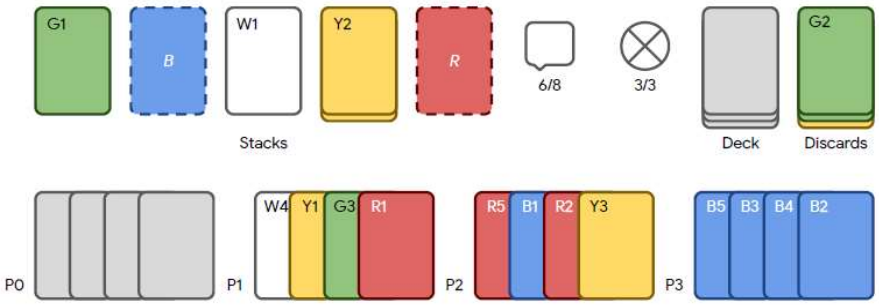
The Hanabi Game

Xiuhan Wang, Jing Wei
Institute for Interdisciplinary Information Sciences, Tsinghua University

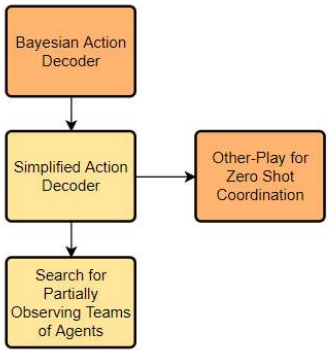
Introduction

Motivation:
- A benchmark of fully cooperative games of imperfect information

Contribution:
- Surveyed a sequence of learning methods and compared their performance
- Used a symmetric network in order to learn symmetries itself



State of the Art

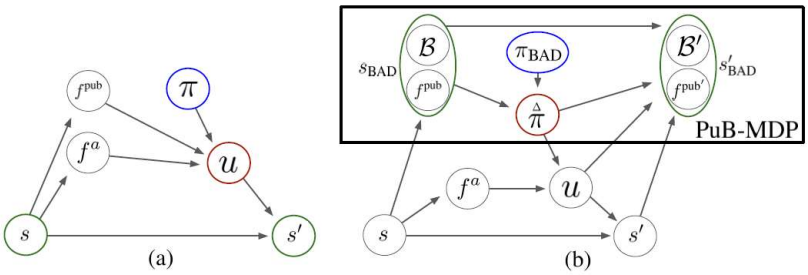


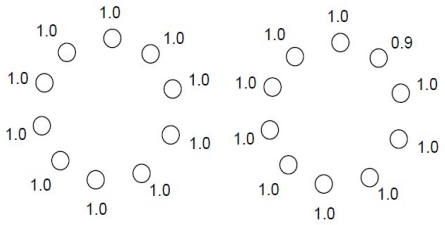
Methods

Bayesian Action Decoder:
- Joint public belief over private features
- Sample deterministic partial policies

Other-Play for Zero Shot Coordination:
- Optimize $\pi^* = \arg\max E_{\phi\sim\Phi} J(\pi^1, \phi(\pi^2))$
- Φ is given as input

Our Method:
- An analog approach to convolutional neural network
- Use symmetric network structures to learn symmetries itself





Details

- A permutation group G: defining symmetries
- For each neuron h in the network and a permutation $\phi \in G$, there should also be a neuron $\phi(h)$
- The neurons can be divided into equivalence classes
- h and $\phi(h)$ share the same bias
- If h takes its input from the neuron g with weight w, $\phi(h)$ takes input from $\phi(g)$ with the same weight

In the specific problem, $G = S_5$ and the construction of the network applies to various types of neurons including LSTM cells.

Experiments

The performances under different training settings:

method	mean score	s.e.m.	perfect ratio
IQL	23.773	0.017	0.436
VDN	23.897	0.018	0.487
SAD	23.980	0.018	0.531
SAD + Aux	24.015	0.018	0.544

Bayesian Action Decoder and its variants perform better than the previous methods!

Experiments for our symmetric network are still running.