Measuring State Utilization During Decision Making in Human-Robot Teams

Saurav Singh (ss3337@rit.edu), Jamison Heard (jrhee@rit.edu)

Rochester Institute of Technology, Rochester, New York, USA

INTRODUCTION

In human robot teams, robots need to adapt to human external states (such as position and velocity) and internal states (such as workload and fatigue)

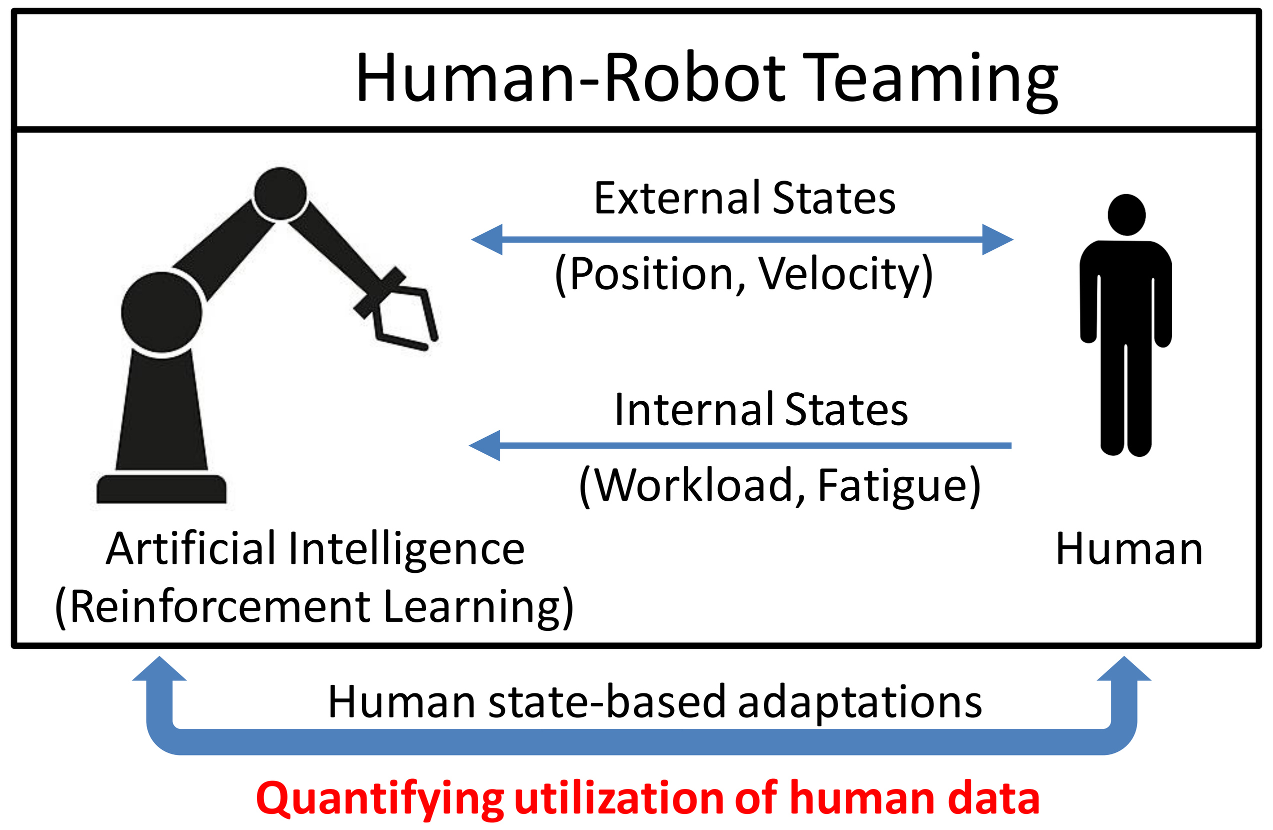


Figure 1: State utilization to quantify utilization of human data by a robot

Problem: How much are adaptive black-box algorithms such as reinforcement learning algorithms relying on human data?

Proposed solution: New state utilization metric that quantifies RL agent's reliance on state features, including human data in human robot teaming scenarios.

Tests:

- Cartpole by OpenAI Gym (Toy Problem)
- NASA MATB-II (Human Subjects Study)

Observed Benefit: Metric indeed increases explainability in multimodal neural networks. Significant potential impact in multimodal data fusion across various application

METHOD

- Observe the difference between output of the decision network with original dataset and permuted dataset.
- A smaller change implies the state feature S_{ni} have low utilization and vice versa.

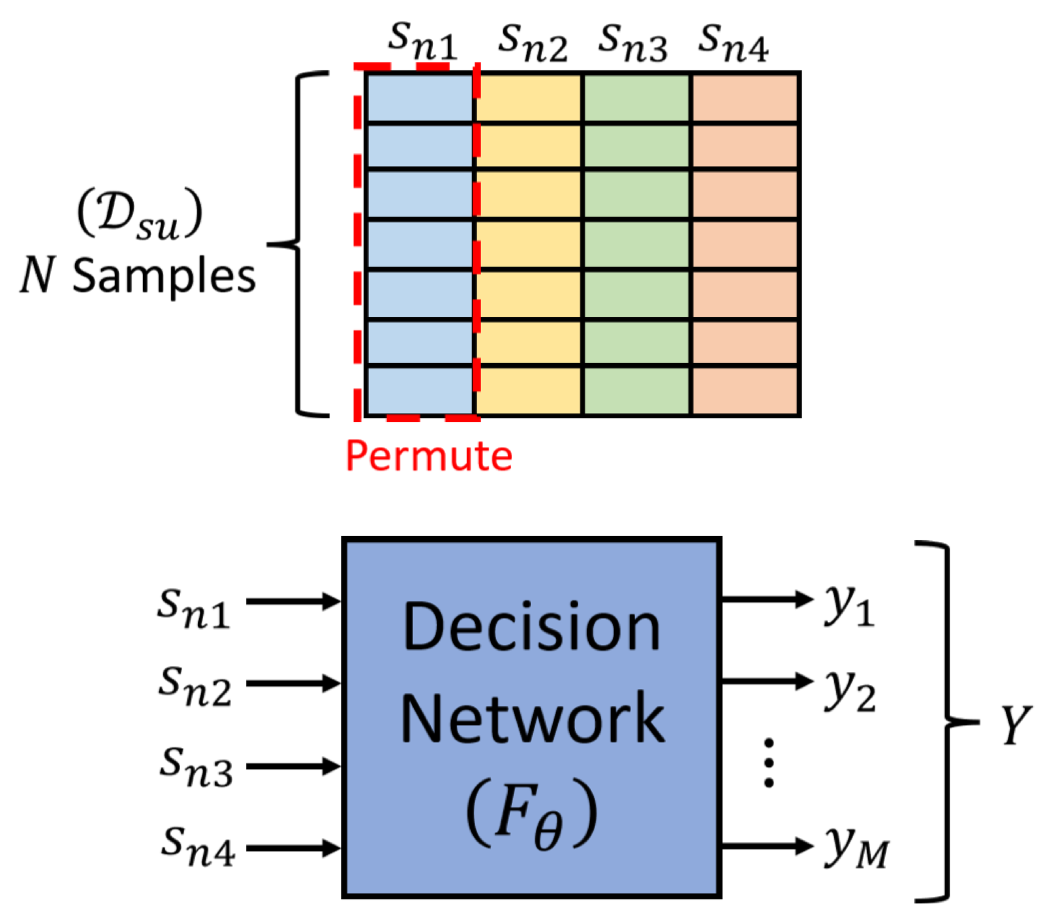
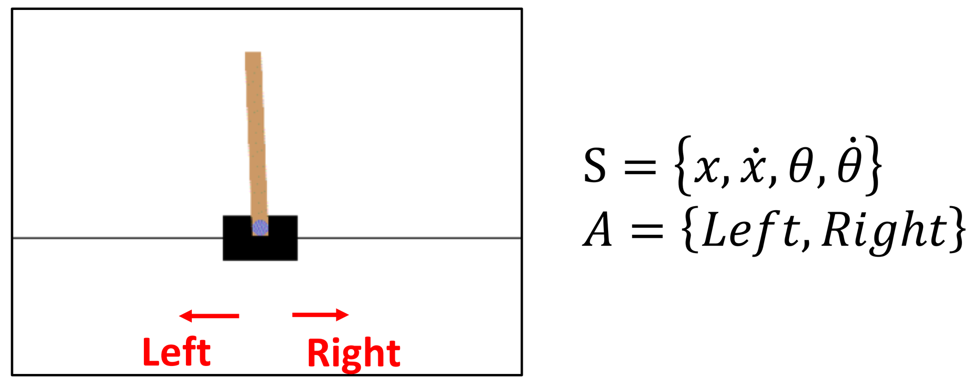


Figure 2: Permuting/shuffling samples of a state feature S_{ni} in the replay buffer to break the association between the state feature S_{ni} and the output Q

Algorithm 1: State Utilization for RL

Initialize the RL decision network F_θ, learned model parameters θ, replay memory $\mathcal{D}_{replay}$;
Sample a batch of data $\mathcal{D}_{su}$ from replay memory $\mathcal{D}_{replay}$;
Compute decision network output Y, Eq. 2;
for *each state feature s_{ni}* **do**
  Randomly permute the samples of state feature s_{ni} while keeping the state features s_{nj}, $j \neq i$ unchanged;
  Compute decision network output Y_i with permuted state feature s_{ni}, Eq. 3;
end
for *each state feature s_i* **do**
  Compute State Utilization (SU_i) using $SU_i = \frac{\|Y_i - Y\|}{\sum_{j=1}^{S} \|Y_j - Y\|}$, Eq. 4;
end

ABLATION STUDIES ON CARTPOLE



$S = \{x, \dot{x}, \theta, \dot{\theta}\}$
$A = \{Left, Right\}$

Figure 3: Cartpole Environment by OpenAI Gym

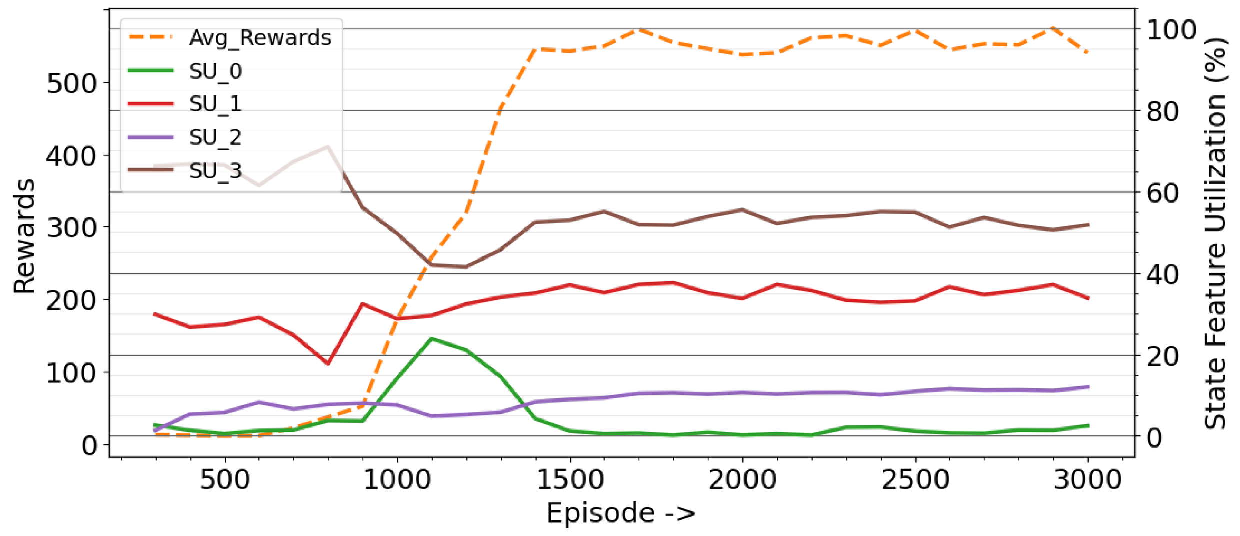


Figure 4: Average Rewards and State Utilization (SU) across episodes for the Cartpole environment solved by a DDQN agent with $S = \{x, \dot{x}, \theta, \dot{\theta}\}$.

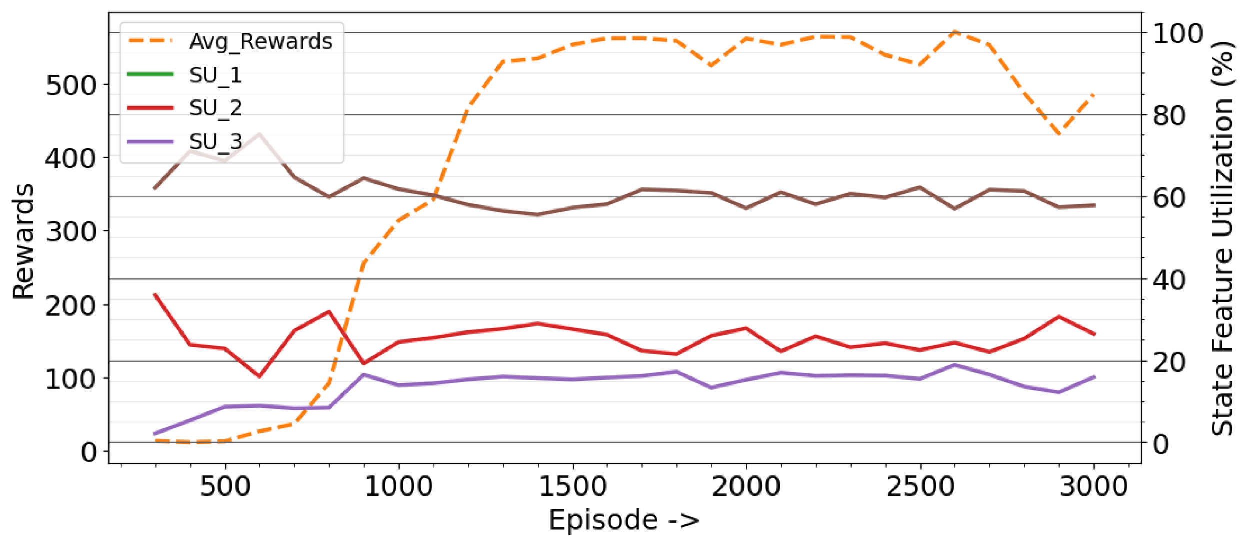


Figure 5: Average Rewards and State Utilization (SU) across episodes for the Cartpole environment without S_{n0} i.e., $S = \{\cancel{x}, \dot{x}, \theta, \dot{\theta}\}$.

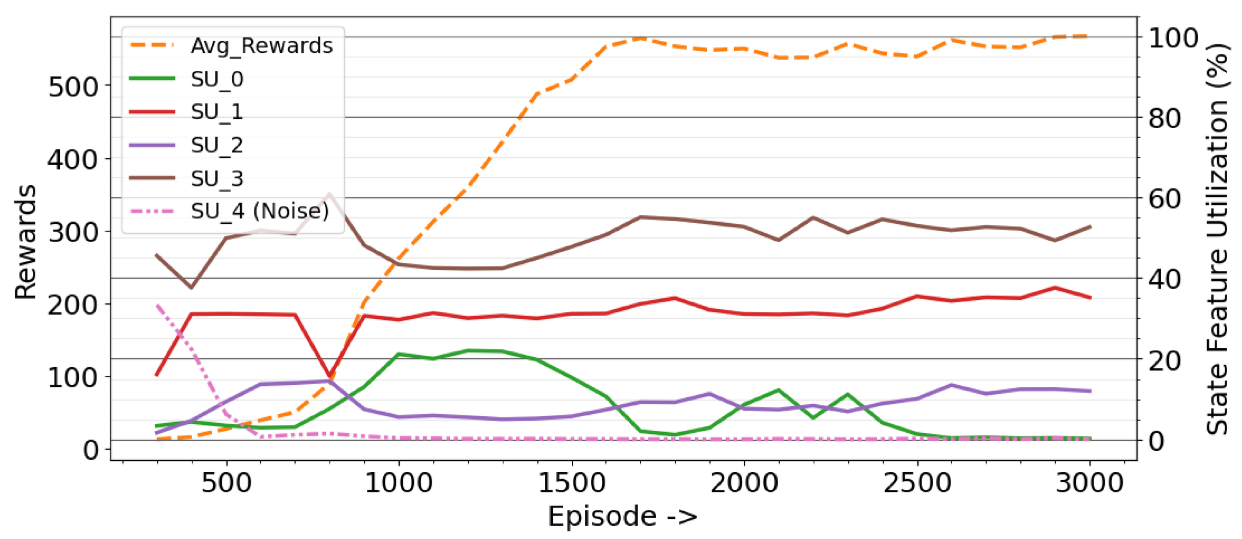


Figure 6: Average Rewards and State Utilization (SU) across episodes for the Cartpole environment with a random noise state feature (S_{n4}), i.e., $S = \{x, \dot{x}, \theta, \dot{\theta}, Noise\}$.

MEASURING UTILIZATION OF HUMAN DATA IN HUMAN ROBOT TEAMING

- SU computed in prior human-robot studies [3][4] showed the RL agent's dependence on human input.

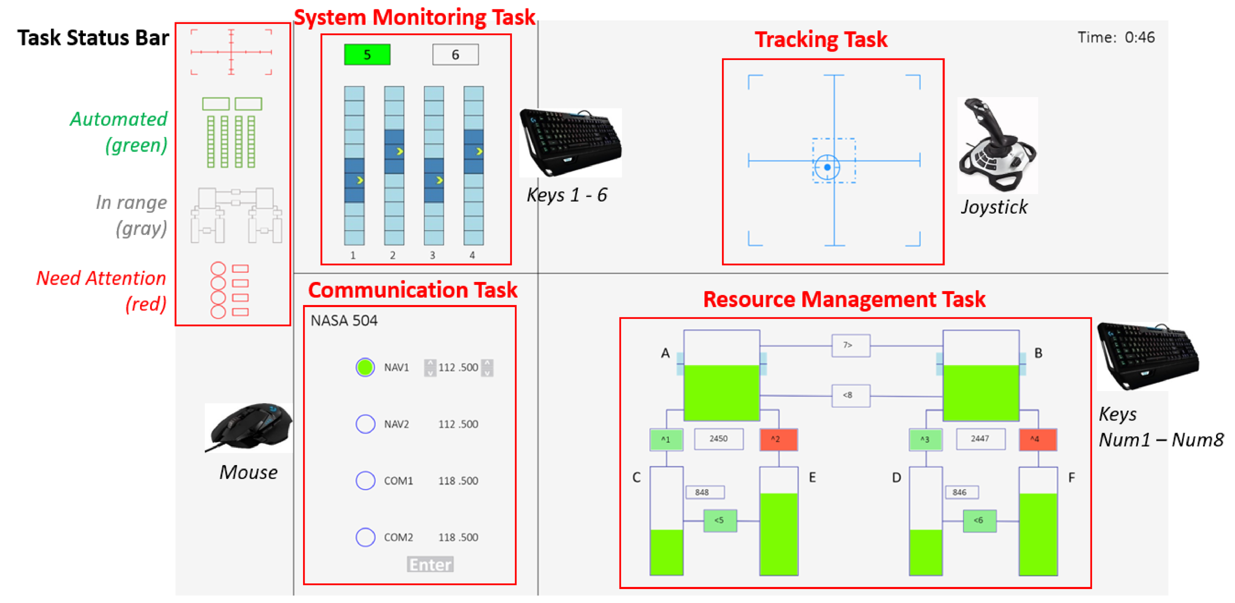


Figure 7: The NASA Multi-Attribute Task Battery-II Env.

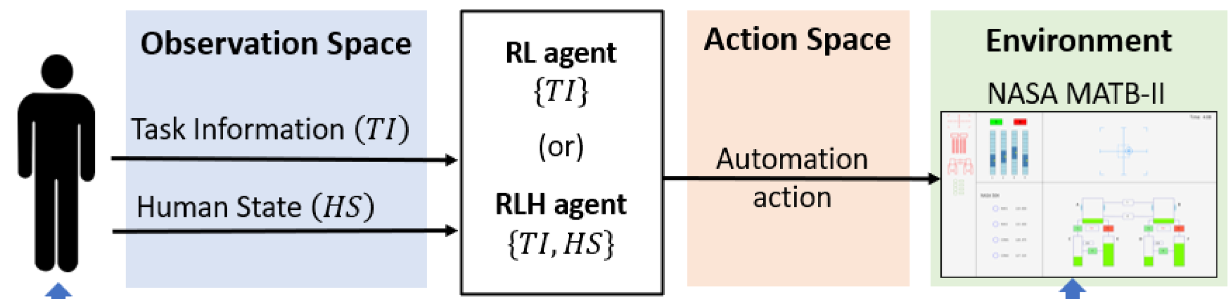


Figure 3: Adaptive Human-Robot Teaming architecture with human state estimates augmented to RL's observation space.

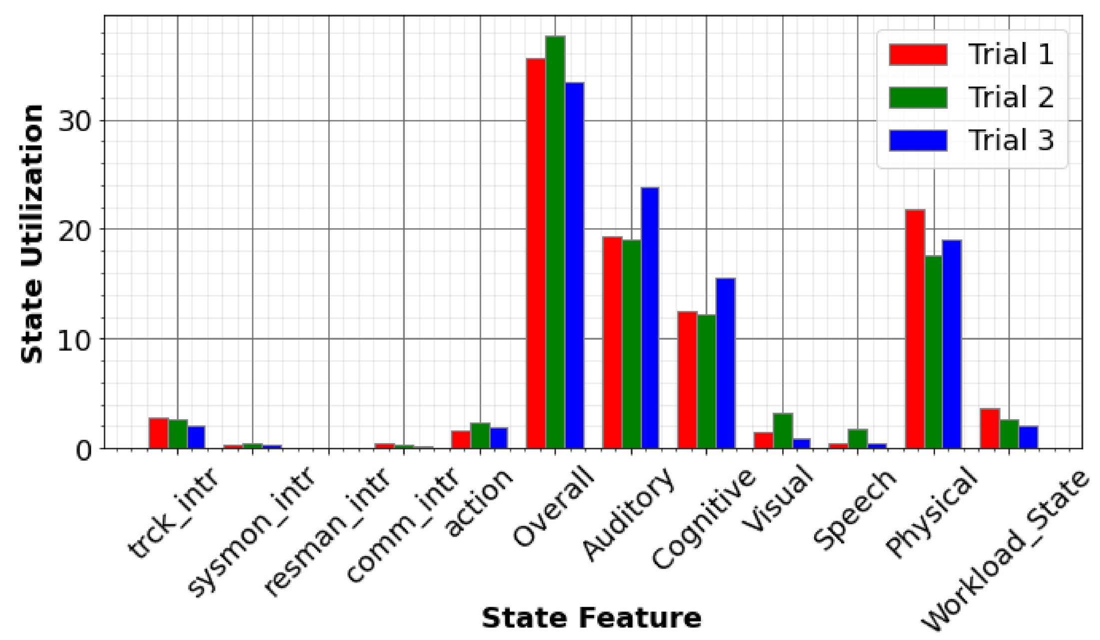


Figure 8: State Utilization for the RLH agent trained on task interaction and human internal states data in the NASA MATB-II experiment.

CONTACT INFORMATION

Saurav Singh
ss3337@rit.edu

Jamison Heard
jrheee@rit.edu

[1] A. Fisher, C. Rudin, and F. Dominici, "All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously," *Journal of Machine Learning Research: JMLR*, vol. 20, 2019.
[2] S. Singh, P. P. Markopoulos, E. Saber, J. D. Lew and J. Heard, "Measuring Modality Utilization in Multi-Modal Neural Networks," *2023 IEEE Conference on Artificial Intelligence (CAI)*, Santa Clara, CA, USA, 2023, pp. 11-14.
[3] S. Singh and J. Heard, "A Human-Aware Decision Making System for Human-Robot Teams," 2022 17th Annual System of Systems Engineering Conference (SOSE), 2022, pp. 268-273.
[4] S. Singh and J. Heard, "Human-aware reinforcement learning for adaptive human robot teaming," in Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction, ser. HRI '22. IEEE Press, 2022, p. 1049–1052.