HW3: Common Challenges

10-403 RECITATION (4/3/20)

General Recommendation

Keras

- Model.fit (batch_size = T)
- Model.train_on_batch (): arbitrary batch size
- Custom loss function (or cross-entropy loss)

Action repeat

- Don’t anneal the number of action (that should be constant)
- Correct transition tuples: $(s_t, a, s_{t+r-1}, \sum_{i=t}^{t+r-1} \gamma^{i-t} R_i)$

General Recommendation

Symptom: the average reward of DQN plunges after solving the task

- You don't have a target network
- Large learning rate
- Small size of replay memory

Symptom: REINFORCE and A2C saturates (e.g. stuck at sub-optimal returns)

- You take the argmax of actions; sample from your policy
- Large learning rate
- You can add the entropy regularization to your loss