2020 29th IEEE International Conference on
Robot and Human Interactive Communication (RO-MAN)
August 31 - September 4, 2020 - Virtual Conference

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 applied an artificial neural network with machine theory of mind (ToMnet+) to learn and predict social preferences based on implicit information from the way agents and social targets interact behaviorally. Our findings have implications for machine applications that seek to infer hidden information structures solely from third-person observation of behaviors. We consider that social machines with such an ability would have an enhanced potential for more naturalistic human-machine interactions.**

I. Introduction

The use of artificially 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]–[6]. 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.

One characteristic driving fluent human social interactions is access to information about the underlying social network of the persons involved, which is often implicit [7], [8]. 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 the 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 from 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, work has shown that artificial neural networks implementing Theory of Mind (e.g. ToMnet [9]) can observe past social interaction outcomes between agents and targets (derived from predetermined interaction rewards) and form internal representations of the agents' hidden false beliefs. In addition, other works have also evaluated the importance of Theory of Mind in machines in order to enhance human-machine interactions [10]–[12].

In this study, we adapted the work in [9] to construct ToMnet+ and evaluated how its ability to represent hidden social networks from observed interactions between agents and targets. Such a demonstration has implications on how neural network models might be engaged to infer deep relational structures in apparently disparate observations across various data problems. Note that this approach is distinct from previous studies on inferencing social relationships using Bayesian algorithms [13]. 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 the performance of ToMnet+ (Fig. 1). These simulated social networks consisted of agents with different inter-

*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 and the Center for Artificial Intelligence and Advanced Robotics, 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).

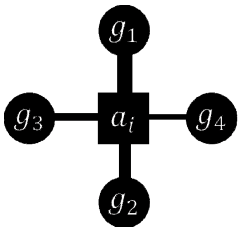
E. Gamborino is with the Center for Artificial Intelligence and Advanced Robotics, National Taiwan University, Taipei, Taiwan. (phone: +886 958 376 105; e-mail: gamborino@ntu.edu.tw).

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).



Target (g_s)	Social Support Score (u_s)
g_1	21
g_2	17
g_3	10
g_4	7

Fig. 1. An example instance of the simulated simple social support network of agent a_i, and four targets, $\{g_s|s \in [1,4]\}$. Connections between the agent and targets represent the degree of social support (u_s) the agent perceives for each target.

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 [14]. In general, people more readily approach and interact with persons in their network whom they perceive as providing them with greater social support [10]–[12], [15], [16]

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 the judgement of ToMnet+ on agent-target social interaction preferences could predict a similar rank order when compared to the social support weights.

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

II. 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 physically (e.g. rehabilitation therapy) or psychologically (e.g. companionship, emotional support). Previous works [3]–[5] have 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 [17] 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. [18] 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, [19] and [20] 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 about and were more willing to interact with the robot after several sessions when it learned their preferences.

In sum, the evidence above 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, [21] 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 [22] 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 [16], [23] . 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) [24]. Such social support networks drive many everyday social interaction decisions between different people.

Several approaches have been applied to index social support networks [25]. These methods range from assessments of the availability of assistive persons to self-ratings of personal levels of social functioning. Of these, the Social Support Questionnaire (SSQ) [14] is one of the most common and well-validated instruments that is also easy to use. Essentially, the SSQ incorporates an objective demographic (number of persons for a given need) and subjective psychological information (satisfaction of support from each person) 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 famous 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. Such inferring of others' beliefs is one of many aspects of Theory of Mind, which involves various specific social, cognitive, and affective processes. Importantly, we propose that to infer social networks explaining the interactions between agents and targets from observations of the interactions likely involves Theory of Mind processes in humans.

[9] 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 to 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.

III. Methodology

A. The Social Game for Simulated Agents

As mentioned, key to our approach to validate ToMnet+'s ability to infer social networks is to first generate ground truth information against which we can assess the system's performance. To this end, we simulated social networks for 30 virtual agents, $\{a_i | i \in [1,30]\}$, each with four different social targets, G=$\{g_s | s \in [1,4]\}$, whom the agent perceives as providing different degrees of social support u_s (Fig. 1; see below). For each agent a_i, we simulated 10,000 2-dim 12×12 grid worlds, $\{w_j | j \in [1, 10000]\}$. In each grid world w_j, we placed n_{target} targets (range from 1 to 4) and $n_{barrier}$ barriers (range from 0 to 50) in random locations (Fig. 2). From the agent a_i's perspective, each target g_s has a social reward value u_s, and a physical distance d_{js}, which is the minimum steps the agent needs to take to approach the target in w_j. The agent could only move vertically or horizontally, but not diagonally. The agent could not move into where the barriers are or out of the boundary of the grid world. Once the agent reaches one of the targets, the trajectory is completed and a new grid world w_{j+1} follows. For each world w_j, the target that the agent a_i approaches in the end (g_j^*) is decided by

$$g_j^* = \arg\min_s (u_s - d_{js}) \quad (1)$$

Specifically, the agent a_i approaches g_j^* using the shortest path in a deterministic way, constituting a grid world trajectory instance τ_j (green arrows in Fig. 2). Note that in our present implementation the agent has full visibility of the grid world. For each virtual agent, the social reward values for its four targets $\{u_s | s \in [1,4]\}$were sampled randomly from a uniform distribution between 0 and 26. This range of u_s [0,26] was chosen to adjust the maximal number of steps to diagonally span a no-barrier grid world without retracing (23) based on the response range of 9 and maximum score of 10 in the social game for human participants (10*23/9 ≈ 26); see also III.B). We also imposed a constraint such that sample u's

for 6 virtual agents had a standard deviation (SD) of 0.1, 12 agents had SD 1.1, and 12 agents had SD 2.1. We imposed this constraint on the SDs to test the robustness of the model (preference inference should be harder for smaller SDs).

B. Social Support Questionnaire

The social networks constructed in the above simulations and assessed in human participant data below were based on the SSQ [14]. The original SSQ is written in English and consists of 27 items evaluating different aspects of social support. In this study, to obtain human agent-target social reward values, we applied two modifications to the SSQ. These include translation to Chinese and simplification to 7 items focusing on more psychological emotional aspects of support. These items, each on a 10-point (1 to 10) Likert scale, asked participants to rate the degree to which their target persons fulfilled the following roles: 1) The person can provide me with social support, 2) I can turn to this person for advice about handling problems, 3) The person cares about me, regardless of what is happening to me, 4) I can count on the person to help me feel better when I am feeling generally down in the dumps, 5) I can count on the person when I need help, 6) I can share my most private worries and fears with the person, 7) The person is important in my life. Participant ratings for these items were then used to compute the social reward value for the human social game. Specifically, we scaled the averaged value of the adapted SSQ ratings for each target by a factor of 23/9, where 23 is the maximal steps to take to span the diagonal of a no-barrier grid world and 9 is the range of the adapted SSQ score. The scaled adapted SSQ values were then assigned to the target as its social reward value with respect to the agent.

C. The Social Game for Human Participants

Twelve human participants (mean age = 26.2 yrs, age range 23 to 32 yrs, 4/8 males/females) played a social game which had a similar format as the game for simulated agents. Participants all gave written informed consent for this study, which was approved by the National Taiwan University local Research Ethic Committee (NTU-REC case no. 201803HS017). Participants played the game via web browser either by mobile phone or computer and all completed at least 150 trajectories.

Before starting the social game, human participants were asked to list four target close friends/family members whom they felt provided them with the most social support. The participants then completed our adapted 7-item SSQ for each of the four target persons. Participants were then presented with a screen showing a grid world with 1-4 targets and barriers, the reward assigned to each target, along with an agent that the participant should control (Fig. 3). The action space was the same as for the virtual agent (i.e., only horizontal and vertical moves). The social reward value of each target was the scaled adapted SSQ score. The current score was shown in the upper left corner. Each step costs one point. The final score for each trajectory round was the social reward value of that target minus the number of steps taken. After reaching a target, another grid world context was then presented to the participant consisting of different target sampling (out of the listed four), spatial configurations, and score set to 0. Note, two participants played the game twice entering distinct sets of targets each time. Thus, we had a total of 14 human data sets.

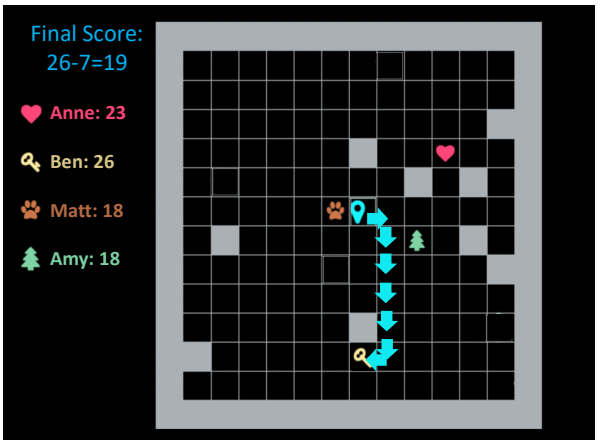


Fig. 3. An example grid world of the social game for humans. In each trial, participants are presented with a 12×12 grid world with 1-4 targets, represented by a heart, a key, a paw, and a tree. The name associated with each target as well as the corresponding social reward, u_s, and the current score are shown in the left. The human participant (map marker) navigates horizontally or vertically to reach their desired target by keypress or clicking buttons on the mobile phones. Once the participant reaches one of the targets, the participant is rewarded with the final score, which is derived via subtracting u_s, by the number of steps taken.

D. ToMnet+

Given either the simulated or human data described above, in order to infer the agents' preferences we applied a ToMnet+ model, extending from [9] (Fig. 2, 4). Specifically, we based our model on section 3.2 of [9] which was designed to infer goal-directed behavior from single shot of new trajectories. In addition, the major extension from the original ToMnet is that we included a "preference inference phase" to the system. The primary trainable parameters reside in the a character network and prediction network components of the model (Fig. 4). One ToMnet+ model was trained for each virtual agent/human. For each agent, ToMnet+ takes two inputs at a time: a trajectory τ_j and a query state q_k. The query state q_k is the shot of the first time-step of another trajectory τ_k (with $j \neq k$). The rationale is that the character network should extract an abstract representation of the agent's preference for targets from τ_j and represent it as a character embedding $e_{char,j}$. The prediction network then takes $e_{char,j}$ as input and predicts the target $(\widehat{g_k^*})$ the agent will approach in *another* trajectory given query state q_k. The model is trained end-to-end with tuples of (τ_j, q_k, g_k^*). The rationale is that once the model is trained, it can (1) extract the agent's preference $(e_{char,j})$ from the trajectory (τ_j) and (2) utilize the preference information $(e_{char,j})$ to predict which target the agent will approach $(\widehat{g_k^*})$.

The character network consumes each trajectory τ_j and outputs the character embedding $e_{char,j}$, which contains the abstract representation of the agent's preference for each target. Each trajectory τ_j is a 4d tensor (10×12×12×11), where 10 is the number of consecutive time steps in the

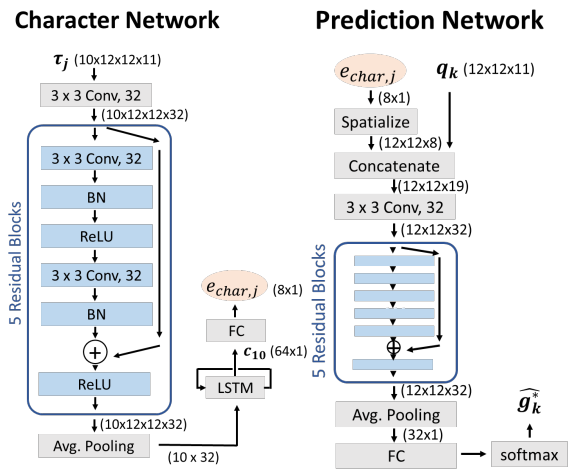


Fig. 4. Schematic diagrams of the character and prediction networks and other components in the ToMnet+ architecture. The number in parentheses indicate the size of tensors at each step. Conv: convolutional layer, BN: batch-normalization, ReLU: rectified linear unit activation function, Avg. Pooling: average pooling layer, LSTM: long short-term memory layer, FC: fully-connected layer, Spatialize: spatialize 2d vector into 3d tensor, Concatenate: concatenate two 3d tensors, τ_j : the trajectory, $e_{char,j}$: the character embedding, q_k the query state, $\widehat{g_k^*}$ the predicted approached target.

trajectory, 12 is both the width and height of the grid world, and 11 is the number of feature channels. Trajectories with more than 10 steps are truncated such that the last 10 steps are preserved, whereas the ones with less than 10 steps are zero-padded before the first step. The 11 binary feature channels include 5 actions (up, down, left, right, reaching the goal), the positions of the 4 targets, the position of the obstacles, and the initial position of the agent. If a target is absent in the trajectory τ_j, its feature plane is zero-padded. Following the design of [9], thirty-two 3×3 convolutional kernels are first applied to each time step (12×12×11) separately to scale the number of channels from 11 to 32. The convolved τ_j (10×12×12×32) is then passed into a resnet [26] with 5 residual blocks, each with 32 channels, batch-normalization, and ReLU nonlinearity. The output from the resnet is a 4d tensor (10×12×12×32), which then passes through a global average pooling layer that collapses the entire spatial dimension into a 2d tensor (10×32), which is a sequence representing the trajectory spatial information in each time step. This sequence with 10 time steps is passed to a single-layer long short-term memory (LSTM [27]) with 64 channels. The last LSTM cell state summarizing all time steps is then extracted with a dense layer to yield an 8-dim character embedding which is $e_{char,j}$.

The prediction network predicts the target $\widehat{g_k^*}$ that the agent will approach in the query states q_k given $e_{char,j}$. The character embedding $e_{char,j}$ is spatialized and concatenated with the query state q_k, which together form a 3d tensor of size 12×12×(11+8). This tensor then passes through 32 3×3 convolutional kernels which scales the number of channels from 19 to 32. The results are fed into a 5-layer resnet, with 32 channels, batch-normalization, and ReLU nonlinearity, followed by a global average pooling layer, and a dense layer to yield 4-dim logits, followed by the output softmax layer to give $\widehat{g_k^*}$. The loss function used was the softmax cross-entropy loss. The ToMnet+ model for each virtual and human agent was trained separately with an 8:1:1 training, validation, and testing split of the tuples (τ_j, q_k, g_k^*). We trained each model with the Adam optimizer [28] with initial learning rate = 10^{-4}, batch size = 16, and number of steps = 10^4.

After the model was trained, it was then used to infer virtual agent/human's preference for each target. For each agent a_i , we fed 100 pairs of (τ_{j+}, q_{k+}) to the trained model. τ_{j+} is a subset of τ_j that had exactly 4 targets to ensure that the generated test character embedding contains information about 4 targets. q_{k+} is a special query state that exists only for preference inference (we called it "inference query state"), where there are no barriers and the agent is placed equidistant from all targets. The exact positions of the 4 targets were shuffled randomly across all pairs. For each pair of (τ_{j+}, q_{k+}), the soft-max probability for each target was averaged across 100 pairs. The average softmax probability is then rank-transformed to get agent a_i's inferred preference ranking $\widehat{pref}$. ToMnet+ was implemented in Tensorflow version 1.12 [29].

IV. Experimental Results

A. Simulated Data

The performance of the trained models was evaluated with the tuples (τ_j, q_k, g_k^*) in the testing set, measured by the accuracy of predicting the virtual agents' final targets g_k^* from the query state q_k given another trajectory τ_j (Fig. 5). The model for each virtual agent reached above 80 % regardless of the SD(u). A Wilcoxon signed-rank test indicated the model accuracies were above chance, W = 465, p = 1.86×10^{-9}. Critically, in the preference inference phase, the model could infer the virtual agents' underlying preference rankings (Fig. 6A). To quantify how well ToMnet+ inferred target preference, we derived Kendall's tau-b (a non-parametric correlation coefficient) for each agent via

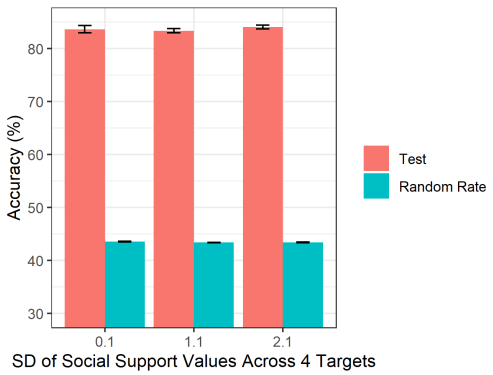


Fig. 5. Accuracy in the test set for models trained with simulated data. Each stack of bars represents virtual agents with different standard deviation (SD) of social support values across 4 targets in the training set. Each red bar is the average model test accuracy in test set (averaged across all the simulated data with the same SD). The blue bar is the average random rate which represents the baseline to be compared with. Random rate for each model is derived for each simulated agent by dividing 100% by the average number of targets in the trajectories. The error bars represent the standard errors.

correlating the ground-truth simulated preference ranking and the inferred preference ranking. Subsequently, we tested whether the median of the distribution of Kendall's tau-b was greater than 0 with a Wilcoxon signed-rank test. The result indicated that the inferred preference rankings significantly correlated with the ground-truth, $W = 390$, $p = 1.61 \times 10^{-5}$.

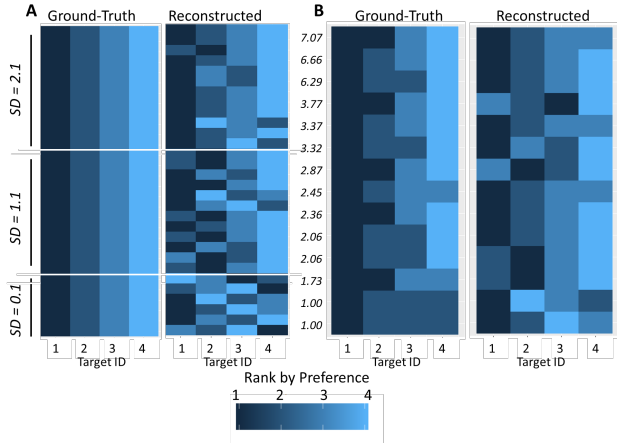


Fig. 6. The ground-truth preference matrix and the reconstructed preference matrix for (A) simulated data and (B) human data. The color of the cell at row i and column j encodes the agent a_i's ranked preference (1-4; 1 being the favorite target and 4 being the least favorite) for target g_j . If there 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 (*i.e.* 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 social support value. The reconstructed preference matrix is constructed by the preference rank inferred by ToMnet+ (see Methods). The labels on the left are, SD(u), the standard deviations of the ground-truth preference scores (before rank-transformation) between the 4 targets.

B. *Human Data*

We evaluated the models trained with human data in the same way as for simulation data. The model accuracy was between 50.0% to 81.6% across all 14 data sets, which is significantly above chance, $W = 105$, $p = .0001$. Moreover, with greater numbers of human training trajectories available, model accuracy improved (Fig 7). The model could reconstruct participants' preference rankings (Fig. 6B). We used the same method as above to quantify how well ToMnet+ reconstructed ground-truth human preference ranking. Inferred preference rankings significantly correlated with the ground-truth human preference rankings, $W = 105$, $p = .00103$.

V. Discussion

A. *Limitations*

It is possible that the human social game implemented in this study was not an entirely accurate probe of the true social preferences of the human participants. The structure of the social game was such that participants could have engaged movements in grid world simply to maximize points in the game. As such, the participant behaviors we tested may not reflect their underlying social networks but merely their ability to adhere to their reported SSQ ratings of targets and engage strategic actions. Nevertheless, we argue that as long

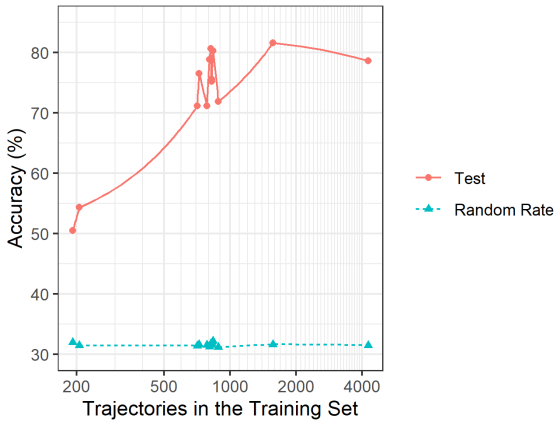


Fig. 7. Accuracy for the test sets as a function of the number of trajectories in the human data training set. Red round dots are the model accuracies for the test set of each human participant. Blue triangles are the random rates, which should be the baseline to be compared against. Random rates were derived for each participant by dividing 100% by the average number of targets in the trajectories. The x-axis is log-transformed for clearer illustration. Each grid along the x-axis represents 100 trajectories.

as participants truthfully reported their social support target details and sought to maximize final scores in the game, the resulting behavior sufficiently allows ToMnet+ to infer participant social support details as reported. We maintain that these do reflect participant social support networks, and critically, these details were hidden from ToMnet+.

The approach we adopted to evaluate ToMnet+ relied on quantized spatial movement in grid world as a proxy for social interaction preferences. This grid world input format limits ToMnet+'s applicability to problems that might not be suitably formulated as such, albeit possible mapping transformations might be found. Also, we note that agents in our grid world were only allowed to interact with targets and targets did not interact with each other. This is certainly not realistic since true human social networks are more dynamic, with all agents/targets co-interacting. Moreover, ToMnet+ maintains a birds eye view of the grid worlds, which is not always available in real world settings. Thus, while our grid world representation was adequate for our proof-of-concept study to infer simple social networks, future extensions must consider other more universal and dynamic formats of value-laden information (e.g. evolving more complex graphs of actions, conceptual data, or even facial expressions).

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. This is further distinct from previous work adopting a Bayesian modeling approach [13]. We demonstrate that an artificial neural network with ToMnet-based architecture can also model

aspects of real hidden social networks reflected in human social preferences.

While this present study investigates an algorithmic approach to infer social networks from social behaviors, our findings might also have implications in neuropsychological research. In principle, the human brain is also a neural network that operates by integrating observations of human interactions to generate internal hypotheses about real social networks [30]. As such, such models of learning and behavior contribute formal theory about information mechanisms at work in human brains. We suggest that such platforms with specified neural network architectures might be used to better understand how the human mind grasps reality.

Critically, this proof-of-concept provides primary demonstration of ToMnet+'s ability to infer deep relational structures via observing social interaction behaviors. We recommend that future work should now begin to evaluate its efficacy in other data problems and examine its application in improving human-machine interactions using virtual agents or robots.

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, doi: 10.1109/TKDE.2016.2569096.

[2] A. Bellogín, I. Cantador, P. Castells, and A. Ortigosa, "Discovering relevant preferences in a personalised recommender system using machine learning techniques," presented at the Proceedings of the ECML-PKDD 2008 Workshop on Preference Learning, 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. B Biol. Sci.*, vol. 374, no. 1771, p. 20180026, Apr. 2019, doi: 10.1098/rstb.2018.0026.

[4] A. K. Ostrowski, D. DiPaola, E. Partridge, H. W. Park, and C. Breazeal, "Older Adults Living With Social Robots: Promoting Social Connectedness in Long-Term Communities," *IEEE Robot. Autom. Mag.*, vol. 26, no. 2, pp. 59–70, Jun. 2019, doi: 10.1109/MRA.2019.2905234.

[5] B. Scassellati *et al.*, "Improving social skills in children with ASD using a long-term, in-home social robot," *Sci. Robot.*, vol. 3, no. 21, p. eaat7544, Aug. 2018, doi: 10.1126/scirobotics.aat7544.

[6] B. Scassellati and M. Vázquez, "The potential of socially assistive robots during infectious disease outbreaks," *Sci. Robot.*, vol. 5, no. 44, p. eabc9014, Jul. 2020, doi: 10.1126/scirobotics.abc9014.

[7] I. Momennejad, A. Duker, and A. Coman, "Bridge ties bind collective memories," *Nat. Commun.*, vol. 10, no. 1, p. 1578, Apr. 2019, doi: 10.1038/s41467-019-09452-y.

[8] R. I. M. Dunbar, A. Marriott, and N. D. C. Duncan, "Human conversational behavior," *Hum. Nat.*, vol. 8, no. 3, pp. 231–246, Sep. 1997, doi: 10.1007/BF02912493.

[9] N. C. Rabinowitz, F. Perbet, H. F. Song, C. Zhang, S. Eslami, and M. Botvinick, "Machine theory of mind," *ArXiv Prepr. ArXiv180207740*, 2018.

[10] Y. Demiris and M. Johnson, "Distributed, predictive perception of actions: a biologically inspired robotics architecture for imitation and learning," *Connect. Sci.*, vol. 15, no. 4, pp. 231–243, Dec. 2003, doi: 10.1080/09540090310001655129.

[11] B. Scassellati, "Theory of Mind for a Humanoid Robot," *Auton. Robots*, vol. 12, no. 1, pp. 13–24, Jan. 2002, doi: 10.1023/A:1013298507114.

[12] C. Breazeal, J. Gray, and M. Berlin, "An Embodied Cognition Approach to Mindreading Skills for Socially Intelligent Robots," *Int. J. Robot. Res.*, vol. 28, no. 5, pp. 656–680, May 2009, doi: 10.1177/0278364909102796.

[13] F. Zanlungo, Z. Yücel, and T. Kanda, "Intrinsic group behaviour II: On the dependence of triad spatial dynamics on social and personal features; and on the effect of social interaction on small group dynamics," *PLOS ONE*, vol. 14, no. 12, p. e0225704, Dec. 2019, doi: 10.1371/journal.pone.0225704.

[14] 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, doi: 10.1037/0022-3514.44.1.127.

[15] T. L. Albrecht and D. J. Goldsmith, "Social support, social networks, and health.," in *Handbook of health communication.*, Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 2003, pp. 263–284.

[16] C. A. Heaney and B. A. Israel, "Social networks and social support.," in *Health behavior and health education: Theory, research, and practice, 4th ed.*, San Francisco, CA, US: Jossey-Bass, 2008, pp. 189–210.

[17] Y. Wu, E. Gamborino, and L. Fu, "Interactive Question-Posing System for Robot-Assisted Reminiscence from Personal Photos," *IEEE Trans. Cogn. Dev. Syst.*, pp. 1–1, 2019, doi: 10.1109/TCDS.2019.2917030.

[18] 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*, Cham, 2019, pp. 365–375.

[19] E. Gamborino and L. Fu, "Interactive Reinforcement Learning based Assistive Robot for the Emotional Support of Children," in *2018 18th International Conference on Control, Automation and Systems (ICCAS)*, Oct. 2018, pp. 708–713.

[20] E. Gamborino, H. Yueh, W. Lin, S. Yeh, and L. 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)*, Oct. 2019, pp. 1–6, doi: 10.1109/RO-MAN46459.2019.8956460.

[21] J. A. Barnes, "Class and Committees in a Norwegian Island Parish," *Hum. Relat.*, vol. 7, no. 1, pp. 39–58, Feb. 1954, doi: 10.1177/001872675400700102.

[22] L. Euler, "LEONHARD EULER AND THE KOENIGSBERG BRIDGES," *Sci. Am.*, vol. 189, no. 1, pp. 66–72, 1953.

[23] K. Wright, "Social Networks, Interpersonal Social Support, and Health Outcomes: A Health Communication Perspective," *Front. Commun.*, vol. 1, p. 10, 2016, doi: 10.3389/fcomm.2016.00010.

[24] 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, doi: 10.1080/03637751.2010.502538.

[25] B. H. Gottlieb and A. E. Bergen, "Social support concepts and measures," *J. Psychosom. Res.*, vol. 69, no. 5, pp. 511–520, Nov. 2010, doi: 10.1016/j.jpsychores.2009.10.001.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2016, pp. 770–778.

[27] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014.

[29] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," in *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, Savannah, GA, Nov. 2016, pp. 265–283, [Online]. Available: https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi.

[30] J. O. S. Goh, H.-Y. Hung, and Y.-S. Su, "Chapter Seven - A conceptual consideration of the free energy principle in cognitive maps: How cognitive maps help reduce surprise," in *Psychology of Learning and Motivation*, vol. 69, K. D. Federmeier, Ed. Academic Press, 2018, pp. 205–240.