

# A Study on Like-Attracts-Like versus Elitist Selection Criterion for Human-like Social behavior of Memetic Multitagent Systems

Xuefeng Chen\*, Yifeng Zeng<sup>†</sup>, Yew-Soon Ong<sup>‡</sup>, Choon Sing Ho<sup>‡</sup>, Yanping Xiang\*,

\*School of Computer Science and Engineering

University of Electronic Science and Technology of China, P. R. China

Email: {cxflovehchina, xiangyanping}@gmail.com

<sup>†</sup>School of Computing

Teeside University, United Kingdom

Email: Y.Zeng@tees.ac.uk

<sup>‡</sup>School of Computer Engineering

Nanyang Technological University, Singapore

Email: {asysong, csho}@ntu.edu.sg

**Abstract**—Memetic multiagent system emerges as an enhanced version of multiagent systems with the implementation of meme-inspired computational agents. It aims to evolve human-like behavior of multiple agents by exploiting the Dawkins' notion of a meme and Universal Darwinism. Previous research has developed a computational framework in which a series of memetic operations have been designed for implementing human-like agents. This paper will focus on improving the human-like behavior of multiple agents when they are engaged in social interactions. The improvement is mainly on how an agent shall learn from others and adapt its behavior in a complex dynamic environment. In particular, we design a new mechanism that supervises how the agent shall select one of the other agents for the learning purpose. The selection is a trade-off between the elitist and like-attracts-like principles. We demonstrate the desirable interactions of multiple agents in two problem domains.

**Keywords**—multiagent systems, memetic automaton, towards human-like behavior

## I. INTRODUCTION

Multiagent system (MAS) has become an important tool to solve complex problems in various domains like robotics [1], computer games [2], e-learning/tutoring [3] and so on. The MAS technologies have successfully overcome the computational intractability of relevant problems. The success lies in the merit that multiple agents are able to collaborate with and learn from each other through effective communication channels. As intelligent agents become more and more involved in our daily life, the agents-based solutions are expected to exhibit human-like characteristics that facilitate the interactions between agents and humans [4]. For example, it is good to have home assistant robots that can take care of housework without messing up our personal life. While playing digital computer games, human players would like to challenge computer players (called non-player characters or game agents) whose behavior follows a human line in game-play. This drives the agent-related research towards the study on human-like behavior of intelligent agents [5]. Memetic MAS (MeMAS) [6],

[7] is one of such frameworks that permits a systematic study on generating human-like behaviors, particularly on agents' social interactions.

MeMAS is a marriage between agent technologies and memetic automaton [8], [9] where the individual and interaction models of a population of agents are meme-inspired design. As with genes in genetics, a meme is synonymous to memetics as being building block of cultural know-how that is transmissible and replicable. A meme is represented internally as the learned knowledge, and externally as the manifested behaviors of agents. Accordingly a meme automaton is defined as an adaptive agent that autonomously acquires increasing level of capability and intelligence through embedded memes evolving independently or via social interactions. In particular, the stochastic evolution of individual memes resembles the varying growth of human in nature, while the transmission during memes' interactions is akin to the human imitation in the real life. Naturally the meme-inspired computation serves as a plausible platform for the design of human-like MAS.

Recently Feng *et al* [7] materialize the MeMAS framework in the form of neural network architecture particularly via the Temporal Difference - Fusion Architecture for Learning and Cognition (TD-FALCON) [10]. The contextualized MeMAS shows fast online learning abilities and considered natural human-like behavior such as imitation for facilitating the problem solving. Adopting a imitate-from-elitist principle, the MeMAS drives agents' learning through the rewarding mechanism - agents always learn from a successful teacher agent in the solution history. However, we notice that in practice, rewarding should not be the sole element to drive the learning of memetic agents, particularly on whom shall be selected for the learning purpose. As one part of social interactions, the notion of *like-attracts-like* has a significant influence on interpersonal attraction which is one of the core principles behind the human learning [11]. Taking this clue, in this paper our interest is to study the learning behavior of memetic agents under a like-attracts-like selection versus the commonly used elitism selection mechanism in the MeMAS

for problem solving.

The principle of like-attracts-like drives agents to learn from the one that has a similar memotype. A memotype encapsulates both an agent's personality and its encounter. In this paper, we investigate a probabilistic representation of memotype that encodes the stochastic actions of memetic agents given world states. We then adapt distance measurements to gauge the similarity between a pair of agent memotypes. The similarity, together with the agent fitness, provide a two-dimensioned criterion to the learning agent for selecting a teacher agent. The criterion trades off the elitist principle with the like-attracts-like principle. By doing this, the memetic agents may acquire more social behavior in their interaction. We demonstrate the performance of the proposed selection mechanism in two domains. One is a task-based problem setting where the MeMAS shows emergent behavior while achieving the goal. The other is a video game where non-player characters trained by the MeMAS interact naturally with human-players online.

The remainder of this paper is structured as follows. Section II provides an overview of the MeMAS based on the TD-FALCON artificial neural network models. Section III details the proposed like-attracts-like selection criterion and integrates it into the agent learning process. Section IV presents the empirical results on the agents behavior in MeMAS based on the like-attracts-like criterion as compared to the commonly used elitism principle, on two game domains. Finally, Section VI concludes the paper with some brief remarks.

## II. BACKGROUND: MEMAS AND TD-FALCON MODEL

In this section, we briefly describe Memetic Multiagent System (MeMAS) followed by its implementation using TD-FALCON models, and refer the reader to [7] for more details.

### A. Memetic Multiagent Systems

Inspired by memetic automaton, MeMAS is designed towards human-like learning and interaction. We depict the architecture of MeMAS in Fig. 1, where a population of memetic agents learn and evolve in a dynamic environment.

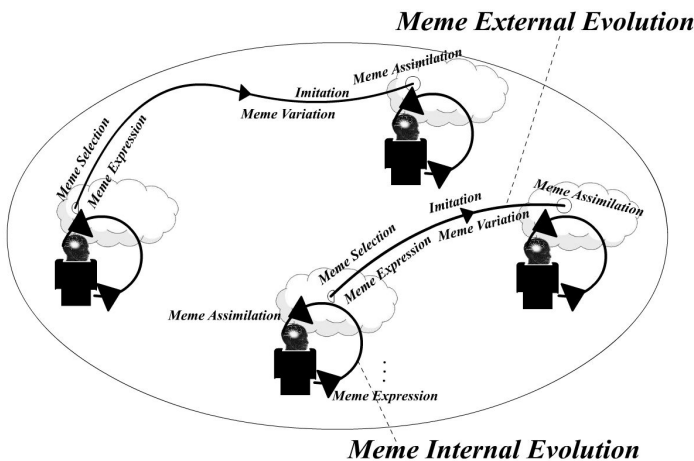


Fig. 1. Illustration of Memetic Multiagent System

The basic data structure of the MeMAS contains *meme representation* that is manipulated by four control functions called *meme expression*, *meme assimilation*, *meme internal evolution*, and *meme external evolution*. The MeMAS uses meme to represent internal knowledge of individual agents that are generalized instructions in agents' mind. *Meme expression* translates the knowledge into a set of behaviors that can be observable or partially observable in the environment. *Meme assimilation* then translates the observed or partially observed behaviors into knowledge that blends into the individual mind-universe. In other words, *meme representation* decides on the type of meme, called *memotype*, while *Meme expression/assimilation* activates/updates the *memotype* in the execution.

Meme evolution, including *meme internal evolution* and *meme external evolution*, is central to the behavioral aspects of memetic automaton. *Meme internal evolution* governs the growing of individual agents mainly through self-learning. *Meme external evolution* models the interactions among agents which is primarily driven by imitation. With regard to imitation, *memetic selection* concerns with whom one imitates, while *meme transmission and variation* relates to how one imitates and what is imitated. During the interactions with other agents, memetic agents pick up a proper teacher agent and then imitate behaviors of others to enhance their skills in an incremental way. We outline the implementation framework of MeMAS in Fig. 2.

### PSEUDO CODE OF MEMETIC MULTIAGENT SYSTEM

1. Initialize  $N$  memetic agents
2. **While** Stopping conditions are not satisfied
3.     **For**  $i=1$  to  $N$  **do**
4.         Compute the probability  $Pr_i$
5.         **If**  $Rand() < Pr_i$
6.             /\* Perform *meme external evolution* \*/
7.             Perform *meme selection* to pick out the teacher agent  $Agt_j$
8.             Perform *meme variation* with the probability  $\tau$  between  $Agt_i$  and  $Agt_j$
9.             Perform *meme transmission* from  $Agt_j$  to  $Agt_i$
10.         **Else**
11.             Perform *meme internal evolution*

Fig. 2. A general framework of MeMAS includes *meme internal evolution* and *meme external evolution*.

In Fig. 2, we generate  $N$  memetic agents that compose the initial MeMAS. The MeMAS evolves gradually until solutions to a task are found (line 2). During an evolution procedure, every memetic agent will undergo either *meme external evolution* or *meme internal evolution*, which relies on the probability  $Pr_i$  (lines 4-5). Normally,  $Pr_i$  is computed according to the agent performance in the history. For example, it could be the success rate defined as the number of times that the agent completes the task successfully. If the agent executes the process of *meme external evolution*, it firstly selects one teacher agent (line 6) and then performs *meme variation* and *meme transmission* (lines 7-8). This decides how the agent interacts with the teacher agent. The general MeMAS framework can be implemented using classical techniques including

genetic programming, fuzzy rules, probabilistic models, and so on.

### B. TD-FALCON

The TD-FALCON (standing for *Temporal Difference - Fusion Architecture for Learning and Cognition*) models a connectionist agent in the form of three-channel neural network architecture. Following the spirit of reinforcement learning paradigm, it associates a world state with agents' actions that lead to a desirable outcome. We show the implementation of memetic agents using the TD-FALCON in Fig. 3.

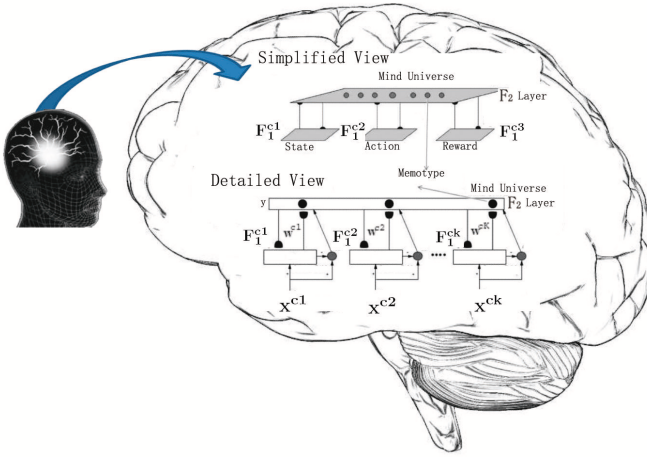


Fig. 3. The TD-FALCON model implements a memetic agent in MeMAS.

A TD-FALCON model contains three components: a sensory field  $F_1^{c1}$ , a motor field  $F_1^{c2}$ , and a feedback field  $F_1^{c3}$ .  $F_1^{c1}$  represents the input of environmental states,  $S = \langle s_1, s_2, \dots, s_n \rangle$ , while  $F_1^{c2}$  models a set of available actions,  $A = \langle a_1, a_2, \dots, a_l \rangle$ , that agents can select and execute given the sensory input.  $F_1^{c3}$  is a reward vector,  $R = \langle r, \bar{r} \rangle$  where  $r$  is the reward signal value and  $\bar{r} = 1 - r$ . In addition, the model has a cognitive layer  $F_2$  for the acquisition and storage of agents' memes. More concretely, the meme encodes a relation among the patterns of  $F_1^{c1}$ ,  $F_1^{c2}$  and  $F_1^{c3}$ .

Let  $x^{ck}$  denote the  $F_1^{ck}$  activity vector for  $k=1,2,3$ , which is weighted by  $w_m^{ck}$  in the  $m$ th neuron in layer  $F_2$ . The output neuron, denoted by  $y$  in  $F_2$ , is the learned patterns (memes) from  $F_1^{c1}$  once the weight training of activity vectors is completed. We compute the activation value  $T_m$  for meme  $m$  in Eq. 1.

$$T_m = \sum_{k=1}^3 \gamma^{ck} \frac{|x^{ck} \wedge w_m^{ck}|}{\alpha^{ck} + |w_m^{ck}|} \quad (1)$$

where the fuzzy AND operation  $\wedge$  is defined by  $p \wedge q = \min(p_i, q_i)$ , and the norm  $|\cdot|$  is defined by  $|p| = \sum_i p_i$  for vector  $p$  and  $q$ . Parameters  $\gamma^{ck}$  and  $\alpha^{ck}$  are specified by users.

The *meme internal evolution* adopts a temporal difference (TD) formulation to estimate reward values for a state-action combination in the repetitive trials. An  $\epsilon$ -greedy action selection scheme is used in MeMAS to balance exploration and exploitation in the agent's self-learning. In the *meme external*

*evolution*, a memetic agent selects the teacher agent,  $Agt_j$  that has similar experience from a pool of elite agents. The selection criterion follows Eq. 2.

$$Agt_j = \operatorname{argmax}_j T_{Best} \times \frac{F(Agt_j)}{F_{Best}} \quad (2)$$

where  $T_{Best} = \max\{T_m | \text{all memes } m \text{ in } F_2\}$ ,  $F(Agt_j)$  the fitness of  $Agt_j$  larger than  $F(Agt_i)$ , and  $F_{Best} = \max\{F(Agt_j) | j \in N\}$ .

Once the teacher agent is decided, the memetic agent enters the *meme variation* and *meme transmission* stages. In MeMAS, meme variations is realized by adding a perturbation component to the statistical merit of each TD-FALCON based teacher agent. This induce a diversity of teacher actions for transmission in the evolution.

### III. PROPOSED LIKE-ATTRACTS-LIKE VERSUS ELITISM PRINCIPLE AS SELECTION CRITERION

One central issue in the *meme external evolution* is on whom shall be selected for the imitation purpose. Imitate-the-elite is one of the most popular strategies for selecting a teacher agent. With such a scheme, the coupled formulation (Eq. 2) thus focuses on the elite pool and measures the experience in terms of fitness values, which has the tendency of a biased selection towards only the elite agents.

As humans are prone to imitate others of a similar type, a memetic agent shall also manifest such human-like social behavior when it is selecting a teacher agent. The selection shall consider not only solution performance in terms of agents' fitness, but also the evolution origin in terms of agents' personal attributes. Thus, in contrast to using an imitate-the-elite scheme, here we consider TD-FALCON agents that adopts the like-attracts-like principle of agents' experiences in MeMAS. To begin, we first define the similarity measurement of agents' memotype, and then specify the parameters involved and their settings.

#### A. Selection Criterion

As defined in the TD-FALCON model, a memotype, denoted by  $Q = (S, A, R)$ , essentially encodes a mapping between input states,  $S$ , and actions,  $A$ , through the reward measurement,  $R$ . It models how a memetic agent responds to a sensory input. This basic definition loses an important connection between a meme and its genetic origin in the concept of meme automaton. Analogous to that a meme is in part regulated by its gene, behavior of memetic agents may be often determined by their original properties. Accordingly, we expand a memotype with agents' attributes denoted by  $\Theta$ . Formally, the augmented memotype is defined as:  $MT = \langle \Theta, Q \rangle$ . The next issue is on how to measure the similarity among agents' memotypes.

As agents' attributes,  $\Theta$ , normally have a numerical scale, we use normalized Euclidean distance to measure the similarity between two attribute sets  $\Theta$  and  $\Theta'$ .

$$S_{ED}[\Theta, \Theta'] = 1 - \frac{\sum_{\theta \in \Theta, \theta' \in \Theta'} \sqrt{\sum_{\theta_{dim} \in \theta, \theta'_{dim} \in \theta'} (\theta_{dim} - \theta'_{dim})^2}}{|\Theta|} \quad (3)$$



where  $\theta_{dim}$  (or  $\theta'_{dim}$ ) is standardized value whose range is  $[0,1]$ , for a single attribute,  $\theta$  (or  $\theta'$ ), in the sets, and  $|\Theta|$  the cardinality of  $\Theta$ . Note that Eq. 3 computes the common set of  $\Theta$  and  $\Theta'$ .

Considering the bounded rationality of an agent, we use probabilistic models to define its behavior. Formally, let  $Pr(A|S)$  be a set of probability distributions over actions given the input of world states. To measure the distance between  $Q$  and  $Q'$ , we resort to Kullback-Leibler (KL) divergence [12] in Eq. 4.

$$D_{KL}[Q||Q'] = Pr_Q(A|S) \ln \frac{Pr_Q(A|S)}{Pr_{Q'}(A|S)} \quad (4)$$

We further adapt Eq. 4 and define a symmetric measurement of similarity between  $Q$  and  $Q'$  in Eq. 5. Note that the similarity value is scaled within  $[0,1]$ .

$$SD_{KL}[Q, Q'] = e^{-\frac{1}{2}\{D_{KL}[Q||Q'] + D_{KL}[Q'||Q]\}} \quad (5)$$

As a memetic agent may be featured by both attributes and behavior, the similarity between agents,  $Agt_i$  and  $Agt_j$ , is subsequently computed in Eq. 6.

$$SIM(Agt_i, Agt_j) = \frac{S_{ED}[\Theta, \Theta'] + SD_{KL}[Q, Q']}{2} \quad (6)$$

Driven by the like-attracts-like principle,  $Agt_i$  may select  $Agt_j$  that has the largest  $SIM(Agt_i, Agt_j)$  value. On the other hand,  $Agt_i$  may also expect to learn from an elite agent that has more sophisticated skills. We use the fitness ratio,  $\frac{F(Agt_j)}{F_{Best}}$ , to measure how well  $Agt_j$  approaches the best one. Consequently,  $Agt_i$  will select an agent as the teacher that has the largest value for the combined measurements of  $SIM(Agt_i, Agt_j)$  and  $\frac{F(Agt_j)}{F_{Best}}$ . Formally, the selection criterion is defined in Eq. 7.

$$Agt_j = \operatorname{argmax}_j K_1 \times SIM(Agt_i, Agt_j) + K_2 \times \frac{F(Agt_j)}{F_{Best}} \quad (7)$$

where  $K_1$  and  $K_2$  are parameters balancing the similarity and elitist factors.

Note that if the selection is solely driven by the like-attracts-like principle,  $Agt_i$  loses chance to explore the entire solution space. More importantly, by learning from a distinct type of agents,  $Agt_i$  may update the memotype particularly on speeding up the co-evolution by recognizing its genetic origin and relating it to other types.

#### B. Parameter Settings

As an agent does not act individually in the environment, its selection on a teacher agent may be influenced by evolutions of other agents in MeMAS. Particularly, the dynamic properties of MeMAS may impact the trade-off between the aforementioned two principles: like-attracts-like and elitist. We make a further step to illustrate the settings of  $K_1$  and  $K_2$ .

Intuitively, the similarity factor may place an important role in the agent's selection if there are dominating groups of similar agents in the MeMAS. Otherwise, the factor may become weak if all agents are equally similar. Under the thoughts of this vein, we may specify  $K_1$  as the diversity value of the MeMAS that measures the uncertainty of different groups of agents for a population of agents.

Resorting to regular clustering techniques like  $k$ -means [13], we group  $N$  agents into  $l$  clusters ( $< C_1, \dots, C_l >$ ) in terms of similarity measurements. Each group contains a number of agents that have similar memotypes. To compute the diversity of agent groups in the MeMAS, we use the information entropy [12] as defined below.  $K_1$  is proportional to the entropy value in Eq. 8.

$$K_1 \simeq - \sum_l \frac{|C_l|}{N} \ln \left( \frac{|C_l|}{N} \right) \quad (8)$$

where  $\frac{|C_l|}{N}$  is the ratio of the size of cluster  $C_l$  to the MeMAS space.

We perceive that the setting of  $K_2$  depends on distributions of agents' skills in MeMAS. Naturally an agent may pick out a teacher agent depending on the similarity of candidate agents if all of them are elite. In other words, the elitist principle may have a small impact on the selection when there is little divergence of the skill levels for all agents. We compute the variance of all agents' fitness values, and let  $K_2$  be the proportion of the variance in Eq. 9.

$$K_2 \simeq \operatorname{Var}[F(Agt_j | j = 1, \dots, N)] \quad (9)$$

*Example 1 (Parameter Setting):* In Fig. 4, the MeMAS contains a set of agents that have different distributions of fitness values and types in terms of similarity. Fig. 4(b) shows that the agents are grouped into three clusters in terms of the similarity. In addition, the fitness values are distributed over an entire scale. Hence, both the similarity and elitist factors have a competitive impact on the selection. As there is only one group of agents in Fig. 4(a), the diversity of MeMAS approaches zero. Consequently, the single factor of the elite is counted in the selection. Similarly, in Fig. 4(c), most agents have similar fitness values, selecting a teacher agent mainly depends on the similarity factor.

Given no prior knowledge on the MeMAS, we normally let  $K_1$  and  $K_2$  be equal in the initial phase. After each evolution, we compute  $K_1$  and  $K_2$  online, and the new values manifest the updated MeMAS state. When  $K_1$  approaches 0, the new selection criterion completely follows the imitate-the-elite strategy as the memotype similarity does not play any role in the selection.

## IV. EXPERIMENTAL RESULTS

We implemented the new selection mechanism (based on Eq. 7) together with the TD-FALCON model in the MeMAS framework. The new mechanism is embedded in Line 6 of Fig. 2, and provides a criterion to the learning agent for selecting a suitable teacher agent. To demonstrate the social behavior of memetic agents, we investigated on two domains: one is the adapted version of minefield navigation

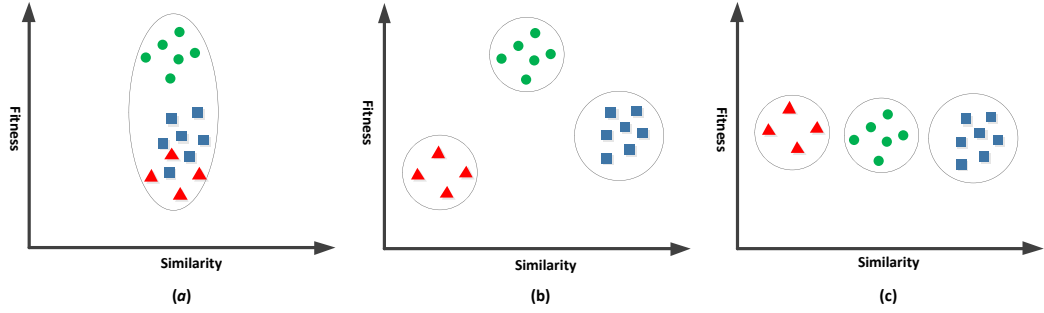


Fig. 4. Distributions of fitness values and types of agents impact the settings of  $K_1$  and  $K_2$ : (a)  $K_1 \approx 0$ ,  $K_2 > 0$ ; (b)  $K_1 > 0$ ,  $K_2 > 0$ ; (c)  $K_1 > 0$ ,  $K_2 \approx 0$ .

problem [10] and the other a 3D interactive game on the homeland defense <sup>1</sup>. We study the like-attracts-like selection criterion (denoted here as MeMAS-E) versus the conventional elitism selection criterion (denoted as MeMAS-C) [7] in the MeMAS framework. It is observed that the memetic agents in MeMAS-E exhibit better human-like social behavior while maintaining the performance on achieving the mission goals. In addition, we also conducted further subject study on both MeMAS frameworks. We invited human players to evaluate several human-like properties of the memetic agents and to play with the agents in the games. The evaluation asserts the social behavior of memetic agents as expected by the human players.

#### A. Minefield Navigation Domain

The first domain is the minefield navigation problem that has been used to assess social behavior of the MeMAS-C [7] previously. In the minefield navigation task, a team of tank agents move across a minefield in order to capture a red flag. A set of flags (with the positions) are randomly generated one by one in the map. The field is filled with land mines that are unknown to the tanks. In the adapted minefield navigation problem that we consider here, we include two types of tanks and also two forms of mines in the field. All tanks have the same navigation actions (turning left/right, moving forward, proceeding diagonally left/right), but differ in their armored level. One type of tank wears a thin armor, as denoted by *Tank1*, hence they can be easily eliminated by any form of mines; while the other possesses a thick armor, as denoted by *Tank2*, and can only be destroyed by the highly explosive mines (represented in red in Figs. 5 and 8). All tanks are equipped with sonar sensors so that they have access to a set of observations on the minefield, including mine detection, agent detection and target bearing. A tank is rewarded with a positive value, and is eliminated when it collides with other agents or is hit by a mine. A screenshot of this domain is shown in Fig. 5.

In the experiment, we have a total of 8 tanks (divided equally for two types) and 40 mines with different explosion levels in a  $30 \times 30$  field. We let the tank agents execute the task every 300 trials of training and continuously perform this for a total of 3000 trials. Each execution is terminated when either all of the tanks are destroyed or they successfully complete the mission. We repeat the simulations for 30 times.

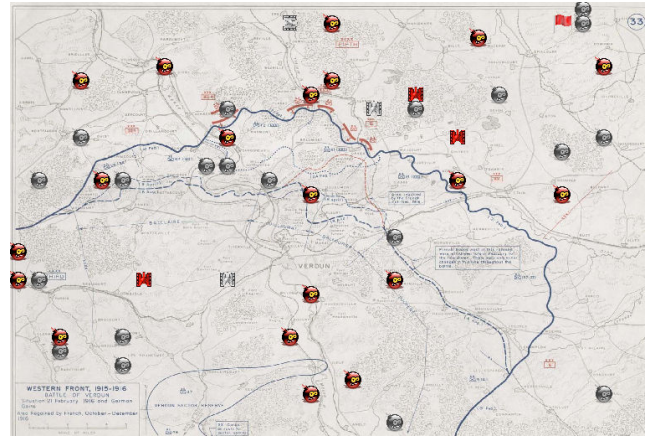


Fig. 5. A scenario of the minefield navigation domain. A team of tanks (*Tank1* is represented by gray rectangle while *Tank2* is represented by red rectangle) aim to capture the flag (red colored) around which mines (gray or red circle) are randomly scattered in the field.

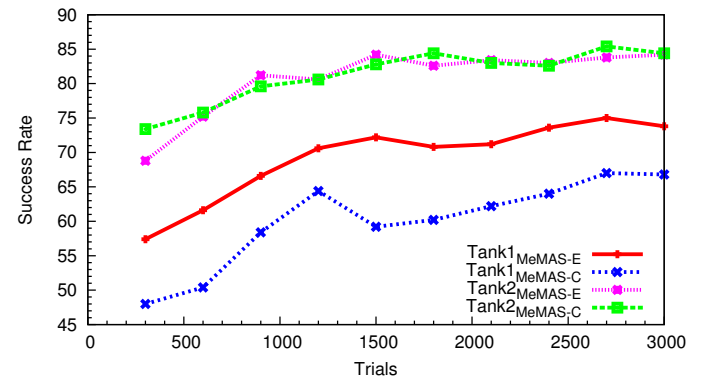


Fig. 6. Average Success Rate Performance of tank agents in both MeMAS-E and MeMAS-C on completing the missions in the minefield navigation domain.

Fig. 6 depicts the average success rates of both types of tanks on completing the missions in the minefield navigation domain. The *Tank2* agents performed competitively in both MeMAS-E and MeMAS-C. *Tank1* agents in the MeMAS-E (based on a like-attracts-like selection criterion), on the other hand, significantly outperforms its counterparts in MeMAS-C (based on a elitism selection criterion) on completing the missions successfully. In particular, in the MeMAS-E, tanks

<sup>1</sup>The entire game package is available upon request.

with thin armor (*Tank1*) have a much higher success rate than those in MeMAS-C. The results thus demonstrate the benefits of the like-attracts-like selection criterion over an elitism based scheme. In the MeMAS-C, we observed that *Tank1* prefers to select *Tank2* as the teacher agent since *Tank2* often succeeds in achieving the goals and becomes the elitist agent. However, this is unhelpful to *Tank1* in completing the goal. The knowledge imitated from *Tank2* actually led to the downfall of many *Tank1*, since their weaker armor property relative to *Tank2* could not survive the mines' devastation forces. Consequently, *Tank1* failed to repeat *Tank2*'s successful experience. On the other hand, by learning from a similar type of tanks within the environment, as in MeMAS-E, *Tank1* is able to truly imitate the appropriate skill of similar counterparts in achieving the robust performances observed by both *Tank1* and *Tank2*.

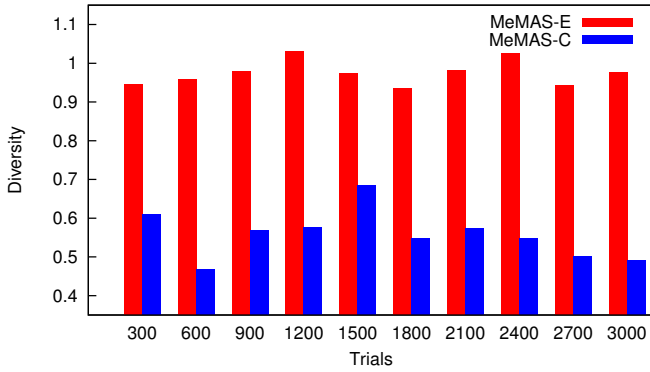
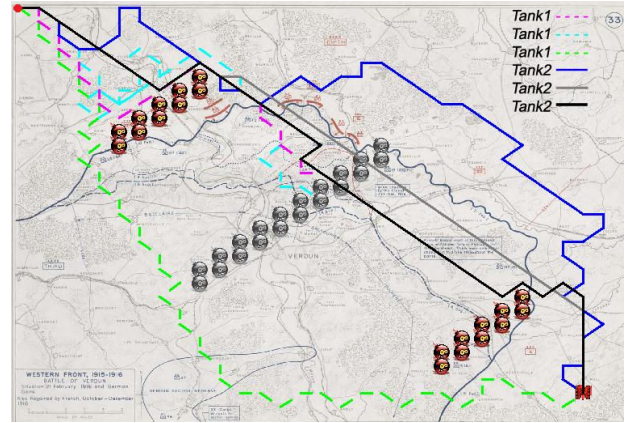


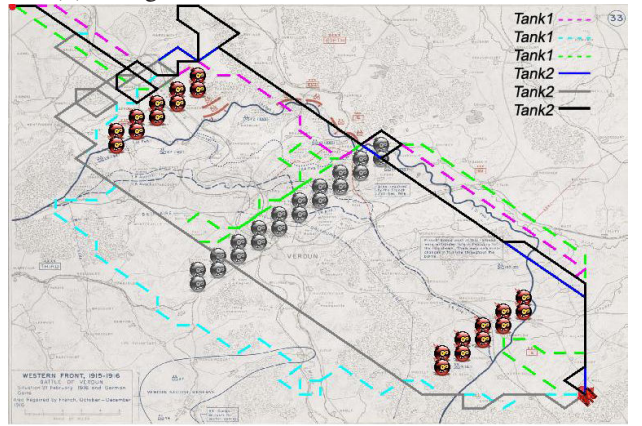
Fig. 7. The MeMAS-E maintains a larger diversity of tanks than the MeMAS-C in the minefield navigation domain.

We also computed the MeMAS diversity using information entropy, based on the type of tanks that evolves over time, and the results are summarized in Fig. 7. We noted the different types of tanks co-existing in the MeMAS-E. As the tank agents learn from both a similar type of tanks and the elitist tanks, they differ not only in the armor property, but also the emerged social behaviors. This is in contrast to the situation in the MeMAS-C where all agents learn from the elitist agents, thereby converging to successful but rather predictable behaviors. Note that high diversity is a desirable property of the MeMAS as it is an indicator of effective co-evolution of memetic agents on solving the task. It also contributes the emergence of human-like social behavior in the MeMAS.

We make a further step to depict a portion of typical navigation routes that the tank agents take in Fig 8. Three tanks of the first type (*Tank1*) follow the broken routes (pink, light blue and green broken lines) while the other *Tank2* use the solid trails (dark blue, black and grey solid lines). We observe that, in the MeMAS-C, most of the tank agents follow the routes in the same portion of the field (up triangle) (Fig. 8 (a)). The selection is mainly due to the fact that *Tank1* opts to follow *Tank2* when choosing *Tank2* as their teacher agent to imitate the routing behavior. Unfortunately, the imitated knowledge turns out to be a devastating path for *Tank1* since most of them are killed by the gray mines. The situation turns to better when the tanks follow the routes experienced by their similar types in MeMAS-E (Fig. 8 (b)). *Tank1* is observed to intelligently bypassing the black mines in the middle. Overall, the MeMAS-E showcases a high diversity of routes in the



(a) Navigation routes of tanks in the MeMAS-C



(b) Navigation routes of tanks in the MeMAS-E

Fig. 8. Tank agents learn to navigate towards the destination. Red dot at the top-left corner is the starting position. Mines with a red cover can hit any type of tank while gray mines can only destroy *Tank1*. *Tank1* follow the broken lines while *Tank2* use the solid lines.

minefield that resembles the desirable social behaviours found in human, thus leading to innovative and appealing ways of successful problem solving.

We further enrolled 18 participants to observe the behavior of the tanks on completing 18 tasks (including both failure and success cases) in the minefield. The observations include how the tanks move through dangerous areas filled with land mines, how they avoid the collision with other tanks, how they predict the unseen targets, and how they individually/collaboratively capture the flag. Firstly, we ask the participants to rate on both the diversity and intelligence of the tanks' actions, and then to rate the human-like performance of the MeMASs. We report the average scores (with the variance) of both MeMAS-C and MeMAS-E in Table I.

TABLE I. AVERAGE SCORES OF THE MEMASS' PERFORMANCE. 5 IS THE HIGHEST AND 1 IS THE LOWEST.

Criteria	Diversity	Intelligence	Human-like
MeMAS-E	3.78(0.73)	3.89(0.43)	3.89(0.32)
MeMAS-C	2.56(0.50)	2.42(0.31)	2.58(0.34)

Table I shows that MeMAS-E outperforms the MeMAS-C on all evaluation criteria. It is a bit surprising that the elitist principle based MeMAS-C does not exhibit much intelligence on solving the tasks. The failure probably results from the



incompatible actions that the tank agents learn from others without being aware of their personal types. The subject study also confirms that the trade-off between the elitist and like-attracts-like principles improves the human-like behavior of the MeMAS.

### B. 3D Interactive Homeland Defense Game

Here we designed a 3D interactive game that actively engages human-players in assessing the performance of the like-attracts-like principle via MeMAS-E. The game is an abstraction of the popular online *Plants v.s. Zombies*. We implemented the game based on the *Unity*<sup>2</sup> engine and integrated the MeMAS framework with the game engine. A game screenshot is shown in Fig. 9.



Fig. 9. Human-players intend to enter/defend the house (homeland) in the middle of the scene. They are building the forts while combating the offenders.

The homeland defense game is a defensive task for a human-player who aims to bring its avatars into a safe house (homeland) while eliminating the offenders. The defender can construct two types of forts, namely an arrow tower and a stone tower, that shoot its enemies within a certain range. Meanwhile, the offenders have two types of arms, light cavalry and heavy infantry. Light cavalry has a high agility and a thin armor, while heavy infantry has a low agility and a thick armor. The light cavalry can be eliminated immediately once it is shot by an arrow. Its health points are reduced by half when the light cavalry is attacked by a stone. In contrast, the heavy infantry can be killed by a stone when it gets shot. The health points are deducted by 5% of its original points when it is hit by an arrow. The game is terminated when either all offenders (non-player characters(NPCs)) are eliminated or all avatars controlled by a human-player have moved into the house.

We collected the data of 200 trials in the game-play and used them to train all of the NPCs. Both MeMAS-E and MeMAS-C were employed to train the NPCs on how they shall react in a game state. We then invited 21 persons (from novice to experienced game players) to play with the trained NPCs in the game. After that, the players were asked to rate the NPCs that have been trained by the MeMAS-E and the MeMAS-C respectively. They evaluated the NPCs based on

the answers to two questions: 1) How tricky and interesting are the routes selected by the NPCs? 2) how intelligent are the actions taken by the NPCs to attack the avatars? Finally, they also scored the NPCs' human-like behavior. We report the average scores (with the variance) in Table II. The last column in the table shows the success rate of the NPCs when they compete with the human-players.

TABLE II. AVERAGE SCORES OF THE MEMASS' PERFORMANCE IN THE HOMELAND DEFENSE GAME. 5 IS THE HIGHEST AND 1 IS THE LOWEST.

Criteria	Route Selection	Attack Mode	Human-like	Success Rate
MeMAS-E	4.24(0.47)	3.60(0.90)	3.88(0.50)	52%
MeMAS-C	3.07(0.72)	2.93(0.72)	3.12(0.50)	29%

The results demonstrate that the MeMAS-E improves the MeMAS-C not only on the human-like behavior of the NPCs, but also on the success rate. One interesting comment from the human-player is on the remarkable behavior of NPCs trained by the MeMAS-E: the heavy infantry moves slowly as most of its types do in the beginning, but it becomes a risk-seeker in the last minute of games when the safe house is protected by a circle of stone forts. This observation reveals the dynamic selection mechanism in the MeMAS-E. In the early stage of the game, the heavy infantry learns the behavior from its similar type as both the light cavalry and the heavy infantry perform smoothly without much interference from the under-developed forts. When the games progress, the light cavalry exploits its strength on the agility to break through the crowding forts thereby achieving more successful experience on hunting the avatars. Consequently, the heavy infantry selects the light cavalry as the teacher agent and acquires necessary skills while updating the type simultaneously.

## V. RELATED WORK

There has been seen a growing line of research on human-like behavior of intelligent agents. The *Soar* architecture with its update provides a cognitive model to develop believable agents and has demonstrated successful experience in computer games [14], [15]. In parallel, the *ICARUS* framework facilitates the implementation of goal-directed agents [16], [17]. Recently, a human-like agent shows believable interactions between users and virtual characters [18], [19]. Human-like agents also contribute the active research on humanoid robot where robots are expected to communicate with humans [20].

Computational intelligence methods have been well exploited to design human-like agents [21], [22], [23]. Hussain and Vidaver [24] used a genetic algorithm based framework to show real-time human-like performance of NPCs in games. Particularly, recent research on memetic automaton [7] has developed a more sophisticated evolutionary model for constructing human-like behavior of intelligent agents. Memetic algorithm has enhanced a personalized agent in an e-learning application [25]. More trials of memetic agents have appeared in the UAV pathfinding, character design in computer games and so on [8]. However, most of the work on memetic agents still employ the elitist principle to design the learning mechanism in the MeMAS.

<sup>2</sup><http://unity3d.com/>

## VI. CONCLUSION

Human-like behavior is a desirable property of multiagent systems that not only improves solutions to complex problems, but also allows the multiagent techniques to be seamlessly engaged in personal business. Recent research on MeMAS has shown promising results on human-like agents as the system naturally simulates the evolution of individuals and a human society. Following the imitate-the-elite principle, the MeMAS still drives the learning agent to select a teacher agent from an elite pool. Recognizing the importance of the like-attracts-like principle in human interactions, we further improve the human-like social behavior of MeMAS by establishing a new selection mechanism for the learning agent. The selection is a trade-off between the aforementioned two principles. Meanwhile, we weight the influence of each principle in a dynamic way, and advise the agent to choose a suitable teacher when all agents evolve over time. The performance is demonstrated in two problem domains.

While research has been conducted on human-like behavior for a long period, few formal methods have been found on quantifying the human-like behavior of intelligent agents. Most of behavioral evaluation still relies on the subject study, which is also the line we follow in this paper. Immediate research can be carried out to develop a quantitative formulation on human-like social behavior. We perceive that the elitist and like-attracts-like principles may imply some important factors, like diversity and intelligence of actions, in the formulation. We will take a further investigation and examine the formulation in various types of problem domains.

## VII. ACKNOWLEDGEMENTS

This work is partially supported by the Multi plAtform Game Innovation Centre (MAGIC) at Nanyang Technological University. MAGIC is supported by the Interactive Digital Media Programme Office (IDMPO) hosted by the Media Development Authority of Singapore. IDMPO was established in 2006 under the mandate of the National Research Foundation to deepen Singapore's research capabilities in interactive digital media (IDM), fuel innovation and shape the future of media. Yifeng and Xuefeng would thank the support from Teesside University Research Fund.

## REFERENCES

- [1] S. Liemhetcharat and M. Veloso, "Modeling and learning synergy for team formation with heterogeneous agents," in *Proceedings of the Eleventh International Conference on Autonomous Agents and Multiagent (AAMAS)*, pp. 365–374, 2012.
- [2] M. V. Lent, J. Laird, J. Buckman, J. Hartford, S. Houchard, K. Steinkraus, and R. Tedrake, "Intelligent agents in computer games," in *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pp. 929–930, 1999.
- [3] H. Lai, M. Wang, and H. Wang, "Intelligent agent-based e-learning system for adaptive learning," *International Journal of Intelligent Information Technologies*, vol. 7, no. 3, pp. 1–13, 2011.
- [4] A. Sloman, "Architectural requirements for human-like agents both natural and artificial. (what sorts of machines can love?)," Report BU-SCS-CSR-98-21, Birmingham Univ. (United Kingdom). School of Computer Science, 1998.
- [5] J. Orkin and D. K. Roy, "Automatic learning and generation of social behavior from collective human gameplay," in *Proceedings of Eighth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pp. 385–392, 2009.
- [6] G. Acampora, J. M. Cadenas, V. Loia, and E. M. Ballester, "A multi-agent memetic system for human-based knowledge selection," *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, vol. 41, no. 5, pp. 946–960, 2011.
- [7] L. Feng, Y.-S. Ong, A.-H. Tan, and X. Chen, "Towards human-like social multi-agents with memetic automaton," in *Proceedings of IEEE Congress on Evolutionary Computation (CEC)*, pp. 1092–1099, 2011.
- [8] Y.-S. Ong, M.-H. Lim, and X. Chen, "Memetic computation - past, present & future [research frontier]," *IEEE Computational Intelligence Magazine*, vol. 5, no. 2, pp. 24–31, 2010.
- [9] X. Chen, Y.-S. Ong, M.-H. Lim, and K. C. Tan, "A multi-facet survey on memetic computation," *Trans. Evol. Comp.*, vol. 15, pp. 591–607, Oct. 2011.
- [10] A. H. Tan, N. Lu, and D. Xiao, "Integrating temporal difference methods and self-organizing neural networks for reinforcement learning with delayed evaluative feedback," *IEEE Transactions on Neural Networks*, vol. 9, no. 2, pp. 230–244, 2008.
- [11] T. B. Gutkin, G. C. Gridley, and J. M. Wendt, "The effect of initial attraction and attitude similarity-dissimilarity on interpersonal attraction," *Cornell Journal of Social Relations*, vol. 11, no. 2, pp. 153–160, 1976.
- [12] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, 1951.
- [13] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Addison-Wesley, 2005.
- [14] J. E. Laird, A. Newell, and P. S. Rosenbloom, "Soar: An architecture for general intelligence," *Artificial Intelligence*, vol. 33, no. 1, pp. 1–64, 1987.
- [15] J. Z. Xu and J. E. Laird, "Combining learned discrete and continuous action models," in *Proceedings of the Twenty-Fifth National Conference on Artificial Intelligence (AAAI)*, pp. 1449–1454, 2011.
- [16] P. Langley and D. Choi, "A unified cognitive architecture for physical agents," in *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI)*, pp. 876–881, 2006.
- [17] D. Choi, T. Konik, N. Nejati, C. Park, and P. Langley, "A believable agent for first-person shooter games," in *Proceedings of the Third Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE)*, pp. 71–73, 2007.
- [18] E. Dourado and C. Martinho, "I mean it!: detecting user intentions to create believable behaviour for virtual agents in games," in *Proceedings of the Ninth International Conference on Autonomous Agents and Multiagent (AAMAS)*, pp. 83–90, 2010.
- [19] M. Thibaux, S. Marsella, A. N. Marshall, and M. Kallmann, "Smart-body: behavior realization for embodied conversational agents," in *Proceedings of the Ninth International Conference on Autonomous Agents and Multiagent (AAMAS)*, pp. 151–158, 2008.
- [20] A. Powers, S. B. Kiesler, S. R. Fussell, and C. Torrey, "Comparing a computer agent with a humanoid robot," in *Proceedings of the Second ACM SIGCHI/SIGART Conference on Human-Robot Interaction (HRI)*, pp. 145–152, 2008.
- [21] Y. Zeng, D. P. Buus, and J. C. Hernandez, "Multiagent based construction for human-like architecture," in *Proceedings of the Ninth International Conference on Autonomous Agents and Multiagent (AAMAS)*, pp. 63–63, 2007.
- [22] A.-H. Tan, Y.-S. Ong, and A. Tapanuj, "A hybrid agent architecture integrating desire, intention and reinforcement learning," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8477–8487, 2011.
- [23] Z. Raeesy, J. Brzostowski, and R. Kowalczyk, "Towards a fuzzy-based model for human-like multi-agent negotiation," *Proceedings of IEEE/ACM International Conference on Intelligent Agent Technology (IAT)*, pp. 515–519, 2007.
- [24] T. S. Hussain and G. Vidaver, "Flexible and purposeful npc behaviors using real-time genetic control," in *Proceedings of IEEE Congress on Evolutionary Computation (CEC)*, pp. 785–792, 2006.
- [25] G. Acampora, M. Gaeta, and V. Loia, "Exploring e-learning knowledge through ontological memetic agents," *IEEE Computational Intelligence Magazine*, vol. 5, no. 2, pp. 66–77, 2010.