19th International Conference on Autonomous Agents and Multi-Agent Systems 2020

Networked Multi-Agent Reinforcement Learning with Emergent Communication

(Extended Abstract)

Shubham Gupta[1] and Rishi Hazra[1] and Ambedkar Dukkipati[1]

(The first two authors have contributed equally)

[1]Department of Computer Science and Automation
Indian Institute of Science, Bangalore – 560012, INDIA
[shubhamg, rishihazra, ambedkar]@iisc.ac.in

Outline

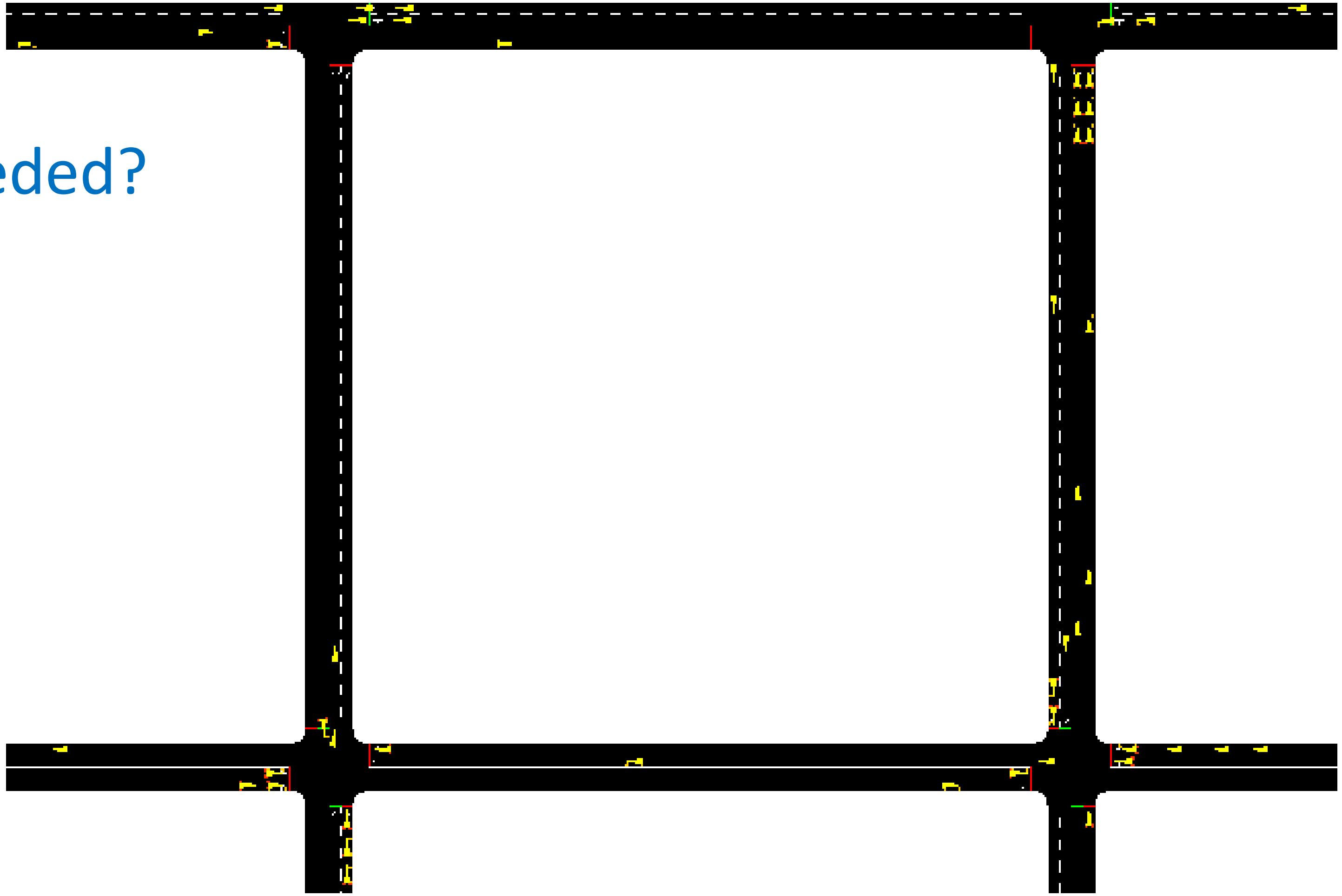
- Emergent communication
- Cooperative multi-agent reinforcement learning (MARL) with emergent communication and networked agents
- Managing traffic controllers
- Analysis of emergent communication

Emergent Communication

Emergent Communication – Motivation I

Why is communication needed?

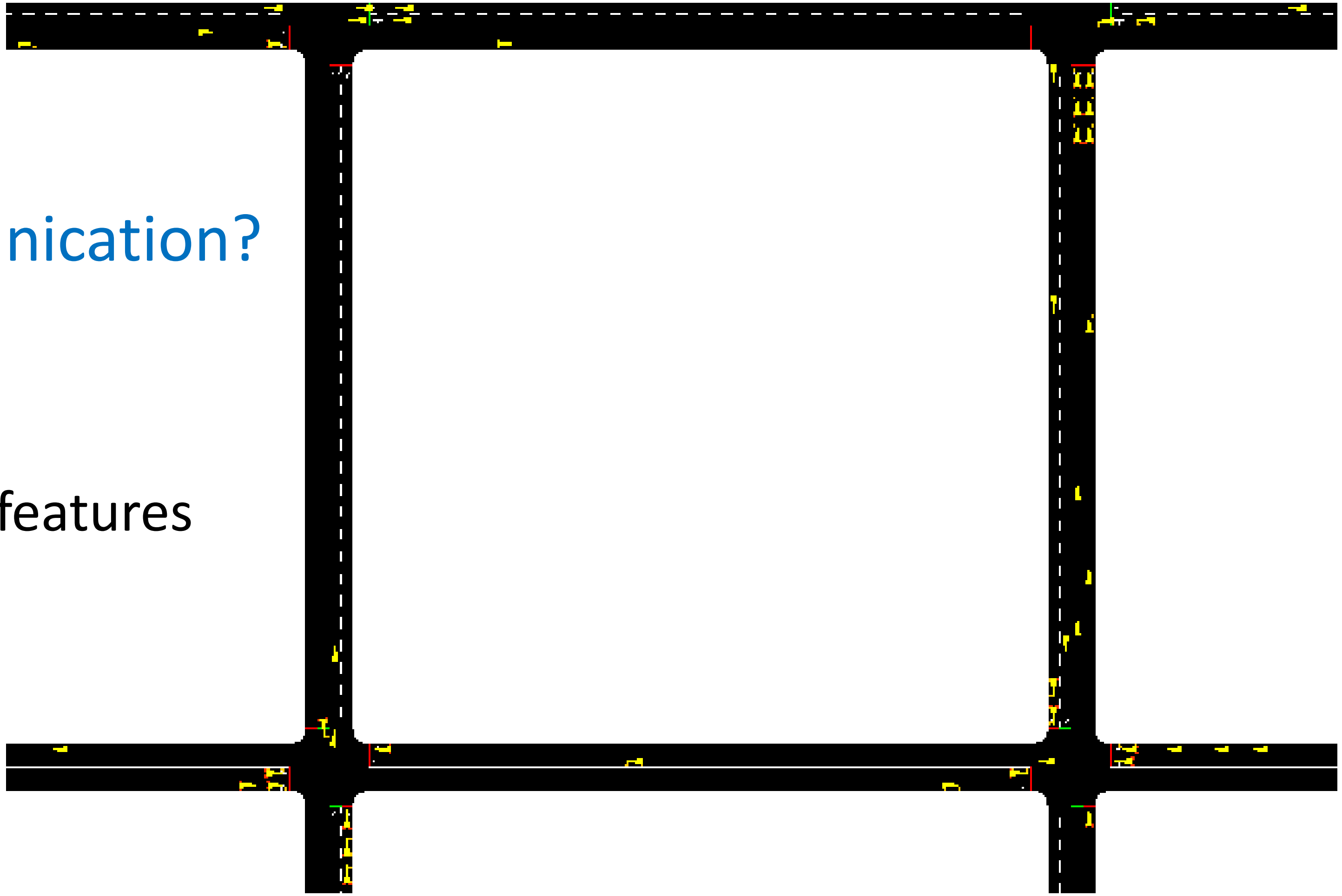
- Partial observability
- Improved cooperation
- Dynamic coordination



Emergent Communication – Motivation II

Why use emergent communication?

- Optimized for the end task
- Think of fixed versus learned features



Emergent Communication – Motivation III

Design choices

- Continuous versus discrete messages
  - Ease of analysis
  - Low bandwidth requirement
- Broadcasting messages versus restricted transmission
  - Agents can only communicate along an underlying network
  - More realistic setting
  - Imposed restrictions influence the emergent communication

Cooperative MARL with Emergent Communication and Networked Agents

MARL – Markov Games

A Markov Game is specified by: $(S, \{A_n, r_n\}_{n=1}^{N}, T, \gamma)$

In our case:

- Agents are connected via an underlying network
- Agents can communicate with each other
- Agents are cooperative

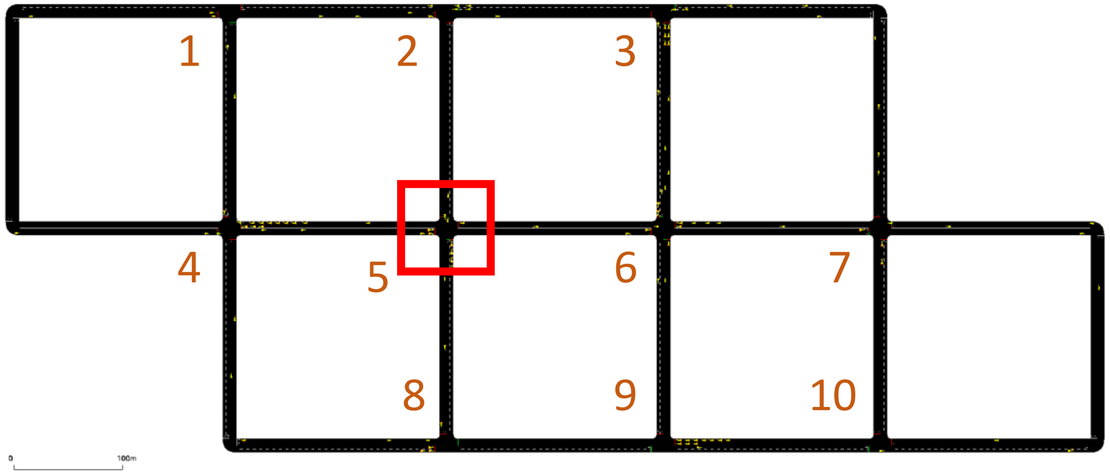
MARL – Related Work

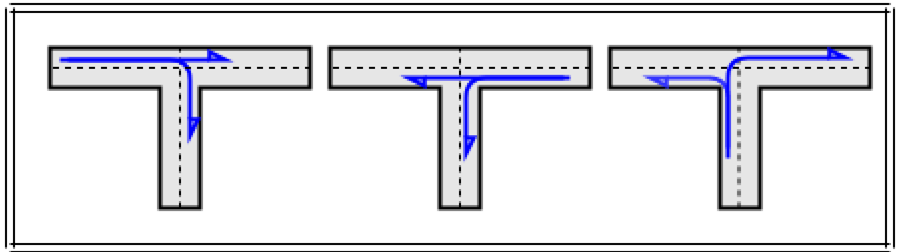
- No communication: [Lowe *et al.*, 2017; Wen *et al.*, 2019]
- Broadcast communication: [Sukhbaatar *et al.*, 2016; Foerster *et al.*, 2016; Mordatch and Abbeel, 2018; Das *et al.*, 2019]
- No analysis of emergent communication: [Zhang *et al.*, 2018; Gupta and Dukkipati, 2020]
- Limited applicability: [Havrylov and Titov, 2017; Cao *et al.*, 2018]
- Traffic: [Miller, 1963; Cools *et al.*, 2013; Wei *et al.*, 2018]

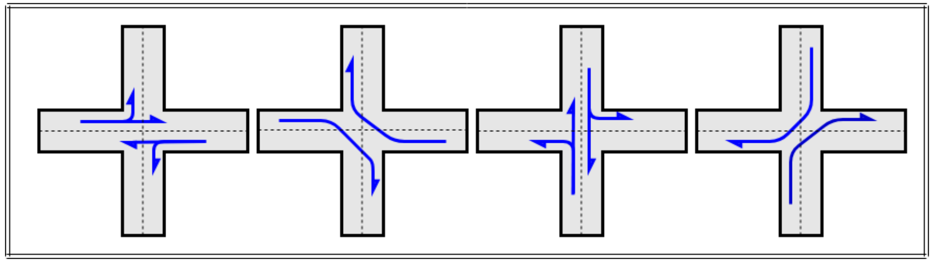
Managing Traffic Controllers

Traffic – Problem Formulation

- Used SUMO simulator [Krajzewicz *et al.*, 2012]
- Local observations
- Junction dependent action spaces
- Cooperative reward structure
- Discrete communication

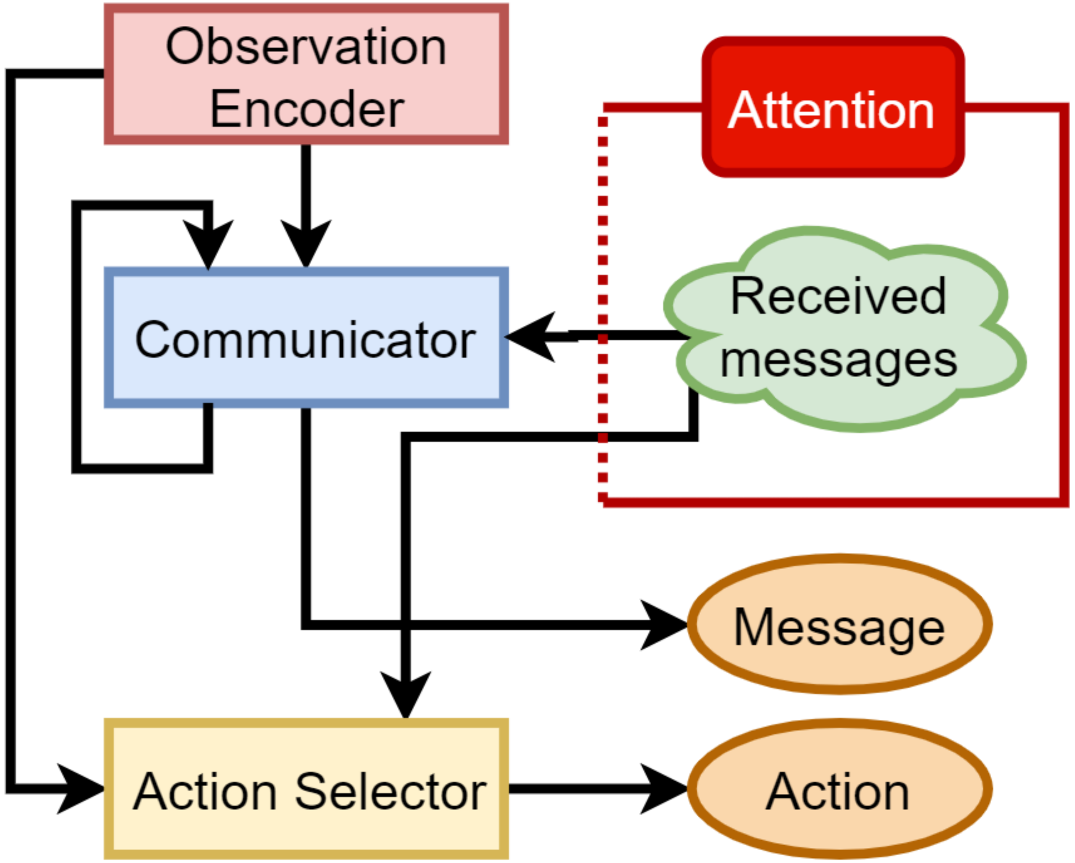






Traffic – Implementation Overview

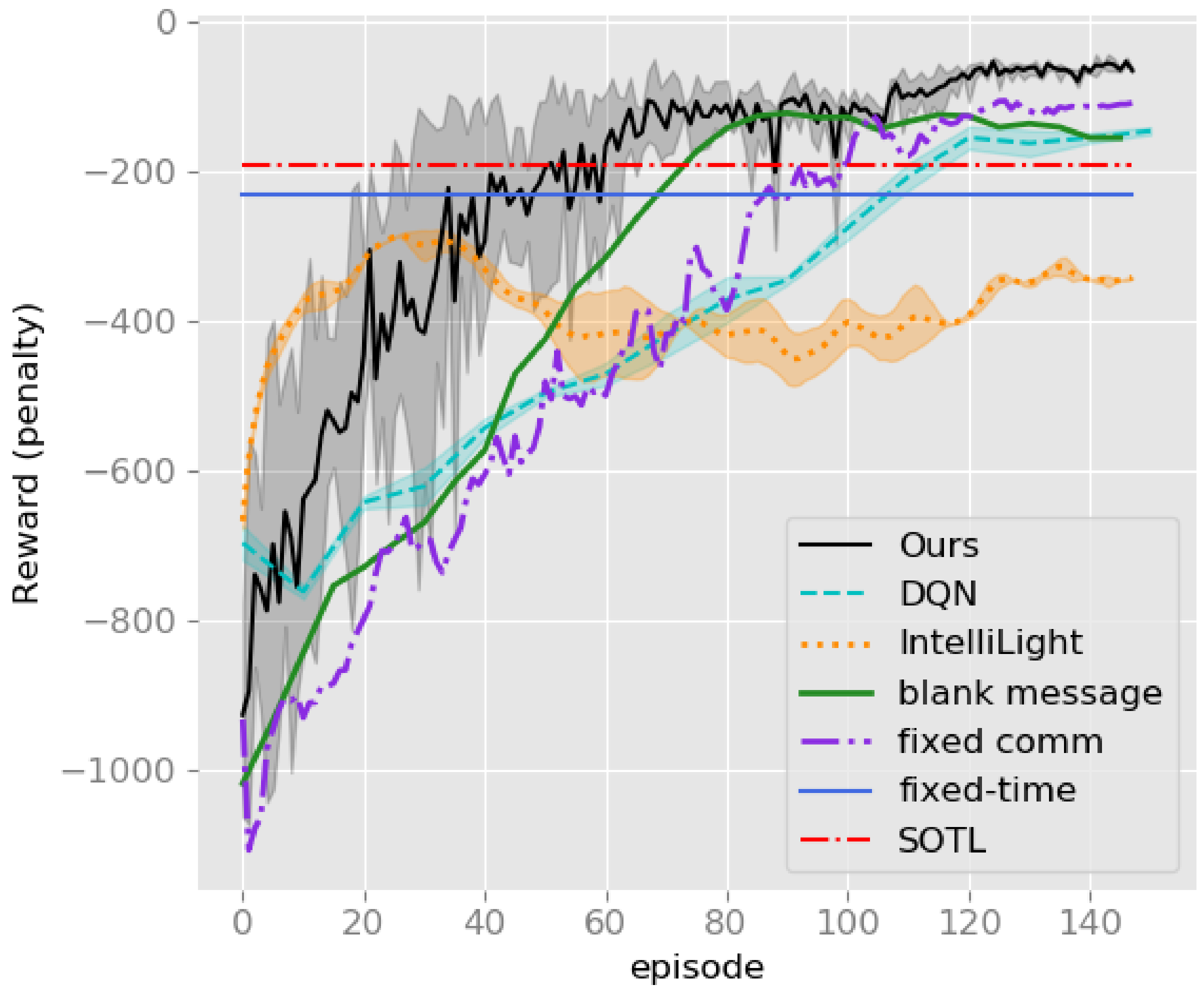
- All modules have been implemented as neural networks
- Gumbel-Softmax used for discrete communication
- Attention mechanism is similar to [Das *et al.*, 2019]



Traffic – Results

Compared with:

- Independent training (DQN)
- No communication based RL methods (blank message, IntelliLight [Wei et al., 2018])
- Fixed communication protocol (fixed comm)
- Heuristics (fixed-time, SOTL)



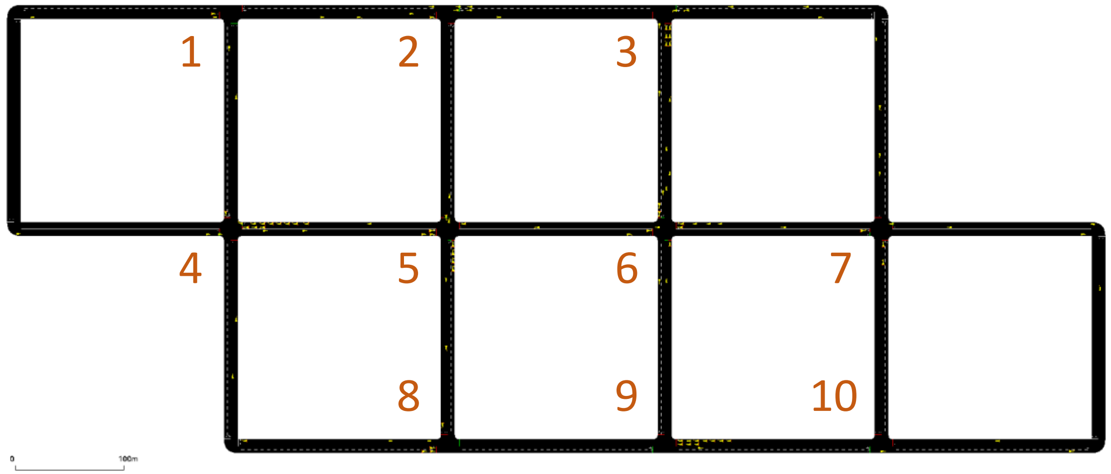
Analysis of Emergent Communication

Communication Analysis

- Utility of communication
- Groundedness
- Effect of network topology

Communication Analysis - Utility

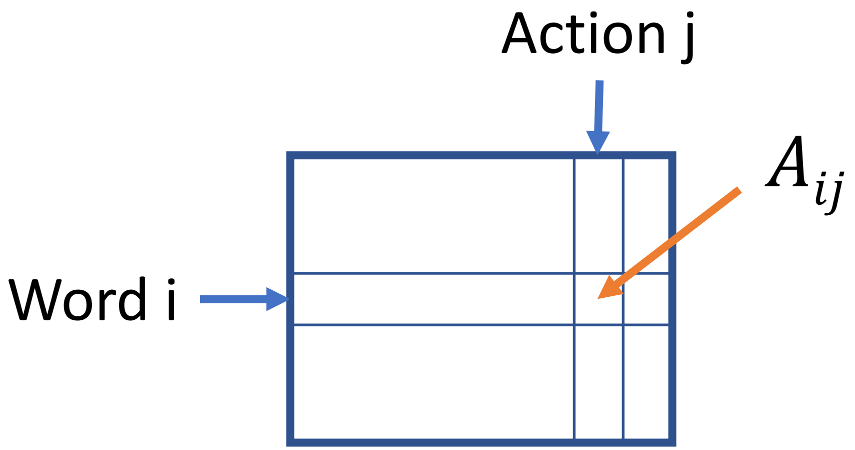
- Blank message baseline does not perform as well
- IntelliLight performs poorly
- Experimented with blind agents
  - No change with one blind agent
  - Performance decreases when two agents are blind

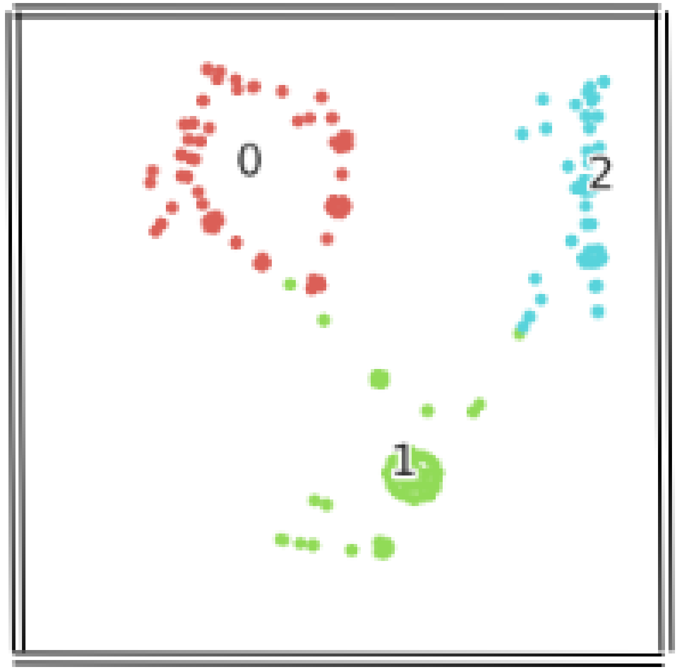


No. of Blind Agents	0	1	2
Reward	-65	-75	-170

Communication Analysis - Grounding

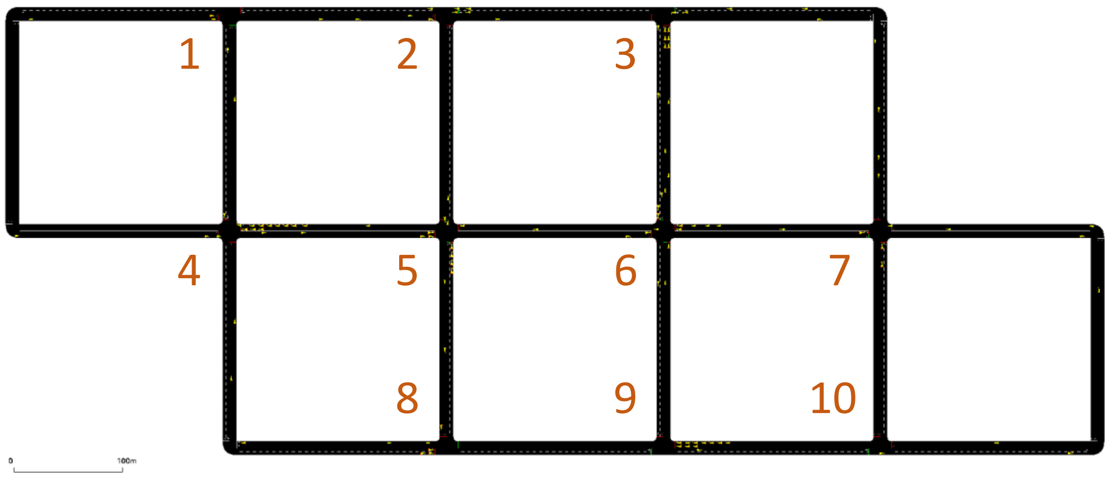
- Take two neighbouring agents and compute the word-action matrix (say A)
- Compute SVD: $A = U\Sigma V^T$
- Rows of U are word-embeddings
- Rows of V are action-embeddings





Communication Analysis – Effect of Network

- Agents sharing a neighbour are called one-hop neighbors
- One-hop neighbors develop the same language
  - Overlap in action-embeddings
  - Similar unigram distributions
- Examples: 1 & 5, 4 & 8 and so on



More results in the paper

Experiments on larger networks

Experiments on robustness of the proposed approach

Summary

- Proposed a general framework for cooperative MARL with:
  - Communication
  - Networked agents
- Achieved state-of-art results on a traffic management problem
- Performed detailed analysis of emergent communication

Thank you!

References

[Lowe et al., 2017] Ryan Lowe, YI WU, Aviv Tamar, Jean Harb, OpenAI Pieter Abbeel, and Igor Mordatch. Multiagent actor-critic for mixed cooperative-competitive environments. In Advances in Neural Information Processing Systems 30, pages 6379–6390. 2017.

[Wen et al., 2019] Ying Wen, Yaodong Yang, Rui Luo, Jun Wang, and Wei Pan. Probabilistic recursive reasoning for multi-agent reinforcement learning. In 7th International Conference on Learning Representations (ICLR), 2019.

[Sukhbaatar et al., 2016] Sainbayar Sukhbaatar, Arthur Szlam, and Rob Fergus. Learning multiagent communication with backpropagation. In Advances in Neural Information Processing Systems 29, pages 2244–2252. Curran Associates, Inc., 2016.

[Foerster et al., 2016] Jakob Foerster, Ioannis Alexandros Assael, Nando de Freitas, and Shimon Whiteson. Learning to communicate with deep multi-agent reinforcement learning. In Advances in Neural Information Processing Systems 29, pages 2137–2145. Curran Associates, Inc., 2016.

References (Contd.)

[Mordatch and Abbeel, 2018] Igor Mordatch and Pieter Abbeel. Emergence of grounded compositional language in multi-agent populations. In Thirty-Second AAAI Conference on Artificial Intelligence, 2018.

[Das et al., 2019] Abhishek Das, Theophile Gervet, Joshua ´ Romoff, Dhruv Batra, Devi Parikh, Mike Rabbat, and Joelle Pineau. Tarmac: Targeted multi-agent communication. Proceedings of the 36th International Conference on Machine Learning, PMLR, 97:1538–1546, 2019.

[Zhang et al., 2018] Kaiqing Zhang, Zhuoran Yang, Han Liu, Tong Zhang, and Tamer Basar. Fully decentralized multi-agent reinforcement learning with networked agents. Proceedings of the 35th International Conference on Machine Learning, PMLR, 80:5872–5881, 2018.

[Gupta et al., 2020] Shubham Gupta and Ambedkar Dukkipati. Winning an Election: On Emergent Strategic Communication in Multi-agent Networks, to appear ini AAMAS 2020.

References (Contd.)

[Havrylov and Titov, 2017] Serhii Havrylov and Ivan Titov. Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In Advances in Neural Information Processing Systems 30, pages 2149–2159. Curran Associates, Inc., 2017.

[Cao et al., 2018] Kris Cao, Angeliki Lazaridou, Marc Lanctot, Joel Z Leibo, Karl Tuyls, and Stephen Clark. Emergent communication through negotiation. In 6th International Conference on Learning Representations (ICLR), 2018.

[Miller, 1963] Alan J. Miller. Settings for fixed-cycle traffic signals. Journal of the Operational Research Society, 14(4):373–386, 1963.

[Cools et al., 2013] Seung-Bae Cools, Carlos Gershenson, and Bart D'Hooghe. Self-Organizing Traffic Lights: A Realistic Simulation, pages 45–55. Springer London, 2013.

References (Contd.)

[Wei et al., 2018] Hua Wei, Guanjie Zheng, Huaxiu Yao, and Zhenhui Li. Intellilight: A reinforcement learning approach for intelligent traffic light control. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery; Data Mining, KDD ’18, pages 2496–2505, 2018.

[Krajzewicz et al., 2012] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker-Walz. Recent development and applications of sumo - simulation of urban mobility. International Journal On Advances in Systems and Measurements, 5(3&4):128–138, 2012.