BEST: A Benchmarked Experiential System for Training

Kevin Gildea, Georgiy Levchuk, Jared Freeman, Aptima Inc.
Shruti Narakesari, Wayne Shebilske, WSU

Presented at 12th CCRTS
Date: June, 21, 2007

www.aptima.com
Woburn, MA ▪ Washington, DC

© 2007, Aptima, Inc.

The Easy Problem

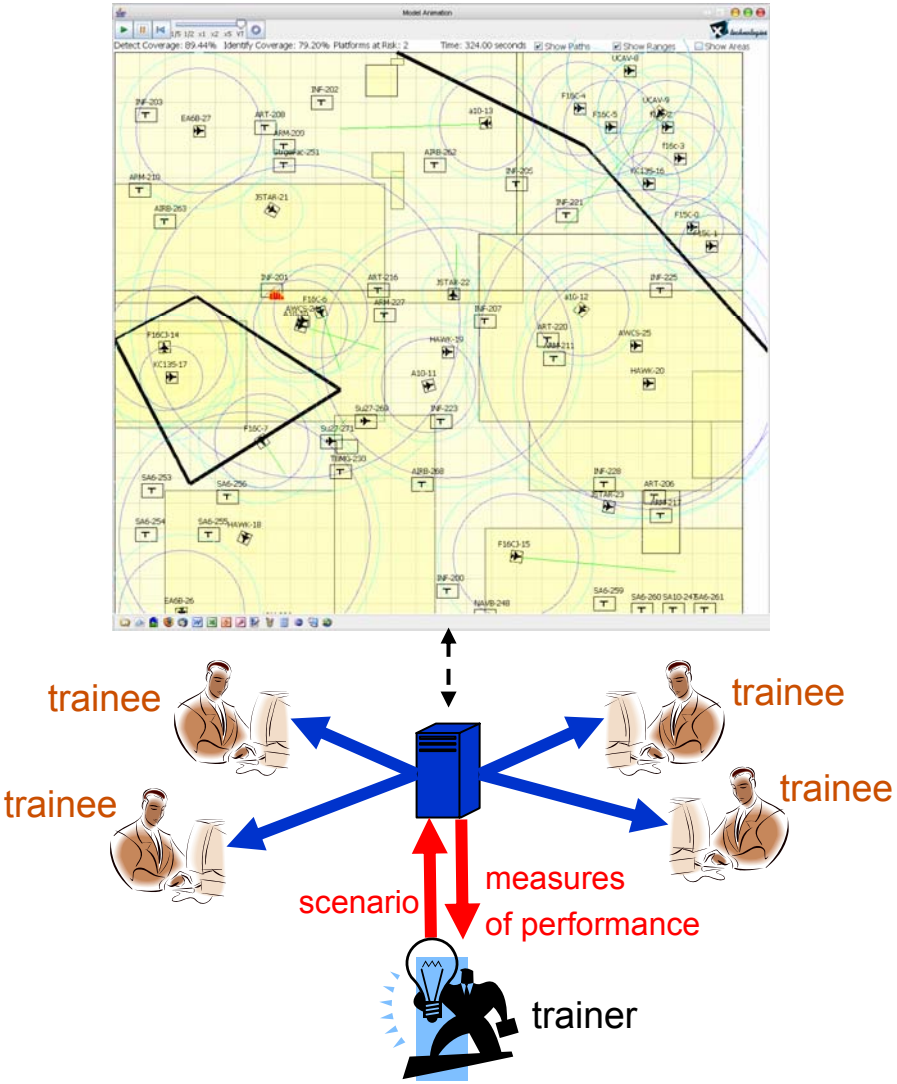
- How do we recognize **expertise**?
  - Performance
- How do we know it is **high**?
  - Benchmarking
- How can we **achieve** high expertise?
  - Training

© 2007, Aptima, Inc.

Our Domain

- **AF Dynamic Targeting Cell**

- Find targets
- Classify/ID targets
- Coordinate assets
- Approve strike packages/execute



The Complication

- **What if the task is ill-structured or complex?**
  - Can make **mistakes in recognizing achieved expertise**
  - Do not know **impact of training**
- **Why?**
  - Multiple types of expertise combine to affect performance
  - Incorrect/biased observer ratings or measures
  - Weak knowledge of expertise dynamics

© 2007, Aptima, Inc.

What Happens to Training?

- You cannot train what you cannot measure
  - How do you give training if you do not know **achieved expertise** and **impact the training** will have?
- Different teams = different learning curves

© 2007, Aptima, Inc.

State of Training Systems

- ITTS systems rarely model team learning, typically model
  - Expert team members (Miller, et al., 2000)
  - Coaches for individuals (Freeman, et al., 2005)
- Qualitative I/O models of team learning are … qualitative so cannot drive training sims
- Quantitative I/O models of team learning do not drive training sims (Kozlowski, et al., 2001)

© 2007, Aptima, Inc.

Our Contributions

- **Definition of expertise & dynamics**

- **Assessment:** benchmarking via optimal mission execution solution

- **Improvement:** intelligent training system
  - ***Optimization <u>in</u> instructional strategy:*** Train the team with expert-selected, annotated, and animated near optimal solutions delivered in feedback
  - ***Optimization <u>of</u> instructional strategy:*** Select scenarios maximizing likelihood of advancing the team most directly to goal expertise using a POMDP model

© 2007, Aptima, Inc.

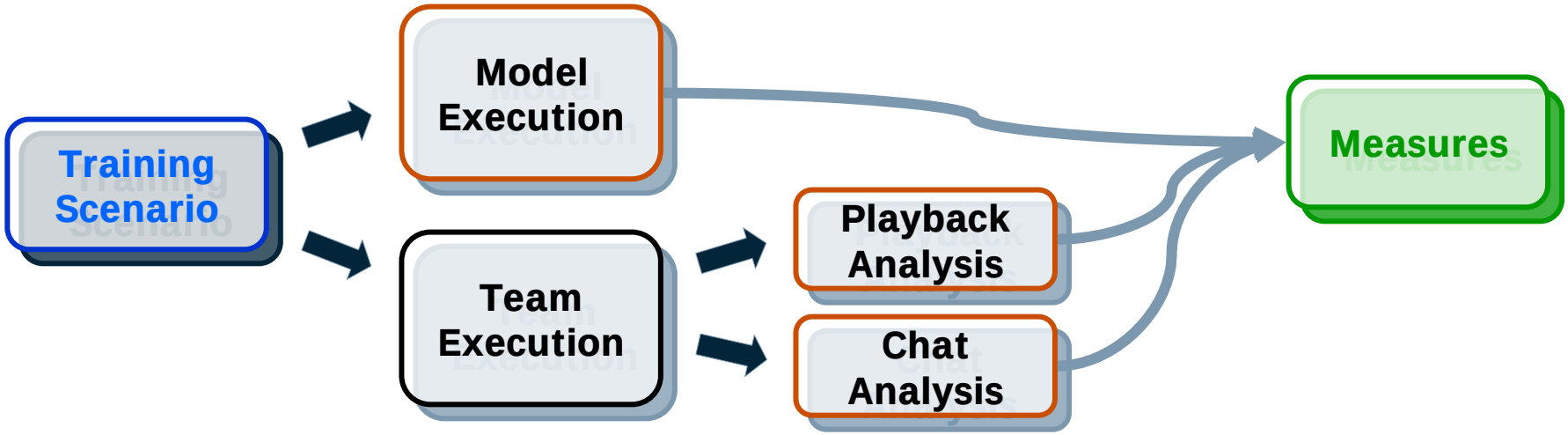
Contribution 1: Expertise State & Mission Representations

- Expertise (1-2 vectors, with 1=inadequate; 2= adequate):
  - ISR maintaining low risk
  - ISR maintaining high coverage
  - ISR nominating and DTC designating TSTs
  - DTC prioritizing TSTs
  - DTC coordinating a strike package plan
- Mission:
  - # Time sensitive targets
  - # Enemy defensive threats

© 2007, Aptima, Inc.

Contribution 2: Assessment

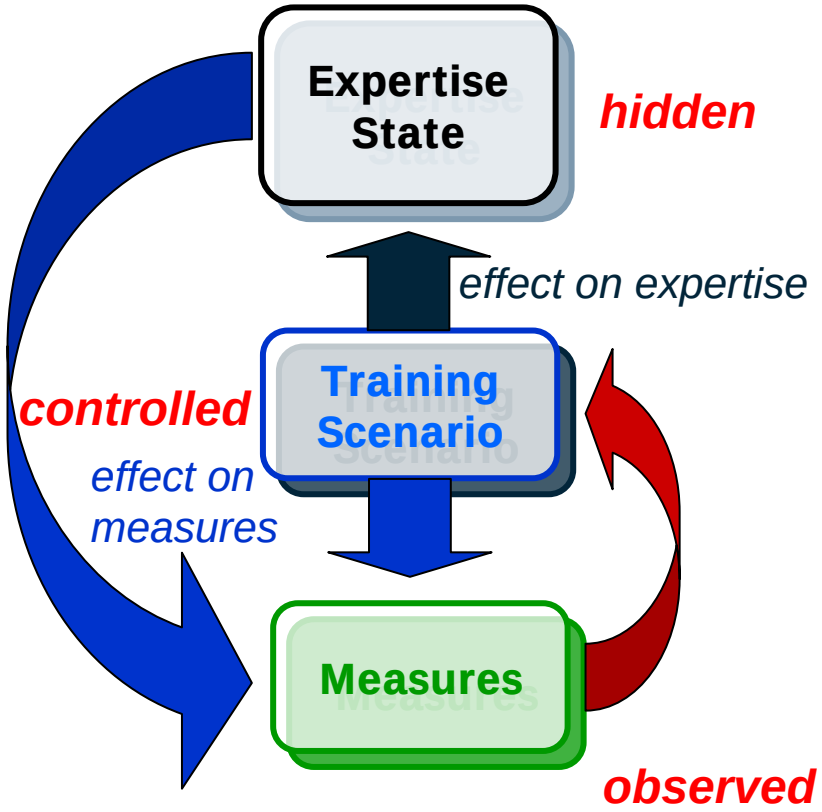
- Scenario Benchmark: Near Optimal Solution Model
  - Agent Model
- Human Performance
  - Analysis of DDD log files
  - “Playback” Model
  - Analysis of chat communications



© 2007, Aptima, Inc.

Contribution 3: Strategy Improvement

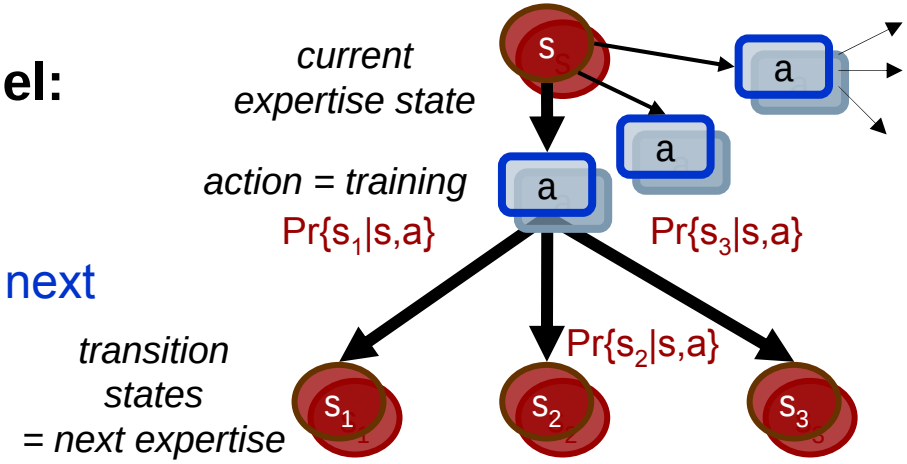
- Training = Planning under uncertainty
  - **State** = discrete expertise
  - **Learning dynamics** = *impact of training*
  - **Control** = training scenarios
  - **State observations** = measures of performance



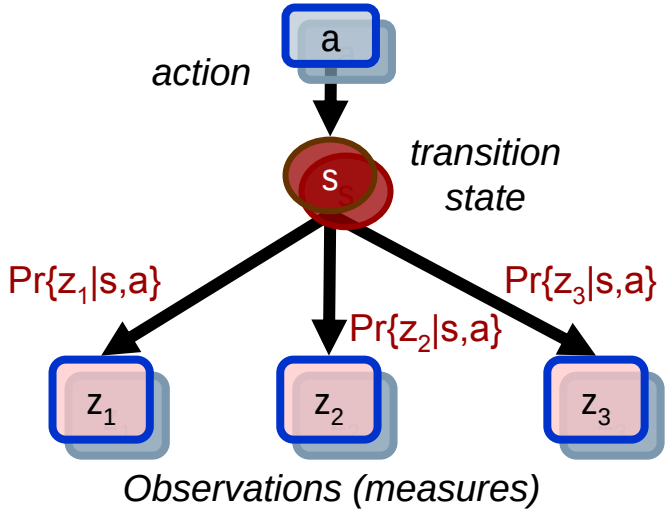
Partially Observable Markov Decision Processes

Model for Optimal Training: POMDP

- **Hidden state transition model:** how the training affects team expertise
  - Variables: Pr (current state | next state, training scenario)



- **Observation model:** how current scenario and team expertise affect what measures can be observed
  - Variables: Pr (observed measure | next state, training scenario)



© 2007, Aptima, Inc.

Comparison to Other Markov Models

Markov Models		Do we have control over the state transitons?	
		NO	YES
Are the states completely observable?	YES	**Markov Chain**	**MDP** Markov Decision Process
	NO	**HMM** Hidden Markov Model	**POMDP** Partially Observable Markov Decision Process

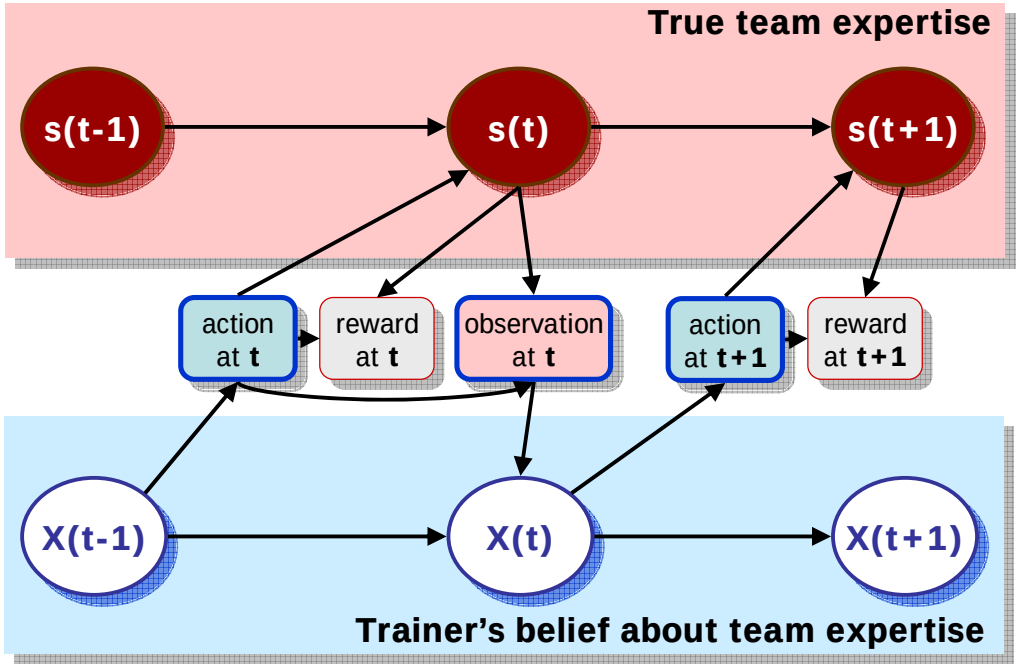
© 2007, Aptima, Inc.

How Does POMDP Work?

- Benefit of training defined using "reward" of visiting expertise state nodes
  - Want to maximize the expected total reward of training
- Can incorporate specific "training path" restrictions
  - Depends on trainer's objectives

or

$$\max E\left[\sum_{t=1}^{\infty} \gamma^t r_t\right]$$



training sessions

$$objective = \max E\left[\sum_{t=1}^{k} r_t\right]$$

expected value

reward at time *t*

How do Trainers Use POMDP? Policy Graph → Lookup Table

POMDP solution is provided as a "policy graph" = look-up table

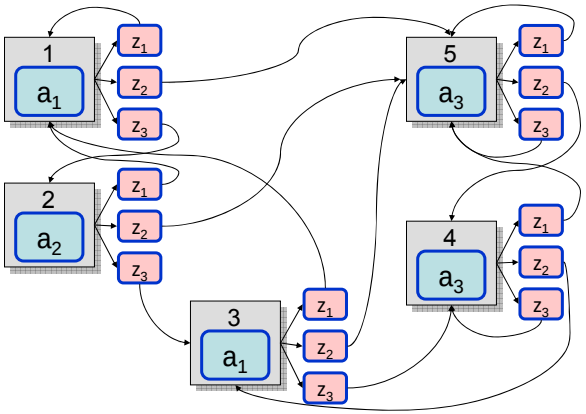
- Policy nodes = abstractions for **belief** about expertise state
  - Policy node is described using

(PolicyNodeID, StepID)

  - pair

- Each node specifies **training scenario** to be given to the team

- For each feasible observation – specify next policy node



POMDP Policy Solution

2. Receive Observation

0. Start

4. Next node

1. Use this scenario

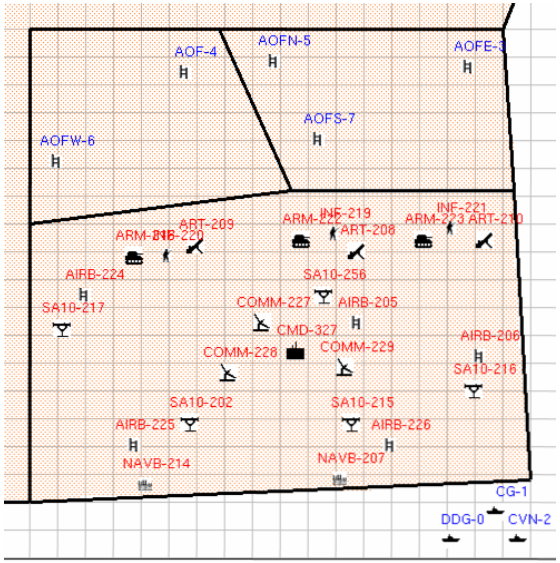
3. Mark next policy node

		What to do?		Next Policy Node ID if Observation =											
StepID	Policy Node ID	Scenario ID	Scenario Name	11111	12111	12211	12221	21111	21211	21221	22211	22212	22221	22222	NextStepID
1	1	6	d1	100	100	X	X	100	X	X	X	X	X	X	2
2	100	7	d2	18	18	23	X	18	23	X	23	X	X	X	3
3	18	7	d2	14	14									X	4
3														X	4
4					0									X	5
4						20	51	X	20	51	20	X	51	X	5
4	179	9	d4	X	X	X	51	X	X	51	X	51	51	19	5
5	19	4	c3r4	X	X	8	30	X	8	30	8	33	30	33	6
5	20	8	d3	X	X	8	30	X	8	30	8	X	30	X	6
5	40	7	d2	5	5	8	X	5	8	X	8	X	X	X	6
5	51	9	d4	X	X	X	30	X	X	30	X	30	30	33	6
6	5	7	d2	3	3	1	X	3	1	X	1	X	X	X	7
6	8	8	d3	X	X	1	2	X	1	2	1	X	2	X	7
6	30	9	d4	X	X	X	2	X	X	2	X	2	2	2	7
6	33	4	c3r4	X	X	1	2	X	1	2	1	2	2	2	7
7	1	8	d3	X	X	X	X	X	X	X	X	X	X	X	8
7	2	9	d4	X	X	X	X	X	X	X	X	X	X	X	8
7	3	7	d2	X	X	X	X	X	X	X	X	X	X	X	8

© 2007, Aptima, Inc.

Empirical Validation: Task & Team

- **Find:** ***Detect and differentiate (T.O.1)*** Time Sensitive Targets (TSTs) ISR
  - Minimize risk to assets
  - Maintain high coverage of targets
  - Nominate & designate TSTs
- **Fix:** Complex communications to detect and differentiate (T.O.1) & ***prioritize (T.O.2)*** TSTs
- **Target:** Complex communications to ***coordinate attack assets (T.O.3)***
- Approve strike package



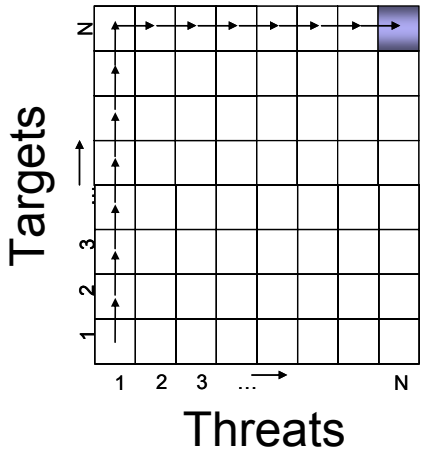
- Team: 7 Ss
  - Participants: ISR, DTC Chief, Ground Track Coordinator (GTC), Attack Coordinator (AC), Target Duty Officer (TDO)
  - Confederates: Senior Offensive Duty Officer (SODO), Chief of Combat Operations (CCO)

© 2007, Aptima, Inc.

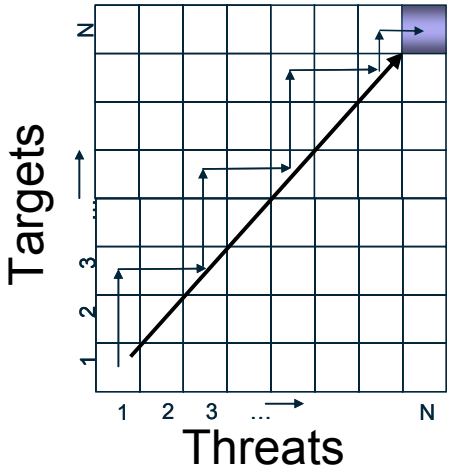
Empirical Validation: Procedure

- Participants: 7 undergraduates
- Phases
  - Phase I: Declarative & procedural training + practice (50 hrs)
  - Phase II: 49 trials + feedback (49 hrs)
    - Consistent enemy
    - Scale-up targets & threats
  - Phase III: 18 trials (18 hrs)
    - Inconsistent enemy
    - Scale-up targets & threats
- Measure
  - Quality of the proposed strike package (TO3) for each TST, determined by expert ratings (96% agreement)
- Design (within Ss)
  - Protocol (POMDP vs. control)
  - Phase (II vs. III)
  - Test (pre vs. post)
  - (Counterbalanced scenarios & order)

Control treatment

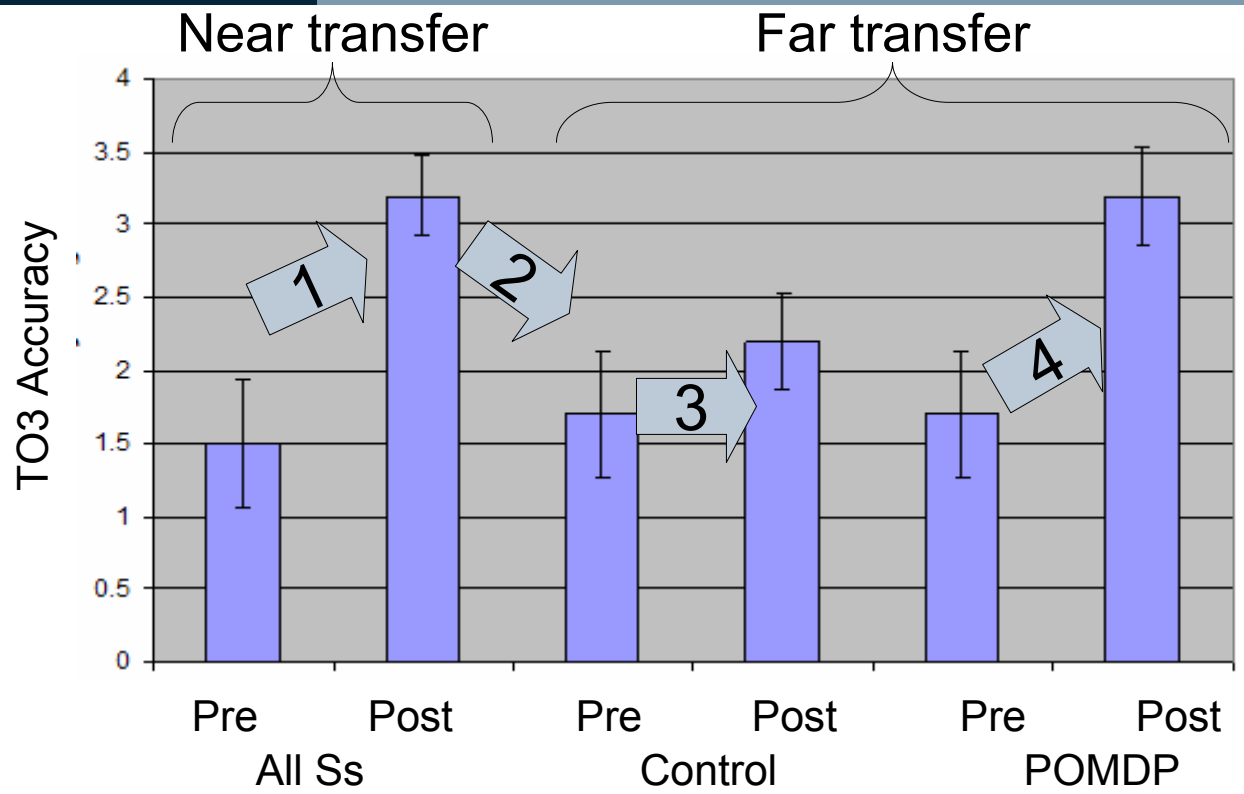


Potential POMDP treatment



© 2007, Aptima, Inc.

Results



1. Teams can learn the task ($p<.01$)
2. Far transfer degrades performance ($p<.01$)
3. Controls learn slowly if at all ($p>.05$)
4. POMDP condition learns rapidly ($p<.01$)

- Phase II training produces mid/high competency operational team

© 2007, Aptima, Inc.

Conclusion

- Aptima is leading research in optimizing instructional strategy
  - Optimization *in* instruction: near-optimal solutions as feedback
  - Optimization *of* instruction: POMDP model driven scenario selection
- Our developed automated intelligent tutoring system significantly utperformed expert-based training solutions

© 2007, Aptima, Inc.

APTIMA®

HUMAN-CENTERED

ENGINEERING