sewts

Weekly Review

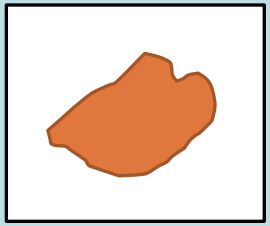
06/04/21

- **Tasks**
- Conduct experiments and evaluate results using GUI and plots ⚠ ***In progress***
- **Problems**
- Results of experiment unexpected
- **To-Do Items for Next Week**
- Compare organization of current repo with RLPYT original repository
- Update states, values, rewards till you get
- Define more reward functions, states, actions for our use case
- **To-Do Later**
- Explore usage of intermediate testing on simulation before sim-to-real transfer
- Define use-case (for different type of towels (colour, texture, etc.) / one type)

sewts

Cloth Manipulation using SAC

06/04/21



Observation
(64*64 img)

Reward
(Overlap with goal state)

SACAgent

Agent

Action
Pick point and place point
From random pixel points
Inside segmented mask

Environment (Mujoco)

RL Problem for obtaining one flat seam

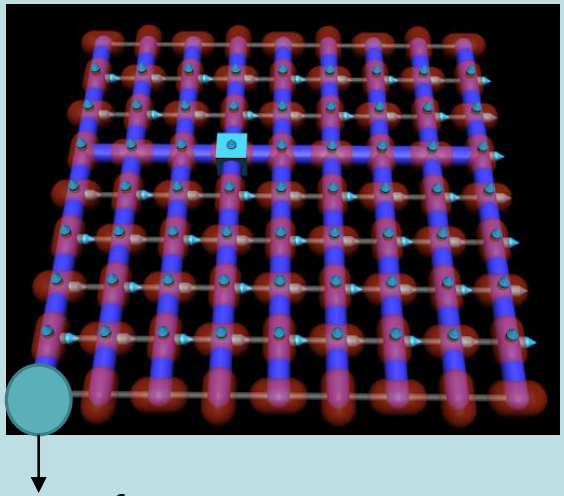
06/04/21

Goal

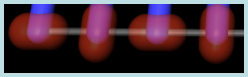
Obtain one flat seam

Given

Cloth in mujoco represented by 64 particles in 8*8 grid



Observations



[x,y,z] positions of 4 points adjacent to corner 1

Actions

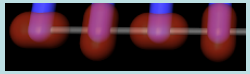
Random [x,y] movement of corner 1

RL Problem for obtaining one flat seam

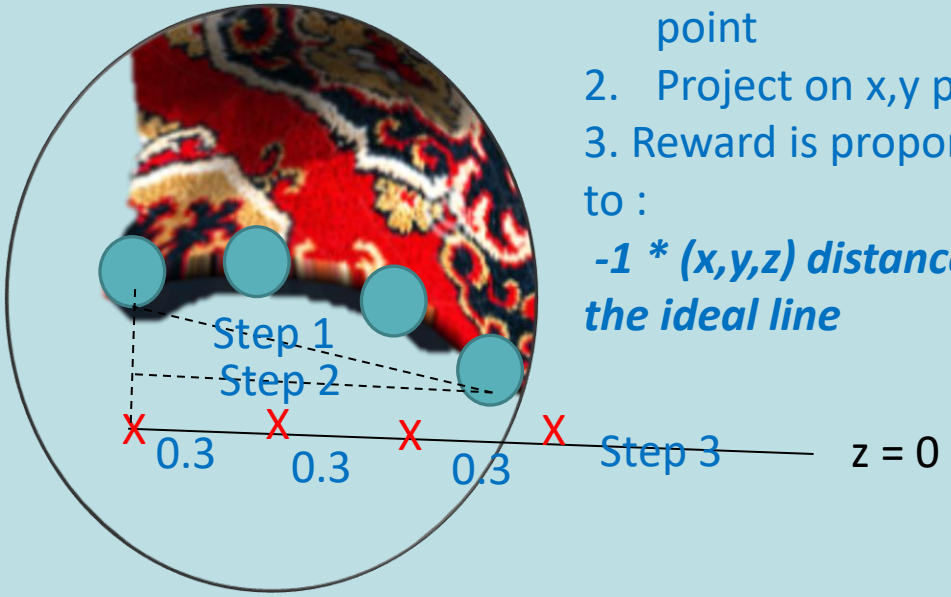
06/04/21

Goal

Obtain one flat seam -> Corner particle + 3 adjacent particles in a straight line



Reward



1. Join 1st point and last point
2. Project on x,y plane
3. Reward is proportional to :

-1 * (x,y,z) distance from the ideal line

sewts

Approach testing now

06/04/21

Algorithm

1. Take random actions for initial N steps
2. Store observations, actions, next observations, rewards
3. Filter the stored data for rewards > -0.010 (Good rewards)
4. Select states and actions from filtered data
5. Define a single layer NN
6. Train the NN on states and actions from filtered data
7. Predict actions for new observations after the initial N steps
8. Check results
9. Update NN network (Add more layers / nodes)
10. Repeat Steps 6-10 till the results are satisfactory

Next Steps

- Include more networks from the SAC implementation
- Train on full SAC implementation

Plan

06/04/21

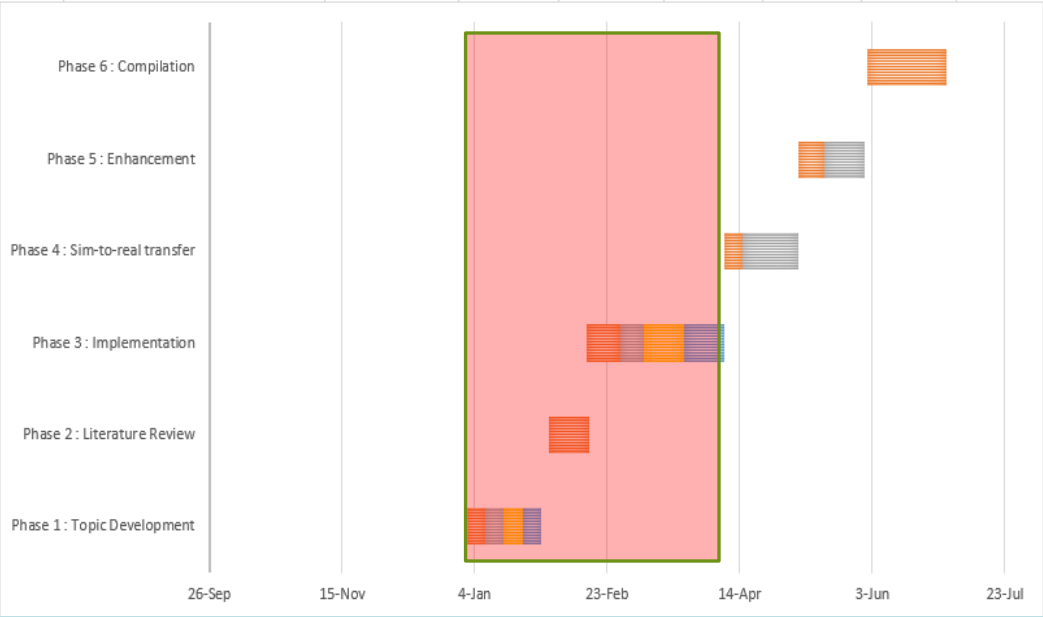
- Phase 3 : Implementation : 52 days (mid Feb- early Apr)

a) Setting up the Reinforcement Learning Platform and Simulation environment : 13 days

b) Prepare a custom implementation taking existing states, actions, rewards : 9 days

c) Redefine actions and rewards for our use case : 15 days

d) Test the pipeline and iterate : 15 days



sewts

THANK YOU