

The logo for Carnegie Mellon University, featuring the text "Carnegie Mellon University" in a white serif font, set against a dark blue background with a grid of colorful lines (red, green, yellow, blue) forming a diamond pattern.

**Carnegie
Mellon
University**

Learning to Defend by Attacking (and Vice- Versa): Transfer of Learning in Cybersecurity Games

July 7th, 2023

Tyler Malloy, Yinuo Du, and Cleotilde (Coty)
Gonzalez



CMU: Dynamic Decision-Making Lab

Coty
Gonzalez

Ngoc
Ngueyn

Jeffrey
Flag

Me

Erin
Bugbee



z

Chase
MacDonald

Baptiste
Prebot

Yinuo
Du

Maria
Ferreria





Background: Attack/Defense Decision Making

Attack/Defense Scenarios

Attacker



Network

CyBORG Game

Round 7/10 Last round: **6.1** Total loss: **0.1**

Subnet	Subnet Name	IP Address	Hostname	Activity	Compromised level
10.0.0.128/28	Sub2 - Enterprise	10.0.0.130	Defender	None	No
10.0.0.128/28	Sub2 - Enterprise	10.0.0.131	Enterprise1	Exploit	User
10.0.0.128/28	Sub2 - Enterprise	10.0.0.135	Enterprise2	None	No
10.0.178.16/28	Sub3 - Operational	10.0.178.23	Op_Just0	None	No
10.0.178.16/28	Sub3 - Operational	10.0.178.19	Op_Server0	None	No
10.0.29.176/28	Sub1 - User	10.0.29.183	User0	None	No
10.0.29.176/28	Sub1 - User	10.0.29.187	User1	None	User

> Select an action: > You chose to:

Monitor Analyze Remove Restore Remove Enterprise2 Next

Defender



End User



Attack/Defense Modeling

Human Attacker



Network

CyBORG Game

Round 7/10 Last round: **0-1** Total loss: **0-1**

Subnet	Subnet Name	IP Address	Hostname	Activity	Compromised level
10.0.60.120/28	Sub2 - Enterprise	10.0.60.130	Defender	None	No
10.0.60.120/28	Sub2 - Enterprise	10.0.60.131	Enterprise1	Exploit	User
10.0.60.120/28	Sub2 - Enterprise	10.0.60.135	Enterprise2	None	No
10.0.178.16/28	Sub3 - Operational	10.0.178.23	Op_Just0	None	No
10.0.178.16/28	Sub3 - Operational	10.0.178.19	Op_Server0	None	No
10.0.29.176/28	Sub1 - User	10.0.29.183	User0	None	No
10.0.29.176/28	Sub1 - User	10.0.29.187	User1	None	User

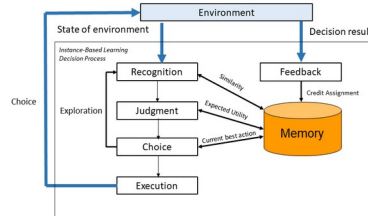
> Select an action: > You chose to:

Monitor Analyze Remove Restore Remove Enterprise2 Next

Human Defender



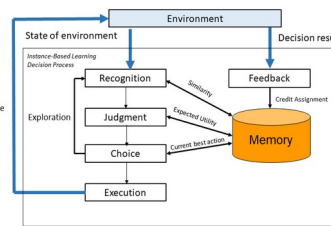
Attacker Model



End User



End User Model



5 Gonzalez, C., Aggarwal, P., Cranford, E. A., & Lebiere, C Design of Dynamic and Personalized Deception: A Research Framework and New Insights for Cyberdefense. In *Proceedings of the 53rd hawaii international conference on system sciences* (Vol. 1834).

Attack/Defense Modeling

Human Attacker



Network

CyBORG Game

Round 7/10 Last round: **0-1** Total loss: **0-3**

Subnet	Subnet Name	IP Address	Hostname	Activity	Compromised level
10.0.60.128/28	Sub2 - Enterprise	10.0.60.130	Defender	None	No
10.0.60.128/28	Sub2 - Enterprise	10.0.60.131	Enterprise1	Exploit	User
10.0.60.128/28	Sub2 - Enterprise	10.0.60.135	Enterprise2	None	No
10.0.178.16/28	Sub3 - Operational	10.0.178.23	Op_Just09	None	No
10.0.178.16/28	Sub3 - Operational	10.0.178.19	Op_Server0	None	No
10.0.29.176/28	Sub1 - User	10.0.29.183	User0	None	No
10.0.29.176/28	Sub1 - User	10.0.29.187	User1	None	User

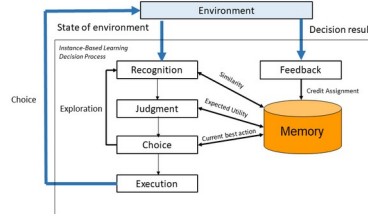
> Select an action: > You chose to:

Monitor Analyze Remove Restore Remove Enterprise2 Next

Human Defender



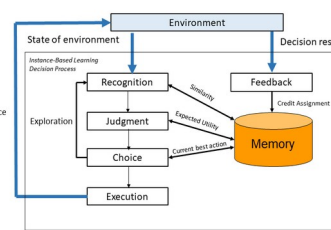
Attacker Model



End User



End User Model



IBL agent predictions

IBL agent predictions

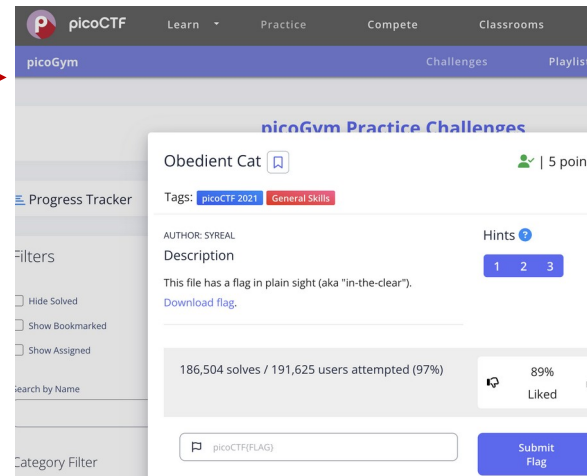
Adaptive and Personalized Defense Strategies

Attack/Defense Real World

Human Participant



Capture The Flag



CTF Designer



How do Humans Learn Cybersecurity?

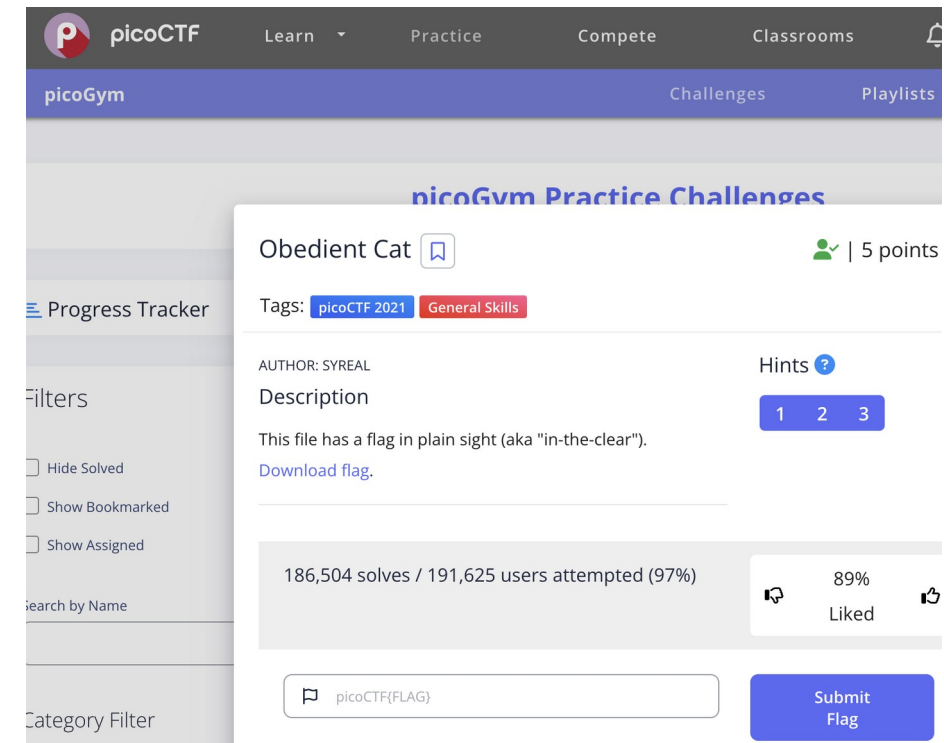
Transfer of Learning.

- Application of experience in one task onto another related task.

Theory of Mind.

- Predicting the beliefs, goals, and behavior of other agents.

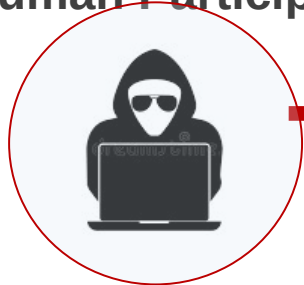
Capture The Flag Challenge



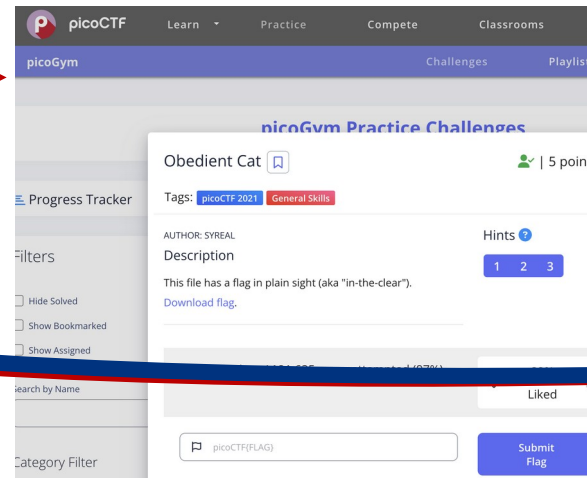
The screenshot displays the picoCTF website interface. At the top, there are navigation tabs for 'Learn', 'Practice', 'Compete', and 'Classrooms'. Below this, the 'picoGym' section is visible, with sub-tabs for 'Challenges' and 'Playlists'. The main content area shows a challenge titled 'Obedient Cat' with a bookmark icon and '5 points'. The challenge is tagged with 'picoCTF 2021' and 'General Skills'. The author is listed as 'SYREAL'. The description reads: 'This file has a flag in plain sight (aka "in-the-clear").' and includes a 'Download flag.' link. Below the description, it shows '186,504 solves / 191,625 users attempted (97%)' and '89% Liked'. At the bottom, there is a text input field containing 'picoCTF{FLAG}' and a blue 'Submit Flag' button.

Attack/Defense Real World

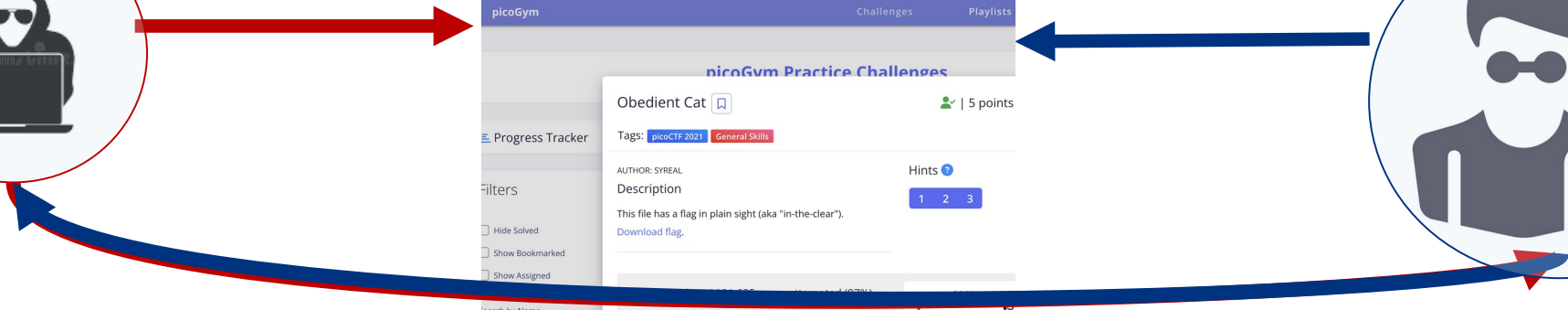
Human Participant



Capture The Flag



CTF Designer

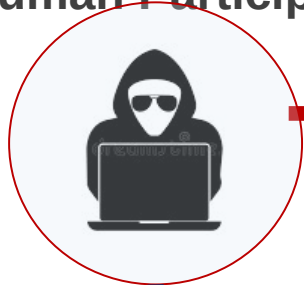


Theory of Mind

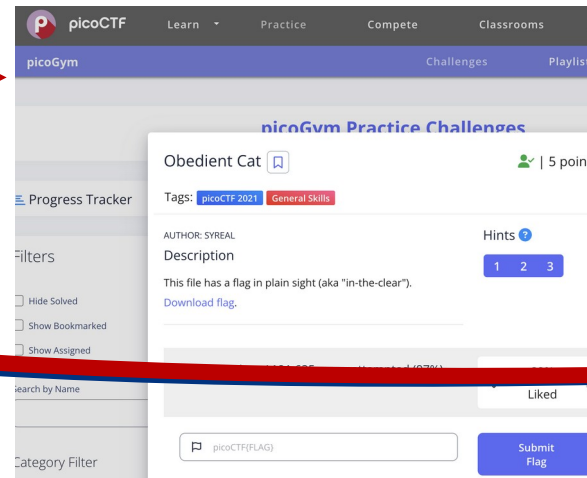
Transfer of Learning

Attack/Defense Real World

Human Participant



Capture The Flag



CTF Designer



Transfer of Learning

Theory of Mind

Cyber Testbeds

Realistic

Abstract

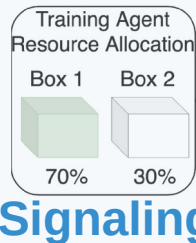
Deception Technique

Phishing Training



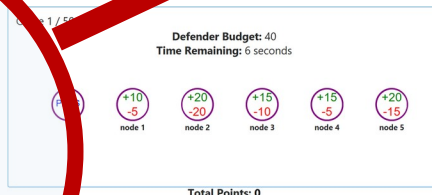
Dynamic Training

Box Game



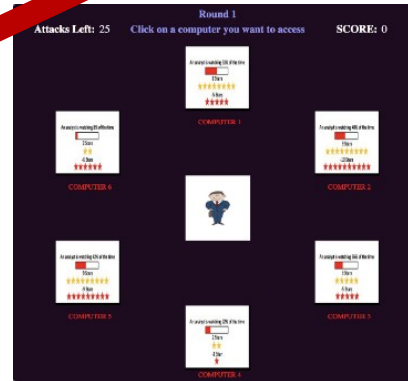
Simplistic

HoneyPot



Decoying

Intruder Attack

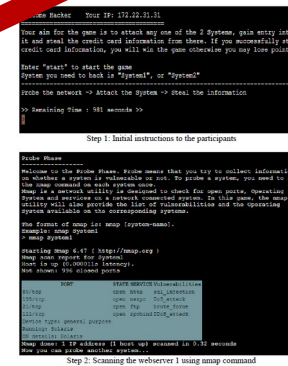


Signaling

Cage: CybOrg

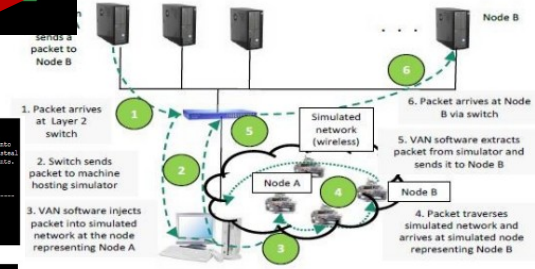
Host ID	Host Name	IP Address	Network	Activity	Operational Level
10.4.0.100/24	hdd	10.4.0.100	hdd	hdd	hdd
10.4.0.100/24	hdd	10.4.0.101	hdd	hdd	hdd
10.4.0.100/24	hdd	10.4.0.102	hdd	hdd	hdd
10.4.0.100/24	hdd	10.4.0.103	hdd	hdd	hdd
10.4.0.100/24	hdd	10.4.0.104	hdd	hdd	hdd
10.4.0.100/24	hdd	10.4.0.105	hdd	hdd	hdd
10.4.0.100/24	hdd	10.4.0.106	hdd	hdd	hdd
10.4.0.100/24	hdd	10.4.0.107	hdd	hdd	hdd
10.4.0.100/24	hdd	10.4.0.108	hdd	hdd	hdd
10.4.0.100/24	hdd	10.4.0.109	hdd	hdd	hdd

Packet



Decoying + Masking

CyberVAN



Masking

Gonzalez, C., Aggarwal, P., Cranford, E. A., & Lebiere, C Design of Dynamic and Personalized Deception: A Research Framework and New Insights for Cyberdefense. In *Proceedings of the 53rd hawaii international conference on system sciences* (Vol. 1834).

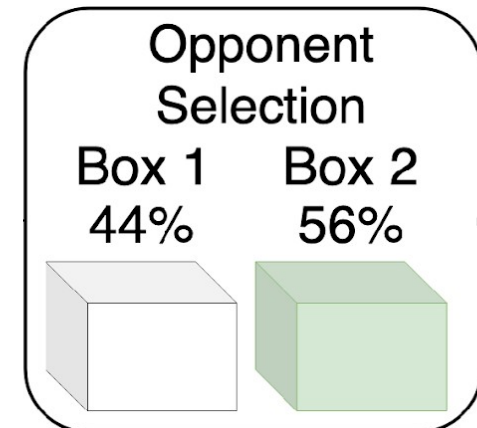
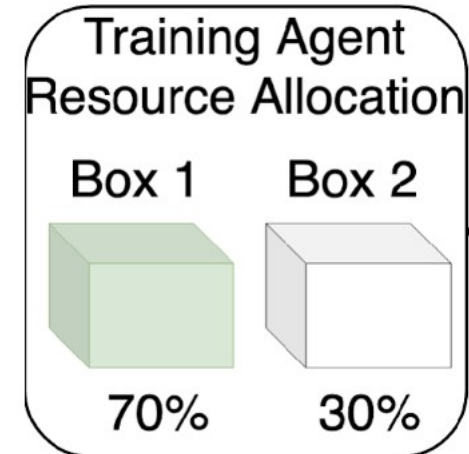
Stackelberg Security Games

Very simple two player game.

- Sequential decision making with 2 or more action options.

Consists of an attacker and defender.

- Defender has limited resources to protect a set of assets.
- Attacker chooses an asset to attack.

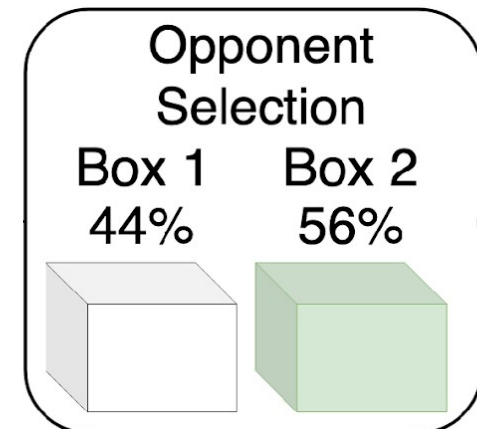
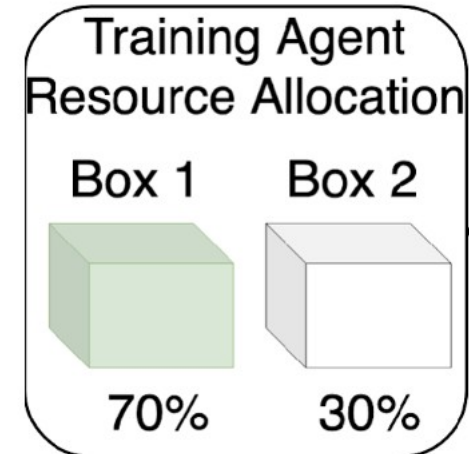


Real-world applications of SSGs

SSGs are used physical security scenarios.

- Opportunistic crime:
 - Fare evasion in public transit.
 - Robbery, vandalism, etc.
- Infrastructure security.
 - Airport passenger screening.
- Illegal hunting and poaching.

SSGs have yet to be applied onto real-world cybersecurity.



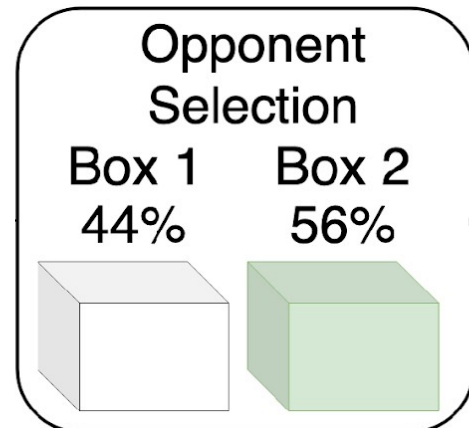
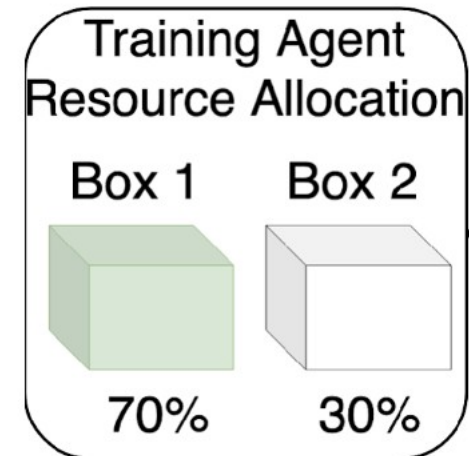
Challenges of applying SSGs to Cybersecurity

SSGs assume long deploy/asses time.

- We consider the **'repeated SSG'**.

Real-world applications of SSGs have **focused on biases** present in human decision making.

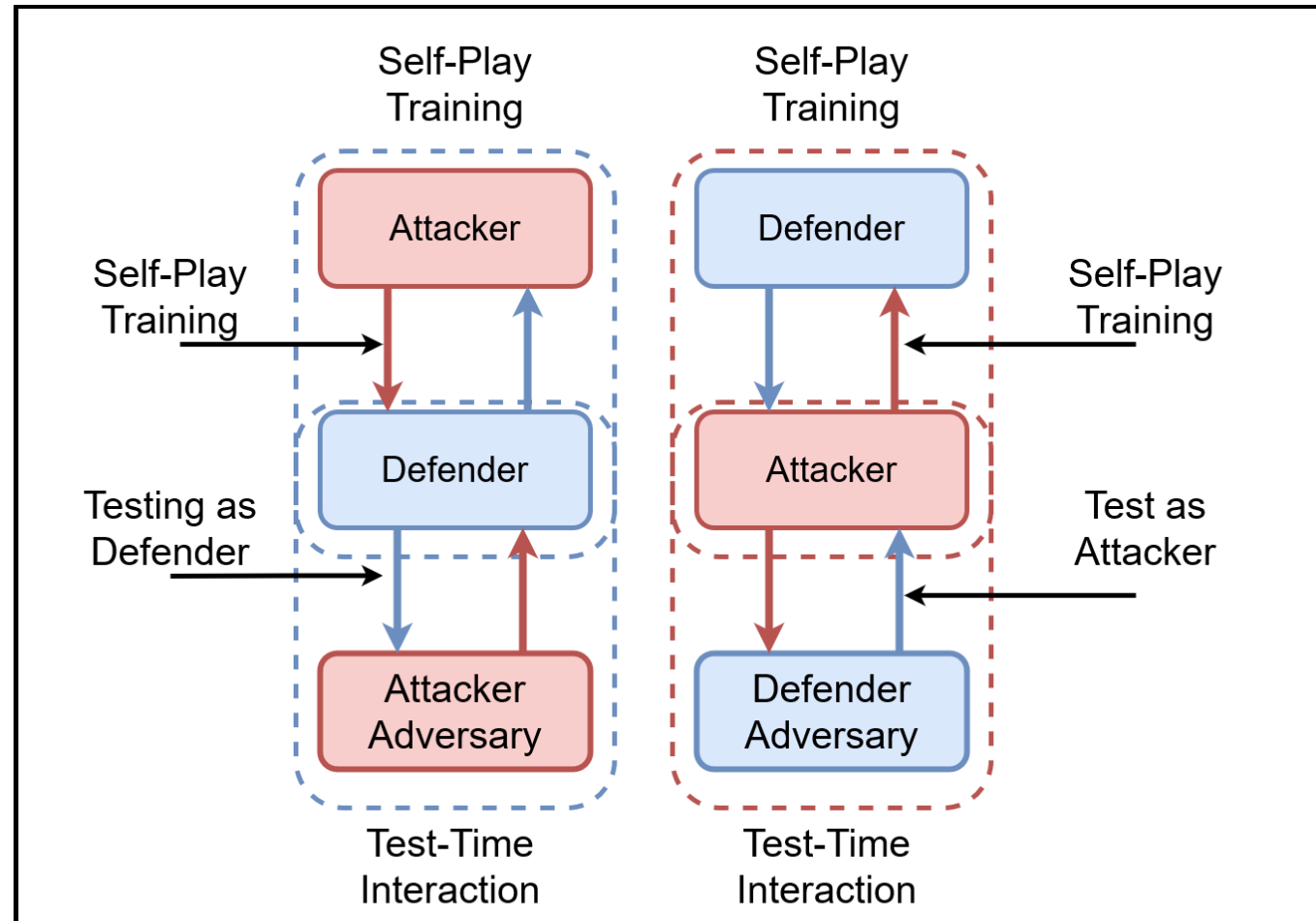
- We additionally consider how humans **overcome these biases.**





Theory of Mind and Transfer of Learning

Traditional Attack/Defense Training



Theory of Mind

Predicting the beliefs, goals, and observations of other agents.

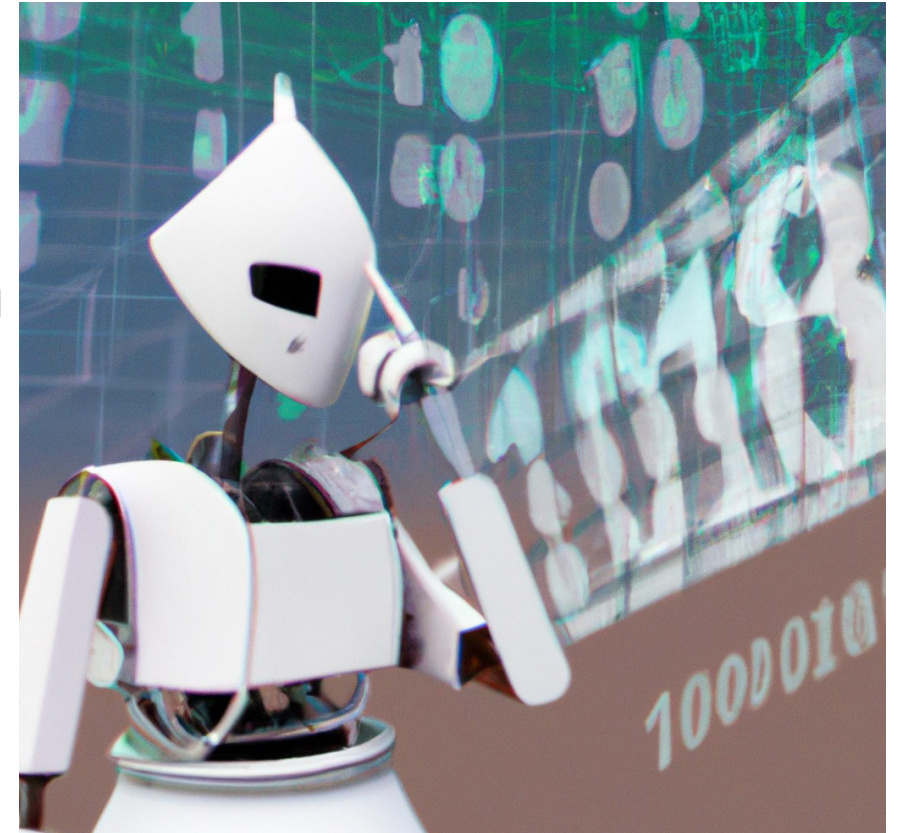
- Can allow agents to account for things they do not directly observe.
- Opponent's behavior can often reveal what they observe and what their goals are.



Transfer of Learning

Using experience in one domain to inform decision making in another.

- Common target for AI research and engineering, related to zero or few-shot learning.
- Many different types, the type discussed here is analogous to *domain transfer*.





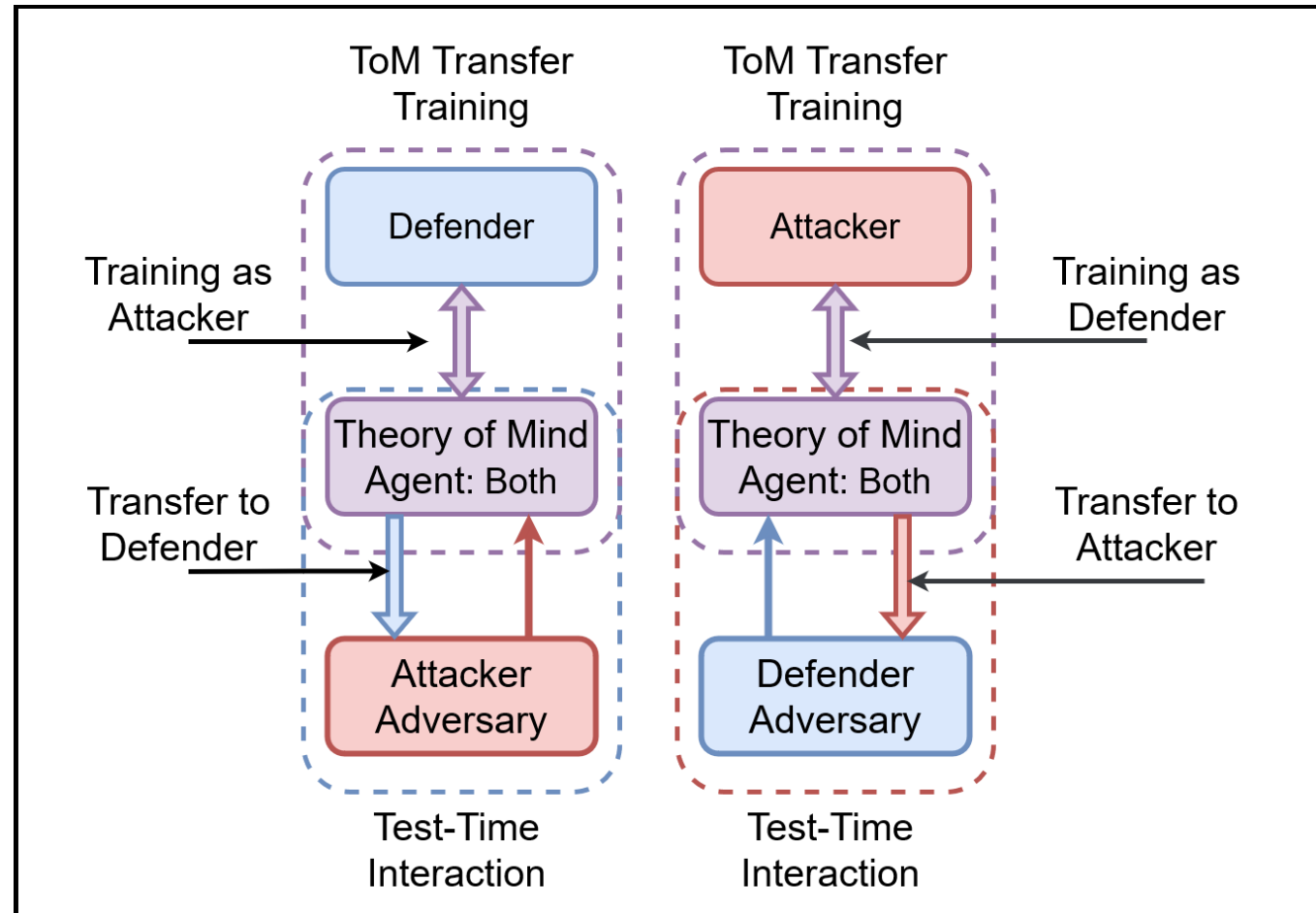
Human Transfer of Learning

In the real world, humans have experience as attackers and defenders (Attack-Defense and Jeopardy CTFs)

Security systems that take into account human-like decision making and learning have been shown to be more effective in real-world situations.

Humans use varied experience, theory of mind, and transfer learning to overcome biases. Carnegie Mellon University

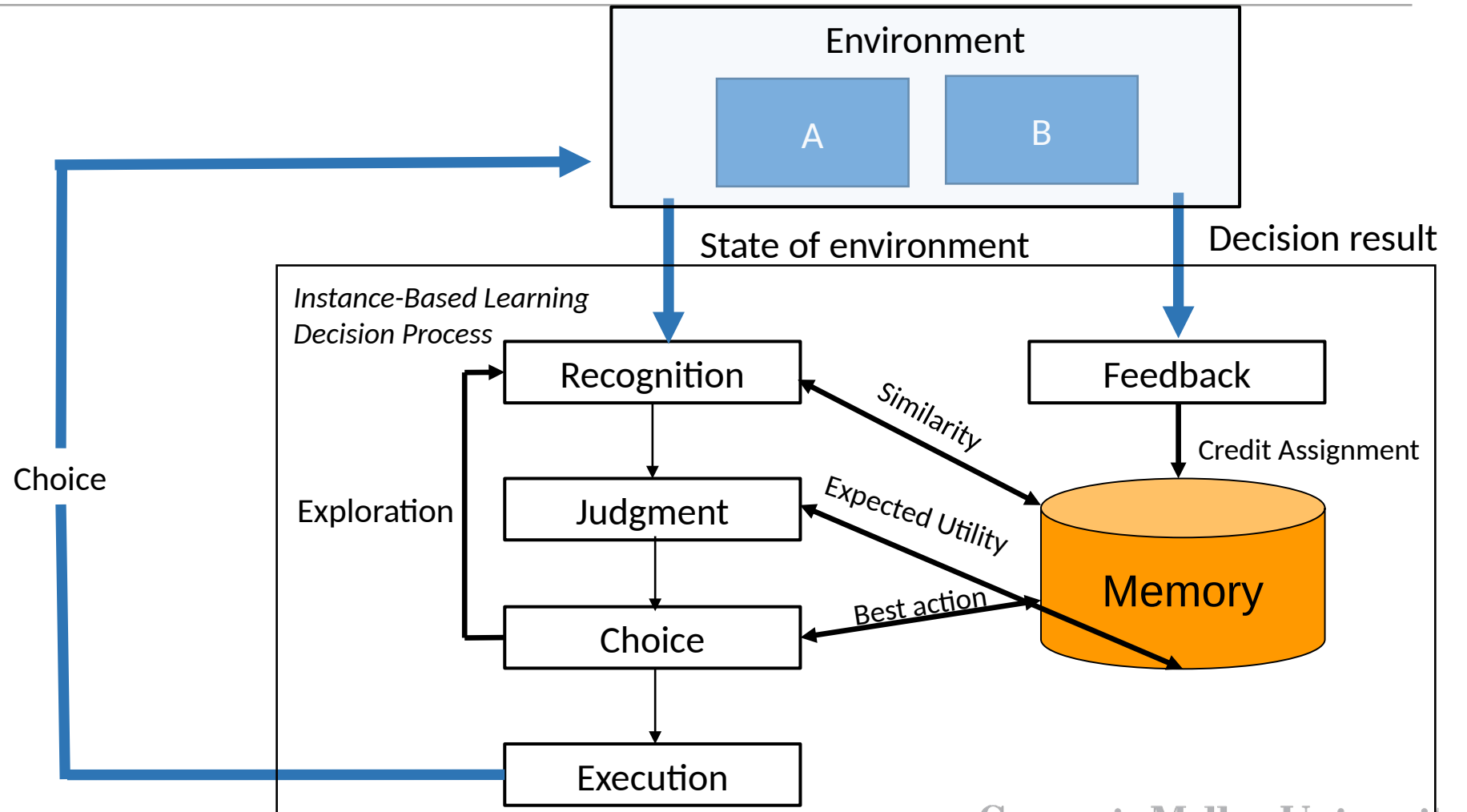
Theory of Mind + Transfer of Learning



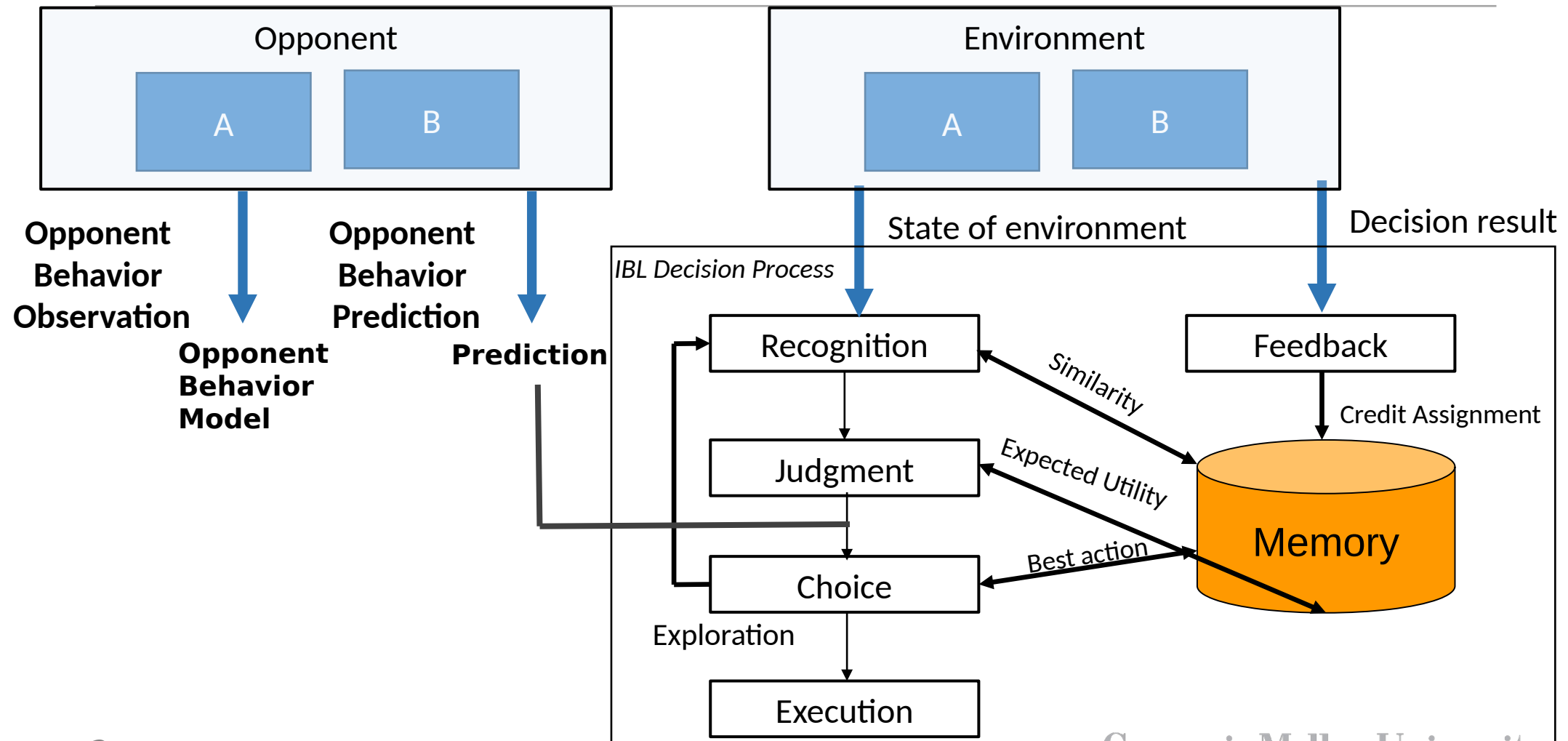


ToM and ToL with Instance Based Learning

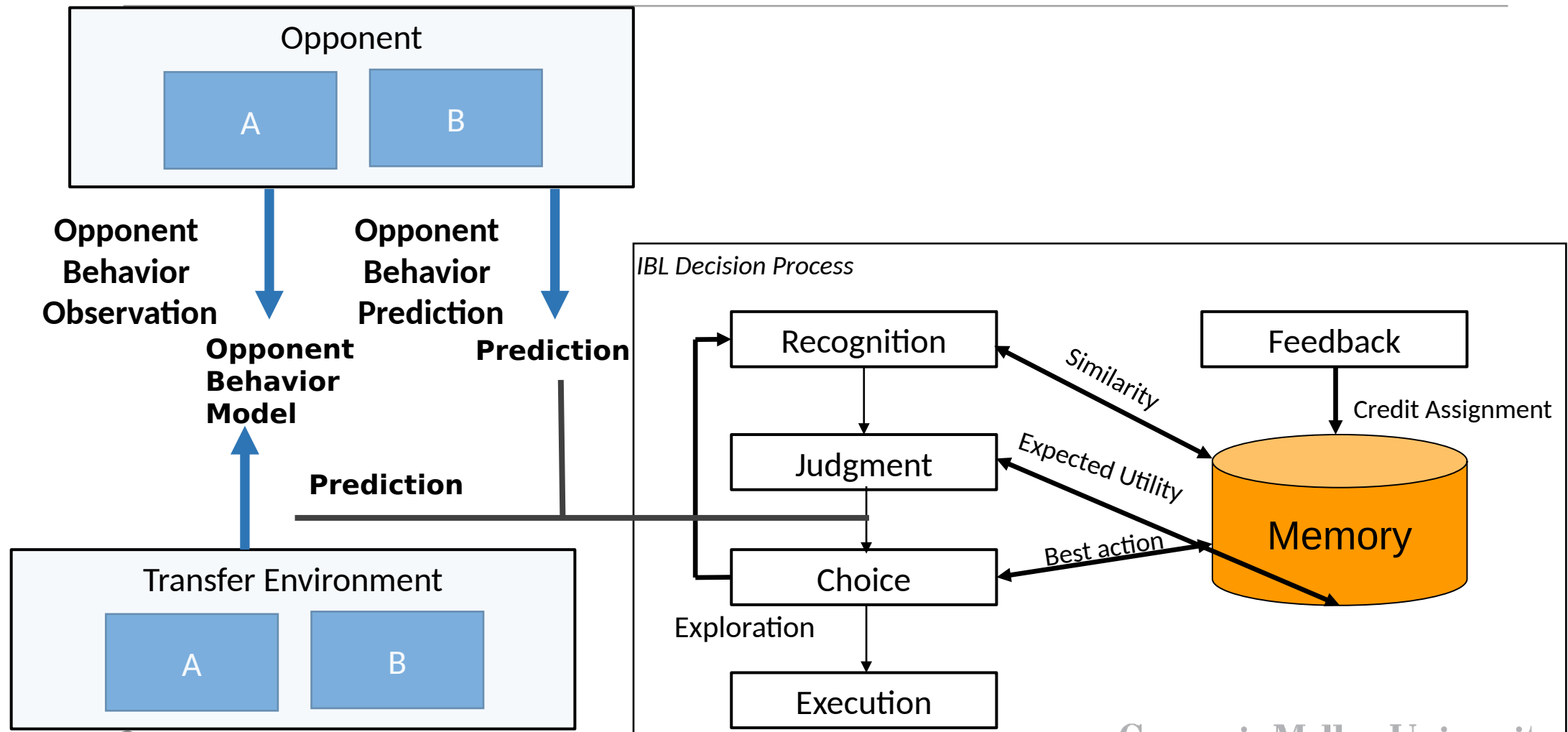
Instance Based Learning Model



Instance Based Learning with ToM



IBL Transfer of Learning with ToM



Comparison Models

The base IBL model predicts BR behavior based on experience.

The IBL + ToM model uses ToM to predict opponent actions and learn a model for Transfer.

Upper Confidence Bound
Model behaves optimally in SSGs.

Basic IBL model


$$\frac{\exp(V_{i,k_i,t}/\tau_v)}{\sum_{k_j=k_1}^{b_n} \exp(O_{i,k_j,t}/\tau_v)}$$

IBL + ToM model

$$\frac{\exp(O_{i,k_o,t}/\tau_o)}{\sum_{k_j=k_1}^{b_n} \exp(O_{i,k_j,t}/\tau_o)}$$

Optimal UCB model

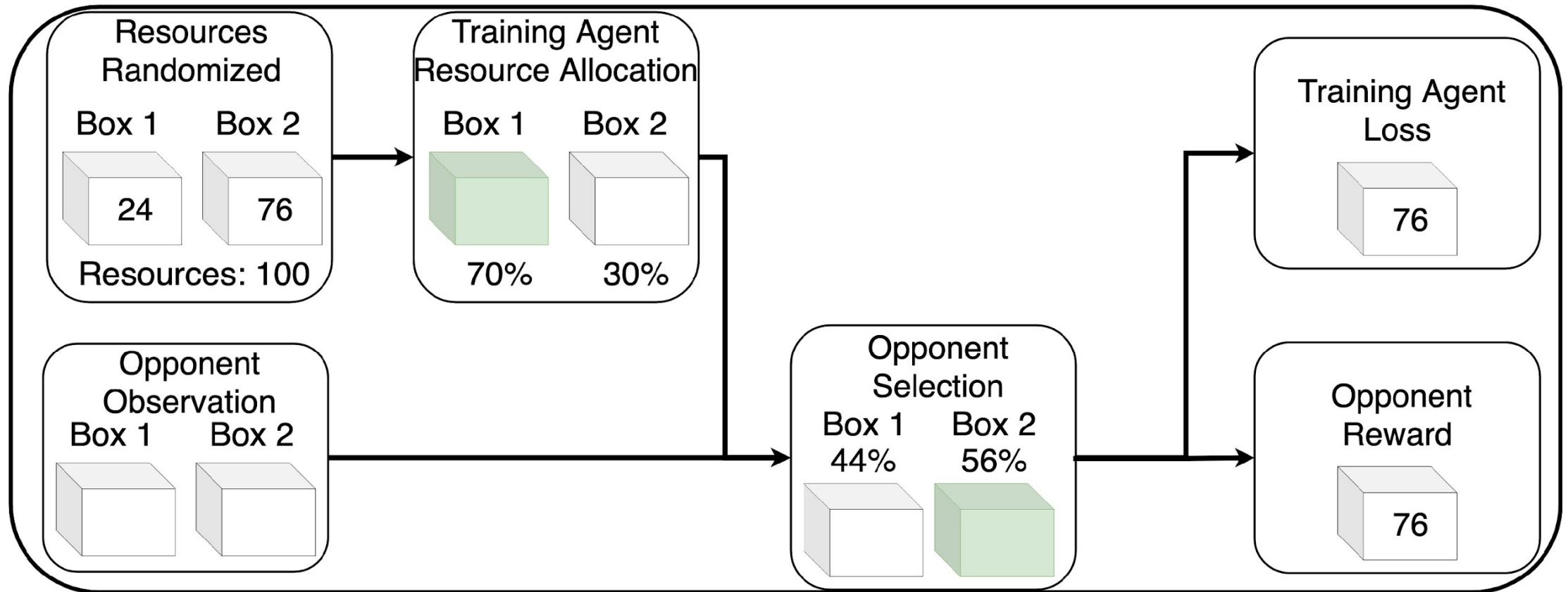
$$\arg \max_a \left[Q_t(a) + c \sqrt{\frac{\ln t}{N_t(a)}} \right]$$



Simulation Experimentation

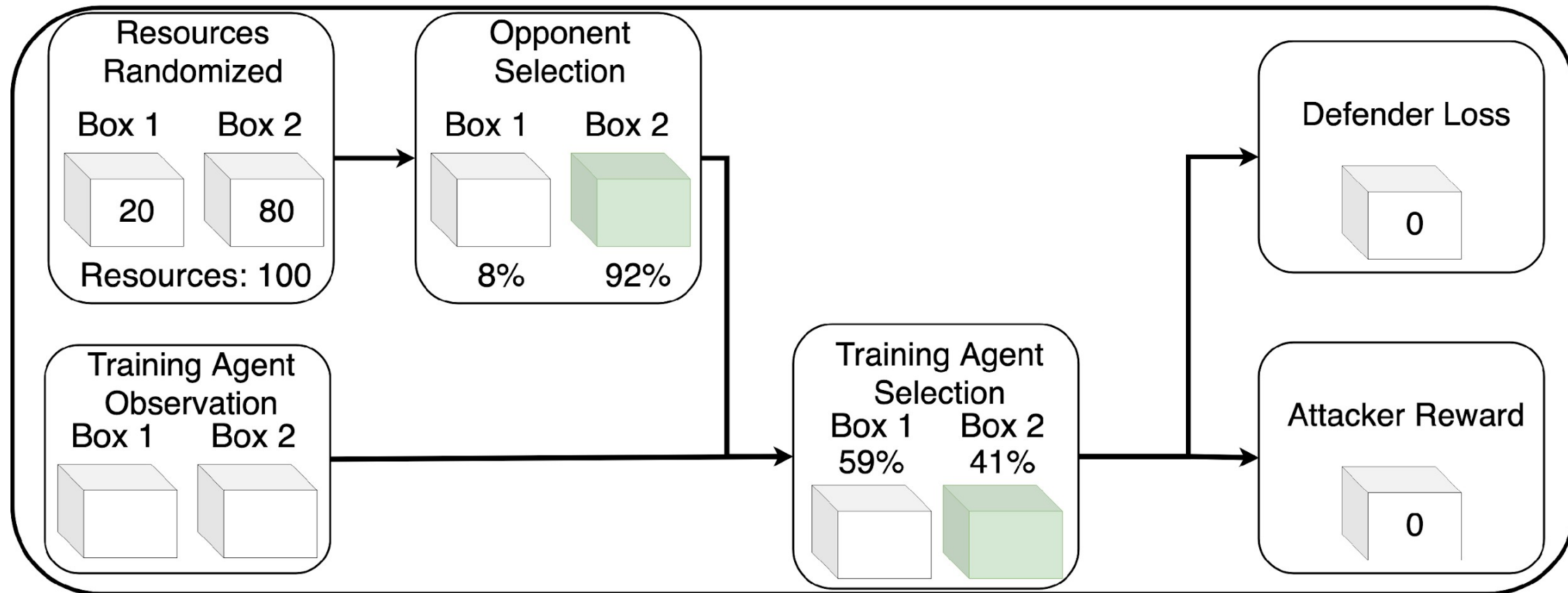
Experimentation Environment

Initial Training Period

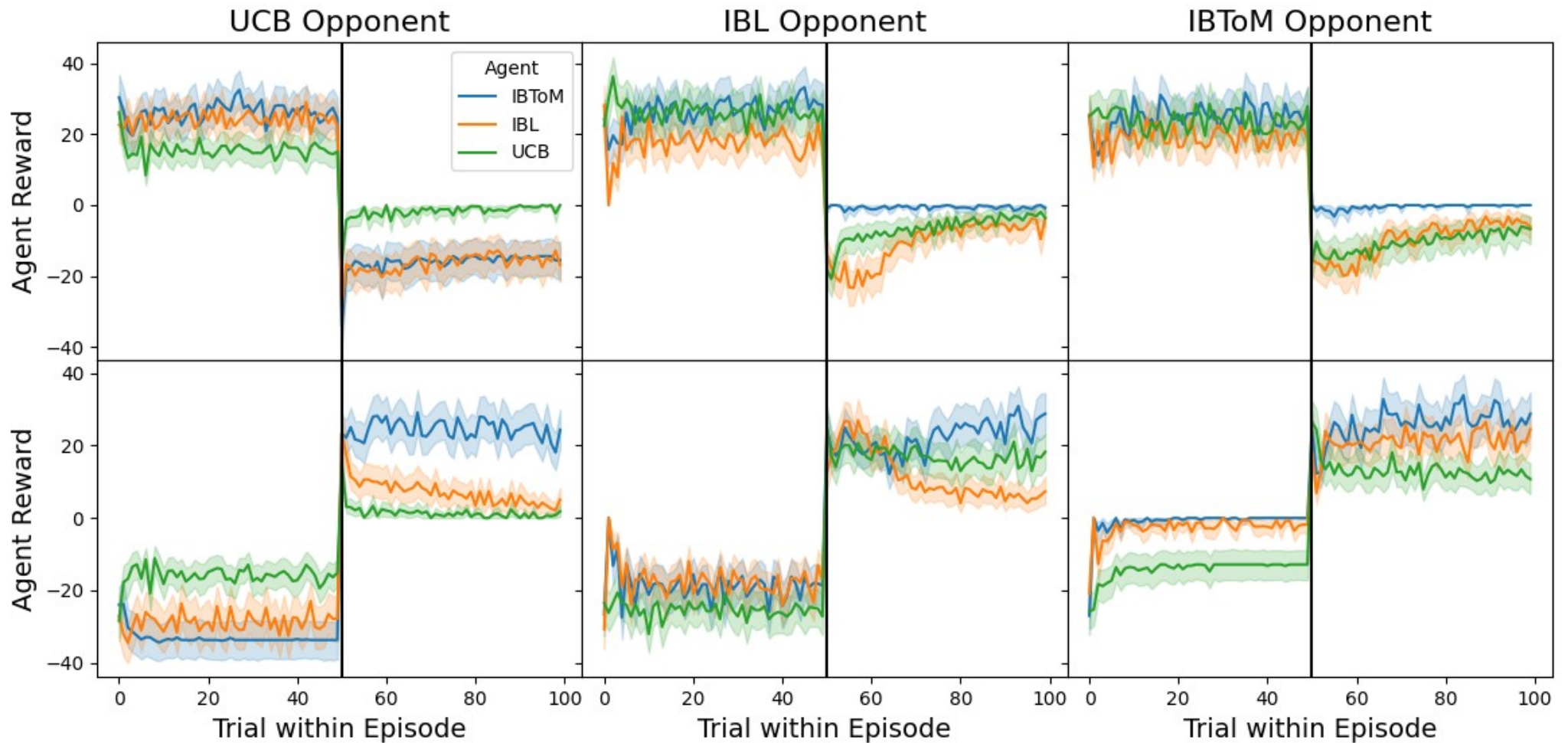


Experimentation Environment

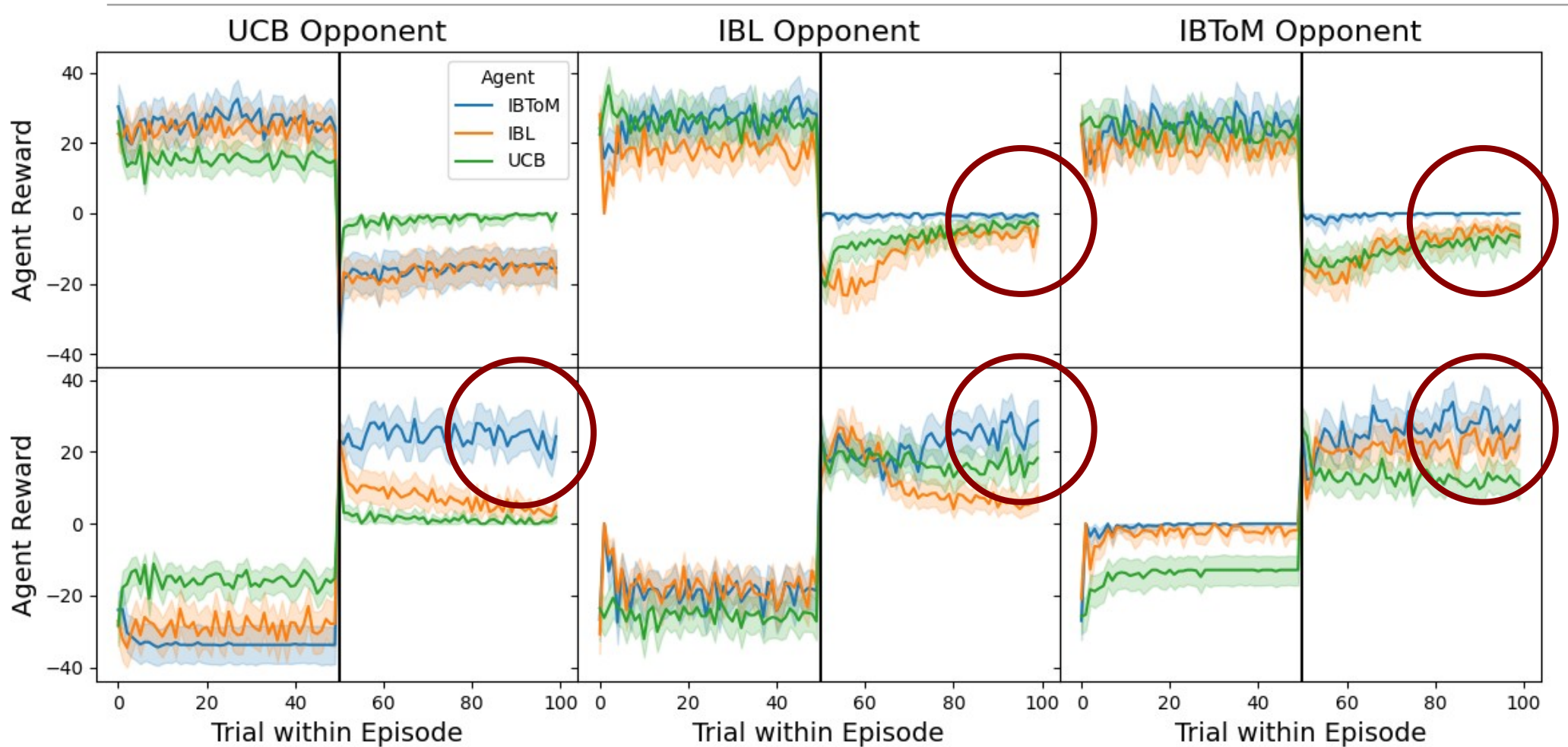
Transfer Learning Period



Improved Transfer through Theory of Mind



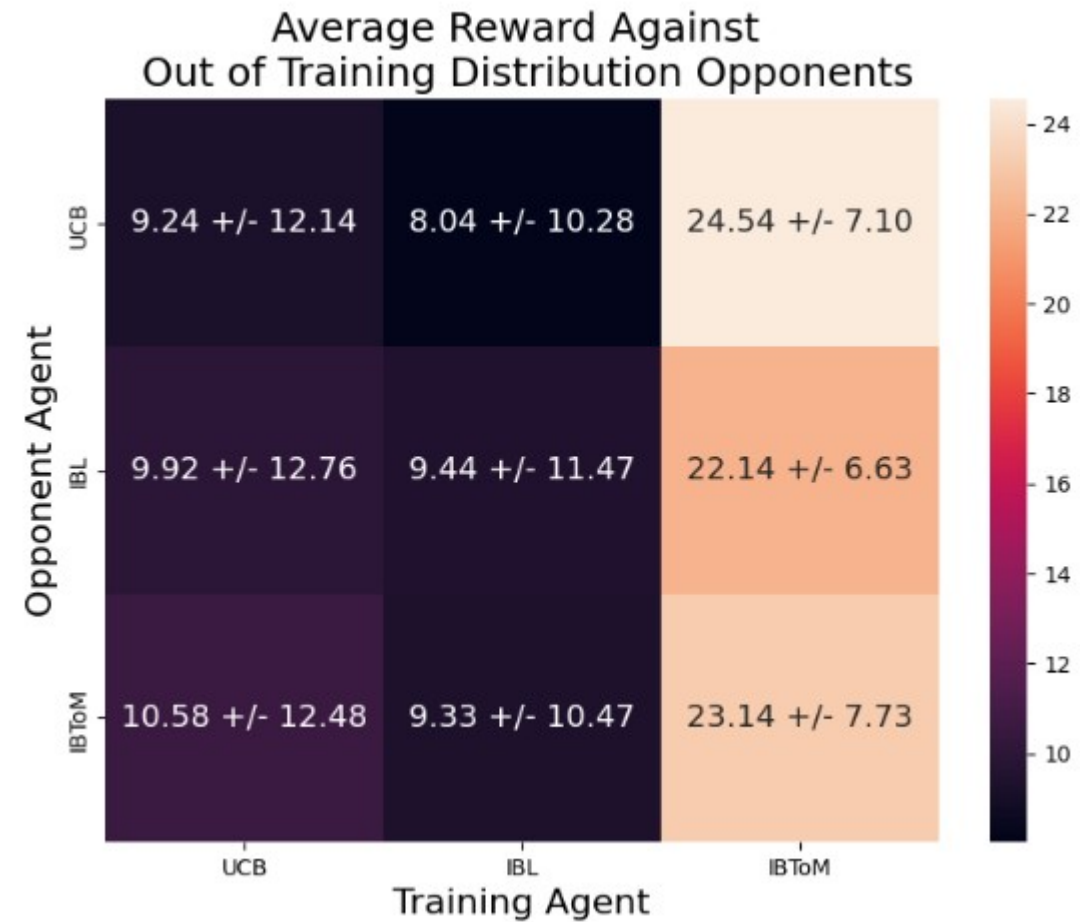
Improved Transfer through Theory of Mind



Improved Performance Against Varied Opponents

We are ultimately interested in performance against a wide range of opponents (real-world applications with humans).

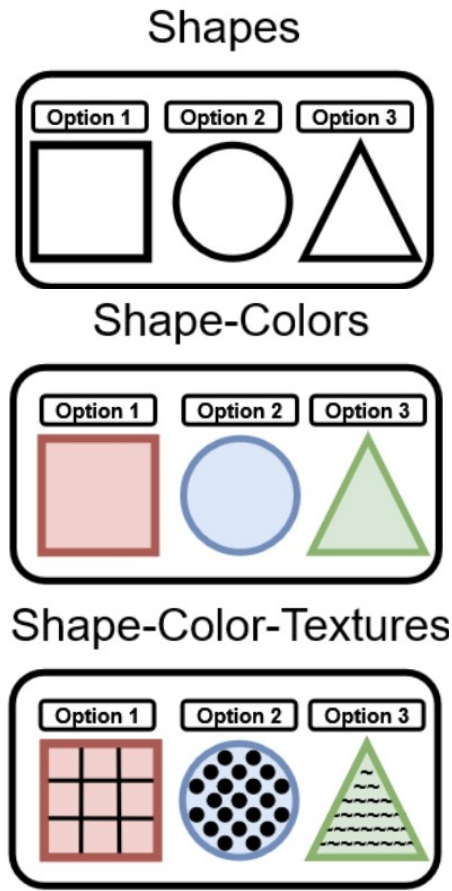
We take trained agents and compete against a randomized sampling of all model types.



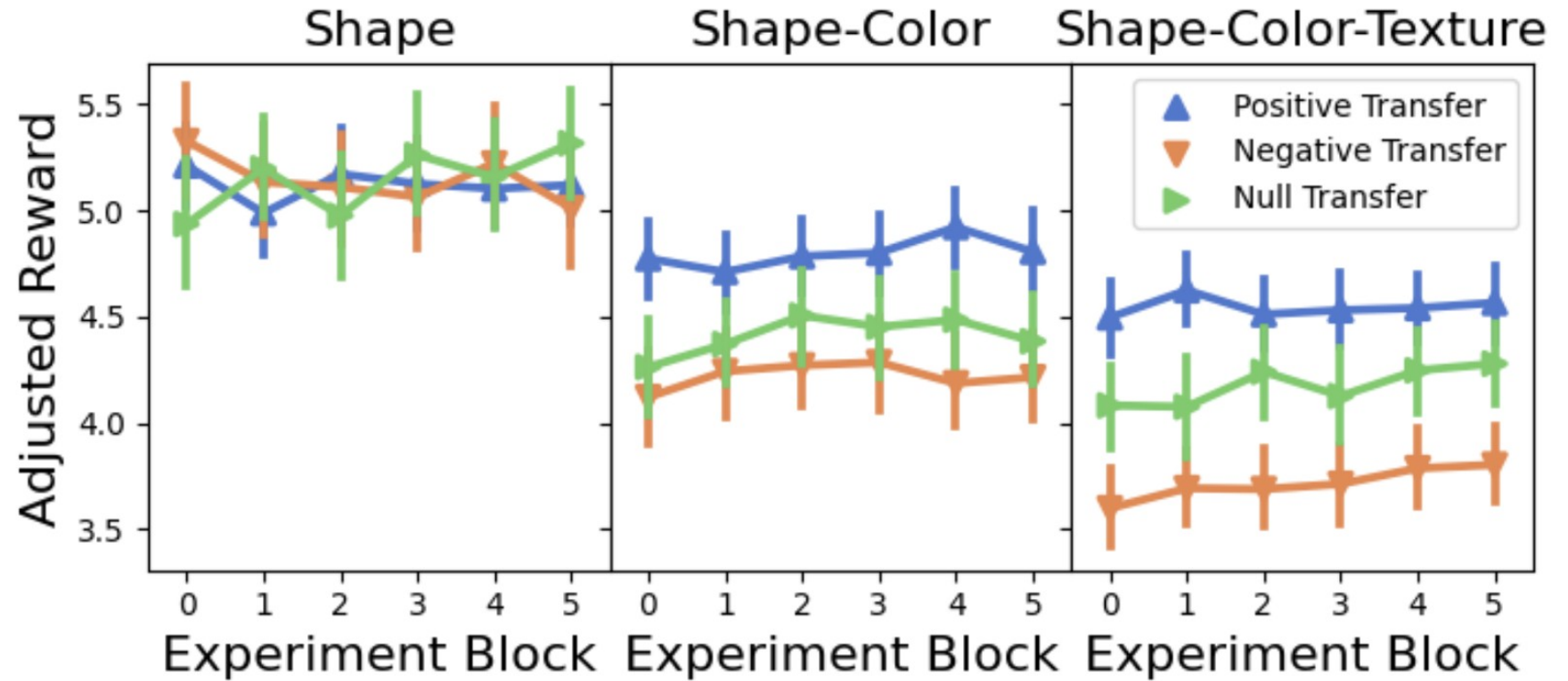


Conclusions and Future Directions

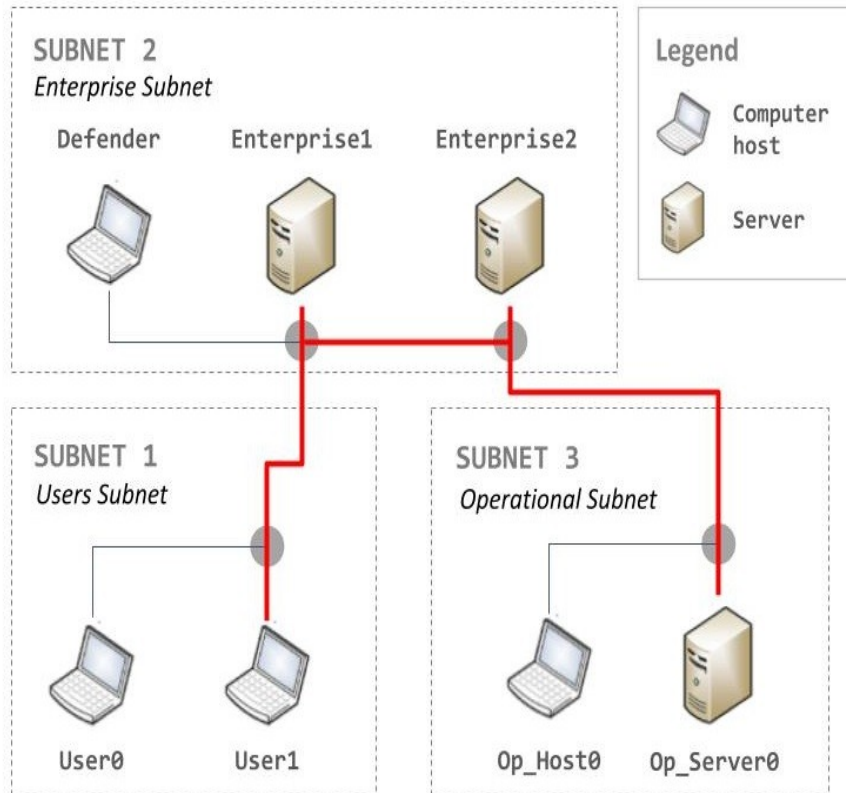
Transfer of Learning: Single Player Games



Participant Adjusted Reward by Experiment Block



More Complex Environments: CyBorg



CybORG Game Help

Round 7/25 Last round: **-0.1** Total loss: **-0.3**

Subnet	Subnet Name	IP Address	Hostname	Activity	Compromised level
10.0.60.128/28	Sub2 - Enterprise	10.0.60.130	Defender	None	No
10.0.60.128/28	Sub2 - Enterprise	10.0.60.131	Enterprise1	Exploit	User
10.0.60.128/28	Sub2 - Enterprise	10.0.60.135	Enterprise2	None	No
10.0.178.16/28	Sub3 - Operational	10.0.178.23	Op_Host0	None	No
10.0.178.16/28	Sub3 - Operational	10.0.178.19	Op_Server0	None	No
10.0.29.176/28	Sub1 - User	10.0.29.181	User0	None	No
10.0.29.176/28	Sub1 - User	10.0.29.187	User1	None	User

> Select an action: Monitor Analyze Remove Restore

> You chose to: Remove Enterprise2 Next

References

- Sinha, Arunesh, et al. "Stackelberg security games: Looking beyond a decade of success." IJCAI, 2018.
- Sutton, Richard S., and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.
- Simon, Herbert A. "Bounded rationality." *Utility and probability* (1990): 15-18.
- Weiss, Karl, Taghi M. Khoshgoftaar, and DingDing Wang. "A survey of transfer learning." *Journal of Big data* 3.1 (2016): 1-40.
- Yang, Rong, et al. "Scaling-up security games with boundedly rational adversaries: A cutting-plane approach." *Twenty-Third International Joint Conference on Artificial Intelligence*. 2013.
- Nguyen, T. N., & Gonzalez, C. (2022). "Theory of mind from observation in cognitive models and humans". *Topics in Cognitive Science*, 14(4), 665-686.
- Gonzalez, C., Aggarwal, P., Cranford, E. A., & Lebiere, C "Design of Dynamic and Personalized Deception: A Research Framework and New Insights for Cyberdefense". In *Proceedings of the 53rd hawaii international conference on system sciences* (Vol. 1834).
- Gonzalez, Cleotilde, Javier F. Lerch, and Christian Lebiere. "Instance-based learning in dynamic decision making." *Cognitive Science* 27.4 (2003): 591-635
- Malloy et al. "Accounting for Transfer of Learning using Human Behavior Models" *Under review*.



Questions?