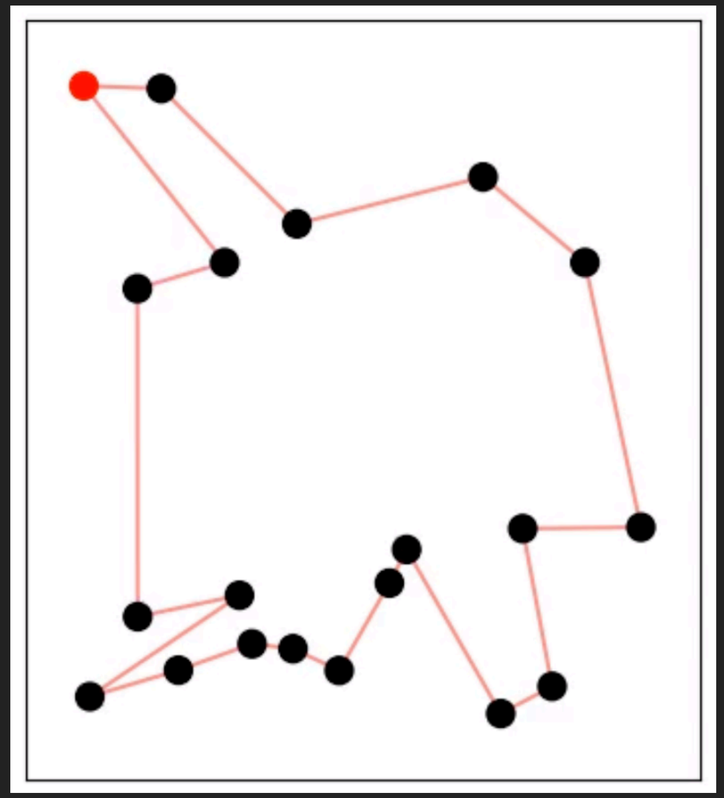
Leibniz Universität Hannover

RL FINAL PROJECT 2022

VRP-GYM

MOTIVATION

- Logistics plays a huge role in globalization
- Reduce CO2 emission with intelligent routing
- Active research area
- NP-Hard problem



OUR PROJECT

- Implemented an easy to extent and well documented VRP gym environment
- Three variants and a state of the art attention based agent[1]
  - TSP: Agent needs to traverse all nodes
  - VRP: Agent traverses all nodes and has to return to depot
  - IRP: Agent needs to traverse all nodes and has to deliver demand

[1]https://arxiv.org/pdf/1803.08475.pdf

ENVIRONMENT

- Observation Space

$$OBS_{TSP,VRP} \in \mathbb{R}^{b \times n \times 4}$$

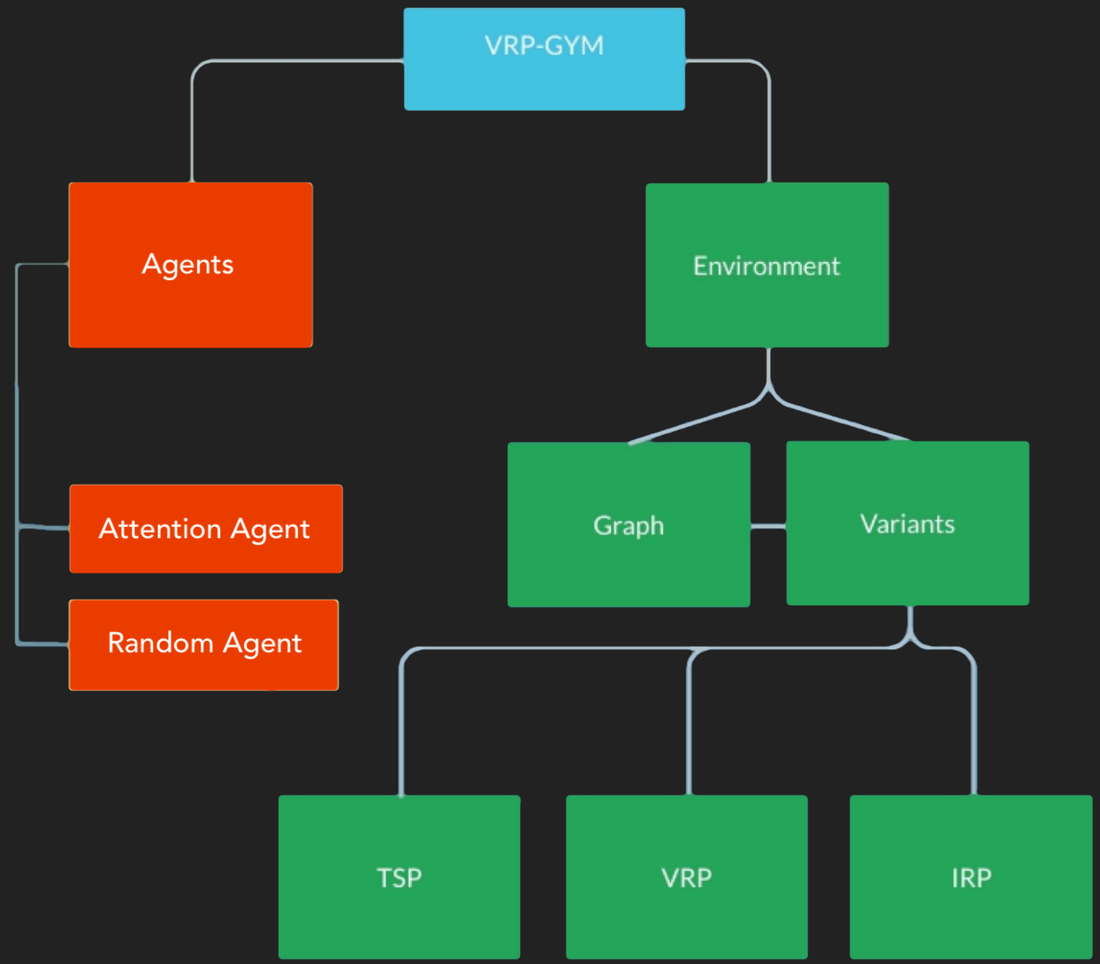
$$OBS_{IRP} \in [\mathbb{R}^{b \times n \times 5}, \mathbb{R}^{b \times 1}]$$

  - Each node entry consists of the coordinates, if it's a depot and if it's visitable. For IRP the demand on each node and the current load is added.

- Action Space
  - Each node in the graph that is currently visitable

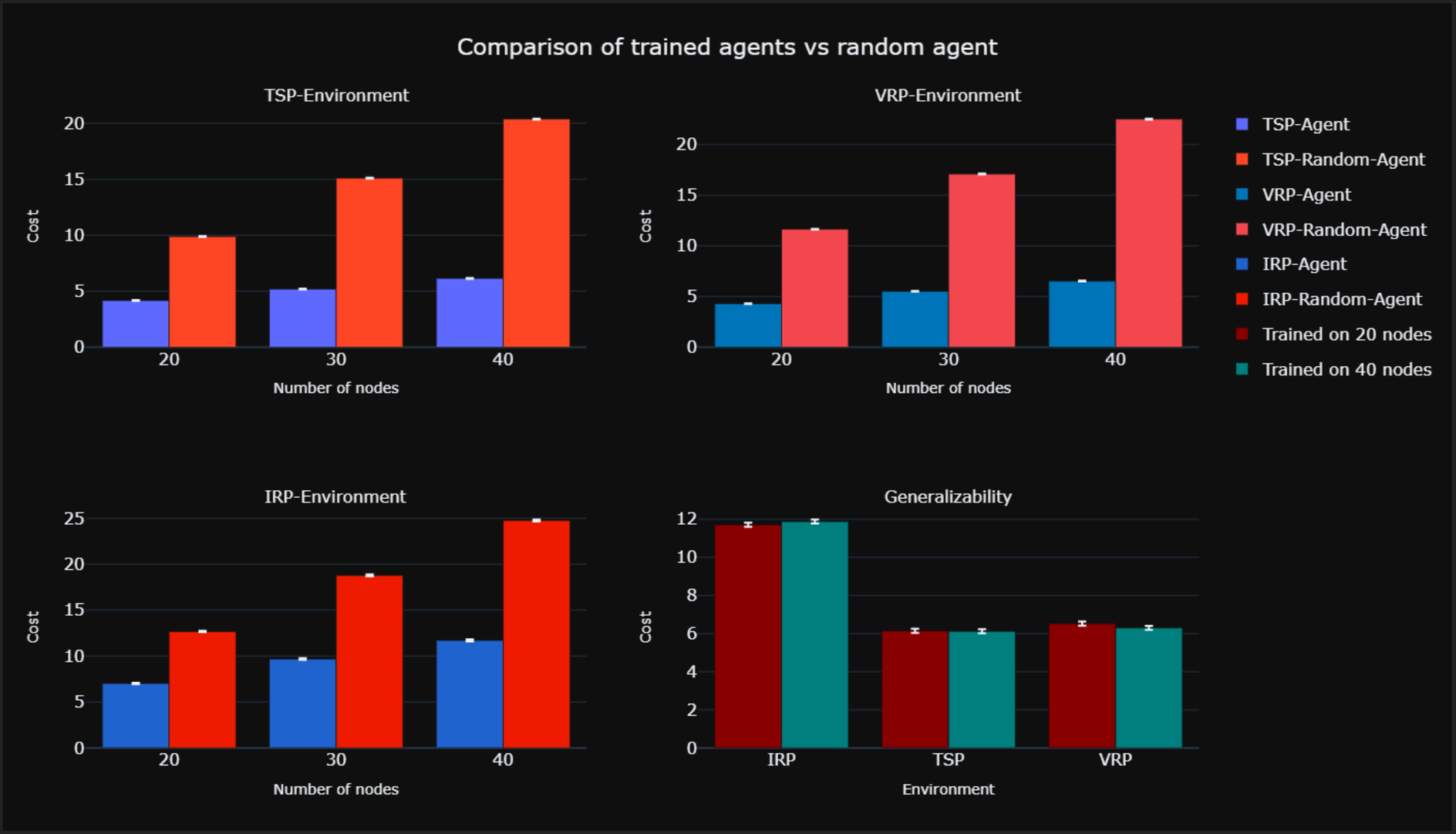
- Reward
  - Negative traversed distance of current step

- Markov Assumption holds true



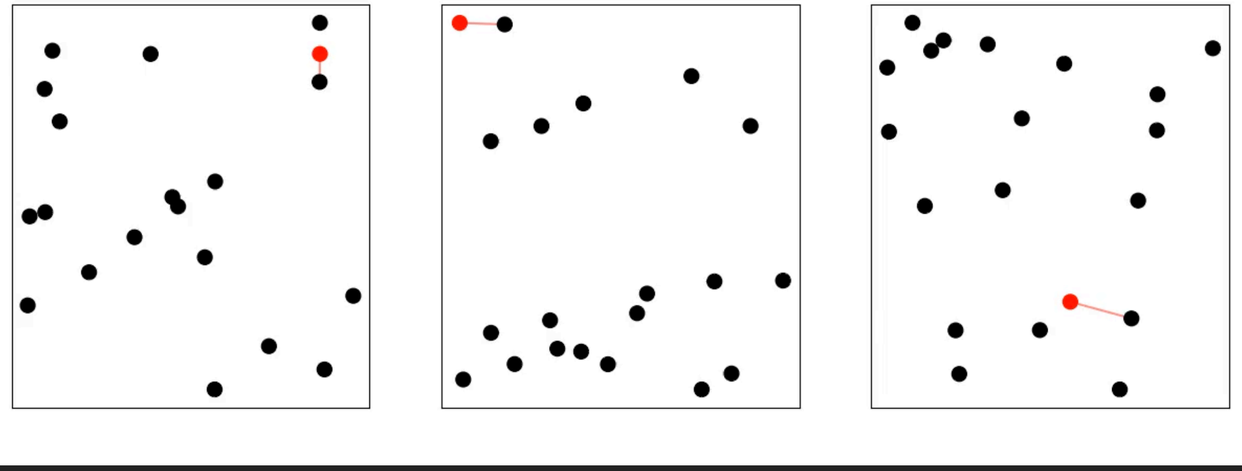
TRAINING EVALUATION

EVALUATION

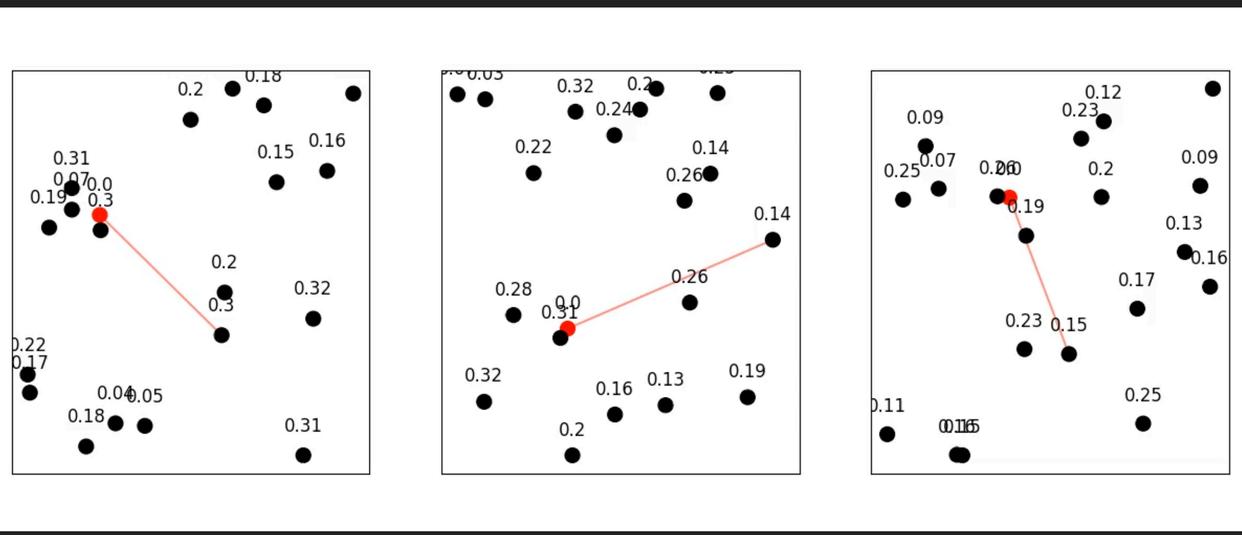

Comparison of trained agents vs random agent
TSP-Environment
VRP-Environment
IRP-Environment
Generalizability
Cost
Number of nodes
Environment
IRP
TSP
VRP
TSP-Agent
TSP-Random-Agent
VRP-Agent
VRP-Random-Agent
IRP-Agent
IRP-Random-Agent
Trained on 20 nodes
Trained on 40 nodes

ATTENTION AGENT ON VRP AND IRP PROBLEM WITH 20 NODES

Vehicle Routing Problem



Inventory Routing Problem



FUTURE WORK

- Implement and evaluate other VRP variants and agents
  - Agent can control multiple vehicles to solve task
- Evaluate on larger graphs (1000+ nodes)
  - Remove solved graph parts and create new embedding
- Benchmark against heuristics
- Optimize efficiency of environment

PIZZA
DELIVERY
USING VRP-GYM
USING
HEURISTICS
imgflip.com

QUESTIONS?