Learning Group-Level Information Integration in Multi-Agent Communication

Extended Abstract

Xiangrui Meng
Peking University, School of Intelligence Science and Technology
Beijing, China
mxxxr@stu.pku.edu.cn

Ying Tan
Peking University, School of Intelligence Science and Technology & Key Laboratory of Machine Perceptron (MOE) & Institute for Artificial Intellignce
Beijing, China
ytan@pku.edu.cn

ABSTRACT

In multi-agent systems, it's hard to make proper decisions for agents due to the partial observability of the environment. Among categories of multi-agent reinforcement learning (MARL) algorithms, communication learning is a common approach to solving this problem. However, existing work focus on individual-level communication which usually leads to significant communication costs. Meanwhile, the group feature couldn't be well captured at the individual level. To tackle these problems, this paper proposes a group-level information integration model called Double Channel Communication Network (DC2Net). In DC2Net, individual and group features are learned in two independent channels. Agents no longer interact with each other at the individual level and all information interaction is carried out in the group channel. This model ensures effective learning of group features while reducing individual-level communication costs. Empirically, we conducted experiments on several environments and tasks. The experimental results show that the DC2Net not only has a better performance compared to other state-of-the-art MARL communication models but also reduces the costs of communication. Furthermore, it's a natural communication topology with the ability in balancing individual and communication learning.

KEYWORDS

Multi-Agent Communication; Multi-Agent Reinforcement Learning; Deep Reinforcement Learning

ACM Reference Format:

Xiangrui Meng and Ying Tan. 2023. Learning Group-Level Information Integration in Multi-Agent Communication: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023,* IFAAMAS, 3 pages.

1 INTRODUCTION

Multi-agent reinforcement learning (MARL) [1] communication is a branch of MARL in which individual-level communication is often used [2, 4, 6]. It means that there is communication between every two agents or some pairs of them. The cost is unbearable when there are a number of agents or a limited communication bandwidth [8]. How to select appropriate partners to communicate with is another challenge that individual-level communication is faced with [3, 7].

Considering the behavior of human cooperation to complete complex tasks, it is a common strategy to reach a group consensus based on personal knowledge [10]. Generally, forming a meaningful group consensus can help individuals to make better decisions [15]. Several work have studied ways to raise group-level information in MARL communication [7, 12, 13]. However, group information is often seen as an aid to individual learning in these work. Due to insufficient attention to group information, group learning and individual learning are often mixed [13], which makes the extraction of group features more difficult. In our opinion, group information is a key point in communication, it needs to be seen as a prominent signal, and independent feature extraction is required at the group level.

In this paper, we propose a communication model named Double Channel Communication Network (DC2Net) which consults individual and group learning in two independent channels. In DC2Net, there are no information interactions between agents at the individual level. Individual information is fed into a common channel for group information integration. This approach avoids the problem of mixed learning while reducing the costs of individual-level communication.

2 MODEL

2.1 Individual Channel

In our model, the individual channel is responsible for the independent learning of individual information. For each agent, the local observation o_i is encoded into a hidden state h_i by an encoder. After the encoding, several controllers are used to extract features of $h = \{h_1, h_2, ..., h_n\}$. Here we use fully connected neural networks as the controllers in the individual channel (FC-I):

$$h_i^{l+1} = f_{FC-I}^{l}(h_i^l) \tag{1}$$

where l indicates the lth layer of the network. Inspired by the skip connection in ResNet[5], we use skip connections to maintain and reuse the individual information in different layers.

$$h_i^{l+2} = f_{FC-I}^{l+1}(h_i^l, f_{FC-I}^{l}(h_i^l)). \tag{2}$$

2.2 Group Channel

The group channel is where individual information is integrated. In our opinion, the order of the individuals should not affect the output of the group channel. Here we use a group pooling operation

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

to integrate the individual hidden states h_i. To show the mechanism of the group pooling, we define a hidden state map H^l which stacks the hidden state h_i^l of each agent i in the lth layer.

$$H^l = concatenate(h_i^l), i = 1, 2, ..., N. \quad (3)$$

More specifically,

$$H^l = \begin{pmatrix} h_{11}^l & h_{12}^l & \cdots & h_{1D}^l \\ h_{21}^l & h_{22}^l & \cdots & h_{2D}^l \\ \vdots & \vdots & h_{id}^l & \vdots \\ h_{N1}^l & h_{N2}^l & \cdots & h_{ND}^l \end{pmatrix} \quad (4)$$

where h_{id}^l indicates the dth dimension of hidden state h_i^l. Here we use the max pooling which is widely used in image processing [9, 11] to obtain the maximum feature in each column of H^l:

$$\tilde{h}^l = \{\max_i h_{id}^l | d = 1, 2, ..., D\} \quad (5)$$

where $\tilde{h}^l$ is the group information gathered by the max pooling.

Since the group channel is learned independently under our premise, we make a gradient truncation of h_i^l. There are no gradients from individual learning in $\tilde{h}^l$ so the update of parameters in the group channel can be independent of the individual channel during the backpropagation. Like the learning process of individuals, we use several fully connected layers to extract features in the group channel (FC-G). Except for the first layer, the input of each FC-G consists of the pooled feature $\tilde{h}^l$ and the output $\tilde{h}^{l-1}$ from the previous layer:

$$\tilde{h}_i^l = f_{FC-G}^l(\tilde{h}_i^{l-1}, pooling(H^l)). \quad (6)$$

The output z of the total group channel will be fed into a decoder with the output h_i^L of the individual channel. The decoder can be an MLP that outputs the Q value, and we use the Q-learning algorithm to train our model.

$$q_i = Decoder(h_i^L + z) \quad (7)$$

3 EXPERIMENT

Traffic Junction Results We evaluate DC2Net on different levels of Traffic Junction (TJ) [13] environment. The quantitative results of success rates can be found in Table 1.

In the easy level TJ, all of the four models have similar performance with success rates of more than 99%. This is because there is only one grid of junction at the easy level. Agents just need to learn to wait when there is a car at the intersection to complete the task. In the medium level TJ, DC2Net reaches 97.39% in success rate. In both easy and medium levels, the performance of CommNet is poor compared to the other three models. We think this is because instead of being effective, the group information generated by averaging operations could affect individual learning. Since these two tasks are not particularly difficult (only one junction), providing too much information in the individual learning process will make it more difficult to capture features. The comparable performance of IQL also proves that excessive communication is not required in these two environments. In the hard level TJ, DC2Net has absolute superiority. DC2NET achieves a winning rate of 97.14% compared with the other three methods reaching about 90%. The poor performance of IQL shows that this task requires communication to speed up learning, and the other two communication methods can obtain a faster convergence speed through communication. Nevertheless, DC2Net achieves overall superiority.

Table 1: Success rates on TJ.

Model	Easy	Medium	Hard
IQL [14]	99.02%	97.02%	89.88%
CommNet [13]	99.22%	92.86%	85.43%
DGN [6]	99.14 %	97.22%	90.45%
DC2Net (this paper)	**99.42%**	**97.39%**	**97.14%**

Does the DC2NET Communicate Less? Empirically, we set up a scenario on hard TJ with limited communication bandwidth (A limited amount of bits per episode during execution. The model can only make decisions based on individual information after the limited communication bandwidth has been exhausted). The experimental results of the three models in this scenario are shown in Fig. 1 (the similarly colored curves are the performance of the same model at full and limited bandwidths). It can be seen that DC2Net maintains favorable performance under the condition of limited bandwidth. However, the performance of the other two communication models has declined significantly under this condition.

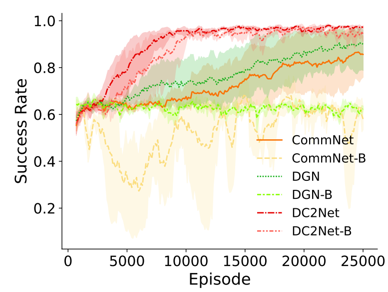

Figure 1: Experimental results on hard level TJ at fixed bandwidth.

4 CONCLUSION

In this paper, we proposed the DC2Net for multi-agent communication by learning in two independent channels which are responsible for individual learning and communication learning. The group-level communication not only effectively captures the group feature but also reduces the costs of communication. Empirically, the DC2Net model outperforms several state-of-the-art MARL communication models on a variety of cooperative multi-agent tasks. The experiment also illustrates that DC2Net can balance individual learning and communication learning in different tasks.

For a broader impact, we demonstrate the effectiveness of independent learning in both individual and group channels for multi-agent communication. The group-level learning manner deserves to be studied more widely.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (Grant No. 62276008, 62250037, and 62076010), and partially supported by Science and Technology Innovation 2030 - "New Generation Artificial Intelligence" Major Project (Grant No. 2018AAA0102301 and 2018AAA0100302).

REFERENCES

[1] Lucian Busoniu, Robert Babuska, and Bart De Schutter. 2008. A comprehensive survey of multiagent reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 38, 2 (2008), 156–172.

[2] Abhishek Das, Théophile Gervet, Joshua Romoff, Dhruv Batra, Devi Parikh, Mike Rabbat, and Joelle Pineau. 2019. Tarmac: Targeted multi-agent communication. In *International Conference on Machine Learning*. PMLR, 1538–1546.

[3] Ziluo Ding, Tiejun Huang, and Zongqing Lu. 2020. Learning individually inferred communication for multi-agent cooperation. *Advances in Neural Information Processing Systems* 33 (2020), 22069–22079.

[4] Jakob N. Foerster, Yannis M. Assael, Nando de Freitas, and Shimon Whiteson. 2016. Learning to Communicate with Deep Multi-Agent Reinforcement Learning. *arXiv:1605.06676 [cs]* (May 2016). http://arxiv.org/abs/1605.06676 arXiv: 1605.06676.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[6] Jiechuan Jiang, Chen Dun, Tiejun Huang, and Zongqing Lu. 2020. Graph Convolutional Reinforcement Learning. *arXiv:1810.09202 [cs, stat]* (Feb. 2020). http://arxiv.org/abs/1810.09202 arXiv: 1810.09202.

[7] Jiechuan Jiang and Zongqing Lu. 2018. Learning Attentional Communication for Multi-Agent Cooperation. *arXiv:1805.07733 [cs]* (Nov. 2018). http://arxiv.org/abs/1805.07733 arXiv: 1805.07733.

[8] Daewoo Kim, Sangwoo Moon, David Hostallero, Wan Ju Kang, Taeyoung Lee, Kyunghwan Son, and Yung Yi. 2019. Learning to Schedule Communication in Multi-agent Reinforcement Learning. *arXiv:1902.01554 [cs]* (Feb. 2019). http://arxiv.org/abs/1902.01554 arXiv: 1902.01554.

[9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60 (2012), 84 – 90.

[10] David G Rand and Martin A Nowak. 2013. Human cooperation. *Trends in cognitive sciences* 17, 8 (2013), 413–425.

[11] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

[12] Amanpreet Singh, Tushar Jain, and Sainbayar Sukhbaatar. 2018. Learning when to communicate at scale in multiagent cooperative and competitive tasks. *arXiv preprint arXiv:1812.09755* (2018).

[13] Sainbayar Sukhbaatar, Arthur Szlam, and Rob Fergus. 2016. Learning Multiagent Communication with Backpropagation. *arXiv:1605.07736 [cs]* (Oct. 2016). http://arxiv.org/abs/1605.07736 arXiv: 1605.07736.

[14] Ardi Tampuu, Tambet Matiisen, Dorian Kodelja, Ilya Kuzovkin, Kristjan Korjus, Juhan Aru, Jaan Aru, and Raul Vicente. 2017. Multiagent cooperation and competition with deep reinforcement learning. *PLOS ONE* 12, 4 (April 2017), e0172395. https://doi.org/10.1371/journal.pone.0172395

[15] Changxi Zhu, Mehdi Dastani, and Shihan Wang. 2022. A survey of multi-agent reinforcement learning with communication. *arXiv preprint arXiv:2203.08975* (2022).