1. **Introduction 1**

**1.1 Contract Bridge 2**

1.1.1 Contract 3

1.1.2 Bidding 4

1.1.3 Card-Play 6

1.1.4 Objective 7

1.1.5 Par Value 8

**1.2 Project Scope 10**

1.2.1 Motivation 10

1.2.2 Level of Autonomy 12

1. **Theoretical Principles 13**

**2.1 Problem Abstraction 13**

2.1.1 Extensive Form Game 13

2.1.2 Markov Decision Process 14

2.1.3 Best Policy 16

2.1.4 Nash Equilibrium 18

**2.2 Reinforcement Learning Concepts 21**

2.2.1 Exploration and Exploitation 21

2.2.2 Delayed Reward 22

2.2.3 Curriculum Learning 23

**2.3 Machine Learning Algorithm 25**

2.3.1 Deep Q-Learning 26

2.3.2 Monte Carlo Reinforcement Learning 28

2.3.3 Counterfactual Regret Minimization 32

1. **Analysis and Algorithm 39**

**3.1 Analysis 39**

3.1.1 Extensive Form Game 39

3.1.2 Markov Decision Process 40

**3.2 Algorithm 41**

3.2.1 Random Walk 41

3.2.2 First Approach (Regret Minimization) 43

3.2.3 Second Approach (MCRL) 47

1. **Results 50**
2. **Architectural Design 53**

**5.1 Information Representation 53**

**5.2 System Architecture 56**

5.2.1 Deal Generator 57

5.2.2 Knowledge Base 57

5.2.3 Game Master 59

5.2.4 Agent 60

1. **Discussion 61**

**6.1 Difficulties 61**

6.1.1 State Space 61

6.1.2 Correlation 62

**6.2 Evaluation 66**

6.2.1 Correct Move 66

6.2.2 Data Efficiency 67

6.2.3 Resources 67

**6.3 Reward 69**

6.3.1 Cooperation 69

6.3.2 Communication 69

**6.4 Suggestion 72**

1. **Conclusion 75**

**References 76**

**Appendices 78**

1. Zhang, H. and Xu, S. (2016). The Face Recognition Algorithms Based on Weighted LTP. *Journal of Image and Graphics*, 4(1), pp.11-14.
2. Sayed, M. (2018). Biometric Gait Recognition Based on Machine Learning Algorithms. Journal of Computer Science, 14(7), pp.1064-1073.
3. S., S. and Kamade, P. (2011). Video OCR for Video Indexing. *International Journal of Engineering and Technology*, 3(3), pp.287-289.
4. Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K. and Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, 362(6419), pp.1140-1144.
5. Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T. and Hassabis, D. (2017). Mastering the game of Go without human knowledge. Nature, 550(7676), pp.354-359.
6. Minh, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M. (2013). Playing Atari with Deep Reinforcement Learning. arXiv:1312.5602v1
7. Hart, S. (n.d.). *Games in extensive and strategic forms*. [S.l.]: [s.n.], pp.20-40.
8. SUTTON, R. (2017). *REINFORCEMENT LEARNING*. 2nd ed. London, England: MIT Press.
9. Mirsoleimani, S., Plaat, A., van den Herik, J. and Vermaseren, J. (2017). An Analysis of Virtual Loss in Parallel MCTS. *Proceedings of the 9th International Conference on Agents and Artificial Intelligence*.
10. Zinkevich, M., Johanson. M., Bowling, M., Piccione, C. (2007). Regret Minimization in Games with Incomplete Information. Proceedings of the 20th Advances in Neural Information Processing System
11. Bengio, Y., Louradour, J., Collobert, R., Weston, J. Curriculum Training. (2009) International Conference on Machine Learning
12. Blackwell, D. (1956). An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1), pp.1-8.
13. John N Tsitsiklis and Benjamin Van Roy. An analysis of temporal-difference learning with function approximation. Automatic Control, IEEE Transactions on , 42(5):674–690, 1997.
14. Long-Ji Lin. Reinforcement learning for robots using neural networks. Technical report, DTIC Document, 1993.
15. Martin Riedmiller. Neural fitted q iteration–first experiences with a data efficient neural reinforcement learning method. In Machine Learning: ECML 2005 , pages 317–328. Springer, 2005.
16. En.wikipedia.org. (2019). *Monte Carlo tree search*. [online] Available at: https://en.wikipedia.org/wiki/Monte\_Carlo\_tree\_search [Accessed 28 Apr. 2019].
17. Rosin, C. D. Multi-armed bandits with episode context. Annals of Mathematics and Artificial Intelligence 61, 203–230 (2011).
18. Brown, N. and Sandholm, T. (2017). Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science*, 359(6374), pp.418-424.
19. fledsu (2019). *关于德州扑克AI中Counterfactual Regret Minimization的介绍*. [online] 知乎. Available at: https://zhuanlan.zhihu.com/p/30438383 [Accessed 29 Apr. 2019].
20. Pi.math.cornell.edu. (2019). *IMPs and Matchpoints*. [online] Available at: http://pi.math.cornell.edu/~belk/impmp.htm [Accessed 29 Apr. 2019].
21. Haglund, B. and Hein, S. (2019). *dds-bridge/dds*. [online] GitHub. Available at: https://github.com/dds-bridge/dds [Accessed 29 Apr. 2019].
22. Hemminga, F. (2019). *atolab/pydds*. [online] GitHub. Available at: https://github.com/atolab/pydds [Accessed 29 Apr. 2019].
23. TensorFlow. (2019). *TensorFlow*. [online] Available at: https://www.tensorflow.org [Accessed 29 Apr. 2019].
24. Keras.io. (2019). *Home - Keras Documentation*. [online] Available at: https://keras.io/ [Accessed 29 Apr. 2019].
25. Geitgey, A. (2019). *Machine Learning is Fun Part 6: How to do Speech Recognition with Deep Learning*. [online] Medium. Available at: https://medium.com/@ageitgey/machine-learning-is-fun-part-6-how-to-do-speech-recognition-with-deep-learning-28293c162f7a [Accessed 29 Apr. 2019].
26. David, E., Netanyahu, N. and Wolf, L. (2017). End-to-End Deep Neural Network for Automatic Learning in Chess. *International Conference on Artificial Neural Networks (ICANN)*, [online] 9887, pp.88-96. Available at: https://arxiv.org/abs/1711.09667.
27. Sądel, B. and Śnieżyński, B. (2017). Online Supervised Learning Approach for Machine Scheduling. *Schedae Informaticae*, 1/2016.
28. Heinrich, J. (2017). Reinforcement Learning from Self-Play in Imperfect-Information Games.