AN INTELLIGENT TIC TAC TOE PLAYER

Assignment 1

Shahmir Raza	34566469
Saeed Ul Hassan	34673092
Ammar Iftikhar	34520497

Table of Contents

1. Description:

The AI Tic Tac Toe game utilizes reinforcement learning (RL) techniques to create an innovative and adaptive gaming experience (Kalra, 2022). Tic Tac Toe, a classic game with simple rules, presents a perfect environment for exploring the capabilities of RL algorithms. Unlike traditional rule-based or heuristic approaches, which may follow predetermined strategies, RL allows the AI to learn and adapt its gameplay through interaction with the environment. The issue space we are focusing on is encouraging an artificial knowledge system to play Tic Tac Toe, a praiseworthy two-player game. Tic Tac Toe is a straightforward yet testing game with an immense request space, making it an exciting issue for PC based knowledge research (Lee et al., 2014). Usually, tending to Tic Tac Toe incorporates methodologies such as minimax estimation with alpha-beta pruning, which can genuinely research the game tree to find the best move. Nevertheless, the chase space becomes enormous as the game advances, making it computationally exorbitant to find the ideal move, especially in extra convoluted games like chess or Go.

One innovative method for managing this issue is using Reinforcement learning (RL) techniques. RL grants the artificial insight expert the ability to learn ideal strategies through association with the environment without requiring unequivocal data on the game rules or heuristics. This can provoke more versatile and adaptable experts who manage different game circumstances and adversaries. The man-caused knowledge structure we propose will use a mix of techniques, including Q-learning and a sort of RL estimation, to get to know the best procedure for playing Tic Tac Toe. Q-learning is a sans model RL computation that sorts out some way to make decisions by surveying the advantage of taking a particular action in a given state. By iteratively playing and reviving its Q-values considering compensations, the PC-based knowledge expert can continuously chip away at its show in the long term. Besides, we can examine additionally created techniques, for instance, Deep Reinforcement Learning (DRL), where cerebrum networks are used to assess the Q-values, taking into account more complicated and nuanced headings (Stember & H Shalu, 2022).

The data expected for planning such an artificial knowledge structure would consolidate innumerable game entertainments, where the recreated insight expert plays against itself or human players. These game multiplications give the expert an essential contribution to learn and chip away at its framework in the long term. By and large talking, by using RL methodologies like Q-learning or DRL, our man-created insight system can offer a more inventive and flexible response for playing Tic Tac Toe, which stood out from standard pursuit-based techniques. This approach might prevail in Tic Tac Toe and other complex games and genuine applications where dynamic in strong circumstances is required (Stember & Hrithwik Shalu, 2020).

2. Background

Tic Tac Toe, generally called Noughts and Crosses, is a model two-player game played on a 3x3 network (Hidayati, 2015). The game's straightforwardness makes it a popular choice for informational purposes and a benchmark for testing different PC-based knowledge computations. Despite its ease, Tic Tac Toe presents captivating moves for mimicked insight structures because of its gigantic interest space and key multifaceted nature. The game starts with an unfilled organization, and players substitute setting their pictures (consistently X or O) in void cells (Inan et al., 2021). The objective is to shape an even, vertical, or corner to corner line of three of one's pictures before the adversary does like this. If all cells are loaded up with close to no player achieving the objective, the game terminations are in a draw. In the standard manner to manage settling Tic Tac Toe, estimations like the minimax computation with alpha-beta pruning are ordinarily used (Plaat, 2024).

The minimax computation explores the entire game tree to find the ideal move for the continuous player, tolerating the foe plays faultlessly. Anyway, this approach becomes computationally exorbitant as the game advances in light of the fanning variable of the game tree. Another ordinary system uses heuristic evaluation capacities to assess the allure of game states. These abilities allow scores to different board plans considering explicit norms, similar to each player's anticipated winning lines. While heuristic appraisal capacities can lessen the pursuit space, they habitually require ace data and may need to be summarized better in different game circumstances. Lately, there has been increasing interest in applying Reinforcement learning (RL) techniques to games like Tic Tac Toe (Kalra, 2022).

RL is an artificial intelligence perspective where an expert sorts out some way to seek after decisions by interfacing with an environment and getting analysis as compensations (Ade, 1997). By acquiring knowledge from trial and error, RL experts can track down ideal frameworks without an express course. One of the key RL computations used in game playing is Q-learning. Q-learning is a sans model RL computation that sorts out some way to make decisions by surveying the advantage of taking a particular action in a given state (Clifton & Laber, 2020). The expert keeps a Q-regard table, where each segment tends to the typical complete pay for taking a specific action in a specific state. Through reiterated joint efforts with the environment, the expert revives its Q-values considering the awards got, consistently chipping away at its technique after some time (Fuchida et al., 2010). Besides, Deep Reinforcement Learning (DRL) techniques have been applied to games like Tic Tac Toe.

DRL unites RL with significant getting, using cerebrum associations to infer the Q-values. This is considered a more incredible and nuanced autonomous bearing, as the expert can acquire from rough game states without relying on high-quality components. RL methodologies like Q-learning and DRL can provide a more inventive and flexible method for playing Tic Tac Toe, which stands out from standard pursuit-based procedures. These PC based knowledge systems may prevail in Tic Tac Toe and in extra confounding games and veritable applications where dynamic in strong circumstances is required (Yuan et al., 2019).

3. Methodology

The way of thinking for cultivating a mimicked insight system to play Tic Tac Toe using Reinforcement learning (RL) methodology like Q-learning and potentially Deep Reinforcement Learning (DRL) incorporates a couple of key stages. Here is a bare essential graph of the technique:

- **Problem Definition:**
  Portray the issue of playing Tic Tac Toe as a help learning task. Demonstrate the game environment, including the state space, action space, and grant structure. Conclude the objective of the PC-based insight trained professional, which is to acquire capability with an ideal plan for making moves that increase its prospects of overwhelming or drawing the match. Environment
- **The course of action:**
  Complete the Tic Tac Toe game environment, including the game board depiction, move endorsement, and game outcome affirmation (win, lose, draw). Describe the state space

depiction conventionally incorporating the continuous board plan. Describe the action space, which contains all the expert's potential moves on the continuous board.

- **Reinforcement Learning Algorithm Selection:**
  Pick a sensible RL computation to learn the best methodology in the Tic Tac Toe environment. Start with principal Q-learning estimation, fitting for discrete movement spaces and even state depictions. On the other hand, additional methods like significant Q-associations (DQN) or procedure tendency strategies for complex state spaces or industrious action spaces should be considered.
- **Agent Implementation:**
  Execute the RL-trained professional, who speaks with the environment by picking exercises considering its learned methodology and reviving its Q-values. Instate the Q-regard table or cerebrum network limits. Execute the movement decision framework, for instance, ε-excited examination to change examination and cheating.
- **Training Process:**
  Train the RL expert by playing against itself or human players. During each game, let the expert select exercises considering the continuous system and update its Q-values using the saw prizes. Use a legitimate award capacity to give analysis to the subject matter expert, for instance, giving a positive pay for winning, a negative honor for losing, and a fair pay for drawing.
- **Evaluation and Fine-Tuning:**
  Survey the show of the pre-arranged expert by assessing estimations like win rate, draw rate, and typical game length. Adjust the RL computation hyperparameters, such as learning rate, markdown component, and examination rate, to foster execution. To find the best blend, investigate various roads regarding state depictions, reward abilities, and examination procedures.
- **Extension to Deep Reinforcement Learning:**
  If fundamental, extend the critical Q-advancing method for managing significant help progressing using mind associations to harsh Q-values. Complete a significant Q-association (DQN) design, which takes the board state as data and results in Q-values for every action. Train the DQN expert using systems like experience replay and target network updates to foster security and mixing. When the recreated insight expert achieves excellent

execution, send it for authentic use cases. Integrate the PC-based knowledge expert into applications or stages where Tic Tac Toe playing limits are needed, such as informative devices, game-playing destinations, or artificial insight partners.

4. AI Techniques:

We should contemplate two mimicked insight systems for assessment: Minimax estimation with alpha-beta pruning and Q-learning. Here is a fundamental examination of every technique and legitimization for picking

- **Minimax Algorithm with Alpha-Beta Pruning:**
  Minimax is a request based estimation regularly used for poorly arranged games like Tic Tac Toe. It researches the entire game tree to find the ideal move by flawlessly tolerating the opponent's plays. Alpha-beta pruning is a methodology to decrease the number of center points surveyed in the minimax search tree, making it more capable. Guarantees ideal play against an optimal adversary. Certainly knew and, for the most part, used in game playing. It is computationally expensive, especially for games with tremendous spreading factors. It requires an examination of the game tree, which may only be viable sometimes. Minimax with alpha-beta pruning is an excellent game-playing philosophy and fills in as a benchmark for evaluating the introduction of various systems like Q-learning.
- **Q-learning:**
  Q-learning is a help-learning computation that learns ideal systems through trial and error. Concerning Tic Tac Toe, the man-made knowledge expert sorts out how to make ideal moves by reviving Q-values related to state-action matches considering compensations. Does not require all-out data on the game tree or foe frameworks. Can manage huge state spaces gainfully. Versatile and flexible to dynamic circumstances. This may call for more prominent speculation to meet diverged from search-based estimations like minimax. May fight to summarize well to hid states or adversaries. Q-learning offers a more versatile and flexible method for managing learning ideal systems in Tic Tac Toe. It does not rely upon careful pursuit and can manage weaknesses naturally in apparent circumstances. While it could call for more noteworthy speculation to blend diverged from minimax, its ability to acquire reality makes it sensible for dynamic and creating game circumstances.

- **Comparison and Justification:**

  Q-learning is picked over the minimax estimation with alpha-beta pruning considering different elements. Q-learning is more fitting for dynamic circumstances like Tic Tac Toe, where the complete game tree may be nonsensical to examine. Q-learning does not require all-out data on the game tree or adversary frameworks, making it more flexible to dark or propelling foes.

Summary:

While Q-learning could request more prominent speculation to join diverged from minimax, its ability to acquire in actuality makes it proper for genuine applications where weaknesses exist. Early on, trial results could incorporate planning Q-learning and executing minimax with alpha-beta pruning experts to play against each other or human players. The display of each and every strategy can be evaluated by considering estimations like win rate, draw rate, and ordinary game length. These results would give trial evidence to help the decision to Progress Q as the team leaned toward a method for playing tic tac toe.

- **Implementation Details:**

  The AI agent encodes the state of the game board, representing possible moves and their corresponding rewards.

- **Data Source:**

  Data for training and testing the AI system can be generated through simulated gameplay or observing human players by playing with them.

- **Data Preprocessing:**

  The AI system may require preprocessing steps to normalize input data and enhance learning efficiency.

- **Comparison with Traditional AI technique:**

  Unlike traditional rule-based approaches, which rely on predefined strategies, the RL-based AI adapts its gameplay based on learned experiences. This adaptability makes it more resilient against diverse opponent strategies and allows continuous improvement.

- **Product Differentiation:**

  Our AI game stands out from traditional Tic Tac Toe implementations by offering an adaptive and challenging AI opponent that continually improves its gameplay through RL techniques. Unlike basic rule-based algorithms, our AI agent can analyze complex game states and make strategic decisions, providing users with a more immersive and dynamic gaming experience.

- **Interpretation of Output:**

  The output of the system is interpreted as the optimal move suggested by the AI agent itseld based on its learned policy and training. This move aims to maximize the AI's chances of winning or achieving a draw, depending on the current game state.

- **Potential Investors:**

  Investors interested in AI gaming applications can benefit from the AI Tic Tac Toe system's unique selling points, such as adaptive gameplay, scalability and complexity, and competitive edge over traditional or rule-based systems.

5. Conclusion:

The AI Tic Tac Toe game leverages reinforcement learning techniques to provide an innovative and adaptive gaming experience. With its ability to learn and adapt strategies based on experience, the AI system offers a competitive edge over traditional approaches. Investors interested in AI gaming applications can benefit from the scalability and versatility of the proposed AI system.

References

Ade, F. (1997). The Role of Artificial Intelligence in the Reconstruction of Man-made Objects from Aerial Images. *Birkhäuser Basel EBooks*, 23–32.

Clifton, J., & Laber, E. (2020). Q-Learning: Theory and Applications. *Annual Review of Statistics and Its Application*, *7*(1), 279–301.

Fuchida, T., Aung, K. T., & Sakuragi, A. (2010). A study of Q-learning considering negative rewards. *Artificial Life and Robotics*, *15*(3), 351–354.

Hidayati, H. (2015). THE USE OF THE TIC TAC TOE GAME IN TEACHING SPEAKING SKILLS. *Linguistics and Elt Journal*, *3*(1).

Inan, M. S. K., Hasan, R., & Prama, T. T. (2021, December 1). *An Integrated Expert System with a Supervised Machine Learning based Probabilistic Approach to Play Tic-Tac-Toe*. IEEE Xplore.

Kalra, B. (2022). Generalized agent for solving higher board states of tic tac toe using Reinforcement Learning. *ArXiv (Cornell University)*.

Lee, T. Y., Mauriello, M. L., Ahn, J., & Bederson, B. B. (2014). CTArcade: Computational thinking with games in school age children. *International Journal of Child-Computer Interaction*, *2*(1), 26–33.

Plaat, A. (2024, March 20). *Research Re: search & Research*. ArXiv.org.

Stember, J. N., & H Shalu. (2022). Reinforcement learning using Deep Q networks and Q learning accurately localizes brain tumors on MRI with very small training sets. *BMC Medical Imaging*, *22*(1).

Stember, J. N., & Hrithwik Shalu. (2020). Deep reinforcement learning to detect brain lesions on MRI: a proof-of-concept application of reinforcement learning to medical images. *ArXiv (Cornell University)*.

Yuan, Y., Yu, Z. L., Gu, Z., Yeboah, Y., Wei, W., Deng, X., Li, J., & Li, Y. (2019). A novel multi-step Q-learning method to improve data efficiency for deep reinforcement learning. *Knowledge-Based Systems*, *175*, 107–117.