

## RESEARCH ARTICLE

# Rational Coordination in Cognitive Agents: A Decision-Theoretic Approach Using ERMM

NOUSHIN MAZHAR<sup>ID</sup> AND MARYAM KAUSAR<sup>ID</sup>

Department of Software Engineering, Foundation University Islamabad, Islamabad 44000, Pakistan

Corresponding author: Noushin Mazhar (noushin.mazhar@fui.edu.pk)

This work was supported by Foundation University, Islamabad.

**ABSTRACT** Over the years, research in multi-agent systems has become increasingly popular. Agents evolve by interacting with their environment and must communicate with other agents in order to do various cooperative tasks. The research aims to provide efficient coordination among cooperative cognitive agents in unpredictable multi-agent situations. Xiang's rational agent model addresses scenarios when no social conventions or predefined communication protocols exist for the agents' interaction and then makes decisions by recursive modeling. We address the deficiencies of the loosely coupled framework and the problem of mispredictions in Xiang's architecture. The solution is based on Lawniczak's Architecture for generic cognitive agents and an enhanced model of Xiang's Recursive Modeling Method for coordinated decision-making in multi-agent situations. We instruct the cognitive agent to learn about other agents from past mispredictions and then consider its best choice. The feedback module is incorporated so agents can learn to maximize their joint expected reward. The model filters the mispredictions and evaluates the error rate. We compare the enhanced method with the Recursive Modeling Method. The results show that mispredictions are corrected from 33% to 10.9% and errors in perception are reduced from 22% to 0.097%, as the system progresses. Overall, the approach demonstrates superior performance. It significantly lowers the rate of mispredictions about other agents' actions and takes 30% to 42% less time and 55.4 % fewer moves than RMM.

**INDEX TERMS** Cognitive agent, multi-agent system, recursive reasoning, agent reasoning, agent-oriented methodologies, loosely coupled framework, decision making, MAS communication, agent coordination, recursive modeling method.

## I. INTRODUCTION

Artificial intelligence is a broad field of study that produces intelligent machines. Technology is constantly trying to mimic human intelligence, which is why AI is currently receiving so much attention [1]. AI has been extensively used across numerous fields, including healthcare systems [2], social media trend prediction [3], citation prediction [4], and multi-agent systems [5].

Agent is a modern computer program that works independently on behalf of the user [6]. However, agent-based systems are rarely stand-alone systems. These systems are usually required to collaborate and interact with other entities (i.e., agents, humans, etc.); such a system is called a multi-

agent system [7]. Though the agents are intelligent, their behavior is restricted within the predefined scenario. The restricted behavior of reflexive agents encourages the need to design and develop cognitive agents. These agents keep track of what happened in the past and can predict what will happen in the future; in short, the most vital part is memory: which utilizes previous knowledge or perception to anticipate what is yet to come. Cognitive agents contain explicit representations for their intentions, beliefs, and goals [8] and can handle the problem at hand by executing "Cognitive Acts". The Cognitive Act is comprised of five things: 1) Perceiving: perceiving information from the environment, 2) Reasoning: reasoning the information perceived using existing knowledge, 3) Judging: judging the gained information, 4) Responding: responding to other cognitive agents or the environment; and 5) Learning: evolving the existing

The associate editor coordinating the review of this manuscript and approving it for publication was Taehong Kim<sup>ID</sup>.

knowledge [9]. Interest is growing in multi-agent systems that motivate coordination and cooperation mechanisms among cognitive agents.

Coordination is considered one of the critical aspects of a multi-agent system; by introducing efficient cooperation in the distributed multi-agent environment, complex tasks can easily be performed beyond individual agents' capabilities. Frameworks for cooperative multi-agent decision-making can be divided into loosely coupled frameworks (LCFs) and tightly coupled frameworks (TCFs). In TCFs, agents communicate through messages, the joint belief of team agents is well defined, and each TCF agent's belief is consistent with the joint belief. However, the cost of setting up the organization of a TCF-based multi-agent system is high. In LCFs, agents do not communicate but rely on observing other agents' actions to coordinate. The decision-making in multi-agent LCF is constrained by making guesses about the actions of other agents. Any misinterpretation can degrade the system's performance. [10].

The motivation behind this research is to propose a mechanism that allows multiple agents to coordinate their actions so that individual actions do not adversely affect the overall system's performance. Effective coordination among agents in dynamic environments can be accomplished by extending the agents' learning ability to know the beliefs of other agents present in their environment. We aimed to build cognitive agents capable of better decision-making in unpredictable multi-agent situations by predicting the actions of other agents.

In this research, we addressed the deficiencies of the loosely coupled framework and propose a communicative loosely coupled framework for cognitive agents. The objective of this study is to extend the decision-theoretic approach of the Recursive Modeling Method (RMM) for coordinated decision-making in a distributed multi-agent environment where each agent models other agents recursively to choose its action [10]. These recursive models represent the agent's information about itself and other agents in the environment. The agent recursively predicts the other agents, thus building the payoff matrices, and this modeling terminates when the agent has no more knowledge. Thus, the no-information model occurs at the leaf nodes of the RMM hierarchy.

In Xiang's agent model [10], an agent's model of other agents can be recursively nested to any depth; the model 1) expands exponentially, and 2) the computational cost is elevated as the number of recursive levels increases. The agent lacks communication and bases its decision on possibilities. The consideration of all the possibilities about the other agent's actions can elevate the computational cost and deepen the RMM hierarchy. We are motivated to make the decision-making process more efficient by introducing communication within a loosely coupled framework. Knowledge sharing is considered soft evidence about other agents' beliefs and is used to update the beliefs of the receiving agent. Knowledge sharing is performed whenever a communication channel is available. Each agent builds an observation

history and calculates mispredictions whenever feedback is available. Using this approach, the chances of mispredictions about each agent's actions are deflated. It will control the misjudgment of the joint action plan, thereby improving future decision-making.

Section II illustrates the related work. The proposed methodology is presented in Section III. Section IV demonstrates the experimental setup and illustrates the results and performance evaluation. Section V discusses the findings of the study. Section VI explains the future work and limitations of the study. Section VII gives a detailed conclusion.

## II. RELATED WORK

Agents are autonomous, goal-oriented, and proactive [11]. The focus of agents is to make decisions. Decisions vary from utility maximization to complex planning problems [12] to provide sustainability solutions [13], [14]. Agents need to consider the functioning of other agents, coordinate with them and act strategically to make a good decision. Over the past few years, immense research has been done on the mechanism of the agent coordination. Decentralized approaches in several kinds of multi-agent systems, e.g., swarm robotics [15], modular robots [16], and sensor networks. Reference [8] are inspired by collective behaviors in nature, where coordinated global behavior involves local interactions.

For a cooperative multi-agent system, new reinforcement learning methods are required to learn decentralized policies. Jakob N. Foerster et al. have proposed counterfactual multi-agent (COMA) policy gradients based on the actor-critic method [17]. This methodology has utilized a centralized critic to calculate the Q-function and to optimize the agents' policies; decentralized actors are used. However, with many agents, it is harder to train centralized critics and challenging to coordinate.

There is an assumption that in distributed algorithms, all agents have equal privileged information. Chih-Han Yu et al. proposed an algorithm in which agents do not have uniform information. Each agent has state variables and can sense neighborhoods passively or actively by communicating with agents [18]. Since the agents lack uniform information, there is a possibility that neighbor's states might have incorrect states. So the agents have to reason on their own to make effective decisions.

Ante Vilenica et al. proposed coordination spaces dedicated to coordination to separate the coordination and application functionality. This work implies different coordination strategies without altering the functionalities of agents [19]. We say that agents are hard-coded to make their decisions in predefined strategies. The capability of agents to predict and reason their environment is suppressed.

In the past, the communication protocol between agents was manually specified, and it was not changed during training. In order to make agents learn about communication among themselves, Sainbayar Sukhbaatar et al. proposed a

neural model [20]. A deep feed-forward network is used to control each agent. Every agent has access to a communication channel through which they receive the summed transmission from other agents. The authors trained the model via backpropagation because of continuous communication, but it has limitations in controlling heterogeneous agents.

Multi-agent coordination has become an important study in intelligent target-tracking systems. In research [21] by Xing et al., it is believed that to control a group of agents, collaboration among them is required. To control agents, track the targets in a cooperative manner. Kalman filter algorithm is optimized for target tracking and the Leader-Follower control method is utilized to keep the agents in control, in which one agent is assigned as a leader and tracks the target, and the other agents follow the leader. It has a significant impact on single-agent control and single-target tracking.

In the game environment, Zhang et al. [22] analyzed a multi-agent collaboration strategy. They proposed a learning method to measure cooperative information between multiple agents. They conducted a Nash equilibrium game strategy on the specific multi-agent game problem of territory defense. The deep Q-learning method was applied to learn the defender's joint defense strategy. In practice, the problem of partial observability in a multi-agent distributed environment needs to be considered.

In a research [23] by Gmytrasiewicz and Durfee, a framework based on decision theory, is presented about agent interactions when no communication protocols are defined. They use the decision-theoretic model of rational decision-making under uncertainty, according to which an agent should make decisions to maximize its expected utility [12], [24], [25].

Ethics and values highly influence human decision-making. Values are considered a criterion for deciding and accomplishing one's task. In recent studies, the methodologies are proposed for AI agents to align with human values [26] and estimate value preferences for human choices to motivate and guide AI agents in the hybrid participatory system [27]. In the future, authors are motivated to provide direct feedback to AI agents or use Natural Language Processing algorithms to identify values automatically.

In another piece of research, authors explored a partial evaluation to improve the efficiency of the MAE agent [28]. It significantly improved centralized decisions but not distributed decisions. Xiang et al. proposed a distributed decision algorithm for Collaborative Decision Networks (CDNs) based on partial evaluation [29]. For agent communication, they considered online decisions over short horizons. The agents decide the best joint action based on their knowledge of the current environment. CDNPE gave significantly optimal results under the pivot probability assumption. More study is underway to check decision optimality as pivot probability is relaxed.

Gmytrasiewicz and Durfee designed an agent to anticipate and model other agents' actions to make its own decision that maximizes its utility. This consideration of the other agents' actions leads to the recursive levels of modeling. This

Recursive Modeling Method (RMM) achieves coordination in unpredictable environments. The RMM agents depict reasonable behavior, but the approach can become greedy, and agents can only act to maximize their utility.

Further, in this research area, the authors partitioned the cooperative multi-agent decision-making frameworks into frameworks where the individual agent contains a single variable and one where the individual agent possesses a complete internal model. However, significant research efforts have been dedicated to SVFs, e.g., [30], and LCFs, e.g., [31], in comparison with those to TCFs [32], [33], [34].

The latter can further be partitioned into tightly coupled frameworks (TCFs) and loosely coupled frameworks (LCFs). TCFs provide the agent's communication through interfaces, and the agent's belief is accordant, without any self-conflict. Whereas in LCFs, the agents do not communicate with each other instead, they observe the actions of other agents to make their own decisions. In [10] Xiang and Hanshar contributed by comparing TCFs and LCFs. From LCFs, they selected a decision theory recursive modeling method (RMM). As agents do not directly communicate but coordinate their actions, each agent observes the other agents within its defined neighborhood. Xiang's decision-making agent model is limited by making guesses about other agents' actions, resulting in the wrong estimation of reward distributions and an incorrect Bayesian update. Another limitation of RMM reported was the presence of several optimal joint action plans promising the same maximum expected reward, and agents may follow the different joint action plans, thus hindering the system's performance.

Since cognitive computing is a multidisciplinary research area that aims to devise decision-making mechanisms and computational models, we are interested in proposing a communicative, loosely coupled framework for Cognitive Agents. According to a survey [35], different authors defined cognitive computing in different times and contexts. A cognitive architecture defines the fundamental structure of an intelligent system. So research in cognitive architecture has also been encouraged for building intelligent agents as competent as humans [9].

### III. PROPOSED METHODOLOGY

In a multi-agent system, agents coexist in an environment and potentially interact with each other to make decisions. Multi-agent decision-making is a distinguishing feature of the MAS. Decision-making can be done in a distributed manner when no central control is available, and each agent can make decisions on its own. For distributed decision-making, an appropriate coordination mechanism is needed. Inspired by Xiang's concept, we provide a communicative, loosely connected framework for a cooperative multi-agent system. In the proposed system, we have extended coordinated decision-making using RMM and embedded it with the feedback module to reduce mispredictions among the agents and enhance their capability to make fewer faulty decisions.

## A. MULTI-AGENT ENVIRONMENT

The common characteristics of the multi-agent system framework are:

- 1) The system contains a finite number of agents and environment states, and each agent can take a defined set of actions.
- 2) The system must operate on a discrete timestamp. The acts carried out by the agents at the same timestamp are all considered to be occurring concurrently.
- 3) The environment is discrete, comprising a grid of cells; each cell is either empty or occupied by an object. The objects are defined according to the application domain.
- 4) The agent's neighborhood is defined, and a cell visited by an agent within the neighborhood is marked "visited".
- 5) The environment is stochastic as the agents' actions are uncertain and partially observable, which means the agent can only perceive its neighborhood.

## B. ASSUMPTIONS

Because of different multi-agent systems, it is logical to define the framework: a finite number of agents, a finite set of actions that the agent can take, and a discrete time step. In a general framework of MAS, we expect:

- 1) All agents can only observe those agents that come into their defined neighborhood and can only perceive the part of their actions, i.e., the other agents' actions, that might be hidden from view.
- 2) An agent can take action based on its preferences in each state and make better decisions than other agents depending on the expected payoff.
- 3) At any time step, the agent's actions are not dictated by any other factor except the current state of the environment.
- 4) For decision-making, agents do not communicate through communication protocols, but each agent decides its action by reasoning about the actions of other agents through recursive modeling.

## C. COORDINATING COGNITIVE AGENTS ARCHITECTURE

This research proposes a coordinating cognitive agent architecture that is based on 1) Lawniczak's architecture for generic cognitive agents [9] and 2) the Enhanced Recursive Modeling Method (ERMM). Considering the layered approach of Lawniczak's architecture, we applied the Enhanced Recursive Modeling Method as a decision-theoretic approach for coordinating multi-agent interactions on the Perceptual, Reasoning, Learning, and Response Layers.

### 1) PERCEPTUAL LAYER

The basic functionality of the Perceptual layer is to perceive the environment in which agents act. In our scenario, the perceptual layer gathers information regarding the

observable neighborhood of an agent. Each agent in recursive modeling reasons about the information perceived from the environment.

### 2) REASONING LAYER

The Reasoning layer logically deals with the perceived information from the environment. Cognitive agents reason about other agents through a hierarchy of recursive models. The model considers all the information available and summarizes all the uncertainties as a set of probability distributions. Payoff matrices are generated to analyze the possible joint actions and their utilities. Using this representation, an agent explicitly explains how the joint agents' actions can influence its utility. The value of each cell represents the reward of each joint plan. The agent chooses the joint plan with the highest expected accumulative reward.

### 3) RESPONSE LAYER

The Response layer executes the agent's action in the joint plan with the highest probability (0 to 1) that maximizes its utility.

### 4) JUDGING LAYER

The Judging layer plays a vital role in enhancing the recursive model. It allows loose communication of the agents using a blackboard architecture. A blackboard communication allows agents interactions and provides soft evidence regarding the observation of the actions of other agents. The error occurs if the prediction goes wrong. A history (knowledge base) is maintained, containing all the mispredictions that have led to errors. It will be used in the future to update the probability distribution to influence actions.

### 5) LEARNING LAYER

Agents consider their experiences when making future decisions. Past knowledge or history helps to update the agent's beliefs. For the wrong predictions, the reward is reduced to the lowest, i.e., zero, to avoid them in the future. The perception of an agent about other agents gets refined, thus reducing the error risk. A joint strategy is computed, which is not only rational for one interaction but might be productive in repetitive situations.

## D. ENHANCED RECURSIVE MODELING METHOD (ERMM)

In multi-agent systems, where there are no predefined communication protocols, an agent hypothesizes other agents' actions by using knowledge about itself, other agents, and the environment. The agents within the common neighborhood comprise a group of size  $g$ , and the planning horizon is a single cell. The agent can perceive its own location and the locations of other agents in its neighborhood. The coordination is purposeful when a certain number of agents exist in the neighborhood. The observable neighborhood is fixed and shown in Figure 1.

It is how an agent models other agents to choose its course of action. It explains the alternative courses of action.

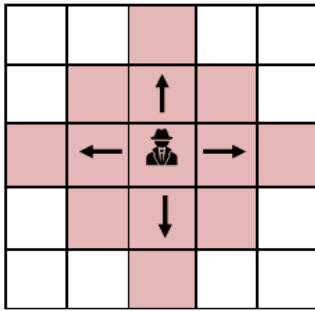


FIGURE 1. Observable neighborhood of a predator.

This reasoning gives rise to the recursive nesting of models. An agent must choose actions that maximize its expected utility. The recursive modeling terminates when the agent's knowledge comes to an end, and its knowledge depth has reached. The Zero-knowledge model shows that the agent has no knowledge about the agent being modeled, as shown in Figure 2. Based on its limited knowledge, the agent gives equal probability to the actions of other agents.

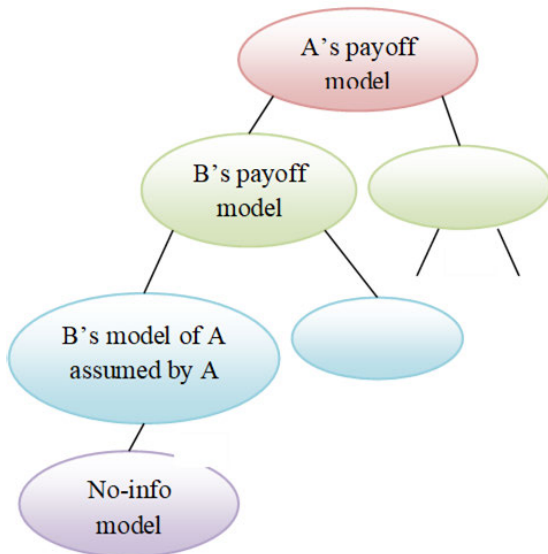


FIGURE 2. Interaction of Agents.

In Enhanced RMM, the payoff matrix is used to represent the agent's actions and preferences. It is a building block to represent the decision-making characteristics of the agents. For a multi-agent system consisting of  $n$  agents, a payoff matrix has  $n$  dimensions for each agent. In the payoff matrix, each dimension is related to each agent. Each cell of the matrix represents the utility or payoff for each action. The agent can take four possible directions (North, South, East, and West). So, the number of cells in the payoff matrix can be calculated as is  $k^n$ . Each entry corresponds to the agent's expected payoff from the joint action plan of all the other agents. As explained in Figure 3, where the y-axis represents

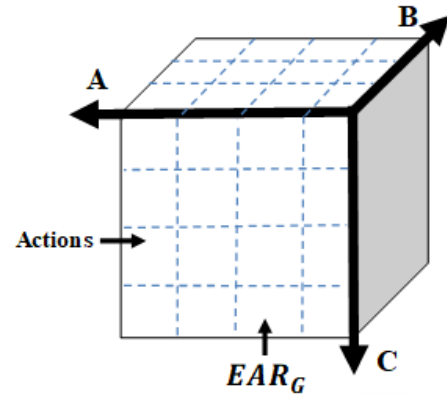


FIGURE 3. Payoff matrix.

TABLE 1. Summary of Notations.

| Notations   | Description                                            |
|-------------|--------------------------------------------------------|
| $G$         | the group of agents, where agent $x \in G$             |
| $mv$        | joint action plan                                      |
| $mv^i$      | joint action plan at $i$ th step by all agents in $G$  |
| $mv^{x,i}$  | action of agent $x$ at $i$ th step                     |
| $t$         | group trajectory                                       |
| $c^i$       | the group configuration after joint action $mv^i$      |
| $ps^{x,i}$  | current position of agent $x$ at $i$ th step           |
| $rw_y(c^j)$ | the reward $y$ gets from the group configuration $c^j$ |

the actions of agent B. Each cell in the matrix contains the payoff of the relevant joint action plan  $mv$  in Group  $G$ .

Let the order of joint actions of agents in  $G$  be,

$$mv = (mv^1 \dots mv^k) \quad (1)$$

Suppose a group trajectory be

$$t = (c^1 \dots c^k) \quad (2)$$

whereas  $c^i$  is the group configuration after joint action plan  $mv^i$

Configuration  $c^i$  comprises each agent position,

$$c^i = \{ps^{x,i} | x \in G\} \quad (3)$$

The payoff or reward is then computed as the expected accumulative reward for group  $G$

$$EAR_G(mv) = \sum_{y \in G} \left\{ \sum_t [P(t|mv) \sum_{j=1}^k rw_y(c^j)] \right\} \quad (4)$$

where the second summation is the sum of all the possible group trajectories.

As the group configuration  $c^i$  only depends on its current configuration and current joint action plan  $mv^i$ , we have

$$P(t|mv) = \prod_{i=1}^k P(c^i, mv^i) \quad (5)$$

As the position of each agent  $x$  in  $c^i$  is based only on its own action  $mv^{x,i}$ , we have

$$P(c^i, mv^i) = \prod_{x \in G} P(p_s^{x,i}, mv^{x,i}) \quad (6)$$

By combining the above equations, (4) becomes

$$EAR_G(mv) = \sum_{y \in G} \left\{ \sum_t \left[ \left( \prod_{i=1}^k \prod_{x \in G} P(p_s^{x,i}, mv^{x,i}) \right) \sum_{j=1}^k rw_y(c^j) \right] \right\} \quad (7)$$

Agents choose the joint plan  $mv$  with the highest EAR.

### 1) FEEDBACK MODULE

Xiang's agent model [10] for decision-making is limited by making guesses about other agents' decisions and states. Its perception of agents is limited, and knowledge is insufficient. Moreover, considering all the possibilities about the other agents' actions takes computational cost and a more deep RMM hierarchy. The misjudgment of the agents' actions can take the system to the misjudgment of the optimal joint plan. Overall the system can pursue less optimal solutions. To overcome this problem, we have introduced the feedback module, as shown in Figure 4.

The feedback module is comprised of four sub-modules:

- 1) Agent Communication
  - 2) Error Evaluation
  - 3) History Build Up
  - 4) Update Agent's Belief
- 1) **Agent Communication**

In previous approaches, the topology of agent communication either does not exist or is tightly coupled. The predictions made by the agents need to be verified. Thus, we allow agent communication to overcome the shortcoming of Xiang's agent model. The ERMM introduces a blackboard architecture for agent communication. As agents choose the joint action with the maximum expected reward, this needs to be shared among other agents. RMM has limitations when multiple optimal joint plans with the same expected reward exist and agents choose different joint plans.

The proposed agent model loosely couples the agents by introducing indirect, loose communication. We have used a common information space to make information available about the agents' actual actions and predictions. The predictions about agents' actions are verified as correct or incorrect. Agents pinpoint their mispredictions, analyze each other, and align their actions to achieve a system goal. This enhancement leads to a

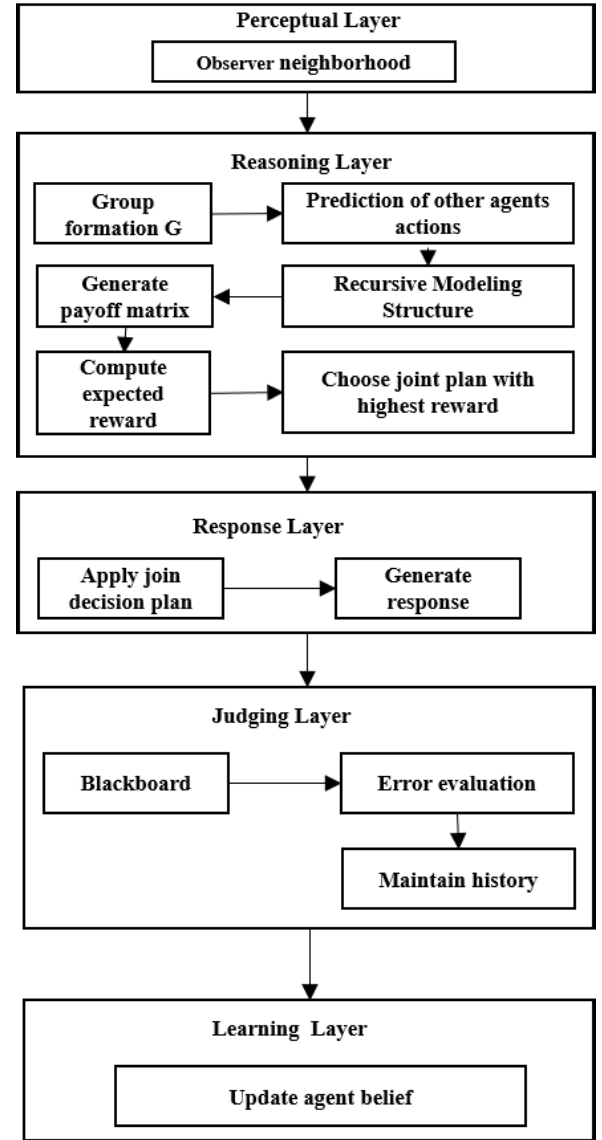


FIGURE 4. Coordinating cognitive agents architecture.

more flexible agent model and helps to handle uncertainty in the joint decision plan.

### 2) Error Evaluation

This study focuses on enhancing the agent model by knowledge sharing to find and evaluate mispredictions. Knowledge is made available about the agents' predictions and actual decisions through the common information space. We compute the time integral of the absolute error using the following equation.

$$Integral\ of\ absolute\ error = \int_0^{\infty} |E(t)| dx \quad (8)$$

$E(t)$  represents the error of an agent's perceptions about other agents' actions. It is the comparison between the predicted decision of the agent and its actual decision.

The error is evaluated at every time step. Error evaluation addresses the mistakes made by the agents about the other agent's actions. It is further used to penalize predictions having the error.

### 3) History Build Up

As an agent explores the environment and interacts with other agents, it builds up its observation history. It is the backup record of all the mispredictions observed and wrong guesses made by the agents. History serves as the knowledge base and contains sufficient knowledge about the previous actions of others. It keeps track of past mispredictions and takes more reliable moves in similar situations.

Based on the observation history, the reward distribution of the agent is updated. For mispredictions, the reward is reduced to zero to avoid the decision. It helps to avoid the wrong optimal joint plan and eliminates the chances of misjudgments of other agents' actions in the future.

### 4) Update Bayesian Belief

The state of the physical environment depends on the actions of other agents. The probabilities of the actions of agents change the environmental states. These probabilities are obtained by solving its recursive models. The changes in the observations need to be updated. After knowledge sharing among the agents and summarizing the agent's observation history, the agent's belief is updated using the Bayes rule. For agent A, its belief  $P(\text{nbp}_A^{B_1} \dots \text{nbp}_A^{B_{g-1}})$  is the reward distribution around the agent and it needs to be updated for the new neighborhood.

Unilateral is the reward collected by the single agent. Now suppose  $\text{nbp}_B^A$  shows the unilateral rewards in the area of agent A that is unobserved to agent B, whereas,  $\text{nbp}_B^A \in \{\text{allLow}, \neg\text{allLow}\}$ . If there is at least one high reward in the unobserved area, then  $\text{nbp}_B^A = \neg\text{allLow}$ . Let  $\text{lmv}_A^B$  be the last movement of agent B observed by agent A. Based on the history of  $\text{lmv}_A^B$ , the probability distribution in the new neighborhood is updated. If the last movement observed turned into a misprediction then the reward  $\text{rw}$  value is set to default or lowest. The reward values are sustained for the correct predictions. For agent A,  $P(\text{nbp}_A^B, \text{nbp}_A^C | \text{lmv}_A^B, \text{lmv}_A^C)$  is updated, making agent A learn about its bad decisions and stay almost close to the optimal beliefs of the environment and other agents.

In the pursuit problem, four cognitive agents named *predators* aim to capture the agent named *prey*. Predators coordinate their actions to surround their prey. The enhanced RMM is implemented on predators to enhance their decision-making capability. The ERMM implementation of predator uses a payoff matrix to plan its movements.

The predators aim to minimize the distance between themselves and the prey. The payoff value depends

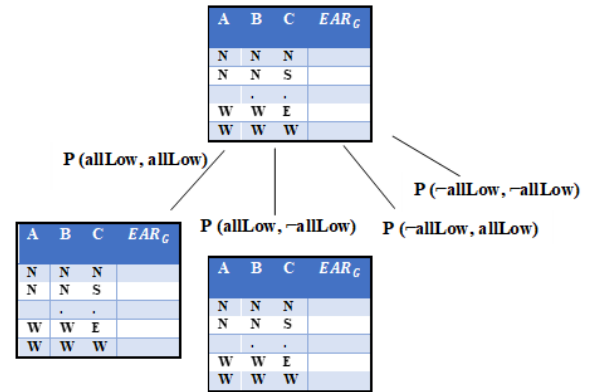


FIGURE 5. Payoff matrix.

on this distance. The shortest path has a high payoff. The environment contains 1 percent obstacles that inhibit predators' movements to those cells containing obstacles.

Agent A and agent C come into the neighborhood of agent B, shown as the shaded portion. We pinpoint the critical problem for agent B to predict and reason the movements of agent A, guessing if A will move towards B or away from B. From B's interpretation, the neighborhood of A is partitioned into the area between B and A and the area behind A. In the area behind A, the reward distribution may be invisible to agent B. The above probability is computed by recursively modeling all the possible cases in A's neighborhood. We specify the unseen area behind A that may have cells with high unilateral reward. If yes, A tends to move away from B. Otherwise, for cooperation, A moves towards B.

For a group  $g$ , each agent contains  $g-1$  unobserved areas, one for each other agent in the group. And they form possible cases of  $2^{g-1}$ . Each case makes a model at the second level of the ERMM tree. Each model is related to B's belief  $P(\text{nbp}_B^{A_1} \dots \text{nbp}_B^{A_{g-1}})$ .

In Enhanced RMM, we have considered three levels of tree hierarchy because reasonable information cannot be logically perceived from the deep levels and gives little performance [36]. Figure 5 represents the ERMM tree with  $g=3$ . The third level is dropped because it comprises Zero models. In the Enhanced RMM tree structure, each cell is filled with  $EAR_g$  and  $P(\text{mv})$  can easily be calculated for each joint plan  $\text{mv}'$  of the other agents. Let  $M$  be a model at the second level, and  $M$  is associated with the probability  $P_M(\text{nbp}_B^A, \text{nbp}_B^C)$ . Because the third level is a Zero model, in  $M$ , each joint plan  $\text{mv}$  has an equal probability of being selected by agents A and C. Thus, the joint plan  $\text{mv}^*$  with the highest  $EAR_g$  value is supposed to be selected by agents A and C.

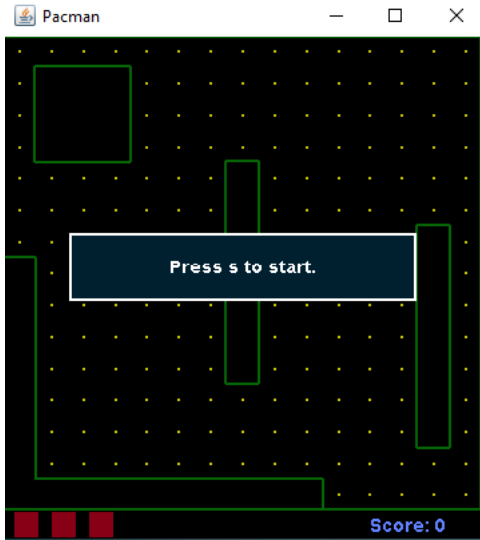


FIGURE 6. Initial screen of Pac-man.

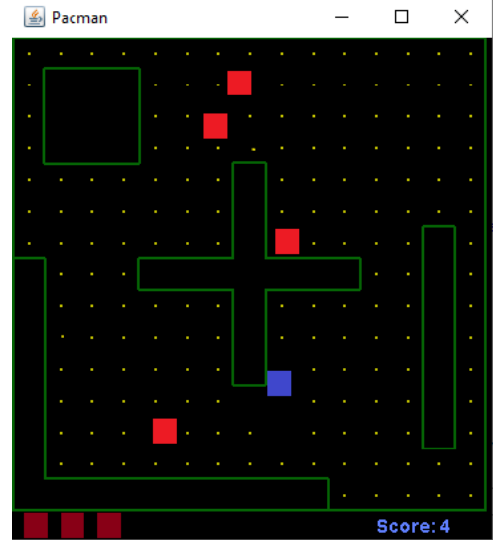


FIGURE 7. Initial screen of Pac-man.

#### IV. EXPERIMENTAL SETUP

##### A. PAC-MAN ENVIRONMENT

Inspired by real-world planning, teamwork, and search operations, we utilize the Pac-man game. We implemented a framework using the Java SE software development kit (JDK 1.6). Pac-man is sufficient to address the issues of multi-agent decision-making because of these characteristics:

- 1) The Pac-man has a constantly varying state in the game. The player is constantly moving according to the strategy at the game's current state.
- 2) State of the game is partially continuous.
- 3) Many board games are deterministic, but the Pac-man environment is stochastic because the ghosts take random moves.

The Pac-man game has four ghosts. Pac-man is script-driven to move in the maze. Ghosts are moving based on ERMM's efficient performance to capture Pac-man. Pac-man has three lives in a game; pills are distributed over the game's matrix, and the Pac-man agent eats these pills. There are variant levels of the game. At a higher level, the number of ghosts increases, increasing the game's difficulty level. Figure 6 shows a screenshot of the start state of the game.

Figure 7 displays an overview of the gameplay. Four ghosts are chasing the Pac-Man agent represented by four red blocks. Pac-Man agent represented by a white block is eating the pills or dots distributed over the maze. The lower left corner shows the number of lives remaining. On the right-hand corner is a score bar displaying the score achieved so far. There are walls in different game areas that act as obstacles for ghost agents and Pac-man.

##### B. MECHANISM OF GHOSTS

First, define the observable neighborhood of each ghost (shaded portion) and then perceive information from the defined neighborhood. The ghost reasoned with the perceived

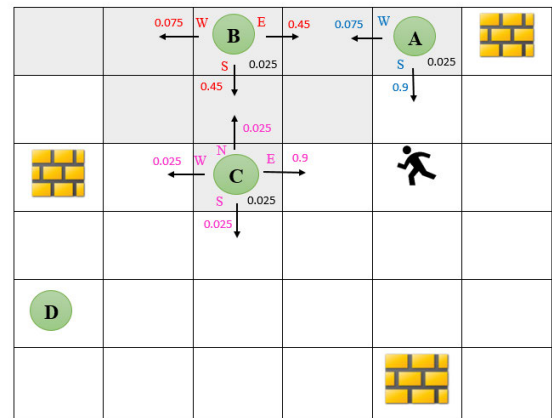


FIGURE 8. Illustration of mechanism of ghosts.

information logically to formulate the strategy to capture Pac-man. Figure 8 explains the mechanism w.r.t. ghost B.

The reward distributed is shown in the neighborhood of each ghost. The cell that takes the ghost closest to Pac-man will have a high reward. Now the ghosts A, B, and C will predict the actions of each other to make their own decisions to capture Pac-man. The computation is shown with respect to the ghost B. As ghosts A and C come within the observable neighborhood of the ghost, B will predict the possible actions of A and C in the multiagent situation. Ghost B deals with all the uncertainties related to A and C. From B's perceptions, the neighborhood of A and C can be partitioned into the observed area between A and B, C and B, and the area behind A, behind C. The reward distribution can be unobserved to agent B in the area behind A and C.

For a group of size 3, each agent has  $3-1$  unseen areas, one for each agent in the group. They set up  $2^{3-1} = 4$  possible cases, as illustrated in Figure 9. Each possible case gives

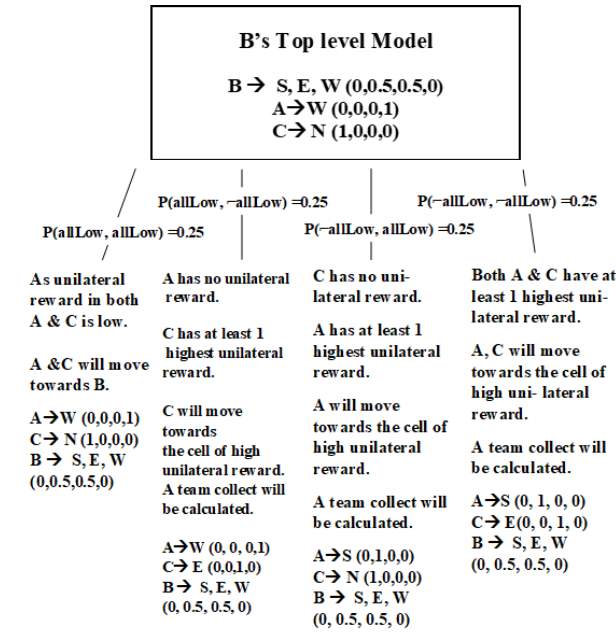


FIGURE 9. Recursive modeling explanation.

a model at the second level of the ERMM tree hierarchy. In the given scenario, B deals with the four uncertainties in the unobserved neighborhood of A and C. The modeling probabilities sum to unity. The third level of RMM forms the No-Information model, which means that the limitation of the ghost's knowledge has arrived and assumes an equal probability of states. The No-information models are omitted from the tree hierarchy for simplicity.

Suppose M is a model at the second level of the ERMM tree with probability  $P_M(\text{nbp}_B^A, \text{nbp}_B^C)$ . As the third level is Zero model and in M, each joint plan mv has an equal probability of being selected by agents A and C. Thus, the joint plan  $mv^*$  with the highest  $EAR_G$  value is supposed to be selected by agent A and C. Consider  $mv^*$  as the joint action plan for agent A and C w.r.t  $mv^*$ . Given that  $P(mv^*) = P_M(\text{nbp}_B^A, \text{nbp}_B^C)$ .

An assumption is made that each model has a unique joint action plan with the highest  $EAR_G$  value. At the second level of the ERMM tree, agents A and C get four joint action plans, each from one model. Each joint action plan is set with the probability of the model from which it is taken. The other joint action plans will be valued equal to zero probability. Figure 10 shows the tabular form of Payoff matrices in the hierarchy of the ERMM tree for ghost B.

The joint plan with highest  $EAR_G$  with  $P(\text{allLow}, \text{allLow}) = 0.25$  is

$$(A, B, C) = (W, W, N) = 0.1375$$

$$\Rightarrow 0.1375 \times 0.25 = 0.03375$$

The joint plan with highest  $EAR_G$  with  $P(\text{allLow}, \neg \text{allLow}) = 0.25$  is

$$(A, B, C) = (W, S, E) = 1.3$$

$$\Rightarrow 1.2 \times 0.25 = 0.3$$

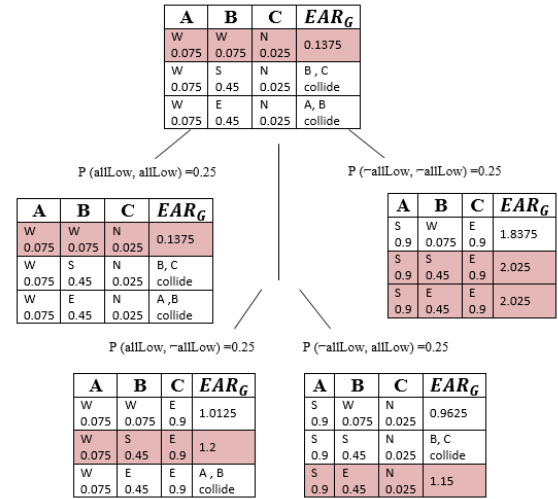


FIGURE 10. RMM tree Hierarchy in ghost B.

The joint plan with highest  $EAR_G$  with  $P(\neg \text{allLow}, \text{allLow}) = 0.25$  is

$$(A, B, C) = (S, E, N) = 1.15$$

$$\Rightarrow 1.15 \times 0.25 = 0.28$$

The joint plan with highest  $EAR_G$  with  $P(\neg \text{allLow}, \neg \text{allLow}) = 0.25$  is

$$(A, B, C) = (S, S, E) = 2.025$$

$$\Rightarrow 2.025 \times 0.25 = 0.50625$$

The highest joint plan derived is from the fourth model (S, S, E). According to which B will take the South direction and for Agent A and C South and East directions are predicted respectively. The feedback module executes these decisions. Agents A, B, and C share knowledge. Agent B evaluates errors in its predictions regarding A and C. If an error comes in predictions and A and C have taken different directions, then history is maintained. The reward distribution is updated to avoid mispredictions in future movements.

### C. RESULTS

This section provides a graphical comparison based on the performance measures. We provide results of the Enhanced Recursive Modeling Method (ERMM) compared to the Recursive Modeling Method (RMM). Integral Absolute Error (IAE) is computed at every time step when the decision is taken. It is expected that with time IAE is reduced and false predictions are deflated. It can be said that the agents continue to learn and make decisions close to optimal. Figure 11 shows the trend of IAE as the agent continues to make decisions.

Initially, IAE is very high because the agent's history is empty. As the agent interacts and shares information in multi-agent environments, it pinpoints the mispredictions. With the progress in history buildup, the agent can learn to make fewer

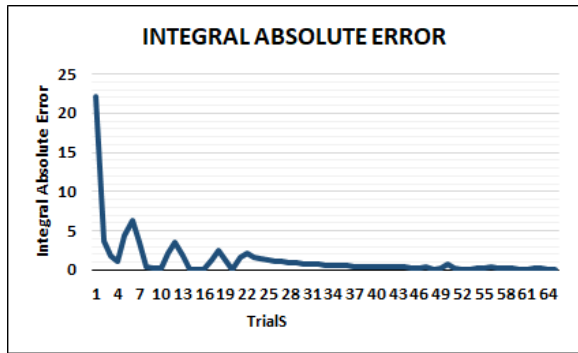


FIGURE 11. Integral absolute error.

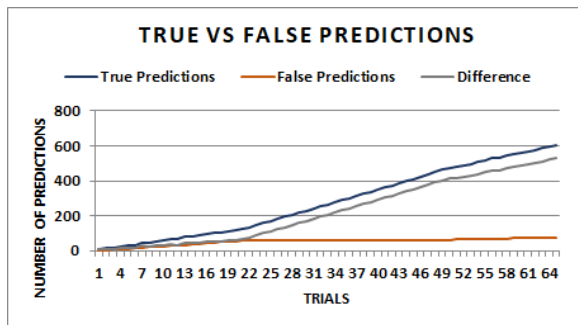


FIGURE 12. True Predictions vs. False Predictions in ERMM.

mispredictions about others. Thus the error is much reduced from 22% to 0.097%, and the agent starts to make more accurate decisions. The corresponding data table is given in Appendix.

Figure 12 illustrates the trend between true and false predictions made by the ERMM ghost agents. Agents tend to make true predictions as their knowledge is enhanced. It shows a reduction in false predictions from 33% to 10.9% as the agent's history builds up. The corresponding data table is given in Appendix.

The research compares the results of Xiang's Recursive Modeling Method (RMM) and Enhanced RMM. The results are based on total moves taken, time taken to catch the Pac-man, pills remaining on the board, and the score achieved by the Pac-man. The experiment is performed by taking a hundred instances of each. One instance comprises three given lives of the Pac-man agent.

### 1) IMPACT ON TOTAL MOVES

The impact of moves taken for ghost agents to kill the Pac-man agent is illustrated in Figure 13. It shows the number of moves or actions taken to achieve cooperation in the game. RMM takes more actions by guessing other agents whereas ERMM takes fewer movements. The average percentage decrease shows that the total moves are reduced by 55.4% using ERMM which is a significant change. The corresponding data table is given in Appendix.

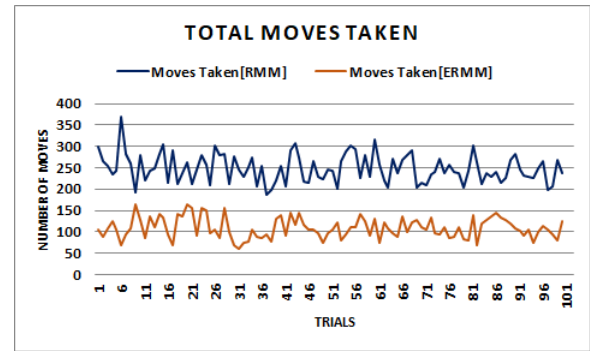


FIGURE 13. Total moves taken.

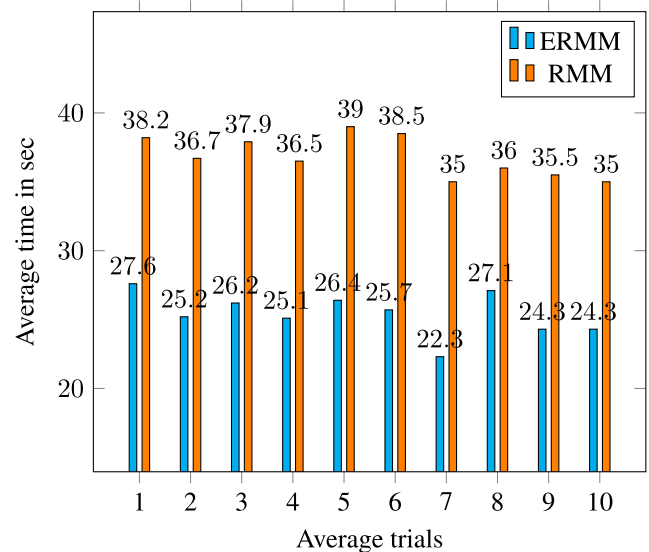


FIGURE 14. Time taken when two ghosts.

### 2) IMPACT ON TIME

The impact of the time ghost agents take to kill Pac-man is illustrated. Time is measured in seconds. Results show that When ERMM is applied on ghost agents, the time taken to kill the Pac-man is reduced significantly as compared to RMM. Figures 14-16 demonstrate the impact on time when the number of ghosts is two, four, and six in RMM and ERMM. When ERMM is applied to two ghosts to kill the Pac-man, the time taken is reduced by 30% as compared to RMM. Similarly, it is reduced to 37% and 42% compared to RMM in the case of four and six ghosts respectively. It is evident that the time taken by ghosts driven by RMM is greater than those driven by Enhanced RMM. This reduction in time indicates the capability of agents to develop coordination for decision-making. The corresponding data table is given in Appendix.

### 3) IMPACT ON PILLS

Another performance measure is the number of pills remaining in the maze which are initially distributed over the maze. Pac-man agent eats the pills during its three given lives. The goal of Pac-man is to eat the pills to make a score whereas ghost agents have a task to capture the Pac-man. If ghost

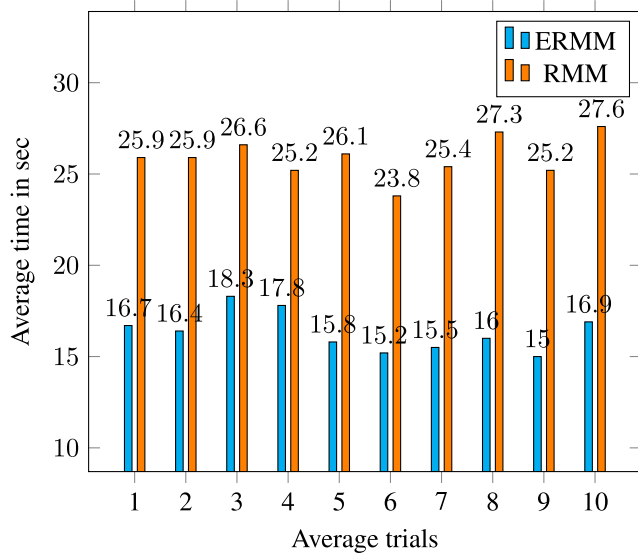


FIGURE 15. Time taken when four ghosts.

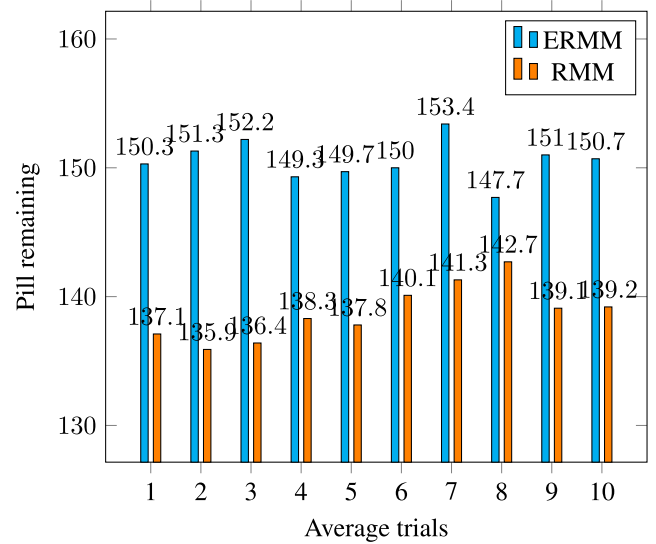


FIGURE 17. Pills Remaining when two ghosts.

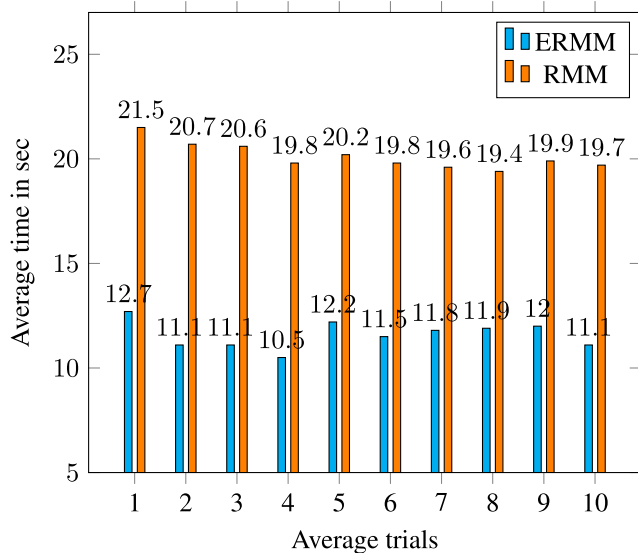


FIGURE 16. Time taken when six ghosts.

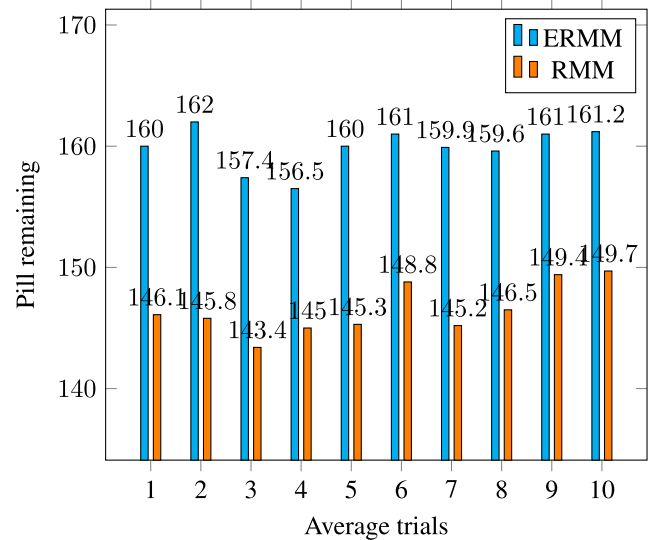


FIGURE 18. Pills Remaining when four ghosts.

agents capture Pac-man quickly then Pac-man will fail to eat the pills then more pills will remain over the maze.

Figures 17-19 demonstrate the impact on pills remaining over the maze when the number of ghosts is two, four, and six in RMM and ERMM respectively. The readings show that when ghost agents run on ERMM the number of pills over the maze is higher as compared to RMM. The high number of pills remaining in the maze shows that the ghosts' performance is good. The results show an approximately 10% average increase in pills remaining on the maze. The increase in the number of pills over the maze indicates the decrease in the score made by Pac-man. It depicts that the decision-making capability of ghost agents has improved in ERMM and has developed team cooperation and coordination to capture

Pac-man in less time. The corresponding data table is given in Appendix.

#### 4) IMPACT ON SCORE

The third performance measure is the score achieved by Pac-man. The score is incremented as the Pac-man eats the pills. If the score is low, ghosts perform effectively to capture the Pac-man. Figures 20-22 demonstrate the impact on the score when the number of ghosts is two, four, and six in RMM and ERMM respectively. The results show the 10% decrease in score when ERMM is applied on ghosts against RMM. The corresponding data table is given in Appendix.

## V. DISCUSSION

Research in the field of decision-making has made tremendous progress. This study is motivated to understand the

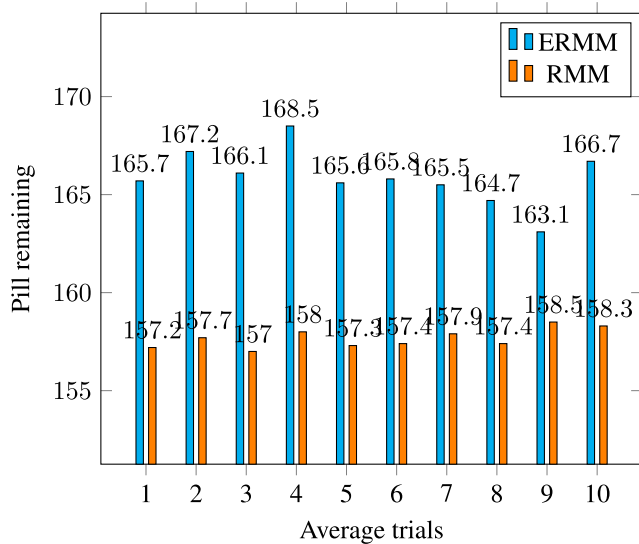


FIGURE 19. Pills Remaining when six ghosts.

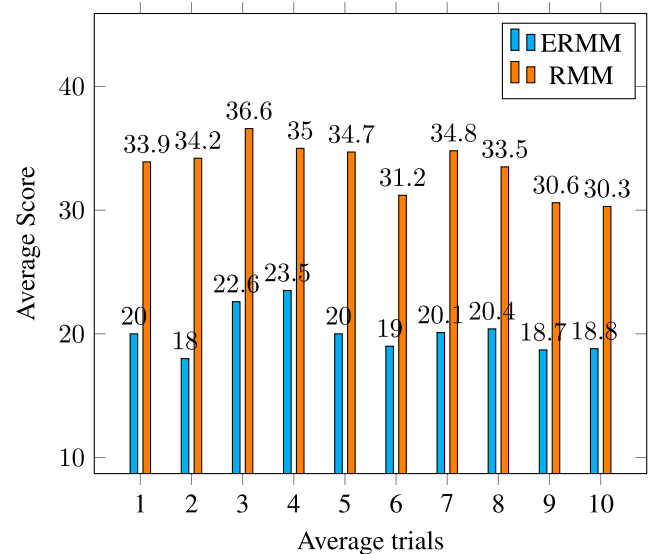


FIGURE 21. Score achieved when four ghosts.

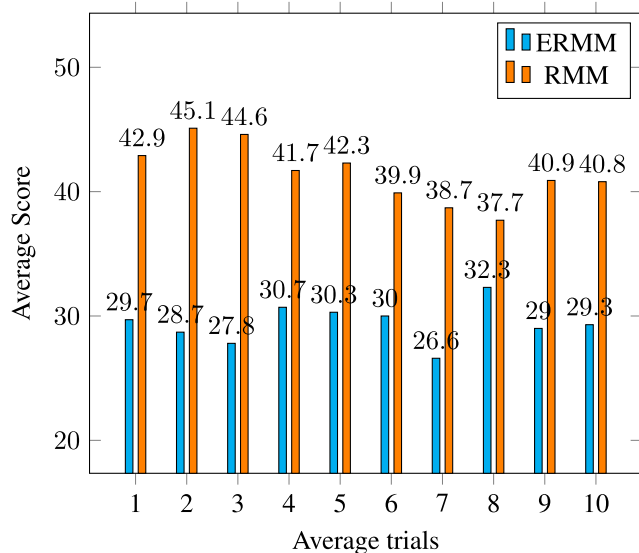


FIGURE 20. Score achieved when two ghosts.

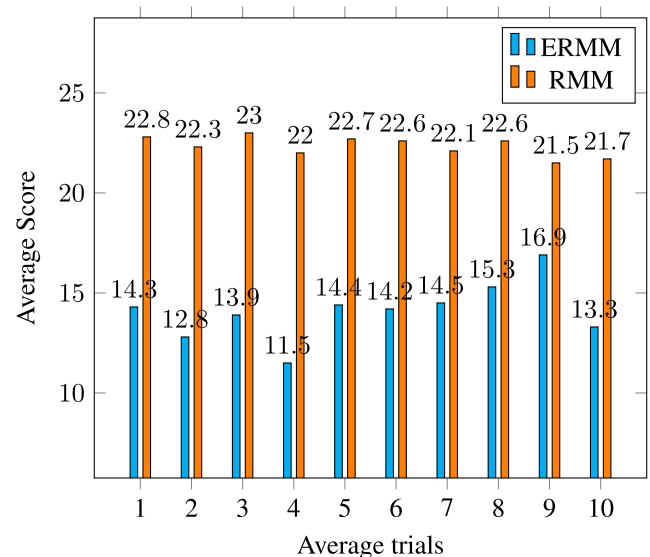


FIGURE 22. Score achieved when six ghosts.

merits of the communicative loosely coupled frameworks (LCFs) for cooperative multi-agent systems. RMM and non-communicative LCFs are limited in modeling agent interactions without enough information. It is true that communication adds cost but in communicative LCF the communication is loosely defined without strict protocols which is why the cost of setting up the multi-agent interaction is low. ERMM is the communicative LCF that allows agent communication.

To illustrate the recursive reasoning food foraging problem [37] and firefighting problems [38] are utilized. RMM is limited to deciding a single action. Whereas, in the firefighting domain agents require sequential actions of reasoning, and here RMM fails to outperform. In ERMM, agents maintain the observation history and update the recursive beliefs

using Bayesian probabilistic theory. It is utilized for sequential decision-making.

In the Pac-man environment, the RMM agent predicts other agents' actions by making guesses and becomes greedy. It only chooses an action that maximizes its own utility. The possibility of incorrect prediction leads to a faulty joint action plan. Thus it is evident from the results that RMM fails to plan and takes more than 55% moves to search the Pac-man and 30 to 42 % more time in capturing the Pac-man than ERMM. In ERMM soft evidence is received via loose agent communication. The false predictions are identified with the help of the observation history. Also, Bayesian belief updates the agent's knowledge about its new neighborhood thus reducing the false predictions to 10% and absolute error to 0.097%.

**TABLE 2.** Integral absolute error and number of predictions.

| Total Predictions | True Predictions | False Predictions | Difference | Error Percentage | Integral Absolute Error |
|-------------------|------------------|-------------------|------------|------------------|-------------------------|
| 12                | 12               | 0                 | 12         | 0                | 22.22222222             |
| 18                | 14               | 4                 | 10         | 22.22222222      | 3.703703704             |
| 27                | 20               | 7                 | 13         | 25.92592593      | 1.851851852             |
| 36                | 26               | 10                | 16         | 27.77777778      | 1.111111111             |
| 45                | 32               | 13                | 19         | 28.88888889      | 4.444444444             |
| 51                | 34               | 17                | 17         | 33.33333333      | 6.349206349             |
| 63                | 46               | 17                | 29         | 26.98412698      | 3.450655625             |
| 69                | 48               | 21                | 27         | 30.43478261      | 0.334448161             |
| 78                | 54               | 24                | 30         | 30.76923077      | 0.265251989             |
| 87                | 60               | 27                | 33         | 31.03448276      | 0.215517241             |
| 96                | 66               | 30                | 36         | 31.25            | 2.083333333             |
| 102               | 68               | 34                | 34         | 33.33333333      | 3.50877193              |
| 114               | 80               | 34                | 46         | 29.8245614       | 1.842105263             |
| 120               | 82               | 38                | 44         | 31.66666667      | 0.11627907              |
| 129               | 88               | 41                | 47         | 31.78294574      | 0.101112235             |
| 138               | 94               | 44                | 50         | 31.88405797      | 0.088731145             |
| 147               | 100              | 47                | 53         | 31.97278912      | 1.360544218             |
| 153               | 102              | 51                | 51         | 33.33333333      | 2.424242424             |
| 165               | 114              | 51                | 63         | 30.90909091      | 1.254651781             |
| 171               | 116              | 55                | 61         | 32.16374269      | 0.058479532             |
| 180               | 122              | 58                | 64         | 32.22222222      | 1.695906433             |
| 190               | 132              | 58                | 74         | 30.52631579      | 2.09494324              |
| 204               | 146              | 58                | 88         | 28.43137255      | 1.579520697             |
| 216               | 158              | 58                | 100        | 26.85185185      | 1.413255361             |
| 228               | 170              | 58                | 112        | 25.43859649      | 1.271929825             |
| 240               | 182              | 58                | 124        | 24.16666667      | 1.150793651             |
| 252               | 194              | 58                | 136        | 23.01587302      | 1.046176046             |
| 264               | 206              | 58                | 148        | 21.96969697      | 0.955204216             |
| 276               | 218              | 58                | 160        | 21.01449275      | 0.875603865             |
| 288               | 230              | 58                | 172        | 20.13888889      | 0.805555556             |
| 300               | 242              | 58                | 184        | 19.33333333      | 0.743589744             |
| 312               | 254              | 58                | 196        | 18.58974359      | 0.688509022             |
| 324               | 266              | 58                | 208        | 17.90123457      | 0.639329806             |
| 336               | 278              | 58                | 220        | 17.26190476      | 0.595238095             |
| 348               | 290              | 58                | 232        | 16.66666667      | 0.555555556             |
| 360               | 302              | 58                | 244        | 16.11111111      | 0.519713262             |
| 372               | 314              | 58                | 256        | 15.59139785      | 0.487231183             |
| 384               | 326              | 58                | 268        | 15.10416667      | 0.45770202              |
| 396               | 338              | 58                | 280        | 14.64646465      | 0.430778372             |
| 408               | 350              | 58                | 292        | 14.21568627      | 0.406162465             |
| 420               | 362              | 58                | 304        | 13.80952381      | 0.383597884             |
| 432               | 374              | 58                | 316        | 13.42592593      | 0.362862863             |
| 444               | 386              | 58                | 328        | 13.06306306      | 0.343764817             |
| 456               | 398              | 58                | 340        | 12.71929825      | 0.326135852             |
| 468               | 410              | 58                | 352        | 12.39316239      | 0.30982906              |
| 480               | 422              | 58                | 364        | 12.08333333      | 0.294715447             |
| 492               | 434              | 58                | 376        | 11.78861789      | 0.371295052             |
| 508               | 450              | 58                | 392        | 11.41732283      | 0.076930039             |
| 522               | 462              | 60                | 402        | 11.49425287      | 0.258297817             |
| 534               | 474              | 60                | 414        | 11.23595506      | 0.734579014             |
| 543               | 478              | 65                | 413        | 11.97053407      | 0.258822358             |
| 555               | 490              | 65                | 425        | 11.71171171      | 0.125258949             |
| 561               | 496              | 65                | 431        | 11.58645276      | 0.106391914             |
| 573               | 506              | 67                | 439        | 11.69284468      | 0.180817186             |
| 582               | 515              | 67                | 448        | 11.51202749      | 0.232566212             |
| 594               | 527              | 67                | 460        | 11.27946128      | 0.387205387             |
| 600               | 530              | 70                | 460        | 11.66666667      | 0.22875817              |
| 612               | 542              | 70                | 472        | 11.4379085       | 0.156294402             |
| 621               | 549              | 72                | 477        | 11.5942029       | 0.219795316             |
| 633               | 561              | 72                | 489        | 11.37440758      | 0.106801949             |
| 639               | 567              | 72                | 495        | 11.26760563      | 0.099521862             |
| 651               | 577              | 74                | 503        | 11.3671275       | 0.155006284             |
| 660               | 586              | 74                | 512        | 11.21212121      | 0.20021645              |
| 672               | 598              | 74                | 524        | 11.01190476      | 0.097450485             |
| 678               | 604              | 74                | 530        | 10.91445428      | 0.097450485             |

**TABLE 3.** Total moves taken by RMM and ERMM.

| Moves Taken[RMM] | Moves Taken[ERMM] | Difference | % Decrease  |
|------------------|-------------------|------------|-------------|
| 299              | 104               | 195        | 65.2173913  |
| 265              | 88                | 177        | 66.79245283 |
| 254              | 108               | 146        | 57.48031496 |
| 235              | 124               | 111        | 47.23404255 |
| 244              | 104               | 140        | 57.37704918 |
| 371              | 69                | 302        | 81.40161725 |
| 283              | 93                | 190        | 67.13780919 |
| 261              | 108               | 153        | 58.62068966 |
| 192              | 164               | 28         | 14.58333333 |
| 280              | 129               | 151        | 53.92857143 |
| 220              | 86                | 134        | 60.90909091 |
| 243              | 136               | 107        | 44.03292181 |
| 249              | 111               | 138        | 55.42168675 |
| 283              | 142               | 141        | 49.82332155 |
| 304              | 134               | 170        | 55.92105263 |
| 214              | 93                | 121        | 56.54205607 |
| 290              | 69                | 221        | 76.20689655 |
| 211              | 141               | 70         | 33.17535545 |
| 237              | 135               | 102        | 43.03797468 |
| 263              | 164               | 99         | 37.64258555 |
| 211              | 155               | 56         | 26.54028436 |
| 245              | 92                | 153        | 62.44897959 |
| 280              | 156               | 124        | 44.28571429 |
| 257              | 149               | 108        | 42.0233463  |
| 209              | 96                | 113        | 54.06698565 |
| 303              | 105               | 198        | 65.34653465 |
| 279              | 85                | 194        | 69.53405018 |
| 282              | 156               | 126        | 44.68085106 |
| 211              | 101               | 110        | 52.13270142 |
| 276              | 69                | 207        | 75          |
| 247              | 60                | 187        | 75.70850202 |
| 228              | 73                | 155        | 67.98245614 |
| 251              | 78                | 173        | 68.92430279 |
| 274              | 105               | 169        | 61.67883212 |
| 206              | 88                | 118        | 57.2815534  |
| 254              | 86                | 168        | 66.14173228 |
| 186              | 95                | 91         | 48.92473118 |
| 197              | 76                | 121        | 61.4213198  |
| 221              | 131               | 90         | 40.7239819  |
| 254              | 138               | 116        | 45.66929134 |
| 207              | 91                | 116        | 56.03864734 |
| 292              | 144               | 148        | 50.68493151 |
| 307              | 117               | 190        | 61.88925081 |
| 275              | 146               | 129        | 46.90909091 |
| 219              | 117               | 102        | 46.57534247 |
| 214              | 104               | 110        | 51.40186916 |
| 266              | 104               | 162        | 60.90225564 |
| 229              | 98                | 131        | 57.20524017 |

**TABLE 3.** (Continued.) Total moves taken by RMM and ERMM.

|     |     |     |             |
|-----|-----|-----|-------------|
| 224 | 73  | 151 | 67.41071429 |
| 246 | 97  | 149 | 60.56910569 |
| 243 | 106 | 137 | 56.37860082 |
| 201 | 121 | 80  | 39.80099502 |
| 266 | 79  | 187 | 70.30075188 |
| 288 | 94  | 194 | 67.36111111 |
| 302 | 110 | 192 | 63.57615894 |
| 293 | 110 | 183 | 62.45733788 |
| 225 | 143 | 82  | 36.44444444 |
| 281 | 124 | 157 | 55.87188612 |
| 229 | 90  | 139 | 60.69868996 |
| 317 | 132 | 185 | 58.35962145 |
| 256 | 75  | 181 | 70.703125   |
| 220 | 123 | 97  | 44.09090909 |
| 204 | 109 | 95  | 46.56862745 |
| 271 | 96  | 175 | 64.57564576 |
| 237 | 89  | 148 | 62.44725738 |
| 267 | 136 | 131 | 49.06367041 |
| 280 | 99  | 181 | 64.64285714 |
| 291 | 122 | 169 | 58.07560137 |
| 204 | 129 | 75  | 36.76470588 |
| 216 | 112 | 104 | 48.14814815 |
| 208 | 106 | 102 | 49.03846154 |
| 235 | 133 | 102 | 43.40425532 |
| 240 | 96  | 144 | 60          |
| 270 | 95  | 175 | 64.81481481 |
| 237 | 110 | 127 | 53.58649789 |
| 257 | 85  | 172 | 66.92607004 |
| 239 | 88  | 151 | 63.17991632 |
| 237 | 110 | 127 | 53.58649789 |
| 204 | 83  | 121 | 59.31372549 |
| 243 | 80  | 163 | 67.0781893  |
| 303 | 138 | 165 | 54.45544554 |
| 264 | 70  | 194 | 73.48484848 |
| 212 | 118 | 94  | 44.33962264 |
| 238 | 127 | 111 | 46.63865546 |
| 230 | 135 | 95  | 41.30434783 |
| 239 | 144 | 95  | 39.74895397 |
| 214 | 134 | 80  | 37.38317757 |
| 226 | 129 | 97  | 42.92035398 |
| 268 | 119 | 149 | 55.59701493 |
| 282 | 108 | 174 | 61.70212766 |
| 246 | 103 | 143 | 58.1300813  |
| 231 | 92  | 139 | 60.17316017 |
| 229 | 105 | 124 | 54.14847162 |
| 225 | 74  | 151 | 67.11111111 |
| 250 | 100 | 150 | 60          |
| 265 | 114 | 151 | 56.98113208 |
| 199 | 106 | 93  | 46.73366834 |
| 206 | 93  | 113 | 54.85436893 |
| 267 | 80  | 187 | 70.03745318 |
| 237 | 125 | 112 | 47.25738397 |
|     |     |     | 55.42848765 |

We focus on the key features of the two frameworks, RMM and ERMM, and provide our empirical comparison between the two decision-theoretic approaches. RMM is considered a general framework for decision-making and filled a void in multi-agent systems for deep reasoning in simple settings. However, ERMM is the collaborative decision-making framework in which the agent tends to have more updated prior beliefs.

## VI. FUTURE WORK AND LIMITATIONS

The recursive frameworks allow high-order reasoning which is utilized in robotics [39], meta games [40], military [38], and transportation problems [41] for modeling the unknown behavior of other agents. We generate results in a constrained environment of Pac-man such as the following:

- 1) The system contains a finite number of agents and environment states, and each agent can take a defined set of actions.
- 2) The system must operate on a discrete timestamp. All the actions taken by the agents at the same timestamp are considered to be taken in parallel.
- 3) The agent's neighborhood is defined, and a cell visited by an agent within the neighborhood is marked "visited".
- 4) The test environment is stochastic as the agents' actions are uncertain and partially observable, which means the agent can only perceive its neighborhood.
- 5) The ERMM considers three levels of the tree hierarchy to avoid a deep structure with less useful information [36].
- 6) The ERMM is evaluated on a game environment like Pac-man. Nowadays, a multi-agent tiger problem has become a standard for multi-agent research to evaluate agent decision modeling [42]. Memetic Multi-Agent System (MeMAS) has emerged as another multi-agent framework to implement human-like characters in complex games [43]. ERMM can be utilized in real-world problems like the firefighting domain and the air defense domain.

The recursive frameworks provide significant results in modeling agents' behavior in multi-agents. These frameworks have interesting contributions in sports in predicting injury [44]. However, outthinking other agents' behavior results in a deeper tree structure, thus increasing the computational complexity. Given the benefit of the recursive reasoning models, we face the challenge of computational complexity as compared to other reasoning models. Studies in recursive modeling know about this additional computational stress [38]. This gives the future direction to encourage the utilization of recursive models easily with less computational complexity.

Meta-reasoning is applied to reduce the computational complexity in machines based on recursive models [45]. Pruning of deep branches of the agent planning tree is suggested but it still needs to be explored for meta-reasoning.

Other improvements like behavior assumptions, constraining recursive models, and investigating optimal conditions would be advantageous in recursive modeling.

## VII. CONCLUSION

The main objective of the research is to address the limitations of loosely coupled agents. We were motivated to improve decision-making performance in cognitive agents using ERMM in Lawniczak's Architecture. We have given the efficiency comparison with Xiang's Architecture of recursive modeling. Xiang's agent model [10] is limited by making guesses about the states of other agents. The inaccuracy in predicting other agents can degrade the system's performance. The misjudgment of the other agents' actions takes the system to the misjudgment of the optimal joint plan. Considering all the possibilities about the other agents' actions takes computational cost and a deeper RMM hierarchy [38]. It leads to overhead in algorithm performance.

In ERMM, agents communicate their perceptions about each other. With communication, agents get sufficient information about other agents' states and decisions, thus resulting in improved coordination. Our cognitive agent model builds up the history based on knowledge sharing. This prominent feature allows agents to coordinate their actions without any wrong interpretation. In ERMM, the agent's knowledge is revised and aligned toward the correct observations. By updating the agent's belief in ERMM, only true possibilities are considered and mispredictions are filtered off from the recursive modeling. Hence, it takes less computational cost and avoids a deep ERMM hierarchy. Thus agents become capable of collaborating on a common joint plan in less time.

In our experimental comparison, we considered the RMM as a non-communicative LCF and ERMM as the representative of a communicative LCF. Communication in ERMM is loosely defined without strict protocols to avoid the cost. The feedback is provided about the prior predictions to date the agent's belief. The results show that the wrong predictions are highlighted, and with progress, error in perception gets reduced to 0.097% in ERMM thus giving superior performance.

TABLE 4. Time taken by ERMM and RMM when two ghosts.

| Average Time[ERMM] | Average Time[RMM] | Difference | % Decrease |
|--------------------|-------------------|------------|------------|
| 27.6               | 38.2              | 10.6       | 27.7       |
| 25.2               | 36.7              | 11.5       | 31.3       |
| 26.2               | 37.9              | 11.7       | 30.9       |
| 25.1               | 36.5              | 11.4       | 31.2       |
| 26.4               | 39                | 12.6       | 32.3       |
| 25.7               | 38.5              | 12.8       | 33.2       |
| 22.3               | 35                | 12.7       | 36.3       |
| 27.1               | 36                | 8.9        | 24.7       |
| 24.3               | 35.5              | 11.2       | 31.5       |
| 24.3               | 35                | 10.7       | 30.6       |
|                    |                   |            | 30.9       |

**TABLE 5.** Score Board of ERMM and RMM when two ghosts.

| Average Score[ERMM] | Average Score[RMM] | Difference | Percentage |
|---------------------|--------------------|------------|------------|
| 29.7                | 42.9               | 13.2       | 30.8       |
| 28.7                | 44.1               | 15.4       | 34.9       |
| 27.8                | 44.6               | 16.8       | 37.7       |
| 30.7                | 41.7               | 11         | 26.8       |
| 30.3                | 42.3               | 12         | 28.3       |
| 30                  | 39.9               | 9.9        | 24.8       |
| 26.6                | 38.7               | 12.1       | 31.2       |
| 32.3                | 37.3               | 5          | 13.4       |
| 29                  | 40.9               | 11.9       | 29.1       |
| 29.3                | 40.8               | 11.5       | 28.9       |
|                     |                    |            | 28.5       |

**TABLE 6.** Pills remaining by ERMM and RMM when two ghosts.

| Average Pills[ERMM] | Average Pills[RMM] | Difference | % Increase |
|---------------------|--------------------|------------|------------|
| 150.3               | 137.1              | 13.2       | 9.6        |
| 151.3               | 135.9              | 15.4       | 11.3       |
| 152.2               | 135.4              | 16.8       | 12.4       |
| 149.3               | 138.3              | 11         | 7.95       |
| 149.7               | 137.7              | 12         | 8.7        |
| 150                 | 140.1              | 9.9        | 7.1        |
| 153.4               | 141.3              | 12.1       | 8.6        |
| 147.7               | 142.7              | 5          | 3.5        |
| 151                 | 139.1              | 11.9       | 8.6        |
| 150.7               | 139.2              | 11.5       | 8.3        |
|                     |                    |            | 8.6        |

**TABLE 7.** Time taken by ERMM and RMM when four ghosts.

| Average Time [ERMM] | Average Time[RMM] | Difference | % Decrease |
|---------------------|-------------------|------------|------------|
| 16.7                | 25.9              | 9.2        | 35.5       |
| 16.4                | 25.9              | 9.5        | 36.7       |
| 18.3                | 26.6              | 8.3        | 31.2       |
| 17.8                | 25.2              | 7.4        | 29.4       |
| 15.8                | 26.1              | 10.3       | 39.5       |
| 15.2                | 23.8              | 8.6        | 36.1       |
| 15.5                | 25.4              | 9.9        | 38.9       |
| 16                  | 27.3              | 11.3       | 41.4       |
| 15                  | 25.2              | 10.2       | 40.5       |
| 16.9                | 27.6              | 10.7       | 38.8       |
|                     |                   |            | 36.8       |

**TABLE 8.** Score Board of ERMM and RMM when four ghosts.

| Average Score[ERMM] | Average Score[RMM] | Difference | % Score |
|---------------------|--------------------|------------|---------|
| 20                  | 33.9               | 13.9       | 41.0    |
| 18                  | 34.2               | 16.2       | 47.4    |
| 22.6                | 36.6               | 14         | 38.3    |
| 23.5                | 35                 | 11.5       | 32.9    |
| 20                  | 34.7               | 14.7       | 42.4    |
| 19                  | 31.2               | 12.2       | 39.1    |
| 20.1                | 34.8               | 14.7       | 42.2    |
| 20.4                | 33.5               | 13.1       | 39.1    |
| 18.7                | 30.6               | 11.9       | 38.9    |
| 18.8                | 30.3               | 11.5       | 37.9    |
|                     |                    |            | 39.9    |

This study provided the concept of coordinated decision-making in Lawniczak's Architecture for cognitive agents.

**TABLE 9.** Pills remaining by ERMM and RMM when four ghosts.

| Average Pills[ERMM] | Average Pills[RMM] | Difference | % Increase |
|---------------------|--------------------|------------|------------|
| 160                 | 146.1              | 13.9       | 9.5        |
| 162                 | 145.8              | 16.2       | 11.1       |
| 157.4               | 143.4              | 14         | 9.8        |
| 156.5               | 145                | 11.5       | 7.9        |
| 160                 | 145.3              | 14.7       | 10.1       |
| 161                 | 148.8              | 12.2       | 8.2        |
| 159.9               | 145.2              | 14.7       | 10.1       |
| 159.6               | 146.5              | 13.1       | 8.9        |
| 161.3               | 149.4              | 11.9       | 7.9        |
| 161.2               | 149.7              | 11.5       | 7.78       |
|                     |                    |            | 9.1        |

**TABLE 10.** Time taken by ERMM and RMM when six ghosts.

| Average Time [ERMM] | Average Time [RMM] | Difference | % Decrease |
|---------------------|--------------------|------------|------------|
| 12.7                | 21.5               | 8.8        | 40.9       |
| 11.1                | 20.7               | 9.6        | 46.4       |
| 11.1                | 20.6               | 9.5        | 46.1       |
| 10.5                | 19.8               | 9.3        | 46.9       |
| 12.2                | 20.2               | 8          | 39.6       |
| 11.5                | 19.8               | 8.3        | 41.9       |
| 11.8                | 19.6               | 7.8        | 39.8       |
| 11.9                | 19.4               | 7.5        | 38.7       |
| 12                  | 19.9               | 7.9        | 39.7       |
| 11.1                | 19.7               | 8.6        | 43.7       |
|                     |                    |            | 42.4       |

**TABLE 11.** Score Board of ERMM and RMM when six ghosts.

| Average Score[ERMM] | Average Score[RMM] | Difference | % Score |
|---------------------|--------------------|------------|---------|
| 14.3                | 22.8               | 8.5        | 37.2    |
| 12.8                | 22.3               | 9.5        | 42.6    |
| 13.9                | 23                 | 9.1        | 39.5    |
| 11.5                | 22                 | 10.5       | 47.7    |
| 14.4                | 22.7               | 8.3        | 36.5    |
| 14.2                | 22.6               | 8.4        | 37.1    |
| 14.5                | 22.1               | 7.6        | 34.3    |
| 15.3                | 22.6               | 7.3        | 32.3    |
| 16.9                | 21.5               | 4.6        | 21.4    |
| 13.3                | 21.7               | 8.4        | 38.7    |
|                     |                    |            | 36.8    |

**TABLE 12.** Pills remaining by ERMM and RMM when six ghosts.

| Average Pills[ERMM] | Average Pills[RMM] | Difference | % Increase |
|---------------------|--------------------|------------|------------|
| 165.7               | 157.2              | 8.5        | 5.4        |
| 167.2               | 157.7              | 9.5        | 6.0        |
| 166.1               | 157                | 9.1        | 5.8        |
| 168.5               | 158                | 10.5       | 6.6        |
| 165.6               | 157.3              | 8.3        | 5.3        |
| 165.8               | 157.4              | 8.4        | 5.3        |
| 165.5               | 157.9              | 7.6        | 4.8        |
| 164.7               | 157.4              | 7.3        | 4.6        |
| 163.1               | 158.5              | 4.6        | 2.9        |
| 166.7               | 158.3              | 8.4        | 5.3        |
|                     |                    |            | 5.2        |

It addressed the deficiencies of the non-communicative loosely coupled framework. The proposed communicative

LCF i.e. ERMM gathered sufficient information about other agents for coordinated decision-making. The ERMM can be utilized in real-world scenarios where planning, search, and well-coordinated teamwork are required such as transportation problems, and air defense systems. The ERMM can prove its applicability in the domains of sports for injury prediction as injuries are recursive in nature. We believe that the Enhanced RMM with feedback can be utilized for the prevention of the injury by identifying the incorrect training techniques. We hope that this work will contribute towards the enhancement of deep reasoning and decision-making in diverse domains.

## APPENDIX

### DATA TABLES OF RESULTS

See Tables 2–12.

## REFERENCES

- [1] N. Sharma, R. Sharma, and N. Jindal, "Machine learning and deep learning applications—A vision," *Global Transitions Proc.*, vol. 2, no. 1, pp. 24–28, 2021.
- [2] O. B. Ayoade, T. O. Oladele, A. L. Imoize, J. B. Awotunde, A. J. Adeyoye, S. O. Olorunyomi, and A. O. Idowu, *Explainable Artificial Intelligence (XAI) in Medical Decision Systems (MDSS): Healthcare Systems Perspective*. Edison, NJ, USA: IET, 2022.
- [3] A. Khan, J. P. Li, N. Ahmad, S. Sethi, A. U. Haq, S. H. Patel, and S. Rahim, "Predicting emerging trends on social media by modeling it as temporal bipartite networks," *IEEE Access*, vol. 8, pp. 39635–39646, 2020.
- [4] K. Abbas, M. K. Hasan, A. Abbasi, U. A. Mokhtar, A. Khan, S. N. H. S. Abdullah, S. Dong, S. Islam, D. Alboaneen, and F. R. A. Ahmed, "Predicting the future popularity of academic publications using deep learning by considering it as temporal citation networks," *IEEE Access*, vol. 11, pp. 83052–83068, 2023.
- [5] R. Calegari, G. Ciatto, V. Mascardi, and A. Omicini, "Logic-based technologies for multi-agent systems: A systematic literature review," *Auto. Agents Multi-Agent Syst.*, vol. 35, no. 1, pp. 1–67, Apr. 2021.
- [6] Z. Akata et al., "A research agenda for hybrid intelligence: Augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence," *Computer*, vol. 53, no. 8, pp. 18–28, Aug. 2020.
- [7] N. Vlassis, "Multiagent systems and distributed AI," in *Intelligent Autonomous Systems*. Amsterdam, The Netherlands: University of Amsterdam, 2003.
- [8] S. Guillaume and B. Charnomordic, "Fuzzy inference systems: An integrated modeling environment for collaboration between expert knowledge and data using FisPro," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 8744–8755, Aug. 2012.
- [9] P. Langley, J. E. Laird, and S. Rogers, "Cognitive architectures: Research issues and challenges," *Cogn. Syst. Res.*, vol. 10, no. 2, pp. 141–160, Jun. 2009.
- [10] Y. Xiang and F. Hanshar, "Comparison of tightly and loosely coupled decision paradigms in multiagent expedition," *Int. J. Approx. Reasoning*, vol. 51, no. 5, pp. 600–613, Jun. 2010.
- [11] M. Wooldridge and N. R. Jennings, "Intelligent agents: Theory and practice," *Knowl. Eng. Rev.*, vol. 10, no. 2, pp. 115–152, Jun. 1995.
- [12] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, vol. 26. Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.
- [13] T.-T. Dang, N.-A.-T. Nguyen, V.-T.-T. Nguyen, and L.-T.-H. Dang, "A two-stage multi-criteria supplier selection model for sustainable automotive supply chain under uncertainty," *Axioms*, vol. 11, no. 5, p. 228, May 2022.
- [14] C.-N. Wang, F.-C. Yang, N. T. M. Vo, and V. T. T. Nguyen, "Enhancing lithium-ion battery manufacturing efficiency: A comparative analysis using DEA Malmquist and epsilon-based measures," *Batteries*, vol. 9, no. 6, p. 317, Jun. 2023.
- [15] M. Dorigo, V. Trianni, E. Şahin, R. Groß, T. H. Labella, G. Baldassarre, S. Nolfi, J.-L. Deneubourg, F. Mondada, D. Floreano, and L. M. Gambardella, "Evolving self-organizing behaviors for a swarm-bot," *Auto. Robots*, vol. 17, nos. 2–3, pp. 223–245, Sep. 2004.
- [16] C.-H. Yu and R. Nagpal, "Self-adapting modular robotics: A generalized distributed consensus framework," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2009, pp. 1881–1888.
- [17] J. Foerster, G. Farquhar, T. Afouras, N. Nardelli, and S. Whiteson, "Counterfactual multi-agent policy gradients," in *Proc. 32nd AAAI Conf. Artif. Intell.*, Apr. 2018, vol. 32, no. 1, pp. 2974–2982.
- [18] C.-H. Yu, J. K. Werfel, and R. Nagpal, "Collective decision-making in multi-agent systems by implicit leadership," in *Autonomous Agents and Multiagent Systems*. New York, NY, USA: Association for Computing Machinery Press, 2010.
- [19] A. Vilenica, A. Pokahr, L. Braubach, W. Lamersdorf, J. Sudeikat, and W. Renz, "Coordination multi-agent systems: A declarative approach using coordination spaces," in *Proc. 20th Eur. Meeting Cybern. Syst. Res. (EMCSR)-Int. Workshop From Agent Theory Agent Implement. (AT2AI-7)*, R. Trappl, Ed. Austrian Society for Cybernetic Studies, 2010, pp. 441–446.
- [20] S. Sukhbaatar and R. Fergus, "Learning multiagent communication with backpropagation," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 2244–2252.
- [21] J. Xing, W. Liu, and M. Kong, "Multi-agent formation control and target tracking," in *Proc. Int. Conf. Control, Autom. Inf. Sci. (ICCAIS)*, Oct. 2018, pp. 244–248.
- [22] H. Zhang, D. Li, and Y. He, "Multi-robot cooperation strategy in game environment using deep reinforcement learning," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2018, pp. 886–891.
- [23] P. J. Gmytrasiewicz and E. H. Durfee, "Rational coordination in multi-agent environments," *Auto. Agents Multi-Agent Syst.*, vol. 3, no. 4, pp. 319–350, 2000.
- [24] W. Jacobs and M. Kiefer, "Robot decisions based on maximizing utility," in *Proc. IJCAI*, 1973, pp. 402–411.
- [25] J. Doyle, "Rationality and its roles in reasoning," *Comput. Intell.*, vol. 8, no. 2, pp. 376–409, May 1992.
- [26] E. Liscio, M. van der Meer, L. C. Siebert, C. M. Jonker, and P. K. Murukannaiah, "What values should an agent align with?" *Auto. Agents Multi-Agent Syst.*, vol. 36, no. 1, pp. 1–32, Apr. 2022.
- [27] L. C. Siebert, E. Liscio, P. K. Murukannaiah, L. Kaptein, S. Spruit, J. van den Hoven, and C. Jonker, "Estimating value preferences in a hybrid participatory system," in *Augmenting Human Intellect*. Amsterdam, The Netherlands: IOS Press, 2021.
- [28] Y. Xiang and F. Hanshar, "Partial evaluation for planning in multiagent expedition," in *Proc. Can. Conf. Artif. Intell.* Cham, Switzerland: Springer, 2011, pp. 420–432.
- [29] Y. Xiang and F. Hanshar, "Multiagent decision by partial evaluation," in *Proc. Can. Conf. Artif. Intell.* Cham, Switzerland: Springer, 2012, pp. 242–254.
- [30] R. Aumann and A. Brandenburger, "Epistemic conditions for Nash equilibrium," *Econometrica, J. Econ. Soc.*, vol. 63, no. 5, pp. 1161–1180, 1995.
- [31] J. Shen and V. Lesser, "Communication management using abstraction in distributed Bayesian networks," in *Proc. 5th Int. Joint Conf. Auto. Agents Multiagent Syst.*, May 2006, pp. 622–629.
- [32] X. An, Y. Xiang, and N. Cercone, "Dynamic multiagent probabilistic inference," *Int. J. Approx. Reasoning*, vol. 48, no. 1, pp. 185–213, 2008.
- [33] Y. Xiang, J. Chen, and W. S. Havens, "Optimal design in collaborative design network," in *Proc. 4th Int. Joint Conf. Auto. Agents Multiagent Syst.*, Jul. 2005, pp. 241–248.
- [34] Y. Xiang, *Reasoning in Multiagent Systems: A Graphical Models Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2002.
- [35] J. O. Gutierrez-Garcia and E. López-Neri, "Cognitive computing: A brief survey and open research challenges," in *Proc. 3rd Int. Conf. Appl. Comput. Inf. Technol., 2nd Int. Conf. Comput. Sci. Intell.*, 2015, pp. 328–333.
- [36] P. J. Gmytrasiewicz, S. Noh, and T. Kellogg, "Bayesian update of recursive agent models," *User Model. User-Adapted Interact.*, vol. 8, pp. 49–69, Mar. 1998.
- [37] M. Chandrasekaran, Y. Chen, and P. Doshi, "On Markov games played by Bayesian and boundedly-rational players," in *Proc. AAAI Conf. Artif. Intell.*, vol. 31, no. 1, 2017, pp. 437–443.
- [38] P. Doshi, P. Gmytrasiewicz, and E. Durfee, "Recursively modeling other agents for decision making: A research perspective," *Artif. Intell.*, vol. 279, Feb. 2020, Art. no. 103202.
- [39] F. Ahmed and K.-Y. Kim, "Recursive approach to combine expert knowledge and data-driven RSW weldability certification decision making process," *Robot. Comput. Integr. Manuf.*, vol. 79, Feb. 2023, Art. no. 102428.

- [40] L. Junren, Z. Wanpeng, Y. Weilin, H. Zhenzhen, C. Shaofei, and C. Jing, "Research on opponent modeling framework for multi-agent game confrontation," *J. Syst. Simul.*, vol. 34, no. 9, p. 1941, 2022.
- [41] M. Zimmermann and E. Frejinger, "A tutorial on recursive models for analyzing and predicting path choice behavior," *EURO J. Transp. Logistics*, vol. 9, no. 2, Jun. 2020, Art. no. 100004.
- [42] Y. Pan, H. Zhang, Y. Zeng, B. Ma, J. Tang, and Z. Ming, "Diversifying agent's behaviors in interactive decision models," *Int. J. Intell. Syst.*, vol. 37, no. 12, pp. 12035–12056, Dec. 2022.
- [43] Y. Hou, M. Sun, W. Zhu, Y. Zeng, H. Piao, X. Chen, and Q. Zhang, "Behavior reasoning for opponent agents in multi-agent learning systems," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 6, no. 5, pp. 1125–1136, Oct. 2022.
- [44] N. F. N. Bittencourt, W. H. Meeuwisse, L. D. Mendonça, A. Nettel-Aguirre, J. M. Ocarino, and S. T. Fonseca, "Complex systems approach for sports injuries: Moving from risk factor identification to injury pattern recognition—Narrative review and new concept," *Brit. J. Sports Med.*, vol. 50, no. 21, pp. 1309–1314, Nov. 2016.
- [45] T. L. Griffiths, F. Callaway, M. B. Chang, E. Grant, P. M. Krueger, and F. Lieder, "Doing more with less: Meta-reasoning and meta-learning in humans and machines," *Current Opinion Behav. Sci.*, vol. 29, pp. 24–30, Oct. 2019.



**NOUSHIN MAZHAR** received the B.S. degree in software engineering from APCOMS, Affiliated with UET Taxila, Pakistan, in 2012, and the M.S. degree in computer science from NUCES-FAST Islamabad, Pakistan, in 2015. Previously, she has worked in the software industry primarily on projects related to web development. In 2015, she was also a Lecturer with the International Undergraduate Program, Roots IVY International University. She joined Foundation University Islamabad as a Lecturer, in 2016. Her research interests include artificial intelligence particularly multi-agent systems, machine learning, and robotic autonomy.



**MARYAM KAUSAR** received the B.S. degree in computer science from NUCES-FAST, Pakistan, in 2010, the M.S. degree in computer science and IT management from The University of Manchester, U.K., in 2011, and the Ph.D. degree from the School of CSE, University of Salford, U.K., in 2018. Since 2019, she has been an Assistant Professor with Foundation University Islamabad, Pakistan. Her research interests include software engineering, requirement engineering, global software development, and agile software development. Her awards and honors include securing funding for developing an augmented reality game using agile methods and being the Chair of the Pakistan Agile Development Society.

• • •