Multi-Agent Environment Standard

Assumption:

Each agent works synchronously.

Member Variable

self.env_type = 'periodic'/'episodic'

Member Functions

reset()

reward_list = step(action_list)

obs_list = get_obs()

reward_list records the single step reward for each agent

action_list records the single step action instruction for each agent

obs_list records the single step observation for each agent

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()
```