

WEEK 9



Pommernan



Fight



BIN CHEN



JIARONG QIU



SAIRAM KAMAL RAJ



SHUWEI SHI



WEI CHENG

USC Viterbi

School of Engineering



Univers. of Southern California

Overview



What is Pommerman?

Four Players: 

Mode: FFA(Free for All)

Action :      

Buff: ammo, bomb strength

Goal: Survive



Prior Research



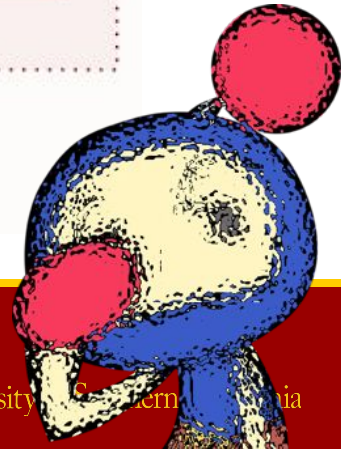
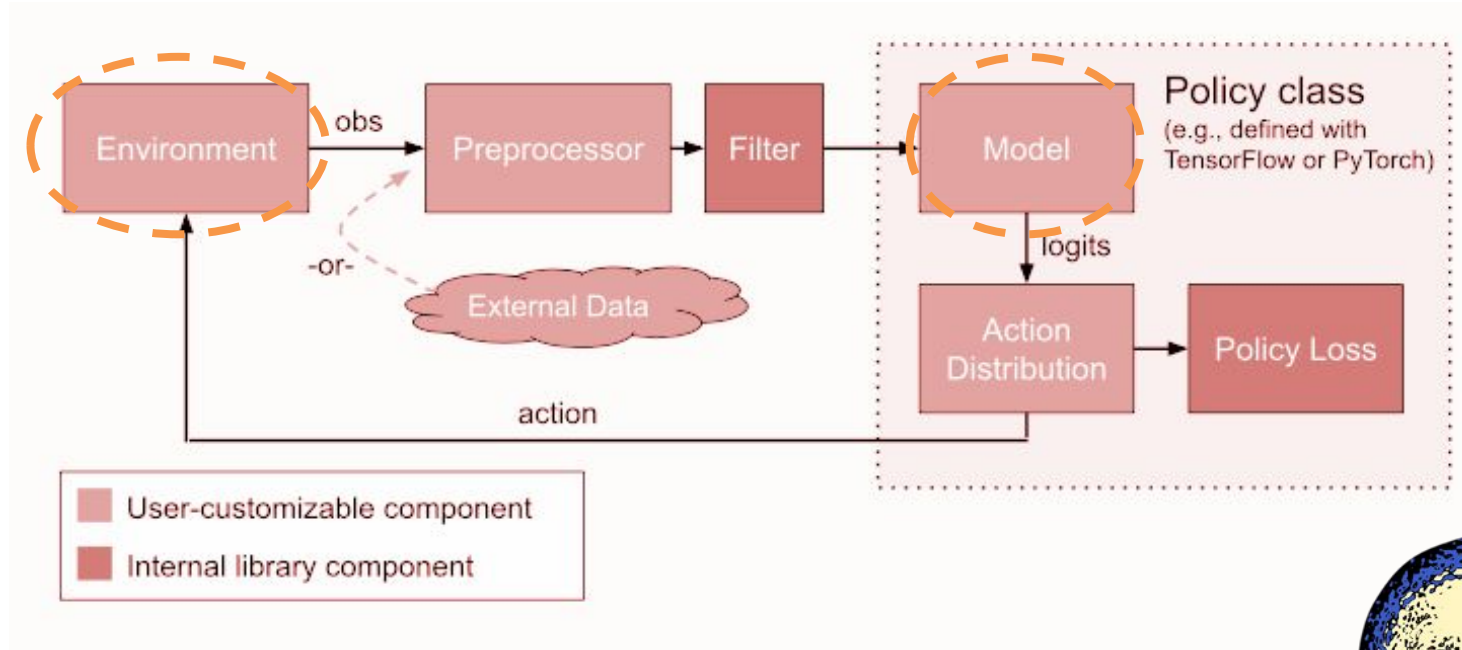
Libraries:

- TensorFlow
- PyTorch
- Gym

RL Algorithms:

- A2C
- DQN
- PPO

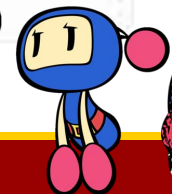
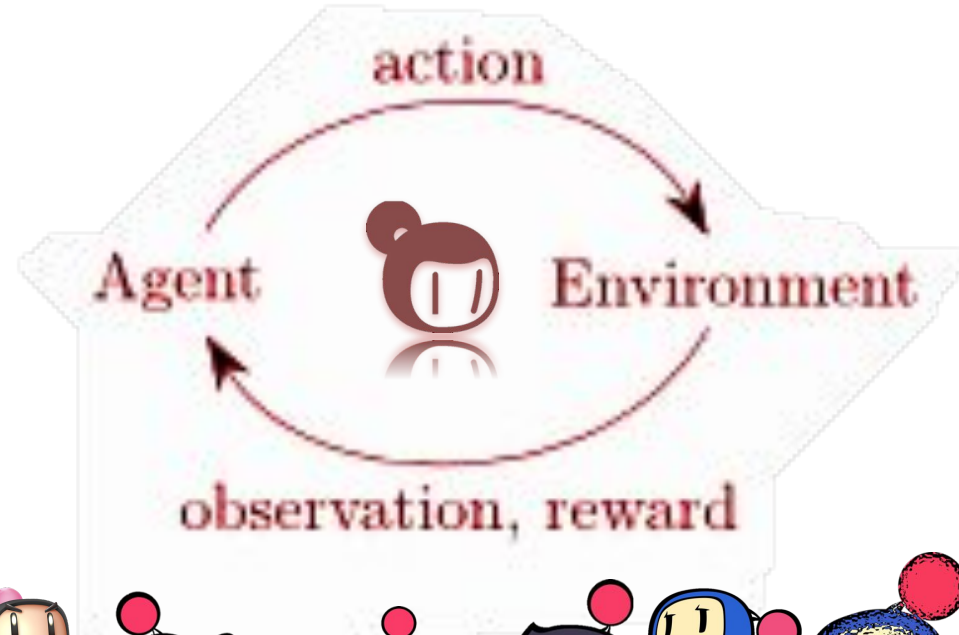
RLlib Framework



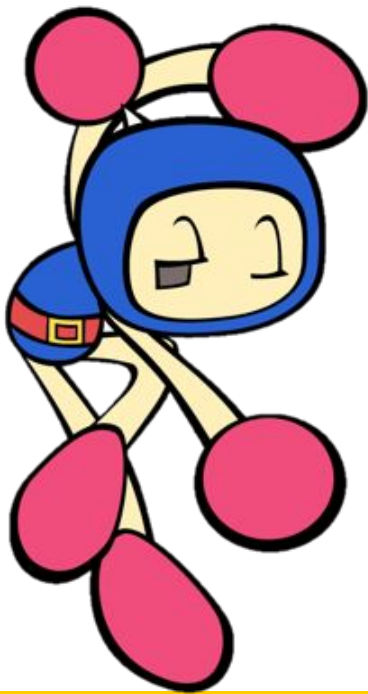
Environment



- Properties
 - observation_space
 - action_space
- Functions
 - reset() -> obs
 - step(action) -> obs, reward, done, info



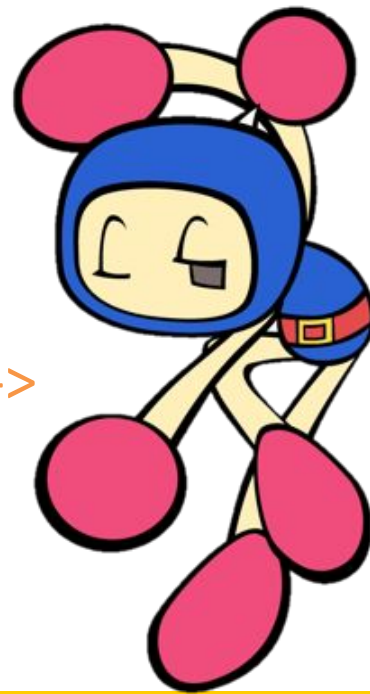
Model



1. Inherit TFModelV2 or TorchModelV2

2. Override Functions

- `forward(input_dict, state, seq_lens) -> logits, state`
- `value_function() -> value`



Training API

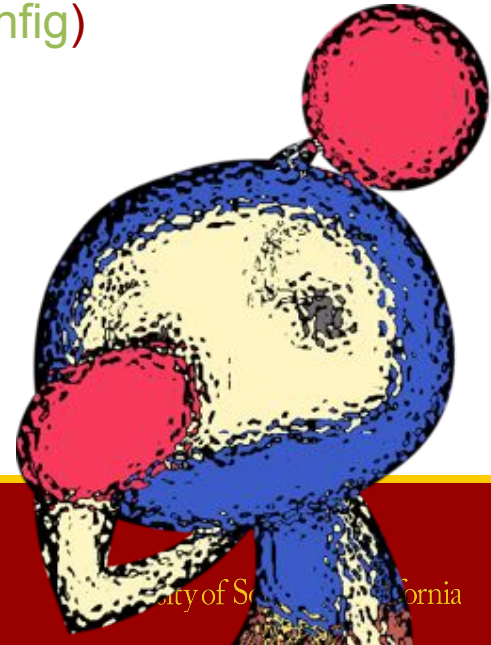


Trainers

- `trainer = a3c.A2CTrainer(env="my_env", config=config)`
- `trainer = ppo.PPOTrainer(env="my_env", config=config)`
- `trainer = dqn.DQNTrainer(env="my_env", config=config)`

Training:

- `for i in range(num_iterations):`
- `trainer.train()`



Pommernan Problem

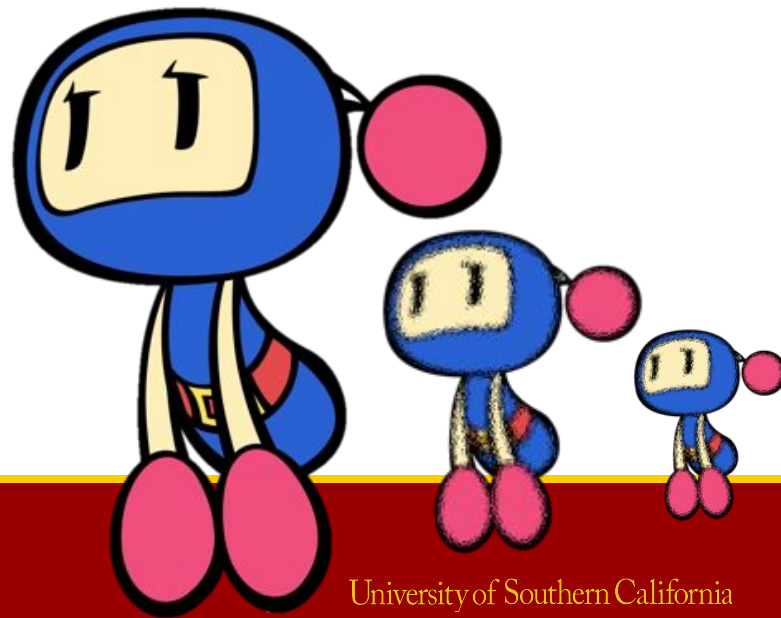


Observations

- Board, Bomb_blast_strength, Bomb_life, Flame_life (11x11)
- Position (row,col)
- Ammo, Blast_strength (1)

Actions :       [0-5]

Reward: Win: +1, Lose: -1

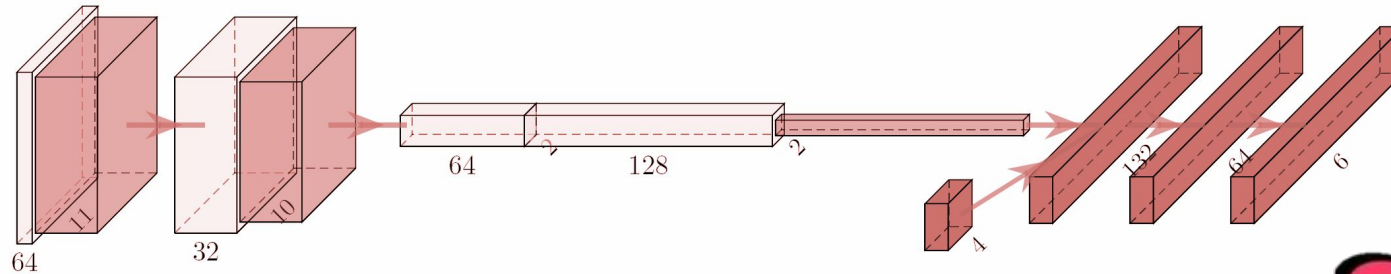




Customized Model



- Fully Connected model
- CNN with LSTM
- CNN derived from VGG-16



Customized Reward Shape



- Base reward each step: 0
- Win: $+$ 200
- Lose: $-$ 200
- Drop bomb: # of wood and enemies around the bomb
- Gain a buff: 3





1. Dummy Agent (epoch=0)
2. Dodging Agent (original reward shape)
3. Offensive Agent (custom reward shape)

Issues



1. Hide in a corner
2. No attempt to defeat other players
3. Overfitting
4. Strategies too SIMPLE



At the end of the semester, we will....



*Build a super agent
that is no longer a
human being. DIO!*

Thank You!



*To be
continued...*

At the end of the semester, we will....



Build a **super**
agent that is no
longer a **human**
being. DIO!

Thank You!



*To be
continued...*

