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

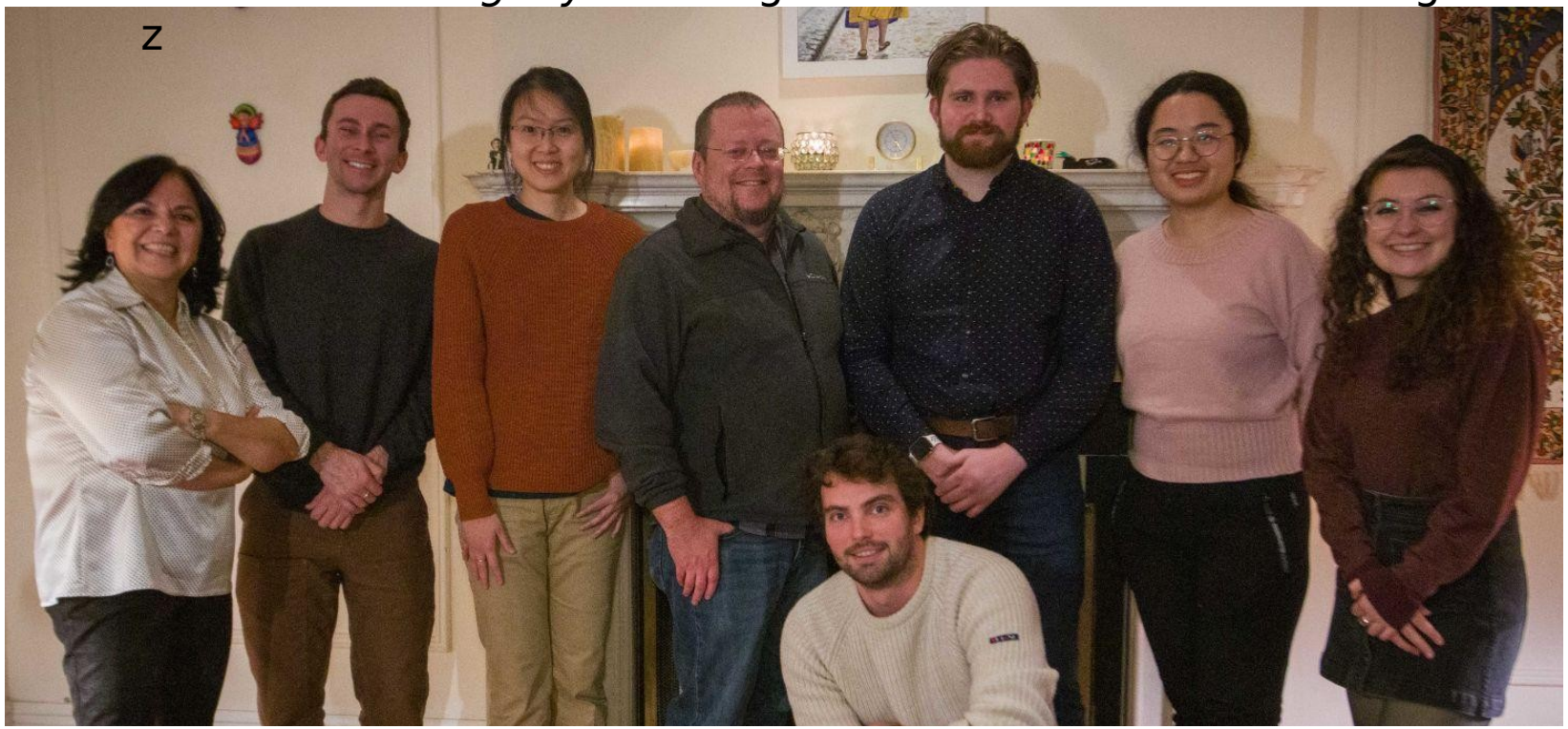
Coty Gonzalez

Ngoc Ngueyn

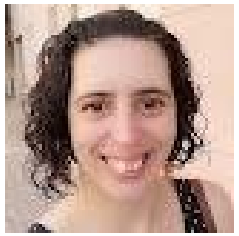
Jeffrey Flag

Me

Erin Bugbee



Maria Ferreria



Chase MacDonald

Baptiste Prebot

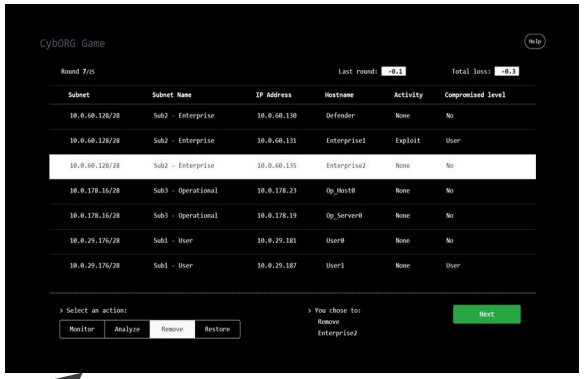
Yinuo Du

Background:
Attack/Defense Decision Making

Attack/Defense Scenarios

Attacker

Network

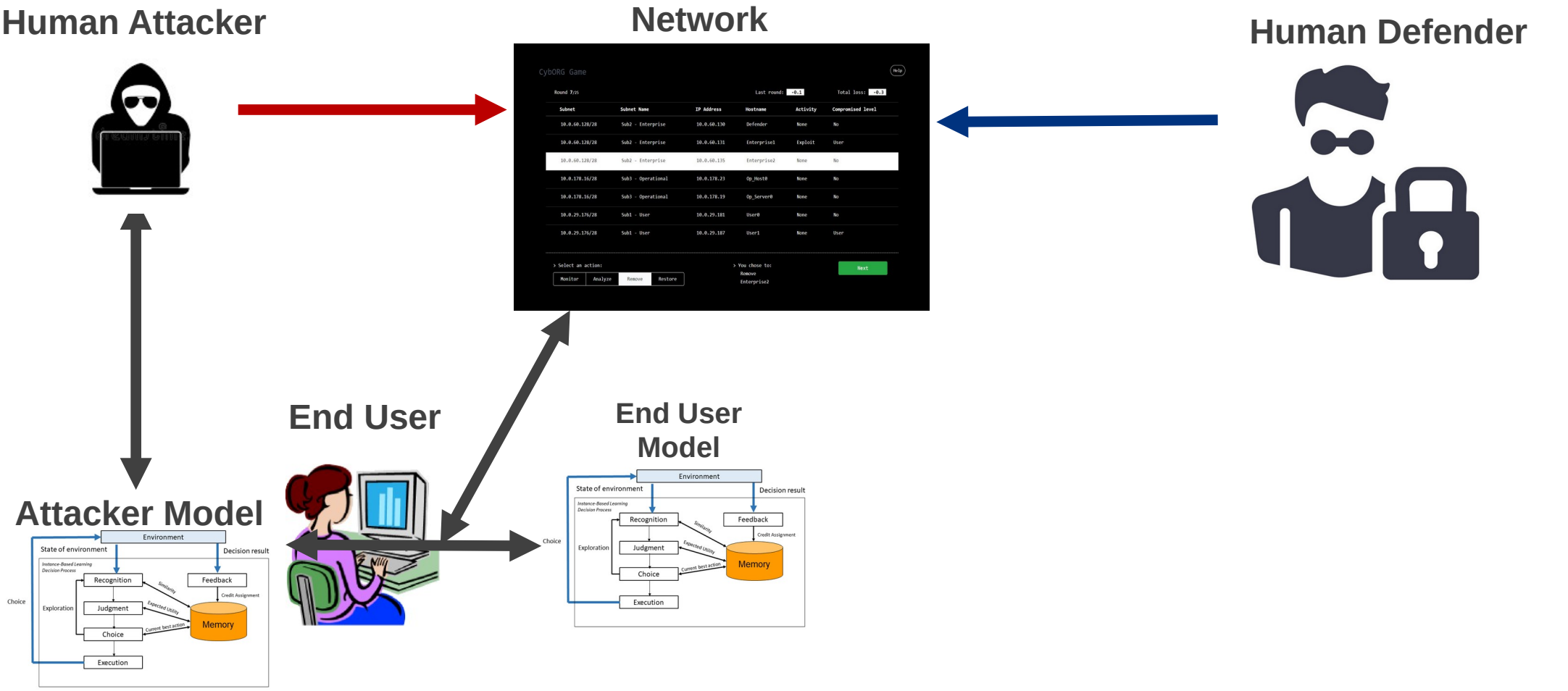


Defender

End User

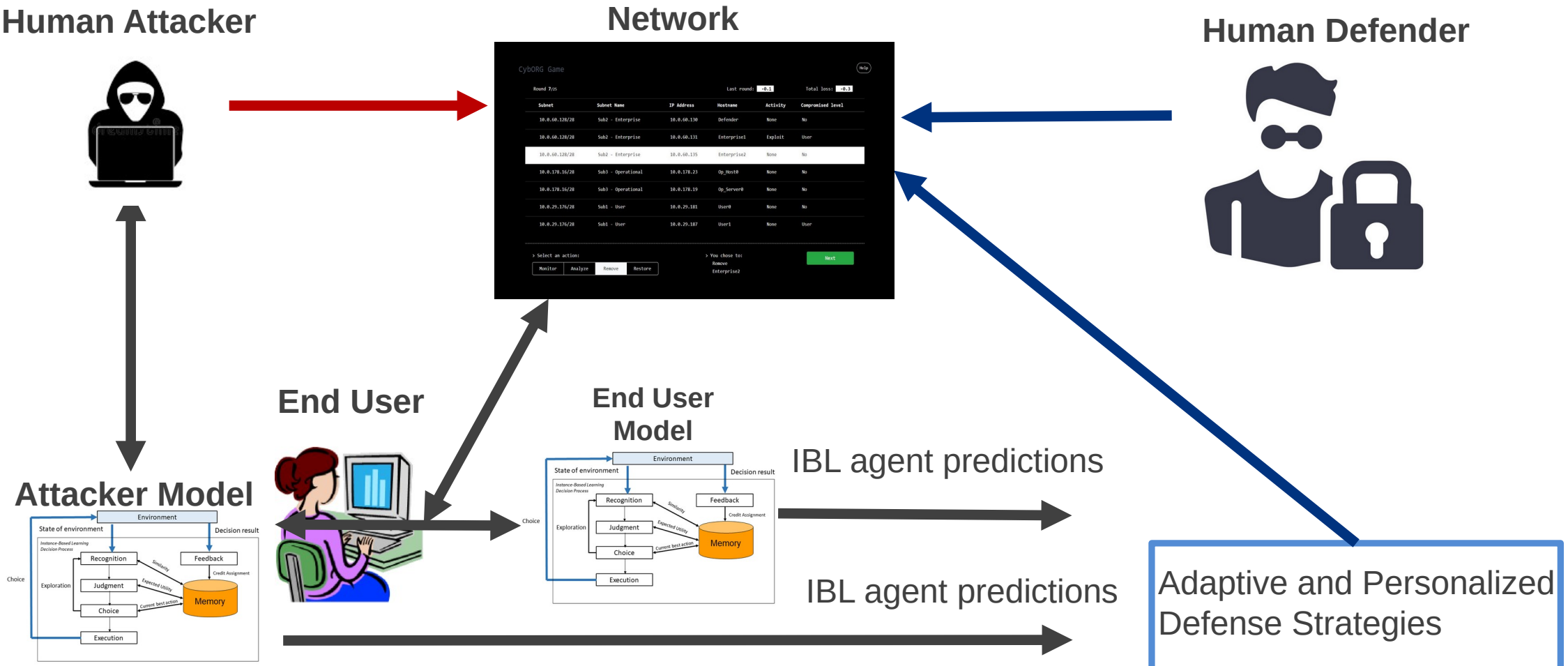
Gonzalez, C., Aggarwal, P., Cranford, E. A., & Lebiere, C Design of Dynamic and Personalized Deception: A Research Framework and New Insights for Cyberdefense. In *Proceedings of the 53rd hawaii international conference on system sciences* (Vol. 1834).

Attack/Defense Modeling



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

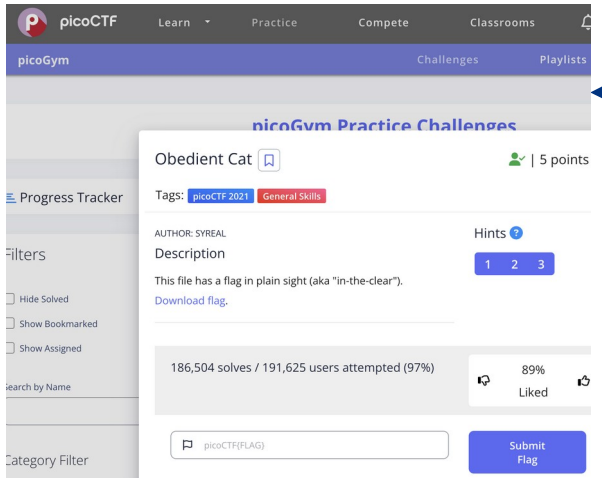


Gonzalez, C., Aggarwal, P., Cranford, E. A., & Lebiere, C Design of Dynamic and Personalized Deception: A Research Framework and New Insights for Cyberdefense. In *Proceedings of the 53rd hawaii international conference on system sciences* (Vol. 1834).

Attack/Defense Real World

Human Participant

Capture The Flag



CTF Designer

How do Humans Learn Cybersecurity?

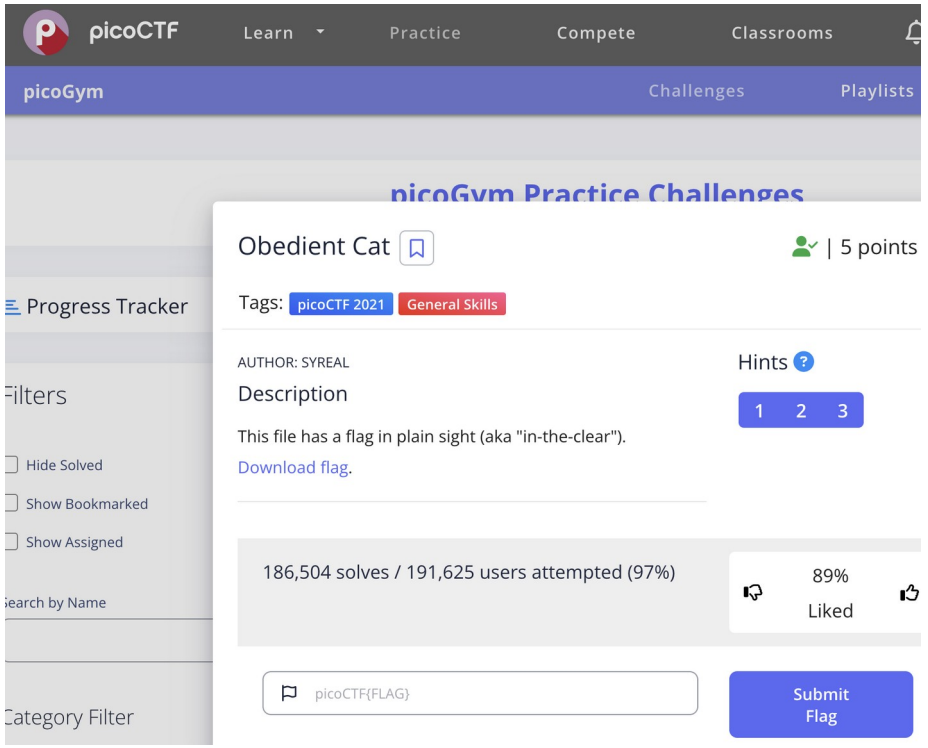
Transfer of Learning.

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

Theory of Mind.

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

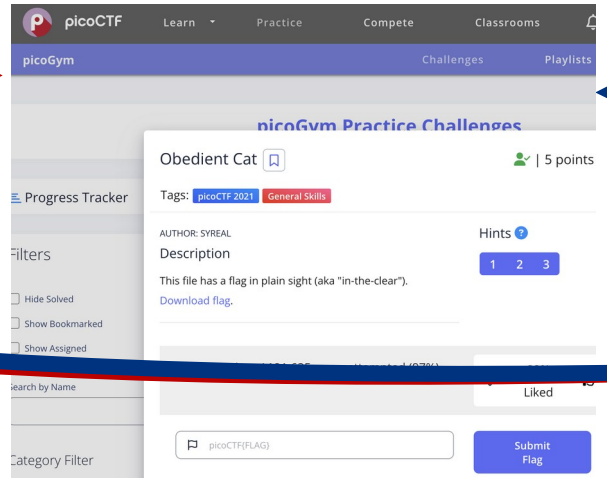
Capture The Flag Challeng



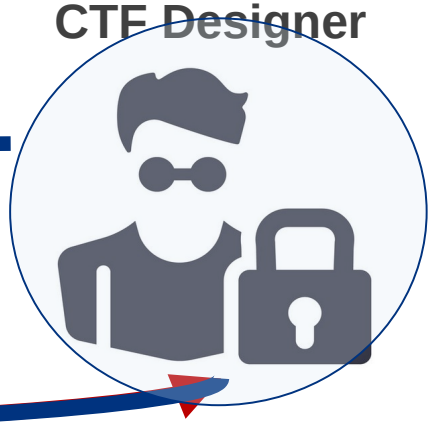
Attack/Defense Real World

Human Participant

Capture The Flag



CTF Designer



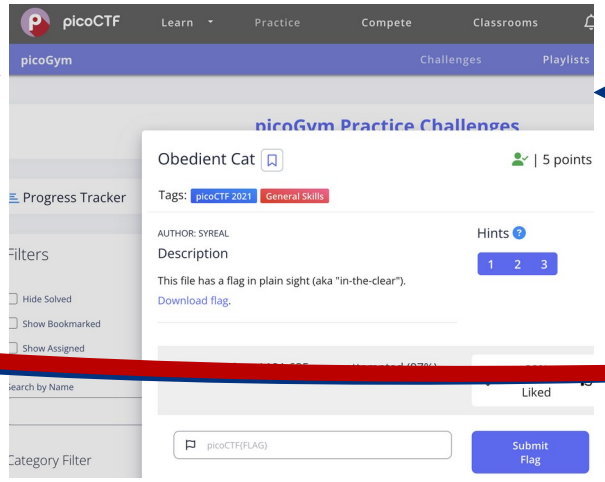
Theory of Mind

Transfer of Learning

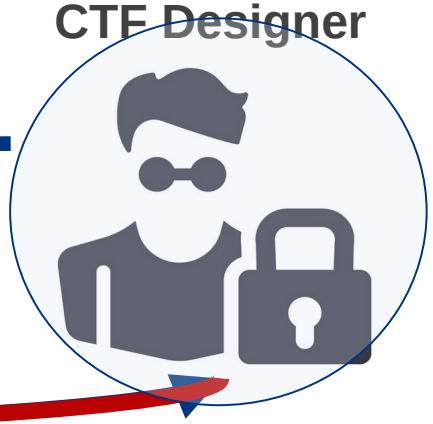
Attack/Defense Real World

Human Participant

Capture The Flag



CTF Designer



Transfer of Learning

Theory of Mind

Cyber Testbeds

Realistic

Abstract

Deception Technique

Simplistic

Complex

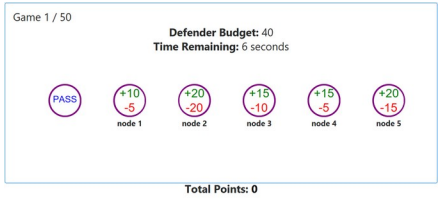
Phishing Training

Dynamic Training

Box Game

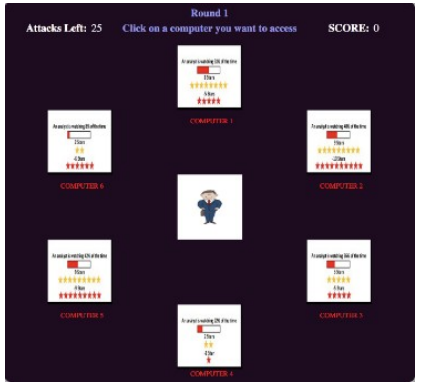
Signaling

Honeypot



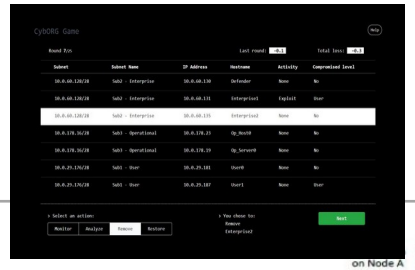
Decoying

Insider Attack

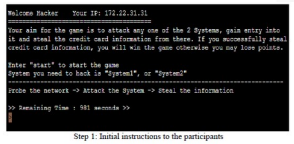


Signaling

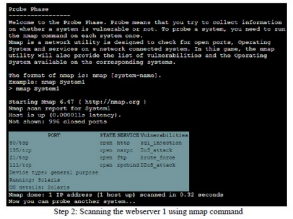
Cage: CybOrg



Hackl



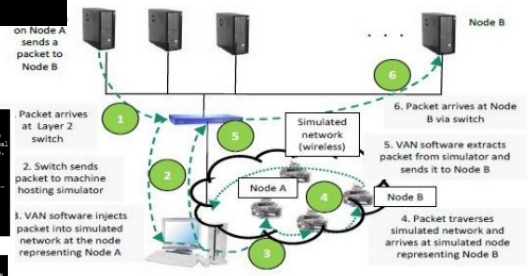
Step 1: Initial instructions to the participants



Step 2: Scanning the webserver 1 using nmap command

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).

Carnegie Mellon University

Cyber Testbeds

Realistic

Abstract

Deception Technique

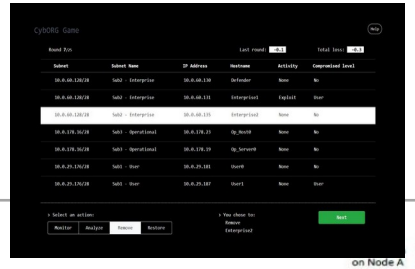
Simplistic

Complex

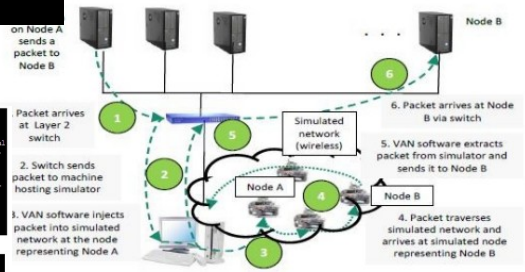
Phishing Training

Dynamic Training

Cage: CybOrg

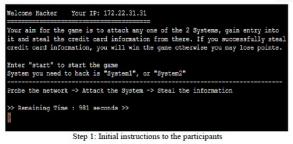


CyberVAN

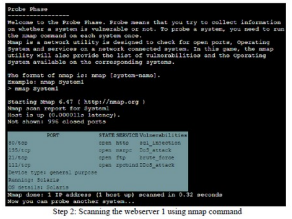


Masking

Hackl



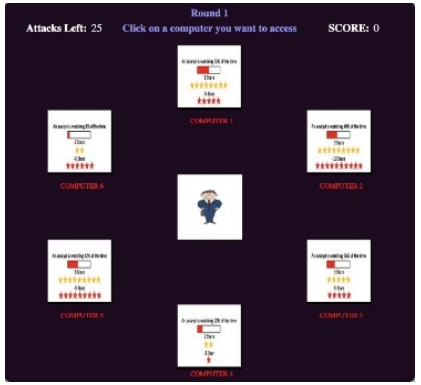
Step 1: Initial instructions to the participants



Step 2: Scanning the webserver 1 using nmap command

Decoying + Masking

Insider Attack

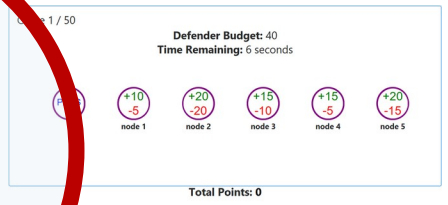


Signaling

Box Game

Signaling

Honeypot



Decoying

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).

Cyber Testbeds

Realistic

Abstract

Deception Technique

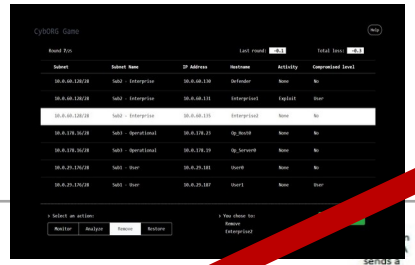
Simplistic

Complex

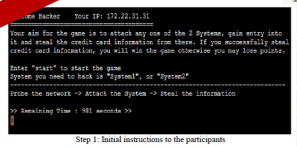
Phishing Training

Dynamic Training

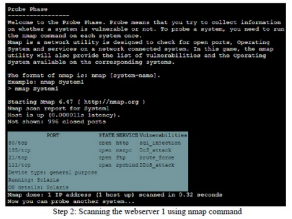
Cage: CybOrg



Hackl



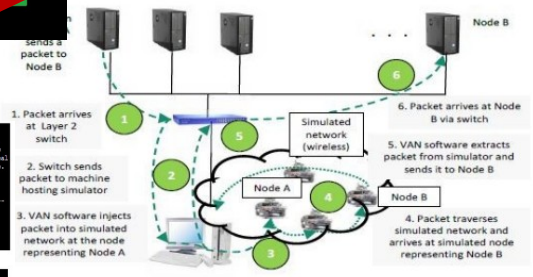
Step 1: Initial instructions to the participants



Step 2: Scanning the webserver 1 using nmap command

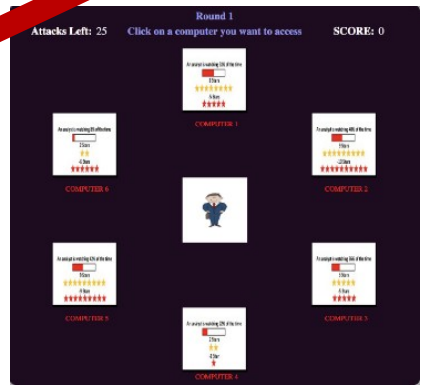
Decoying + Masking

CyberVAN



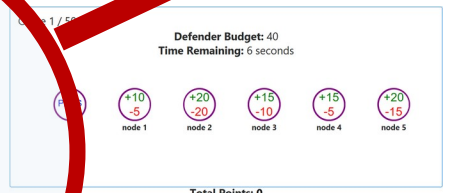
Masking

Insider Attack



Signaling

Honeypot



Decoying

Box Game

Signaling

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).

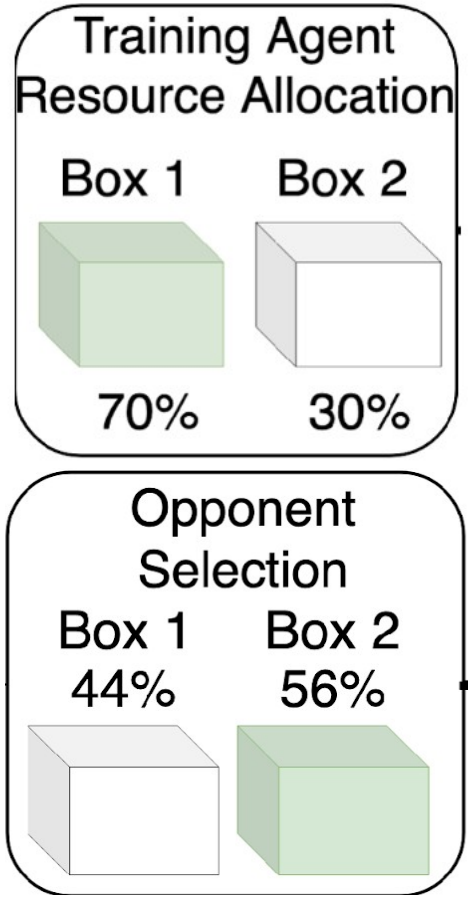
Stackelberg Security Games

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.



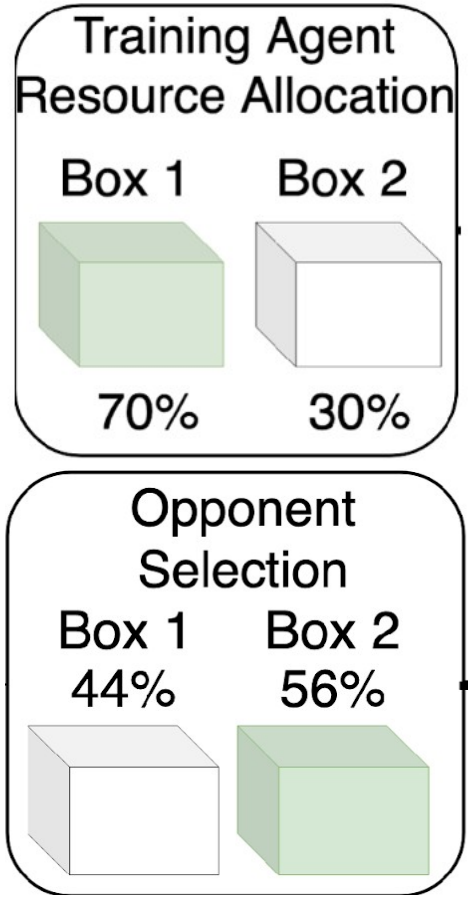
Stackelberg Security Games: Looking Beyond a Decade of Success
Arunesh Sinha, Fei Fang, Bo An, Christopher Kiekintveld, Milind Tambe

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.



Stackelberg Security Games: Looking Beyond a Decade of Success
Arunesh Sinha, Fei Fang, Bo An, Christopher Kiekintveld, Milind Tambe

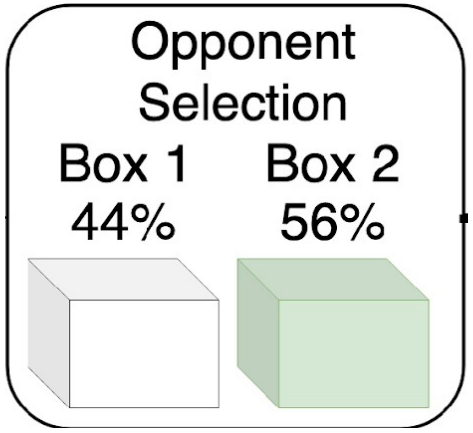
Challenges of applying SSGs to Cybersecurity

SSGs assume long deploy/asses time.

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

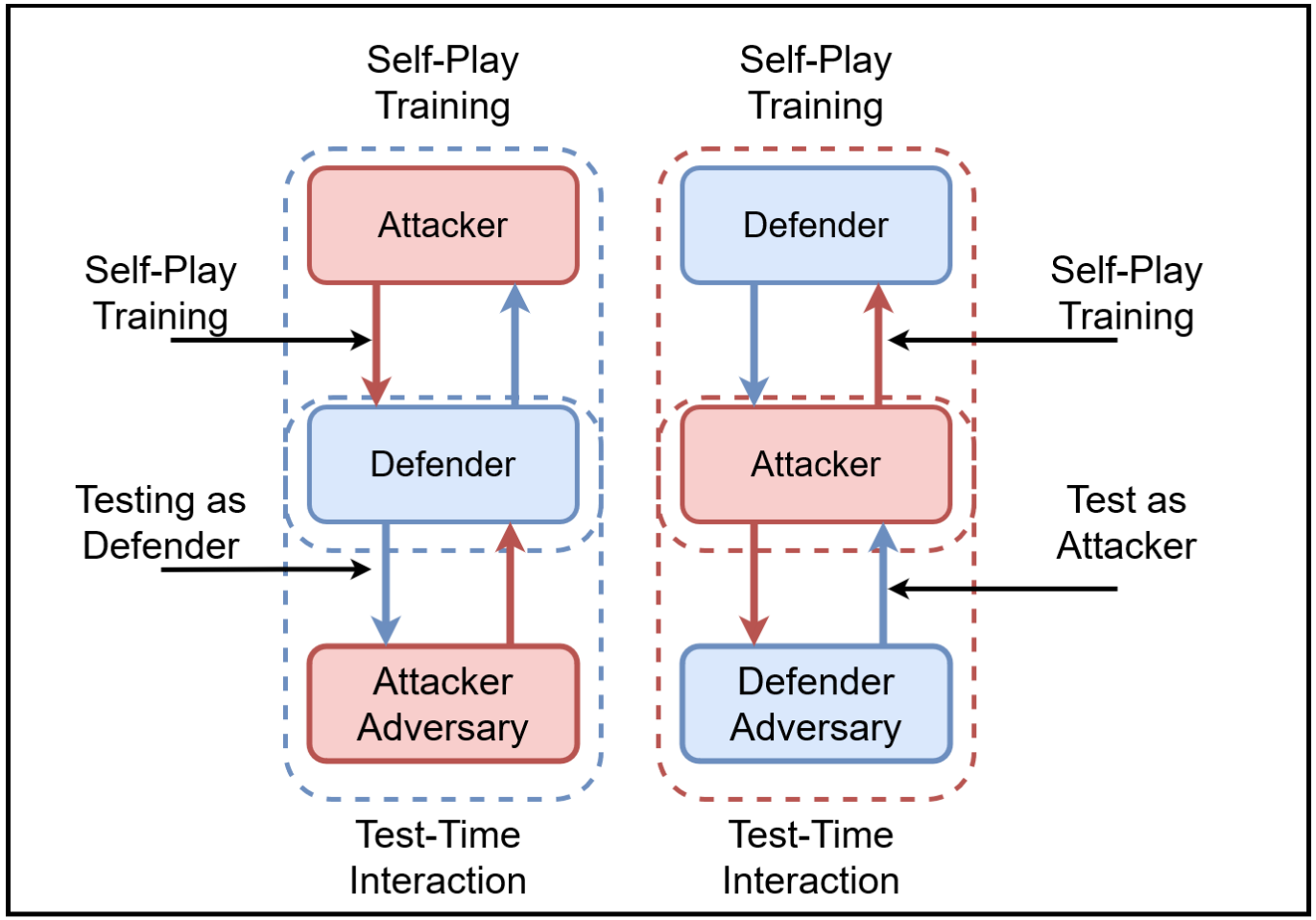
Real-world applications of SSGs have **focused on biases** present in human decision making.

- 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.

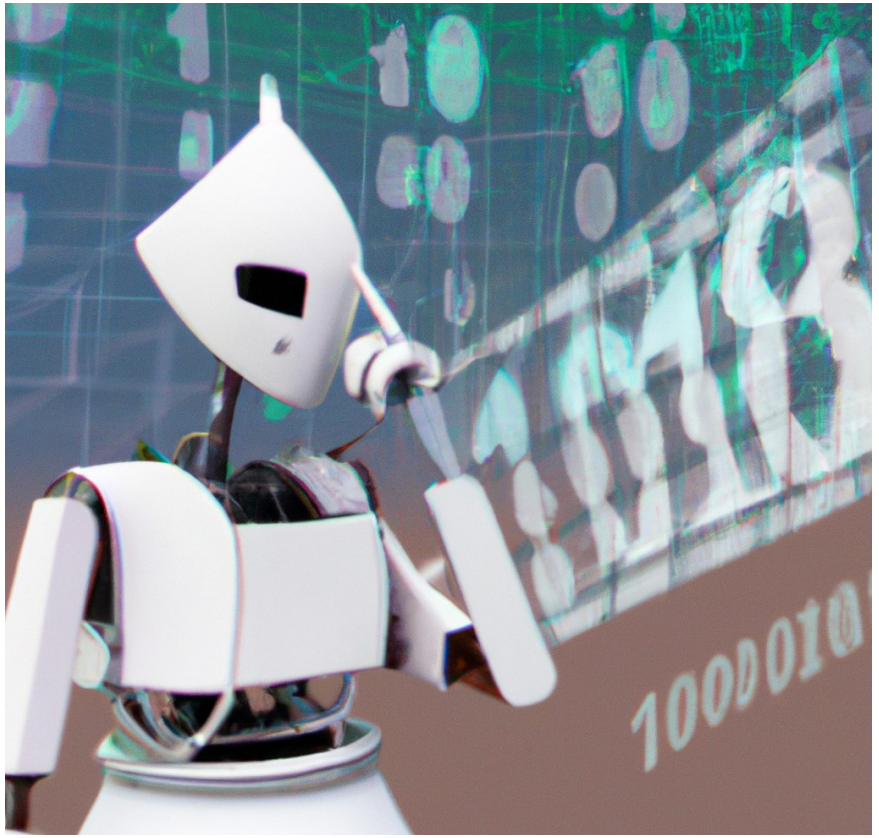
- Can allow agents to account for things they do not directly observe.
- Opponent’s behavior can often reveal what they observe and what their goals are.

Theory of mind from observation in cognitive models and humans.
Nguyen, T. N., & Gonzalez, C.

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*.



A survey of transfer learning.
Weiss, K., Khoshgoftaar, T. M., & Wang, D.

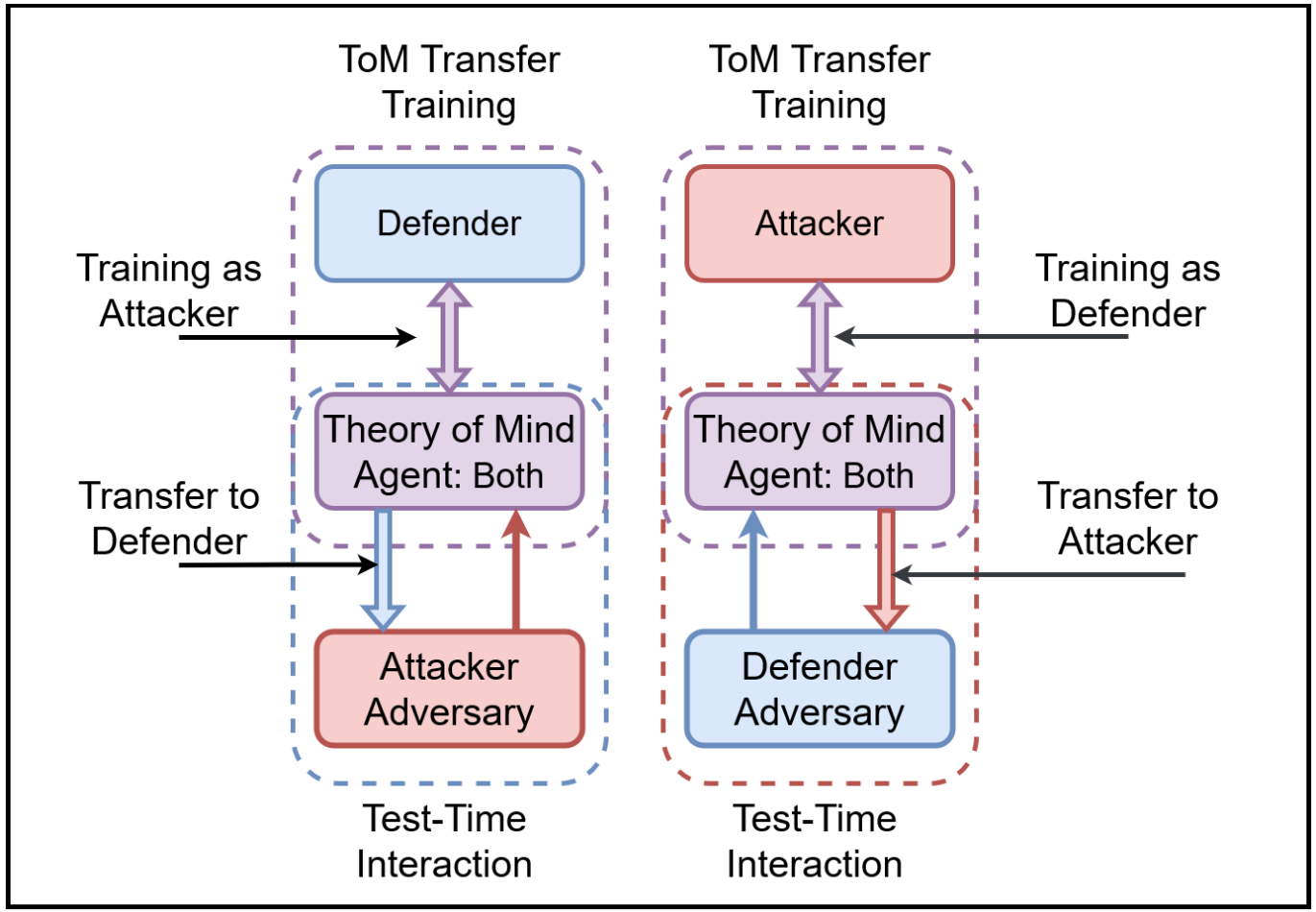
Human Transfer of Learning

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.

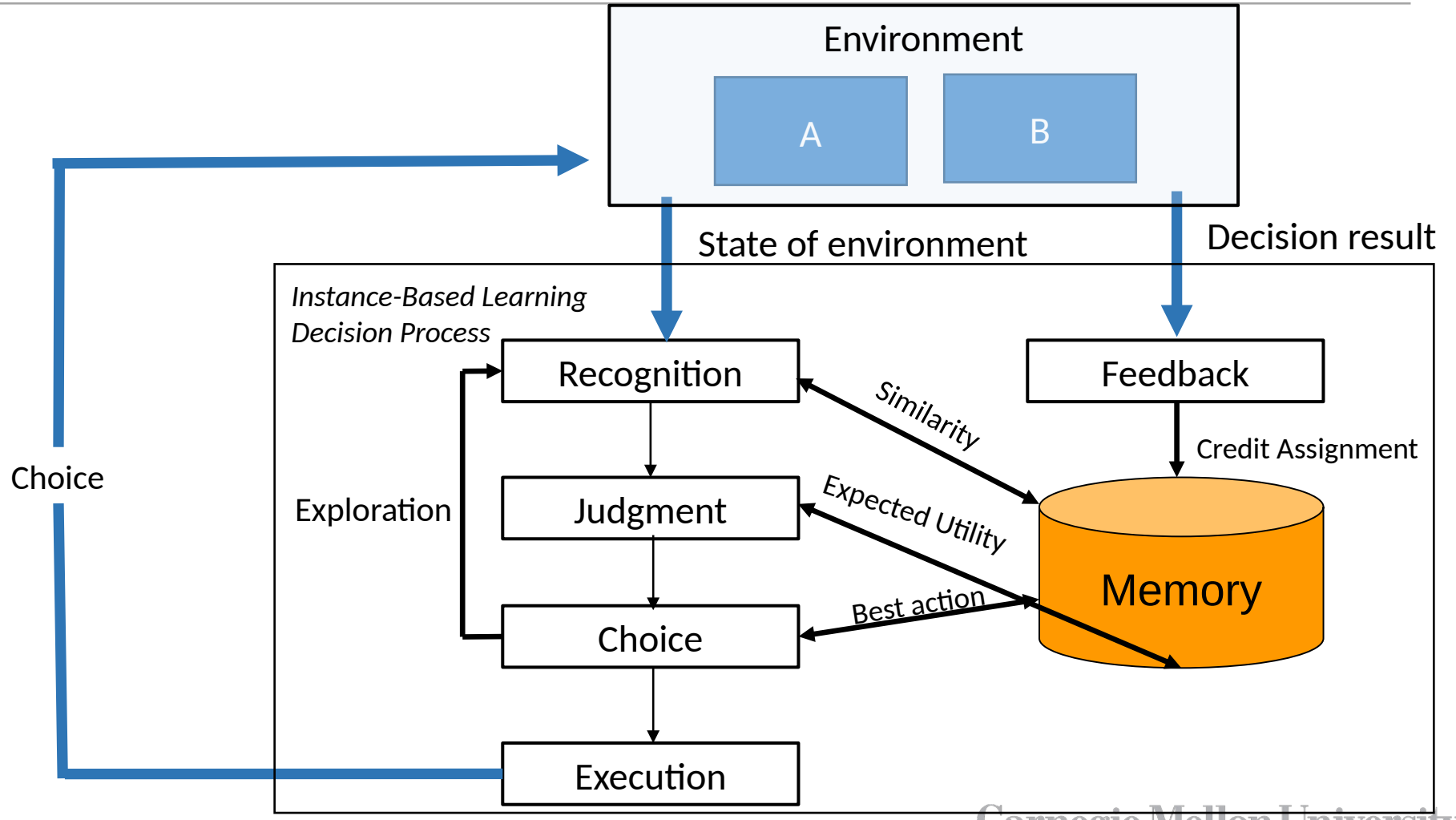
Humans use varied experience, theory of mind, and transfer learning to overcome biases.

Theory of Mind + Transfer of Learning

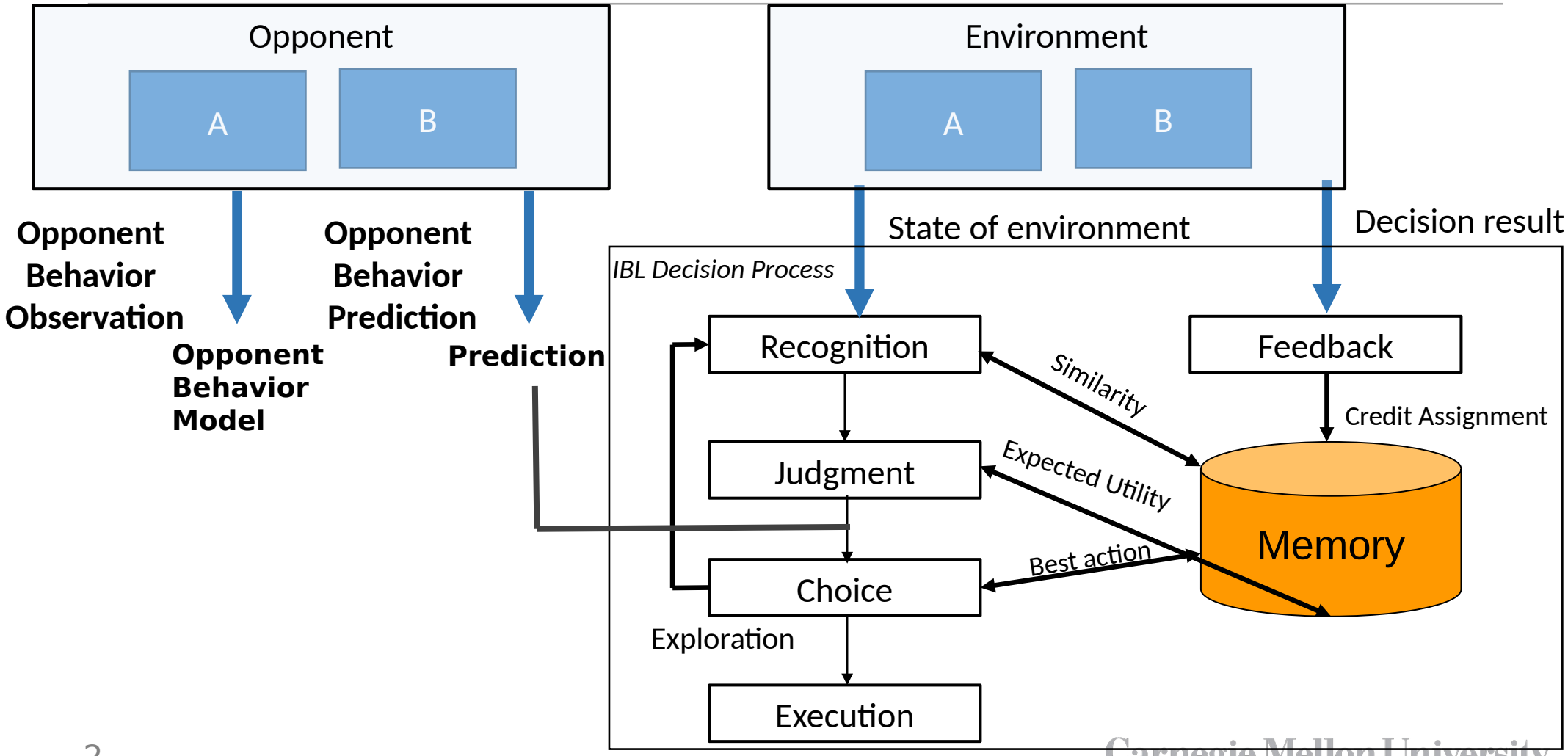


ToM and ToL with Instance Based Learning

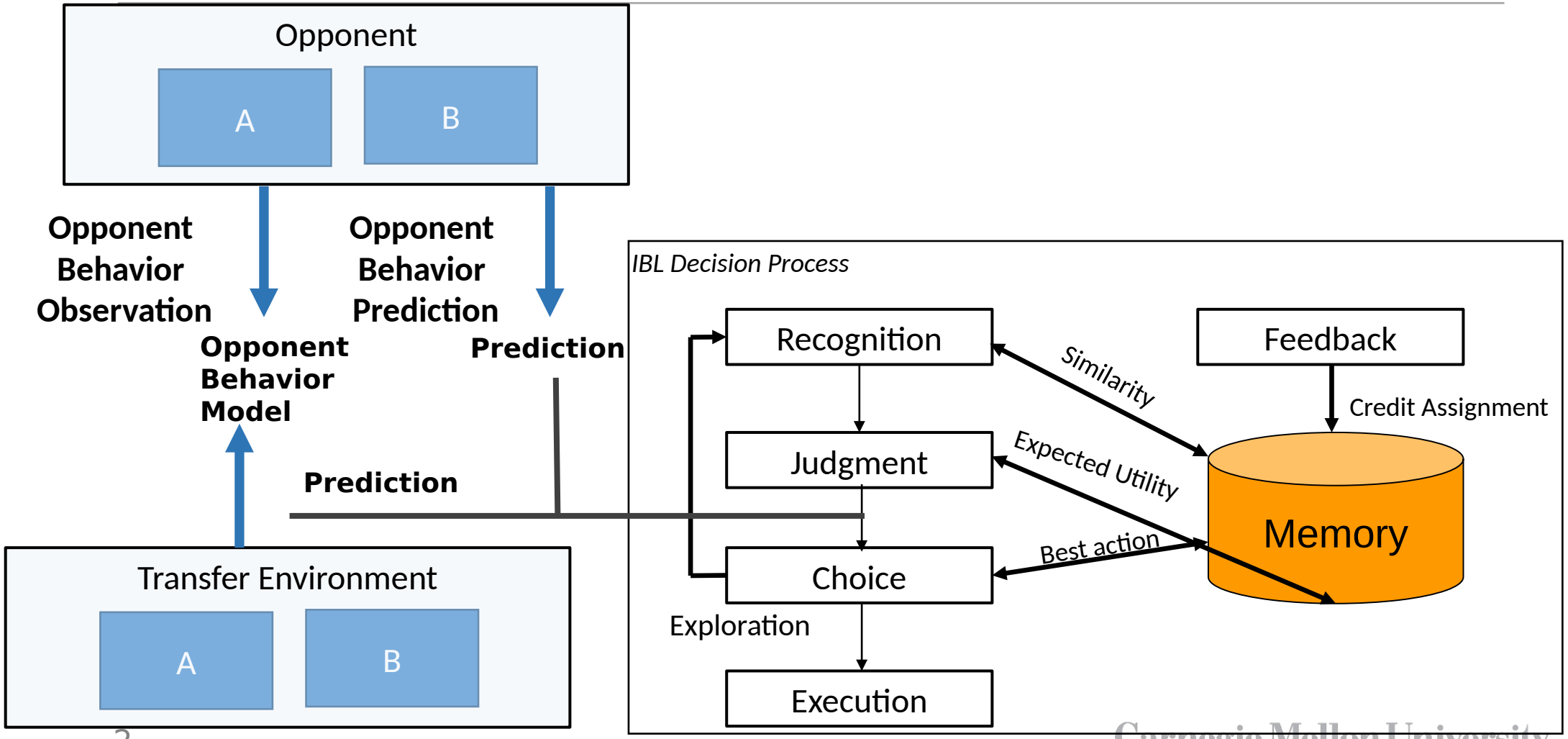
Instance Based Learning Model



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

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

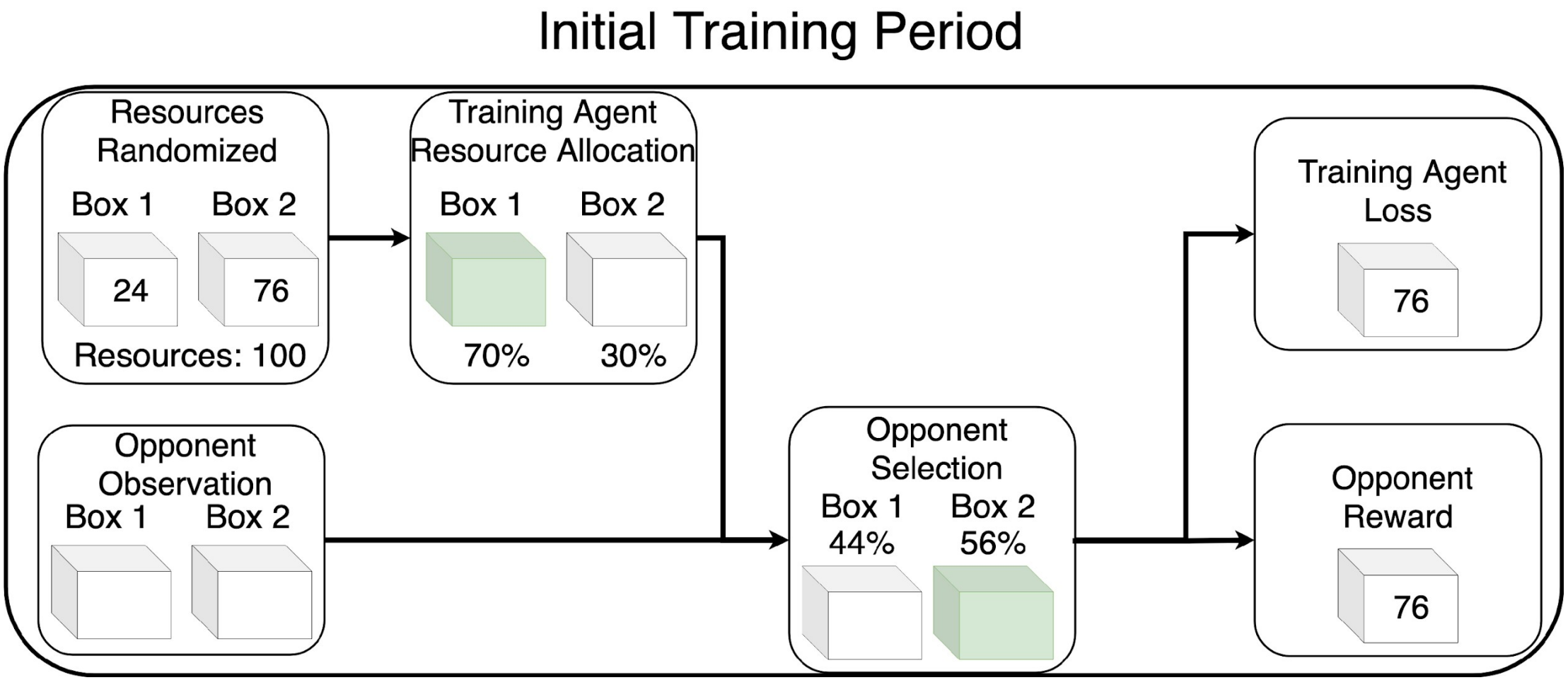
Optimal UCB model

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

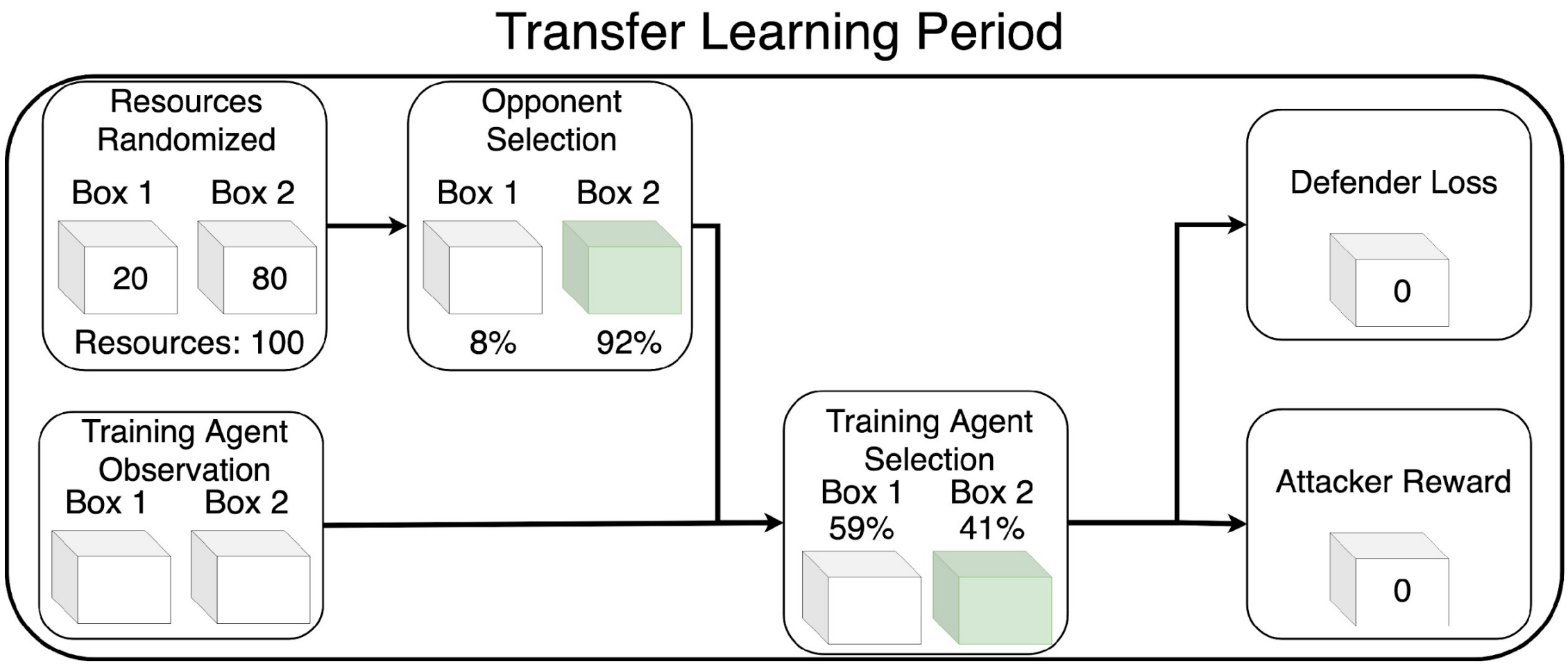
Reinforcement learning: An introduction RS Sutton, AG Barto

Simulation Experimentation

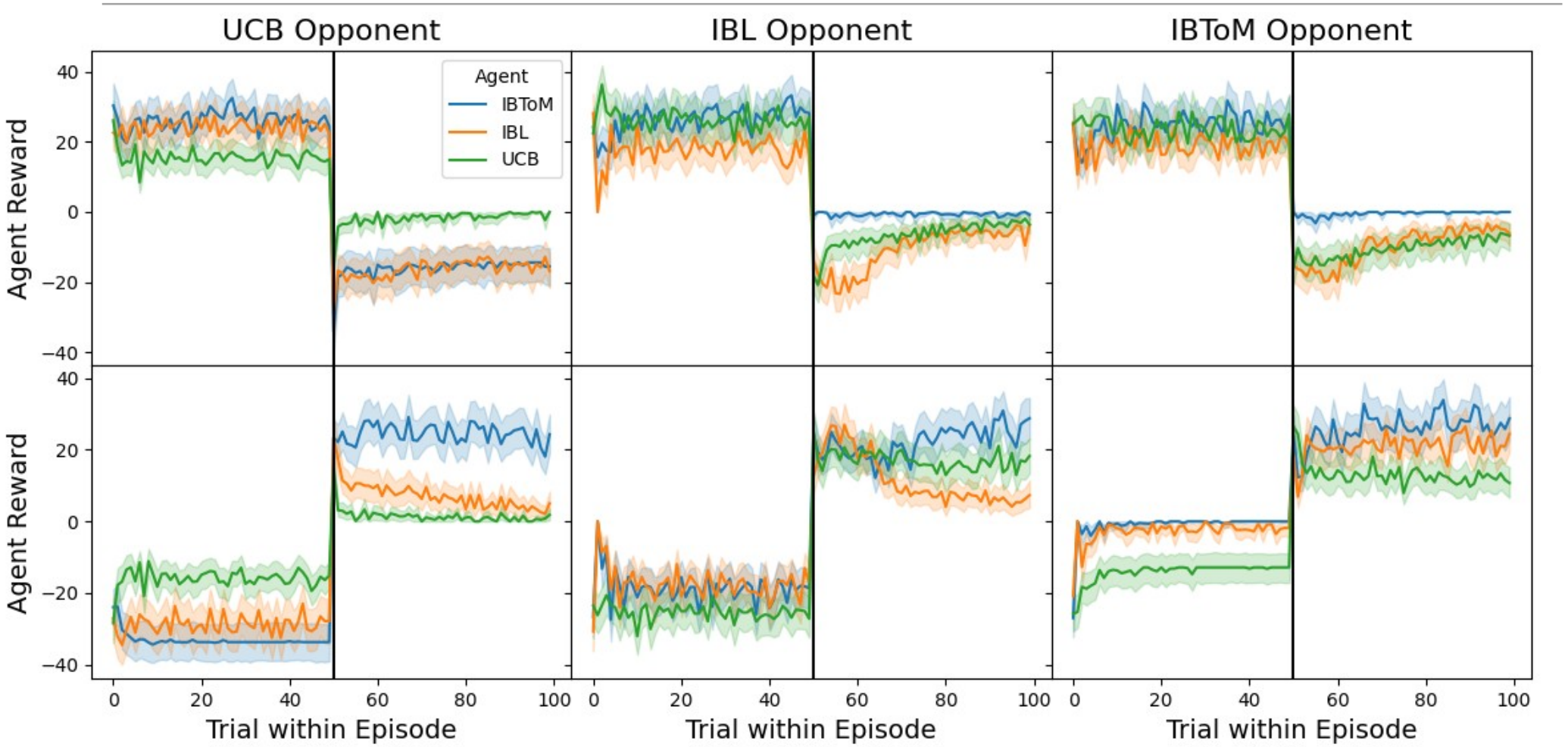
Experimentation Environment



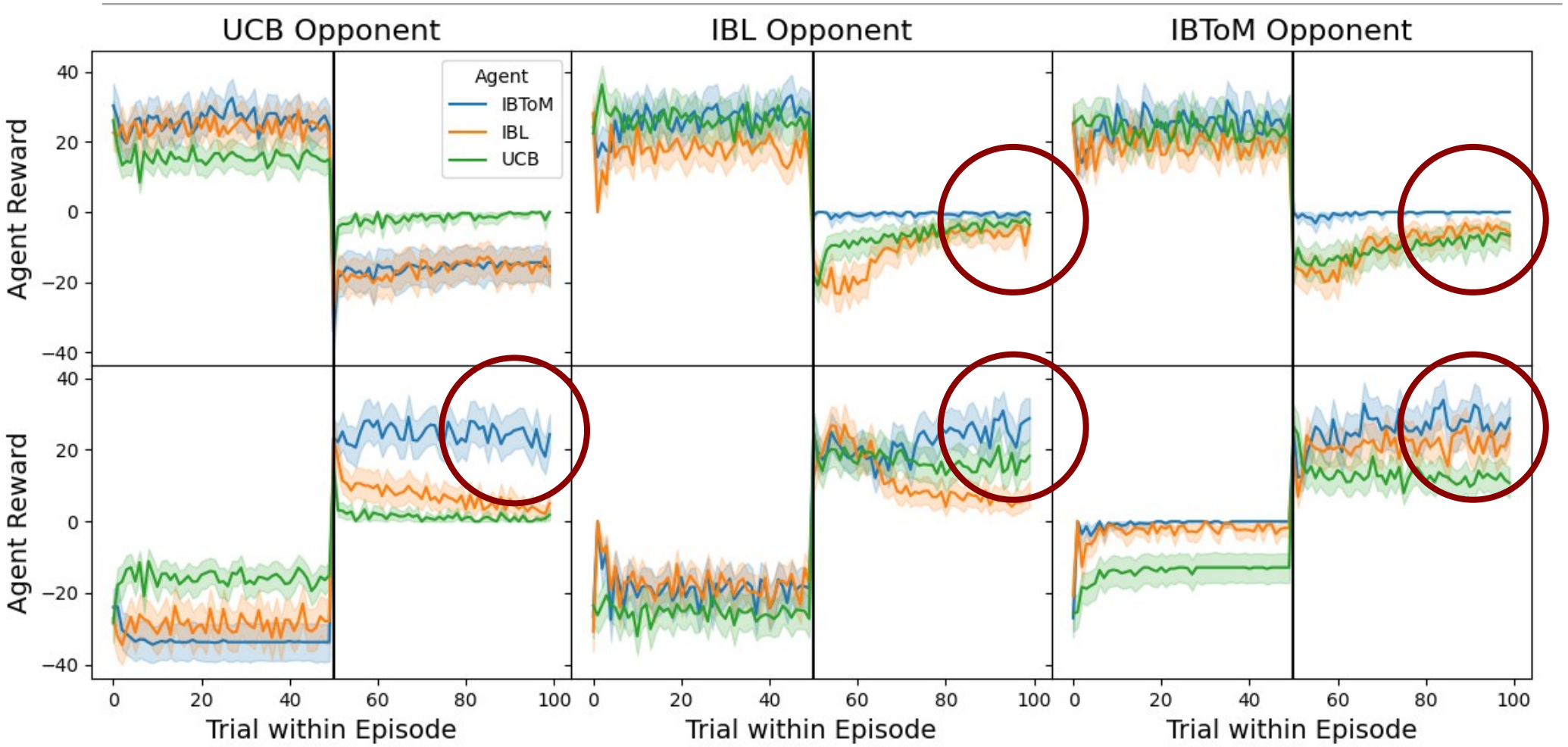
Experimentation Environment



Improved Transfer through Theory of Mind



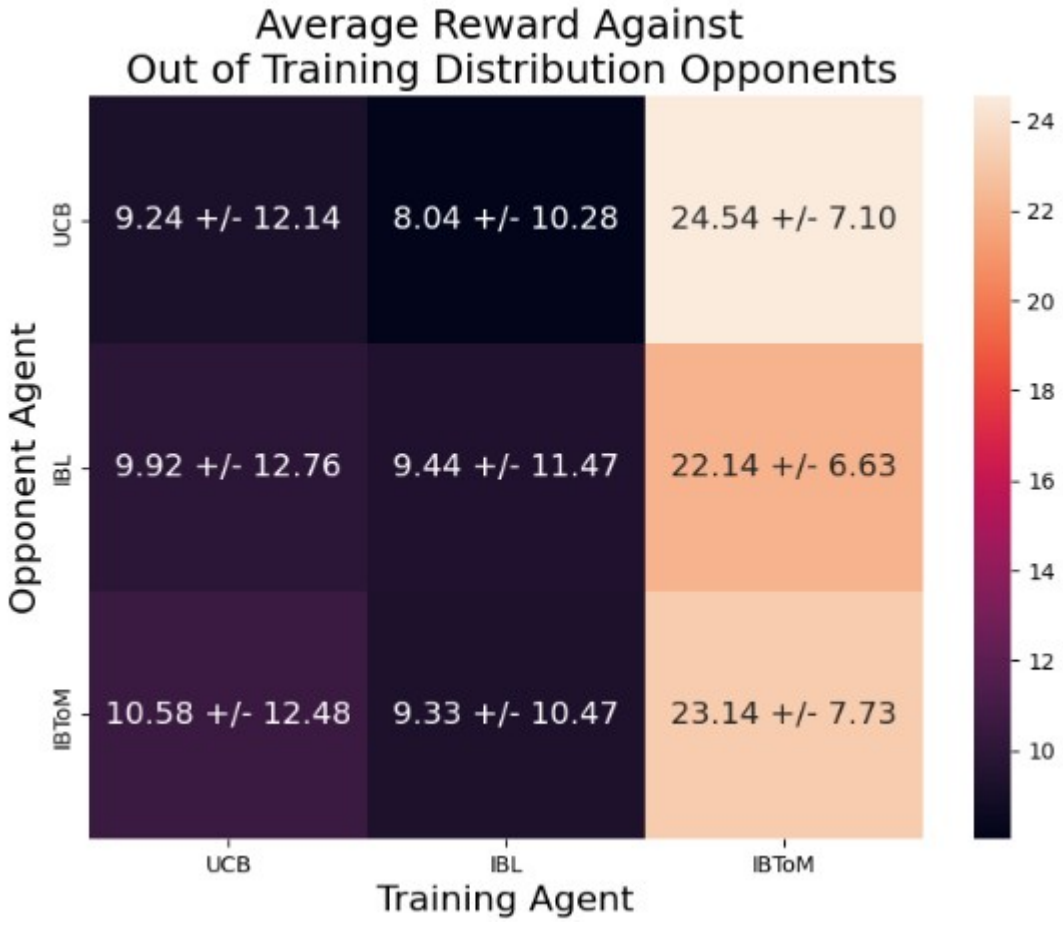
Improved Transfer through Theory of Mind



Improved Performance Against Varied Opponents

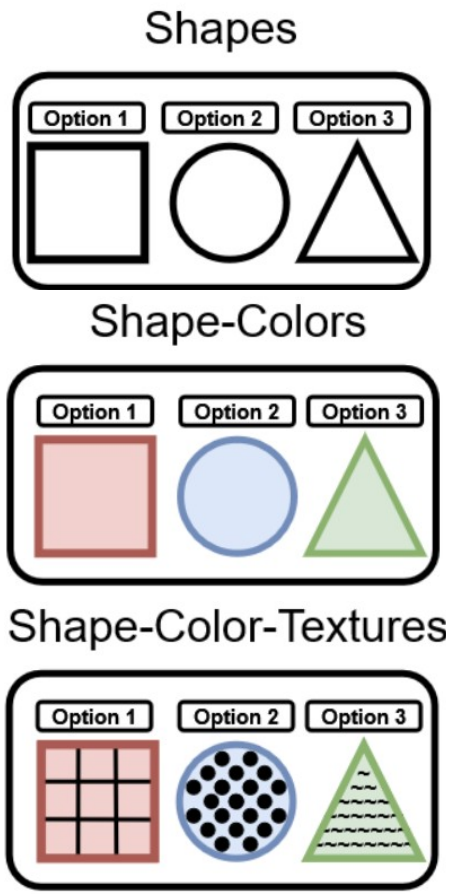
We are ultimately interested in performance against a wide range of opponents (real-world applications with humans).

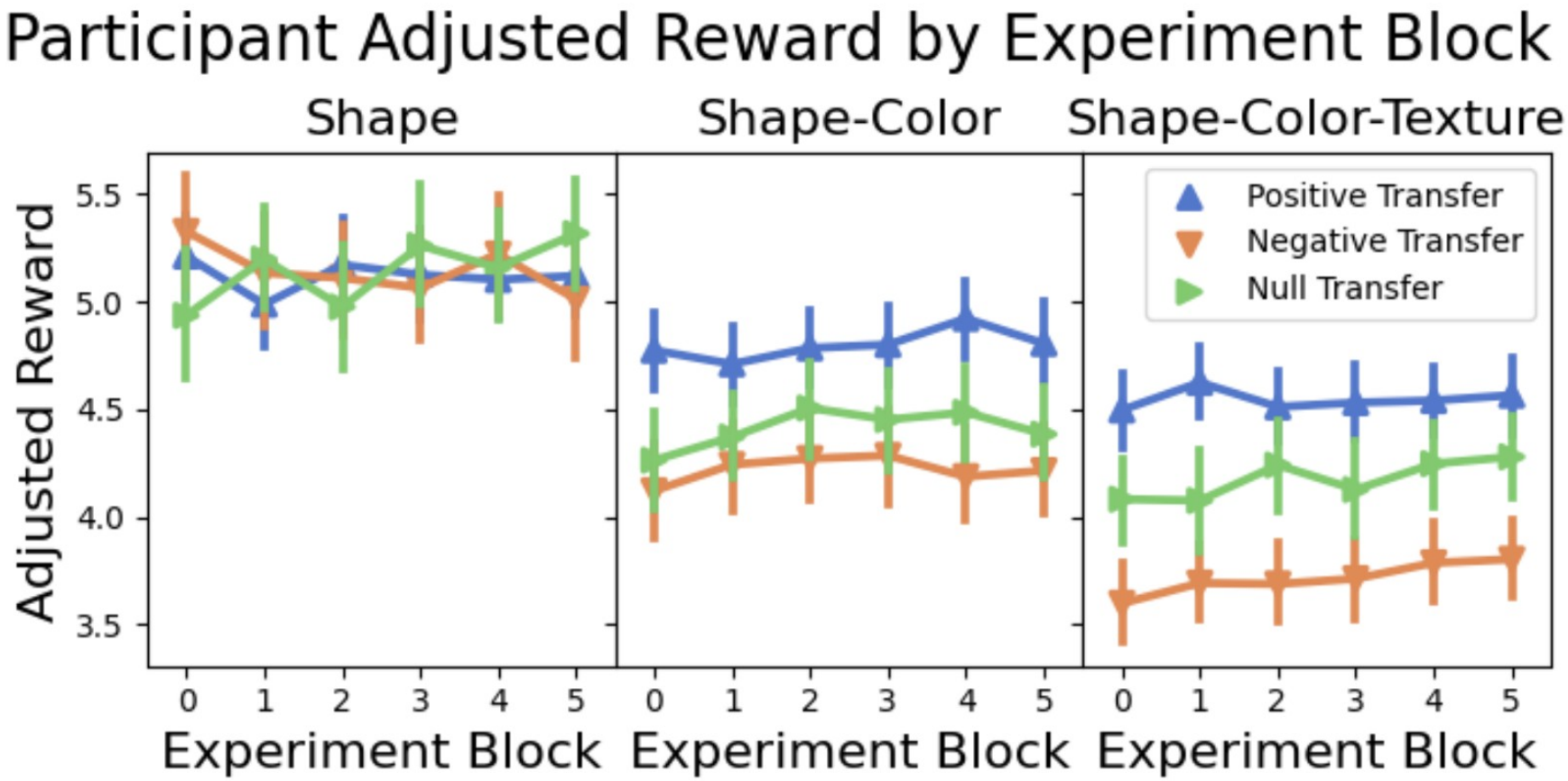
We take trained agents and compete against a randomized sampling of all model types.



Conclusions and Future Directions

Transfer of Learning: Single Player Games

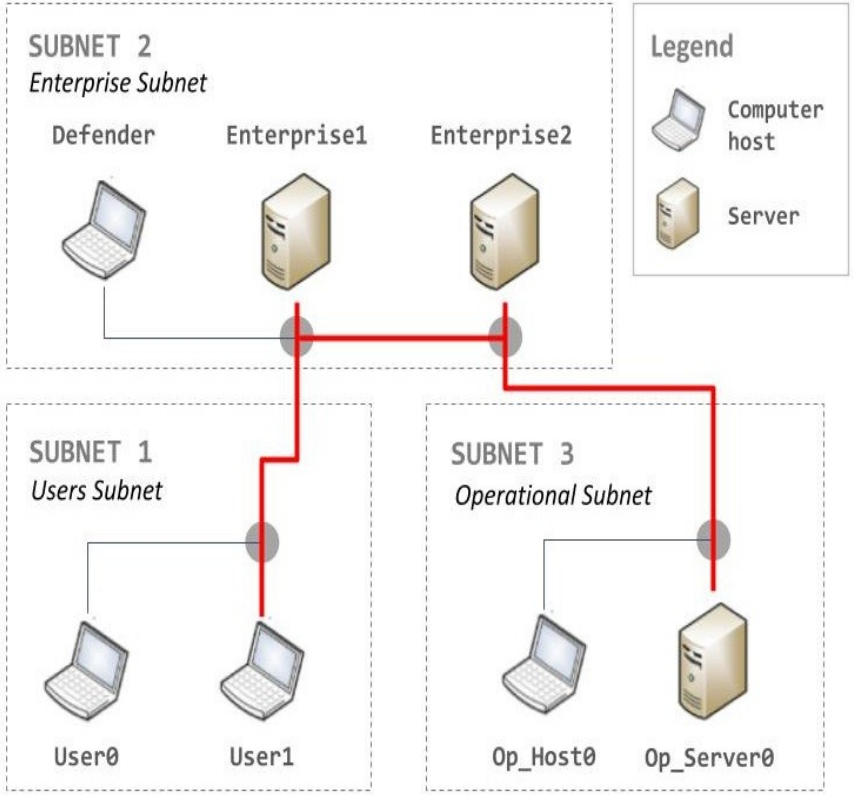


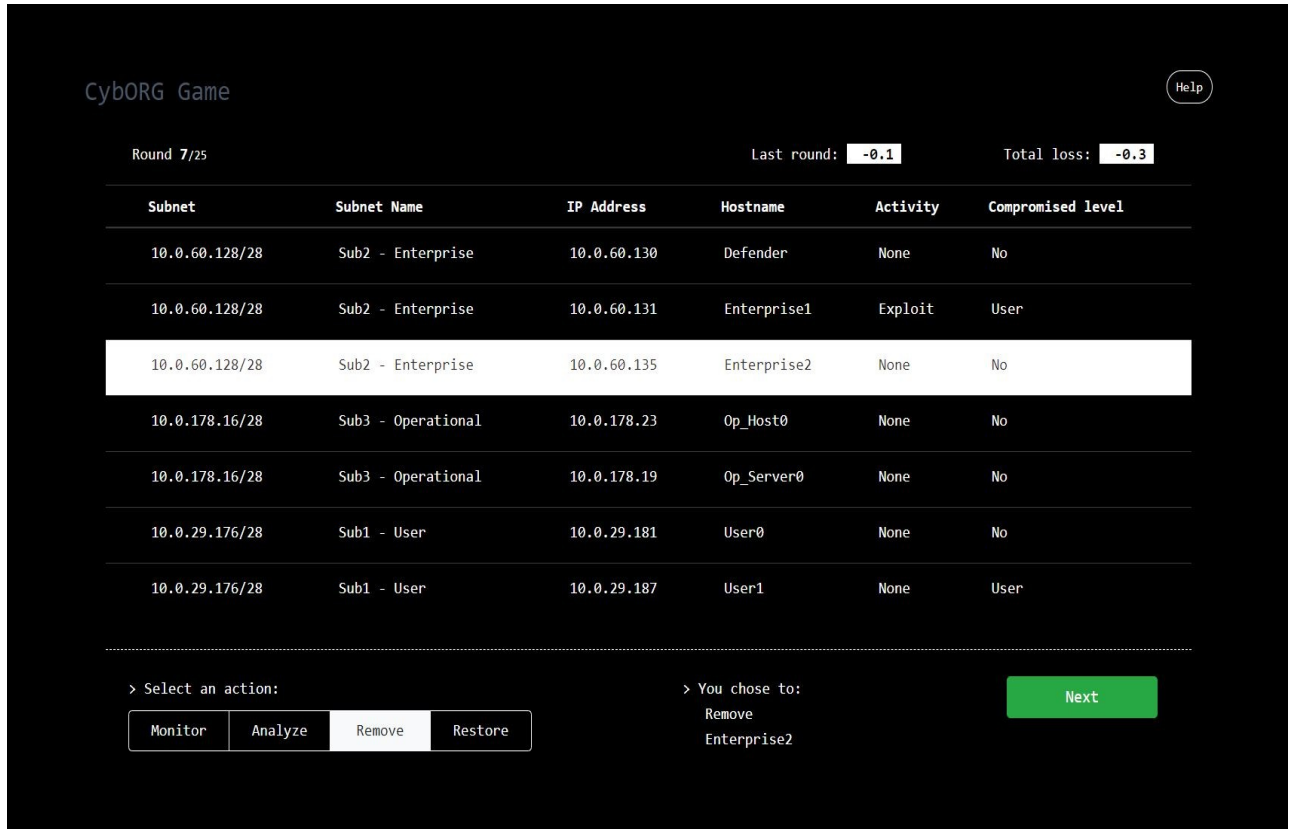


Carnegie Mellon University

Malloy et al. "Accounting for Transfer of Learning using Human Behavior Models" *Under review.*

More Complex Environments: CyBorg





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?