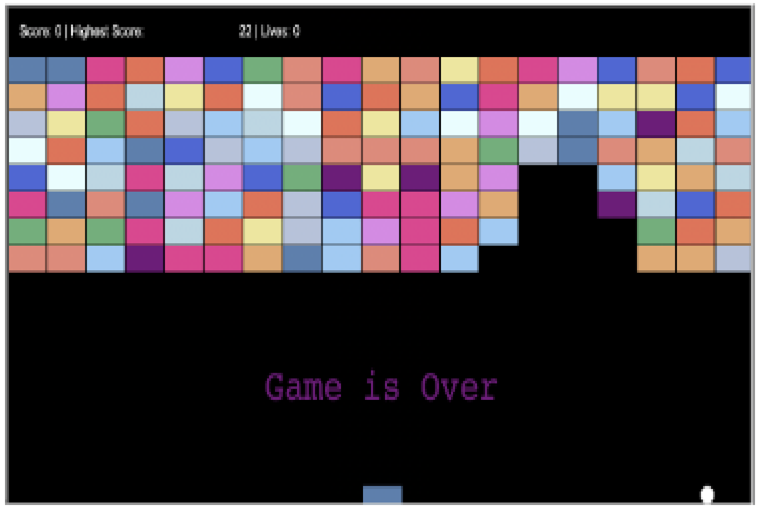
Learning Generalised Policies using Graph Neural Networks and Symbolic Representations

By Rishabh Bhatia

M.Sc Data Science

AI advancements

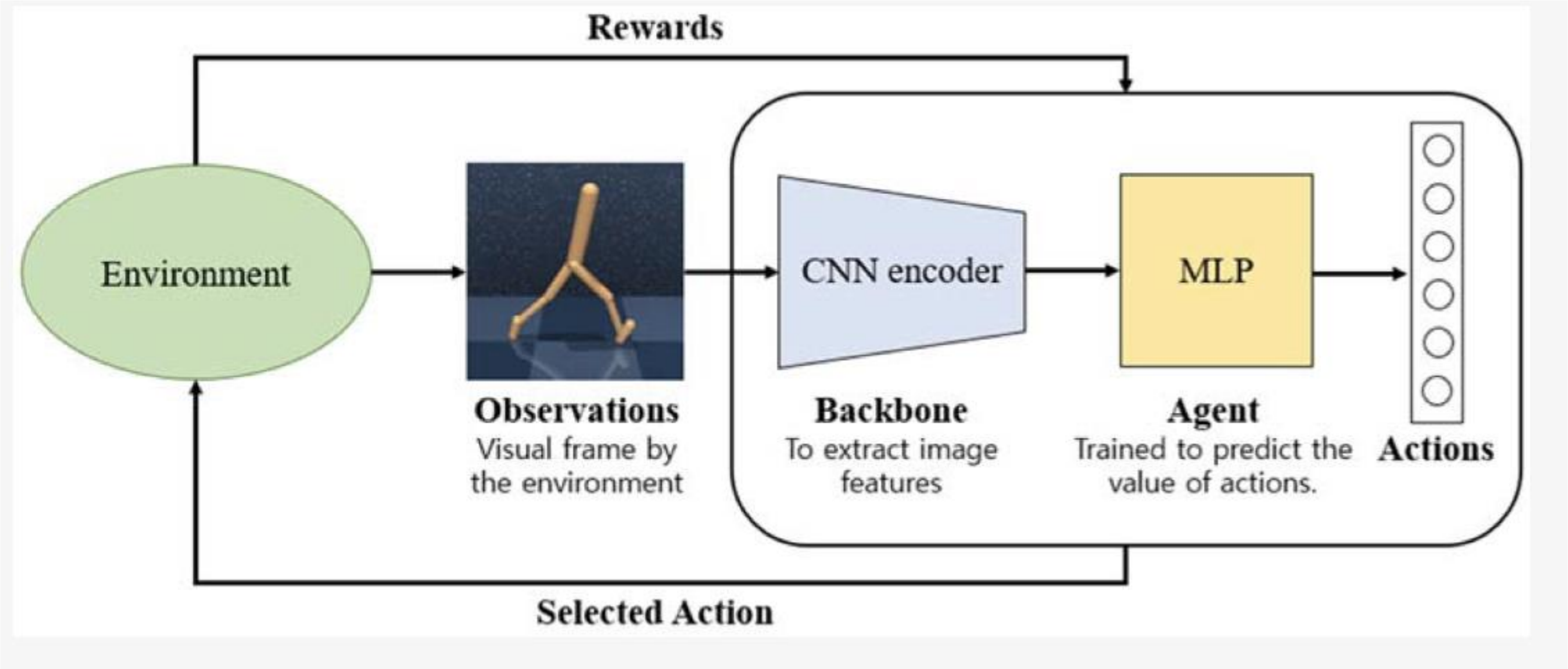
Generalisation



(a) Breakout Instance 1

(b) Breakout Instance 2

General Reinforcement Learning Pipeline

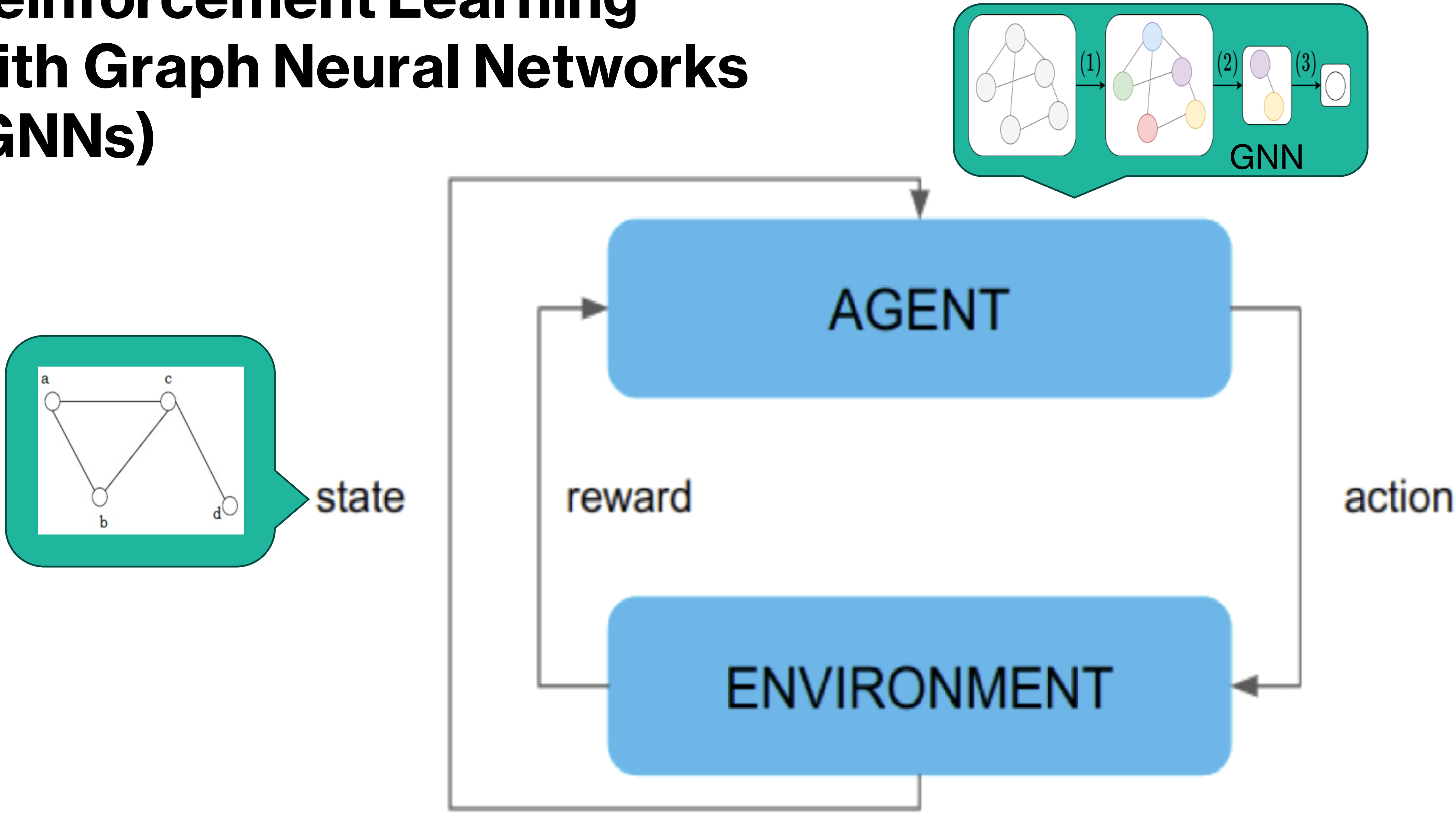


Source: Park et al, 2023

Limitations of Pixel Learning

- Depend on appearance of objects.
- Don't take into account relational dynamics.
- Current approaches don't perform well on out of distribution environments.

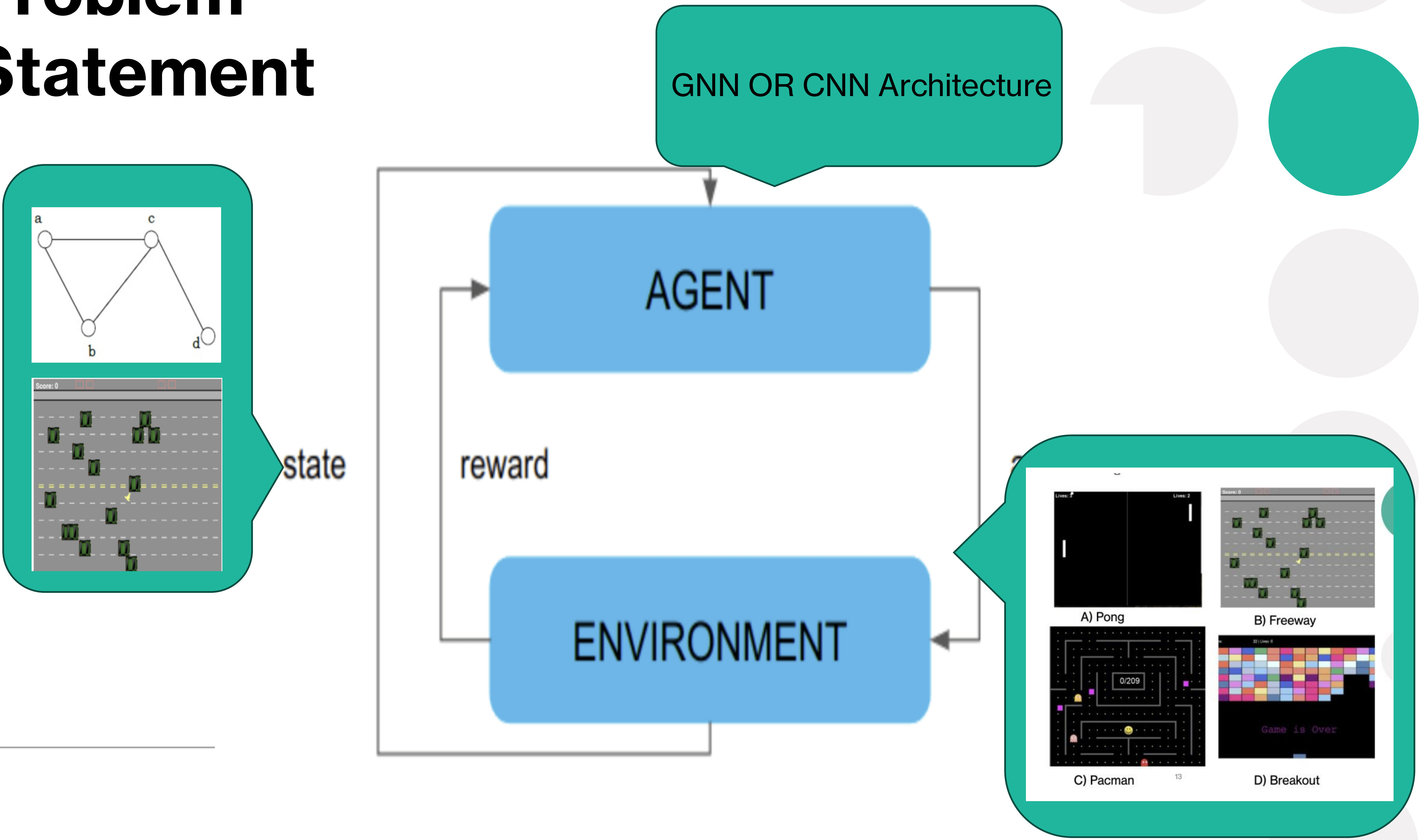
Reinforcement Learning with Graph Neural Networks (GNNs)



Motivation for Graph Representation

- Rich Representation.
- Relational reasoning.
- Transferability.

Problem Statement



Research Questions and Hypothesis

Research Questions

- **RQ1:** Performance and generalisation comparison of GNN-based vs. CNN-based agents in diverse game environments.
- **RQ2:** Effects of various graph-based state representations on GNNs' performance and generalisation in games.
- **RQ3:** Influence of different GNN architectures on AI performance and generalisation across game types.

Hypotheses

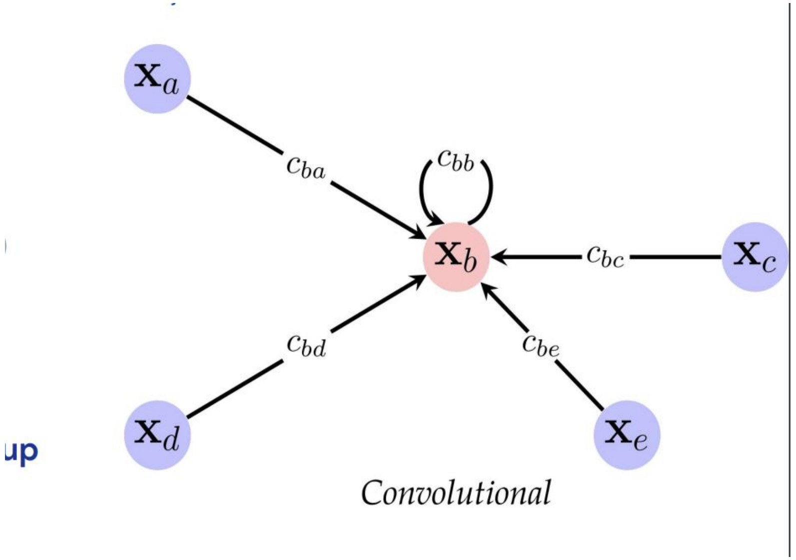
- **H1:** GNN-based agents will outperform CNN-based agents by effectively modeling relational data. The specific graph state representations and GNN architectures will further enhance performance and generalization.

GCN (Graph Convolutional Network)

- Usually, the weights directly depend on Adjacency Matrix.
  - ChebyNet (Defferrard et al.)
  - GCN (Kipf & Welling)
  - SGC (Wu et al.)

- Useful for homophilous graphs and scaling up
  - When edges encode label similarity

$$\mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j \in \mathcal{N}_i} c_{ij}\psi(\mathbf{x}_j)\right)$$



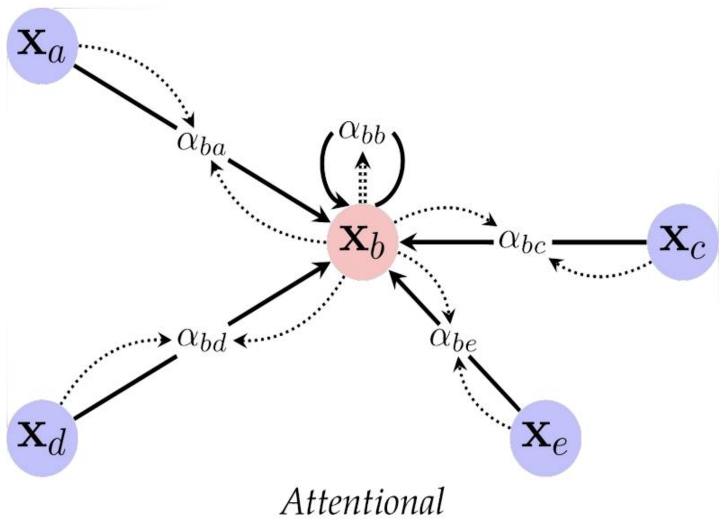
Source: Petar V, (2021) https://petar-v.com/talks/GNN-Wednesday.pdf

10

Graph Attention network (GAT)

- Features of neighbours aggregated with implicit weights (via attention)
- Attention weight computed as $\alpha_{ij} = a(x_i, x_j)$
  - MoNet (Monti et al.)
  - GAT (Veličković et al.)
  - GaAN (Zhang et al.)
- Useful as "middle ground" w.r.t. capacity and scale
- Edges need not encode homophily
- But still compute scalar value in each edge

$$\mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j \in \mathcal{N}_i} a(\mathbf{x}_i, \mathbf{x}_j)\psi(\mathbf{x}_j)\right)$$



Source: Petar V, (2021) https://petar-v.com/talks/GNN-Wednesday.pdf

Optimisation Algorithm

- Will use popular optimisation algorithm (PPO).
- Model Free, On-policy.
- Balances exploration and exploitation.

Custom Games

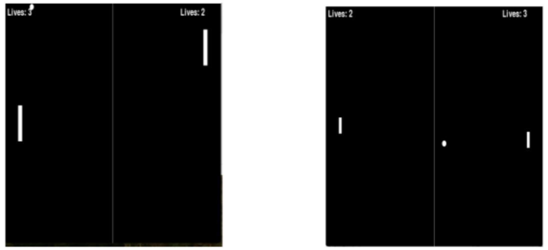
- Ability to extract symbolic representation easily.
- Can vary the properties for the game easily to get new flavours of games.

A.4 Breakout Game

(a) Breakout Instance 1

(b) Breakout Instance 2

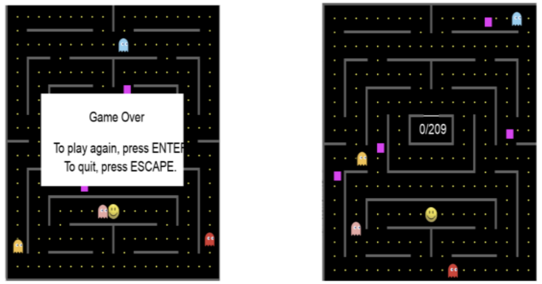
A.1 Pong Game



(a) Pong instance 1

(b) Pong instance 2

A.3 Pacman Game



(a) Pacman instance 1

(b) Pacman instance 2

Proximity-Based graph

- Edges for close objects.
- Defined by distance d.
- Removes irrelevant information and focuses on most important information.

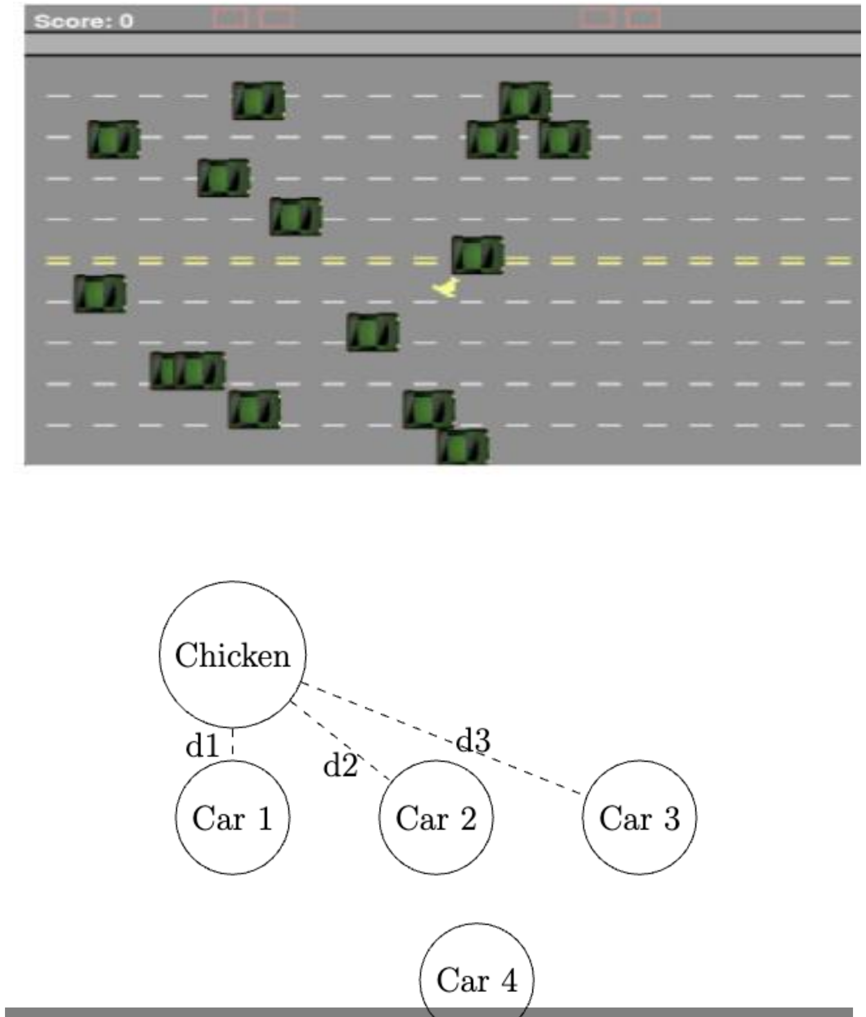
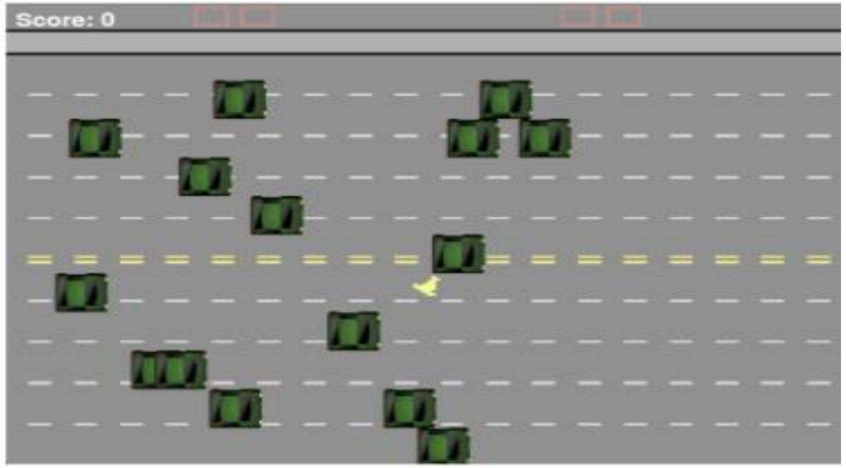


Fig. Example Graph for Freeway

Predicate based graph

- Graph where nodes represent objects and atoms true in a state.
- Edges represent predicates.
- Focuses on relevant interactions.
- Following Stahlberg et al. 2022.

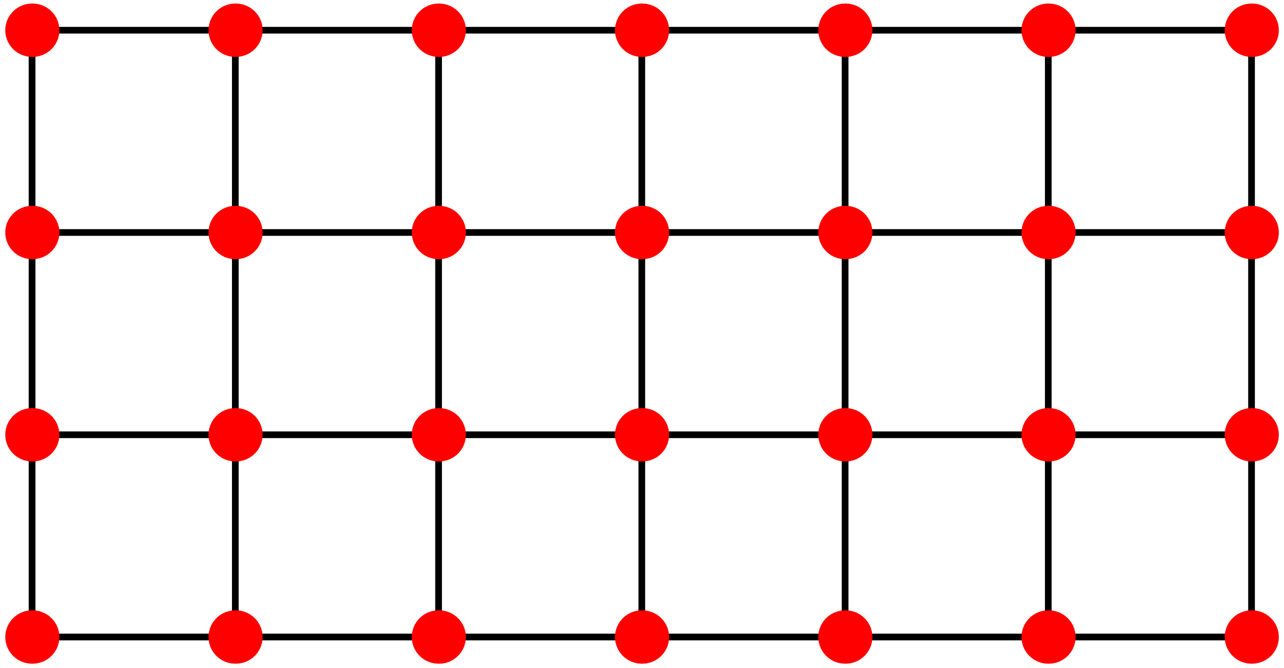


Predicates For Freeway

- CarOnLane(Car,Lane)
- ChickenOnLane(Car, Lane)
- LaneNextToLane(Lane1, Lane2)

Grid Based Graph

- Game Screen as grid.
- Nodes for grid cells.
- Edges connect adjacent cells.



Evaluation Procedure

Game	Parameter	Training Range	Testing Range
Pong	Ball Speed (units/tick)	5-10	8-15
	Paddle Size (% of height)	20-30	10-20
	Scoring Rules (points)	5-10	15-20
Freeway	Number of Lanes	3-4	5-6
	Vehicle Speed (units/tick)	4-7	6-10
	Cars per Lane	2-3	4-5
	Chicken Speed (units/tick)	3	2-4
Pacman	Maze Layouts	3 from 4	Include 4th
	Number of Ghosts	3-4	5-6
	Ghost Speed (% of Pac-Man's)	90-110	100-120
	Power-up Placements	3-5 per maze	2-8 per maze
Breakout	Brick Arrangement Patterns	2 from 3	All 3
	Ball Speed (units/tick)	4-6	3-10
	Paddle Size (% of width)	20-25	15-30
Shooting Game	Projectile Speed (units/tick)	6-9	5-11

Evaluation Metrics

Metrics Overview:

- Average Reward per Episode: Evaluates agent's decision-making effectiveness over time .
- Average Game Score: Directly measures agent's success in achieving game objectives .

Objective:

- Compare these metrics across CNN and GNN architectures across training and test sets to identify strengths in learning and generalisation in game strategies.

Research Questions and Hypothesis

Research Questions

- **RQ1:** Performance and generalisation comparison of GNN-based vs. CNN-based agents in diverse game environments.
- **RQ2:** Effects of various graph-based state representations on GNNs' performance and generalisation in games.
- **RQ3:** Influence of different GNN architectures on AI performance and generalisation across game types.

Hypotheses

- **H1:** GNN-based agents will outperform CNN-based agents by effectively modeling relational data. The specific graph state representations and GNN architectures will further enhance performance and generalisation.