- High level vs. low level operator models
  - Manual vs. supervisory control models
  - Stochastic vs. deterministic
  - Variable performance within and across humans
- No one size fits all
- For decentralized UV sim, probabilistically modeled:
  - Reaction times (planning & comms), error rates, attention allocation schemes, management strategies

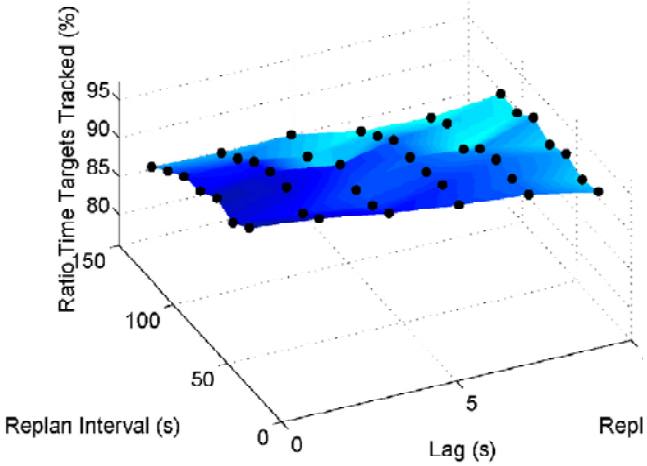


Investigating joint system performance

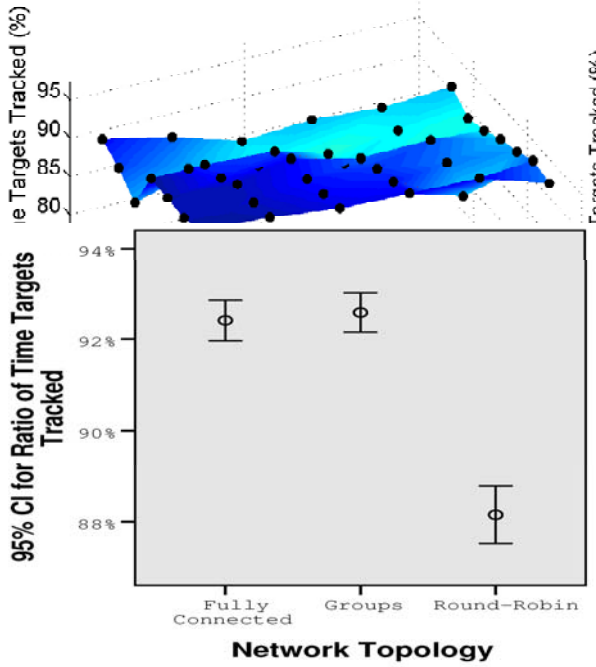


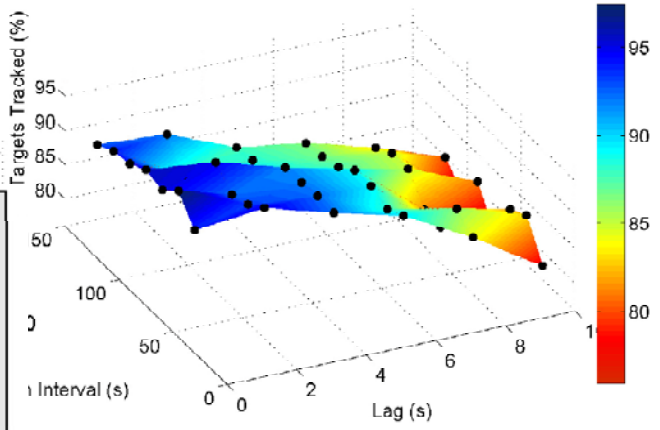






(a) Fully Connected Network





(c) Round-Robin Network

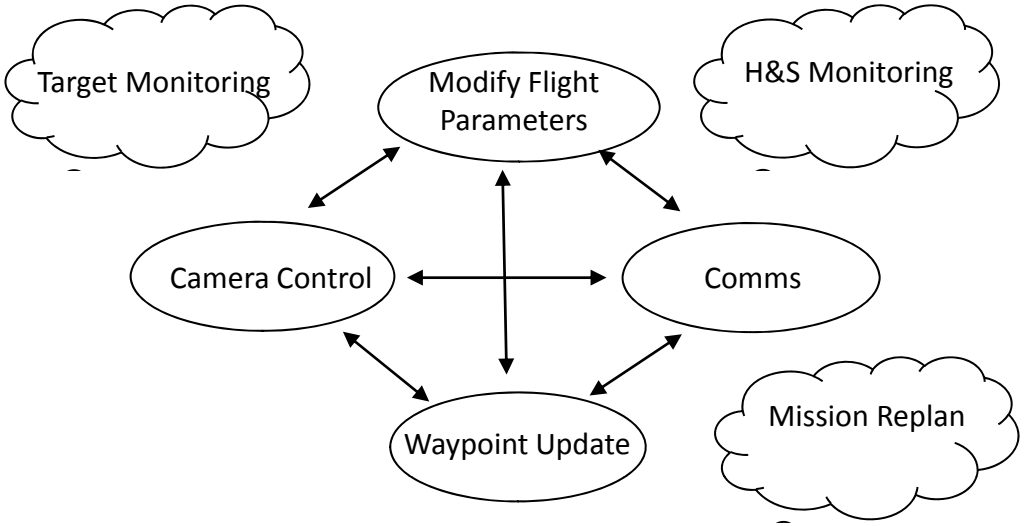
Automation helping supervisors monitor and assess personnel performance

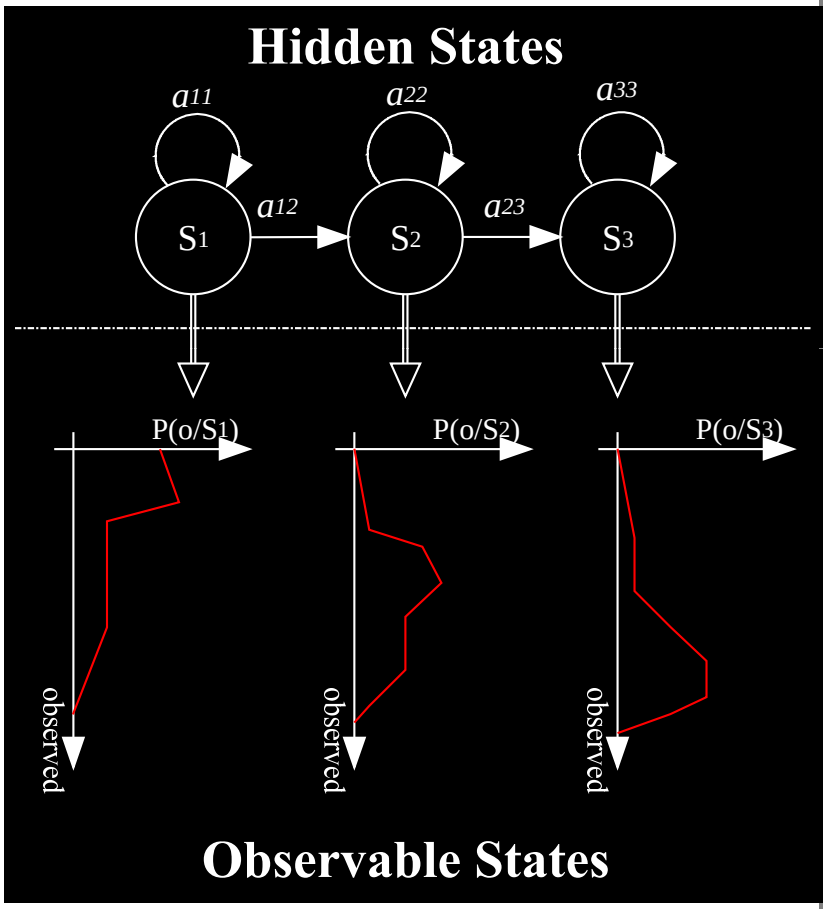
Motivation

- Supervisors of controllers of automated systems need assistance in determining when to intervene
- Such supervisory control human behavior is generally procedure oriented and event driven, with a temporal component.

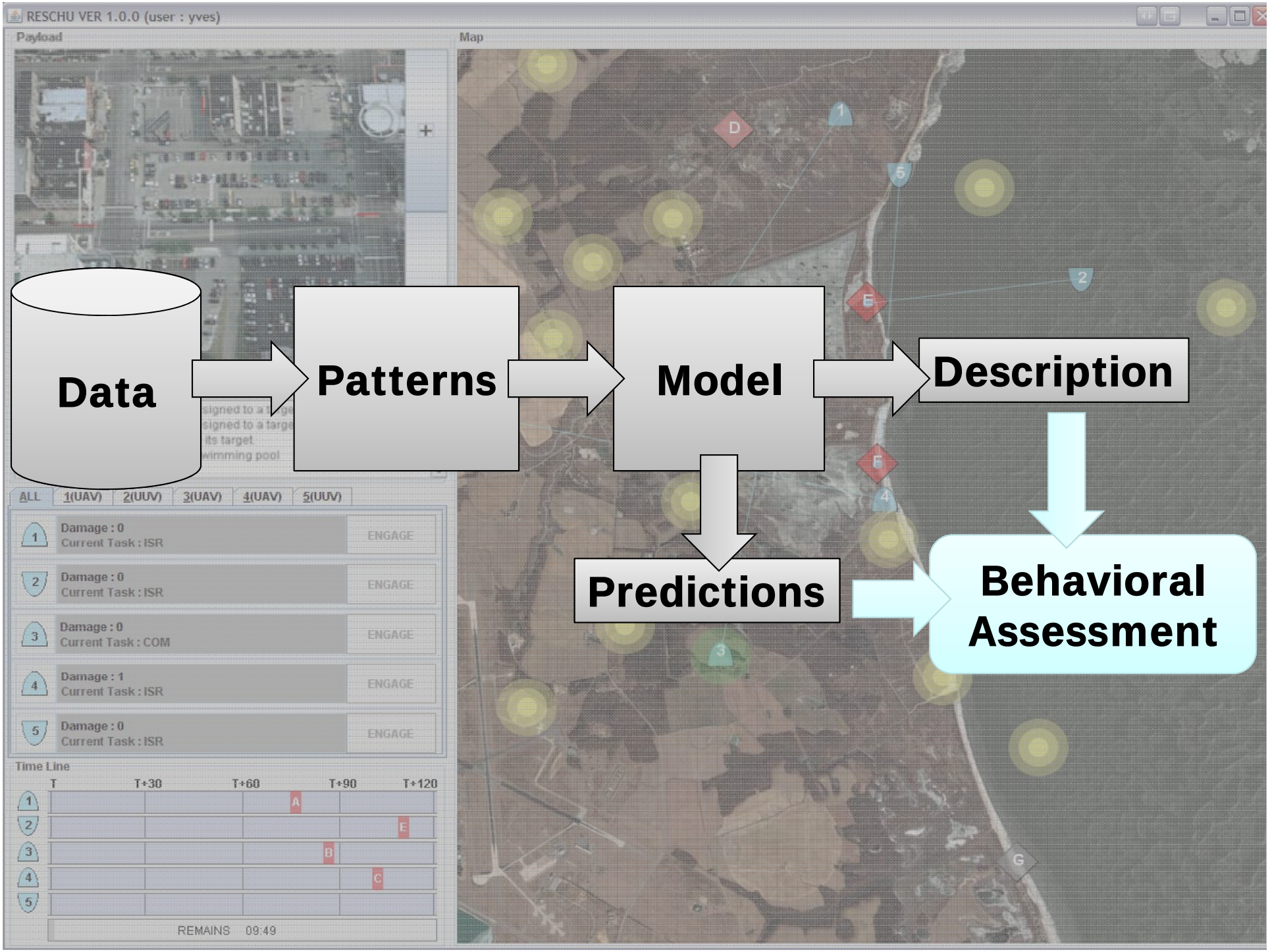
- Machine learning can be used to identify the patterns in the event/time series
- Adherence to procedures as patterns can be computed from a normative model

HMMs in UAV Control Example

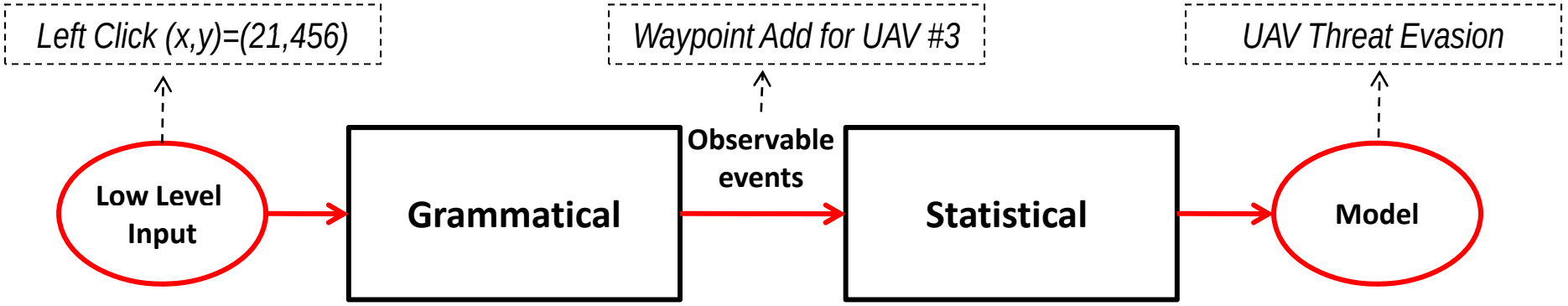




- Observables vs. Hidden States
  - Hidden states are not directly visible, but variables influenced by them are.
- Action vs. Task
- Event vs. Behavioral Pattern
- Doubly stochastic Bayesian Process
- Supervised vs. unsupervised


RESCHU VER 1.0.0 (user : yves)
Payload
Map
Data
Patterns
Model
Description
Predictions
Behavioral Assessment
ALL
1(UAV)
2(UUV)
3(UAV)
4(UAV)
5(UUV)
Damage : 0
Current Task : ISR
ENGAGE
Damage : 0
Current Task : ISR
ENGAGE
Damage : 0
Current Task : COM
ENGAGE
Damage : 1
Current Task : ISR
ENGAGE
Damage : 0
Current Task : ISR
ENGAGE
Time Line
T
T+30
T+60
T+90
T+120
REMAINS 09:49

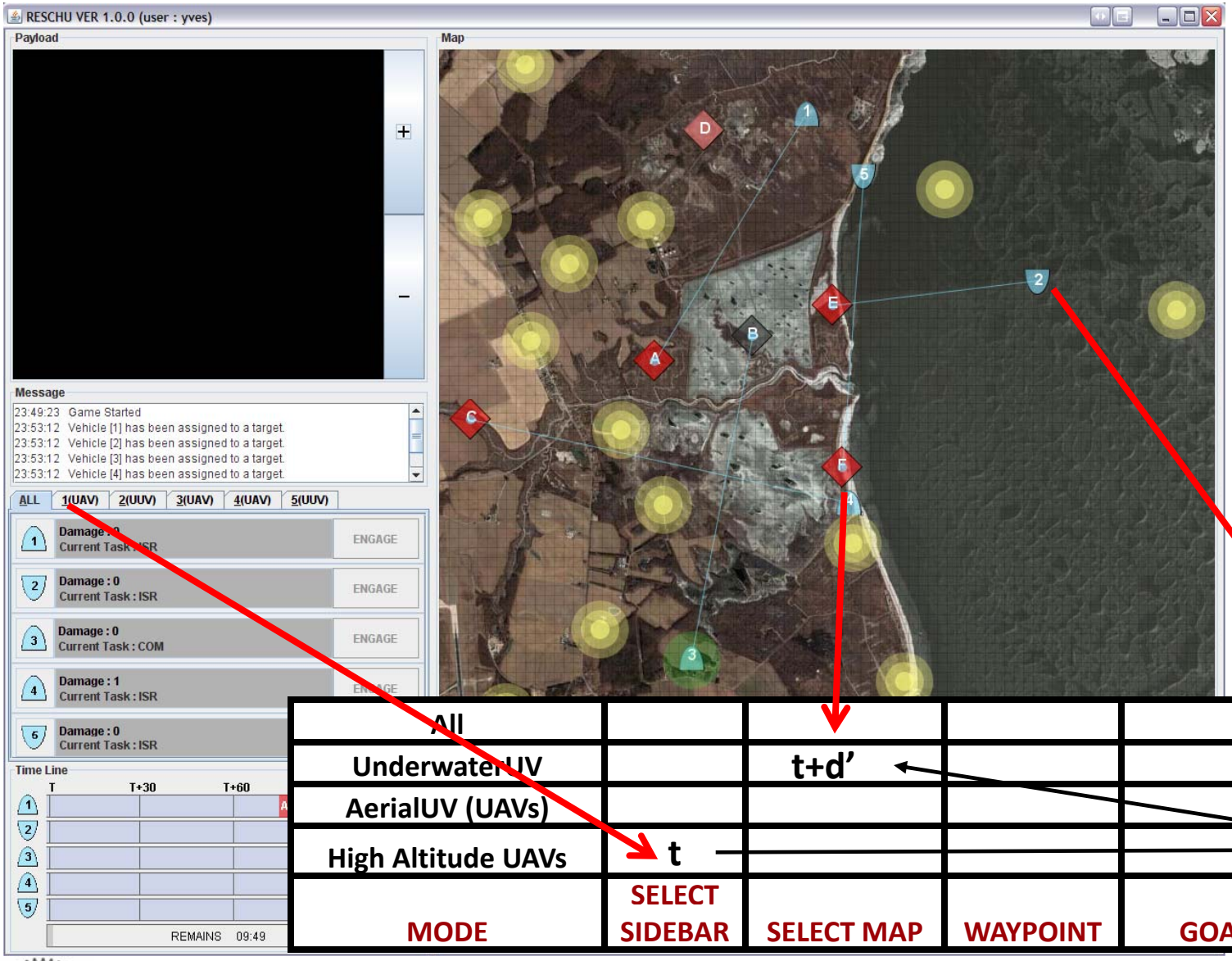
Behind the (Modeling) Curtain…



- **Grammar**= transforms real-world events into observable events usable by the pattern recognizer (state space reduction), obtained by Cognitive Task Analysis

- **Pattern Recognizer** = recognition and prediction
  - "Simple": Hidden Markov Models (HMMs)
  - Time dependent: Hidden Semi-Markov Models (HSMMs)

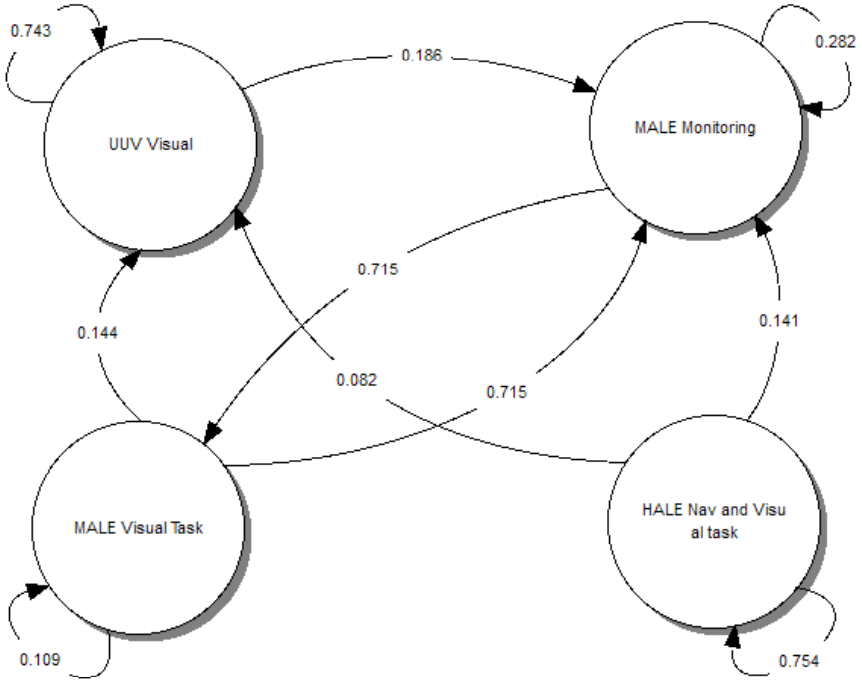
Grammar

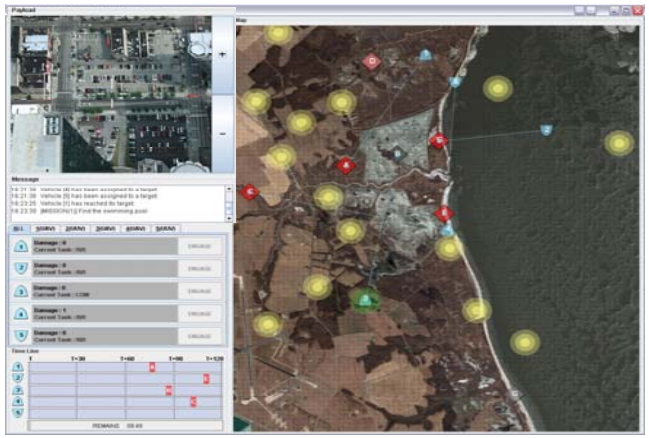


MODE	SELECT SIDEBAR	SELECT MAP	WAYPOINT	GOAL	VIZ/ENGAGE
All					
UnderwaterUV		t+d’			
AerialUV (UAVs)					
High Altitude UAVs	t				t+d

HMM Results

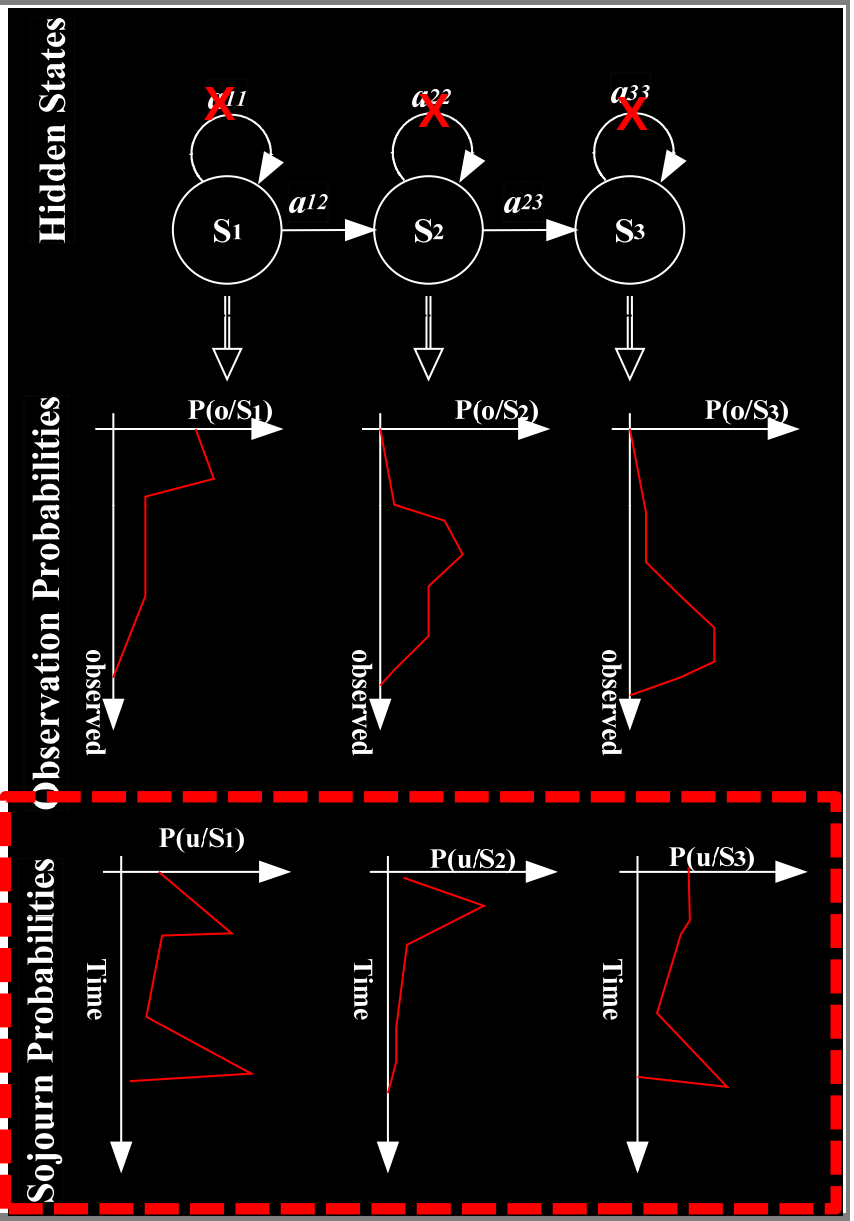
- HMMs can predict next state transition with ~80% accuracy
- Nice but not particularly useful in the real world...
  - Time is critical





Hidden Semi-Markov Models

- HMM + sojourn time distribution
- There are no self-transitions
- **Semi**-Markov regime
  - Relaxing the Markov assumption
- Issues
  - Complexity
    - Both in computation and generalizability
  - Which is more important, time of transition or next likely state?
  - How to turn such a model into a useful decision support tool?



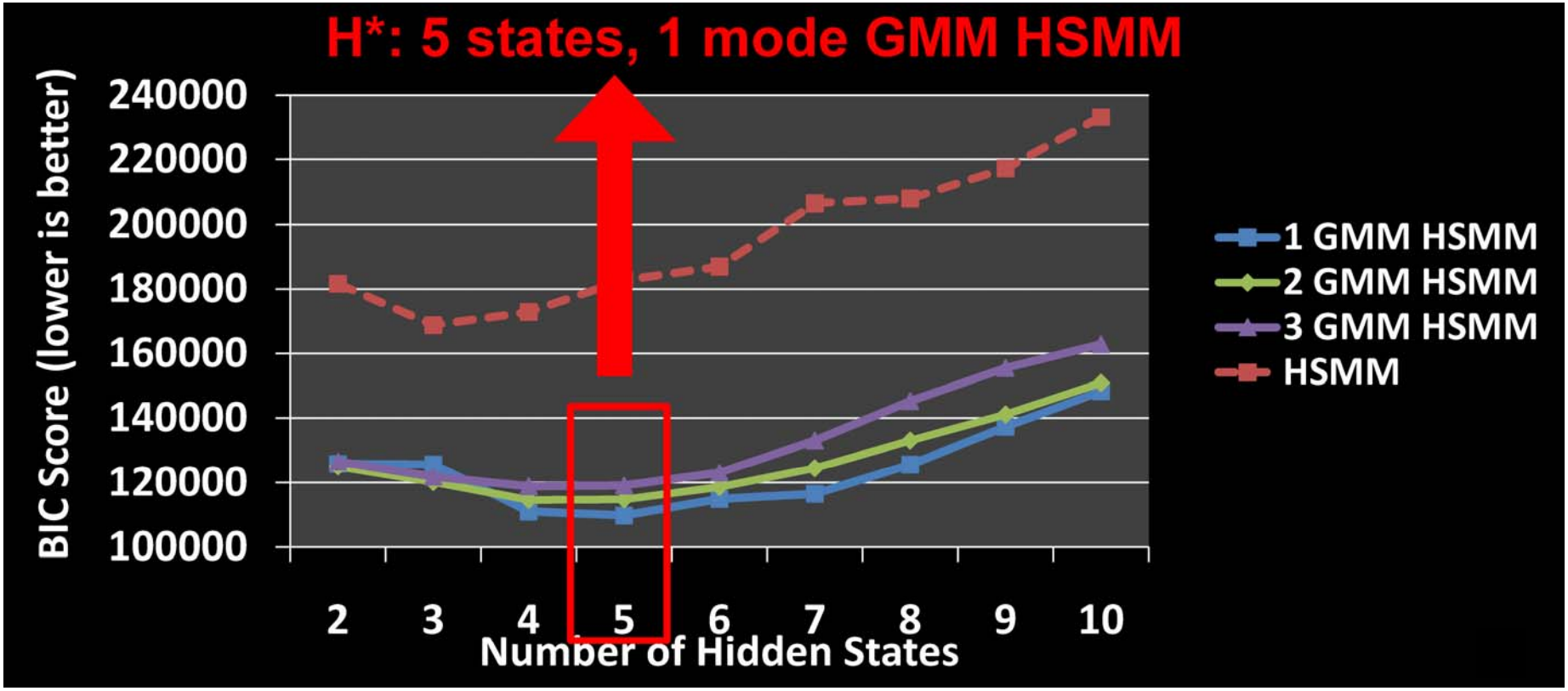
Parameter Reduction & Model Selection

- Bayesian Information Criterion

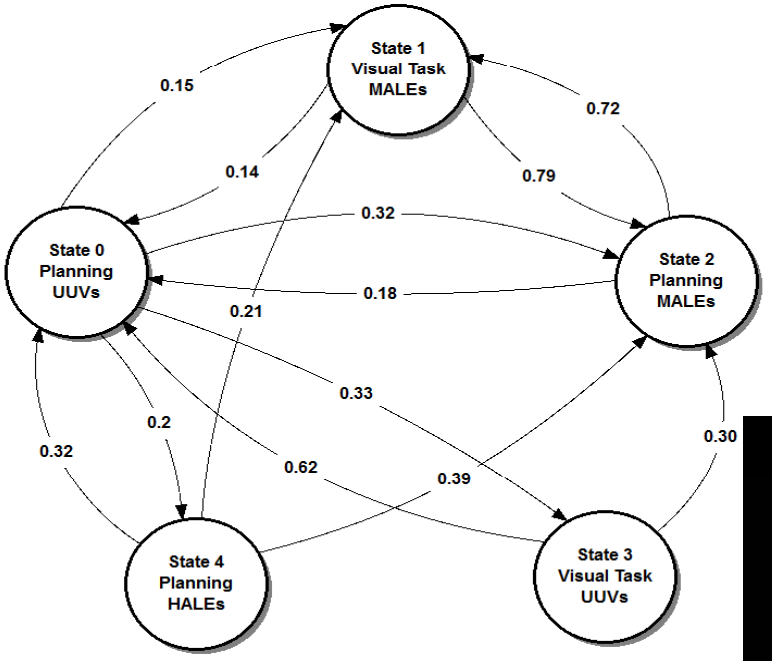
$$BIC = -2\log L(H) + P\log(K)$$

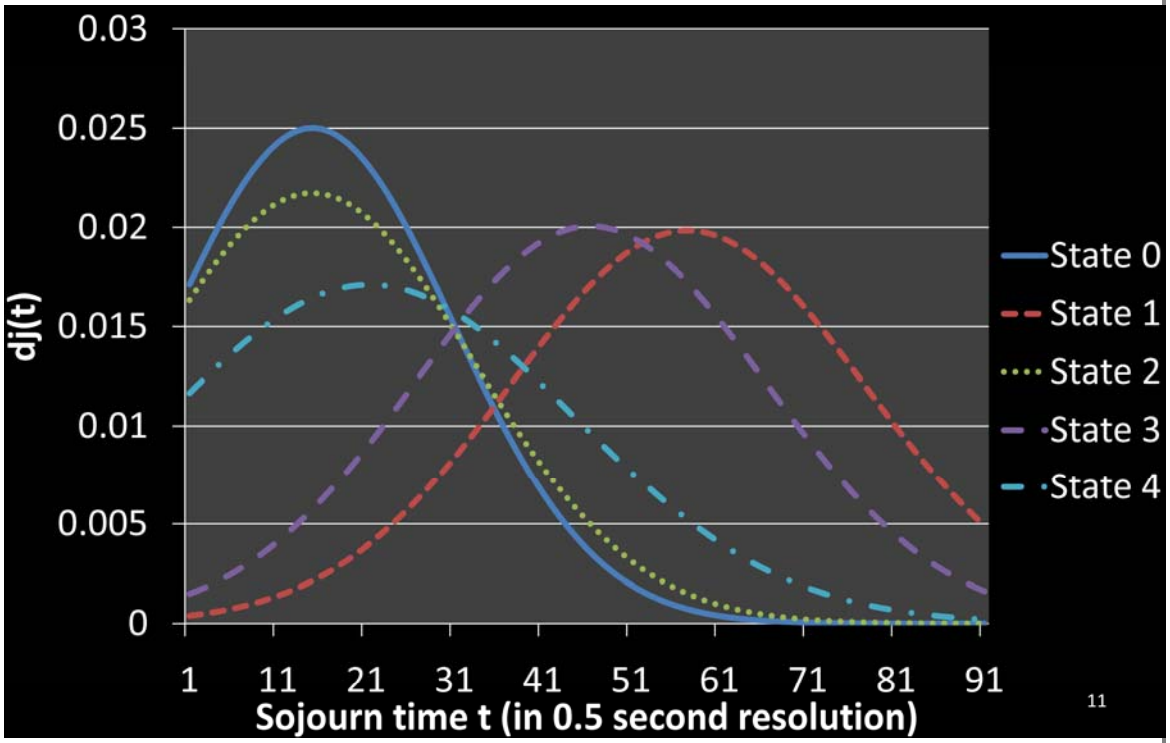
L(H): max log lik
P: num free params
K: num obs

- Gaussian Mixture Models



States and Sojourns for the 1 mode GMM model





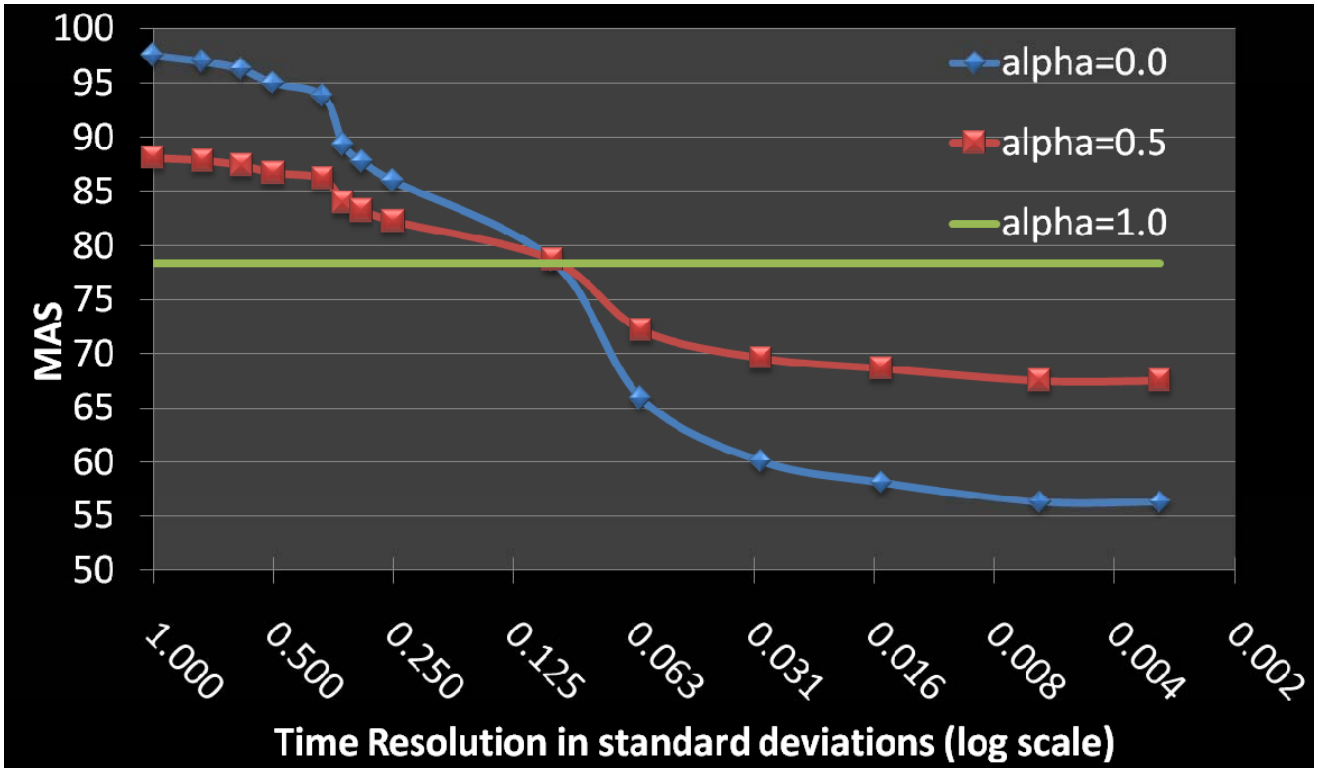
11

Prediction Metric of HSMMs

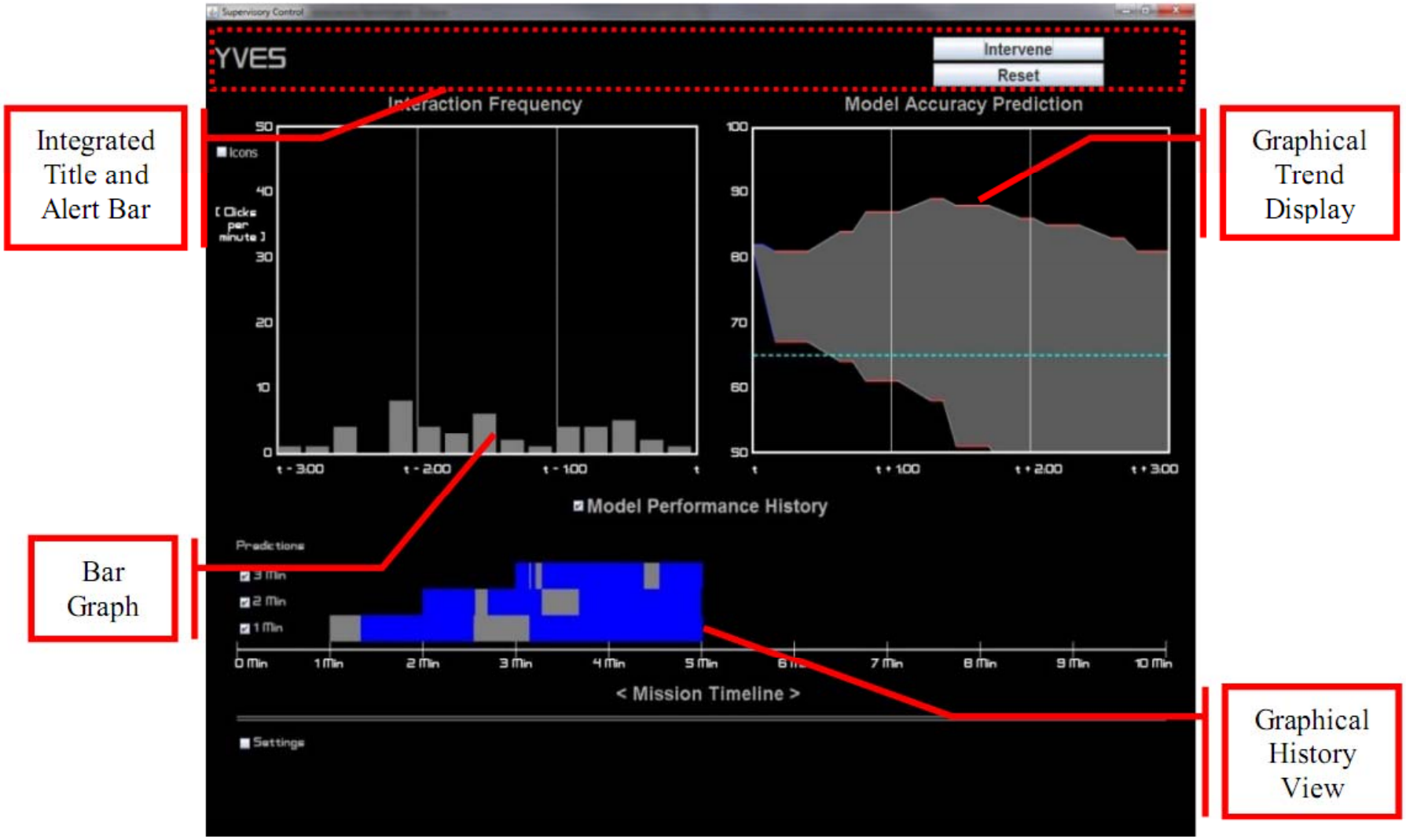
- Model Accuracy Score

$$MAS(t) = \frac{\sum_{i=t,\ldots,t-n} \alpha \times Quality(i) + (1-\alpha) \times Timing(i)}{n}$$

- Inherently a cost-benefit & potentially subjective metric

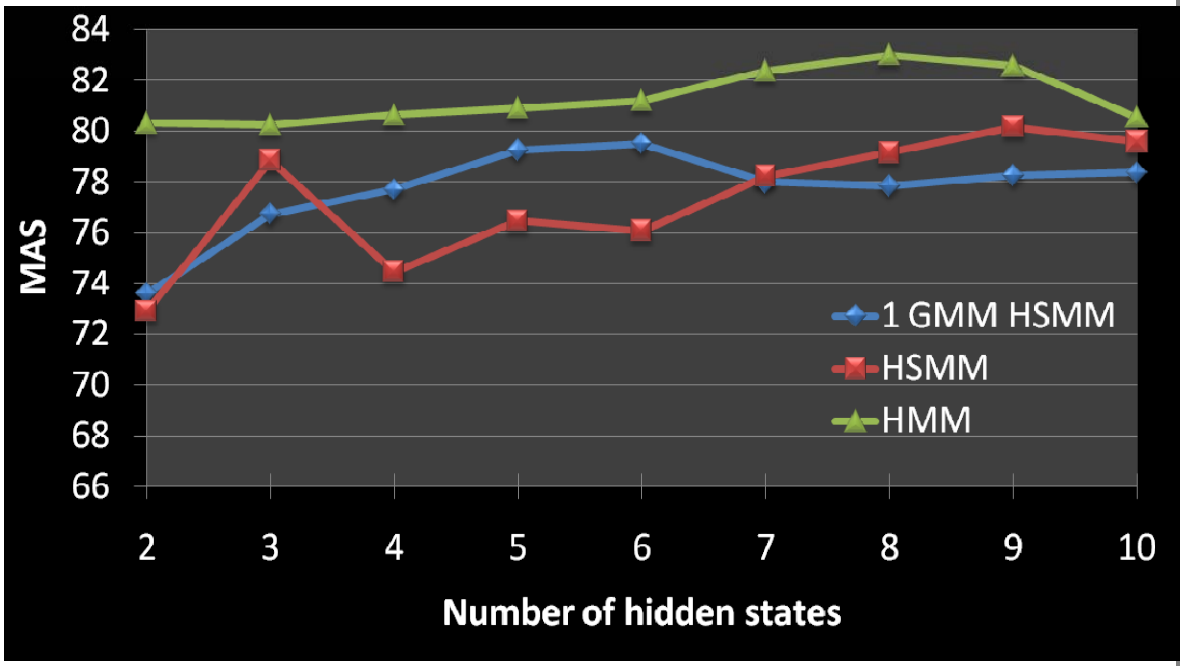


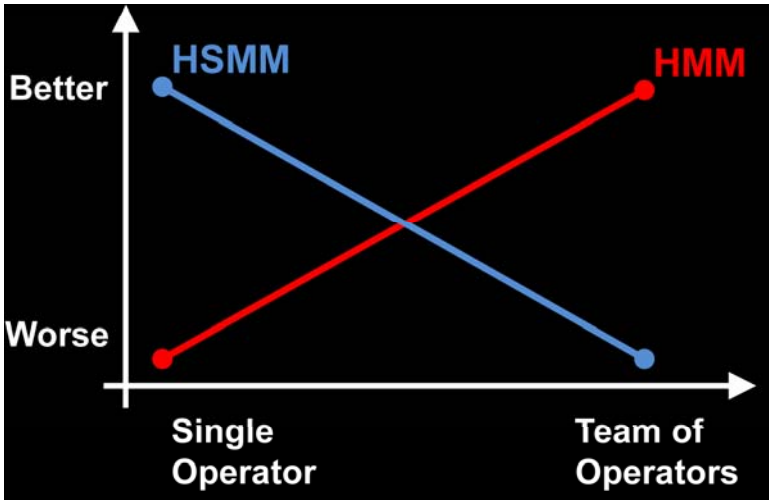
Real-Time Decision Support?



The path ahead?

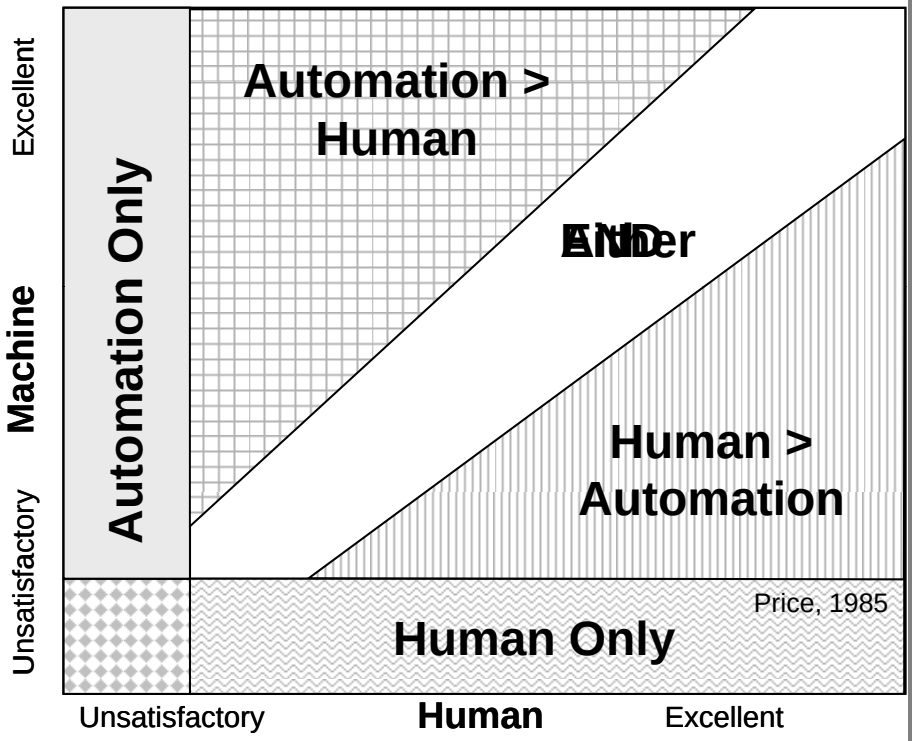
- Real-time vs. off-line use
  - Model sensitivity
  - Prediction horizons
  - Alpha weighting in real time
  - Decision support viability
- Individuals vs. teams of operators – can this scale?
- Training evaluation
- New PhD





Meta-Analysis

- Algorithms and humans can be complementary in many settings
- Humans are good at:
  - Filling in data/information gaps
  - Assessing goodness/badness of off-nominal predictions
- Automation is good at:
  - Handling large quantities of data quickly
  - Detecting trends not immediately observable to humans
- Uncertainty and planning time horizons are the key factors to consider



Questions?