Carnegie Mellon University

Research @ Dynamic Decision Making Laboratory

Carnegie Mellon University
Department of Social & Decision Sciences
http://www.cmu.edu/ddmlab

Human-Machine Collaborations

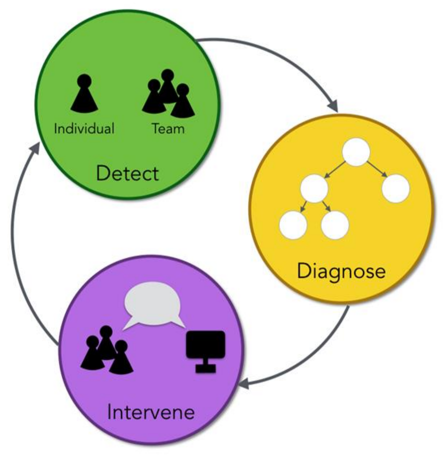
Funding Source: Defense Advanced Research Projects Agency (DARPA) and Air Force Research Laboratory (AFRL).

Internal Collaborators:
- Thuy-Ngoc Nguyen
- Nhat Phan
- Erin Bugbee
- Chase McDonald
- Don Morrison

External Collaborators:
- Anita Woolley, CMU
- Henny Admoni, CMU
- Leslie Blaha, AFRL
- Erin McCormick, AFRL

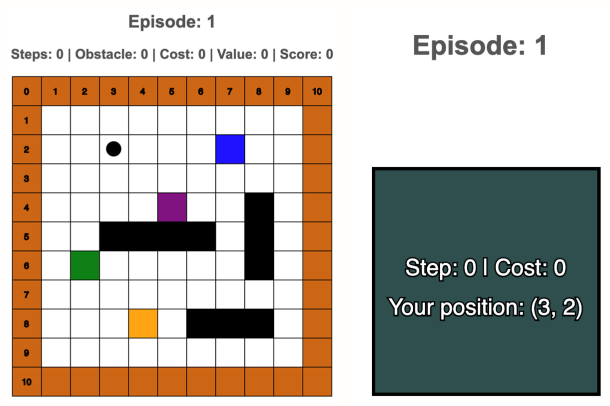
Long-term goal: *Design synthetic coaches that would have Machine Theory of Mind to support team work and enhance team collaboration.*

- Develop a process of coaching in Human-Machine teams.
- Use this coach to perceive individual cognitive states and team social states.
- Understand the role of humans and other agents in the context of the task environment.
- Diagnose team success to design interventions to improve teamwork.



Credit Assignment: Developing Human-like AI Agents

Goal: *To investigate which temporal credit assignment mechanisms can account for behavior under different levels of uncertainty in goal-seeking navigation tasks.*

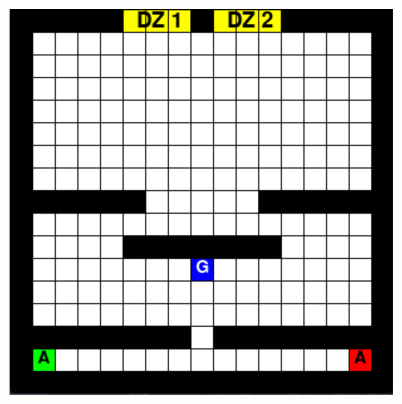


- Provide humans with different levels of uncertainty, represented by different visual representations of the same tasks.
- Develop cognitive models with different credit assignment mechanisms.
- Compare each of the mechanisms with collected human data.

Learning in Cooperative Multiagent Systems Using Cognitive and Machine Models

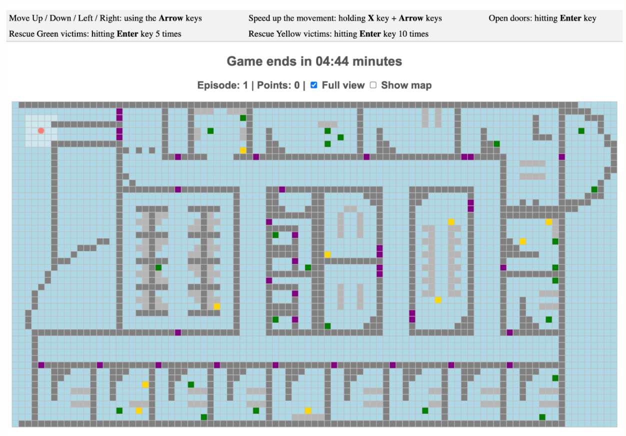
Goal: *To develop cognitive machine models for stochastic scenarios in cooperative multiagent systems (CMS).*

- Introduce three models: Greedy, Hysteretic, and Lenient Multiagent IBL models for CMS
- Conduct experiments on four stochastic scenarios of Coordination Multiagent Object Transportation Problem
- Compare our models with three deep reinforcement learning models



Understanding the Effect of Structural Complexity and Uncertainty on Human Learning

Goal: *To examine the effect of structural complexity and uncertainty on human performance in a simulated search and rescue mission.*

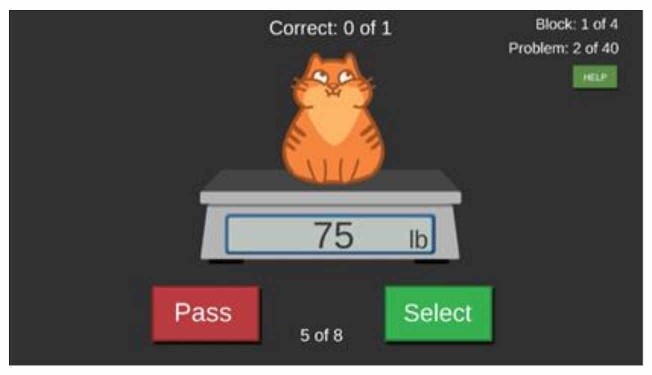


- Develop an interactive simulated search and rescue mission called Minimap.
- Provide human subjects with different degrees of structural complexity coupled with uncertainty in Minimap.
- Analyze different aspects of human behavior in terms of various metrics.

Cognitive Models of Behavior in Sequential Decision Tasks

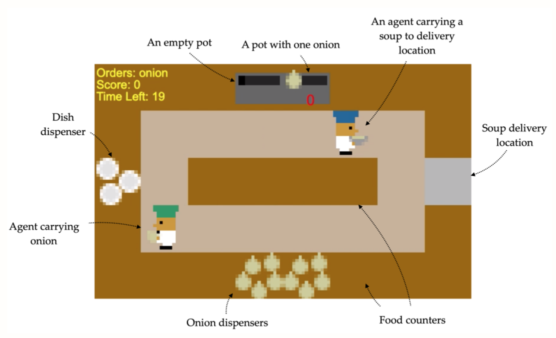
Goal: *To understand how people make sequential decisions in various tasks involving balancing exploration and exploitation, and to develop cognitive models of their behavior in these tasks.*

- Analyze human behavior in experiments in which participants decide when to stop exploring by selecting an option in a sequence of options to maximize reward.
- Develop cognitive models of human behavior in sequential tasks (e.g. Optimal Stopping, Balloon Analog Risk Task), and show generalization from one task to another.



Cognitively Aware Reinforcement Learning

Goal: *To investigate how cognitive models can be used in tandem with reinforcement learning (RL) agents to learn policies that complement human behavior.*



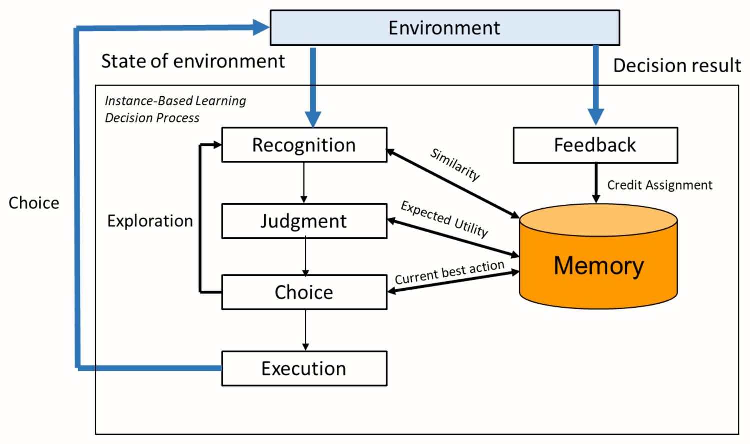
- We incorporate cognitive models into the RL training and testing pipelines to see how such models can improve performance in cooperative tasks.
- We test these models with human proxies and real humans, and analyze their behavior using collaborative fluency metrics, to see how well they learn collaborative policies.

General Research Goals

Our research aims to understand learning and decisions from experience in dynamic decision environments. Our work relies on a theory of learning from experience called ***Instance-Based Learning Theory (IBLT)*** and on other theories and ideas, mostly from cognitive psychology.

In our research, we address questions such as:

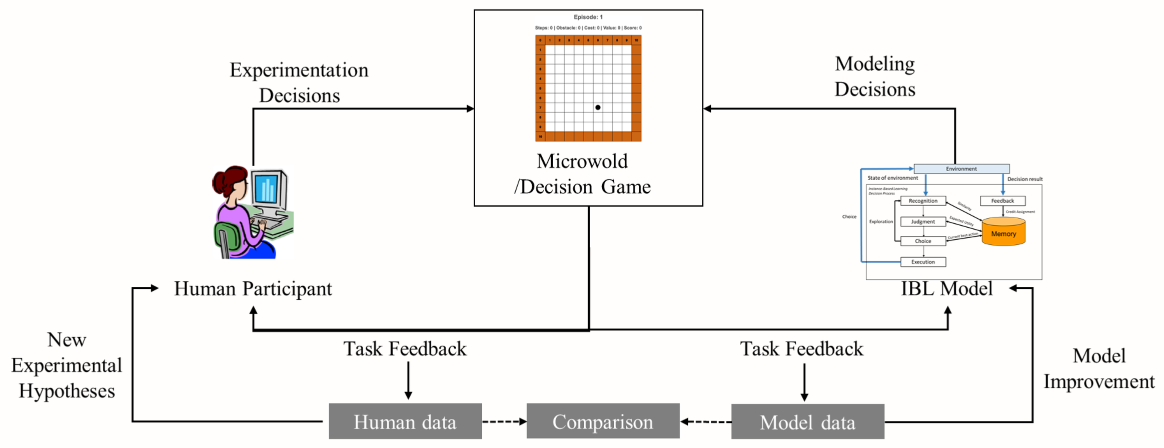
- How does experience influence our decisions?
- What kinds of and how much experience produces better performance and better adaptation to novel environments?
- How does experience transfer to new situations?



General Methods

Our research approach includes laboratory experiments and cognitive models, which form a learning cycle that compares human data from experiments against theory-informed data from computational models.

- We collect behavioral data using complex, dynamic simulations.
- Our experiments often involve extended practice to help us understand how experience develops, changes, and transfers to new situations.



- We create cognitive models that rely on IBLT and the mechanisms proposed in the ACT-R cognitive architecture to represent and predict human behavior in decision making tasks.
- Our theory and methods are applied to many domains that deal with the prediction of behavior in complex systems.

Computational Models

The DDMLab is constantly upgrading its systems for computational modeling. A fantastic new addition is SpeedyIBL, a Python library that allows to create single or multiple IBL agents with fast processing and response time without compromising the performance.

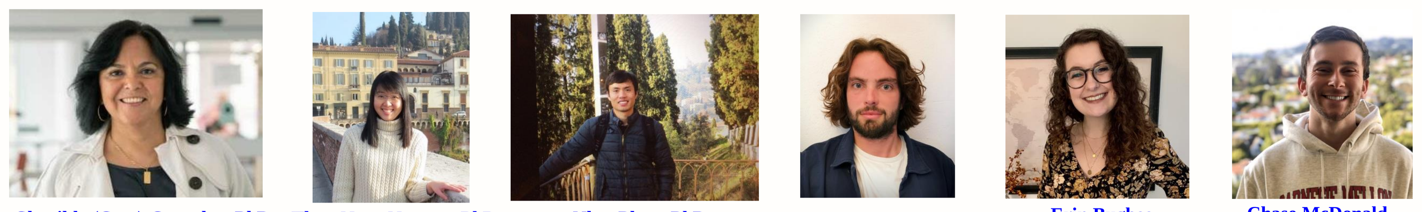
SpeedyIBL

- SpeedyIBL utilizes fast processing and response time without compromising performance compared to the traditional implementations.
- SpeedyIBL can be used to create IBL agents that can do a wide range of decision games such as Binary Choice, Insider Attack, Minimap, Ms. Pac-man, Fireman, and Cooperative Navigation tasks.

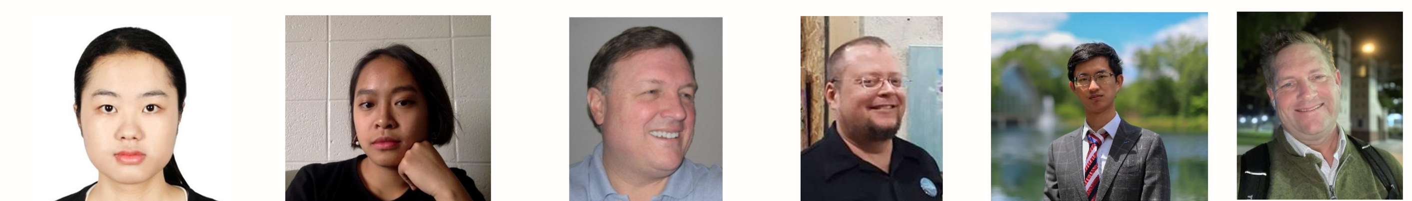
Current Lab Members

The DDMLab is a group **fully funded by grants** from research institutions such as National Science Foundation, Army Research Labs, Army Research Office, Defense Threat Reduction Agency, and others.

Our small but productive group is comprised of researchers from different fields, including Behavioral Decision Research, Psychology, Engineering, and Computer Science.



Cleotilde (Coty) Gonzalez, PhD Research Professor, Director of the DDMLab | **Thuy-Ngoc Nguyen, PhD** Post-Doctoral Fellow | **Nhat Phan, PhD** Post-Doctoral Fellow | **Baptiste Prebot, PhD** Post-Doctoral Fellow | **Erin Bugbee** PhD Student | **Chase McDonald** PhD Student



Yinuo Du PhD Student | **Jiajia Hu** PhD Student | **Don Morrison** Senior Research Designer | **Jeffrey Flagg, MS** Research Associate II | **Tony Xi** Undergraduate Research Assistant | **Ed Matlack** Visiting Scholar

Contact Us

Dynamic Decision Making Laboratory
Department of Social and Decision Sciences
Porter Hall 223-G
Pittsburgh, PA 15213

Phone: (412) 268-6242
ddmlab@andrew.cmu.edu
http://www.cmu.edu/ddmlab

Behavioral Cybersecurity

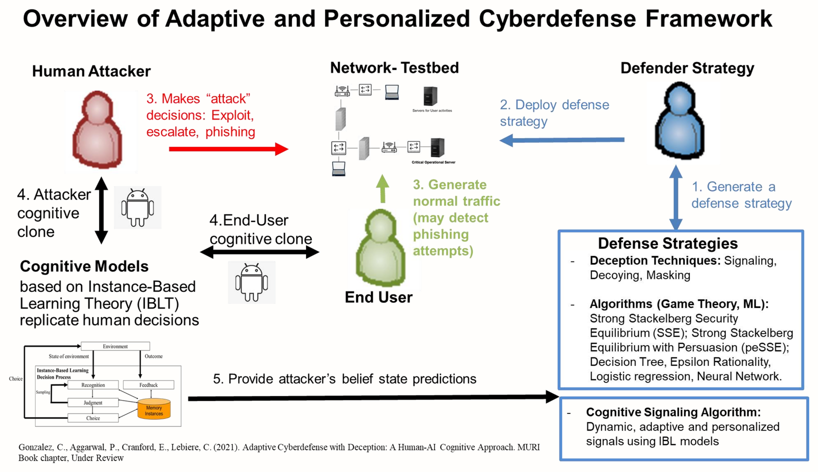
Funding Source: Army Research Laboratories - Collaborative Research Alliance (ARL-CRA) and Army Research Office - Multi University Research Initiative (ARO-MURI) on cyber deception and MURI-AUS on Human-Bot cyber defense teams.

Internal Collaborators:
- Baptiste Prebot
- Yinuo Du
- Tony Xi
- Ed Madlack

External Collaborators:
- Edward Cranford, CMU
- Christian Lebiere, CMU
- Palvi Aggarwal, UTEP
- Kuldeep Singh, UTEP
- Milind Tambe, Harvard
- Christopher Kiekintveld, UTEP
- Prasant Mahopatra, UC Davis
- Prashanth Rajivan, U of Wash

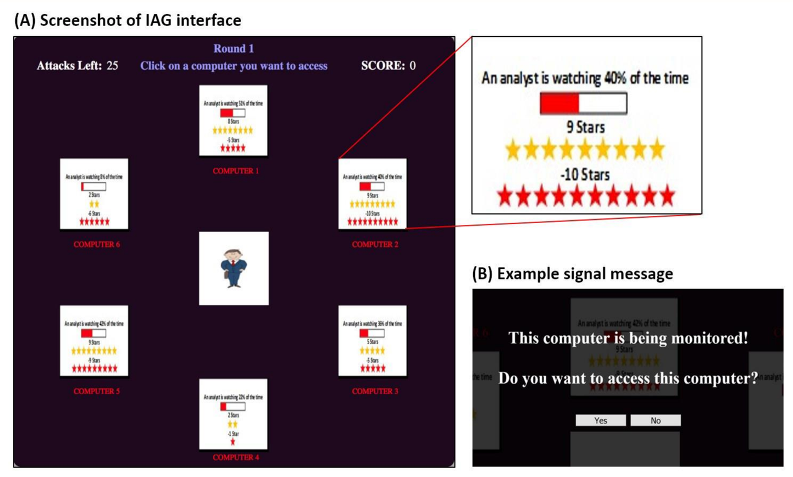
Long-term goal: *To design effective personal, dynamic, and adaptive defense techniques informed directly by dynamics of human behavior, emergent cognitive biases, and psychological deception strategies.*

- Design defense algorithms using Stackelberg Security Games (SSG) and signaling theory.
- Design the task using different experimental games.
- Human attackers interact with different experimental games.
- Develop cognitive model of attackers to learn their behavior.
- Adapt the SSG defense algorithms.



Deception Through Signaling and Masking

Goal: *Design dynamic and personalized deception strategies using cognitively-informed algorithms for defense.*

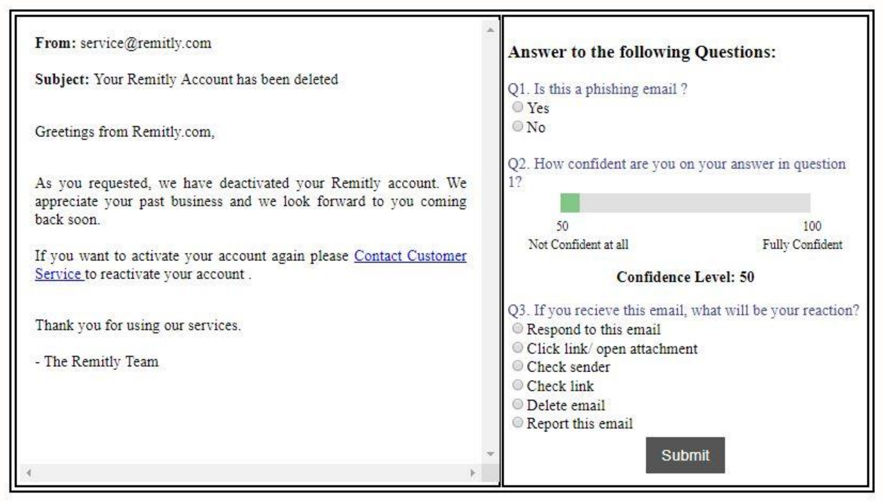


- Defenders strategically reveal information to the attackers to influence their decisions.
- Defenders can use a combination of truthful and deceptive signals to protect unprotected resources.
- Defenders can also use masking strategies to manipulate features of real machines.
- Cognitive algorithms learn the attacker's behavior and inform game theoretic models to adapt the defense.

Understanding the Learning of End-Users in Phishing Training

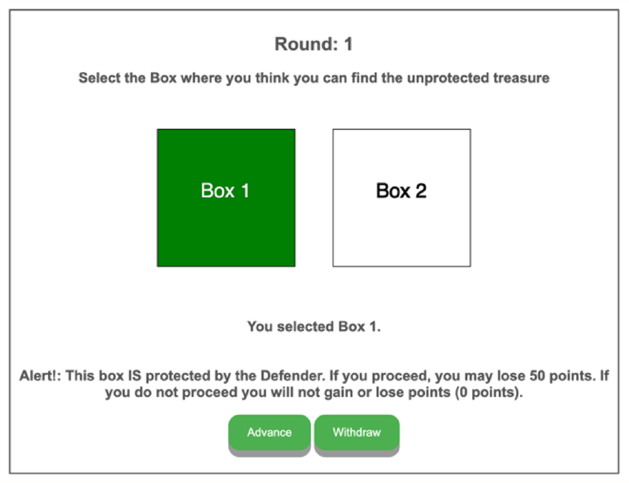
Goal: *To determine the effect of cognitive factors on the detection of phishing emails through experiential learning.*

- Train end-users with different frequency, recency and content of phishing emails.
- Provide different kinds of feedback during training.
- Test their detection capabilities after training.
- Develop cognitive models of end-users to predict their actions ahead of time.



Defense Strategies in a Repeated Binary Choice Task

Goal: *To design defense strategies to influence human choices in the box game, a repeated binary choice task.*

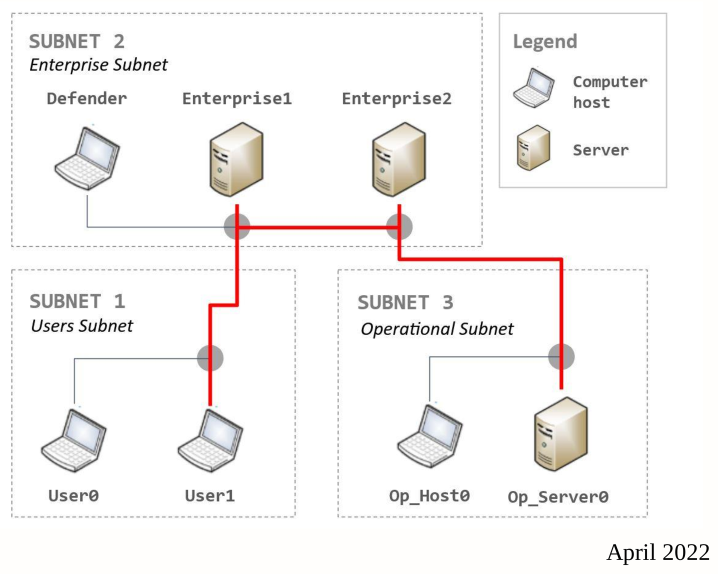


- Attackers repeatedly attempt to find a treasure in one of two boxes.
- Defenders provide a potentially deceptive signal about the protection of the chosen box.
- Attackers decide whether to advance or withdraw, then observe the outcome.
- Defenders use defense strategies informed by cognitive algorithms, in which the attacker's behavior is used to adjust the strategy dynamically.

Towards Human-AI Collaboration in Autonomous Cyber Operations

Goal: *To study the integration of IBL models for improving Trust and Mental Models sharing in Human-Autonomy teams for cybersecurity.*

- Develop a framework for Human-AI collaboration research for cyber-defense.
- Develop Cognitive models of Human defenders to predict their decisions.
- Test Human defenders in an Interactive Defense Game.
- Test Human-AI (ML & IBL) collaboration in similar scenario.



April 2022