1. The article is headlined “Counterfactual multi-agent policy gradients”

2. The article is written by students from University of Oxford. It was published in the journal “Artificial Intelligence ” in 2017.

3. The main idea of the article is to create multi-agent reinforcement learning method that is able to solve many real-world problems, such as network packet routing or the coordination of autonomous vehicles. The authors present counterfactual multi-agent policy gradient method to address these ones.

4. The authors start by telling that many complex real-world problems are naturally modeled as cooperative multi-agent systems. However, currently designed single-agent reinforcement learning algorithms are poor on such tasks. Therefore, the article describes MARL algorithm that is efficient and solves some valuable problems like centralized learning, decentralized execution and credit assignment between agents. The authors write that centralized learning allows to learn common critic that is able to estimate value of environment’s state and can be used by agents. Thus, it is possible to learn such agents to be independent what solves decentralized execution problem. Further the authors say that there remains the problem where cooperative agents should share global reward between each other. According to the text, the authors use difference rewards mechanism to overcome this problem where the agent’s value is estimated between selected action and some default action, and if the estimate is positive it means that agent’s action is more valuable and it should get more rewards. In conclusion, the authors compare their approach with other algorithms and report that the approach works 20% better.

5. I find the article important because it contains state-of-the-art algorithm is able to learn independent agents that action cooperatively.