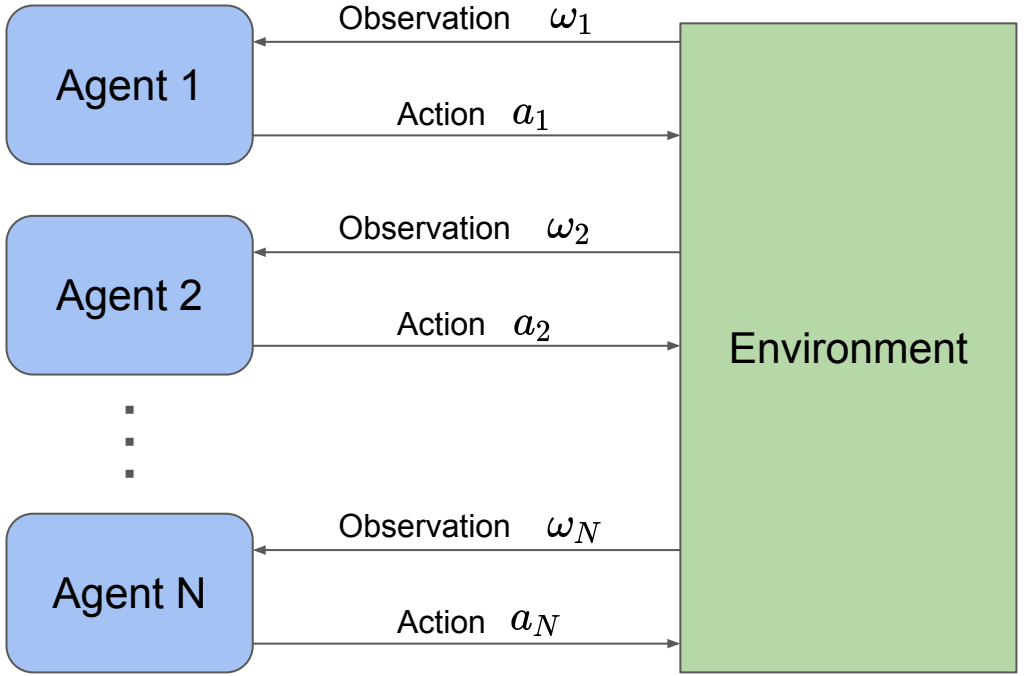
Learning to Communicate in Multi Agent Settings

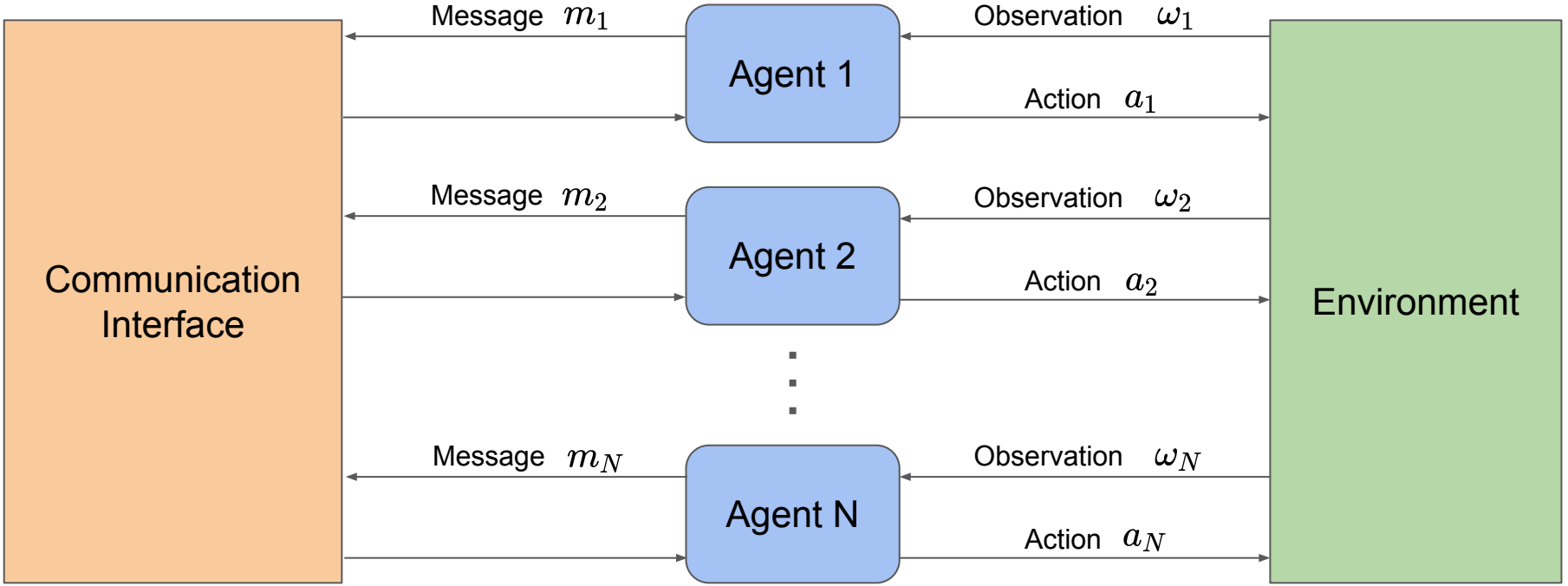
Joel Oskarsson, joeos014@student.liu.se

Reinforcement Learning

- Partial observability
- Collaborative



Communication

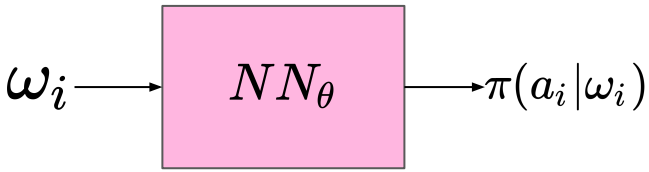


Deep RL Algorithms

- Goal: Learn a policy $\pi(a_i|\omega_i)$

Policy Gradient Methods (e.g. REINFORCE)

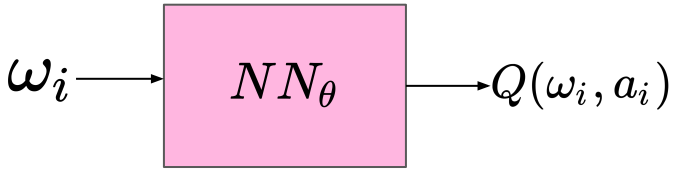
- Policy is Neural Network: $\pi(a_i|\omega_i) = NN_\theta(\omega_i)$



[Sutton *et al.*]

Deep Q-learning (DQN)

- Q-function is Neural Network:
  $Q(\omega_i, a_i) = [NN_\theta(\omega_i)]_{a_i}$



[Mnih *et al.*]

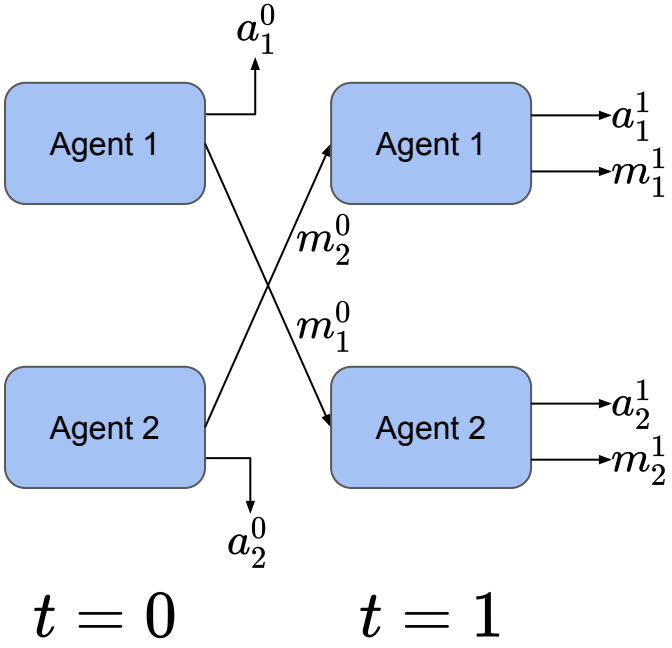
Discrete Messages

- Binary messages $m_i \in \{0,1\}^L$

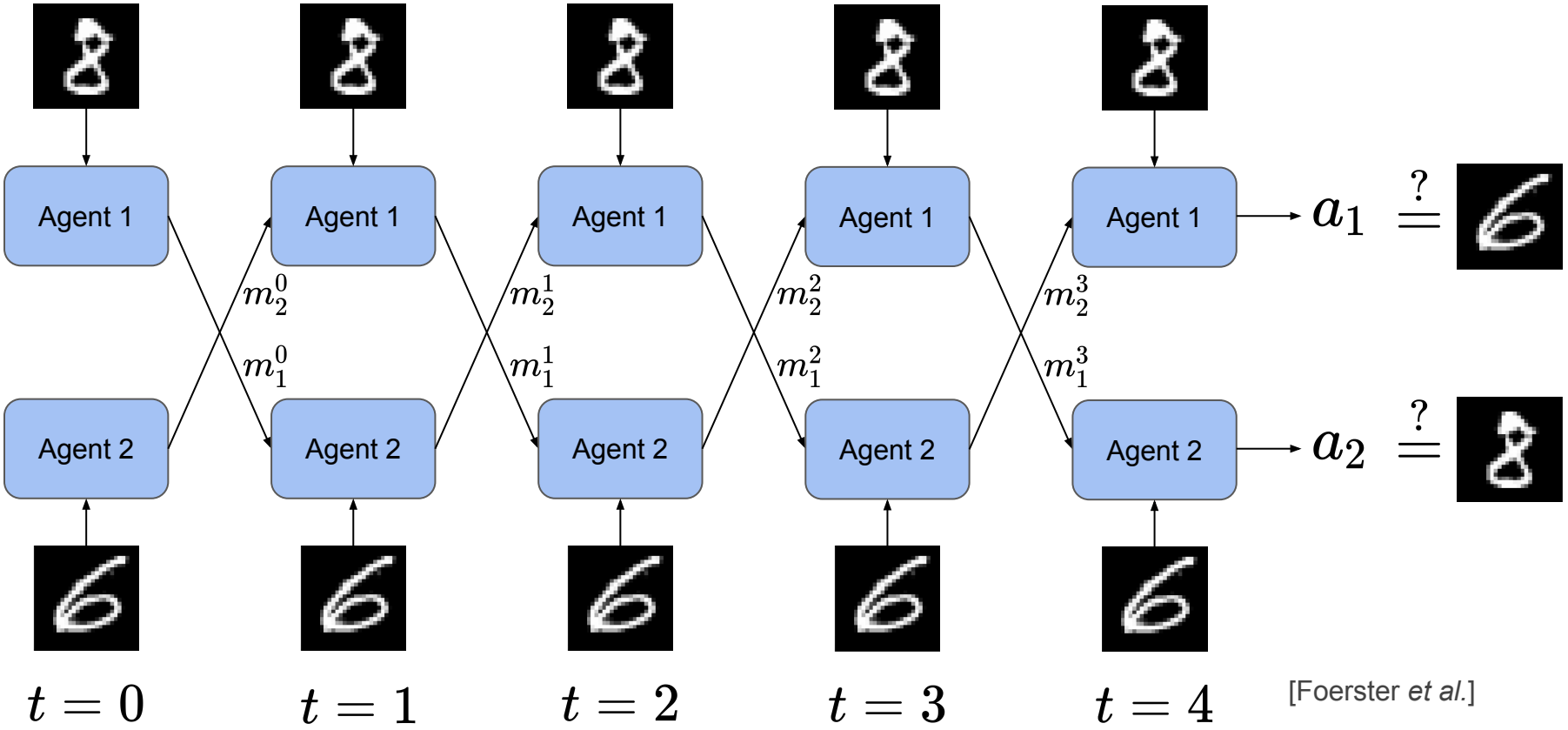
Differentiable Inter-Agent Learning (DIAL)

- DQN for actions: $Q(\omega_i, a_i)$
- Messaging trained implicitly

[Foerster *et al.*]



MNIST Game



[Foerster *et al.*]

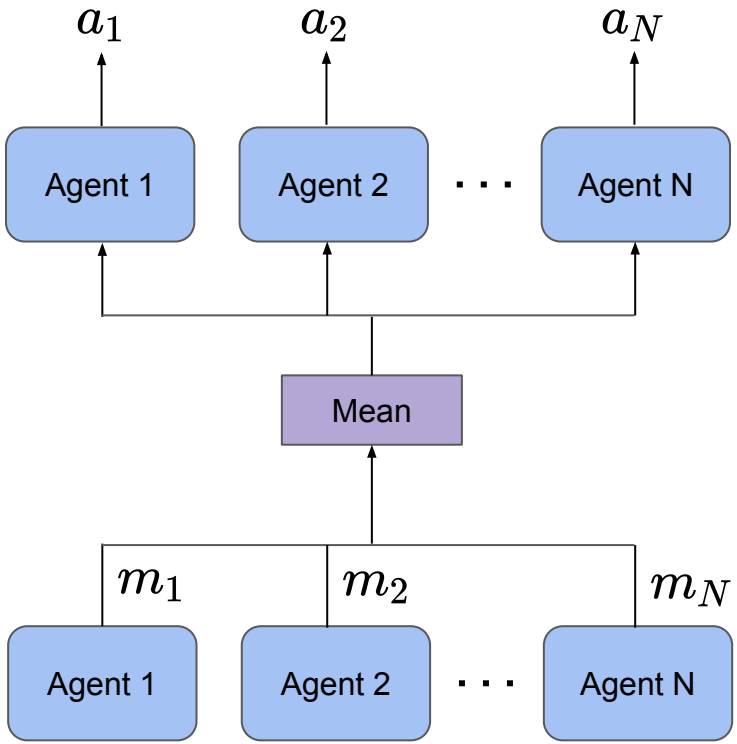
Continuous Messages

- Real-valued messages $m_i \in \mathbb{R}^L$

Communication Neural Net (CommNet)

- REINFORCE
- Dynamic amount of agents

[Sukhbaatar *et al.*]



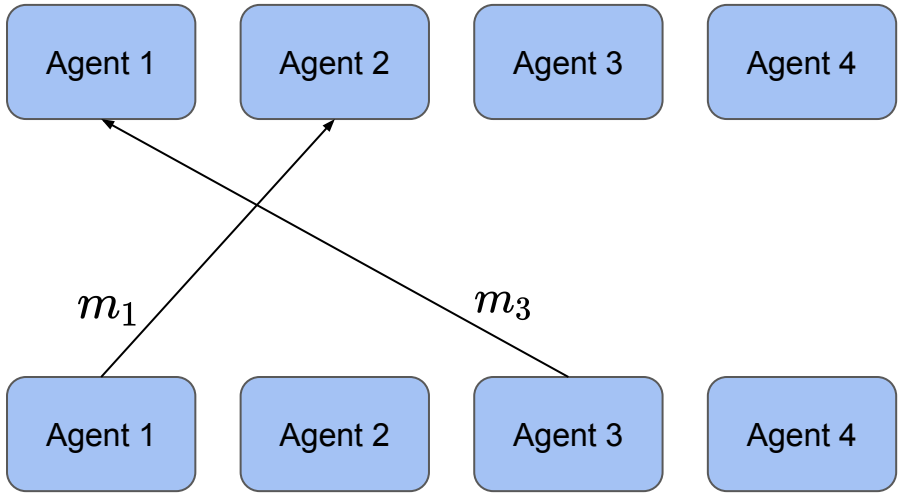
Extensions

IC3Net [Singh *et al.*]

- Learning <u>when</u> to communicate

TarMAC [Das *et al.*]

- Targeted communication



Thanks for listening!

Questions are welcome!

References

- [Sutton et al.] Richard S Sutton, Andrew G Barto. *Introduction to reinforcement learning*, volume 2. MIT press Cambridge, 1998.
- [Mnih et al.] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. *Human-level control through deep reinforcement learning*. Nature, 518(7540):529, 2015.
- [Foerster et al.] 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
- [Sukhbaatar et al.] 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.
- [Singh et al.] Amanpreet Singh, Tushar Jain, and Sainbayar Sukhbaatar. *Individualized controlled continuous communication model for multiagent cooperative and competitive tasks*. In International Conference on Learning Representations, 2019.
- [Das et al.] Abhishek Das, Th´eophile Gervet, Joshua Romoff, Dhruv Batra, Devi Parikh, Mike Rabbat, and Joelle Pineau. *Tarmac: Targeted multi-agent communication.* In International Conference on Machine Learning, pages 1538–1546, 2019.