# Informal Review of Multi-Agent Reinforcement Learning Literature

Sean Carver @ Data Machines Corporation

May 7, 2021

## 1 May 6, 2021

### 1.1 Positioning COMMS and JAMMERS to Defeat Each Other

Quite by accident, I discovered a body of literature that I deem very relevant to my COMMS and JAMMERS problem.

This work has an intersection with Multi-Agent Systems in that all the drones are autonomous systems moving to achieve some goal under partial information. It seems that the authors I have read so far do not specifically use reinforcement learning in their solution but their solution begs to be extended to an RL framework. (That said, see [1] for a non-adversarial MARL-example.)

Also, as far as I know, so far, no one [I am not certain of this, yet], specifically talks about the problem I am interested in with COMMS and JAMMERS—at least not in the way I have framed it—but the relevance should be clear.

The problem they discuss is how to control a sensor network (in my case, move the UAVs) in order to bring in the most information from the sensors. They use information-theoretic criteria as their objective functions. Beyond information theoretic criteria there are other objectives that do similar things.

I'll list 2 scantly cited papers [2, 3]. One more cited but less recent paper is [4]. There are certainly more papers available on Google Scholar, but I haven't yet dug into the literature yet. The first paper (2007) was found in [5], a bound collection of papers and not available otherwise. The second is a more recent IEEE paper (2018) that is not available for free, and I have only seen the abstract—but looks even more relevant. There are dozens more, but I would have to read more than the titles to determine relevance, which I will do, of course, in time.

The choice facing me now is to go after *depth* to solve the COMMS/JAMMERS problem or *breadth* with to understand the full scope of the MARL field—or some combination. Actually, maybe it is two separate projects, so I will await direction from stakeholders.

## 2 May 7, 2021

Reading (2010) "Particle filter based information-theoretic active sensing," [6]. Concerns controlling robot(s) where the control objective is to gather sensory information ("information gain"). This objective is described as minimizing the entropy of an estimate distribution—the probability distribution for the target location. Their example is target tracking by a UAV with a camera mounted on a fixed-wing autonomous aircraft and a moving target on the ground.

This (2010) paper [6] cites two more papers for background on the methods they develop: (2004) [7] and (2009) [8]. The first (2004) [7] is a review paper by leaders of the field (that I recognized) that seems to just give an explanation needed to understand the fundamentals of the particle filter. The second (2009) [8] also looks very helpful for putting all this together. The (2010) paper [6] seems to be a direct extention of the (2009) paper [8]—the difference being that the (2010) paper does a receding horizon control (RHC) to make it possible to model fixed-wing aircraft that take a long time to turn as well as mounted cameras that they claim lead to non-minimum phase behavior. The paper [6] uses a single aircraft as their example (though they say their method is more general) whereas [8] uses a swarm of maneuverable (in one time step) UAVs.

The (2009) [8] paper cites [7] as a reference for Receeding Horizon Control (RHC).

**All three of these papers look like papers I would want to read closely.**

# References


[1] Enrico Testi, Elia Favarelli, and Andrea Giorgetti. Reinforcement learning for connected autonomous vehicle localization via UAVs. In *2020 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor)*, pages 13–17. IEEE, 2020.

[2] David Nicholson, Sarvapali D Ramchurn, and Alex Rogers. Information-based control of decentralised sensor networks. In *Defense Industry Applications of Autonomous Agents and Multi-Agent Systems*, pages 15–32. Springer, 2007.

[3] Anna Guerra, Nicola Sparnacci, Davide Dardari, and Petar M Djurić. Collaborative target-localization and information-based control in networks of UAVs. In *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE, 2018.

[4] Ben Grocholsky, Alexei Makarenko, Tobias Kaupp, and Hugh F Durrant-Whyte. Scalable control of decentralised sensor platforms. In *Information Processing in Sensor Networks*, pages 96–112. Springer, 2003.

[5] Michal Pechoucek, Simon G Thompson, and Holger Voos. *Defense Industry Applications of Autonomous Agents and Multi-Agent Systems.* Springer, 2008.

[6] Allison Ryan and J Karl Hedrick. Particle filter based information-theoretic active sensing. *Robotics and Autonomous Systems*, 58(5):574–584, 2010.

[7] Christophe Andrieu, Arnaud Doucet, Sumeetpal S Singh, and Vladislav B Tadic. Particle methods for change detection, system identification, and control. *Proceedings of the IEEE*, 92(3):423–438, 2004.

[8] Gabriel M Hoffmann and Claire J Tomlin. Mobile sensor network control using mutual information methods and particle filters. *IEEE Transactions on Automatic Control*, 55(1):32–47, 2009.