Inferring Knowledge from Behavior in Search-and-rescue Tasks

Scott Cheng-Hsin Yang, Sean Anderson, Pei Wang, Chirag Rank, Tomas Folke, and Patrick Shafto
Department of Mathematics & Computer Science, Rutgers University–Newark

RUTGERS UNIVERSITY | NEWARK

shaftolab.com

Introduction

Theory-of-Mind inference is natural for humans but poses significant computational challenges. This inference can be cast as a problem of inverse planning. The core difficulty then is the exponential growth of paths to consider in planning given a mental state. We tackle this problem in a search-and-rescue task implemented in Minecraft [1]. **Our goal is to infer differences in knowledge from participants' continuous-time trajectories?** By abstracting the spatiotemporal state space and reward function together, we elicit natural decision points, on which we compare the participants' behavior to myopic rational agents [2] of varying knowledgeability. Collectively, abstraction and rational agent analysis yield successful inference of participants' knowledge states and reveal distinct patterns of their exploratory behavior.

Experiment

Task: Each participant engaged in a search-and-rescue task in a simulated environment (Minecraft MALMO). The environment was a collapsed office building consisting of rooms and corridors. Participants were tasked with saving victims located inside some of the rooms (never in a corridor). Victims varied in the severity of their injuries. Triaging yellow (green) victims took 15 (7.5) seconds and added 30 (10) points. The goal was to accumulate as many points as possible in 10 munutes.

References & Acknowledgments

[1] Huang, L., Freeman, J., Cooke, N., Cohen, M., et al. (2021). Using humans' theory of mind to study artificial social intelligence in minecraft search and rescue. Technical report.

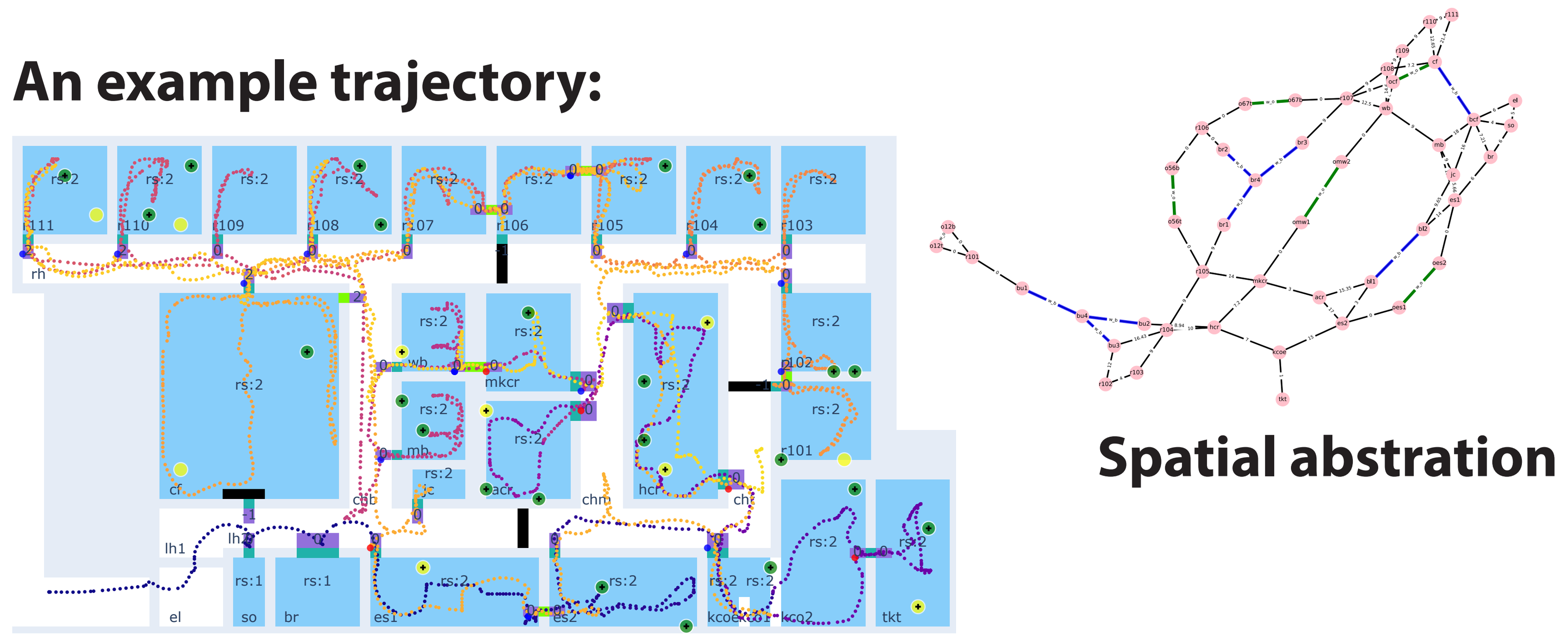
[2] Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. Nature Human Behaviour, 1(4), 1–10.

This project was supported by DARPA grant HR00112020039.

Experiement (cont.)

Knowledge condition (signal vs no signal): All participants had a rescue device that indicated the presence of victims in front of the entrance of a room. The device beeped twice if there was at least one yellow victim inside the room, beeped once if there was no yellow but at least one green victim, and remained silent if there was no victim. Participants in the *signal* condition were told the meaning of the beeps, while those in the *no-signal* condition were not. There were 36 participants in the no-signal condition and 18 in the signal condition.

An example trajectory:



Spatial abstraction

Abstraction: The reward distribution naturally led to a *spatial abstraction* which lifts the original Euclidean state space to a graph where the nodes represent rooms and the edges represent the paths connecting rooms. The room-based topology inspired a new *room-based utility* function, as opposed to a typical reward function that assigns positive values to known victim locations and a small negative value to all other spatial states. The room-based utility inspired a further *temporal abstraction* that replaces continuous time by time of room exists. These *decision points* marked the times when a decision had to be made.

Model

Let K denote the knowledge condition (signal or no signal). The posterior probability that a participant with trajectory S being in knowledge condition K is given by Bayes' Rule:

$$P(K|S) = \frac{P(S|K)P(K)}{\sum_{K'} P(S|K')P(K')},$$

The prior P(K) is set to be the (empirical) prior probability of a participant being in the different knowledge conditions. The likelihood P(S|K) is given by:

$$P(S|K) = \prod_t P(s_t|K) = \prod_t \frac{e^{\beta \cdot U^K_{i=s_t}}}{\sum_j e^{\beta \cdot U^K_j}},$$

The t indexes the decision points, which are the room exists. The s_t is the room selected to be visited next at decision point t. The β (set to 5) is an exploration parameter which trades off exploiting the choice with highest utility against making a random choice.

The expected utility U for entering room i under knowledge condition K is given by:

$$U^K_i = \sum_{\{n_y, n_g\}} P^K_i(n_y, n_g | I) R(n_y, n_g, t_i),$$

$P^K_i(n_y, n_g|I)$ is knowledge-state-dependent probability of room i containing n_y yellow victims and n_g green victims given all information I gathered thus far. $R(n_y, n_g, t_i)$ is the utility (points per time) of a room if that room has n_y and n_g green victims and takes time t_i to travel to.

Model summary: The inference of knowledge condition (K) is based on how similar a participant's choices of room visits (S) are to those of the rational agent in the signal condition, relative to those in the no-signal condition.

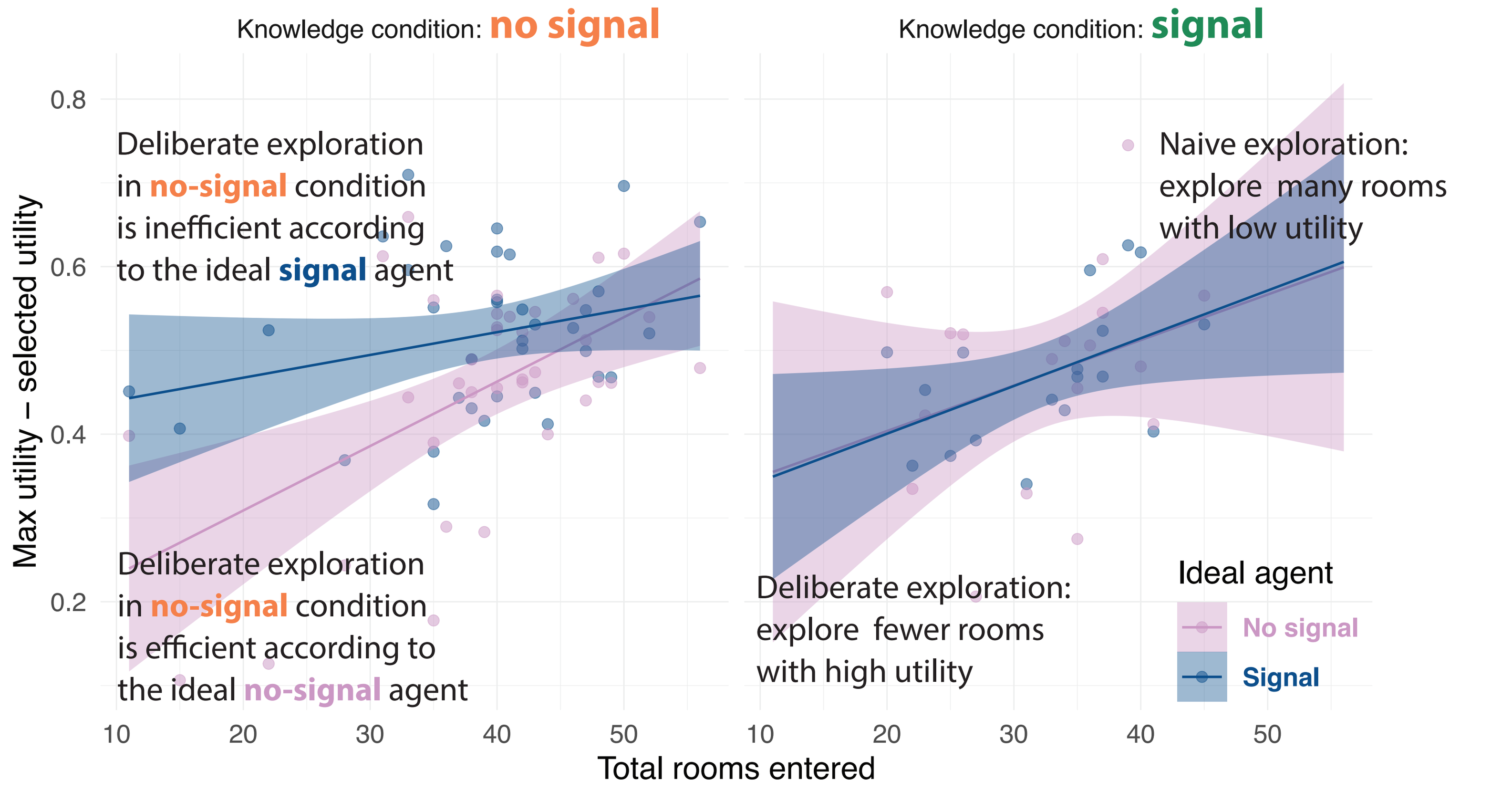
Results

Main question: Can we infer the assigned knowledge condition from the participant's continuous-time trajectory? **Yes!**

Truth \ Inference	No signal	Signal
No signal (36)	0.611 (22)	0.389 (14)
Signal (18)	0.278 (5)	0.722 (13)

This is a confusion matrix from maximum-a-posteriori estimates. Rows indicate ground-truth conditions, and columns inferred conditions. Numbers in parentheses indicate the number of participants in the conditions. Fisher's exact test suggests that the classification accuracy of the model exceeds what would be expected by chance (OR=3.98, p=.04).

More analysis: Analysis shows distinct exploratory behaviors.



Conclusions: We successfully inferred knowledge state from continuous trajectories. We did so by: (1) devicing an abstraction over spatiotemporal states and the reward; (2) surfacing natural decision points; (3) building rational agents with differing knowledge; and (4) conducting rational agent analyses. The analyses also highlight distinct exploratory behaviors.