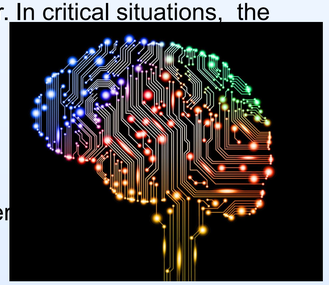
Eliciting an Algorithm to Replicate Human Trust in Automation in the Domain of Compliance

Jayson Boubin, Maj. Christina Rusnock, PhD, Dr. Michael Miller The Air Force Institute of Technology, Department of Systems Engineering and Management

Introduction

As the use of technology rises across the world, the value of automation becomes more apparent. Tasks that do not require human decision making, or are too complicated or esoteric for the average user to understand, may be automated. The human and automation become two parts of a team. With adaptive automation, the human performs their tasks, and if certain conditions are met, the automation will trigger. In critical situations, the human must trust that, when triggered, the automation will perform its task correctly. This type of trust is called compliance. If the user Is 100% compliant, they always assume the automation has correctly triggered, and thus accept and respond to the automation. If they are 100% noncompliant, they always assume the automation has falsely triggered, and thus ignore the automation's actions. We believe that compliance with automation could change with taskload. To fully understand the change of compliance based on taskload, we have used a game entitled Space Navigator, which is equipped with multiple types of automated agents. We measured the compliance with said automated agents as taskload increases.



Questions

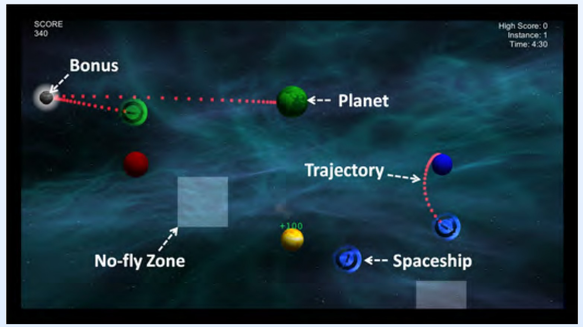
How can we measure compliance in the domain of Space Navigator?

How does compliance change with taskload?

Can we create an algorithm which replicates Human Compliance when teamed with Space Navigator's automated agents?

Experiment Design

Human trials were performed on Space Navigator to allow us to fully understand taskload's contribution to compliance with various types of automated agents. Each type of automated agent would wait a set amount of time, and if a ship had not received a route, the ship would be routed by one of three algorithms. The types of algorithms included:

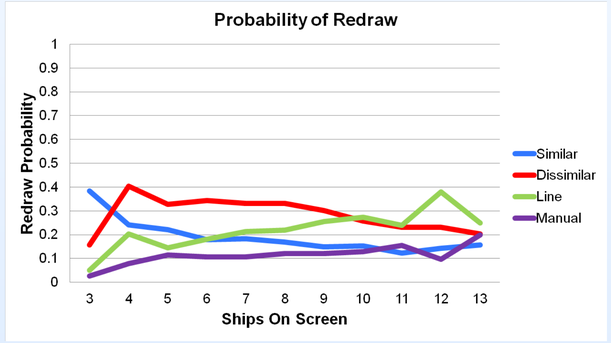


- A straight line drawn to the planet
- A line drawn by the nearest neighbor algorithm
- A line drawn in a random direction

Taskload is defined in this situation as the number of ships on the screen at any given time. We defined compliance as the rate at which a player redraws a route for a ship that has already been routed, either by themselves or by the automation. With 100% compliance meaning they never redraw, and zero compliance meaning all routes are redrawn. We analyzed the data, looking for the percentage of ships that are redrawn after given a route for each taskload level. We then compared these rates for the three automation types against the user's baseline rate when in manual mode

Results

Through our analysis, we were able to find probabilities for the redrawing of ships based on the amount of ships on the screen. We were able to reduce these probabilities to functions which, if given the number of ships on screen and the automation type, provides an empirical percentage of compliance.



Explaining 70-90% of all ships redraw behavior, our elicited functions provide realistic redraw probabilities for each automated agent type for any taskload up to 14 ships on screen.

Regression: Redraws as a factor of Ships On Screen				
Agent	**Similar**	**Dissimilar**	**Line**	**Manual**
P-Value	**0.00104**	**0.008326**	**0.007145**	**0.01918**

We used linear regression to analyze the significance of the trends using redraws as a factor of ships on screen. Based on this analysis it is evident that compliance changes over time with taskload at a significant rate for all agents, human and automated.

Functions and R^2 values

	Functions	R^2 values
Similar	**R = -0.0136T + 0.2224**	**0.9027**
Dissimilar	**R = -0.0176T + 0.374**	**0.7808**
Line	**R = 0.018T + 0.1464**	**0.7934**
Manual	**R = 0.0065T + 0.0961**	**0.698**

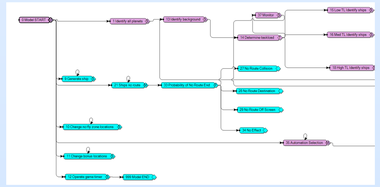
These functions were obtained by performing a least squares regression on the portion of the data where the most ships were concentrated (5-11 ships on screen). R corresponds to the per cent chance of a redraw occurring with that agent type, T corresponds to the taskload, meaning the number of ships on screen at that time.

Conclusion

Our analysis has established an algorithm that quantifies compliance-based trust.
The data that provided the algorithm has some interesting trends. Compliance with Similar and Dissimilar automated agents increase, as expected, with a rise in taskload. This could be due to the user being overloaded and having no choice but to comply, or that increased workload makes these automated agents strategies more viable.
Compliance with straight line automated agents and draws made by the user decreases with increased taskload. This may be due to a decrease in the viability of straight line automation as a strategy and the users own ability to accurately draw effective routes in periods of high taskload /activity.
The results show that compliance is dependent both on the automated agents behavior and on the taskload experienced by the user. This knowledge can be used to create automation which considers taskload to implement appropriate behavior and strategies.

Future Work

Our algorithm will be implemented in a Discrete Event Simulation of Space Navigator. This verified algorithm will allow us to create a human-performance simulation which properly replicates and predicts the trust users experience throughout changes in taskload.



We will create a baseline simulation for user trust and will be able to observe changes in performance when applying different trust levels.
We will also create algorithms that reflect the change in trust over time as the users familiarize themselves with the automation types. These algorithms will also allow us to model the learning process of trust. These techniques can also be used to study compliance on automation and its change with taskload and strategy.

Acknowledgements

The authors gratefully acknowledge the support of the Air Force Office of Scientific Research, Computational Cognition and Machine Intelligence Program, which partially funded this research.