

## Bab 5

# Simpulan dan Saran

Bab 5 membahas kesimpulan dari Tugas Akhir dengan judul "Studi Implementasi Metode *Multi-Agent Reinforcement Learning* Untuk Optimasi Luas Cakupan Jaringan Sensor" serta saran yang dapat dikembangkan untuk kedepannya. Kesimpulan diambil dari data simulasi yang dilakukan pada Bab 4.

### 5.1 Simpulan

Pada penelitian Tugas Akhir ini, telah dilakukan perancangan algoritma *single agent reinforcement learning* dan *multi-agent reinforcement learning* untuk menemukan luas daerah *coverage* optimal berdasarkan posisi awal agen berada. Aspek yang harus diperhatikan agen dalam melakukan pembelajaran seperti mempertimbangkan daya yang digunakan untuk mencapai koordinat dengan *coverage* optimal dan gerakan yang diambil agen agar tidak bergerak terlalu jauh maupun terlalu dekat dengan agen lainnya sudah terjawab menggunakan metode pembelajaran MARL. Agen dapat menghemat daya dengan menekan jumlah *steps* yang digunakan serta meminimalkan pergantian arah gerak. Agen juga bergerak tidak terlalu dekat ataupun terlalu jauh dengan agen lainnya.

Kedua agen menunjukkan kemajuannya dalam pembelajaran. Pada SARL, agen dapat memilih koordinat dengan *coverage* optimal terbaik berdasarkan faktor-faktor di atas. Pembelajaran yang dilakukan juga sudah bergerak menuju optimal. Hal tersebut menunjukkan bahwa agen memang belajar dan bukan hanya mengandalkan metode pemilihan aksi secara acak. Pada MARL, agen sudah berhasil bernegosiasi dan menghindari terjadinya *overlap* antara kedua agen. Beberapa agen merubah koordinat terbaik menurut agen tersebut demi mendapatkan hasil yang lebih baik bagi kedua agen. Hal tersebut membuktikan bahwa negosiasi yang dilakukan agen dapat berjalan dengan baik.

## 5.2 Saran

Beberapa saran yang dapat dilakukan untuk mengembangkan algoritma *multi-agent reinforcement learning* yang dilakukan adalah:

1. Meningkatkan kompleksitas *cell* yang tadinya berbentuk kubus menjadi *hexagon tessellation* sehingga kemungkinan aksi yang dilakukan agen menjadi lebih kompleks.
2. Meningkatkan algoritma negosiasi untuk mendapatkan nilai yang lebih tinggi apa bila kedua agen tidak menemukan *nash equilibrium*.
3. Implementasi pada *quadrotor* dan pembelajaran yang dapat dilakukan secara *online* pada *quadrotor*.
4. Melonggarkan aturan-aturan pada agen, seperti membiarkan agen membentuk *overlap* apabila luas *coverage* yang dihasilkan kedua agen lebih luas dibandingkan jika tidak ada *overlap*.
5. Merancang algoritma agar agen dapat mengetahui luas optimal pada suatu lingkungan.
6. Merancang algoritma pembelajaran yang membuat agen mampu untuk menyelesaikan tugasnya tanpa harus menunggu agen lainnya.
7. Mengoptimalkan algoritma, nilai *Q-value*, dan *joint Q-value* untuk menekan *computational cost* yang dibutuhkan.

# Daftar Pustaka

- [1] A. Ahmed, M. Nagai, C. Tianen, and R. Shibasaki, "Uav based monitoring system and object detection technique development for a disaster area," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, pp. 373–377, 2008.
- [2] S. D'Oleire-Oltmanns, I. Marzloff, K. D. Peter, and J. B. Ries, "Unmanned aerial vehicle (uav) for monitoring soil erosion in morocco," *Remote Sensing*, vol. 4, no. 11, pp. 3390–3416, 2012.
- [3] P. Gong, X.-B. Tang, X. Huang, P. Wang, L.-S. Wen, X.-X. Zhu, and C. Zhou, "Locating lost radioactive sources using a uav radiation monitoring system," *Applied radiation and isotopes*, vol. 150, pp. 1–13, 2019.
- [4] A. T. Oyelami, A. S. Akinade, and K. C. Obianefo, "Development of a real-time framework for farm monitoring using drone technology," *IAES International Journal of Robotics and Automation*, vol. 9, no. 4, p. 244, 2020.
- [5] D. Zorbas, L. D. P. Pugliese, T. Razafindralambo, and F. Guerriero, "Optimal drone placement and cost-efficient target coverage," *Journal of Network and Computer Applications*, vol. 75, pp. 16–31, 2016.
- [6] U. Ozdemir, Y. O. Aktas, A. Vuruskan, Y. Dereli, A. F. Tarhan, K. Demirbag, A. Erdem, G. D. Kalaycioglu, I. Ozkol, and G. Inalhan, "Design of a commercial hybrid vtol uav system," *Journal of Intelligent & Robotic Systems*, vol. 74, no. 1, pp. 371–393, 2014.
- [7] A. A. Mian and W. Daobo, "Modeling and backstepping-based nonlinear control strategy for a 6 dof quadrotor helicopter," *Chinese Journal of Aeronautics*, vol. 21, no. 3, pp. 261–268, 2008.
- [8] A. Breitenmoser, M. Schwager, J.-C. Metzger, R. Siegwart, and D. Rus, "Voronoi coverage of non-convex environments with a group of networked robots," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, 2010, pp. 4982–4989.
- [9] S. S. Ge and Y. J. Cui, "Dynamic motion planning for mobile robots using potential field method," *Autonomous robots*, vol. 13, no. 3, pp. 207–222, 2002.

- [10] H. M. La, “Multi-robot swarm for cooperative scalar field mapping,” in *Robotic Systems: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2020, pp. 208–223.
- [11] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, “Coverage control for mobile sensing networks,” *IEEE Transactions on robotics and Automation*, vol. 20, no. 2, pp. 243–255, 2004.
- [12] H. X. Pham, H. M. La, D. Feil-Seifer, and A. Nefian, “Cooperative and distributed reinforcement learning of drones for field coverage,” *arXiv preprint arXiv:1803.07250*, 2018.
- [13] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” *Science*, vol. 349, no. 6245, pp. 255–260, 2015.
- [14] Z. Zhao and H. Liu, “Spectral feature selection for supervised and unsupervised learning,” in *Proceedings of the 24th international conference on Machine learning*, 2007, pp. 1151–1157.
- [15] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [16] J. Hu and M. P. Wellman, “Nash q-learning for general-sum stochastic games,” *Journal of machine learning research*, vol. 4, no. Nov, pp. 1039–1069, 2003.
- [17] A. Geramifard, T. J. Walsh, S. Tellex, G. Chowdhary, N. Roy, and J. P. How, “A tutorial on linear function approximators for dynamic programming and reinforcement learning,” *Foundations and Trends® in Machine Learning*, vol. 6, no. 4, pp. 375–451, 2013.
- [18] E. Alpaydin, *Introduction to Machine Learning*, 3rd ed., ser. Adaptive Computation and Machine Learning. Cambridge, MA: MIT Press, 2014.
- [19] O. Simeone, “A very brief introduction to machine learning with applications to communication systems,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 648–664, 2018.
- [20] L. Zhou, P. Yang, C. Chen, and Y. Gao, “Multiagent reinforcement learning with sparse interactions by negotiation and knowledge transfer,” *IEEE transactions on cybernetics*, vol. 47, no. 5, pp. 1238–1250, 2016.
- [21] M. A. Khan, H. Tembine, and A. V. Vasilakos, “Game dynamics and cost of learning in heterogeneous 4g networks,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 198–213, 2011.
- [22] F. S. Melo and M. Veloso, “Learning of coordination: Exploiting sparse interactions in multiagent systems,” in *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. Citeseer, 2009, pp. 773–780.

- [23] Y.-M. De Hauwere, P. Vrancx, and A. Nowé, “Learning multi-agent state space representations,” in *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, 2010, pp. 715–722.
- [24] P. Vrancx, Y.-M. De Hauwere, and A. Nowé, “Transfer learning for multi-agent coordination.” in *ICAART (2)*. Rome, Italy, 2011, pp. 263–272.
- [25] C. Claus and C. Boutilier, “The dynamics of reinforcement learning in cooperative multiagent systems,” *AAAI/IAAI*, vol. 1998, no. 746-752, p. 2, 1998.
- [26] T. Sandholm, “Perspectives on multiagent learning,” *Artificial Intelligence*, vol. 171, no. 7, pp. 382–391, 2007.
- [27] M. L. Littman, “Markov games as a framework for multi-agent reinforcement learning,” in *Machine learning proceedings 1994*. Elsevier, 1994, pp. 157–163.
- [28] J. Hu, M. P. Wellman *et al.*, “Multiagent reinforcement learning: theoretical framework and an algorithm.” in *ICML*, vol. 98. Citeseer, 1998, pp. 242–250.
- [29] M. L. Littman *et al.*, “Friend-or-foe q-learning in general-sum games,” in *ICML*, vol. 1, 2001, pp. 322–328.
- [30] A. Greenwald, K. Hall, and R. Serrano, “Correlated q-learning,” in *ICML*, vol. 3, 2003, pp. 242–249.
- [31] D. Fudenberg and J. Tirole, *Game theory*. MIT press, 1991.
- [32] A. Rapoport, “Prisoner’s dilemma,” in *Game theory*. Springer, 1989, pp. 199–204.
- [33] J. F. Nash Jr, “Equilibrium points in n-person games,” *Proceedings of the national academy of sciences*, vol. 36, no. 1, pp. 48–49, 1950.
- [34] Y. Hu, Y. Gao, and B. An, “Multiagent reinforcement learning with unshared value functions,” *IEEE transactions on cybernetics*, vol. 45, no. 4, pp. 647–662, 2014.