Universally Expressive Communication in Multi-Agent Reinforcement Learning

Matthew Morris, Thomas Barret, Arnu Pretorius

InstaDeep Ltd.

Graph Decision Networks (GDNs)

Allowing agents to share information through **communication** is crucial for solving complex tasks in multi-agent reinforcement learning (MARL). Many of the most successful methods can be captured within the following framework:

1. At each time step, define an attributed graph $G = (V, E, a)$ with nodes $V(G) := \{\text{all agents}\}$, edges $E(G) := \{(i, j) \mid \text{agent } i \text{ is communicating with } j\}$, and for all agents i, the node i is labeled with the observation of i.
2. Pass this graph through a GNN f which outputs values for each node
3. Pass each resulting node value through the actor network of the corresponding agent

We refer to communication methods that fall within this paradigm as **Graph Decision Networks (GDNs)**. If the actor networks used shared weights (very common), then any GDN reduces to a **Graph Neural Network (GNN)** node labelling problem.

Limited Expressivity

There is a plethora of research into the expressivity of GNNs. In particular, the expressivity of standard GNNs corresponds to that of the **1-WL algorithm**. These limitations are thus also inherited by GDNs. However, we can also use insights from GNNs to go beyond these expressivity limits.

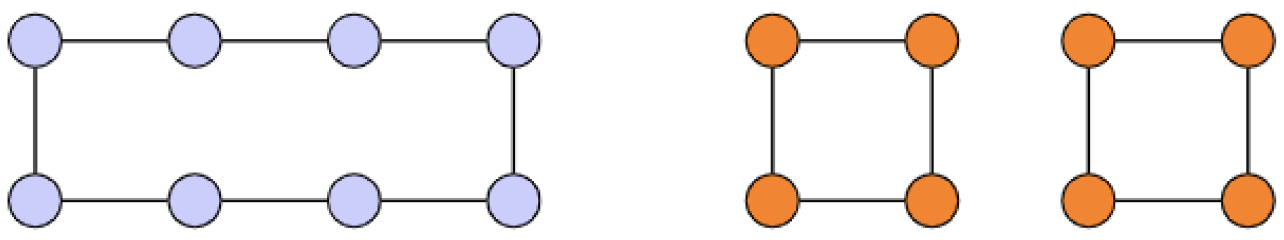

Figure 1: A pair of graphs indistinguishable by 1-WL

Furthermore, even GDNs which can capture any equivariant graph function are unable to perform **symmetry breaking**: i.e. they cannot solve problems which require similar agents that have identical observations to take different actions.

Universally Expressive Communication

Ideally, MARL communication methods should possess all the following properties:

1. Universal expressivity for equivariant graph functions
2. Symmetry breaking for coordination problems
3. Computational efficiency

We apply two existing GNN augmentations to GDNs to achieve this, both of which come with minimal extra computational cost. We first prove that standard GDNs are unable to produce different actions for identical agents with the same observations.

Theorem 1. *Given a GDN f, observations $O = \{o_1, ..., o_n\}$, and communication graph G such that nodes i and j are similar in G and $o_i = o_j$, then $f(O)_i = f(O)_j$.*

We prove that **concatenating noise** to the agent observations achieves universal approximation and the solving of symmetry-breaking coordination problems.

Theorem 2. *Let $n \geq 1$ and let $f : G_n \to \mathbb{R}^n$ be equivariant. Then for all $\epsilon, \delta > 0$, there is a GNN with RNI that (ϵ, δ)-approximates f.*

Theorem 3. *Let $n \geq 1$ and consider a set T, where each $(G, A) \in T$ is a graph-labels pair, such that $G \in G_n$ and there is a set of target labels $A_k \in A$ for each orbit $r_k \in R(G)$, with $|A_k| = |r_k|$. Then for all $\epsilon, \delta > 0$ there is a GNN with RNI g which satisfies:*

$$\forall (G, A) \in T \ \ \forall r_k \in R(G), \{g(G)_i \mid i \in r_k\} \cong_{\epsilon,\delta} A_k$$

Giving each agent a **unique ID** in its observations enables the same.

Theorem 4. *Let $n \geq 1$ and let $f : G_n \to \mathbb{R}^n$ be equivariant. Then for all $\epsilon > 0$, there is a GNN with unique node IDs that ϵ-approximates f.*

Theorem 5. *Let $n \geq 1$ and consider a set T, where each $(G, A) \in T$ is a graph-labels pair, such that $G \in G_n$ and there is a set of target labels $A_k \in A$ for each orbit $r_k \in R(G)$, with $|A_k| = |r_k|$. Then for all $\epsilon > 0$ there is a GNN with unique node IDs g which satisfies:*

$$\forall (G, A) \in T \ \ \forall r_k \in R(G), \{g(G)_i \mid i \in r_k\} \cong_{\epsilon} A_k$$

Evaluation

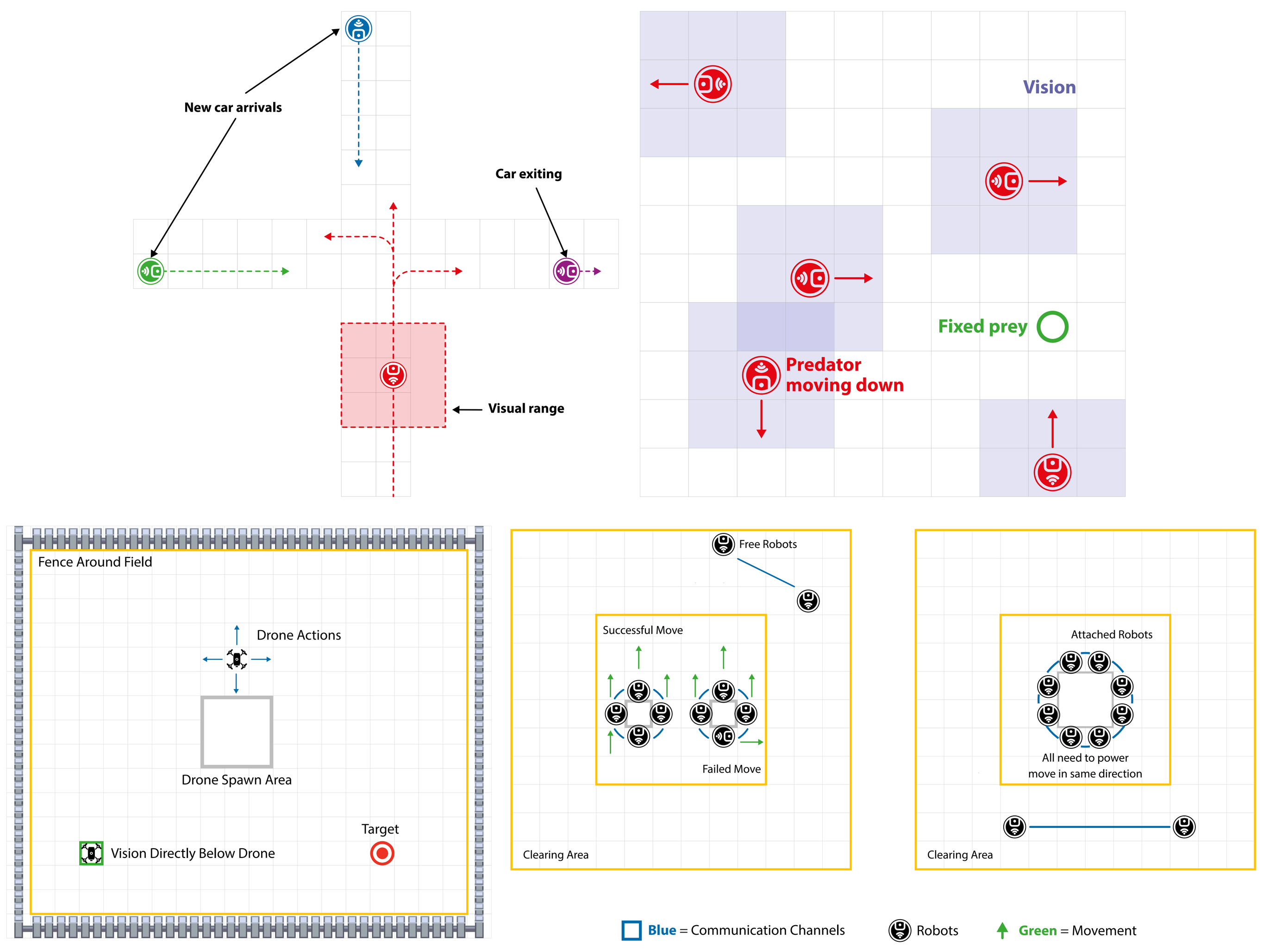

Figure 2: Environments: Traffic Junction (benchmark), Predator-Prey (benchmark), Drone Scatter (designed to test symmetry breaking), Box Pushing (designed to test expressivity beyond 1-WL)

We use 6 successful MARL communication algorithms for our experiments: CommNet, IC3Net, TARMAC, TARMAC-IC3Net, DGN, and MAGIC. We consider augmentations of each of them with random noise and unique node IDs in the observations.

Conclusions

- Unique IDs or smaller RNI augmentations can typically be applied without detriment on standard environments, but do not readily provide improved performance
- On environments where more complex coordination is required, these augmentations are essential for strong performance and both outperform the baselines
- RNI is best suited to environments requiring increased expressivity
- Unique IDs is best suited to environments requiring symmetry-breaking
- Future work: other insights into GNN architectures can be leveraged in GDNs

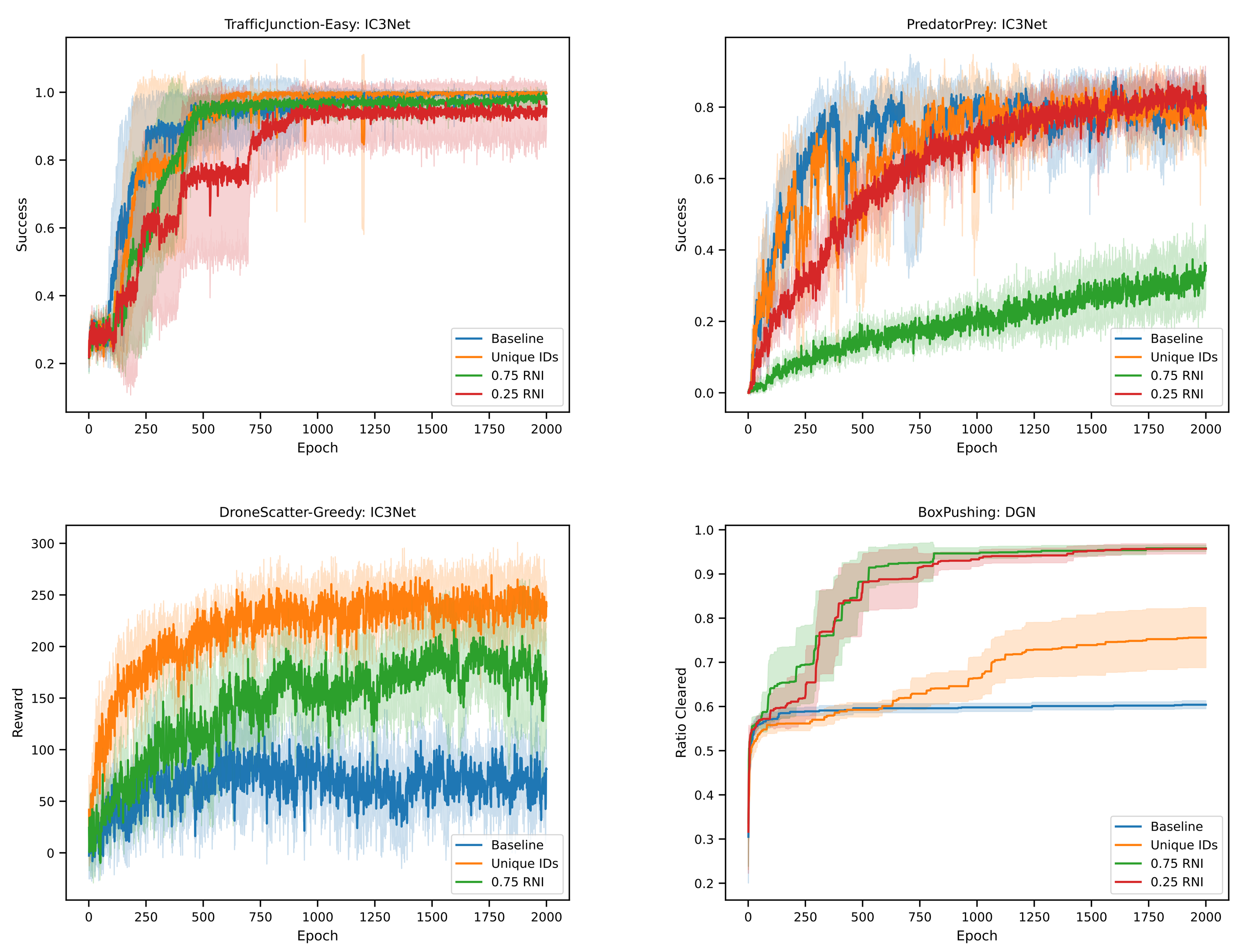

Figure 3: Some of the results across each environment

References

[1] Ralph Abboud et al. "The surprising power of graph neural networks with random node initialization". In: *IJCAI* (2021), pp. 2112–2118.

[2] George Dasoulas et al. "Coloring graph neural networks for node disambiguation". In: *IJCAI* (2020), pp. 2126–2132.

[3] Christopher Morris et al. "Weisfeiler and Leman go Machine Learning: The Story so far". In: *arXiv preprint arXiv:2112.09992* (2021).

[4] Yaru Niu, Rohan Paleja, and Matthew Gombolay. "Multi-Agent Graph-Attention Communication and Teaming". In: *AAMAS*. 2021, pp. 964–973.