國立臺灣大學醫學院腦與心智科學研究所

碩士論文

Graduate Institute of Brain and Mind Sciences

College of Medicine

National Taiwan University

Master Thesis

將機器心智理論模型應用於複雜社交網路

A Machine Theory of Mind for Processing

Complex Social Networks

賴彥伶

Yen-Ling Lai

指導教授：吳恩賜 博士

Advisor: Joshua Oon Soo Goh, Ph.D.

中華民國112年8月  
August 2023

致謝

在即將完成碩士論文的此刻，必須對所有曾經幫助過我的人致上深深的感謝，在眾人的協助下，撰寫論文的崎嶇之路上少了許多顛簸，也增添了許多色彩。

首先，感謝我的指導教授吳恩賜老師，在知識的傳遞以及引導邏輯思辨的能力上不遺餘力，讓我有機會一窺學術研究的奧秘。並且營造了最棒的實驗室環境，讓實驗室的成員們適得其所，能無所顧慮投入研究。感謝我的口試委員，傅立成老師與葉素玲老師，基於不同的專業領域，對我的研究提出非常有意義的建議與討論，使這份論文更趨完善。並且感謝實驗室的成員們，有昀軒協助度過研究的瓶頸，芝嘉與苡蓁的互相陪伴，以莊一向很有幫助的建議，煜翔幫忙解決各種疑難雜症，啟權與俊毅的協助，宛倫、軒齊、鈺軒、雅婷、宛儒、志瑜、易修、景伃、Ashley、麟涵、珺崴、一欣、兩位庭萱、之邑的提點與幫忙，以及芷華與守辰在協助實驗室事務上不遺餘力，感謝所有實驗室的成員們一同努力，才有這一份論文的產生。感謝AI計畫的所有成員，讓我有機會能參與其中。也感謝曉雯時常提醒我們該注意的事項，使在學期間更加順利。

接著，感謝我的父母將我養育成人，並給我這個機會踏入認知科學的領域，這段日子我們經歷了好多事情，但很開心我們一家人能一一克服難關，完成論文的此刻，當然要與您們一同分享這份喜悅。感謝我的外公、外婆、阿姨、姨丈、叔叔、伯伯、伯母、嬸嬸、表哥、堂姐、堂妹、表姐、表弟給予我恰好的關愛與溫暖。也感謝鈺凱以及三隻貓貓囉囉、嘎嘎、胖胖，支持我度過難關。感謝朋友們出現在我的生命中，各方面陪伴著我度過充滿壓力與挑戰的碩班生活。

最後，謹向所有曾經幫助、關心過我的人致上謝意，感謝您們協助我完成碩士論文。

賴彥伶 2023.8

中文摘要

社交網路定義了人類的人際關係。然而人際關係是抽象且需要觀察才能了解的，因此若要描繪社交網路的結構，需要藉由觀察社交行為方能建構。建構社交網路的結構對人類而言十分自然，我們汲取社交行為中的資訊並藉此平順地與人互動。無法與人類自然的產生社交行為成為發展流暢的人機互動的一大挑戰。在這篇研究中，我們認為要讓機器擁有社交能力的必要條件是，機器必須有能力藉由觀察社交行為推測社交網路的結構。在先前的研究中，我們設計一機器心智理論模型（此機器心智理論模型的雛型來自於 Google Deepmind），此模型可藉由觀察代理人（agent）的社交行為推測出該代理人對環境中其他四個人的友好程度。然而在更加複雜的社交網路中，此模型的表現仍然未知。在觀察富有動態變化的社交行為時，模型需要整合更多層面的資訊去推測社交網路的結構。在這篇研究中，我們擴展了機器心智理論模型，擴展後此模型具備能夠解讀更動態的五位代理人組成的社交行為。我們設定ㄧ組社交網路並定義移動規則，使五位代理人遵照移動規則於24x24的網格世界中自由移動，並將移動之最後10個步驟（包含最終位置）記錄起來，用於訓練此機器心智理論模型。擴展模型的過程中，為了符合訓練資料的需求，我們將模型原先的張量（tensor）從12\*12\*11\*10擴大為12\*12\*31\*10，深度增加至31是為了正確傳遞代理人在網格世界中的移動資訊。我們使用了具有五位代理人但未被訓練的網格世界做為測試模型學習的結果，並使學習後的模型輸出對五位代理人最終位置的預測。為了評估模型學習的效果，我們註記每一組預測結果中，具有相鄰代理人的情形，並疊加了所有1000個測試網格世界的預測結果。另外，我們將此模型同步測試於已學習社交網路的測試網格世界資料，以及另外100組未被此模型學習的社交網路的網格世界資料，此模型可以正確解讀與分辨所學習的社交網路與另外100組社交網路的不同。這些結果顯示，此機器心智理論模型得以解讀抽象而且更加複雜的動態社交網路關係，並有潛力更廣泛應用於社交機器人（socially assistive robots）。

關鍵詞: 人工神經網路，社交網路，心智理論，社交機器人，多代理人

Abstract

Social networks define the structure of human inter-relationships. Critically, such structures are abstract and hidden such that to learn about them requires the ability to infer hidden states from observations of social preference behaviors. Human beings do this very naturally in using inferred social knowledge for smooth daily communication and interaction with each other. A key challenge in human robot interaction (HRI) is to imbue machines with similar ability to engage with humans naturally. In this study, we consider that in order to design machines that interact effectively with humans, the artificial intelligence driving these machines should also be able to infer social networks from observations of human social behaviors. We previously designed a Machine Theory of Mind (ToMNet+; modified based on principles applied in Google DeepMind projects) that displayed the ability to observe social interaction behaviors and infer simple human social networks consisting of one agent with distinct socializing preferences for four targets. However, how this model performs for more complex social structures remains unclear. Critically, more dynamic social structures driving social behaviors requires the model to incorporate more dimensions of information towards inferring social networks. In this present study, we expanded ToMNet’s structure and built ToMNet 2.0. ToMNet 2.0 has the capability to infer more complex human social networks consisting of five dynamically interacting agents each with its own set of social preferences for the other four agents in the social network. Such a social preference setting drives all five agents to move dynamically in 24 x 24 grid world over 10 time steps based on a movement formula, which is then utilized as ToMNet 2.0’s training data. To meet the social interaction information dimensionality of this dynamic grid world data, we adjusted ToMNet 2.0 to accept input tensor sizes from 12\*12\*11\*10 (in the original model, 11 channels coded the 5 agents positions and barriers) to 12\*12\*31\*10, the increased depth of the tensor size to 31 conveys the additional information needed to code the dynamic moving agents. At test, ToMNet 2.0 evaluates the start states of agent locations in novel test grid worlds and predicts the final locations of the five agents relative to each other. To evaluate ToMNet 2.0’s performance, we summed the number of times agents clustered with others at the final states in 1000 start state permutations. We compared ToMNet 2.0’s performances for trained social networks and 100 untrained permuted social networks against the ground truth final state determined by the movement formula (details provided in methodology). These results show that ToMNet 2.0 is able to dissociate the hidden social preferences of the five agents from the 100 untrained social networks. Thus, ToMNet 2.0 is able to utilize additional social preference information compared with ToMNet, which raises the potential for ToMNet 2.0 to be applied in socially assistive robots problems more broadly.

Keywords: artificial neural networks, social networks, Theory of Mind, socially assistive robots, multi agent

Contents

[致謝 i](#_Toc142566592)

[中文摘要 ii](#_Toc142566593)

[Abstract iii](#_Toc142566594)

[Contents v](#_Toc142566595)

[List of Figures vi](#_Toc142566596)

[List of Tables vii](#_Toc142566597)

[Introduction 1](#_Toc142566598)

[A. Significance & Motivation for Improving Social Machines 1](#_Toc142566599)

[B. Current Implementations of Social Machines 2](#_Toc142566600)

[C. Toward Theory of Mind in Machines 4](#_Toc142566601)

[D. ToMNet+ 6](#_Toc142566602)

[E. ToMNet 2.0 8](#_Toc142566603)

[F. Hypothesis 9](#_Toc142566604)

[Methodology 11](#_Toc142566605)

[A. The social game for simulated agents 11](#_Toc142566606)

[B. Projecting social interaction scenario into grid world 12](#_Toc142566607)

[C. Grid collection to simulation data 16](#_Toc142566608)

[D. ToMNet 2.0 model structure 17](#_Toc142566609)

[Experimental Results 20](#_Toc142566611)

[A. Evaluating ToMNet 2.0 performance 20](#_Toc142566612)

[B. ToMNet 2.0 outputs and results on training data 22](#_Toc142566613)

[C. Random rate comparison 27](#_Toc142566614)

[Discussion 30](#_Toc142566615)

[A. Implications 30](#_Toc142566616)

[B. Application on human social interaction data 31](#_Toc142566617)

[C. Limitations 32](#_Toc142566618)

[D. Conclusions 32](#_Toc142566619)

[References 33](#_Toc142566620)

List of Figures

[Figure 1 12](#_Toc142565143)

[Figure 2 15](#_Toc142565144)

[Figure 3 19](#_Toc142565145)

[Figure 4 21](#_Toc142565146)

[Figure 5 23](#_Toc142565147)

[Figure 6 26](#_Toc142565148)

[Figure 7 29](#_Toc142565149)

List of Tables

[Table 1 24](#_Toc142565308)

Introduction

1. Significance & Motivation for Improving Social Machines

Improving Human Robot Interactions (HRI) for the purposes of utilizing machines as social companions, such as social robots, in human daily life is of great interest and importance across many technological research areas and applications [1]. A general consensus is that a *social* machine should have the ability to communicate messages or engage cooperative actions with other agents and also demonstrate autonomy in its choices of messages or actions [2]. This is a theoretical ideal that is extremely difficult to practically realize in a machine. Nevertheless, it is a required goal if one is to proliferate social machines effectively in normative human society; for example, in the cases of robots assisting older adults with physical or psychological needs or having robots as part of the users’ homes performing daily tasks [3].

Imagine the scenario of an older adult living in a family while other members are often not immediately present on site. A companion robot who has the capacity to be always present on site and has the ability to effectively communicate broad conversation topics may be an appreciated psychological assistance for the older person’s needs. This might also have the effect of reducing the influence of age-related neurocognitive degeneration in the older adult and thereby alleviate caregiver burdens on the family [4]. By contrast, companion robots without effective social capacities may potentially cause conflicts between family members or friends. For example, if the robot is unable to perceive underlying unhappiness between two family members, and continues to engage in conversation with one of them about the other's situation, it is likely to create a barrier in that family member's interaction with the robot. Thus, the ability to interpret the underlying reasons for observed overt human social interaction behaviors is a critical function required in social machines. This notion forms a driving application motivation for this present study to investigate and improve the ability of an artificial neural network model algorithm in implementing social processing in machine contexts.

1. Current Implementations of Social Machines

At this point, it is important to highlight a distinction in the field between social machines and socially assistive machines [5]. Several studies have focused on the ability of machines to process social signals and engage social actions, and evaluated how these abilities affect the machine’s efficacy in forming social connections with humans. Such social machines also seek to meet social goals, for instance, by continuously interacting with users and responding properly to the environment [5]. In general, social machines are primarily trained to engage outputs that are mapped (albeit using various sophisticated methods) to input feature-level indices derived from data on eye gaze, facial expressions, gestures, movements, and speech. For example, [6] reviewed several studies showing how robots that expressed emotions in a dynamic manner were more accepted by humans than robots expressing emotions in a more static manner. These findings suggest that social machines with richer emotion input and output mappings are better at achieving social goals. Moreover, robots that exhibited more complex emotional repertoires via mimicry of participants’ expressions and gestures also engaged more successful interactions with patients [7]. Finally, we also note the CARESSES project, in which robots were embedded with cultural competence in their verbal behaviors (i.e., mentioning culturally specific conversation topics or terms) to improve their social ability and as a result reduced user loneliness [8].

Thus, the above studies have shown that social robots with the abilities outlined above in the context of experimental or intermittent usage do already help to improve mental statuses in participants experiencing loneliness, depression, or stress [9, 10], fulfilling application goals and societal needs. However, [10] reported various barriers to the use of social robots in normative life. A key limitation is that social robots, despite being quite advanced, are still far from being ready to be implemented for example in people’s homes. The critical barrier is that in everyday settings, compared to the experimental or controlled and limited contexts where social robots have been evaluated, there are in fact much more complex and dynamic interactive behaviors that need to be incorporated and interpreted for social machines to be autonomously deployed.

Critically, in this present study, we consider then that to be socially *assistive* machines, the machine cannot merely map input emotion features to output actions. That is, they cannot only be good emotion feature signal processors (albeit very powerful ones) or good action actuators (albeit very life-like). They need to have an internal representation model of the external information structure that drive human social behaviors to match the true complexity of human social experience. To this end, we suggest that the social machine requires Theory of Mind to be truly socially assistive [11]. This forms the conceptual motivation driving this present study.

1. Toward Theory of Mind in Machines

Theory of Mind (ToM) is the ability of an agent to internally instantiate a model of other agent’s beliefs or intentions (or mental models), and is a key cognitive ability of at least higher order living organisms such as humans. Several human Theory of Mind (ToM) aspects are believed to be developed by the age of 5 years, such that ToM in childhood is a critical component to allow children to attribute mental states to others that aids the child’s social skills, important for normative developmental societal integration and interaction [12, 13]. Similarly, developing machines that have the capacity for ToM has become increasingly important and essential when interactions are growing more elaborate between machines and human [11]. However, [14] reflects that there is a the lack of trust in humans for machines, which might be caused by current models focusing on reproducing exhibited behavior (i.e., input-output mapping) rather than being driven by an internal model of others’ mental states (i.e., ToM). Thus, imbuing machines with ToM stands as a critical barrier to enhancing social connections between human and machines or robots.

Models based on multiagent deep reinforcement learning (MDRL) architectures are potential candidate algorithms to feasibly embed ToM in machines. A survey of recent multiagent learning (MAL) [15] points out Machine Theory of Mind [16] could be classified as a series of multiagent deep reinforcement learning (MDRL) models that have the ability to perform agent-modeling-agents functions. These models are inspired from either deep reinforcement learning (DRL) or MAL structures. Examples of early related work include the Deep Reinforcement Opponent Network (DRON) and Deep Policy Inference Q-Network (DPIQN), which are models that learn their action policies through observing opponents or teammates. Other algorithms in this category include the Self Other Modeling (SOM) approach that estimates the opponents goals in the agent’s self policy, the Neural Fictitious Self-Play (NFSP) and Deep Cognitive Hierarchies (DCHs) that combine game theory and Multiagent learning (MAL), the Minimax Multiagent Deep Deterministic Policy Gradients (M3DDPG) approach that applies the concept of minimax in game theory and MAL, the Learning with Opponent-Learning Awareness (LOLA) approach that anticipates the updated policies of other agents to maximize reward, and the Deep BPR+ approach inspired by BPR+, which is a Bayesian policy reuse framework with policy distillation. The above mentioned MDRL models do somewhat involve agent-modeling-agents functions, however there are still gaps in these approaches to ToM ability in machines.

To this end, Deep Bayesian Theory of Mind Policy (Bayes-ToMoP) applied theory of mind concepts in agent-modeling-agents tasks. Bayes-ToMoP has a similar setting to BPR+ and is said to have the ability to process high-level reasoning strategy and also better self-playing performance. However, Bayes-ToMoP is trained based on the reward of the opponent, constructed as a value based DQN model design [15], which may limit the implementation of social scenario applications.

It has been suggested in [15] that models focusing on recursive reasoning approaches may be a key factor in effective models with machine Theory of Mind to represent the mental states in others. However, the use of recursive learning usually requires numerous training data. Critically, machine Theory of Mind models have been ascribed as a meta-learning models [16]. Indeed, meta-learning is also regarded as a critical feature of mature machine ToMs [11]. Taken together, meta-learning algorithms which are designed to learn from a limited amount of training data, may be candidate models to continuously learn and update concepts in contexts of being embedded in human social interactions.

1. ToMNet+

The original ToMNet model as described by the Google DeepMind team was applied in grid world simulations. The model can be divided into two parts, the learning (character net) and predicting (prediction net) parts. The character net picks up agents’ movement interaction behaviors provided by grid world simulations as inputs. Subsequently, the prediction net predicts the outcome agent positions or movements given the information acquired by character net and the initial positions of novel grid world start positions of agents. This algorithm structure is quite reasonable if we regard the prediction net as reducing entropy by learning some sort of information about the agents embedded in the learned simulations.

ToMNet has achieved comparable human level performance when tested on its capacity to engage theory of mind functions, such as in tests involving partial observability and the classical Sally-Anne test of false belief mental models in others [16]. These test performances reveal ToMNet’s difference in comparison to the previously mentioned algorithms, and highlights the possibility for implementing ToMNet in social machines. Several different variations of ToMNet applications have emerged due to it’s potential in enhancing machine social skills. These include testing ToM models on Few-shot Language Coordination [17], testing ToMNet’s performance in comparison to infants [18]. However, ToMNet has not yet been trained or tested in simulations with embedding of data on social interaction behaviors between agents driven by the agents’ preferences [16]. Thus, there remains a knowledge gap for demonstrating if ToMNet function can be used to a certain degree to represent the internal underlying mental states of others or some hidden information structures purely through observed behavioral interaction inputs.

To bridge this gap, in ToMNet+ [19], we applied a modification of ToMNet [16] onto social scenarios designed to evaluate the ability of such algorithms to identify embedded social preference information structures. Our ToMNet+ study utilized two different sets of training data, including the simulation data and real human generated data. Each data set had the same format of input grid worlds with an agent approaching one of four targets, which were scattered in various locations in each grid world instance. Distance travelled was applied as a penalty such that ToMNet+ essentially learns associative functions that weigh the cost of approaching different targets in grid world against the social utility of approaching each target. Critically, the targets are construed as the central agent’s relatives with different social preference weightings (i.e. the central agent has a social preference score with respect to each target agent). In the simulated grid worlds, the agent was triggered to move by calculating the difference between social scores and the path distance (i.e., the minimal number of steps to target agents). We showed that ToMNet+ was able to recognize social preferences embedded in the grid world movement behaviors both in simulations and human generated data sets. However, one severe limitation was that the grid world stimuli restricted ToMNet+ to gaining only unidirectional social information. That is, ToMNet+ was only trained on scenarios where one central agent had social preferences for four other target agents. Thus, it is still not clear whether ToMNet+ can be adapted to learn bi-directional social preferences between the five agents involved in the grid world simulations. This is important as it reflects more ecologically valid situations in daily human social life.

The results in ToMNet+ stands as proof-of-concept for implementing simple social interaction structures in relatively similar simulations as in ToMNet [16]. From here, therefore, we can expand a range of evaluations and broad applications. In this present research, we expanded the stimuli settings to test the performance of ToMNet 2.0 to distinguish more complex social structures that are closer to real life.

1. ToMNet 2.0

In this present study, our purpose is to evaluate if a modification of ToMNet+, ToMNet 2.0, is able to similarly predict social behavioral outcomes from observed social interactions constructed based on more complex, dynamic, and ecologically valid social structures. Based on this reasoning, we first set our embedded social structures such that all agents in the social network maintained bidirectional social preferences. i.e., each agent has different preferences for the other four agents. In addition, we defined the interactive movements of these agents in grid world simulations as five-bodied problems where the five agents moves dynamically akin to bodies orbiting each other due to their gravitational forces acting on each other differently at different time steps. i.e., from the start states, the agents move based on the resultant weightings of the influences of the distances between them and their social preferences for each other. Subsequently, the agent positions are updated and the distances between them have changed, driving a new resultant weighting that effects the next set of movements, and so on. We decided to use a five agent case because this is the minimal group size in which it is possible to have variable sizes of social clustering of agents (1 vs. 4, 2 vs. 3, 2 vs. 2 vs. 1, etc.), which increase the richness of the training data. Moreover, we extended the simulation grid size to provide different information inputs (see Methods). Overall, the key goal in this study is to evaluate if ToMNet 2.0 has the ability to observe the grid world social interaction behaviors of five agents, which are driven by bidirectional social preferences applied in a dynamic manner, and subsequently predict how the agents will behave in novel grid world situations.

1. Hypothesis

Formally, we hypothesize that ToMNet 2.0, which is augmented from ToMNet+ in terms of the ability to accept increased dimensions of input information, after being trained on variable observations of how five agents interact in grid world, will be able to produce predicted outcomes of how five agents will interact in novel untrained grid world scenarios. The requisite for this ability is that the grid world interactions of the five agents stem from a real underlying driving social network structure which is hidden with respect to ToMNet 2.0. In this manner, ToMNet 2.0 can be said to maintain an internal representation model of external hidden information structures — having a form of Theory of Mind. If successful, ToMNet 2.0 might bring us one step closer towards a more effective socially assistive machine.

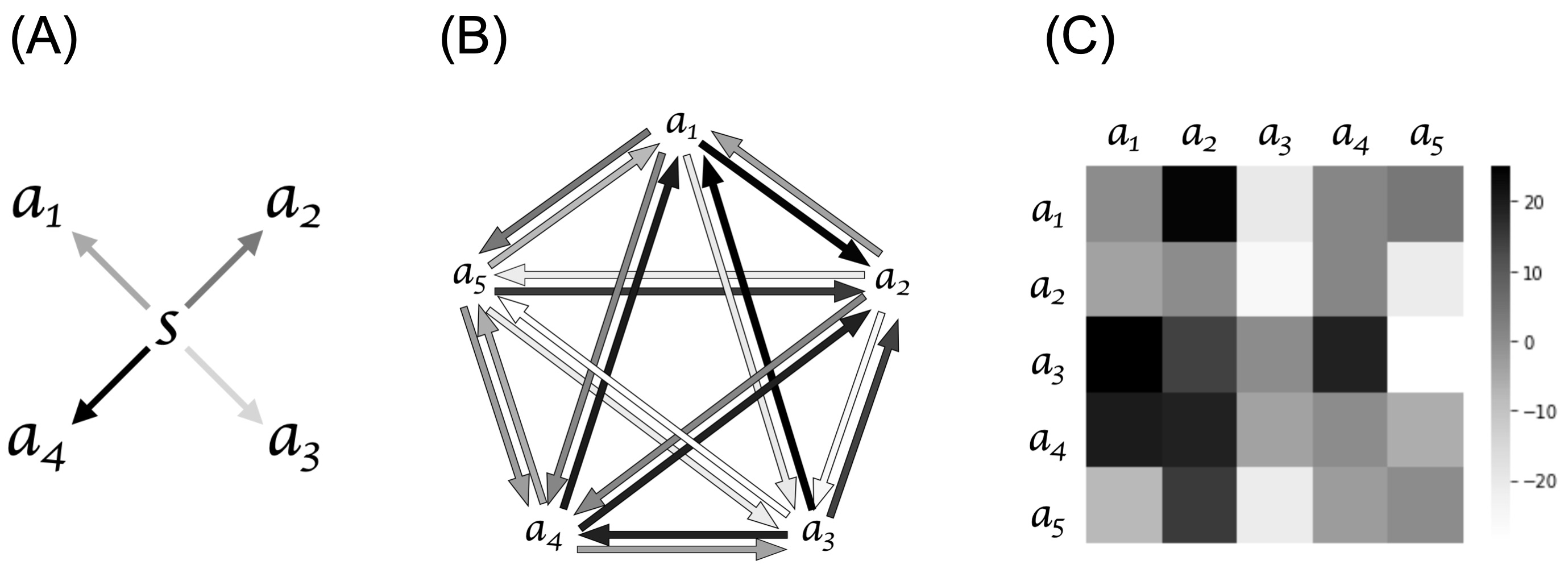
Methodology

1. The social game for simulated agents

In our simulation approach for agent-to-agent interactions, social actions are driven by embedded social structures. This principle was applied in our previous work on ToMNet+ [19], which was modified from Google DeepMind’s ToMNet model [16]. In that study, a given agent, *s*, had different social utility scores for other four goal agents, *a1*, *a2*, *a3*, *a4* (Fig. 1A). These agents were then placed in various locations in a gridworld and then the agent, *s*, moved towards one of the goal agents based on the social utility score and path distances. The agent movement loss functions are detailed below. This setting, however, instantiates a mono-directional social structure representation of one subject agent’s preferences toward four goal agents.

In this present study, we implemented dynamic 5-agent interaction simulations and applied ToMNet 2.0, which is modified from ToMNet+ to evaluate the models ability to represent more complex and ecologically valid social interactions. Social scores were applied to 5 agents (*a1*, *a2*, *a3*, *a4*, *a5*), with each agent having a randomized social score for the other four agents (Fig. 1B & C). The social score range was set at [40, -40] for arbitrary simulation purposes. Other score ranges could be set in subsequent studies. We will focus on our current setting here with some further range inferences described below in section B. Further, the means of total social scores differ for each training set used such that the means were set at 0, 16 and -8 in three different data sets. These different mean social scores between training set mimics social networks in which the net social preference is neutral, attractive, or repulsive, respectively. The given social scores along with the distances between agents in the grid then affects all agents’ movements in a manner described in section B. This social score setting is preserved through each entire simulation data set. Thus, the main difference between social structures used in ToMNet+ and ToMNet 2.0 is that the social preferences fed to ToMNet 2.0 are bidirectional with all five agents having both efferent and afferent social preferences with each other.

Figure 1 (A) represents the ToMNet+ social structure that projects S has different social preference between a1, a2, a3, and a4. The scale of grey arrow indicates the preference score of S for the other four agents. In Figure 1 (B) the bidirectional arrows refer to agents having different social scores for each other. The scale of the grey arrows represents the social scores that are also shown in Figure 1 (C), with white indicating social scores at -30 and black indicating social scores at 30. This 5-agent social structure has a neutral net social preference (equal balance of attraction and repulsion).

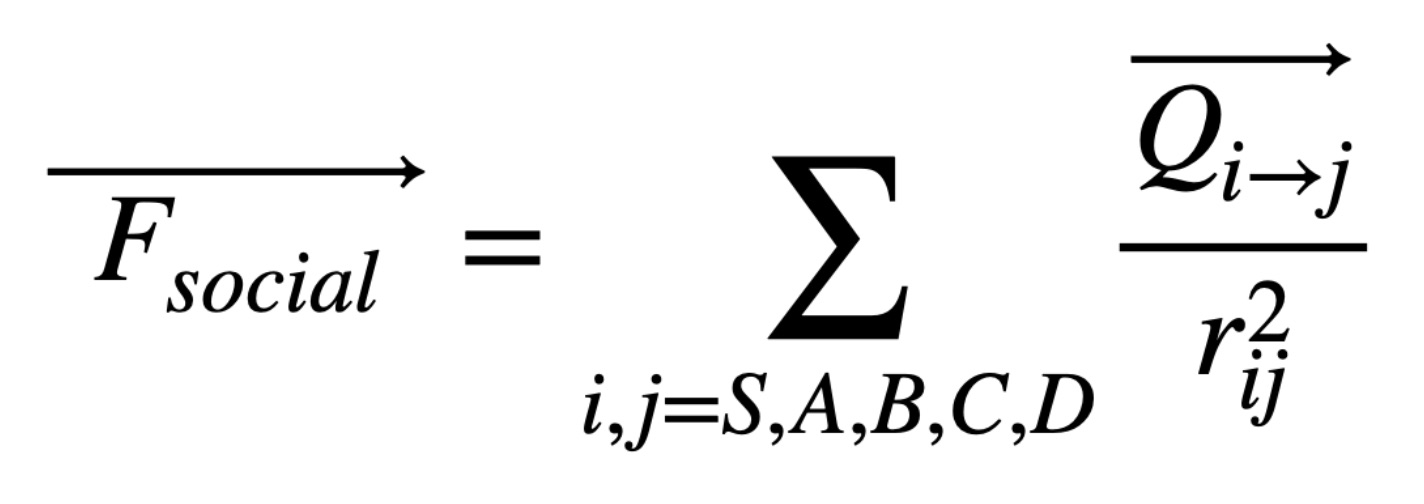


1. Projecting social interaction scenario into grid world

Grid world simulations that contained agents' movements were used to represent humans social interaction structure, this is an application of allowing topological maps capturing various information relationships, which in this case was social network information[20].

In the grid world simulations, we set up a 24 x 24 grid world with a 12 x 12 focal area in the central. ToMNet model observes the 12 x 12 focal area, while the 24 x 24 grid contributes information and continuous effect of agents dynamic movements that is not observable for ToMNet. In each simulation, four agents were randomly located in the central (12 x 12) area, and moved freely in the grid (24 x 24) later on, with one agent fixed in the center of the grid (Fig. 2A). The fixation on one of the agents was a practical implementation to generally keep most of the agents’ migrations in the 12 x 12 grid field of view rather than constraining the migrating flow of agents within some boundaries. If this were not applied, the centroids of agent locations may migrate out of the focal area while converging. This may lead to some simulation trials possibly having no agents within the 12 x 12 grid only after a few movement iterations. All five of the agents were sequentially fixated on in a counterbalanced manner across all simulation trials.

In each simulation trial, once the agents have been located in the grid world, the social preferences drive the movement of agents (*a1*, *a2*, *a3*, *a4*, *a5*) in the grid. With the starting state as 0, n represents the state after *n* steps of movement. From n to n+1 we calculate the movement force (*Fsocial*) for every agent in the n position. Calculating the movement force (*Fsocial*) includes the information of distance (*r*) between each agent at n and the social score (Qij). The movement force (*Fsocial*) on n is then calculated based on the formula stated below which, for those interested who notice, is similar to the electrostatic force that follows Coulomb’s law:



Given the movement force for each agent, we then determined all the agent’s movement together at n to n+1 (Fig. 2). *Fsocial,* which is the sum of movement forces from all other agents on one specific agent, is a vector that indicates the movement direction of the agent in the grid. Note that because the grid is discrete rather than continuous, we applied approximated motion directions such that if a *F*social unit vector was less than 0.5, the agent does not move. If it was over 0.5, the agent was triggered to move in the resultant direction by one unit in the grid. At times, we observed the phenomena of agents orbiting each other; for example *ai* and *aj* orbiting each other, leading to the result of *ai* and *aj* holding the same coordinates in the grid on n and n+1. We set flags to interrupt such infinite orbiting when occuring, since this continuous term does not provide further social information in our simulation.

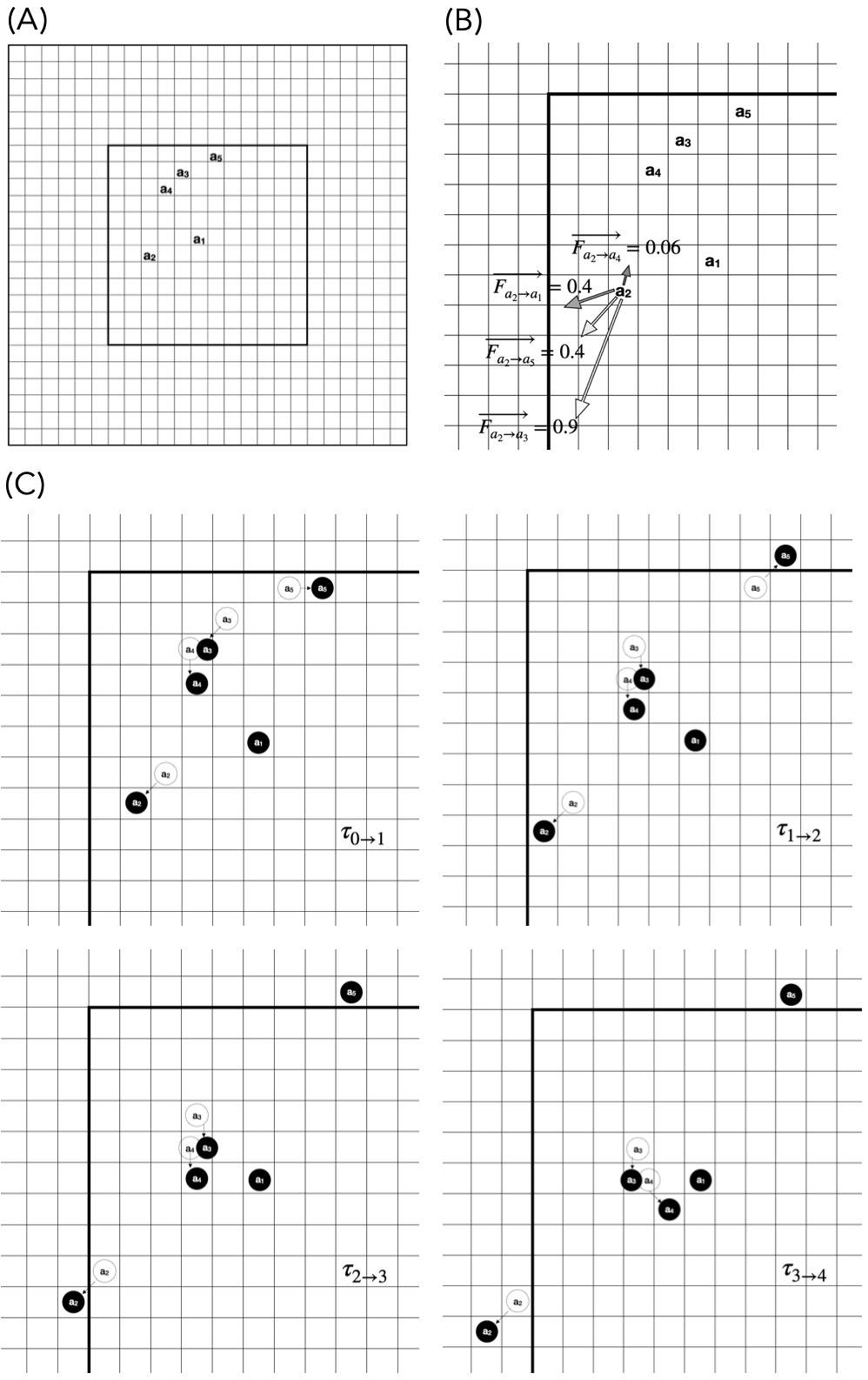


Figure 2 (A) shows one simulation trial as time step 0 , which represents the initial state of agents. Figure 2 (B) is an example of calculating a2’s Fsocial. The arrows labeled with different represents the movement forces a1-5 acting on a2. The arrow’s grey scale represents the social score and has the same scheme as in Figure 1 (B). Due to the sum of Fsocial, a2 moves both left and down in 1. In Figure 2 (C) the unfilled circles represent the agents’ places at , while the black circles represent the agents’ places at . This figure displays how the agents move in the 24\*24 grid from .



1. Grid collection to simulation data

The simulation data is constructed by agents initial positions and the agents movements in the grid. The initial positions of the agents in each grid were set randomly. With the random starting positions and the rule of movements described above, a set of data was generated based on the social structure between agents.

Each training set of this model was composed of 10,000 different grids as a set of simulation data. Grids were compiled into tensors for model input, with each tensor (10\*trajectory) including the final state and the last 10 movements of the grid. If there were less than 10 movements, then the first few steps would be zero padded. In set 1 (social score mean = 0) of the three training sets, 6936 of 10000 tensors were zero padded; in set 2 (social score mean = 16), 9894 of 10000 tensors were zero padded; in set 3 (social score mean = -8), 8747 of 10000 tensors were zero padded.

Each tensor was constructed from 10 trajectories (12\*12\*31), while the trajectories represented the movements in the grid. The initial position was first recorded in 5 layers, then the movements of the five agents were recorded by up, down, left, right, and goal separately. This results in 5 layers for each agent and a total of 25 layers to record all agents’ movements. An additional layer which is recently all zero padded allows further implementation for adding obstacles into the grid. However, this is not currently of use in this present study.

1. ToMNet 2.0 model structure

ToMNet 2.0 model is trained by batches, based on [16] setting the batch size as 16, 16 sequences of randomized tensors are input jointly to the model. ToMNet 2.0 consists of two separate parts (Fig. 3). The first part is the character network where the input consists of the tensors of simulation data channels (e.g. directional motions of each agent at each time step) mentioned above. The output is the character embedding (echar), which is an abstraction of the information given in the input tensors. The character network first inputs the tensor (10\*12\*12\*31) into a 1-layer 3x3 convnet, expanding the tensor into 10\*12\*12\*32, then passes through a 5-layer resnet with 32 channels, batch normalization and ReLu nonlinearities. The resnet is followed by average pooling and leads dimension reduction into 10\*32. The 2-D tensor then passes through LSTM with 1024 channels and a fully connected layer, this returns a linear output that represents the category of all the output possibilities (echar). Note that the channel amount and matrix dimensionality mentioned above those set as 1024 denotes the possible sets of result for the grid output.

The other part of the ToMNet 2.0 model is the prediction network. This is a parallel structure to character network that predicts the result of untrained grids. The input of prediction network is echar from character network, and an untrained grid named query state (qk), this grid contains only the initial state but without the moving steps (12\*12\*31). The echar is spatialized before concatenating with query state (qk), then passed through a 1-layer 3x3 convnet, transferring the tensor into 12\*12\*32. The tensor is then passed into a 5-layer resnet with 32 channels having the similar setting in character network. Followed by a 3x3 convnet and average pooling layer the dimension is reduced into 32\*1, after the fully connected layer and softmax the linear output indicates the prediction of qk, which is labelled as gk in Fig. 3. The output gk after the softmax is a probability tensor with size of 1024\*1 indicates the prediction output of the untrained grid.

According to [19], the model’s training, validation, testing, proportion is 8:1:1, and was trained by applying Adam optimizer, leaning rate = 10-4, and training set = 104. The model works under the environment of python = 3.6.10 and tensorflow = 1.12.0.

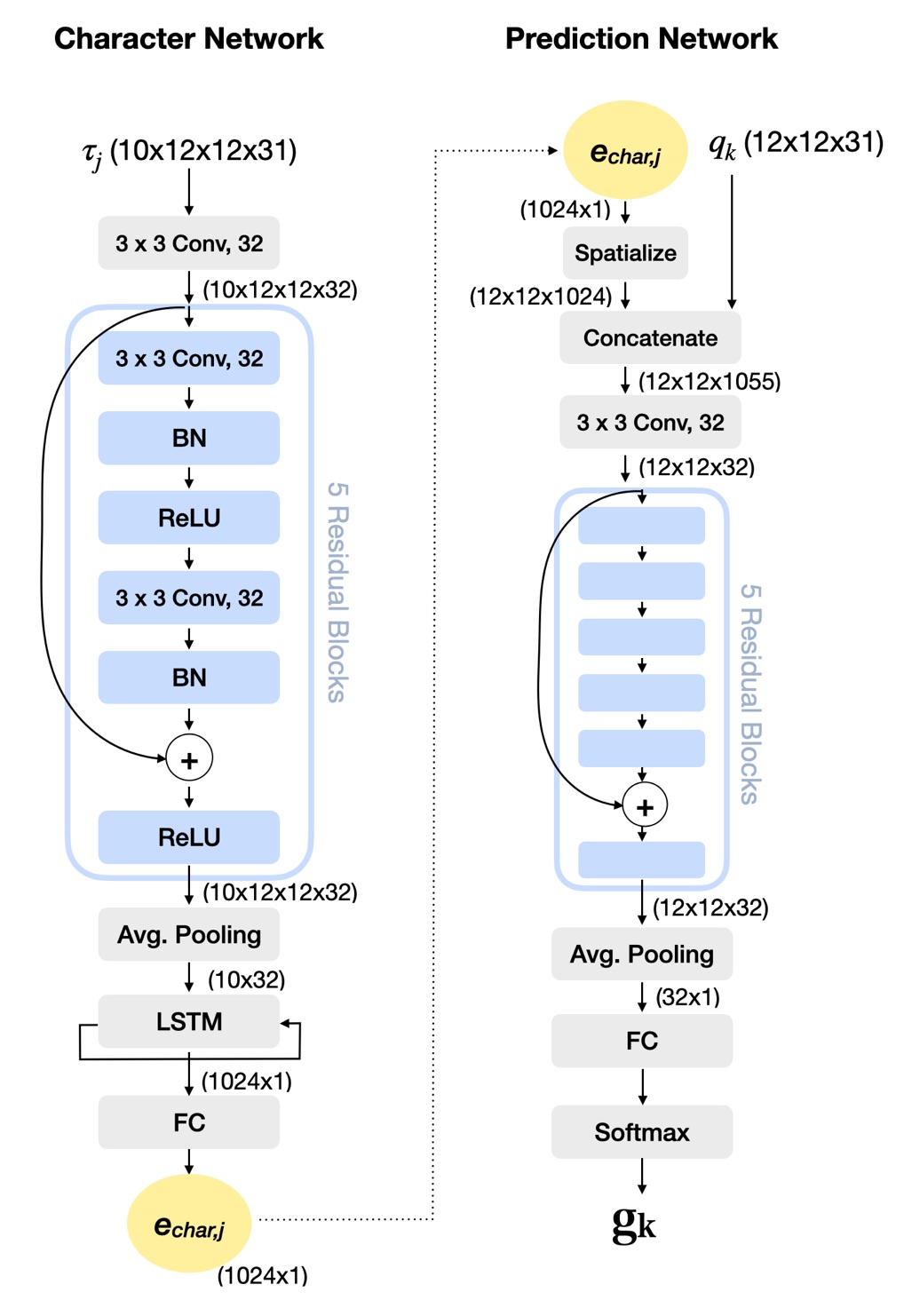


Figure 3. A schematic structure of ToMNet 2.0. The parentheses in the figure indicates the tensor size. Conv: convolutional layer, BN: batch normalization, ReLU: rectified linear unit activation function, Avg. Pooling: average pooling layer, LSTM: long short-term memory layer, FC: fully connected layer, : input trajectory, echar,j: character embedding, qk: query state, gk: prediction result.



Experimental Results

1. Evaluating ToMNet 2.0 performance

The performance of ToMNet 2.0 is evaluated by two different aspects, one examines the model performance by learning curve, the other focus on the prediction result of untrained 1000 grids to evaluate ToMNet 2.0’s performance. Evaluating model’s performance by the learning curve is a common measure in machine learning and will be further elaborated in section B.

Evaluating ToMNet 2.0’s achievement on learning hidden social network structure, we extracted various testing results of ToMNet 2.0 after training in predicting the outcomes of 1000 untrained grids. Considering that grid world simulations essentially index the spacial relation information between the five agents (ai), we measured the model’s learning performances by recording the ‘contact or not’ instances. Such instances were noted down as 1 when the specific two agents settles clustered (in adjacent coordinate locations), while a recording of 0 was logged when the final settling locations for two agents were not adjacent, e.g., a1 and a4 settled with contact = 1 in Fig. 4A. With this setting, each grid could be transformed into a 0 or 1 10 binary digit table contains 10 integers (Fig. 4B & C). We also apply this 10 binary digit table format to the ground truth, for further comparison between ToMNet 2.0’s prediction. Here the ground truth stands for obtained by implementing letting the grid world movement force algorithm play out based on the same start locations of the 1000 untrained grids.

The format of 10 digit table representing each grid terminal state is then applied to various different evaluation. We first summed all 1000 untrained trial ToMNet 2.0 outputs of 10 binary digits table into a single 10 integer vector counting table as our testing results, by summing up the 1000 grid’s prediction result, we get to evaluate the trends of agents adjacent frequency rather than only comparing individual grid prediction with ground truth result. The ground truth grids results are also summed up and the comparison results are further discussed in section B. The 10 integer vector counting table is then applied to measure the RMSE between prediction and ground truth which is further discussed in section C.

We also calculated the accuracy by matching the 10 digit table of prediction with the ground truth. If the 10 digit table of each single prediction matches the ground truth the correctness = 1. Summing up the correctness of all 1000 grids and divide into 1000 equals the accuracy of the prediction set.

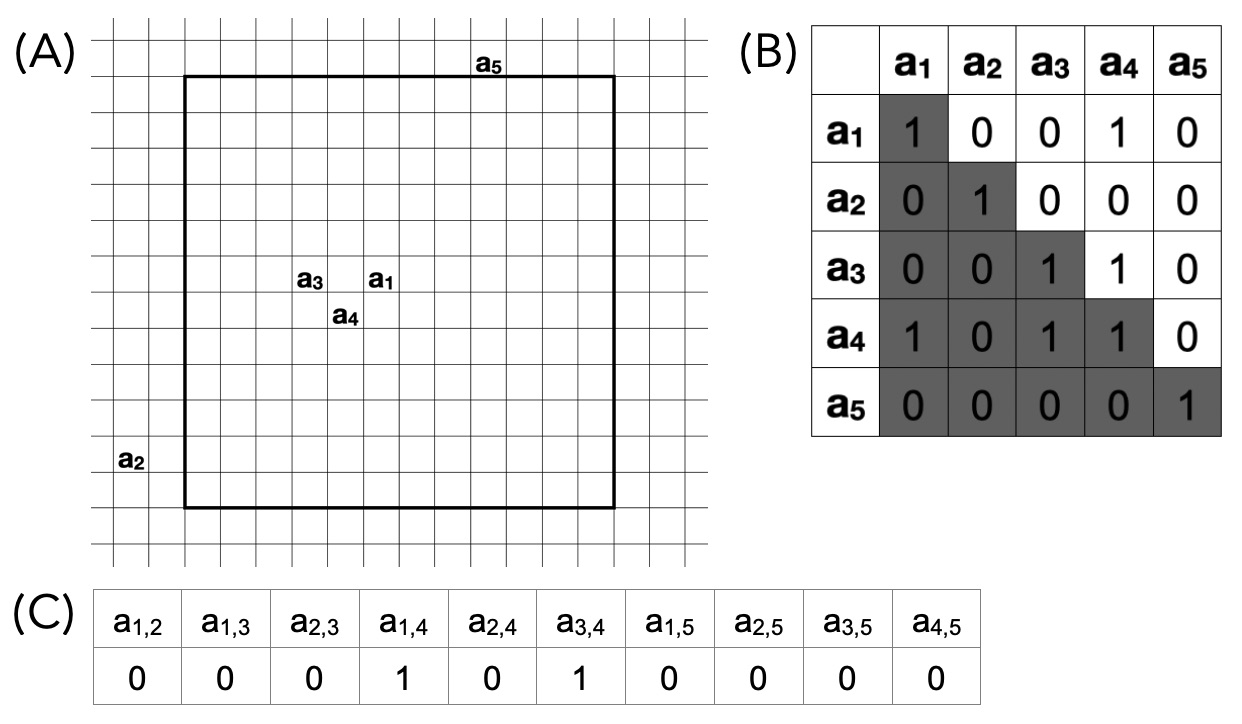


Figure 4. In (A) shows the final state of this grid. (B) The values of the table expresses the convergence of agents, 1 stands for settle with clustered, 0 on the contrary. Both a1 and a3 neighboring a4, and are marked 1 in the table. (C) Then the 10 values are transformed into a 0 or 1 10 binary digit table.

1. ToMNet 2.0 outputs and results on training data

In this study, we trained ToMNet 2.0 separately with three different social preferences set as mentioned in the methodology section (Fig. 5). Following the evaluation format mentioned in section A, table 1 shows the resulting 10 integer vectors obtained from ToMNet 2.0’s predictions for the start locations of the 1000untrained test instances for the three different data set types, comparing with ground truth. Overall, the trends suggests that ToMNet 2.0 did approach having an internal belief of the complex social structures governing the five agents’ movements in grid world. To formally assess this, we used Spearman rank order correlations, which is suitable for comparing ToMNet 2.0’s testing results against the ground truth. A Spearman rank order correlation is performed to test the monotonic similarity of two variables.

The results of training set 1 is stated in Table 1A, this set of simulations has the mean social score = 0 with SD = 9, min = -16 and Max = 18. The social preferences were learned by ToMNet 2.0, which reflects the models ability to decode the embedded social preference in such distribution. A Spearman rank order correlation was further performed according to the results of Table 1A, and revealed a significant correlation with ground truth, the results were = 0.87, p = 0.001.



In sets 2 & 3 it separately informs two different extreme conditions. For set 2, the agent’s social score mean = 16 with the SD = 12, min = 0 and Max = 34, in the grid simulation this leads to agents high frequency settling in contact, while in real life projects a group of people without negative social preferences and who all are attracted to each other. A Spearman rank order correlation was further performed according to the results of Table 1B, which shows significant correlation with ground truth, the results were = 0.75, p = 0.011.



For set 3, the agent’s social score mean = -8 with the SD = 10, min = -25 and Max = 5, this reflects a situation where agents rarely contact in this set of simulation, in real life indicates a group of people who rarely interact. A Spearman rank order correlation was further performed according to the results of Table 1C, and shows significant correlation with ground truth, the results were = 0.9, p = 0.0004.

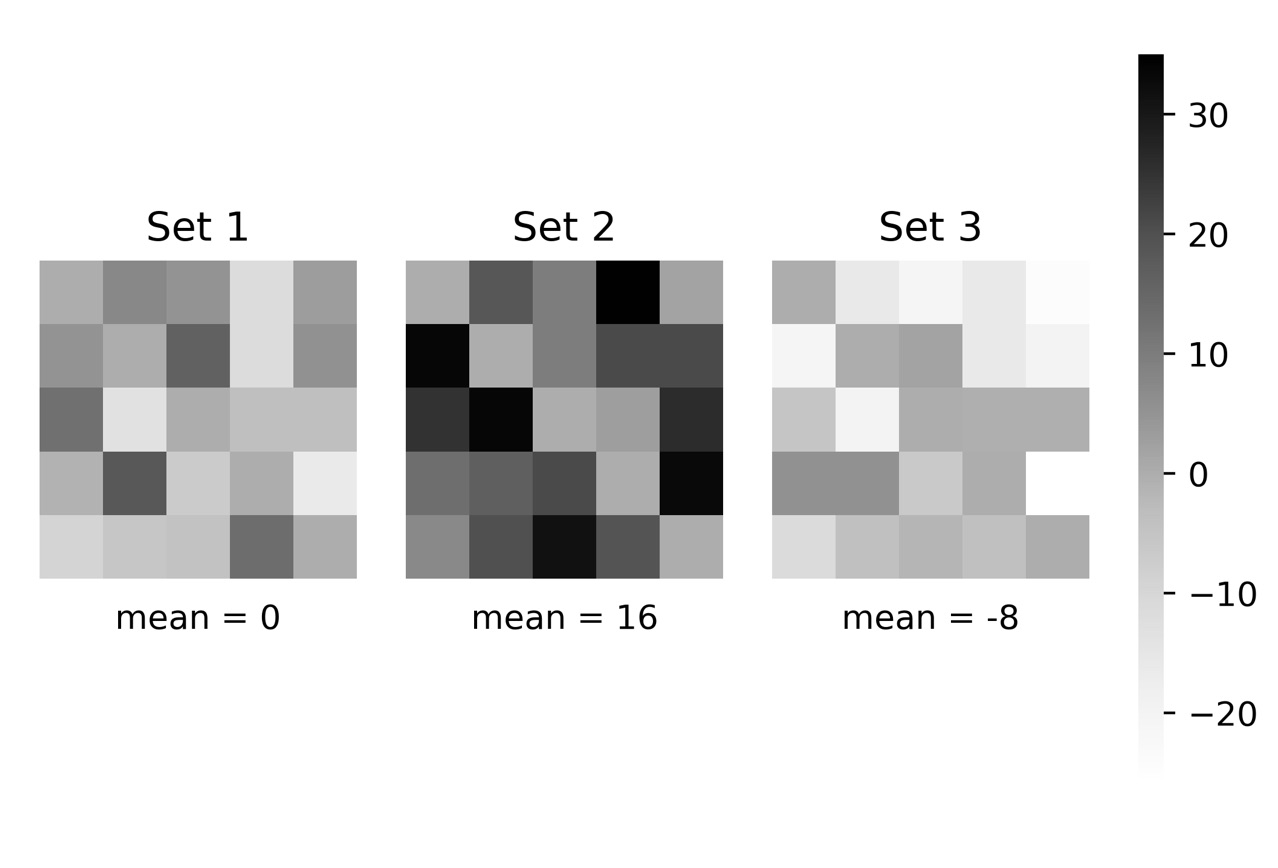


Figure 5. The social preference heat map of the three training sets.

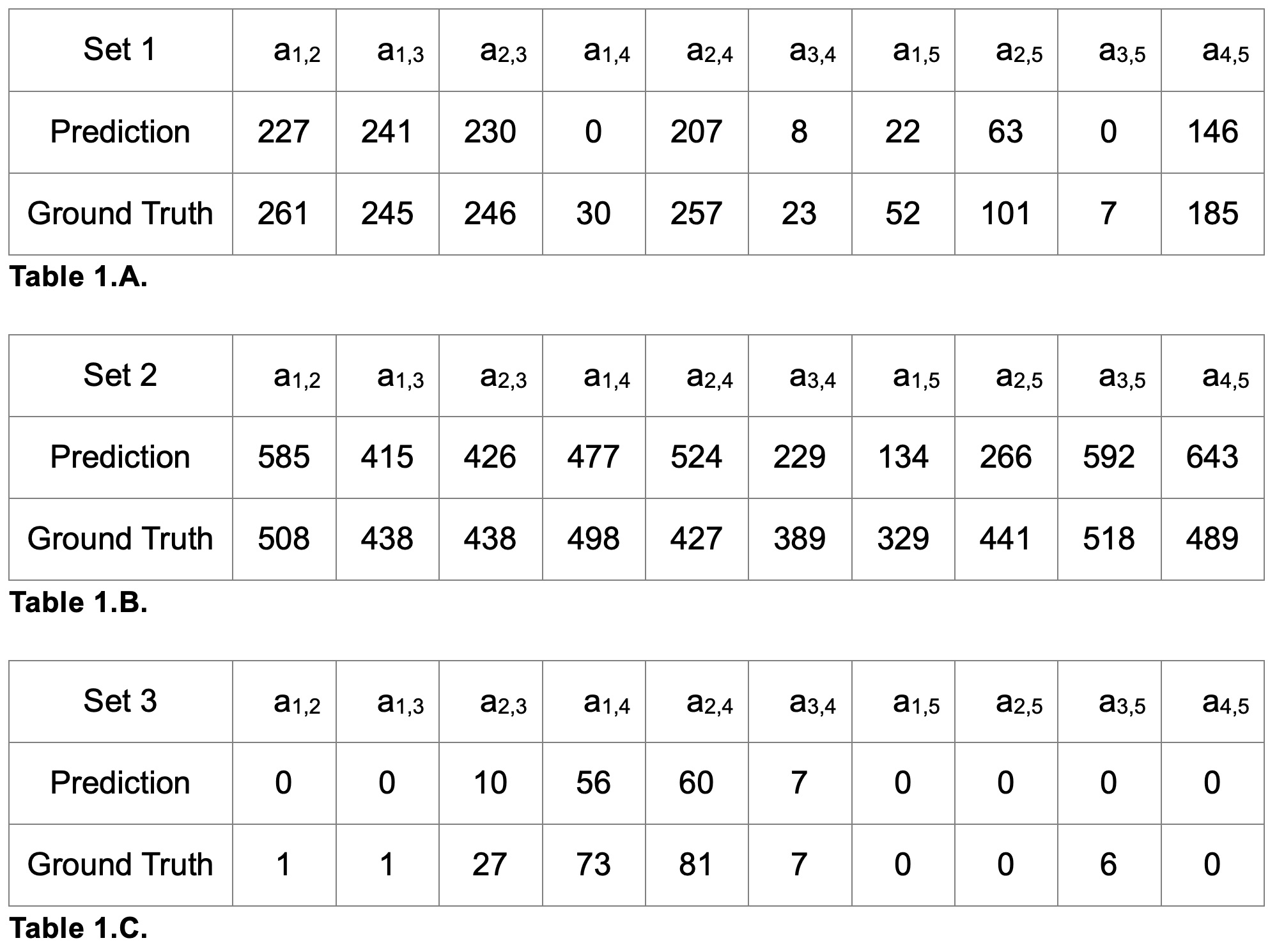


Table 1. The summary of three sets of testing on ToMNet 2.0.

Next, in order to evaluate the model’s performance, we graphed the learning curve of each set of training (Fig. 6). The learning curve is plot by training error and validation error over training steps. In our model we apply the top 1 error rule to calculate training and validation error, which is the percentage of time that the classifier did not give the correct class the highest score. In Fig. 6A & C the model is regarded as a good fit training with some noise, however in Fig. 6B the error rate remains steadily at a high level, the possible reason here may relate to how the model evaluates the error. The accuracy is calculated by the prediction of the final states of the grid world, this indicates that the correctness is only counted while ToMNet 2.0 predicts the exact correct outcome of each grid world simulation. However, there are limitations in this current model to pick up such dynamic but precise information during current simulations, meanwhile, similar performance might occur in real life social interaction, since in an attractive interacting scene, human may not have sufficient variation to distinguish other’s social preferences.

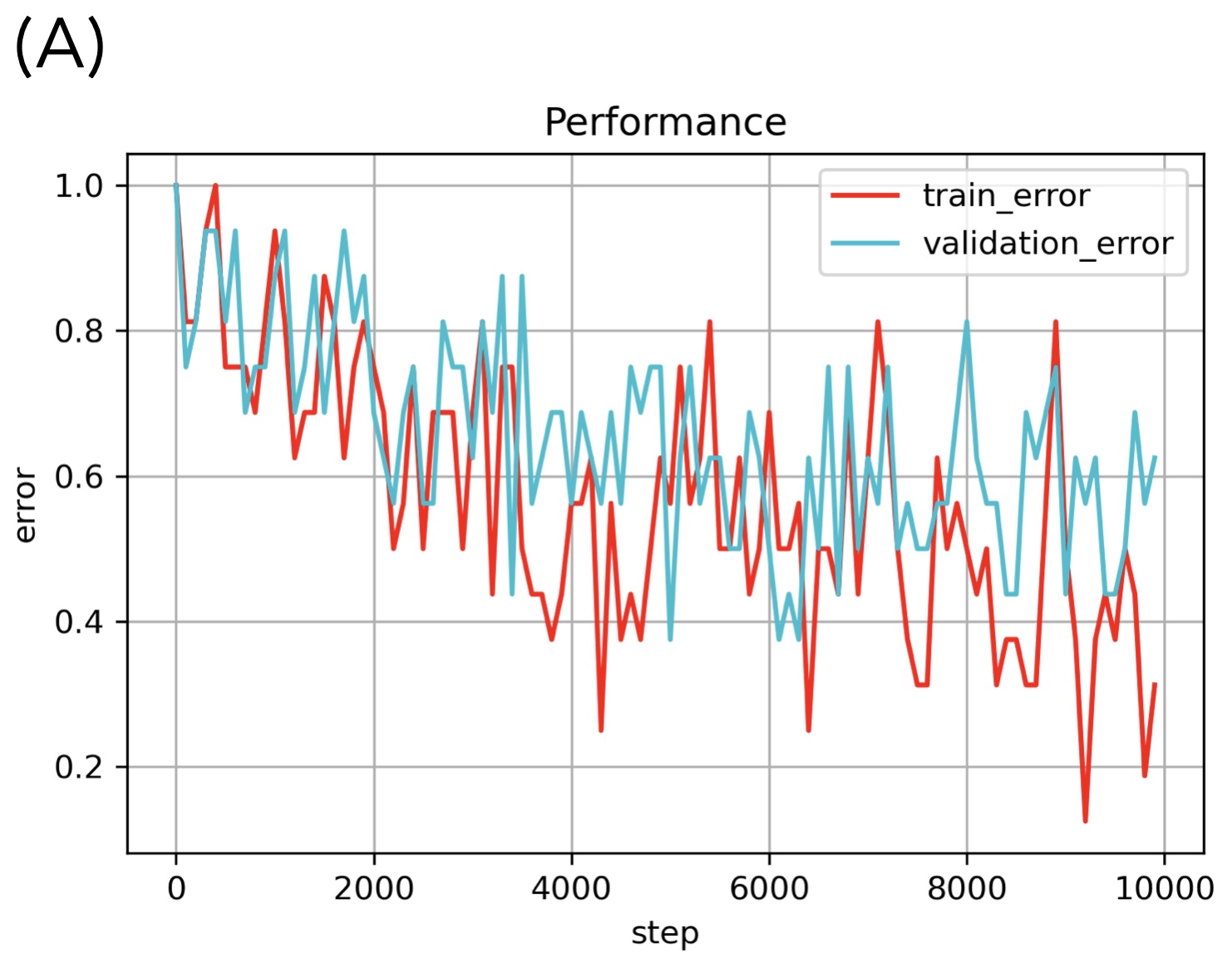
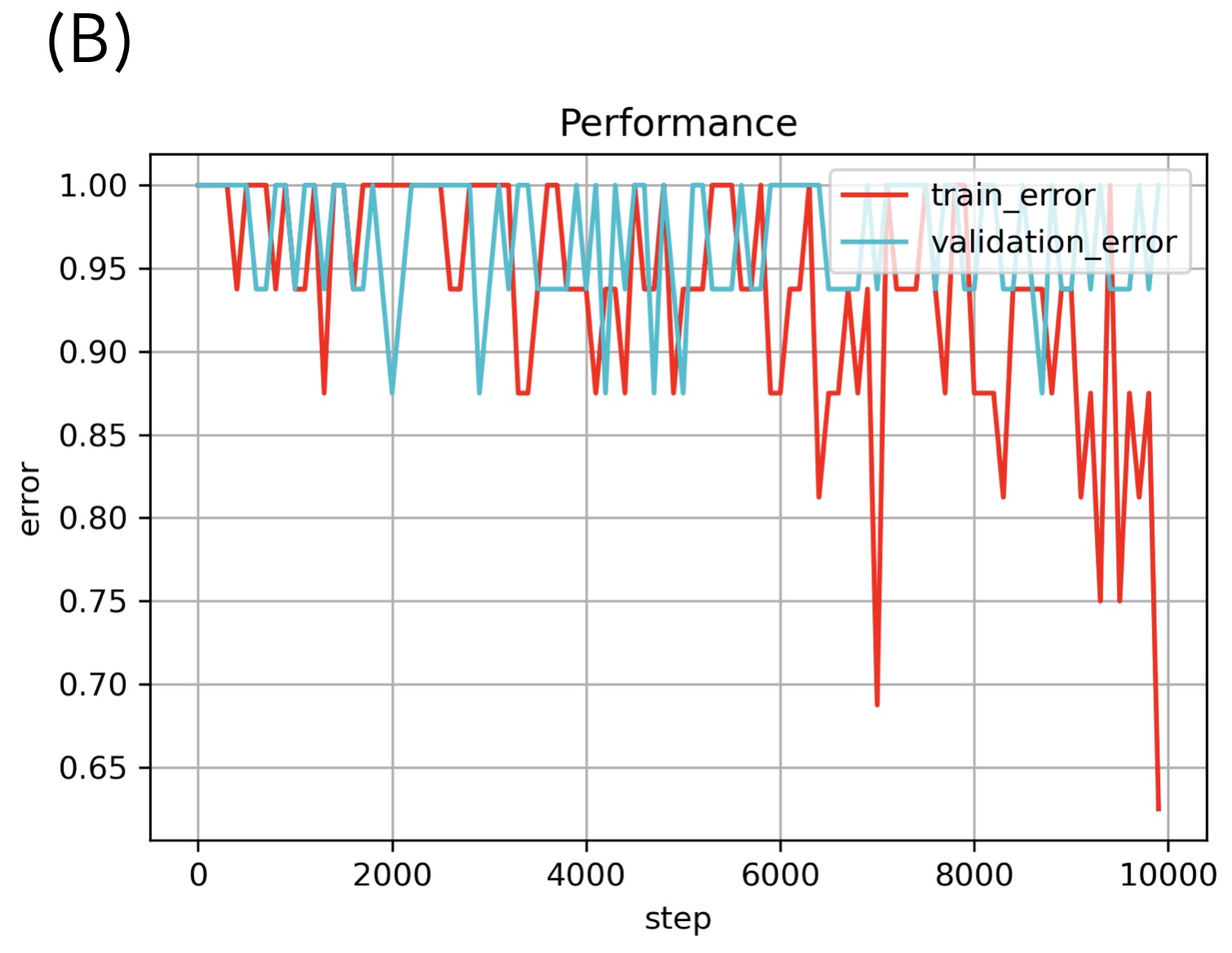
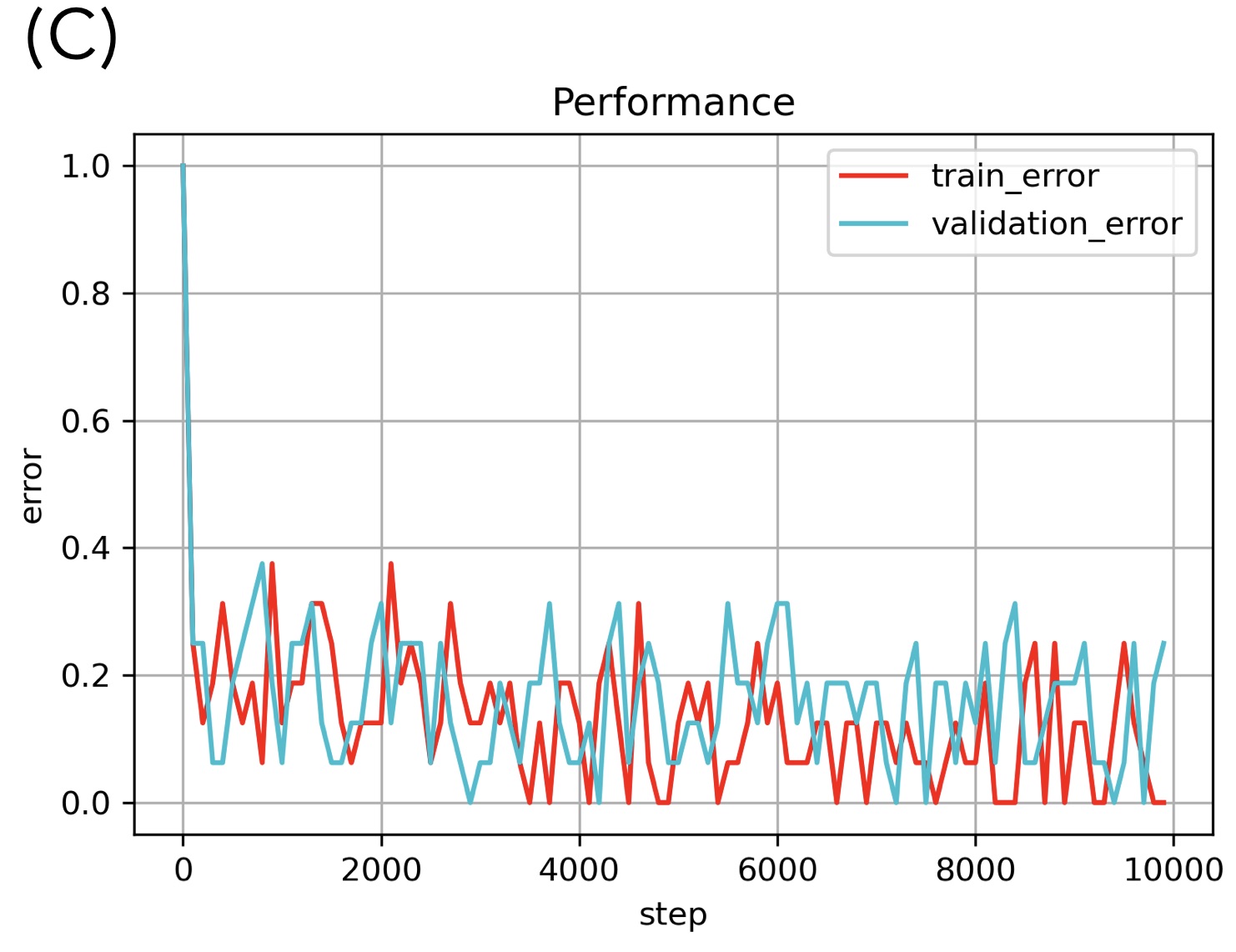


Figure 6. The training performance evaluated by learning curve.

**Fig. 6 (A) presents a properly trained model with an acceptable error and without over/under training. Fig. 6 (B) states such social structures may need an increment of model complexity. Fig. 6 (C) reveals the information embedded may be too simple this leads to early settlement but without overfitting.**

1. Random rate comparison

To further evaluate the model performance, we tested the three sets of trained model with other sets of grids, in which social scores were permuted (i.e., the social preferences had similar ranges but the specific associations were not the same as the one which the model was trained for). We performed the direct 10 integer vector counting, the RMSE against ground truth, and the accuracy as our measurement, details are stated in section A.

In Fig. 7A & C the performances were well distinguished for the trained social structures and indicated ToMNet 2.0 learned the information embedded in the training sets. In Fig. 7B the analysis states that ToMNet 2.0 still gain partial information and forms a similar but imperfect embedded preference through training.

The analysis of RMSE is aimed to measure the difference between ground truth, prediction, and permuted sets. The pattern plotted in Fig. 7 indicates that RMSE may not be a very precise evaluation whether ToMNet 2.0 is well trained, since all three sets seems to hold similar patterns at Fig. 7. However note the value of RMSE at set 2 is significantly grater than the other two sets, which is a sign of the imperfect uptake of preference information.

The analysis of accuracy is a relatively rigorous evaluation of ToMNet 2.0’s performance, since accuracy is calculated by evaluating the testing of each individual grid world movement trial. Thus, we should be aware of this demand in evaluating the accuracy of this model when decoding Fig. 7. Critically, note that ToMNet 2.0 is expected to learn the social network structure rather than the precise ground truth of each grid. In conclusion, only focusing on the poor accuracy in Fig. 7 may lead to disregarding the partial information ToMNet 2.0 gained throughout the training process. Rather, we suggest that when evaluating ToMNet 2.0, the summed 10 integer counting vector should also be referred to instead.

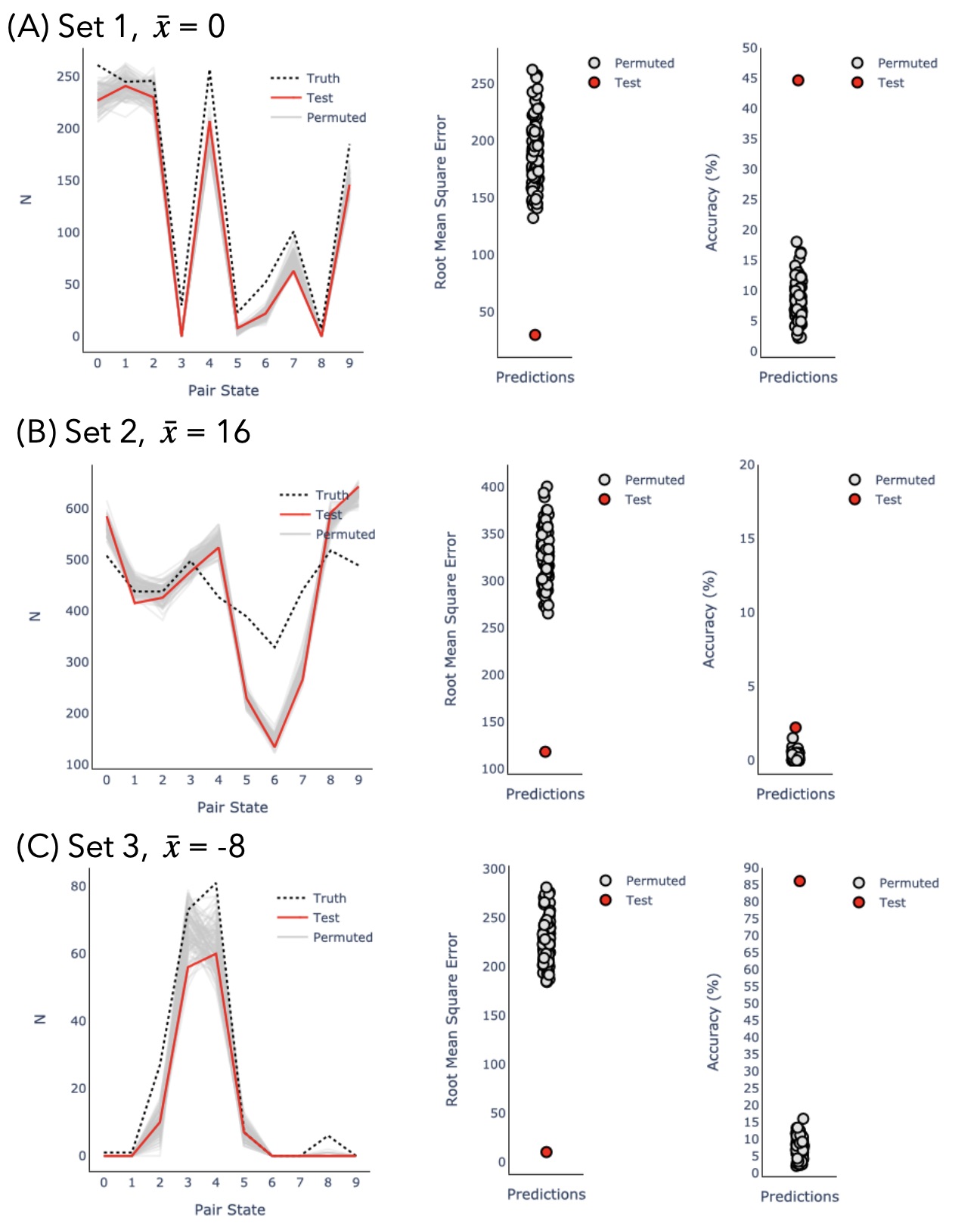


Figure 7. Trained model testing with random social setting grids.

In Fig. (7), the red dot represents a trained ToMNet 2.0’s prediction of the testing data with the same social setting. Other grey dots are the trained ToMNet 2.0’s prediction of the randomly generated social settings. RMSE is calculated by the sum of all 1000 grids prediction results, while accuracy is calculated by judging the correctness of all 1000 grids.

Discussion

1. Implications

According to the results, these findings suggest that the designed structure of ToMNet 2.0 learned from the social preference based training set, and has the ability to decode implicit social preferences. This achievement remains absent in the previous mentioned models. While this result indicates that ToMNet 2.0 have the ability to learn information embedded in a dynamic social interaction, to become an indispensable implementation in socially assistive machines simulation in a more realistic format should further be tested. Since ToMNet 2.0 is able to tell the embedded five-bodied social preference, it becomes the first step to optimize social robots applied in multi-users scenario. With some adjustments on the model design to reduce the required amount of training materials or broaden the acceptance of input formats may benefit the field of socially assistive robots more widely.

In the results where agents are attracted to each other (set 2), seemingly the model was underfitted, even though the summed up of prediction results correlates with the ground truth. Thus, ToMNet 2.0 still gained some information about agent’s social structure despite limited variance in agent outcomes. In human social interaction, people may not be as precise as required in our modeling. In other words, if the aim is to develop models like ToMNet 2.0 that operate without numerous training data or even with a more abstract model setting, the application of this approach should be quite broad.

The ToMNet model was proven to pass Sally Anne test in previous studies, further more in this research, ToMNet 2.0 has the ability to form perceptions of abstract concepts such as multi agent social network. By contrast, ChatGPT as a Large Language Model (LLM) which was also reported to pass some levels of theory of mind (ToM) tests still had limitations in achieving theory of mind. Recent research successfully detect ToM on GPT-4 reports the accuracy raises due to producing longer-form answers. However, the author mentioned further testing on distinguishing inferential reasoning and GPT-4’s ToM should be performed [21]. According to the ToM ability tests performed on GPT-4, LLM’s requires script inputs, while social interaction contains not only verbal input but also lots of non-verbal stimuli. When implementing such stimulus, ToMNet 2.0 simply requires stimuli transformation into social scores, while LLM’s would request for a script describing the non-verbal stimuli. Indeed, there were some similar achievements for both ToMNet 2.0 and GPT-4 when testing machine’s theory of mind, nonetheless to approach the abstract conceptualization of human social network, ToMNet 2.0 seems a slightly better option than GPT-4.

1. Application on human social interaction data

Our intention in the near future is to apply ToMNet 2.0 on human social interaction data. In the social interaction game five participants are recruited, they are required to play a board game named Avalon where each participant role plays a specific character. The character participants role playing forms loyalties and adversaries between each other whilst the game quest flows. Meanwhile we record participants head orientation, face emotion expressions, eye gaze and speech to reflect the dynamic social preference during the game. The collected data is preprocessed and prepared for training and testing on ToMNet 2.0, such dynamic human social interaction data application stands for a precious attempt on evaluating model’s ability to process human social network.

1. Limitations

ToMNet 2.0 has several limitations. First, the input format is limited to grid format, which restricts the application on a more realistic social stimulus. Second, the social structure is fixed as constants, however in realistic human social preference changes by time. Third, the output matrix limits the information of social networks. Social interactions are bidirectional in the initial social network settings, however the output matrix can only provide a sum up result of the social interaction. Last, in the case when social score > 0, the performance seems not as accurate as other cases, which may be reasonable both in stimulation and realistic. When discuss about simulation in the case of > 0, the agents converges due to the Fsocial setting. While in realistic social scenario, a group of friends interacting may be the example of > 0, an observer may not be able to tell the individuals degree of favor between the group.



1. Conclusions

ToMNet 2.0 is proven to comprehend a more complex social network, however this is restricted in the present settings of this model, there are lots of further attempts worth to apply on such models to aid machines understand social scenario more efficiently.

References

1. Feil-Seifer, D. and M.J. Mataric, *Human RobotInteraction.*

2. Duffy, B.R., *Anthropomorphism and the social robot.* Robotics and Autonomous Systems, 2003. 42(3-4): p. 177-190.

3. Henschel, A., G. Laban, and E.S. Cross, *What Makes a Robot Social? A Review of Social Robots from Science Fiction to a Home or Hospital Near You.* Curr Robot Rep, 2021. 2(1): p. 9-19.

4. Whelan, S., et al., *Factors Affecting the Acceptability of Social Robots by Older Adults Including People with Dementia or Cognitive Impairment: A Literature Review.* International Journal of Social Robotics, 2018. 10(5): p. 643-668.

5. Matarić, M.J., Scassellati, B., *Socially Assistive Robotics. In: Siciliano, B., Khatib, O.* Springer Handbook of Robotics., 2016.

6. Stock-Homburg, R., *Survey of Emotions in Human–Robot Interactions: Perspectives from Robotic Psychology on 20 Years of Research.* International Journal of Social Robotics, 2021. 14(2): p. 389-411.

7. Johanson, D.L., H.S. Ahn, and E. Broadbent, *Improving Interactions with Healthcare Robots: A Review of Communication Behaviours in Social and Healthcare Contexts.* International Journal of Social Robotics, 2020. 13(8): p. 1835-1850.

8. Papadopoulos, C., et al., *The CARESSES Randomised Controlled Trial: Exploring the Health-Related Impact of Culturally Competent Artificial Intelligence Embedded Into Socially Assistive Robots and Tested in Older Adult Care Homes.* Int J Soc Robot, 2022. 14(1): p. 245-256.

9. Pu, L., et al., *The Effectiveness of Social Robots for Older Adults: A Systematic Review and Meta-Analysis of Randomized Controlled Studies.* Gerontologist, 2019. 59(1): p. e37-e51.

10. Koh, W.Q., et al., *Barriers and facilitators to the implementation of social robots for older adults and people with dementia: a scoping review.* BMC Geriatr, 2021. 21(1): p. 351.

11. Langley, C., et al., *Theory of Mind and Preference Learning at the Interface of Cognitive Science, Neuroscience, and AI: A Review.* Front Artif Intell, 2022. 5: p. 778852.

12. Astington, J.W. and J.M. Jenkins, *Theory of mind development and social understanding.* Cognition & Emotion, 1995. 9(2-3): p. 151-165.

13. Premack, D. and G. Woodruff, *Does the chimpanzee have a theory of mind?* Behavioral and Brain Sciences, 2010. 1(4): p. 515-526.

14. Williams, J., S.M. Fiore, and F. Jentsch, *Supporting Artificial Social Intelligence With Theory of Mind.* Front Artif Intell, 2022. 5: p. 750763.

15. Hernandez-Leal, P., B. Kartal, and M.E. Taylor, *A survey and critique of multiagent deep reinforcement learning.* Autonomous Agents and Multi-Agent Systems, 2019. 33(6): p. 750-797.

16. Neil C. Rabinowitz, et al., *Machine Theory of Mind.* ArXiv Prepr., 2018. arXiv:1802.07740.

17. Hao Zhu, G.N., Yonatan Bisk, *Few-shot Language Coordination by Modeling Theory of Mind.* 2021. arXiv:2107.05697.

18. Stojnic, G., et al., *Commonsense psychology in human infants and machines.* Cognition, 2023. 235: p. 105406.

19. Yun-Shiuan Chuang, et al., *Using Machine Theory of Mind to Learn Agent Social Network Structures from Observed Interactive Behaviors with Targets.* 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2020.

20. Goh, J.O.S., 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*, K.D. Federmeier, Editor. 2018, Academic Press. p. 205-240.

21. Shima Rahimi Moghaddam, C.J.H., Boosting Theory-of-Mind Performance in Large Language Models via Prompting. 2023. arXiv:2304.11490