

Motivation

Question de recherche.

Comment apprendre à communiquer ?

Travail précédent. Notre revue [1] généralise le processus de communication avec GNN et met en évidence ses défis :

- Générer le signal
- Sélectionner les destinataires
- Interpréter les signaux
- Intégrer les informations acquises

Contexte

MARL. Plusieurs agents apprennent à résoudre des tâches par interaction répétée avec l'environnement.

MARL avec communication. Apprendre à mieux se coordonner avec les autres et à prendre en compte les différents points de vue.

Communication basée GNN. Un GNN, invariant au nombre de voisins, agrège et interprète les messages, et capture de l'information structurelle locale et globale.

GNN dynamique. Un D-GNN permet d'apprendre à partir de la dynamique d'un graphe temporel, par exemple pour prédire des liens futurs.

GNN dynamique pour le MARL avec communication

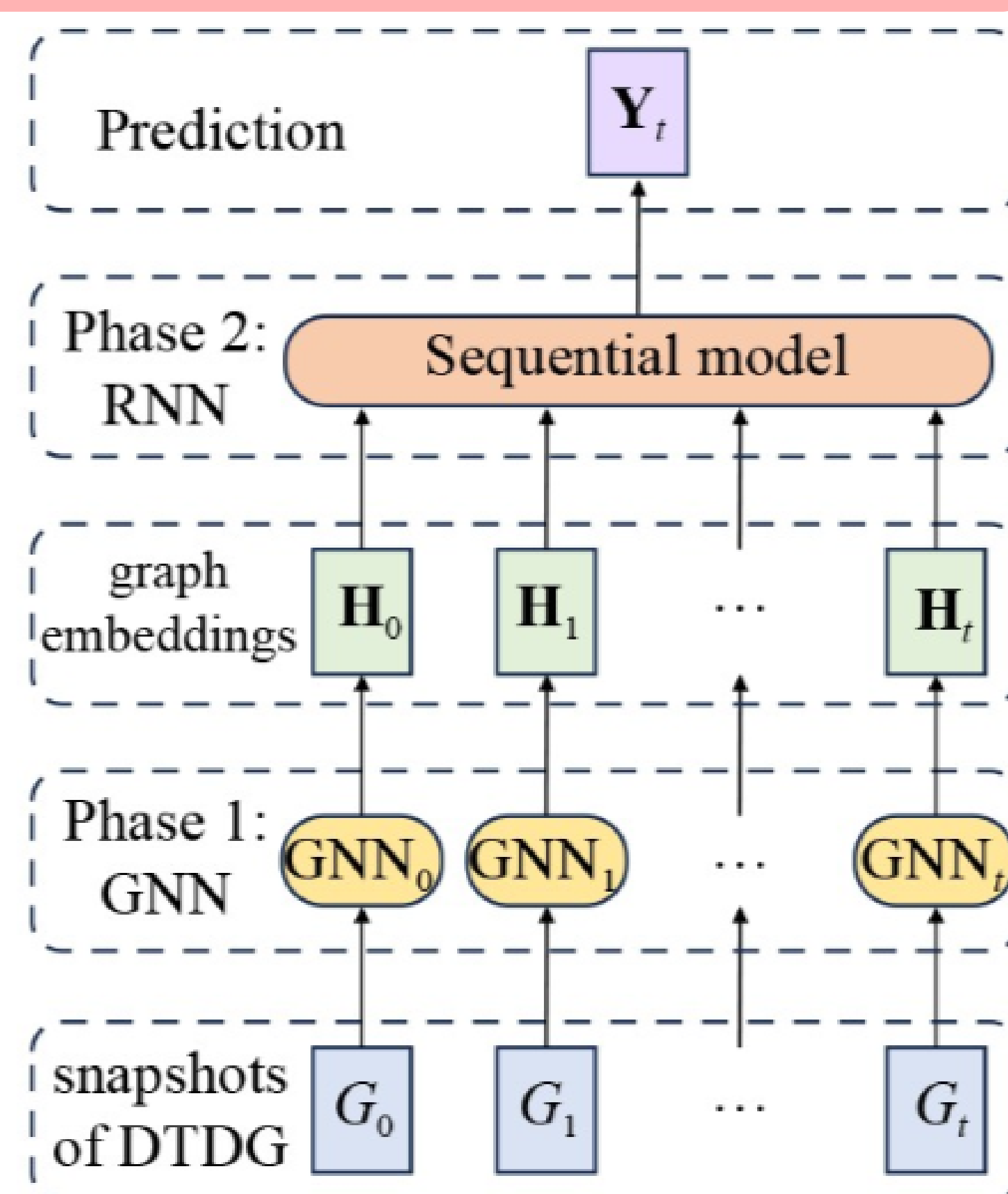
Idée défendue

Doter les agents d'une mémoire de leurs observations partielles permet souvent de meilleurs comportements. De la même façon, en considérant la dynamique du graphe de communication avec un D-GNN, les agents apprennent de meilleurs comportements, à partir de l'historique des représentations enrichi par la communication.

Approche. Évaluation du modèle D-GNN comme politique, entraîné à l'aide de l'algorithme MAPPO.

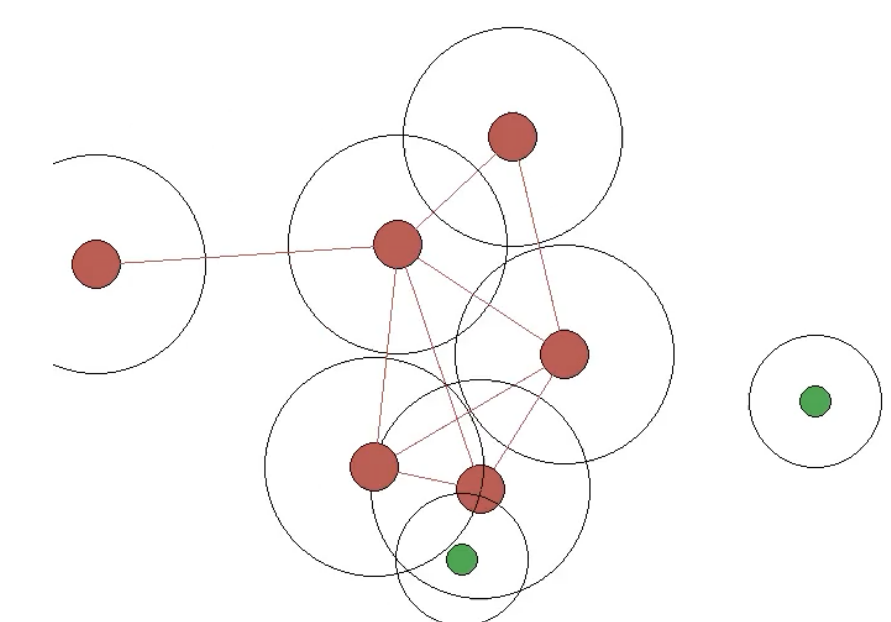
Tableau des modèles comparés.

Modèles	Description
MLP	Multi-layer Perceptron
RNN	Recurrent Neural Networks (mémorisation)
GNN	Graph Neural Networks (communication entre agents)
R-GNN	RNN sur les représentations locales, puis GNN pour la communication
D-GNN	GNN pour la communication puis RNN sur les représentations apprises par le GNN. Prends en compte la dynamique des comm.



Stacked dynamic GNN de Y Zheng [2]

Tâche proie-prédateur



Environnement. 6 agents prédateurs. 2 proies stationnaires. Portée de perception et de communication limitée.

Actions. {stay, up, down, left, right}

Observations. Position, vitesse et pour chaque agent observable :

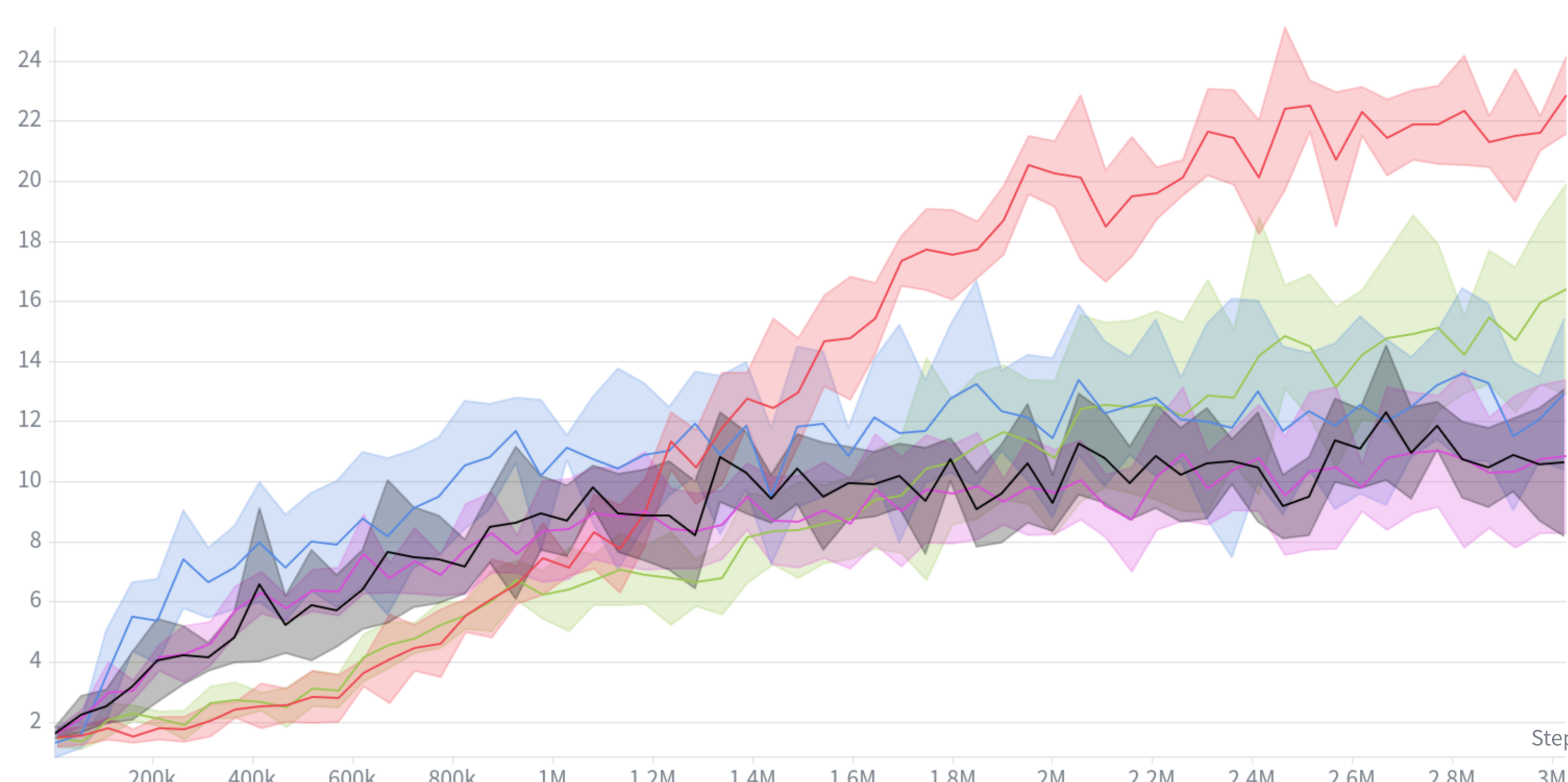
- type d'agent (prédateur/proie)
- position
- vitesse

Récompenses. Capture d'une proie si ≥ 2 agents dans le rayon de capture

Coopérative. Maximiser la somme des récompenses

Validation empirique

- MLP | - RNN | - GNN | - R-GNN | - D-GNN



Courbe d'apprentissage : évolution des retours moyens avec écart-type

Chaque méthode est testée sur 5 graines aléatoires différentes.

La communication par GNN permet aux agents de mieux se coordonner pour capturer les proies.

Dans notre cas, un RNN sur les observations n'est pas nécessaire.

Avec D-GNN, les agents mémorisent leurs communications passées, exploitant à la fois la structure et l'évolution du graphe de communication.

Le GNN dynamique permet aux agents de mieux se coordonner pour capturer les proies, malgré l'observabilité partielle.

Conclusion

Limitations.

- Étude limitée à une seule tâche
- Comparaison avec l'état de l'art

Travaux futurs.

- Tester la robustesse de D-GNN face aux contraintes de communication
- Transfert d'apprentissage

Références

- [1] V Cuzin-Rambaud, L Matignon, M Morge. "A Survey of Multi-Agent Deep Reinforcement Learning with Graph Neural Network-Based Communication". RJCIA@PFIA, 2026, pp.44-53.
- [2] Y Zheng et al. "A Survey of Dynamic Graph Neural Networks". Frontiers of Computer Science, vol. 19, no. 6, 2024.