Using Machine Theory of Mind to Learn Agent Social Network Structures from Observed Interactive Behaviors with Targets

Yun-Shiuan Chuang, Hsin-Yi Hung, Edwinn Gamborino, Joshua Oon Soo Goh, Tsung-Ren Huang, Yu-Ling Chang, Su-Ling Yeh, Li-Chen Fu, *Member, IEEE*

*Abstract*— Human social interactions are laden with behavioral preferences that stem from hidden social network representations. In this study, we trained a neural network (NN) to learn and predict human social preferences based on implicit information from the way agents and social targets behaviorally interact with each other. Our findings have implications for machine applications that seek to infer hidden information structures solely from third-person observation of behaviors.

# INTRODUCTION

# The use of artificial intelligent machines that dynamically interact with people is proliferating across many aspects of human life such as in interactive virtual agent service platforms [1], [2] and socially assistive robots [3]. Nevertheless, the efficacy of such social machines is limited by the naturalness of their interactions with users. Specifically, human-machine interactions are typically hampered because the interactive actions engaged by machines are often contextually aberrant and do not fit human social behavioral norms. Thus, learning algorithms that help social machines display more human-like contextually relevant interactive behaviors should enhance their intended functionality.

\*This research was supported in part by the Ministry of Science and Technology of Taiwan (MOST 107-2634-F-002-018), National Taiwan University, Center for Artificial Intelligence & Advanced Robotics, and Joint Research Center for AI Technology and All Vista Healthcare.

Y.-S. Chuang is with the Graduate Institute of Brain and Mind Sciences, National Taiwan University, Taipei, Taiwan (e-mail: yunshiuan.chuang@gmail.com).

H.-Y. Hung is with the Graduate Institute of Brain and Mind Sciences and the Center for Artificial Intelligence and Advanced Robotics, National Taiwan University, Taipei, Taiwan (e-mail: r05454001@ntu.edu.tw).

E. Gamborino is with the Center for Artificial Intelligence and Advanced Robotics, National Taiwan University, Taipei, Taiwan. (phone: +886 958 376 105; e-mail: gamborino@ntu.edu.tw).

J.O.S. Goh is with the Graduate Institute of Brain and Mind Sciences, Department of Psychology, Neurobiology and Cognitive Science Center, and the Center for Artificial Intelligence and Advanced Robotics, National Taiwan University, Taipei, Taiwan. (phone: +886 2 2312 3456 ext. 88022; e-mail: joshuagoh@ntu.edu.tw).

T.-R. Huang and Y.-L. Chang are with the Department of Psychology and the Center for Artificial Intelligence and Advanced Robotics, National Taiwan University, Taipei, Taiwan. (e-mail: [ychang, trhuang] @g.ntu.edu.tw).

S.-L. Yeh are with the Department of Psychology, the Graduate Institute of Brain and Mind Sciences and the Center for Artificial Intelligence and Advanced Robotics, National Taiwan University, Taipei, Taiwan. (e-mail: suling@g.ntu.edu.tw).

L.-C. Fu is with the Center for Artificial Intelligence and Advanced Robotics, the Department of Electrical Engineering and the Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan (e-mail: lichen@ntu.edu.tw).

# One characteristic driving fluent human social interactions is access to information about the underlying social network of persons involved, which is often implicit [4], [5]. For example, suggesting a friend call Bill Gates about a lunch appointment would be absurd unless one knew that he is a common acquaintance. Also, one might talk about sensitive matters with a sibling amongst other family members but avoid such topics with the sibling when amongst colleagues. In these above scenarios, knowledge of the collocutor’s social network is implicitly required to determine if a given interactive action would be contextually pertinent or not. Similarly, naturalistic human-machine social interactions would require machines to represent and leverage on information about the underlying social networks governing behavioral interactions between humans.

# Social networks, however, are abstract constructs in human minds that are hidden with respect to third-party observers such as another person or a machine. A social network exists only because the persons involved preferentially interact with each other in specific ways. As such, the nature of social connections between persons must be inferred from observations of their interactive behaviors. Critically, other work has shown that artificial neural networks implementing Theory of Mind (e.g. ToMnet [6]) can observe past social interaction outcomes between agents and targets (derived from predetermined interaction rewards) and form internal representations of agents’ hidden false beliefs.

# In this study, we adapted the work in [6] to construct ToMnet+ and evaluated how its operation might represent hidden social networks underlying observed interactions between agents and targets. Such a demonstration has implications on how neural models might be engaged to infer deep relational structures in apparently disparate observations across various data problems. In addition, as mentioned, an artificial neural network developed along these lines might also be integrated as a dynamic module in social virtual agents or robots to enhance human-machine interactions across various functional contexts.

# Core to our approach in this work is simulating plausible social networks that constitute ground truth against which to assess performance of ToMnet+ (Fig. 1). These simulated social networks consisted of agents with different inter-personal connection weights to targets. Importantly, connection weights were based on the range of scores from the Social Support Questionnaire (SSQ) commonly used in Psychology to evaluate real human social dependencies on specific persons [7]. In general, people more readily approach and interact with persons in their network whom they perceive as providing them with greater social support [8], [9].

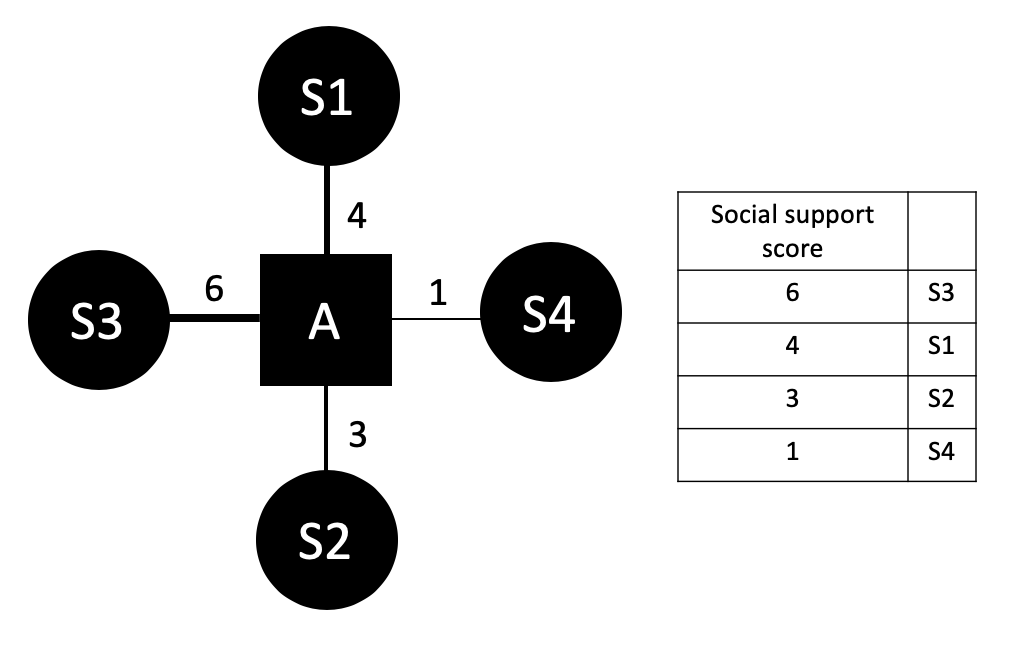


Fig. 1. An example simulated social support network of agent, *A*, and four targets, *S*. Connections between the agent and targets are modulated by the degree of social support the agent perceives receiving from each target.

# Simulated social support networks were thus used to generate sets of agent interactions with targets in different social contexts from which ToMnet+ learned. To test for a hidden social support network representation, we asked ToMnet+ which target an agent would preferentially interact with over various novel combinations of social contexts. We considered that the rank order of agent-target social support weights captures the base topology of our simple simulated social networks. As such, the goal is to determine if ToMnet+ judgment of agent-target social interaction preferences during test has a similar rank order as the social support weights.

# In the following, Section II expands on the notion of social support and its influence on human social interaction as well as considers relevant findings on machine learning of human social preferences. Section III covers our methodology regarding the SSQ, simulation generation, additional real human social interaction data acquisition for ecological validation, and ToMnet+ architecture and implementation. Section IV reports ToMnet+ performance results for the simulation data as well as human data. Section V discusses the findings and conclusion.

# Background and Related Works

## A. Machine Learning of Social Interaction Preferences

Both in the literature and in commercial applications, there are many instances of artificial intelligence being used to learn user preferences in order to provide personalized, or targeted, services from different types of data – visual, verbal, metadata – and with different machine learning approaches.

In the field of virtual agents, Recommender Systems are ubiquitous on web platforms and in our personal computing devices (smartphones and others). These artificial agents learn human preferences and adjust their service accordingly with the objective of maximizing the time the user spends using the platform, which often translates to increased profits from advertisements. These systems often rely on hard metrics (e.g. usage time, number of items consumed), content meta-data (e.g. tags, title, author) and user-generated data (e.g. ratings, engagement in social feedback systems) to train the learning structures that adjust the nature of the content feed shown to the user [1], [2].

Socially Assistive Robotics is a discipline where a robot –defined as an embodied intelligent agent – provides a service to its user, either physical (e.g. rehabilitation therapy) or psychological (e.g. companionship, emotional support). Previous research [3] has indicated that autonomous cognitive and social profiling of the user are key to deploying social robots in environments outside of the laboratory (e.g. in hospitals, schools or at home). Given the recent explosion of social media and ad-based media consumption platforms, data to train such systems is now abundant, but not often made publicly available. To this end, recent work engaged robots that interpret the behavior of their users through implicit cues in their facial expression and body gestures to infer mental states, personalities and emotions and, using this information, use a decision making process to determine how to best interact with a specific user. In [10] a platform for robot-assisted photo reminiscence for the elderly was introduced. Reminiscence Therapy is an intervention commonly used to alleviate neurodegenerative impact in patients with Mild Cognitive Impairment. It most often relies on the evocation of memories from the users’ lifetime through discussion of personal effects such as photographs. The authors implemented a variety of deep neural networks and a knowledge based inference process to generate questions that are directly related to the visual content of each photograph. Results showed that participants felt more engaged when the robot asked questions related to the photograph and their speech than when not. [11] presented a method to promote trust between humans and robots. The approach was based on Natural Language Understanding techniques where, when the user discloses their vulnerability to the robot, the system could infer the underlying feelings and desires of the user in order to provide relevant and effective emotional support. Finally, [12] and [13] discuss the importance of learning user preferences through interactions with an emotional support robot for children. The system presented performed a variety of actions (e.g. play videos, tell jokes), and learned user preferences by assessing emotional reactions from facial expressions with an Interactive Reinforcement Learning algorithm. The results revealed that people gave more positive feedback to and were more willing to interact with the robot after several sessions when it learned their preferences.

In sum, the evidence shows that social machines that infer user mental states through their implicit affective behaviors can engage actions better catered to user preferences, which in turn improves user experience and usage of the machine. However, as mentioned, selection of appropriate human-like social interactive behavioral preferences requires that the system also infers the user’s social context. Thus, the system proposed here aims to infer the social network of the user from naturalistic observation of the users’ behaviors around others in order to refine interactions between virtual agents or robots with users across various social situations.

## B. Social Support Networks

The role of social networks in modulating human inter-personal interaction behaviors has been extensively studied in sociology and psychology. To our knowledge, [14] was the first study to use the term “network” and apply its concept on a small Norwegian community to characterize how pairs of persons were socially related to each other. Specifically, two persons in the social network might be friends, each with their own other sets of friends, some of which might know each other or not. In order to achieve certain goals in the social network, each person interacts with specific sets of others, resulting in the formation of sub-classes of social function. Social networks are thus graphs with specific topologies [15] that emerge from tracing out the paths of relationships between a given person and how that person socially interacts with all other persons in the network.

Importantly, as mentioned, social network structure or topology also determines the sort of social interactions a person preferentially engages in or not. Critically, a person’s social network mediates the ease of obtaining support from others for certain needs [9], [16]. For instance, a baby obtains food from parents more readily than from siblings, and seeks out siblings for other purposes (despite similar physical proximities for both). Thus, a given person maintains several different classes of social support networks for different needs (e.g. emotional, financial, health) [17]. Such social support networks drive many everyday social interaction decisions between different people.

Several approaches have been applied to index social support networks [18]. These methods range from assessments of the availability of assistive persons to self-ratings of personal levels of social functioning. Of these, the SSQ [7] is one of the most common and well-validated instruments that is also easy to use. Essentially, the SSQ incorporates objective demographic (number of persons for a given need) and subjective psychological (satisfaction of support from each person) information across different types of support into its social network characterization. Its test format is also straightforward and systematic in a manner that is very suitable for the purposes of this study (see Methodology). Thus, we frame our simulated and real human social network structures adapting from the SSQ. We then construe social interaction preferences as a function of the differential satisfaction with received support across persons in one’s social network.

## C. Machine Theory of Mind

A key challenge in machine social network learning is the requirement to infer the hidden social connections from third-person observations of interaction behavior between agents and targets. This is a classic Theory of Mind problem, which entails the psychological mechanisms underlying a person’s ability to represent a model of other’s beliefs. For instance, in the Sally Anne test of Theory of Mind, the subject, experimenter, and a confederate together view a doll being placed in a box. The confederate then leaves the room after which the experimenter hides the doll in another second box. When the confederate later returns, subjects with Theory of Mind should not be surprised that the confederate looks for the doll in the first and not the second box. That is, the subject infers the confederate’s false belief from the logical association of observed sequences of events and behaviors. Likewise, social networks explaining the interactions between agents and targets are inferred from observations of the interactions.

A screenshot of a video game

Description automatically generated

Fig. 2. System overview schematic. Agent, *A*, interacts with targets, *S*, by approaching or avoiding them in grid world. On the left, three example grid world interaction context trajectories are shown (green arrows delineate paths taken). Agents’ movements to targets depended on which target provided greater social support in its social network penalized by the distance required to travel to reach the target. Interaction trajectories were fed to ToMnet+ as learning data. On the right, example new test grid worlds are presented (top; different combinations of *S* are drawn during testing) for which ToMnet+ predicts the trajectory agents take (bottom; blue arrows).

[6] presents ToMnet’s ability to represent an agent’s false beliefs. ToMnet observes past social interactions of an agent with targets and encodes character embeddings representing which targets an agent prefers over these histories. Integrating these character embeddings with internal state representations, ToMnet predicts which social actions an agent would perform with respect to targets in new given contexts. Importantly, the authors also applied random changes to target states in the social context that were hidden to the agent. For example, a target might be removed from the context, with this information known to ToMnet but not the agent. Despite this, ToMnet still predicted agent actions vis-à-vis the agent’s status quo as if targets were present, thereby displaying its inference about the agent’s false belief. Because of its ability to derive hidden states from observations, in this proposed system, we apply a modification of ToMnet to infer social networks through observations of how agents interact with targets.

# Methodology

## A. System Overview

We simulated a social network of a virtual agent (*A*) with four social targets (*s*) based on the degree of social support the agent perceives about each target. We then placed agents and selected targets in locations in a grid world randomly generated with barriers (Fig. 2). Using these social agent parameters, we simulated agent movement behaviors to targets by contrasting the normalized physical distances, , which is the normalized number of steps between agent and targets, and normalized social support, ,where . Then,

such that the agent was set to move toward the social target with largest , constituting a grid world trajectory instance (green arrows in Fig. 2). We simulated 10,000 grid world trajectories in total. In addition, a real human participant also indicated four friends (targets) with social support information obtained using questionnaires adapted from the SSQ. The participant was then placed in randomly generated grid world scenarios and was tasked to reach one of the targets with minimal moves, yielding 9,830 real human trajectory data. We then applied a modified ToMnet+ consisting of a deep-learning architecture with a character network and a prediction network. The character network consumes the agent’s past grid world trajectories to summarize the acting agent’s “character”. The prediction network then consumes the 8-dimensional character vector and, with a given new grid world state, predicts which target the agent will approach. The model was applied using a 8:1:1 training, validation, and testing split on simulated and human data separately using Tensorflow [19]. The trained model was used to infer the underlying preference of the simulated/human agents by feeding the model with start grid world states with four targets equidistant from agents. The probability of approaching each target output using softmax were agents preference scores for each target.

## B. Social Support Questionnaire

The SSQ consists of 27 items [7]. In this study, we simulated a social network of a virtual agent (A) with four social targets (s) based on the degree of social support the agent perceives about each target as in the SSQ.

## C. ToMnet+

ToMnet+ is inspired by ToMnet [8]. ToMnet+ is composed of a character network and a prediction network. The character network consists a residual network (resnet) [] with 5 layers followed by an average pooling layer and then a single-layer long-short term memory network (LSTM) []. The input of the character network is a batch of trajectories, Each trajectory is a 4d tensor of size 10×12×12×11, where 10 is the number of consecutive steps in the trajectory, 12 is the width and height of the grid world. 11 is the number of features. Trajectories that have more than 10 steps are truncated such that the last 10 steps are preserved, whereas the ones with less than 10 steps are padded with zeros (padded before the first step). The 11 features include 4 actions (up, down, left, right), the positions of 4 targets, the position of the obstacles, and the initial position of the agent. xxxxxx

## D. Social Game

….

# Experimental results

## A. Simulation Data

Performance of the trained models in predicting the agents’ final targets from test grid world start states is shown in Table 1 for both simulated and human data. Model performances as measured by the accuracy in predicting the agent’s final target given test start states in grid world. S1~S4 are 4 potential targets across the grid world instances, each with different personal and associative social parameter values constituting their underlying social network structures. Preference scores are normalized prediction probabilities (see Methods).

Because there were 4 potential targets in the test grid worlds, the chance prediction was 25%. Critically, the models which observed nothing more than the agents’ grid world behaviors were capable of inferring the agents’ underlying preference rankings.

1. Multi-target preference prediction performance

| Data Type | Final Target Accuracy | Predicted Preference Score | True Preference Score |
| --- | --- | --- | --- |
| Simulated | 80.14% | S4 (1.16)  >S3(0.83)  >S2(0.76)  >S1(0.62) | S4 (1.15)  >S3(0.79)  >S2(0.72)  >S1(0.72) |
| Human | 73.26% | S4 (0.71)  >S3(0.28)  >S2(0.01)  >S1(0.00) | S4 (0.95)  >S3(0.83)  >S2(0.83)  >S1(0.83) |

*A picture containing colorful, flying, kite

Description automatically generated*

Figure X. Accuracy in the test set as a function of the standard deviation (SD) of social support values across 4 targets in the training set for simulated data. Each red round dot is the average model test accuracy in test set (averaged across all the simulated data with the same SD). The blue triangle is the average random rate which should be the baseline to compared with. Random rate for each model is derived for each subject by dividing 100% by the average number of targets in the trajectories. The error bars represent the standard errors.

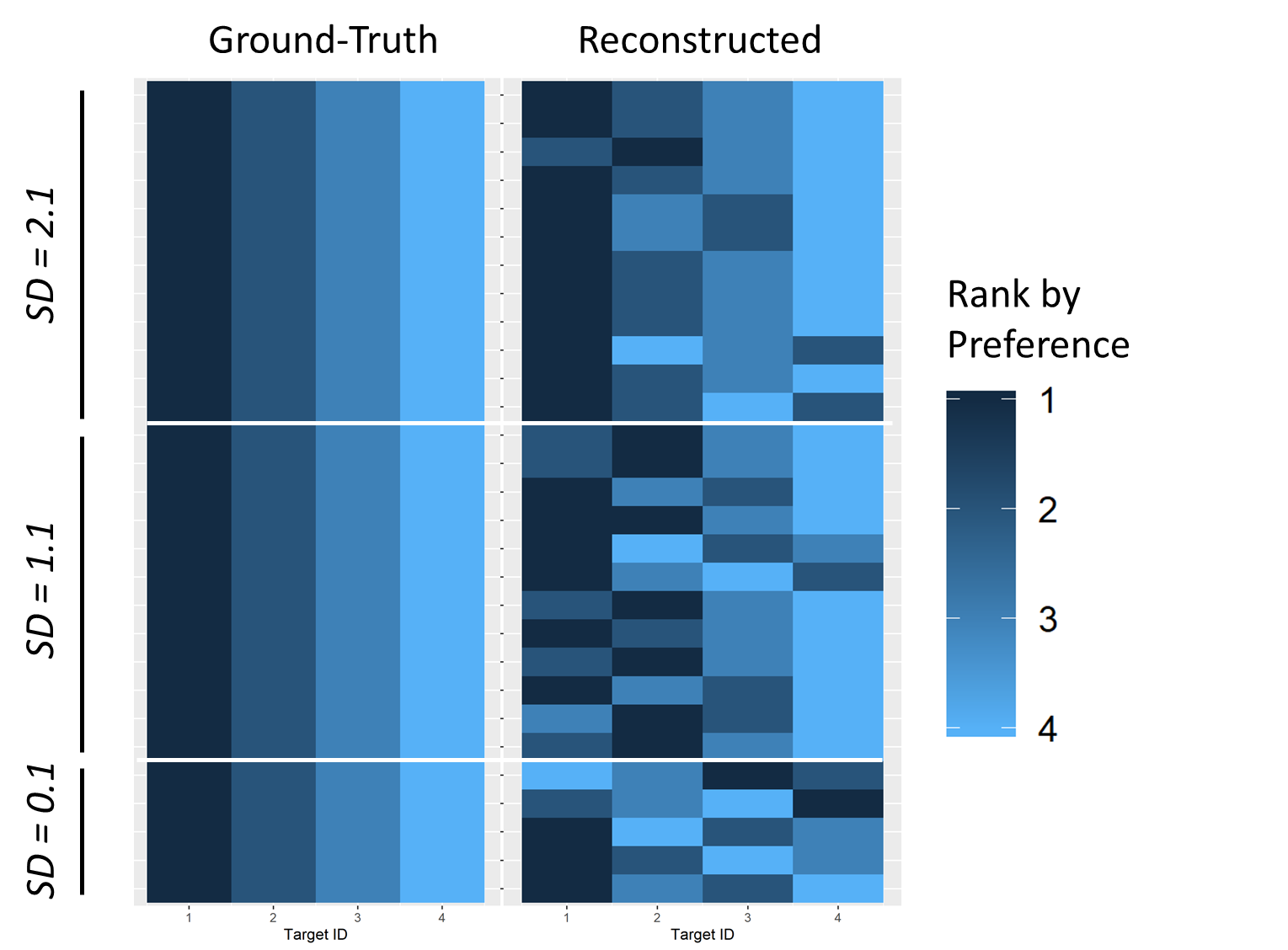
**

Figure Simulation-preference-matrix. The ground-truth preference matrix and the reconstructed preference matrix for simulated data. Each row is a subject and each column is a target that the subject interacts with. The color of the cell\_ij encodes the subject\_i’s ranked preference (1-4; 1 being the favorite target and 4 being the less favorite one) for target\_j. If there are tie(s) in the preference rank among targets (e.g., two or more targets share the same preference score), the targets with ties are assigned the average rank value (two targets share the second place in the preference score will have the rank value of 2.5). The ground-truth preference matrix is constructed by the rank-transformed simulated social support value of each target. The reconstructed preference matrix is constructed by the rank-transformed predicted preference score inferred by ToMnet+. The labels on the left are the standard deviations of the ground-truth preference scores (before rank-transformation) between the 4 targets.

## B. Human Data

A close up of a map

Description automatically generated

Figure X Accuracy in the test set as a function of the number of trajectories in the training set for human data set. Each red round dot is the model accuracy in test set for each human subject. The blue triangle is the random rate, which should be the baseline to compared with. Random rates are derived for each subject by dividing 100% by the average number of targets in the trajectories. The x-axis is log-transformed for clearer illustration. The label besides each red dot is the subject ID.

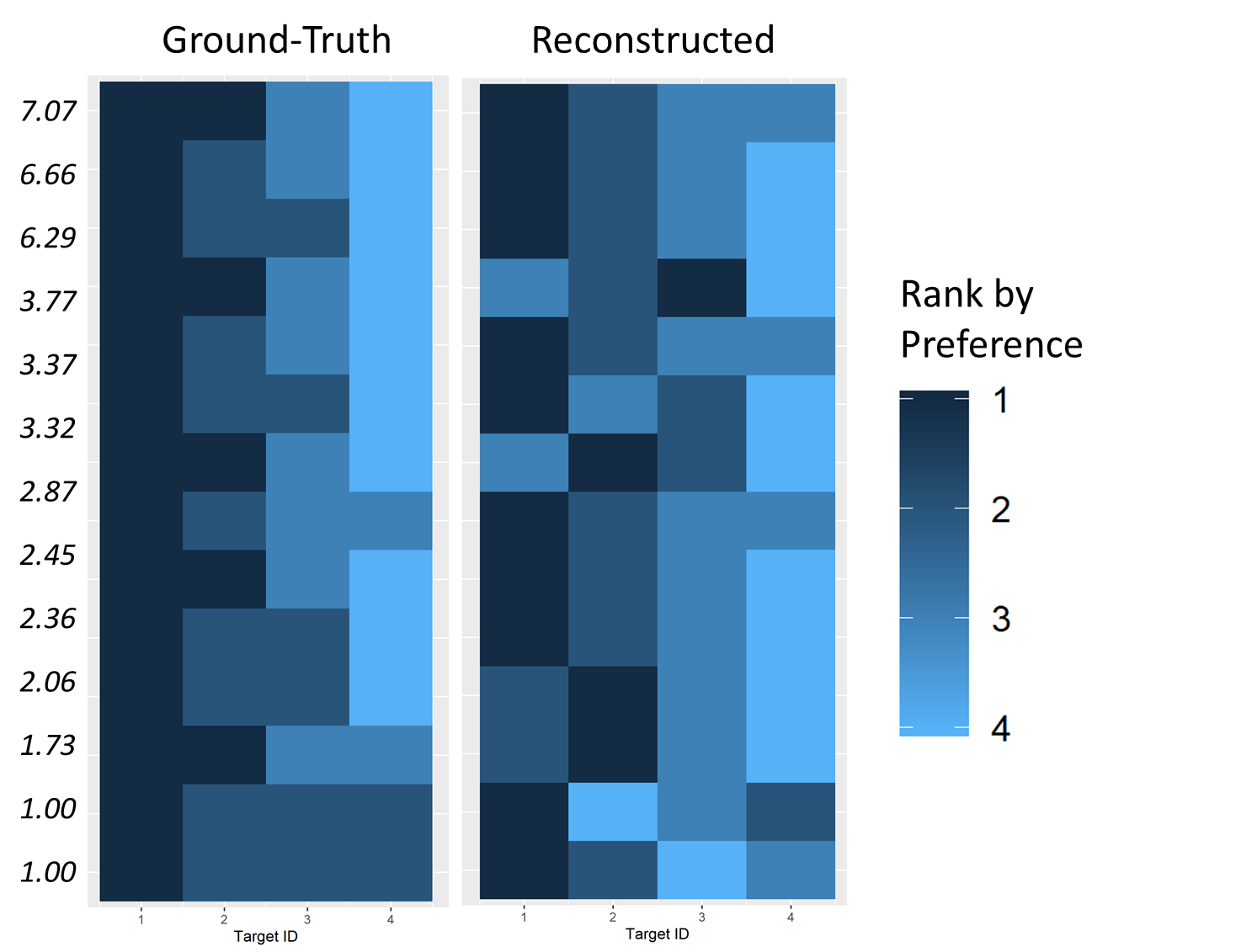


Figure Simulation-preference-matrix. The ground-truth preference matrix and the reconstructed preference matrix for simulated data. Each row is a subject and each column is a target that the subject interacts with. The color of the cell\_ij encodes the subject\_i’s ranked preference (1-4; 1 being the favorite target and 4 being the less favorite one) for target\_j. If there are tie(s) in the preference rank among targets (e.g., two or more targets share the same preference score), the targets with ties are assigned the average rank value (two targets share the second place in the preference score will have the rank value of 2.5). The ground-truth preference matrix is constructed by the rank-transformed simulated social support value of each target. The reconstructed preference matrix is constructed by the rank-transformed predicted preference score inferred by ToMnet+. The labels on the left are the standard deviations of the ground-truth preference scores (before rank-transformation) between the 4 targets.

# Discussion

## A. Limitations

## B. Conclusions

Our findings highlight the potential of machine applications that infer implicit human preferences from third-person behavioral observation data. This is distinct from most current applications that are focused on dissociating explicit signals (e.g. recognizing emotional categories from facial expressions). This is also distinct from the previous study, which used ToMnet to extract preference from simulated agents without hidden associative structures. We demonstrate that a NN such as ToMnet can also model real hidden social networks reflected in human social preferences. Aside from artificial intelligence and robotics applications, our findings also have implications in neuropsychological research. In principle, the human brain is also a neural network, albeit more complex, that operates by integrating observations of how other humans behaviorally interact to generate an internal hypothesis about real social networks [20]. As such, it is intriguing to consider such model implementations of learning and behavior as formal theory about the information mechanisms at work in human brains. With this initial platform, future work expanding on different formats of behavioral information and NN architectures can then be used to better understand how the human mind grasps reality.

References

[1] Z. Zhao, H. Lu, D. Cai, X. He, and Y. Zhuang, “User preference learning for online social recommendation,” *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 9, pp. 2522–2534, Sep. 2016.

[2] A. Bellogín, I. Cantador, P. Castells, and Á. Ortigosa, “Discovering Relevant Preferences in a Personalised Recommender System using Machine Learning Techniques,” 2008.

[3] H. Gunes, O. Celiktutan, and E. Sariyanidi, “Live human-robot interactive public demonstrations with automatic emotion and personality prediction.,” *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 374, no. 1771, p. 20180026, Apr. 2019.

[4] I. Momennejad, A. Duker, and A. Coman, “Bridge ties bind collective memories.,” *Nat Commun*, vol. 10, no. 1, p. 1578, Apr. 2019.

[5] R. I. Dunbar, A. Marriott, and N. D. Duncan, “Human conversational behavior.,” *Hum. Nat.*, vol. 8, no. 3, pp. 231–246, Sep. 1997.

[6] N. Rabinowitz, F. Perbet, F. Song, C. Zhang, S. M. A. Eslami, and M. Botvinick, “Machine Theory of Mind,” presented at the Proceedings of the 35th International Conference on Machine Learning, 2018, vol. 80, pp. 4218–4227.

[7] I. G. Sarason, H. M. Levine, R. B. Basham, and B. R. Sarason, “Assessing social support: The Social Support Questionnaire.,” *J. Pers. Soc. Psychol.*, vol. 44, no. 1, pp. 127–139, 1983.

[8] T. L. Albrecht and D. J. Goldsmith, “Social support, social networks and health. Teoksessa TL Thompson, AM Dorsey, KI Miller & R. Parrott (toim.) Handbook of health communication,” 2003.

[9] C. A. Heaney and B. A. Israel, “Social networks and social support,” *Health behavior and health education: Theory, research, and practice*, vol. 4, pp. 189–210, 2008.

[10] Y.-L. Wu, E. Gamborino, and L.-C. Fu, “Interactive Question-Posing System for Robot-Assisted Reminiscence from Personal Photos,” *IEEE Trans. Cogn. Dev. Syst.*, pp. 1–1, 2019.

[11] X. Guo, Y.-C. Huang, E. Gamborino, S.-H. Tseng, L.-C. Fu, and S.-L. Yeh, “Inferring Human Feelings and Desires for Human-Robot Trust Promotion,” in *Cross-Cultural Design. Methods, Tools and User Experience: 11th International Conference, CCD 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part I*, vol. 11576, P.-L. P. Rau, Ed. Cham: Springer International Publishing, 2019, pp. 365–375.

[12] E. Gamborino and L.-C. Fu, “Interactive Reinforcement Learning based Assistive Robot for the Emotional Support of Children,” presented at the 2018 18th International Conference on Control, Automation and Systems (ICCAS), 2018.

[13] E. Gamborino, H.-P. Yueh, W. Lin, S.-L. Yeh, and L.-C. Fu, “Mood Estimation as a Social Profile Predictor in an Autonomous, Multi-Session, Emotional Support Robot for Children,” in *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2019, pp. 1–6.

[14] J. A. Barnes, “Class and committees in a norwegian island parish,” *Human Relations*, vol. 7, no. 1, pp. 39–58, Feb. 1954.

[15] L. Euler, “Leonhard Euler and the Koenigsberg Bridges,” *Sci. Am.*, vol. 189, no. 1, pp. 66–72, 1953.

[16] K. Wright, “Social networks, interpersonal social support, and health outcomes: A health communication perspective,” *Front. Commun.*, vol. 1, Oct. 2016.

[17] K. B. Wright and C. H. Miller, “A Measure of Weak-Tie/Strong-Tie Support Network Preference,” *Commun Monogr*, vol. 77, no. 4, pp. 500–517, Dec. 2010.

[18] B. H. Gottlieb and A. E. Bergen, “Social support concepts and measures.,” *J Psychosom Res*, vol. 69, no. 5, pp. 511–520, Nov. 2010.

[19] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, *TensorFlow: A System for Large-Scale Machine Learning*. 2016.

[20] J. O. S. Goh, H.-Y. Hung, and Y.-S. Su, “A conceptual consideration of the free energy principle in cognitive maps: How cognitive maps help reduce surprise,” vol. 69, Elsevier, 2018, pp. 205–240.