Multi-Agent Environment Standard

Assumption:

Each agent works synchronously.

Member Functions

reset()

reward_list, done = step(action_list)

obs_list = get_obs()

reward_list records the single step reward for each agent, it should be a list like [reward1, reward2,.....]. The length should be the same as the number of agents. Each element in the list should be a integer.

done True/False, mark when an episode finishes.

action_list records the single step action instruction for each agent, it should be a list like [action1, action2,.....]. The length should be the same as the number of agents. Each element in the list should be a non-negative integer.

obs_list records the single step observation for each agent, it should be a list like [obs1, obs2,.....]. The length should be the same as the number of agents. Each element in the list can be any form of data, but should be in same dimension, usually a list of variables or an image.

Typical Monte Carlo Procedures

```
# reset environment by calling reset()
# get initial observation get_obs()
for i in range(max_MC_iter):
    # get action_list from controller
    # apply action by step()
    # record returned reward list
    # record new observation by get_obs()
    if done:
        env.reset()
```