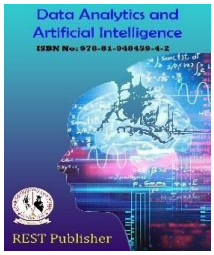
Data Analytics and Artificial Intelligence

Vol: 4(1), 2024

REST Publisher; ISBN: 978-81-948459-4-2

Website: http://restpublisher.com/book-series/daai/

DOI: https://doi.org/10.46632/daai/4/1/7



Predicting Logical Reasoning Performance in Game-Based Puzzles Using Multi-Layer Perceptron Neural Networks

Venkata Kalyan Polamarasetty

Independent Researcher, USA

Corresponding Author Email: ven.office18@gmail.com

Abstract: *Strategic reasoning in game theory intersects with logical structures in a variety of domains, including computation, argumentation, social networks, and multi-agent systems. This study examines the relationship between puzzle complexity, temporal constraints, player expertise, and logical reasoning performance using a large dataset of 200 player records. The research uses multi-layer perceptron (MLP) neural networks to model and predict logical performance outcomes based on three primary parameters: puzzle difficulty (scale 1-10), time limit in seconds, and player experience in years. Correlation analysis reveals significant relationships between these variables and logic scores, with puzzle difficulty demonstrating a strong negative correlation (r = -0.60) and player experience showing a strong positive correlation (r = 0.64) with performance outcomes. The MLP model achieves strong predictive accuracy with R^2 values of 0.744 for the training data and 0.800 for the test data, indicating excellent generalization ability without overfitting. Performance metrics including low MSE, RMSE, and MAE values confirm the model's reliability in predicting logical reasoning outcomes. The findings demonstrate that while increased puzzle complexity presents significant cognitive challenges, accumulated player experience and sufficient time allocation significantly improve logical performance. This study combines theoretical foundations in game logic – including dynamic epistemic logic, coalitional game theories, and language game structures – with empirical machine learning methods. This integration provides a quantitative framework for understanding strategic reasoning in game-theoretic contexts and provides practical insights for designing adaptive gaming systems, educational puzzle platforms, and cognitive assessment tools that optimize challenge levels based on individual player characteristics.*

Key Words: *Game theory logic, Multilayer perceptron, Puzzle difficulty optimization, Strategic reasoning performance, Cognitive game analysis.*

1. INTRODUCTION

Strategic reasoning permeates contemporary life, where individuals navigate interconnected "games" across diverse social contexts each with distinct players, rules, objectives, and preferences. These interactions shape personal success (wins) or setbacks (losses), influencing future engagements and the evolution of adaptive strategies. [1] Graph games, played on directed/undirected graphs with annotated vertices, have wide-ranging applications in computation, argumentation, social networks, and warfare. To illustrate, consider two examples (for foundational logic games, see [2]). Lorenz's framework, though seemingly straightforward, relies on auxiliary rules—e.g., atomic formulas are immune to attacks and require prior confirmation by Player 0. This ensures strategies validate logical consistency, not situational truth. [3] Building on van Bentham's work, we explore path-making games and dynamic arenas, particularly edge-elimination in finite sabotage games and their logics. [4] Wittgenstein's language games, while debated, share semantic parallels with logical games, bridging philosophy, language, and game theory (cf. Hidaka et al.). [5] We analyse two coalition game logics that model analogous scenarios: agents (player's/system components) acting on shared states, transitioning between them—either as abstract multiplayer games or open systems with competing devices. [6] Dynamic Epistemic Logic (DEL) action models capture agent perceptions, recently extending to infinite multiplayer games with solvable cases. However, DEL's classic update dynamics remain limited to turn-

Copyright@ REST Publisher

based interactions. [7] To achieve its objectives, game theory must develop a structured framework of models and formal languages capable of capturing both distinctions and commonalities within operational systems. In this context, Bauman's game theory serves as a classificatory science. Given its objectives and methodology, game theory shares significant conceptual ground with logic, where the latter provides essential formal foundations through its systems of languages and models. [8] Contemporary research is increasingly focused on developing formal methodological models to analyse social practices and interactions among rational agents. These frameworks seek to identify logical structures underlying everyday social dynamics. When conceptualized as games, rational behaviour involves examining how agents' enduring strategic influences shape eventual outcomes. [9] Logical and linguistic dependencies emerge from multiple sources. Truth-values may remain indeterminate due to incomplete information about objects or inherent partiality in the objects themselves. Natural language often reflects this partiality through statements with incomplete preferentiality, as linguistic semantics frequently lack full definitional precision. [10] The following discussion presents Wittgenstein's language game approach - a logical methodology that simultaneously preserves and transcends aspects of Russell an and Tractarianism logical analysis. This method expands logical inquiry beyond calculus-based limitations while addressing certain shortcomings in early Wittgenstein an and Russell and approaches. [11] Preference structures constitute fundamental components of human behaviour, manifesting in diverse phenomena: from individual rational choice in action theory to demonic logic's optimal world obligations, and collective decision-making in social choice theory. More abstractly, preference orderings organize possible worlds along plausibility gradients in conditional logic and non-monotonic reasoning systems.[12] Significant logical research has developed this direction, including alternating temporal logic (Allure, Hen zinger, and Superman 1998), which provides selective metrics for analysing turn-based game outcomes between players and their environment.[13] The complexity classification problem for action types concerns the existence of finite DEL action sequences that achieve epistemic logic objectives. While this resembles solving games of skill with epistemic goals, it remains unexplored in strategic, adversarial contexts. [14] We introduce PRISM-games, a model checking tool for stochastic multi-player games that enables modelling, automated verification, and strategy synthesis for probabilistic competitive/cooperative systems. Using an enhanced Reactive Modules language for modelling and an extended ATL (Raptly) for property specification, the framework incorporates probabilistic operators, reward metrics, quantitative properties, and precision boundaries. [15]

2. MATERIALS AND METHODS

Materials:

Puzzle Difficulty: This parameter measures the complexity of a puzzle, usually expressed on a numerical scale. Lower values indicate simpler, more straightforward challenges, while higher values indicate more complex, multi-step problems that require advanced reasoning. Puzzle difficulty affects cognitive load, decision-making strategies, and overall player engagement. It also interacts with other factors, such as experience and time constraints, to shape a player’s problem-solving approach and the ultimate performance outcomes in the game.

Time Limit Sec: This parameter specifies the maximum time allowed in seconds to complete a puzzle or task. Shorter limits create time pressure, increasing the need for quick thinking and rapid decision-making, while longer limits allow for careful analysis and strategic reasoning. The time limit directly affects performance, stress levels, and the balance between speed and accuracy, shaping both game strategy and measurable outcomes in a gaming or testing environment.

Player Experience Yrs: This parameter indicates the number of years a player has spent actively engaging in similar puzzles or games. More experience is often associated with better-developed strategies, pattern recognition, and problem-solving skills. It reflects accumulated knowledge, the ability to face different challenge types, and the ability to adapt under varying difficulty and time constraints. Player experience can significantly affect logical performance, which can help mitigate the effects of puzzle complexity or tight time limits during gameplay.

Logic Score: This is a numerical representation of a player's logical reasoning ability, measured by accuracy, consistency, and efficiency in solving puzzles. Higher scores indicate stronger analytical skills, structured thinking, and effective decision-making under given constraints. The Logic Score combines multiple performance aspects, including speed, accuracy, and adaptability, to provide a comprehensive measure of a player's cognitive performance on tasks including reasoning, deduction, and problem-solving within a game or test framework.

Copyright@ REST Publisher

Optimization Techniques

Multi-Layer Perceptron: A multilayer perceptron (MLP) is a type of feedforward neural network consisting of multiple layers of interconnected neurons. It serves as a fundamental model in deep learning and is well suited for capturing complex nonlinear patterns in data. MLPs are commonly used in tasks such as classification and regression.

3. ANALYSIS AND DISCUSSION

The dataset "Games and Logic" contains numerical records representing players' performance under various puzzle conditions. Each record has three input variables—Puzzle Difficulty, Time Limit Sec, and Player Experience Yrs and one output variable, Logic Score. The values have wide ranges: Puzzle Difficulty varies from approximately 1 to almost 10, Time Limit Sec extends from a few dozen seconds to almost 600 seconds, and Player Experience Yrs ranges from less than a year to almost 15 years. The main performance measure, Logic Score, varies from zero to about 74. From the patterns in the data, it is clear that performance is not simply a matter of more time or more experience. For example, some players with more experience and longer time limits still recorded Logic Scores of zero, indicating that puzzle difficulty or other unmeasured factors strongly influence outcomes. Conversely, there are instances where moderate experience and time constraints yield Logic Scores above 60, indicating that skill, strategy, and perhaps puzzle type play a major role. Interestingly, very low or very high puzzle difficulties do not consistently correlate with excellent scores. This may indicate that both very easy and very challenging puzzles can suppress performance, although for different reasons. Zero scores are scattered throughout the dataset, showing that no single condition guarantees failure.

TABLE 2. Descriptive Statistics

	Puzzle Difficulty	Time Limit Sec	Player Experience Yrs	Logic Score
Count	200	200	200	200
Mean	5.40452	319.4885	7.4187	17.6541
Std	2.562589	169.52239	4.496608	18.866508
Min	1.013	32.4	0.1	0
25%	3.3555	170.225	3.6075	0
50%	5.2305	321	7.08	11.56
75%	7.67175	466.75	11.725	29.8825
Max	9.933	599.5	14.88	74.35

Table 2 summarizes the descriptive statistics for the games and logic dataset. The variables show considerable variability, reflecting a variety of player conditions and performance levels. Puzzle difficulty averaged 5.40 on a scale of 1–10, with a moderate spread (SD ≈ 2.56), indicating that it included both easy and very challenging puzzles. The time range was wide (32.4 to 599.5 seconds), with an average of about 319 seconds, indicating that the dataset included both time-pressured and relaxed situations. Player experience years ranged from 0.1 to almost 15 years, with an average of 7.42 years, showing a balanced mix of novice and experienced players. The logic score was highly variable (0 to 74.35), with a mean of 17.65 and a median of 11.56, indicating that many players scored relatively low, while a few scored exceptionally high.

Copyright@ REST Publisher

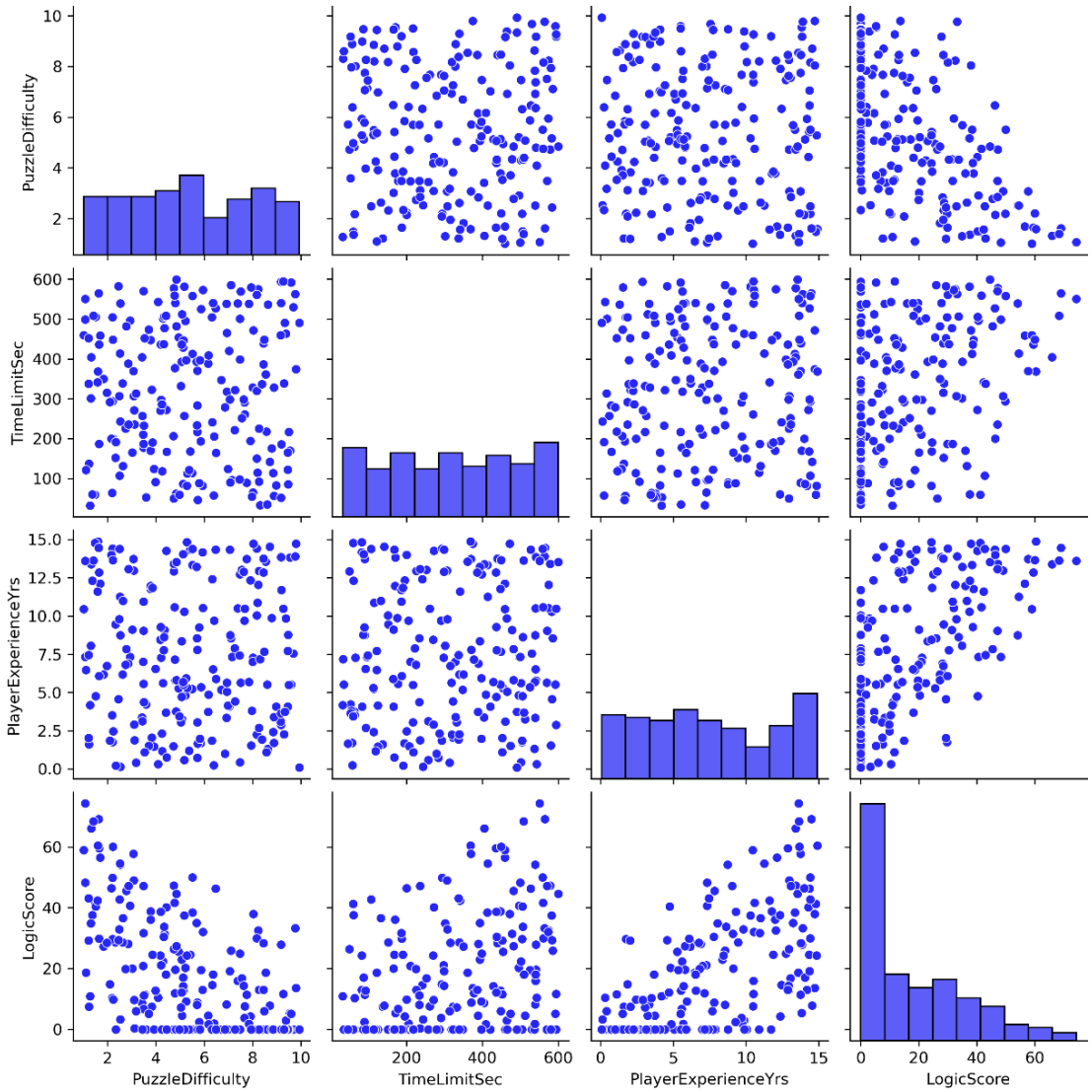


FIGURE 1. Scatter plot of various games and logic

Figure 1 presents a scatterplot illustrating the relationships between puzzle difficulty, time limit sec, player experience years, and logic score in the game and logic dataset. The diagonal histograms reveal that puzzle difficulty and time limit sec are relatively evenly distributed, while player experience shows a slight increase towards higher experience levels. The logic score is skewed towards lower values, confirming the previous descriptive statistics. The scatterplots indicate a weak positive correlation between player experience and logic score, indicating that experience can modestly improve performance. Conversely, puzzle difficulty shows a small negative relationship with logic score, indicating that harder puzzles tend to lower scores. The other parameter relationships appear weak or scattered, without strong linear trends.

Copyright@ REST Publisher

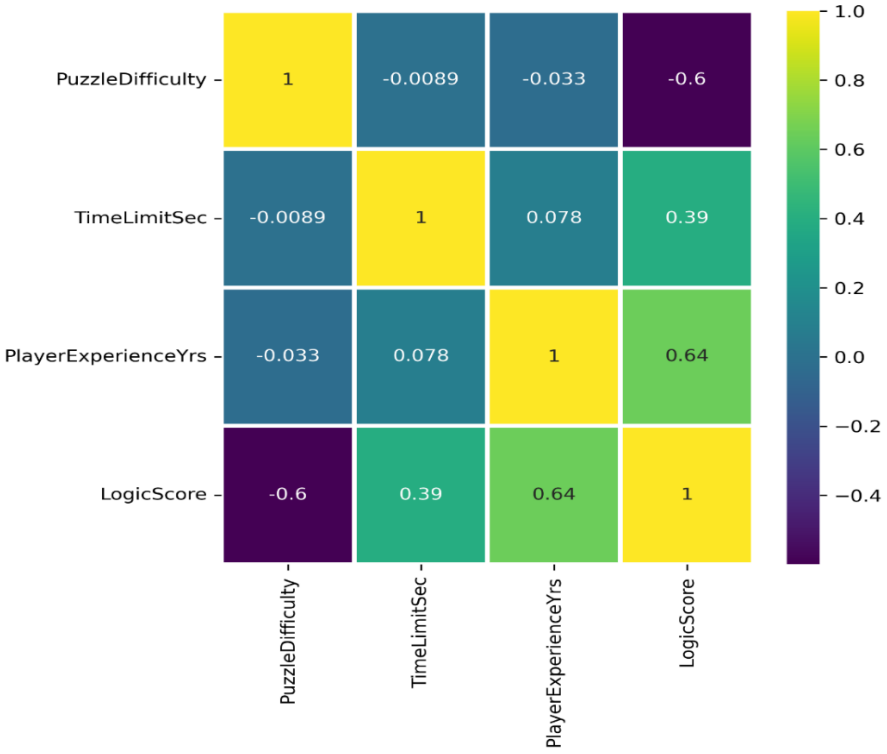


FIGURE 2. Correlation heat map between the process parameters and the responses

Figure 2 illustrates the correlation heat map for the Games and Logic dataset, which shows the relationships between puzzle difficulty, time limit, player experience, and logic score. A strong negative correlation (r = -0.60) is observed between puzzle difficulty and logic score, indicating that as puzzles become more challenging, player performance on logic declines. In contrast, player experience in years exhibits a strong positive correlation with logic score (r = 0.64), indicating that experienced players generally achieve better results on logic problem-solving tasks. Time limit in seconds also exhibits a moderate positive correlation with logic score (r = 0.39), indicating that giving players more time contributes to higher performance. The correlations between the input parameters—puzzle difficulty, time limit, and player experience—are relatively weak, with values close to zero, indicating that these parameters are largely independent in the dataset. This interaction structure highlights that while puzzle difficulty presents a significant challenge, player experience and adequate time allocation can significantly improve logical performance.

Multi-Layer Perceptron

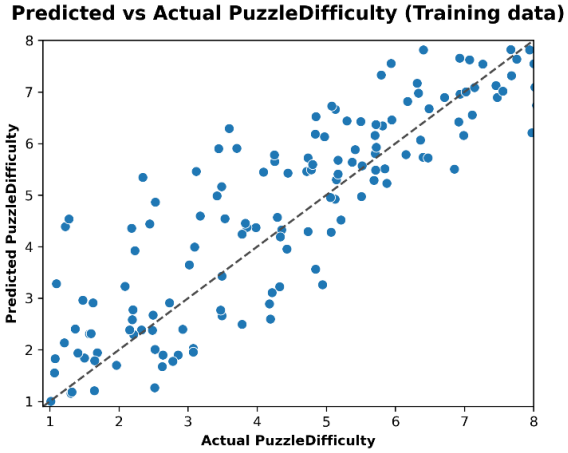


FIGURE 3. Multi-Layer Perceptron Puzzle Difficulty (Training data)

Copyright@ REST Publisher

Figure 3 presents the relationship between the predicted and actual puzzle difficulty values generated by a multi-layer perceptron (MLP) model on the training dataset. The scatterplot reveals a strong positive correlation, indicating that the model effectively learned the underlying patterns in the training data. Most of the data points are closely aligned with the diagonal dashed reference line, indicating perfect prediction accuracy. This alignment indicates that the MLP model successfully approximates the actual difficulty levels for most of the training samples. However, some degree of scatter is visible around the line, reflecting small prediction errors. These deviations, when present, appear relatively small and do not show systematic biases across the range of difficulty. The model performs consistently for both low and high difficulty levels, indicating good generalization within the training set and no significant overfitting at this point. The distribution of points also shows that the predicted values rarely deviate significantly from the actual scores, further demonstrating the consistency of the learning process.

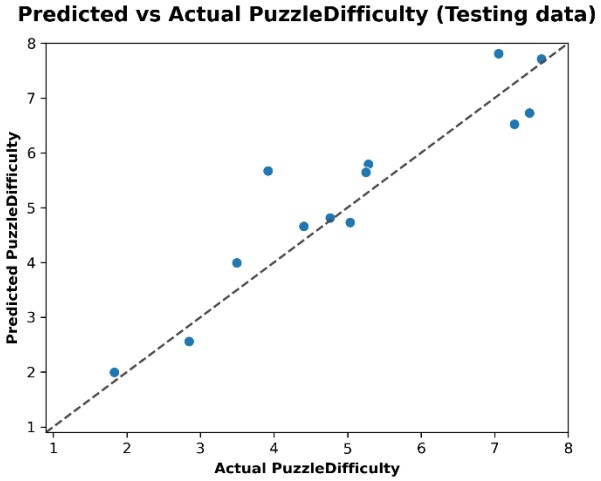


FIGURE 4. Multi-Layer Perceptron Puzzle Difficulty (Testing data)

Figure 4 illustrates the predicted and actual puzzle difficulty values for the test dataset using the Multi-Layer Perceptron (MLP) model. The scatter plot provides insight into how well the model generalizes to unseen data. Compared to the training performance, the test results show a moderate alignment with the diagonal reference line, indicating excellent prediction accuracy. Although many data points lie close to this line, indicating accurate predictions, a significant spread is observed across the plot. This spread reflects the presence of prediction errors, indicating that the model's performance on the test set is less consistent than on the training data. The reduced number of points in the test set highlights the limited size of the test samples, which may increase apparent variability. Despite this, the general upward trend confirms that the model reasonably captures the relationship between input features and difficulty levels. Some underestimation and overestimation can be observed, especially for medium-level difficulty values, where predictions deviate from actual observations. The figure suggests that while the MLP model shows acceptable predictive ability on new data, its performance is not as robust as it was during the training phase. This indicates potential room for improvement through hyper parameter tuning, feature refinement, or increasing the sample size.

TABLE 2. Multi-Layer Perceptron Puzzle Difficulty (Training, Testing Data)

Data	Symbol	R2	EVS	MSE	RMSE	MAE	MaxError	MSLE	MedAE
Train	MLP	0.744484	0.744494	1.633481	1.278077	0.975556	3.369118	0.057258	0.726915
Test	MLP	0.800341	0.822448	1.247993	1.117136	0.818938	3.178309	0.018583	0.548401

Table 2 provides performance metrics for the Multi-Layer Perceptron (MLP) model for predicting puzzle difficulty on both the training and test datasets. The R^2 values indicate reasonably strong predictive performance, with the model reaching 0.744 on the training data and slightly higher at 0.800 on the test data. Since the test R^2 exceeds the training value, this indicates that the model generalizes well and is not overfitting. The Explained Variance Score (EVS) follows a similar trend, with the test EVS (0.822) outperforming the training EVS (0.744), further confirming the consistent generalization capabilities. The MSE and RMSE values for both datasets are relatively low, indicating that the average prediction errors of the model are moderate. Notably, the test MSE (1.2479) and RMSE (1.1171) are

Copyright@ REST Publisher

slightly lower than those of the training set, indicating improved error control on the unobserved data. Additionally, the MAE and MedAE values for both datasets are less than 1, which highlights that the typical prediction deviations are small. The Max Error values indicate that large errors occur rarely, but are within acceptable limits. The table demonstrates that the MLP model performs reliably, with strong consistency across training and testing metrics, making it a suitable model for predicting puzzle difficulty.

4. CONCLUSION

This research successfully combines theoretical game logic frameworks with empirical machine learning methods to analyze and predict logical reasoning performance in puzzle-solving environments. The multi-layer perceptron model demonstrates exceptional predictive ability, achieving better performance on test data ($R^2 = 0.800$) compared to training data ($R^2 = 0.744$), indicating strong generalization without overfitting. Low error metrics across MSE, RMSE, and MAE confirm the model's reliability in capturing complex nonlinear relationships between player characteristics and performance outcomes. Correlation analysis reveals important insights into factors influencing logical performance. The strong negative correlation between puzzle difficulty and logic scores ($r = -0.60$) underscores the significant cognitive demands imposed by complex challenges. Conversely, player experience exhibits a strong positive correlation ($r = 0.64$), demonstrating that accumulated expertise significantly improves problem-solving skills. The moderate positive correlation of time constraints ($r = 0.39$) indicates that sufficient temporal resources contribute meaningfully to performance, although to a lesser extent than experience. Notably, the dataset reveals the occurrence of zero scores across a variety of conditions, indicating that no single parameter guarantees success or failure. This complexity suggests that logical performance emerges from complex interactions between puzzle properties, player skills, and contextual constraints. The findings are consistent with theoretical frameworks discussed in the literature, including dynamic epistemic logic, coalitional game strategies, and preference structures in rational decision making. Practical implications extend to several domains: adaptive educational systems can improve challenge difficulty based on learner experience; game designers can balance puzzle difficulty with player progress; and cognitive assessment tools can account for individual differences in expertise and time management. Future research should examine additional variables such as puzzle type, cognitive load measures, and real-time strategy adaptation, which should further refine predictive models. Expanding the dataset and incorporating advanced deep learning frameworks could improve model performance and reveal deeper insights into the cognitive mechanisms underlying strategic reasoning in game-theoretic contexts.

REFERENCES

[1]. Bulling, Nils, Valentin Goranko, and Wojciech Jamroga. "Logics for reasoning about strategic abilities in multi-player games." In Models of strategic reasoning: logics, games, and communities, pp. 93-136. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016.

[2]. van Benthem, Johan, and Fenrong Liu. "Graph games and logic design." In Knowledge, Proof and Dynamics: The Fourth Asian Workshop on Philosophical Logic, pp. 125-146. Springer Singapore, 2020.

[3]. Blass, Andreas. "A game semantics for linear logic." Annals of Pure and Applied logic 56, no. 1-3 (1992): 183-220.

[4]. Rohde, Philipp. "On games and logics over dynamically changing structures." PhD diss., Aachen, Techn. Hochsch., Diss., 2005, 2005.

[5]. Goranko, Valentin. "Coalition games and alternating temporal logics." In Theoretical Aspects Of Rationality And Knowledge: Proceedings of the 8 th conference on Theoretical aspects of rationality and knowledge, vol. 8, no. 10, pp. 259-272. 2001.

[6]. Maubert, Bastien, Sophie Pinchinat, François Schwarzentruber, and Silvia Stranieri. "Concurrent games in dynamic epistemic logic." In Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence, pp. 1877-1883. 2021.

[7]. Berwanger, Dietmar. "Games and logical expressiveness." PhD diss., Bibliothek der RWTH Aachen, 2005.

[8]. Ghosh, Sujata, and Ramaswamy Ramanujam. "Strategies in games: A logic-automata study." In European Summer School in Logic, Language and Information, pp. 110-159. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.

[9]. Abramsky, Samson. "Sequentiality vs. concurrency in games and logic." Mathematical Structures in Computer Science 13, no. 4 (2003): 531-565.

[10]. Sandu, Gabriel, and Ahti Pietarinen. "Partiality and games: propositional logic." Logic Journal of IGPL 9, no. 1 (2001): 101-121.

[11]. Kuusela, Oskari. "The method of language-games as a method of logic." Philosophical Topics (2014): 129-160.

Copyright@ REST Publisher

[12]. Van Benthem, Johan, Sieuwert Van Otterloo, and Olivier Roy. "Preference logic, conditionals and solution concepts in games." (2005).
[13]. Ramanujam, Ramaswamy, and Sunil Easaw Simon. "Dynamic Logic on Games with Structured Strategies." In KR, pp. 49-58. 2008.
[14]. Maubert, Bastien, Sophie Pinchinat, and François Schwarzentruber. "Reachability games in dynamic epistemic logic." arXiv preprint arXiv:1905.12422 (2019).
[15]. Chen, Taolue, Vojtěch Forejt, Marta Kwiatkowska, David Parker, and Aistis Simaitis. "PRISM-games: A model checker for stochastic multi-player games." In Tools and Algorithms for the Construction and Analysis of Systems: 19th International Conference, TACAS 2013, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2013, Rome, Italy, March 16-24, 2013. Proceedings 19, pp. 185-191. Springer Berlin Heidelberg, 2013.

Copyright@ REST Publisher