Reinforcement Learning Approach for Intelligent Page Replacement

Mallikarjuna Rao Tota

Email id: malli.2813@gmail.com

Designation: Research Scholar

Dept: Department of Computer Science and Informatics

College Name: University College of Engineering and Technology, Mahatma Gandhi University, Nalgonda, 508254

Rekha Redamalla

Email id: rrekhardy@gmail.com

Designation: Professor

Dept: Department of Computer Science and Informatics

College Name: University College of Engineering and Technology, Mahatma Gandhi University, Nalgonda, 508254

Abstract

The research focuses on the concept of reinforcement learning approach for intelligent page replacement. The improvement in “Reinforcement learning (RL)” can lead to enhanced intelligent page replacement with the help of enabling systems to improve learning of optimal page eviction policies by experience instead of relying on static algorithms such as FIFO or LRU. The objectives of the research include, to develop a reinforcement learning agent for enhanced page selection, comparison of performance with Optimal Page Replacement benchmark algorithm, adaptability under the varying workload conditions along with long-term learning efficiency in lowering page faults. “Explanatory research design”, “Positivism research philosophy” and “Deductive research approach” has been discussed. This study uses the secondary data approach to analyse reinforcement learning based page replacement strategies. Through thematic analysis based on peer-reviewed literature, key dimensions such as adaptability, performance optimisation, computational complexity and implementation feasibility were reviewed. The research concludes that reinforcement learning models can be used for improving page hit ratios and reducing page faults by dynamically learning from memory accessibility patterns. However, there are still significant challenges related to issues such as scalability, state-space representation and runtime overhead. Overall, reinforcement learning shows a lot of promise as an adaptive alternative to the traditional heuristic based page replacement techniques used in modern operating systems.

Keywords: Reinforcement Learning, Page Replacement, Intelligent Page Replacement.

1. Introduction

1.1 Context and Background of the Study

“Reinforcement Learning (RL)” for intelligent page replacement assists in optimising memory management through training an agent to make “victim page” selection decisions that assists in increasing long-term rewards such as lower page fault rates considering the current memory states. Unlike the fixed algorithms such as FIFO or LRU, RL-based approaches learn to adapt to different workloads in real-time that positively impacts in enhancing DRAM hit rates as well as reducing workload runtime. According to the ideas of Shi *et al.* (2023), the replacement strategy is reflected as a “Markov Decision Process (MDP)” in which the agent focuses on observing the memory state, receiving feedback along with taking actions. On the contrary to this, Lan *et al.* (2022), approaches often use “value-based DQN (Deep Q-Network)” or “Deep Reinforcement Learning (DRL)” to learn optimal strategies from the experience devoid of prior knowledge of the memory access patterns. Besides this, it can be noted that Reinforcement Learning agents assist in lowering cache miss rates through enhancing learning of complex patterns along with this providing efficient I/O processing instead of traditional algorithms. Apart from this, Reinforcement Learning driven systems can aid to achieve superior performance in comparison to the heuristic methods, specifically in the scenarios of hybrid main memory.

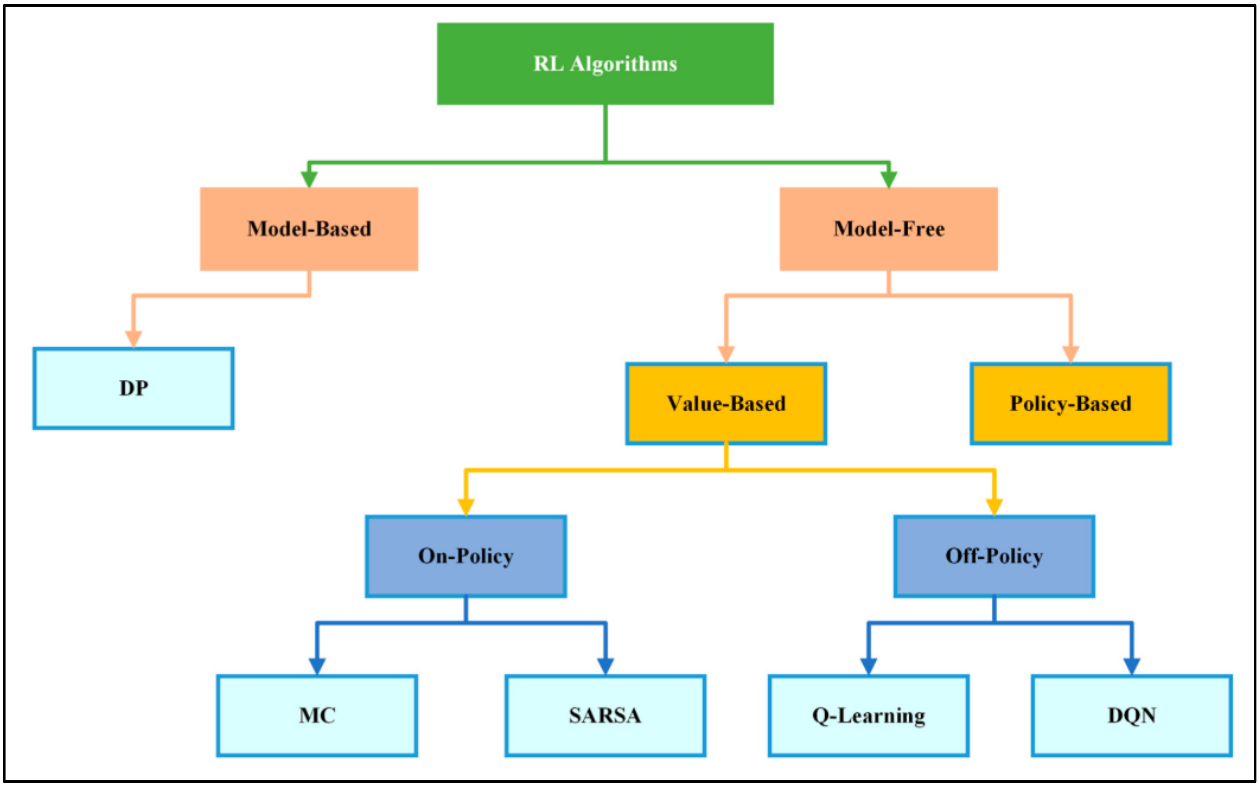

Figure 1.1: Reinforcement learning algorithms

(Source: Al-Hamadani *et al.* 2024)

1.2 Problem Statement

The problem statement of this research is that static, traditional page replacement algorithms such as "First-In, First-Out (FIFO)" and "Least Recently Used (LRU)" often fail to adapt to the diverse and dynamic memory access patterns of the different modern applications. In addition to this it can lead to increased page faults, suboptimal performance along with higher I/O latency. The noteworthy problem can be the inability of the prior defined, rule-based policies to dynamically learn as well as self-tune in response to the switching of application memory access behaviours such as switching iterative access patterns and between common.

1.3 Aim and Objectives of the Study

Aim

The aim of this research is to analyse the reinforcement learning approach for intelligent page replacement.

Objectives

- To develop a reinforcement learning agent for dynamic page selection.
- To compare performance with Optimal Page Replacement benchmark algorithm.
- To evaluate adaptability under varying workload conditions.
- To assess long-term learning efficiency in reducing page faults

1.4 Research Questions

Q1: What are the ways to develop a reinforcement learning agent for dynamic page selection?

Q2: What is the comparison between performance with Optimal Page Replacement benchmark algorithm?

Q3: How does adaptability under varying impacts workload conditions?

Q4: What are the long-term learning efficiency in reducing page faults?

1.5 Significance of the study

The respective research is significant to enhance understanding regarding the memory management systems to automatically adapt to the rapidly changing access patterns of application memory in real time. In addition to this, it can also lead to lower page faults along with noteworthy performance improvements in data-intensive, modern applications. The respective research is also significant in enhancing knowledge regarding the traditional page replacement policies such as "Least Recently Used (LRU)" that work appropriately for some applications. However, the traditional page replacement policies perform poorly with others, such as "exhibiting iterative memory access" patterns such as graph processing or scientific applications (Wu *et al.* 2024). Furthermore, the research is also helpful in analysing different challenges along with strategies to resolve.

1.6 Brief Review of Relevant Literature

Based on the perspectives of Kim and Yeom (2022), stated that "Reinforcement Learning (RL)" for intelligent page replacement is important as it assists to shift memory management from rule-based heuristics, static such as "Least Recently Used - LRU" to self-learning, dynamic systems. In addition to this, that assists in adapting with the changing patterns of memory access. However, Thyagaturu *et al.* (2022), mentioned that this approach also allows operating systems to lower the page faults as well as lower the execution time in real-time and eliminating the need for prior knowledge of the characteristics of applications. On the contrary, Elias *et al.* (2024), has highlighted that the "Optimal Page Replacement algorithm (OPT or MIN)" is reflected as a theoretical benchmark as it assists to guarantee minimum attainable number of page faults for the provided sequence of the memory references.

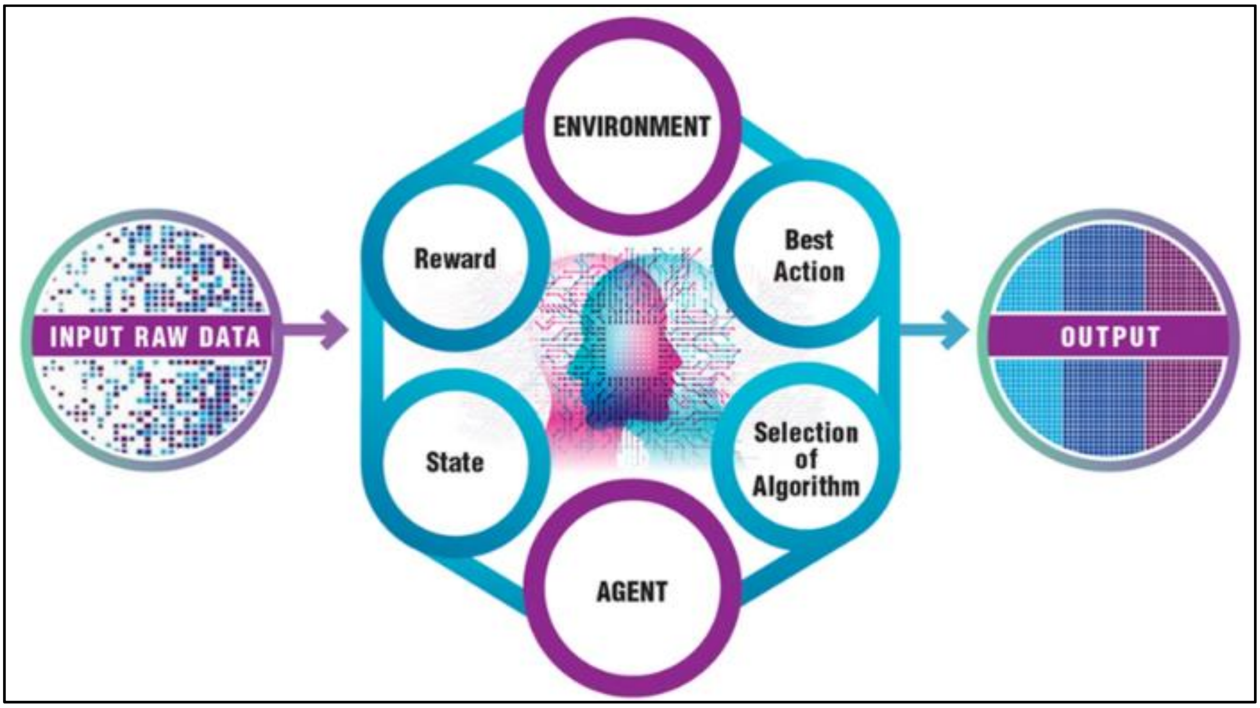

Figure 1.2: Reinforcement Learning

(Source: Kleinstreuer and Hartung, 2024)

Besides this, its performance is also considered as the ideal against all the practical algorithms. Apart from this, Kleinstreuer and Hartung (2024), mentioned that adaptability under the varying workload conditions reflects the ability of teams, individuals or technical systems in order to adjust processes, approaches along with allocation of resources. In addition to this, it assists to effectively manage the fluctuations in demand, avoid negative outcomes such as system failure or burnout followed by maintaining performance. Furthermore, adaptability can be reflected as the crucial skill in the dynamic environments that are characterised by rapid uncertainty and change (Shafizadeh, 2024).

2. Methodology

2.1 Research Onion

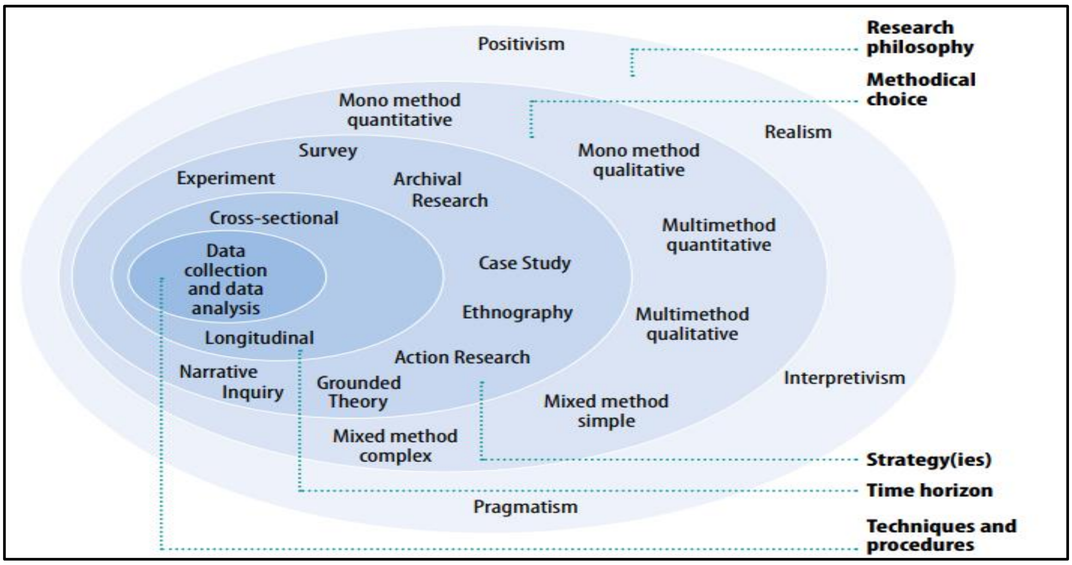

Figure 2.1: Research Onion

(Source: Saunders and Tosey, 2012)

2.2 Research Design and Methods Used

"Explanatory research design" is used in this research as it assists in identifying root causes, filling research gaps, enhancing predictive capabilities, high flexibility along with supporting informed decision-making. "Positivism research philosophy" assists in generalising the findings with maintaining a well-defined structure (Ali, 2024). In this research, "Positivism research philosophy" along with "Deductive research approach" is used that positively impacts in reducing biases in research such as analysing reinforcement learning agent for the dynamic page selection or other concepts.

2.3 Sample Selection and Data Collection Procedures

This research uses a secondary data collection paradigm to scrutinise the application of RL in intelligent page - replacement. Academic sources were identified systematically from reputable scholarly databases, including, among others, from the following: Academic sources were searched in the scholarly databases, or databases that allow more in-depth search possibilities and more detailed, comprehensive search functionalities. A purposive sampling strategy was used to ensure quality of peer-reviewed journal articles and conference proceedings were included.

The selection criteria focused on the research assembled during the previous ten years that dealt with reinforcement learning algorithms, adaptive caching mechanisms, operating system memory management and performance evaluation metrics such as page fault rate, hit ratio and computation overhead. Articles that offered empirical comparisons between classical page replacement algorithms and approaches based on RLs were given top priority.

2.4 Tools and Instruments Used

The frameworks of structured literature review and thematic extraction matrices were the main instruments used in this research. A comparative analysis table was designed to catalogue important variables from the selected studies including RL model type, state-action representation, reward formulation, learning strategy, and performance reported results.

In addition to this, there were synthesis tables built to compare traditional heuristic-based algorithms with adaptive RL-driven policies. These tools helped to organise the work systematically, improve traceability of evidence and ensure that the analytical work process was methodologically transparent.

2.5 Data Analysis Methods

This research methodology focuses on thematic analysis of the qualitative results obtained from literature selection. The analysis was conducted in a sequential manner that included familiarisation with the data, initial coding, theme identification, theme refinement and structured reporting (Naeem *et al.* 2023).

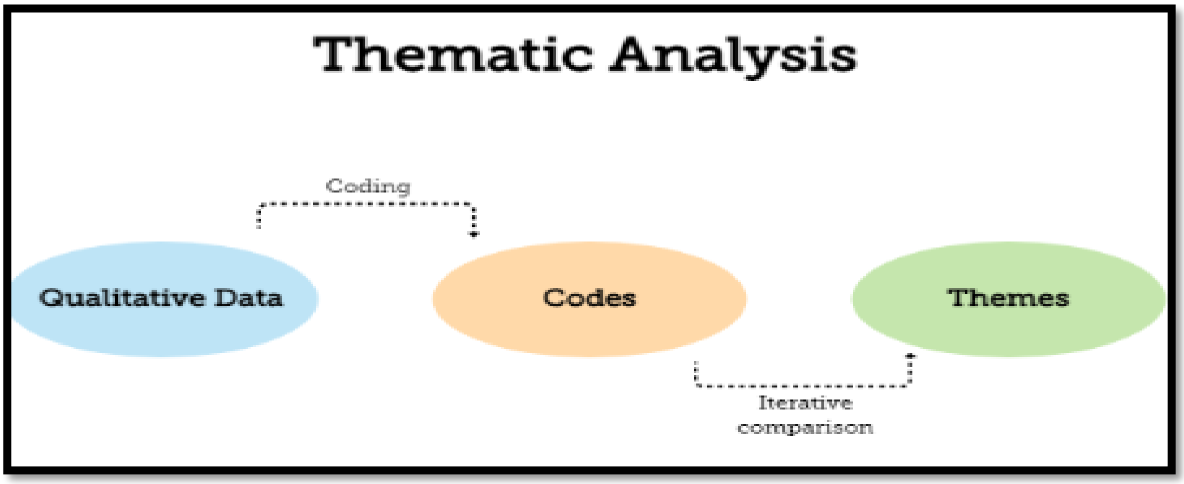

Figure 2.2: Thematic analysis

(Source: As inspired by Naeem *et al.* 2023)

The themes that are selected for the research mainly emerged around adaptive learning capability, optimisation of performance, challenging scalability, state space complexity and computational overhead. Through this approach, recurring trends and conceptual associations could be found, allowing to systematically compare the strategies for static page replacement with RL-based intelligent models.

2.6 Ethical Considerations

All ethical standards were maintained during the research process (Tripathi and v Chaturvedi, 2023). Only publicly accessible academic sources were utilised and appropriate referencing practices followed to ensure academic integrity. No proprietary system data, confidential data, or human participant data were performed or involved in this study.

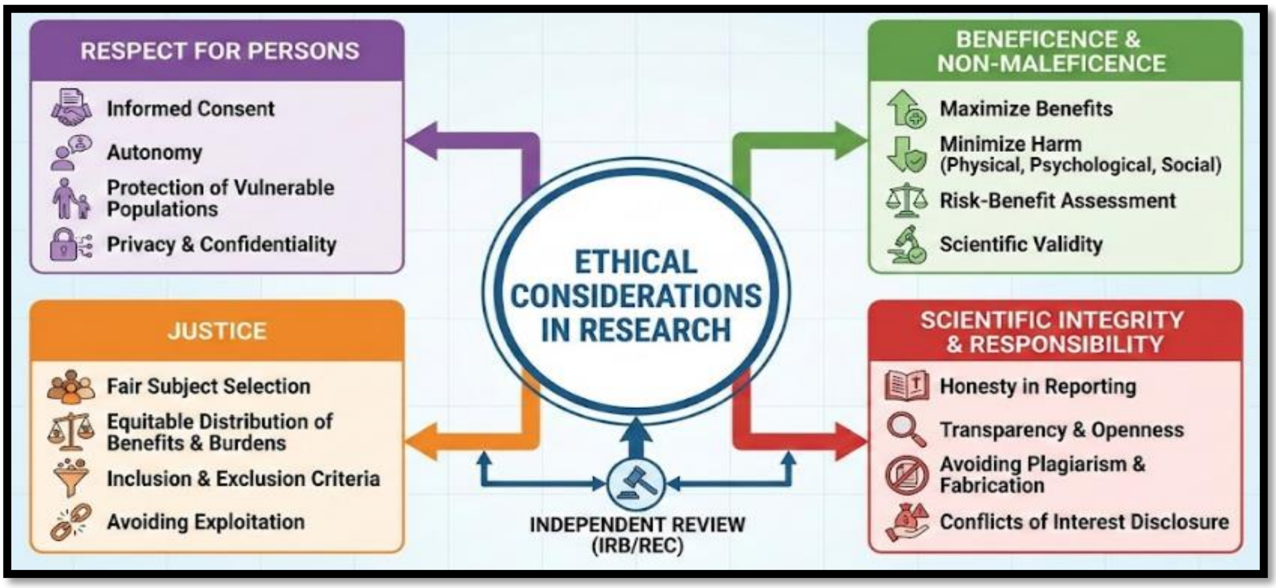

Figure 2.3: Ethical considerations

(Source: As inspired by Tripathi and v Chaturvedi, 2023)

2.7 Research Timeline

Tasks	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Initialising the research						
Identification research Aim & Objectives						
Searching literature and screening data						
Data extraction and thematic coding						
Comparative analysis and synthesis						
Compilation and refinement of research						

Table 1: Research Timeline

(Source: self-made)

3. Results

3.1 Research Findings

The thematic analysis of selected literature led to the identification of four major themes relating to the usage of RL in intelligent page replacement systems

Authors	Themes	Description
Moerland *et al.* 2023 Souza and Freitas, 2024	Adaptive Learning and Dynamic Decision-Making	The findings underline the importance of reinforcement learning to enable the dynamic adaptation to changing memory access patterns. RL-based models learn the best replacement procedures through constant interaction with the system environment (Moerland *et al.* 2023). Thus RL-based model differs from the traditional algorithms such as Least Recently Used (LRU) or First In First Out (FIFO) which works with predefined heuristics, By modelling page replacement as a Markov Decision Process (MDP) The RL agents observe system states such as page reference history and buffer occupancy (Souza and Freitas, 2024). After that these agents select actions that maximise long term rewards. This adaptive learning capability helps in the system to adapt well with the workload variability and reducing unnecessary page evictions leading to better memory utilisation.
Kim and Yeom, 2022 Ni *et al.* 2023	Performance Improvement in Terms of Page Fault Reduction	According to the findings when policies learned from RL learning are implemented improvement in performing metrics is clear. Compared to classical replacement strategies RL reduces page fault rates and increases in page hit ratios (Kim and Yeom, 2022). The reward structure for RL frameworks is set up in such a way to penalise page faults and to reward correct decisions. In addition, the agent helps to refine the policy so that memory swaps incur smaller costs (Ni *et al.* 2023). This performance optimization is especially visible when the access pattern(s) of the

		environment is/are non-stationary and static algorithms do not generalise well.
Liu *et al.* 2024 Wang *et al.* 2022	State-Space Complexity and Computational Overhead	Despite performance benefits, major issues regarding scalability and computational complexity (Liu *et al.* 2024). All these issues are pointed out by several authors. RL models require thought when it comes to formulating state, action space, reward mechanisms. In large scale memory systems, the state space can grow exponentially resulting in a growth increase in the training time and resource consumption. Deep reinforcement learning methods overcome some of these shortcomings using the approximation of functions (Wang *et ai,* 2022). However, then it comes with additional computational costs. Such considerations cause concern when considering real-time applicability in operating systems kernels, where above all you want to be efficient.
Chen and Heydari, 2024 Zhang *et al.* 2022	Implementation Feasibility and System Integration Challenges	The efficient embedding of RL models into existing operating systems architectures is also very important. Studies are showing that placing the learning agents in subsystems for managing memory requires a trade-off to be reached between adaptable units and stable systems (Chen and Heydari, 2024). Issues related to training on convergence, exploration-exploitation trading off and runtime overhead must be addressed. Moreover, hybrid approaches, which combine heuristics algorithms and RL driven changes, are often suggested to improve feasibility and keep performance consistent (Zhang *et al.* 2022).

Table 2: Thematic Analysis

(Source: self-made)

Overall, the results show that reinforcement learning is a promising but complex way to approach intelligent page replacement that promises to not only be adaptive and provide better performance but also introduces new challenges in the design and implementation of the algorithm.

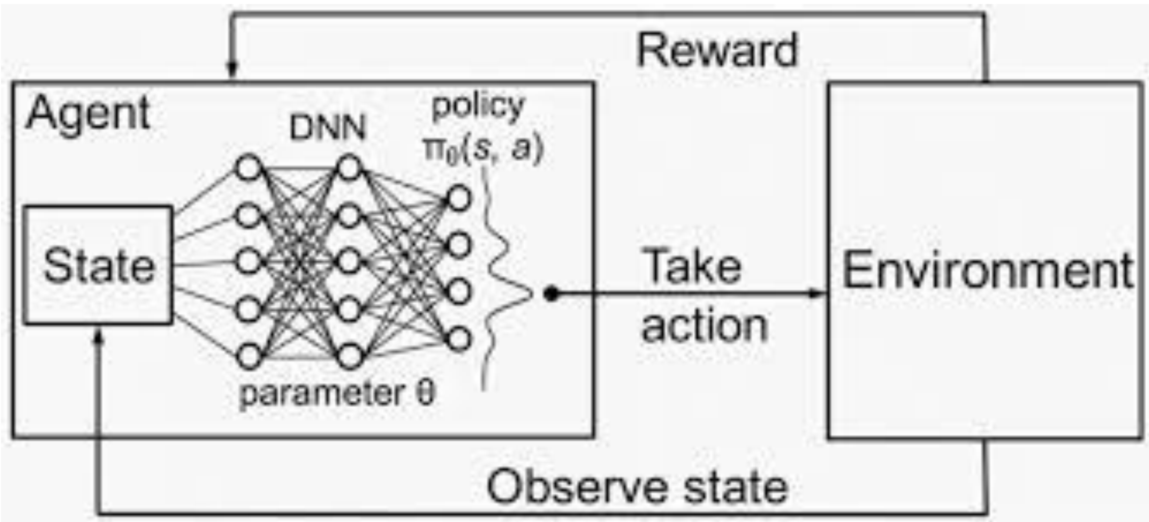

Figure 3.1: Deep reinforcement learning

(Source: As inspired by Wang *et ai,* 2022)

3.2 Analysis and Interpretation

The reinforcement learning approach of intelligent page replacement was analyzed based on the discussion of comparative performance studies and simulation-based studies. The interpretation of secondary data was done through the comparison of reinforcement learning models and the traditional algorithms Least Recently Used, First-In-First-Out, and the theoretical model Optimal Page Replacement. The important parameters of evaluation were page fault rate, hit rate of the cache, latency of the memory access and the system throughput.

According to the interpretation of findings, reinforcement learning redefines page replacement as a decision-making problem that occurs sequentially instead of a heuristic task at a specific point (Cui *et al.* 2024). Through exposure to memory access patterns as well as reward feedback to reduce page faults, the learning agent eventually evolves its replacement policy. This adaptive measure allows the system to react well to the varying workload properties such as non-stationary and bursty memory accesses, which tend to be challenging to conventional algorithms.

In several workload traces, the average page fault rates and better cache efficiency of approaches that used reinforcement learning were lower than the heuristic-based approaches. In contrast to other algorithms which are not dynamic and only utilize recency or arrival order, reinforcement learning is able to measure the long-term implications of the action performed on the replacement (Wang *et al.* 2023). Nonetheless, the analysis further indicates that the performance increase is also affected by the design of state representation and reward engineering. Increasing the state spaces can also lead to an increase in computational complexity and can have an influence on the speed of convergence. Although these limitations exist, the general view is that reinforcement learning offers a scalable and adaptive architecture to current memory management systems especially in dynamically computing systems where workload variability restraints the validity of traditional replacement approaches.

4. Discussion

Adaptive Learning and Dynamic Decision-Making

The initial theme that has been observed in the thematic analysis is adaptive learning and dynamic decision-making. The page replacement algorithms that were traditionally used, including Least Recently Used and the First-In-First-Out use a fixed set of heuristic rules that do not change during execution (Moerland *et al.* 2023). These strategies make assumptions of predictable access patterns of the memory and rely heavily on the locality of time. Though perfect in predictable situations, they cannot cope with workload characteristics that vary.

The reinforcement learning presents a dynamic decision-making machine whereby the page replacement process is governed by the interaction between an agent and the environment whose sequence is dynamic. The agent can learn optimal replacement policies using methods like Q-learning where the agent is provided with reward signals in the form of reduced page faults and an enhanced use of the cache. In the long run, the model changes its policy to optimize the cumulative rewards over time, as opposed to short-term ones (Souza and Freitas, 2024). This adaptability facility helps the system to react effectively to changing patterns of access, bursty loads and heterogeneous computing requirements. With the increased use of cloud platforms and distributed systems in modern computing environments, continuous learning and adaptation becomes a very important benefit compared to fixed heuristic approaches.

Performance Improvement in Terms of Page Fault Reduction

The second theme revolves around better performance, especially on the reduction of the page fault rates. Page faults are very harmful to the system efficiency since they initiate expensive disk input and output processes. The theoretical ideal of reducing page faults is the Optimal Page Replacement which is supposed to know the future memory references. Reinforcement learning gets close to optimal behavior without the need to know something in the future (Kim and Yeom, 2022). The learning agent improves its strategy by continually reconsidering the costs of replacement decisions in the long run to minimize the number of page faults in the long run. Simulation studies have shown empirically that policies based on reinforcement learning have a higher hit ratio in cache, and lower average page fault rate than traditional algorithms. The long-term reward optimization system enables the system to trade-off between the instantaneous replacement requirements with future performance (Ni *et al.* 2023). Consequently, reinforcement learning boosts throughput, latency and efficiency in using memory in dynamic workloads.

State-Space Complexity and Computational Overhead

The third theme deals with the complexity of state-space and overhead. Although reinforcement learning is flexible and offers better performance, it presents serious modeling problems. The management of memory incorporates a myriad of variables; the page references, frequency of access, recency data, and system context (Liu *et al.* 2024). When these factors are represented in a structured state space the results may be high-dimensional models.

As the state space increases in size, learning convergence is computationally intensive. Programs of advanced reinforcement learning, particularly with the use of neural networks, demand more processing and consume more memory. This can be a performance-degrading factor when the decision latency would be too high. The operating systems used in real-time require and demand fast and deterministic responses and efficiency is a big issue. Thus, effective reinforcement learning should be performed with meticulous state abstraction, reward engineering as well as optimization of learning algorithms to maintain that the computational overhead is manageable.

Implementation Feasibility and System Integration Challenges

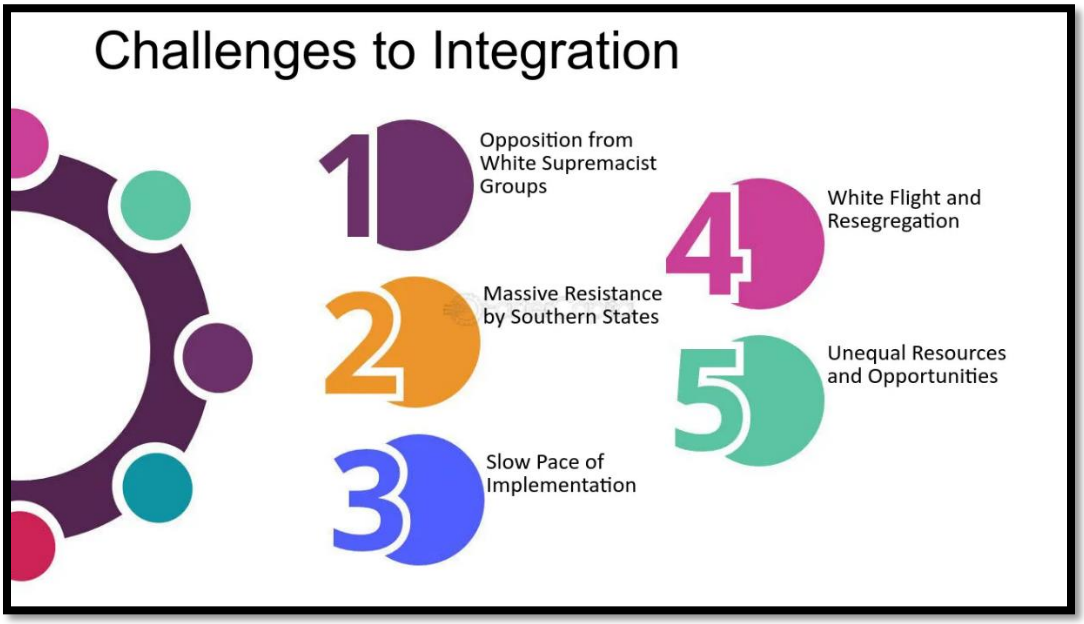

Figure 3.2: Integration Challenges

(Source: as inspired by Wang *et al.* 2022)

The fourth theme emphasizes feasibility of implementation and integration difficulties of systems. This is because traditional page replacement policies are entrenched in the operating system kernel because they are easy to understand and can be executed deterministically. These architectures need major system-level adjustments to incorporate reinforcement learning mechanisms. Difficulties in the designing of reward functions that are in line with system performance goals, stability during exploration periods, and performance at different workloads are some of the challenges (Wang *et al.* 2022). Also, the large-scale environment (when it comes to virtualised or cloud infrastructures) requires the use of scalable learning frameworks that can handle distributed memory systems. The hybrid methods between heuristic baseline strategies and adaptive reinforcement learning elements could present a viable trade-off between intelligence and stability. Despite the reinforcement learning showing a significant promise of an intelligent page replacement, the effectiveness of its implementation is dependent on effective architecture, computation, and integration considerations (Chen and Heydari, 2024). Hence, all the four themes prove that reinforcement learning is an effective means of improving adaptability and performance in memory management systems and at the same time adds complexity and implementation hurdles that must be well handled.

5. Conclusion

The current research involved the reinforcement learning methodology of intelligent page replacement in the contemporary memory management systems. It becomes clear after using thematic analysis that reinforcement learning provides a paradigmatic change of the fixed, rule-based replacement policies to adaptive, learning-oriented decision-making models. In contrast to classical algorithms like Least Recently Used and First-In-First-Out, reinforcement-based learning allows the constant-time policy improvement through the interaction with the ever-changing memory access dynamics. Such flexibility makes reinforcement learning an attractive alternative to complicated and unpredictable computing conditions.

The results show that page replacement strategies, which are based on reinforcement learning, are effective to minimize the page fault rates and improve the percentage of hits in the cache memory. Learning agents by modeling memory management as a sequential decision-making problem will learn near-optimal behavior like Optimal Page Replacement without the need to have prior knowledge of the future memory references. This capability greatly promotes system throughput and resource usage especially in cloud computing, distributed architectures and large-scale virtualized systems.

Nonetheless, there are also practical difficulties related to the application of reinforcement learning mentioned in the study. There are large state spaces, computational costs, the reward function structure, and the implementation within the existing operating system kernels, which pose significant technical challenges. The trade-offs between flexibility and computational performance should be well addressed to provide real-time responsiveness and stability. Although reinforcement learning has advantages in terms of long-term optimization, to implement it in practice, lightweight models are needed, and the state representation should be optimized, and frameworks should be based on balancing between intelligence and operational feasibility. In general, reinforcement learning is a proactive model of intelligent page replacement. It is in line with the changing demands of modern computing infrastructures, which need to have adaptive and scale resources management mechanisms. Even though the thematic results require additional empirical validation and system-level implementation experiments, the thematic results affirm the fact that the reinforcement learning is highly likely to transform memory management strategies in the next-generation operating systems.

5.1 Summary of Key Findings

The paper shows that reinforcement learning contributes to adaptability in page replacement by facilitating the dynamic decision-making process based on the cumulative reward maximization. It continues to enhance performance by minimizing page faults as well as enhancing efficiency of the cache under heavy workload conditions. The strategy estimates the optimal replacement behavior without making unrealistic assumptions that reference to the future. Meanwhile, the computational complexity and integration issues are considered as important considerations in practice.

5.2 Limitations of the Study

The study based on the analysis of secondary data and the synthesis of themes rather than the implementation of the primary experiments. The differences in the simulation environment, workload traces and reinforcement learning architectures restrict the direct comparability of results. There is no actual kernel-level deployment, limiting the inferences about the possibilities of operation in a production system. Besides, variations in the hardware settings and the system limitations can affect the performance results.

5.3 Recommendations for Future Research

Future research should focus on lightweight reinforcement learning models that can be used in real time operating systems. With more focus should be put on effective state abstraction methods and powerful reward functions that are congruent with system goals. Generalizability would be enhanced by empirical validation by practicing the implementation in cloud and edge computing environments. The study of hybrid methodology would be useful to present a balanced and scalable solution to smart page replacement in the current computing systems by further investigating hybrid methods that incorporate heuristic stability with adaptive learning mechanisms.

Reference List

1. Al-Hamadani, M.N., Fadhel, M.A., Alzubaidi, L. and Harangi, B., 2024. Reinforcement learning algorithms and applications in healthcare and robotics: a comprehensive and systematic review. *Sensors*, *24*(8), p.2461. https://www.mdpi.com/1424-8220/24/8/2461
2. Ali, I.M., 2024. A guide for positivist research paradigm: From philosophy to methodology. *Idealogy Journal*, *9*(2). http://www.idealogyjournal.com/ojs/index.php/idealogy/article/download/596/380
3. Chen, Q. and Heydari, B., 2024. The SoS conductor: Orchestrating resources with iterative agent-based reinforcement learning. *Systems Engineering*, *27*(4), pp.715-727. https://incose.onlinelibrary.wiley.com/doi/pdfdirect/10.1002/sys.21747
4. Cui, T., Yang, X., Jia, F., Jin, J., Ye, Y. and Bai, R., 2024. Mobile robot sequential decision making using a deep reinforcement learning hyper-heuristic approach. *Expert Systems with Applications*, *257*, p.124959. https://www.sciencedirect.com/science/article/pii/S0957417424018268
5. Elias, M., Kaplan, H., Mansour, Y. and Moran, S., 2024. Learning-augmented algorithms with explicit predictors. *Advances in Neural Information Processing Systems*, *37*, pp.97972-98008. https://proceedings.neurips.cc/paper_files/paper/2024/file/b1a72c79ac7512df8d7b573f38143ac4-Paper-Conference.pdf
6. Kim, H. and Yeom, H.Y., 2022, November. LPR: learning-based page replacement scheme for scientific applications. In *Proceedings of the 23rd International Middleware Conference Industrial Track* (pp. 36-42). https://dl.acm.org/doi/pdf/10.1145/3564695.3564777
7. Kleinstreuer, N. and Hartung, T., 2024. Artificial intelligence (AI)—it's the end of the tox as we know it (and I feel fine). *Archives of toxicology*, *98*(3), pp.735-754. https://link.springer.com/content/pdf/10.1007/s00204-023-03666-2.pdf
8. Lan, Q., Pan, Y., Luo, J. and Mahmood, A.R., 2022. Memory-efficient reinforcement learning with value-based knowledge consolidation. *arXiv preprint arXiv:2205.10868*. https://arxiv.org/pdf/2205.10868
9. Liu, D., Ren, F., Yan, J., Su, G., Gu, W. and Kato, S., 2024. Scaling up multi-agent reinforcement learning: An extensive survey on scalability issues. *IEEE Access, 12*, pp.94610-94631. https://ieeexplore.ieee.org/iel8/6287639/6514899/10550936.pdf
10. M. Moerland, T., Broekens, J., Plaat, A. and M. Jonker, C., 2023. Model-based reinforcement learning: A survey. *Foundations and Trends in Machine Learning, 16*(1), pp.1-118. https://www.emerald.com/ftmal/article-pdf/16/1/1/11135021/2200000086en.pdf
11. Naeem, M., Ozuem, W., Howell, K. and Ranfagni, S., 2023. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *International journal of qualitative methods*, *22*, p.16094069231205789. https://journals.sagepub.com/doi/pdf/10.1177/16094069231205789
12. Ni, T., Ma, M., Eysenbach, B. and Bacon, P.L., 2023. When do transformers shine in rl? decoupling memory from credit assignment. *Advances in Neural Information Processing Systems, 36*, pp.50429-50452. https://proceedings.neurips.cc/paper_files/paper/2023/file/9dc5accb1e4f4a9798eae145f2e4869b-Paper-Conference.pdf
13. Saunders, M. and Tosey, P., 2012. Research design. *Academia*. https://www.academia.edu/download/32495174/Research_30_-_Research_Onion_Layers_58.pdf
14. Shafizadeh, H., 2024. Decision-Making Under Uncertainty: How Organizations Adapt to Environmental Changes?. *Journal of Resource Management and Decision Engineering*, *3*(1), pp.4-10. http://journalrmde.com/index.php/jrmde/article/download/39/33
15. Shi, W., Guo, S., Cong, X., Sheng, W., Yan, J. and Chen, J., 2023. Frequency agile anti-interference technology based on reinforcement learning using long short-term memory and multi-layer historical information observation. *Remote Sensing*, *15*(23), p.5467. https://www.mdpi.com/2072-4292/15/23/5467
16. Souza, M.A. and Freitas, H.C., 2024. Reinforcement learning-based cache replacement policies for multicore processors. *IEEE Access*, *12*, pp.79177-79188. https://ieeexplore.ieee.org/iel8/6287639/6514899/10547236.pdf
17. Thyagaturu, A.S., Shantharama, P., Nasrallah, A. and Reisslein, M., 2022. Operating systems and hypervisors for network functions: A survey of enabling technologies and research studies. *IEEe Access*, *10*, pp.79825-79873. https://ieeexplore.ieee.org/iel7/6287639/6514899/09844719.pdf

18. Tripathi, S. and v Chaturvedi, R., 2023. Exploring ethical considerations in research: Guidelines and practices. *Adhyayan: A Journal of Management Sciences,* *13*(01), pp.41-45. https://www.smsjournals.com/index.php/Adhyayan/article/download/3147/1486
19. Wang, W., Li, B., Luo, X., and Wang, X. (2023). Deep reinforcement learning for sequential targeting. Management Science, 69(9), pp.5439-5460. https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=3487145
20. Wang, X., Wang, S., Liang, X., Zhao, D., Huang, J., Xu, X., Dai, B. and Miao, Q., 2022. Deep reinforcement learning: A survey. *IEEE Transactions on Neural Networks and Learning Systems, 35*(4), pp.5064-5078. https://ieeexplore.ieee.org/abstract/document/9904958/
21. Wu, M., Isaacman, S. and Bhattacharjee, A., 2024, September. Characterizing emerging page replacement policies for memory-intensive applications. In *2024 IEEE International Symposium on Workload Characterization (IISWC)* (pp. 284-294). IEEE. https://par.nsf.gov/servlets/purl/10549440
22. Zhang, Y., Bai, R., Qu, R., Tu, C. and Jin, J., 2022. A deep reinforcement learning based hyper-heuristic for combinatorial optimisation with uncertainties. *European Journal of Operational Research, 300*(2), pp.418-427. https://people.cs.nott.ac.uk/pszrq/files/EJOR21-drl-hh.pdf