Sequential Decision Making Under Uncertainty

CDT Edward Klein
United States Military Academy
Department of Behavioral Science & Leadership
Brian J. Stankiewicz, Ph.D.
University of Texas, Austin
Department of Psychology & Center for Perceptual Systems

Agenda

- Format of sequential decision making with uncertainty tasks
  - Partially Observable Markov Decision Process
- Seek and Destroy Problem
- Previous Studies
  - Static Model
- Current Study
  - Dynamic Model

Examples of Sequential Decision Making With Uncertainty

- Medical Diagnosis
- Spatial Navigation in an uncertain environment
- Seek & Destroy

Partially Observable Markov Decision Processes

- **State (S)**
  - True state is **hidden** , and cannot be directly observed
- **Action (A)**
  - The decision maker possesses a set of available actions
- **Observation (O)**
  - Actions may result in O, and help determine the true state of the system
- **Transition (T)**
  - Execution of an action may move the system from one state to another

Partially Observable Markov Decision Processes

- **Reward (R)**
  - All actions have costs and rewards
    - These rewards can be considered as time resources, monetary, personnel, etc.
- **Beliefs (B)**
  - The decision maker's current belief about the true state of the system can be represented in terms of likelihood
  - Actions modify this belief state

POMDP: Computing Expected Value

- Using a POMDP we can generate a policy graph for a **Sequential Decision Making Under Uncertainty Task**
  - Policy Graph provides us with the optimal action given a belief.
- Using a POMDP we can compute the optimal performance or theoretical best.
  - This Golden Standard provides a standard for comparison
  - By comparing human behavior to the optimal **Expected Reward** we can get a measure of human **efficiency**
  - This comparison allows the identification of the observer’s cognitive limitations

Seek & Destroy

- Has the same basic structure as the Tiger Problem

Table: Structure of Seek & Destroy Task

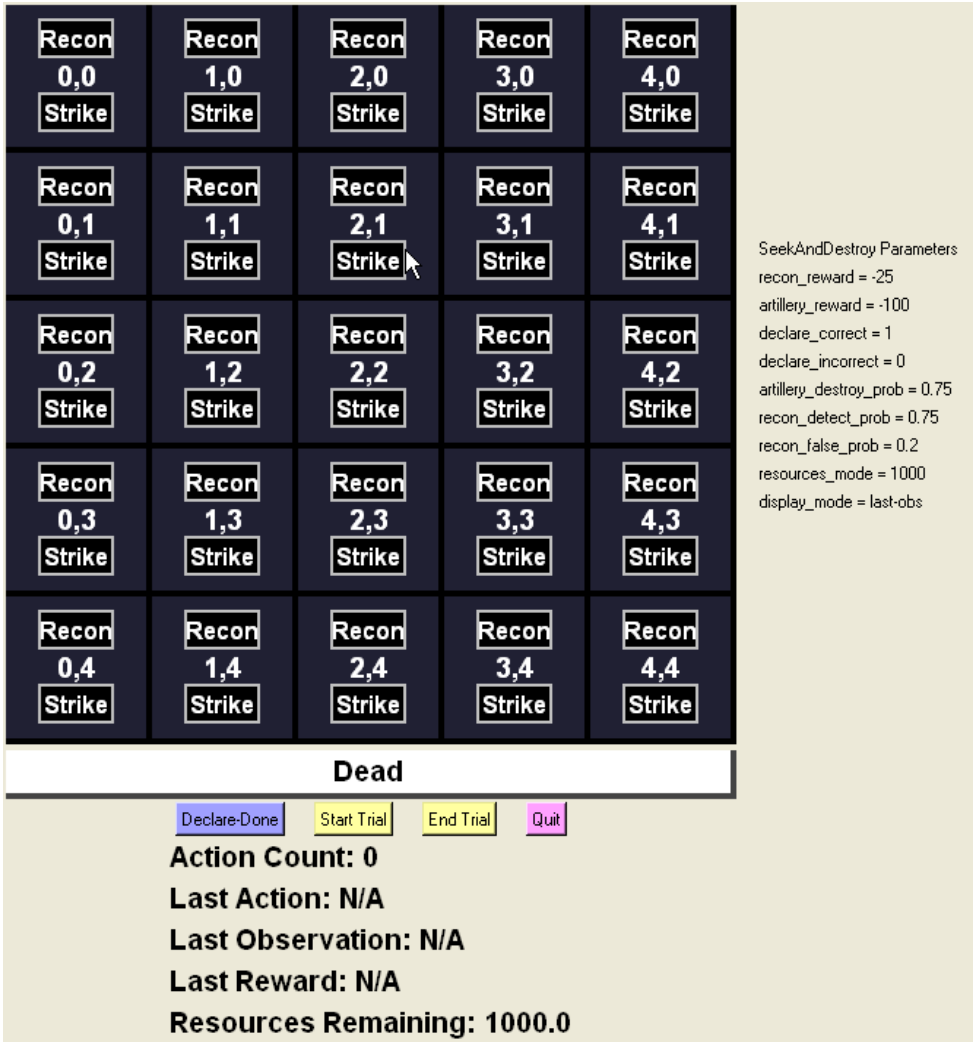
States	Possible locations of enemy
Hidden state	Current position of enemy
Actions	Recon, Arty, Declare “finished”
Observation	Enemy Sighted, Enemy not Sighted
Transitions	Enemy at specific grid -> Enemy Destroyed
Rewards	Cost of UAV recon, Cost of Arty strike, time, success v. failure
Beliefs	Hypothesis about the enemy’s current position

Seek & Destroy (Static)

- Task
  - Localize and Destroy the enemy
- Conditions
  - Utilize “noisy” recon to find the enemy
  - Attack with imperfect weapons to destroy the enemy
- Standards
  - Use no more than available resources to complete the mission

Seek & Destroy (Static)

- Seek & Destroy problem space
  - 5x5 arena
  - Send Recon or Artillery to 25 locations
  - When enemy is destroyed or resources are depleted, declare done.



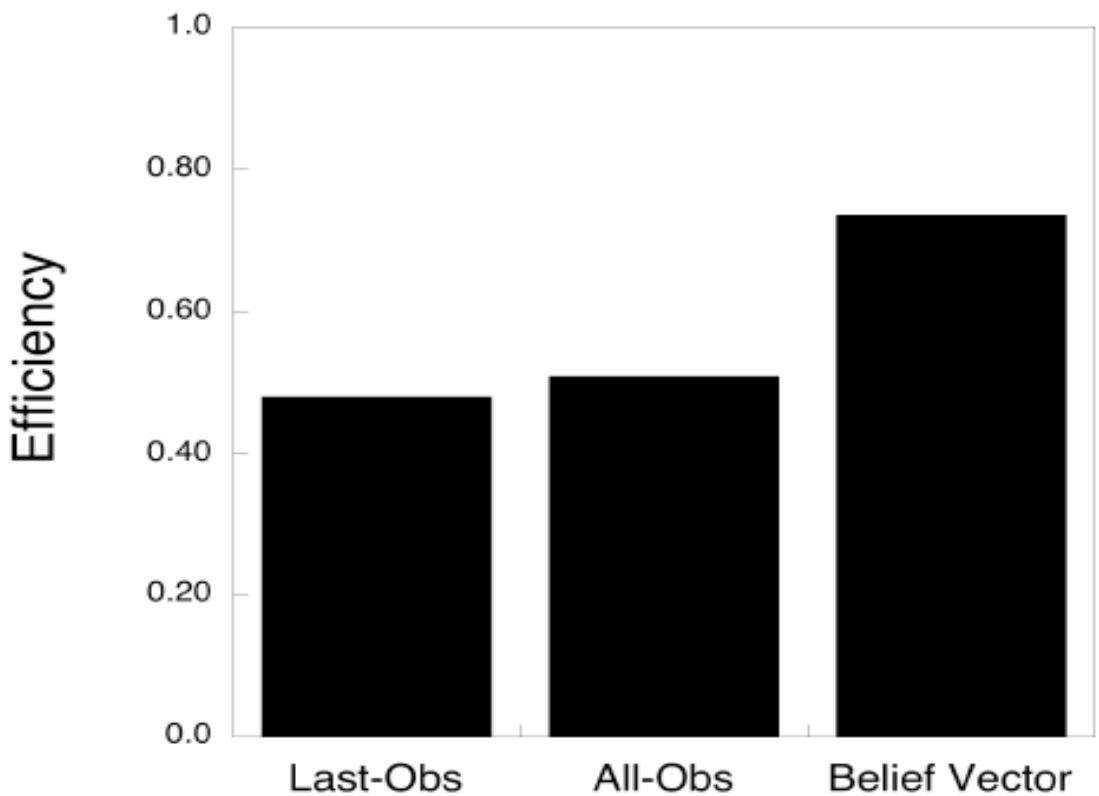
Seek & Destroy (Static)

- Design
  - Independent Variables:
    - Resources available
    - Viewing Conditions:
      1. Only Last Observation (no aids)
      2. All Observations (memory aid)
      3. Belief State (cognitive crutch for computing probability of true state)
  - Dependent Variables:
    - Average reward over time
    - Likelihood of success

Seek & Destroy (Static)

Results: Preliminary findings show increased performance in the belief vector condition

Graph: Performance on Seek & Destroy Task measured in efficiency



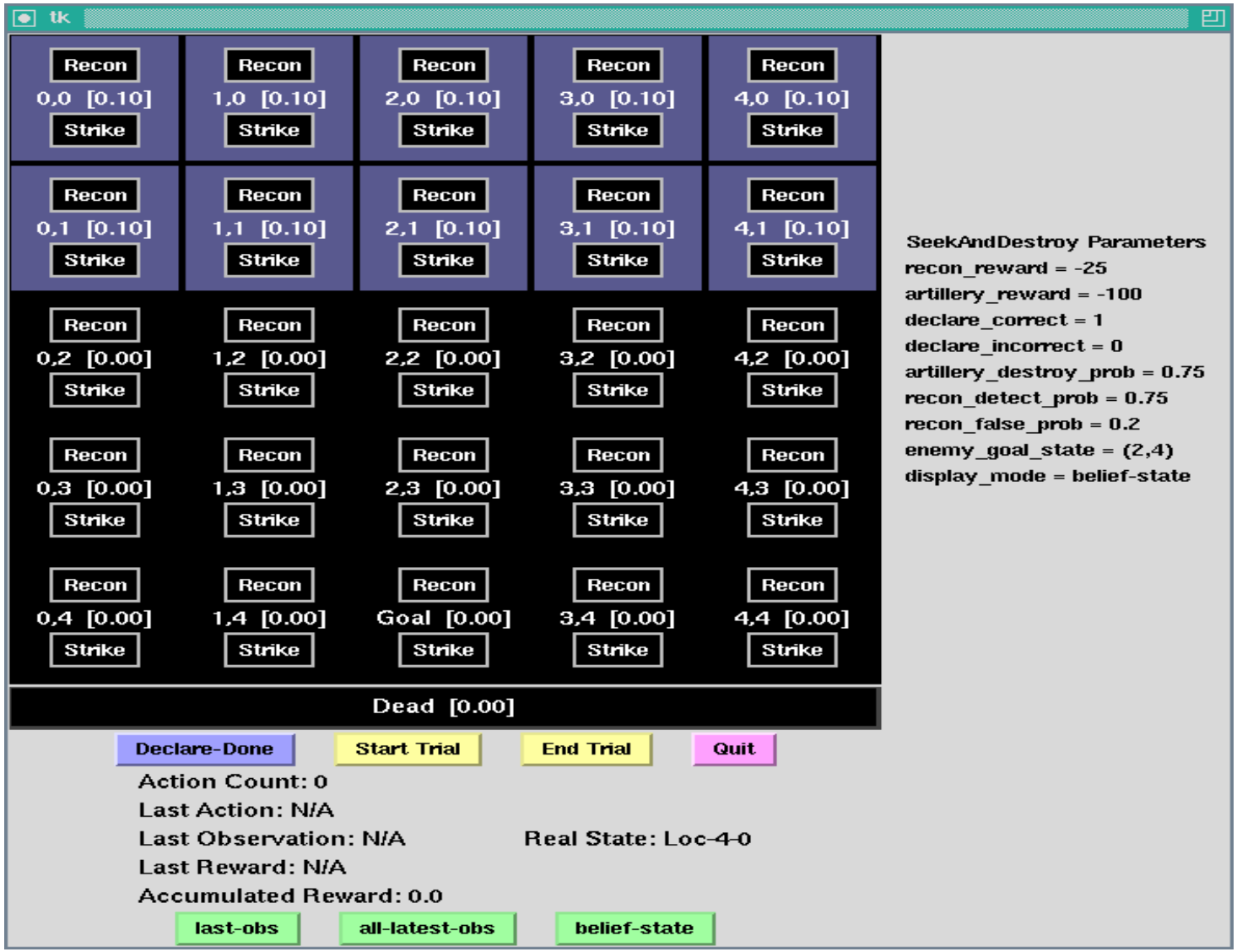
Seek & Destroy (Dynamic)

- Current Study
  - Utilizes a similar problem structure to the previous Seek & Destroy task
    - However, in this case the enemy is dynamic
- Design
  - The enemy will begin above a specific point in the arena (anywhere in the two top rows)
  - The enemy is attempting to move to a specific point in the arena (point 2,4)

Seek & Destroy (Dynamic)

- Design
  - Independent Variables:
    - Viewing Conditions: This scenario utilizes the three viewing conditions of the previous experiment
      1. Only Last Observation
      2. All Observations
      3. Belief State: In this condition, the probability of movement as well as accurate recon are factored into display
  - Dependent Variables:
    - Average reward over time
    - Likelihood of success

Seek & Destroy (Dynamic)



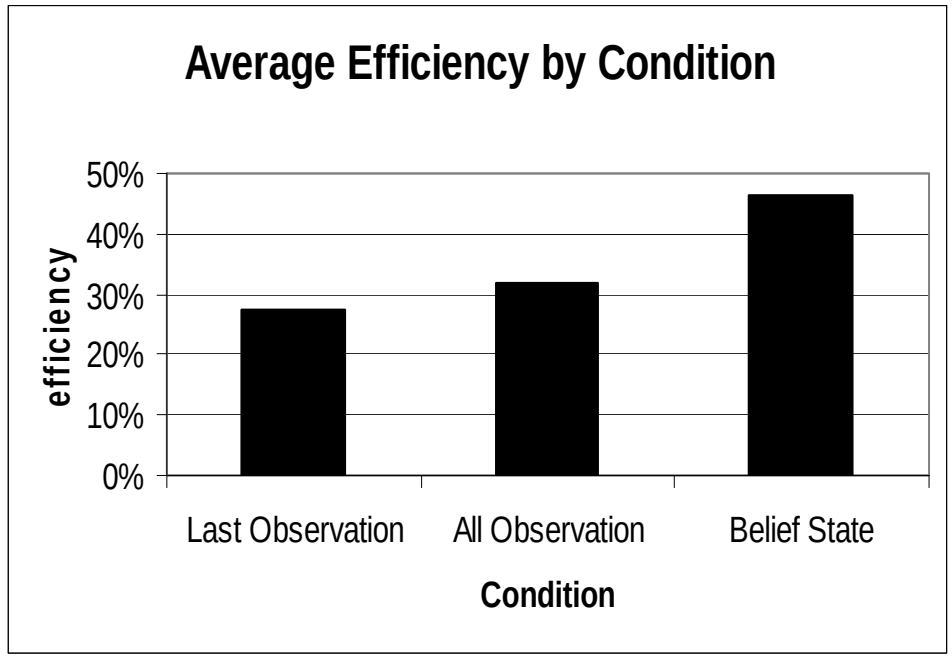
Seek & Destroy (Dynamic)

- Subjects
  - 6 undergraduate and graduate students enrolled at the University Texas at Austin
- Goals of this research
  - Generate a more realistic scenario for analysis
  - Determine the effectiveness of the belief state in more realistic situations
  - Analyze human problem solving in different scenarios

Seek & Destroy (Dynamic)

Results

- The average efficiency across all participants shows a similar trend to what we saw in the static case



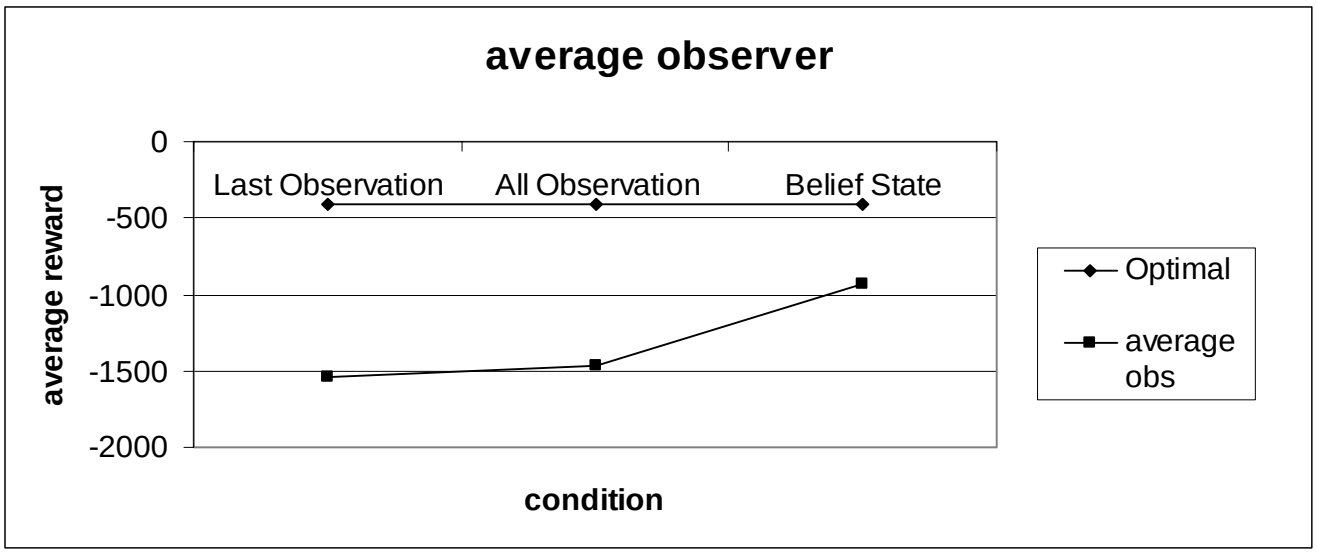
Graph: Performance on Seek & Destroy Task measured in efficiency

Seek & Destroy (Dynamic)

Results

- The average reward over time also demonstrates the strength of the belief vector condition

Graph: Performance on Seek & Destroy Task measured in average reward



Seek & Destroy (Dynamic)

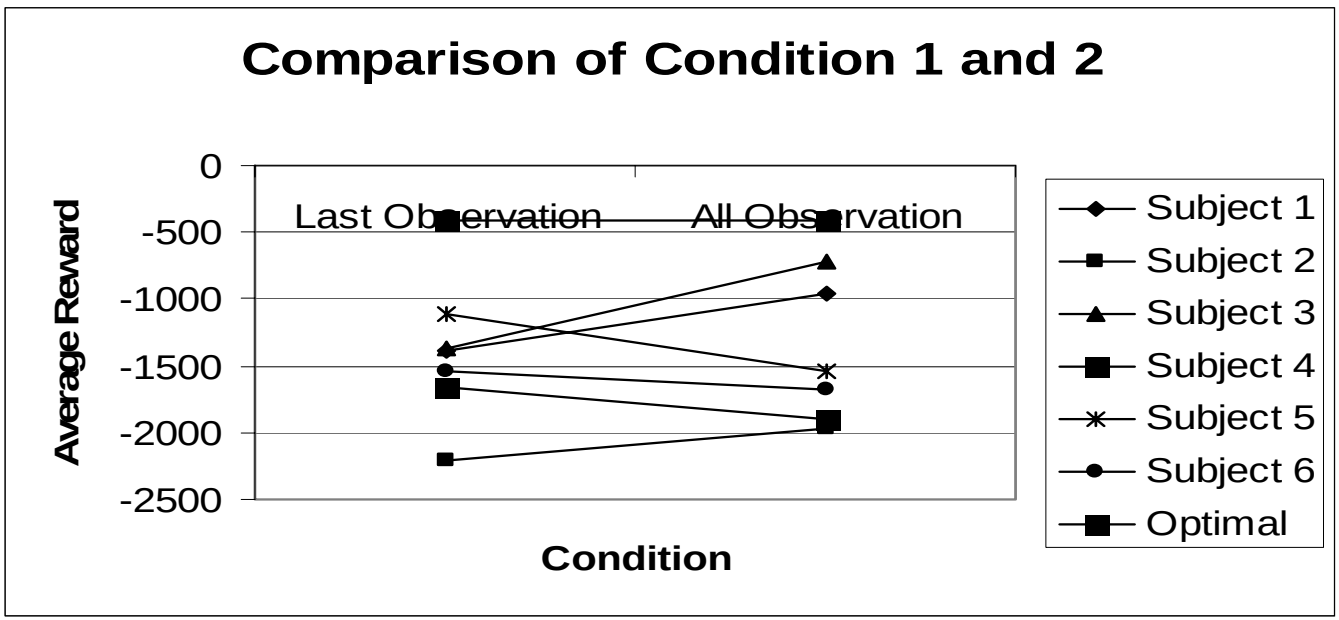
Results

- The average reward over time also demonstrates the strength of the belief vector condition

Results by condition

- While 50% of subjects saw enhanced performance with the memory aid, the other 50% experienced a decline in performance.

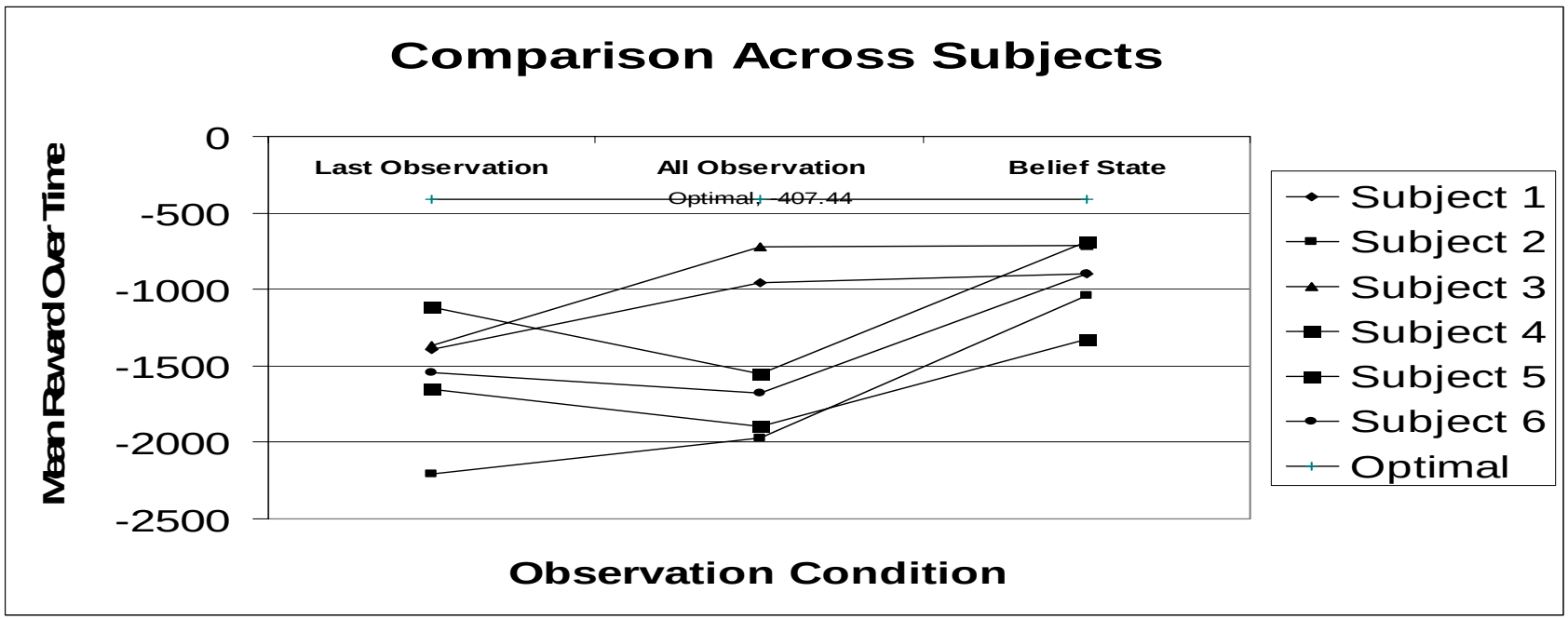
Graph: Performance on Seek & Destroy Task measured in average reward



Seek & Destroy (Dynamic)

Results by condition

- However, with the addition of the belief vector, all subjects saw an improvement over the initial condition.



Graph: Performance on Seek & Destroy Task measured in average reward

Seek & Destroy (Dynamic)

- Future concerns
  - What is causing the two very patterns of performance across the conditions
    - Research should aim at determining if the change in performance was a factor of individual strategy or differences in individual cognitive strengths

Future Research Questions

- Is the addition of a “Declare-incomplete” option beneficial to the researchers knowledge?
- Several manipulations of the scenario have been considered.
  - The addition of another recon device with less area and less noise
  - The addition of a mobile recon patrol
  - The addition of multiple enemies

Summary

- Use of POMDP model in Sequential Decision Making Under Uncertainty
- Tiger Problem
- Seek & Destroy
  - Static
  - Dynamic

Decision Making Under Uncertainty

Thank you for your attention
Are there any questions at this time

CDT Edward Klein
United States Military Academy
Department of Behavioral Science & Leadership
Brian J. Stankiewicz, Ph.D.
University of Texas, Austin
Department of Psychology & Center for Perceptual Systems

Resources

Cassandra, Anthony. “Acting in Partially Observable Stochastic Domains.”

Stankiewicz, Brian. “Localizing the cognitive bottleneck in human sequential decision making under uncertainty by using the ideal observer analysis.”

Stankiewicz, Brian. “Lost in Virtual Space: Studies in Human and Ideal Spatial Navigation.”