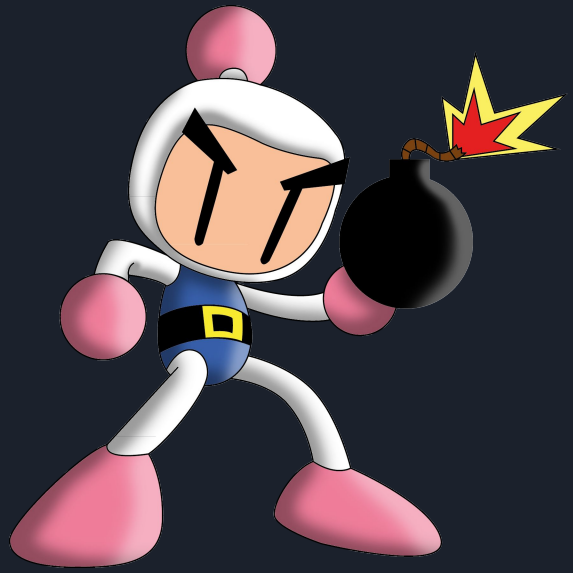


Pommerman

A multi-agent learning game

Team

Allen Shen zhunshen@usc.edu

Jameson Thai jamesont@usc.edu

Kevin Cheng kevinich@usc.edu

Lance Newby lnewby@usc.edu

Muhammad Usman Shahid mshahid@usc.edu

Myron Kwan myronkwa@usc.edu

Raja Zain Abbas abbas@usc.edu

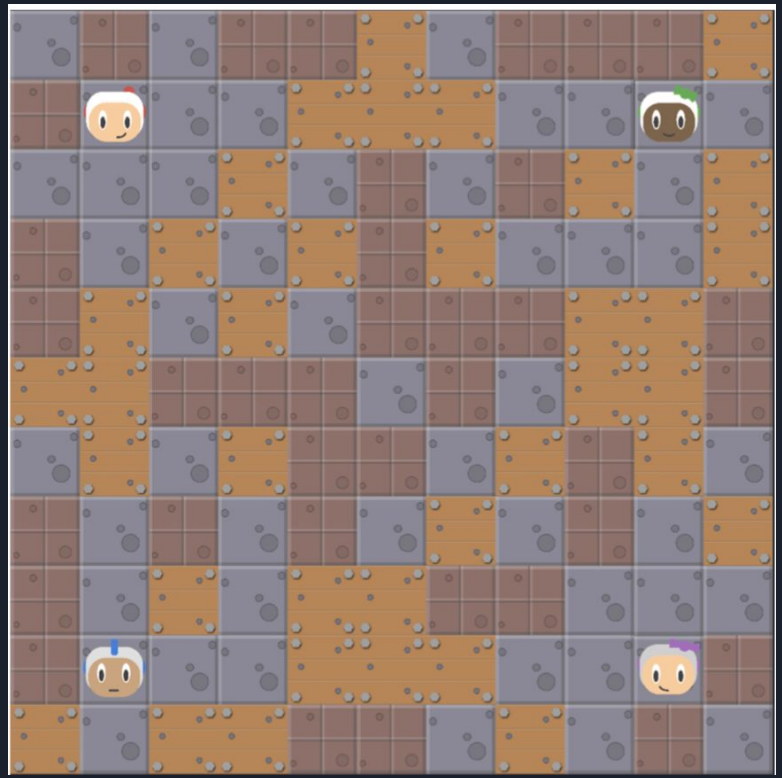
What's Pommerman?

A multi-agent environment based on the classic console game Bomberman.

Pommerman consists of a set of scenarios: 1 vs 1, team vs team, Free-For-All)

Players are cooperative and competitive at the same time.

Game Information



Actions

In any given turn, an agent can choose from one of six actions:

- **Stop:** This action is a pass.
- **Up:** Move up on the board.
- **Left:** Move left on the board.
- **Down:** Move down on the board.
- **Right:** Move right on the board.
- **Bomb:** Lay a bomb.
- **Message:** Send a fixed size discrete vector to your teammate. Happens only in communicative games.

Power-ups

Half of the wooden walls have hidden power-ups that are revealed when the wall is destroyed:

- **Extra Bomb:** Picking this up increases the agent's ammo by one.
- **Increase Range:** Picking this up increases the agent's blast strength by one.
- **Can Kick:** Picking this up allows an agent to kick bombs. It does this by running into them. The bomb then travels in the direction that the agent was moving at a speed of one unit per time step until they are impeded either by a player, a bomb, or a wall.

Observation Data

- **Board:** 121 Ints. The flattened board. All squares outside of the agent's purview will be covered with the fog value (5).
- **Position:** 2 Ints, each in [0, 10]. The agent's (x, y) position in the grid.
- **Ammo:** 1 Int. The agent's current ammo.
- **Blast Strength:** 1 Int. The agent's current blast strength.
- **Can Kick:** 1 Int, 0 or 1. Whether the agent can kick or not.
- **Teammate:** 1 Int in [-1, 3]. Which agent is this agent's teammate. For FFA, this will be -1.
- **Enemies:** 3 Ints, each in [-1, 3]. Which agents are this agent's enemies. If this is a team competition, the last Int will be -1 to reflect that there are only two enemies.
- **Bombs:** List of Ints. The bombs in the agent's purview, specified by (X int, Y int, BlastStrength int).
- **Message:** Fixed-size list of Ints. The message from the agent's teammate. This happens only in communicative environments.

Methods

- ➔ Planning (path, power-ups)
- ➔ Opponent/teammate modeling (team battle, Free-For-All)
- ➔ Communication (team battle)

Algorithms

- ➔ Convolutional neural network CNN
- ➔ Proximal Policy Optimization PPO
- ➔ Advantage Actor Critic (A2C), a synchronous deterministic version of A3C
- ➔ Scalable trust-region method for deep reinforcement learning using Kronecker-factored approximation ACKTR
- ➔ Generative Adversarial Imitation Learning GAIL

Tools

- ➔ Anaconda3
- ➔ Python 3.7
- ➔ OpenAI Gym
- ➔ Tensorflow/Keras
- ➔ Pytorch
- ➔ Google Colab
- ➔ Docker
- ➔ Flask

Goals

- ★ Phase One
  - Read paper/articles
  - Learn new algorithms(PP0, A3C etc)
  - Learn DL frameworks(Pytorch vs Tensorflow)
  - Train new agents
- ★ Phase Two
  - Compete with other teams
  - Learning new strategies
  - Optimize our model
- ★ Phase Three
  - Push source code /documentation
  - Release benchmarks
  - Publish a paper/article for new findings
  - Participate Pommerman Competitions in NIPS 2019 (Fight On)

Collaboration

https://ml4games.slack.com

Join to study and discuss together for 599, even we’re not in the same team.