AN AGENT-BASED MODEL TO UNDERSTAND A SIMPLE THEORY OF MIND: BELIEF REPRESENTATION SYSTEMATIC APPROACH (BRSA)

Zahrieh Yousefi①✉ , Dietmar Heinke①, Ian Apperly①, Peer-Olaf Siebers②

①School of Psychology, University of Birmingham, Birmingham, UK
②School of Computer Science, University of Nottingham, Nottingham, UK
z_yousefi@yahoo.com

Abstract. The natural ability of humans to mentally process others' behavior is central to their social life. This mental process relies largely on understanding the essential constituents of others' behavior such as their beliefs, desires and goals which is called Theory of Mind (ToM). ToM is the ability to attribute mental states to others, and to understand, predict, infer and justify people's actions by reasoning about their mental states such as their beliefs, desires and intentions.
This study presents an original agent-based model for two simple levels of ToM ability. It validates the general set of basic processes underlying a simple ToM ability called Belief Representation Systematic Approach (BRSA). BRSA has already been introduced in earlier study by the authors on an agent-based model for a false belief task. BRSA with its four cognitive phases including Perception, Memory, Reasoning beliefs and desires, and Expressing others' beliefs and desires in an action is a reliable cognitive framework for the processes of a simple ToM task. BRSA is capable of analysing ToM tasks within a network of cognitive structure. This cognitive structure is capable of constructing and deconstructing ToM general processes in a classified scheme. BRSA is an applicable analytical methodology that is able to filter unrelated abilities from ToM tasks, leading to more transparency in the results of the experimental task. BRSA has the potential to exhibit more complex ToM tasks through adjusting new features within its phases.
Key Words: Theory of Mind, Minimal Theory of Mind, Belief Representation, Agent-Based Model.

1 Introduction

People infer others' mental states in everyday life. They make inferences about others' unobserved mental states from the observed behavior. These inferences are often precise and indispensable for humans' social life. Infants start from tracking eye direction effortlessly and automatically while adults are competent in flexible and complex social reasoning (Apperly & Butterfill 2009). From the simple to the complex, mental processes influence the actions people take to reach their goals. People's beliefs, desires, emotions, and other mental states are a dependable guide to their future actions. How do humans infer others' mental states? What are the basic processes through which we can understand others' mental states? How can one construct connections with the mental states of others? How does understanding others' beliefs and desires improve one's performance in terms of achieving his or her goals? A large body of research addressing such questions consistently suggests that humans are capable of Theory of Mind (ToM), the ability to attribute mental states such as beliefs and desires, and to understand, predict and infer people's actions by reasoning about their mental states. Likewise, they may reason backwards to infer others' beliefs and desires from their actions (Baker, Saxe & Tenenbaum 2009).
Theory of mind research has recently been significantly increased through various fields. However, the areas of confusion and gaps are increasing and lack of standardization in the literature has been identified by researchers (e.g. Sharrock & Coulter 2009; Apperly 2012; Schaafsma, Pfaff, Spunt & Adolphs 2015). The underlying processes of ToM are still under considerable debate. For example, some studies of ToM experimental tasks and methods have included abilities that are more than ToM ability (Bloom & German, 2000) and such studies have therefore misinterpreted the ToM processes. The problem arises because some researchers tacitly regard ToM as a single, indivisible process, some consider it as a single brain network and other researchers combine varieties of ToM into one process (Schaafsma et al. 2015). Schaafsma et al. (2015) argue that a scientific concept of ToM requires a set of simpler processes rather than its current definition as the essence of a mental representation of minds, which does not permit an easy breakdown into its basic components. Consequently, they suggest the reconstruction of a concept of ToM with the necessary links to its more basic processes. To achieve this, Schaafsma and her colleagues propose two steps: breaking down ToM and its associated concepts into cognitive components that describe more basic processes, and then reassembling these basic blocks into different features of ToM.
The objective of this paper is to develop a structured understanding of the principles of a simple ToM process. For this purpose, an agent-based model has been implemented. Agent-based models (ABMs) provide a computational platform to simulate individual or collective autonomous entities as agents and replicate their actions. Agent-based simulations are able to evaluate the effect of agents' interactions within the environment. ABMs are capable of representing the rules and relationships between individuals' mental states, their actions, the environment and the procedure that they use to infer others' mental states on the micro level, while the macro level represents the aggregated effect of these actions, for example, how successful individuals perform in the environment. Cognitive science can greatly benefit from ABMs in the analysis of cognitive processes and social aspects of cognition through agents' interactions (Sun 2006). These potential characteristics of ABMs offer a reliable framework for understanding the underlying cognitive processes of ToM.
This paper offers a systematic set of basic processes underlying a simple ToM ability by validating Belief Representation Systematic Approach (BRSA) in the false belief task model (Yousefi, Heinke, Apperly & Siebers 2018) as the basic structure for a simple ToM process . "BRSA is a simple and robust approach that breaks down false belief tasks into four fundamental cognitive phases, including Perception, Memory, Reasoning beliefs and desires, and Expressing others' beliefs and desires in an action" (Yousefi et al. 2018). We argue that BRSA is capable of reconstructing at least simple ToM levels by reassembling a new feature in the reasoning phase and reasoning phase.

1.1 Theory of Mind (ToM)

Theory of mind was defined as a system of inferences about the mental states of others that are not directly observed and it can be used as a theory to make predictions about others' behavior; it was first developed by Premack and Woodruff in 1978. In an article they asked "Does the chimpanzee have a theory of mind?" and they suggested that one has a theory of mind if he/she could impute mental states to self and others (Premack & Woodruff 1978). While psychologists frequently approve this definition as a standard, researchers Sharrock and Coulter (2009, p.69) argue that "Any such studies can only efficiently address the question of what specific capacities may legitimately be ascribed to non-human primates".

Despite nearly forty years of extensive research in ToM, "the appearance of consensus on what theory of mind is, and how it could be studied is misleading" (Apperly 2012, p.2). Apperly suggests three distinct approaches to study ToM: the conceptual domain, the cognitive perspective and the social competence that can vary across individuals. The cognitive perspective involves the architecture and process of ToM; it clarifies the way that belief representation works at a cognitive level. For example, the role of language in the false belief task process or the way that people inhibit their own perspective and take the perspectives of others into account (Apperly 2012). The conceptual approach explains questions such as whether a child has belief representation, or at what age this concept emerges, or how children acquire ToM concepts (Apperly 2012). The conceptual domain also includes conceptual knowledge of others' beliefs and desires that interconnects with the behavior (Apperly 2012). Apperly also suggests that future research will benefit from clearly expressing what aspect of ToM we wish to measure, because each of these directions considers different questions that require different approaches.

This study mainly concentrates on the cognitive perspective of ToM by representing a systematic approach for the processes of a simple ToM ability in a virtual society. It also examines conceptually which agents are capable of ToM ability and why. Thus, this section continues with a succinct literature review on the cognitive domain of ToM.

1.2 The Cognitive Perspective

Experimental research shows that language, memory, and executive functions are critical in ToM competence (e.g. Apperly et al. 2007; Hughes 1998; Marcovitch et al. 2015). The link between executive function and ToM ability, and particularly the role of inhibitory control in children and adults has been studied extensively. For example, Russell (1996) argues that deficits in executive control in autistic children hinder the emergence and expression of their ToM ability while Wellman (2014) states that ToM development corresponds with executive function but not directly.

The study by Leslie and Polizzi (1998) offers a model to pass a false belief task. They suggest that ToM mechanism (ToMM) nominates a true belief as the belief state. Hence, to pass the false belief task, it is necessary to inhibit the default content of the belief, which is true, and change the focus of attention to the alternative belief. In case of inhibition failure, the default content will be allocated to the belief state, which is inaccurate in the case where the target has a false belief. The concept of selection processing (SP) is "to select the most plausible belief content from a small set of plausible candidates" (Leslie, German, & Polizzi 2005, p.51). Leslie et al. (2005) explain SP as an automatic process, which is associated with ToMM and attributes beliefs and desires to the agent. The concept of inhibition is identified through the algorithms of the model in this paper and will be considered in the discussion section.

Another characteristic of ToM relates to the level of complexity involved in the task. For example, there is a distinction between higher-level and lower-level processes of ToM in relation to the level of inferences, storing and using information involved (Apperly 2011). Reasoning is an important part of higher-level ToM processes.

1.3 Two Systems Account

Apperly and Butterfill (2009) have advocated the 'two systems' proposition, as a parallel cognitive construction ToM. The nature of these two distinct systems rests on a compromise between flexibility and efficiency. The first system is fast and cognitively efficient and capable of tracking others' registration of an object rather than belief representation as such, but is inflexible and limited. They suggest that such a system may account for the success on some ToM tasks by human infants, some non-human animals such as chimpanzees and human adults under cognitive load. The second system is associated with cognitively demanding but flexible and slow processing and exists in human adults, parallel with the first system. System one is automatic whereas system two is non-automatic and requires reasoning.

1.4 Minimal ToM

Butterfill and Apperly (2013) proposed a novel approach to minimal ToM to represent the ability to track the perception, knowledge and beliefs of others. Clearly, this is an elaborated version of the first system in their account of two systems. Butterfill and Apperly developed a distinctive minimal form of ToM cognition that involves representing "belief-like" states without any cognitive demands or conceptual sophistication. Their argument starts with a fundamental question: "what could someone represent that would enable her to track, at least with limits, other's perception, knowledge states and beliefs including false beliefs?" (Butterfill & Apperly 2013, p.1) They then formulated four principles to answer this question. The first principle relates to "a minimal grasp of goal-directed action" (Butterfill & Apperly 2013, p.10) that one might understand the goal from bodily movements such as tracking others' visual direction or a change in gaze direction. In the second principle, Butterfill and Apperly (2013) introduce two terms, "field" and "encountering" as basic characteristics of perception. The concept of field relates to a set of objects at any given time that is determined by factors such as proximity, eye direction and barriers. The concept of encountering involves the relation between the agent and an object in the agent's field. According to Butterfill and Apperly (2013), to execute a goal-directed action in pursuit of an object, it is necessary to encounter it first. The third principle, "registration", introduces a new notion of a belief-like concept. Registration is the relation between an agent, an object and its location. The agent registers the location of an object as it encounters the object. A correct registration is a precondition for a successful goal-directed action. Butterfill and Apperly state one's correct registration of an object becomes incorrect by moving or destroying the object in her/his absence. One example for an application of the registration principle is the scrub-jays re-cache food experiment by Clayton, Dally and Emery (2007). In this experiment, scrub-jays have only chosen to re-cache the food that was previously stored in the presence of competitors who had seen where they cached it. Butterfill and Apperly suggest that scrub-jays understand that competitors' correct registration of food results in stealing their cached food in a successful goal-directed action. Therefore, they re-cache the food to prevent competitors from correctly registering its new location (Butterfill & Apperly 2013). The fourth principle involves a shift to thinking of a successful registration as a causal factor for the agent's action. For example, in the non-verbal false belief task by Onishi and Baillargeon (2005), infant subjects and an observer watch while an object is placed in a black box. In the absence of the observer, the object is moved to a white box. When the observer comes back, infants' looking times indicate that they correctly expect that the observer will reach into the black box. Onishi and Baillargeon suggest that the infants are ascribing beliefs about the object to the observer. However, Butterfill and Apperly's alternative explanation is that infants track the registered location of the black box as a cause of action.

2 Mental State Model (MSM) Methodology

An agent-based model for a simple ToM ability is implemented in Repast Simphony ((The Repast Suite, 2017), called Mental State Model (MSM). MSM clarifies the basic processes of inferring others' mental states in a simple ToM ability. MSM comprises interactions between three types of agents, representing different capabilities of understanding others' desires and beliefs in the environment with their heterogeneous set of rules: Infer agents, MinToM agents and Control agents. The environment consists of a grid of 50 by 5. The neighbourhood (Ngh) of an agent is the area around it where the extent of its X-axis and Y-axis are equal. Field of movement consists of the vicinity cells around agents, the first neighbourhood cells, in which agents are able to move. Field of view refers to the neighbourhood in which an agent is able to perceive the environment within the fourth neighbourhood (the square of 9 by 9 cells with agent's location in the center). The goals that agents are required to achieve are called targets, illustrated by green cells in the environment. The number of targets in the environment remains constant through each time step of the simulation. Agents move to a cell which is not occupied by other agents.

2.1 Mental States (MS)

Two mutually exclusive mental states are designed for agents; an Active mental state represents the necessity of achieving a target whereas a Passive mental state signifies that agents are not able to achieve any target in the current time step. At the start of the simulation all agents are Active. However, their mental state changes subject to achieving a target. Unless an agent achieves a target, its mental state remains Active. The mental state of an Active agent that achieves a target changes to Passive. When an agent's mental state becomes Passive, it remains Passive for two time steps.

2.2 Control Agents' Strategy

Control agents as simple reactive agents observe their neighbourhood to search for targets. They are able to recognise their own mental states. Active Control agents move to the nearest cell that contains a target while the Passive Control agents move to a cell which does not contain a target but where there is a target in the vicinity of the cell.

Control agents' perception include a target sensor and memory sensor meaning that they collect information about the targets in their field of view and use the information at the current time step, shown in Table 1. Control agents' reasoning to select a target is based on assessment of the shortest distance between agent and target, thus, they move towards the nearest target. Figure 1 shows the arrow and box diagram of the Control agents.

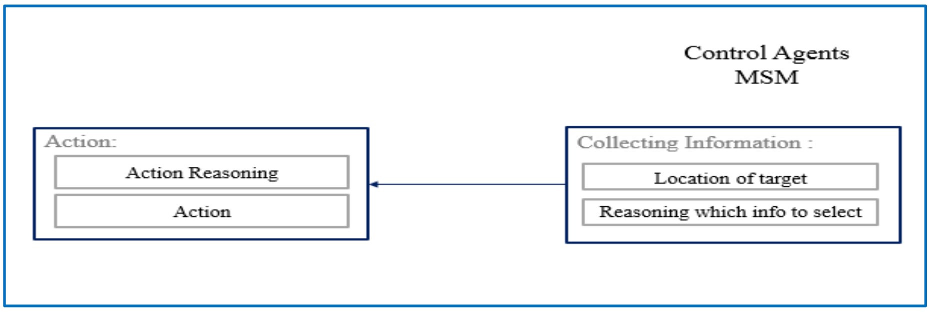

Figure 1. Control agents' arrow and box diagram

Table 1. The agents' perceptions, understanding Mental State (MS) and memory capabilities

Agent	Perception (sensor)		Understanding MS			Memory		
	Target	Agent	Own MS	Tracking others' field	others' MS	Sensor Memory	Short-term Memory	Long-term Memory
Control	✓		✓			✓		
MinToM	✓	✓	✓	✓		✓	✓	
Infer	✓	✓	✓	✓	✓	✓	✓	✓

2.3. Minimal Theory of Mind Agents' (MinToM) Strategy

MinToM agents are reactive agents that collect information about the target and also other agents. They register the location of the target from others' perspectives to track their field of view in the current time step and move towards a target. Figure 2 shows the arrow and box diagram of the MinToM agents; MinToM agents' procedure includes collecting information, registering others' perspective and an action to move towards a target.

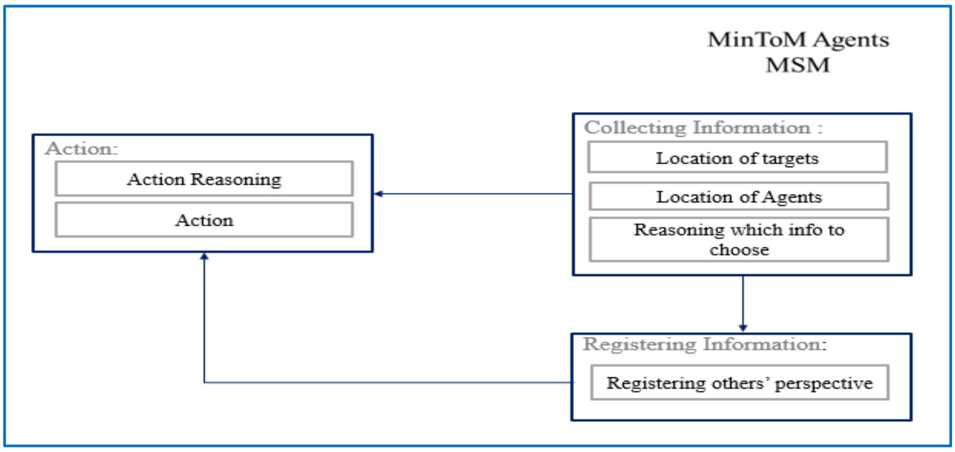

Figure 2. MinToM agents' arrow and box diagram

MinToM agents move on the basis of two factors, the availability of targets and their own mental states. In the case where there is no target in the field of movement and there is more than one target in its field of view, Active MinToM agents move towards a target with the minimum number of agents around it. On the other hand, Passive MinToM agents move towards an empty cell in the vicinity of a target which is surrounded by the minimum number of competitors.

MinToM agents are able to track others' field of view and store the information in their short-term memory. Short-term memory facilitates agents' calculations and keeps the information for processing in each time step.

3.4 Infer Agents' (Infer) Strategy

Infer agents observe other agents' actions and infer their desire from their actions. They collect information about the location of targets, other agents and their actions. They store information, reason and act based on others' perspectives. When they observe that an agent achieves a target, they infer that agent becomes Passive. Otherwise, they assume the agent is Active. Their inferences are based on their perception from their neighbourhood. They store and remember other agents' inferred mental states, until the agent moves out of the Infer agent's field of view. Infer agents' strategies are similar to MinToM agents' strategies if they do not have the related information necessary to infer the mental states of others.

Suppose ms is the variable for Infer agents' MS, NoOfTargets is the number of targets in the first Ngh of Infer Agent, expandFieldOfView is the function that searches for targets one Ngh wider. Also TargetWithMinActive is a function that calculates and returns which target has the minimum number of Active agents around it whereas TargetWithMaxActive is a function that calculates and returns which target has the maximum number of Active agents around it.

The pseudo code of cases are as follows:

```
Case 1) If (ms = Active & NoOfTargets = 1) Then
        Moveto (target)
        Achieve (target)
Case 2) If (ms = Active & NoOfTargets = 0) Then
        expandFieldOfView()
        set output of TargetWithMinActive () to minTarget
        Moveto (A cell in Ngh of minTarget)
Case 3) If (ms = Active & NoOfTargets > 1) Then
        TargetWithMaxActive()
        set output of TargetWithMaxActive () to maxTarget
        Moveto (maxTarget)
        Achieve (maxTarget)
Case 4) If (ms = Passive & NoOfTargets = 1 Then
        Moveto (A cell in Ngh of target)
Case 5) If (ms = Passive & NoOfTargets = 0) Then
        expandFieldOfView()
        set output of TargetWithMinActive () to minTarget
        Moveto (A cell in Ngh of minTarget)
Case 6) If (ms = Passive & NoOfTargets > 1) Then
        set output of TargetWithMinActive () to minTarget
        Moveto (A cell in Ngh of minTarget)
```

2.5. Infer Agents Processes (IAP)

Infer agents processes (IAP) consist of four phases shown in an arrow and box diagram in Figure 3: Collecting Information, Recording Information, Reasoning Process of Beliefs and Desires, and Expressing Others' Mental States (Actions as Output) as follows:

- **Collecting Information Phase.** Infer agents collect information about other agents' perspectives from their field of view as well as the location of food and other agents. They observe other agents' actions in reaching a target and store this information in their long-term memory. Collecting Information is a dynamic and online procedure, which is parallel with the changes of the world over time and highly interconnected with other phases.

- **Recording Information Phase.** The memory demand is indispensable in order to understand others' mental states. Infer agents store other agents' desires and beliefs regarding the targets in their memory. They are able to store the information by exploiting three relevant types of memory. The first type, sensory memory, relates to simple information from the neighbourhood such as the location of a target in the current time step. The second type, short-term memory, is designed to store simple calculated information about different agents' perspectives in the current time step, for example, in registering and tracking others' field of view. The third, long-term memory, stores the inferred information about other agents' desires and beliefs for future access. The two main distinctions between these types of memory relate, firstly, to the length of time that information remains in the memory and secondly to the volume and complexity of their content, for example, information about others' perspectives.

- Reasoning Process of Beliefs and Desires Phase. This phase defines a central information processing unit for Infer agents. They initially choose the agents with the same beliefs about a shared target and infer their desires in each time step. Once the Infer agent perceives that another agent achieves a target, it concludes that agent's mental state was Active until it achieved its target. Thus, the agent's mental state changes to Passive in the current time step. This enables Infer agents to infer others' desires. They inhibit their own desires towards the target and temporally retrieve the other agents' perspectives from memory. The reasoning phase for Infer agents involves the following five generic subroutines:

- Inferring others' desires from their observed behavior
- Processing its own desire and belief
- Self-perspective inhibition of its own desire
- Retrieving information about the subject agents' mental states (from memory)
- Subject agents' desire and belief process

By the end of this phase, Infer agents' reasoning about the other agents' mental states are completed.

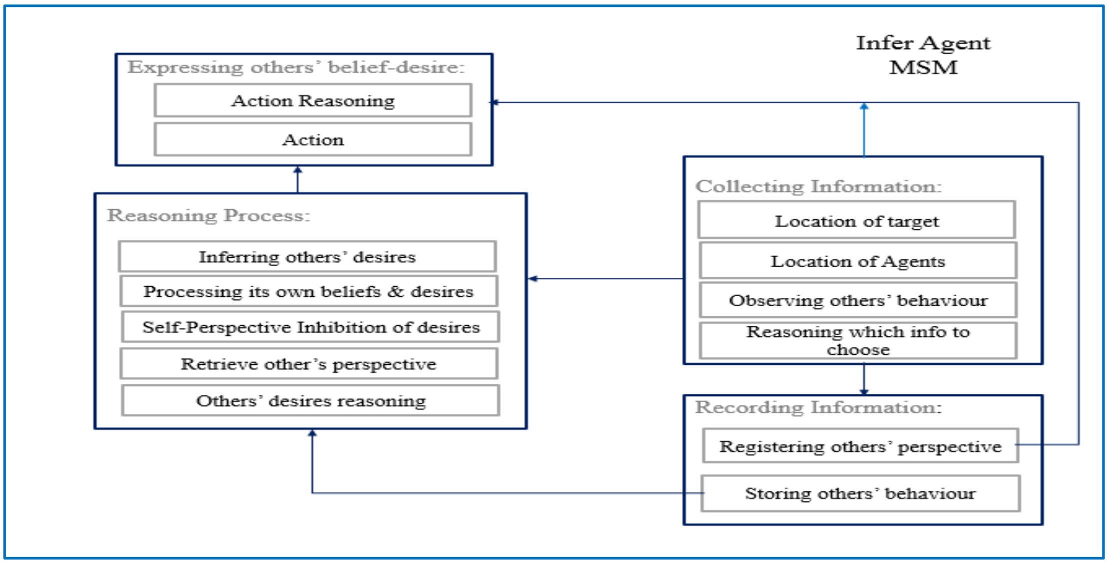

Figure 3. Arrow and box diagram of Infer Agents' Processes (IAP)

- **Expressing Others' Mental States (Actions as Output) Phase.** This phase of the MSM is about deciding on actions by considering the mental states of others. There is a delicate but important distinction between understanding others' mental states and using (or expressing) this understanding in Infer agents' actions. In the first three IAP phases, agents understand others' mental states and in the expressing phase they use and express this understanding. In two identical situations in the environment, Infer agents act differently based on whether the other agent's mental state is Active or Passive. Infer agents use the information resulting from their ToM ability in this phase and express that use through their actions. A level of inevitable reasoning in this phase is based on the principles of rational action to achieve their goals. This level of reasoning combines the concept of understanding others' mental states through an action.

2.6 MSM Implementation

The MSM design comprises two abstract classes: Society and Agent. Agents are moving to achieve their targets in a grid space where Society contains all of the other classes including Agent, Target, and Grid. Figure 4 shows the MSM class diagram including the classes, the relationships between them and their methods. Society class is a container for objects of other classes. Agent class is an abstract class including three main subclasses: Infer MinToM Agent, Agent and Control Agent.

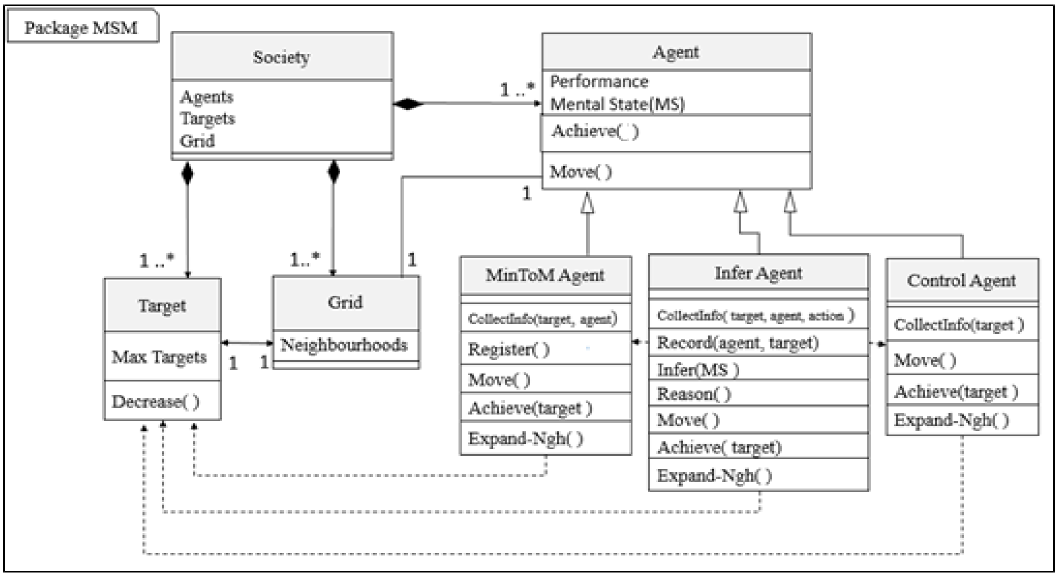

Figure 4. Class diagram of MSM

3. Results

The parameters and their values which have been systematically selected to consider extreme values is shown in Table 2. The simulation is run 10 times for each parameter setting, which obtains reliable patterns and diverse results. The time steps is set to 1000.

Table 2. Parameters and the values

Number of agents (N)	Number of targets (T)
50	50
400	400
800	800
1200	1200
1600	
2000	

3.1 Agents' Performances

The total number of time steps in which Active agents fail to achieve a target is calculated to present the performance measurement. The agents' performances have been normalized to accomplish a standard scale. The results demonstrate a general pattern; Infer agents perform constantly more efficiently than the MinToM agents and similarly MinToM agents' performances are higher than Control agents. The results also demonstrate that the largest performance differences occur between MinToM agents and Infer agents.

The density of the number of targets, number of agents and the ratio between them is critical to determine how agents use their strategies. Figure 5 demonstrates the agents' performances in an extreme situation and shows that when the number of targets = 50, and the number of agents = 50, the difference in agents' performances are not significant. As the number of agents increases to 2000, the differences are greater, yet not significant. The number of targets is very low in comparison to the world and number of agents, thus, it is often less likely that the agents use their strategies.

In extreme situations in which the density of the targets is much higher than the number of agents, they simply achieve targets without exploiting their strategies. For example, in Figure 8, due to the excessive number of targets in the environment, agents are able to simply achieve targets without using any of their ToM abilities. Thus, the performance of agents does not directly reflect their ToM abilities. As the number of agents increases, the performance differences between Infer and MinToM agents increases in almost all situations. For example, Figure 6 shows that with 400 targets, the Infer agents' performances improves due to the increased number of chances to apply their strategy. Therefore, the performance differences between Infer and MinToM agents increases. Figure 7 also demonstrates the pattern of increasing performance differences between agents as they are able to apply their ability to infer others' desires and beliefs.

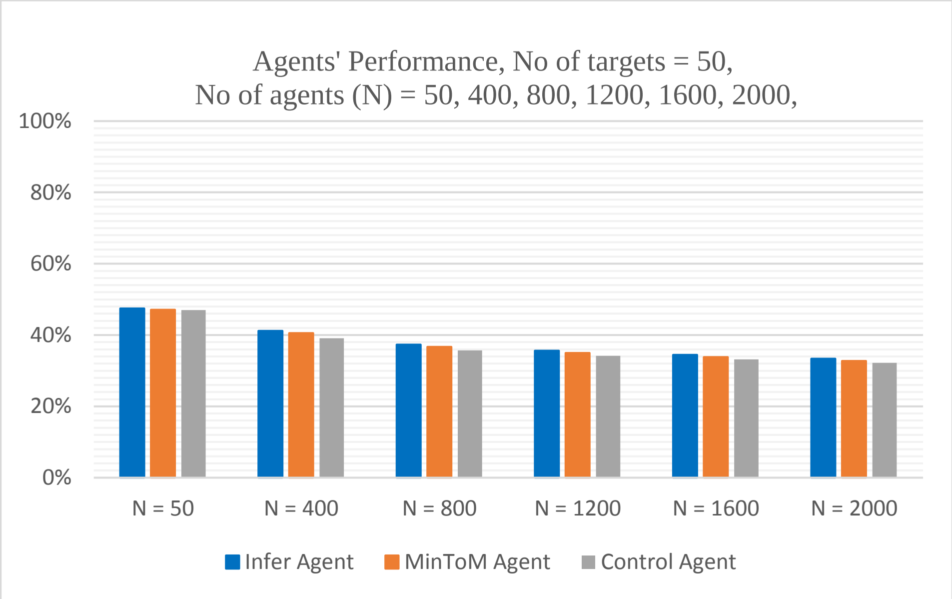

Figure 5. Performance of agents, No of targets = 50

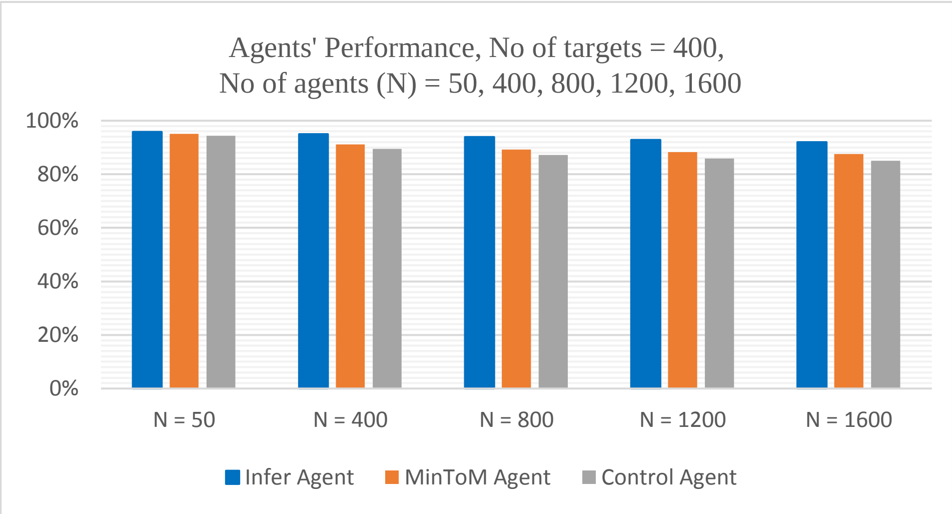

Figure 6. Performance of agents, No of targets = 400

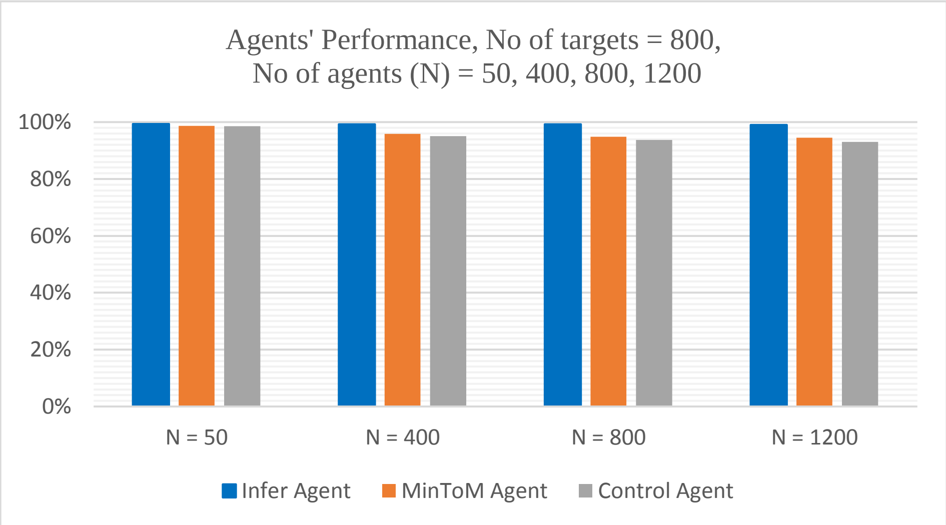

Figure 7. Performance of agents, No of targets = 800

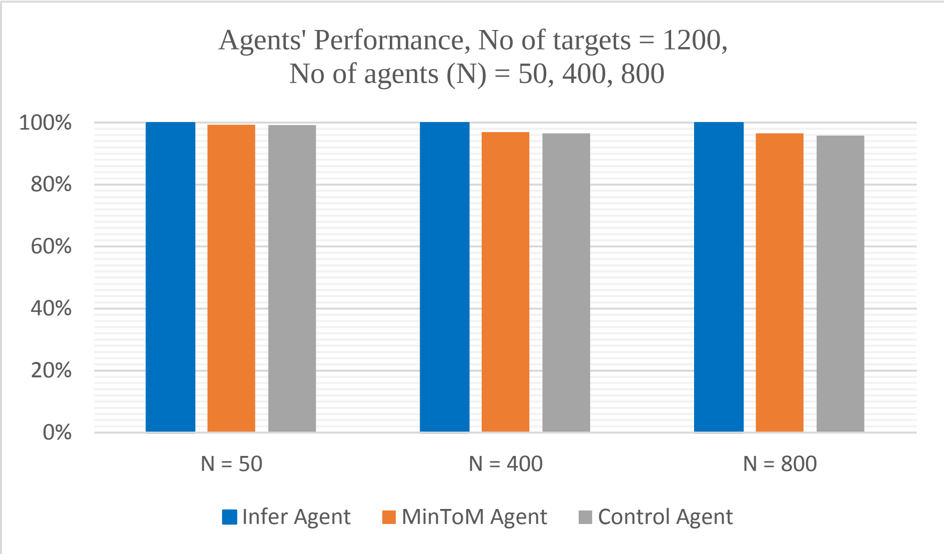

Figure 8. Performance of agents, No of targets = 1200

3.2 The Cost and Resources of ToM

The ability of Infer agents to understand others' mental states demands more resources and costs. These resources consist of a network of perception and attention (first phase of IAP diagram), three types of memory for storing information (second phase of IAP diagram), inhibitory control and reasoning resources (third phase of IAP diagram) which are demonstrated in Table 1.

Moreover, in order to calculate the cost of inferring others' mental states, the isolated processing time of the agents is computed for each type of agent. The results demonstrate that Infer agents have the slowest processing time; possible reasons for this are the reasoning time involved, additional information to process and also that the inhibition step consumes more time in retrieving information from memory. Hence, based on the IAP diagram, the infer system comes at the cost of time and use of interconnected resources including memory, perception, reasoning and inhibitory control and extra information.

4 Discussion

The Mental State Model (MSM) consists of three types of agents with two mutually exclusive mental states regarding the target: Active or Passive. Agents, including Control, MinToM and Infer agents, move to achieve a target when they are in an Active state, whereas they cannot achieve any target while in a Passive state. Infer agents are able to infer other agents' mental states from observing their behavior in achieving a target. MinToM agents register other agents' perspectives regarding the location of targets. Control agents are able to only consider the location of targets.

MSM presents the fundamental cognitive units of simple ToM processes through the IAP diagram. This set of basic processes addresses some of the ambiguity behind the ToM tasks by breaking a task into the standard phases. It is also able to simplify and explain some of the complex characteristics within the tasks through four phases: perception, memory, reasoning beliefs and desires, and action.

Collecting Information Phase. The first phase involves agents' perception (collecting information from the environment) which mainly corresponds to information about others' mental states. This phase is consistent with the second principle of minimal ToM, field of view and encountering (Butterfill & Apperly 2013) that enables agents to watch other agents or objects in their field of view by encountering them. This phase associates to reasoning to determine what information regarding the target is more important to select and in what priority. For example, the information about the nearest targets and the agents with a shared target has the highest priority. Infer agents collect information by observing other agents' actions as they are reaching a target, this is analogous with literature that one of the social information gathering skills in human infants in real life is decoding the social environment information and discerning the information about an entity such as an object (Baldwin & Moses 1996). IAP diagram shows that the perception phase is interconnected with all phases as a dynamic online access to the information.

Recording Information Phase. The second phase involves storing the mental state information into agents' memory. MSM elucidates memory demands on Infer agents to code and decode agents' desires and beliefs. Memory is indispensable in inferring others' mental states. Similarly, there is consistence in agreement regarding the role of memory in ToM processes in the literature, e.g. memory is one of multiple domains in social understanding processes (Mitchell, Macrae & Banaji 2004). As the agents' ToM abilities develop on the micro level, the required memory advances from a simple sensory memory to short-term memory and long-term memory. MinToM agents need sensory memory and short-term memory to register the location of the target and track others' field of view in the current time step. In addition to these, Infer agents need a long-term memory to store others' desires and beliefs.

Reasoning Process of Beliefs and Desires Phase. The third phase, the reasoning process of beliefs and desires, encompasses five subroutines in agents: Inferring others' desires from their observed behavior, Processing its own desire and belief, Self-perspective inhibition of its own mental states, Retrieving information about the others' mental states, Subject agents' desire and belief process. Agents start reasoning about their own beliefs and desires; they then temporary inhibit their own perspective regarding the target and take protagonists' perspectives into account by retrieving the information from memory. The IAP diagram suggests the reasoning phase acts as a central information-processing phase for ToM competence. In addition, the contribution of the executive function to belief representation (Apperly 2011) has been elucidated by this phase; the role of 'memory' identified in retrieving the information subroutine, the role of 'inhibitory control' presented in the self-perspective inhibition subroutine, and the role of 'reasoning' revealed in the selective process of belief and desire subroutine.

Belief representation is accomplished by the end of the reasoning phase. However, agents have not exhibited any evidence of this understanding in their behavior yet.

Expressing Beliefs and Desires of Others Phase. The fourth phase relates to expressing the understanding of others' mental states as an action or behavior. It means performing an action by using a mental representation (Hughes 2011). In real life, the action of this phase might be as simple as an eye gaze and eye tracking or a more complicated action that requires additional reasoning. Based on the IAP diagram, there is a distinction between having ToM ability and using this ability in an action. This distinction would prevent the confusion in ToM measurement tests in the literature, such as the linguistic ability required for the false belief task in children.

Moreover, the IAP diagram is consistent with the developmental literature; it demonstrates a network of resources for Infer agents with ToM competence, including perception, memory, self-perspective inhibition and reasoning resources.

4.4 ToM Ability and Agents' Performances

The results of simulation consistently suggest that there is a strong relationship between the agents' ToM capability and higher performance in competitive society. As the ability of ToM develops from the simple level of understanding their own mental states to the level of tracking others' field of view and to the more complicated level of constructing inferences about other agents' mental states, the performance of agents effectively rises. These findings are consistent with ToM's impact on social situations in real life. However, the Infer agents' high performance comes at the cost of time as they use more time to collect information, reason, inhibit their own perspective and retrieve the information regarding others' mental states.

4.3 A Comparison Between Simple ToM Processes and BRSA

The IAP diagram in Figure 3 demonstrates a simple ToM process within four basic cognitive phases, compatible with the Belief Representation Systematic Approach (BRSA). BRSA encompasses four phases consisting of collecting information as perception, recording information in memory, a reasoning process of belief and desires, and expressing the mental state of others as an action, illustrated in Figure 9.

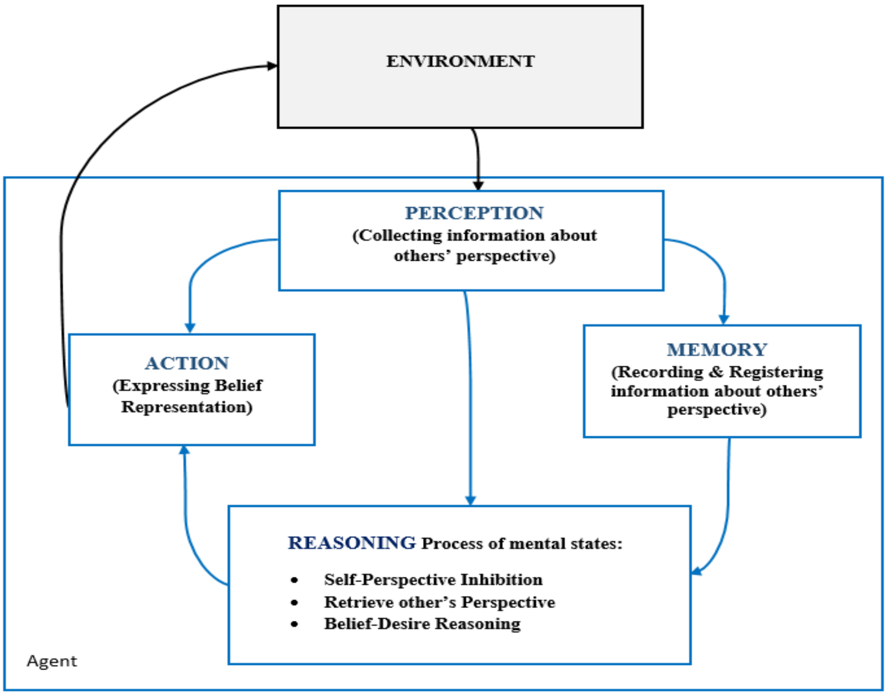

Figure 9. Belief Representation Systematic Approach (BRSA) (Yousefi et al. 2018)

The underlying process of simple ToM phases, as we might expect, are equivalent to the BRSA phases. BRSA reveals the basic fundamental process for belief representation; it is flexible enough though for modification in more complex scenarios. However, the more complex nature of the processes in a simple ToM process, rather than in a false belief task, demands the reasoning phase entails a new subroutine adjustment: the inferring others' metal states subroutine.

BRSA is a reliable systematic approach for ToM processes, and would contribute in analysing ToM tasks within a network of cognitive structure. For example, BRSA would be an analytical methodology for distinguishing behavioral tasks from ToM tasks. This cognitive structure enables constructing and deconstructing ToM processes; it also is capable of filtering unrelated abilities that mislead the results of experimental tasks. BRSA delivers a classification to the ToM research that reduces ambiguity and confusion. BRSA is flexible enough to accommodate more complex ToM tasks through adjusting new features within its phases.

4.4 Presumptions of Mental States (Biases)

Infer agents have their presumption about others' mental states. The presumptions are made when agents lack any reasons to think in a different way; these default assumptions (or biases) constitute the most suitable and productive way to make generalisations and often they are not correct (Minsky 1988). Infer agents are capable of updating their presumptions regarding others' mental states as they observe other agents achieving targets. At the start of the simulation, Infer agents assume other agents' mental states are Active, which is correct. However, this changes over time, and the presumptions are not necessarily accurate anymore and will be updated by Infer agents' online inferences. Also, MinToM agents' presumptions are that all agents are Active.

5. Conclusions

This study develops an effective cognitive framework for a simple ToM process. This framework validates the Belief Representation Systematic Approach (BRSA). BRSA encompasses four phases consisting of collecting information as perception, recording information in memory, a reasoning process of beliefs and desires, and expressing the mental state of others in an action. BRSA is a reliable standard framework for ToM processes, and would contribute to analysing ToM tasks within a network of cognitive structure. This cognitive structure enables constructing and deconstructing ToM processes in a general classified scheme. BRSA is an applicable analytical methodology and is also capable of filtering behavioral tasks and the unrelated abilities from ToM tasks leading to more transparency in the results of experimental tasks.

BRSA is flexible enough to accommodate more complex ToM tasks through adjusting new features within its phases. Therefore, the future research will benefit from BRSA as a basic and standard structure to analyse ToM processes.

References

Apperly, I. (2011). *Mindreaders: The Cognitive Basis of "Theory of Mind".* Hove: Psychology Press.

Apperly, I. (2012). What is "theory of mind"? Concepts, cognitive processes and individual differences. *The Quarterly Journal of Experimental Psychology, 65,* 825–839.

Apperly, I. A. & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*(4), 953–970.

Apperly, I. A., Back, E., Samson, D. & France, L. (2007). The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task. *Cognition, 103*(2), 300–321.

Apperly, I. A., Samson, D., Chiavarino, C., Bickerton, W.-L. & Humphreys, G. W. (2007). Testing the domain-specificity of a theory of mind deficit in brain-injured patients: Evidence for consistent performance on non-verbal, "realityunknown" false belief and false photograph tasks. *Cognition, 103*(2), 300–321.

Baldwin, D. A. & Moses, L. J. (1996). The Ontogeny of Social Information Gathering. *Child Development, 67*(5), 1915–1939.

Baker, C. L., Saxe, R. & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*.

Butterfill, S. & Apperly, I. (2013). How to construct a minimal theory of mind? *Mind and Language, 28*(2), 606–637.

Clayton, N. S., Dally, J. M. & Emery, N. J. (2007). Social cognition by food-caching corvids. The western scrub-jay as a natural psychologist. *Philosophical Transactions of the Royal Society B, 362*, 507–552.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*, 233–253.

Hughes, C. (2011). *Social Understanding and Social Lives: From Toddlerhood through to the Transition to School.* Hove: Psychology Press

Leslie, A. M. & Polizzi, P. (1998). Inhibitory processing in the false belief task: Two conjectures. *Developmental Science, 2*(1), 247–253.

Leslie, A. M., German, T. P. & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology, 50*(1), 45-85.

Marcovitch, S., O'Brien, M., Calkins, S., Leerkes, E., Weaver, J. & Levine, D. (2015). A longitudinal assessment of the relation between executive function and theory of mind at 3, 4, and 5 years. *Cognitive Development.*

Minsky, M. (1988). *The Society of Mind.* New York: Simon & Schuster.

Mitchell, J., Macrae, C. & Banaji, M. (2004). Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *The Journal of Neuroscience, 24*(21), 4912–7.

Onishi, K. H. & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255–258.

Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 515–526.

Russell, J. (1996). *Agency: Its role in mental development*. Hove: Erlbaum (UK), Taylor & Francis.

Schaafsma, S., Pfaff, D., Spunt, R. & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences, 19*(2), 65-72.

Sharrock, W. & Coulter, J. (2009). Theory of mind: A critical commentary continued. In Leudar, A. Costall. (eds.) *Against Theory of Mind.* Basingstoke: Palgrave Macmillan.

Sun, R. (2006). *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation.* New York: Cambridge University Press.

Wellman, H. (2014). *Making Minds. How Theory of Mind Develops.* NewYork: Oxford University Press.

Yousefi, Z., Heinke, D., Apperly, I., Siebers, P.O. (2018) An Agent-Based Model for False Belief Tasks: Belief Representation Systematic Approach (BRSA). In: Thomson R., Dancy C., Hyder A., Bisgin H. (eds.) *Social, Cultural, and Behavioral Modeling.* SBP-BRiMS 2018. Lecture Notes in Computer Science, vol 10899. Springer, Cham. Springer, Cham. pp.111–126.