```
(pid=23257)         'agent_1': np.ndarray((2,), dtype=int64, min=12.0, max=12.0,
mean=12.0)},
(pid=23257)    2: { 'agent_0': np.ndarray((2,), dtype=int64, min=5.0, max=5.0,
mean=5.0),
(pid=23257)         'agent_1': np.ndarray((2,), dtype=int64, min=5.0, max=5.0,
mean=5.0)},
(pid=23257)    3: { 'agent_0': np.ndarray((2,), dtype=int64, min=8.0, max=8.0,
mean=8.0),
(pid=23257)         'agent_1': np.ndarray((2,), dtype=int64, min=8.0, max=8.0,
mean=8.0)},
(pid=23257)    4: { 'agent_0': np.ndarray((2,), dtype=int64, min=4.0, max=4.0,
mean=4.0),
(pid=23257)         'agent_1': np.ndarray((2,), dtype=int64, min=4.0, max=4.0,
mean=4.0)},
(pid=23257)    5: { 'agent_0': np.ndarray((2,), dtype=int64, min=7.0, max=7.0,
mean=7.0),
(pid=23257)         'agent_1': np.ndarray((2,), dtype=int64, min=7.0, max=7.0,
mean=7.0)},
(pid=23257)    6: { 'agent_0': np.ndarray((2,), dtype=int64, min=11.0, max=11.0,
mean=11.0),
(pid=23257)         'agent_1': np.ndarray((2,), dtype=int64, min=11.0, max=11.0,
mean=11.0)},
(pid=23257)    7: { 'agent_0': np.ndarray((2,), dtype=int64, min=14.0, max=14.0,
mean=14.0),
(pid=23257)         'agent_1': np.ndarray((2,), dtype=int64, min=14.0, max=14.0,
mean=14.0)}}
(pid=23257) 2019-10-05 11:32:58,026        INFO sampler.py:305 -- Info return from
env: { 0: {'agent_0': {}, 'agent_1': {}},
(pid=23257)    1: {'agent_0': {}, 'agent_1': {}},
(pid=23257)    2: {'agent_0': {}, 'agent_1': {}},
(pid=23257)    3: {'agent_0': {}, 'agent_1': {}},
(pid=23257)    4: {'agent_0': {}, 'agent_1': {}},
(pid=23257)    5: {'agent_0': {}, 'agent_1': {}},
(pid=23257)    6: {'agent_0': {}, 'agent_1': {}},
(pid=23257)    7: {'agent_0': {}, 'agent_1': {}}}
(pid=23257) 2019-10-05 11:32:58,026        INFO sampler.py:403 -- Preprocessed
obs: np.ndarray((2,), dtype=int64, min=12.0, max=12.0, mean=12.0)
(pid=23257) 2019-10-05 11:32:58,027        INFO sampler.py:407 -- Filtered obs:
np.ndarray((2,), dtype=float64, min=0.0, max=0.0, mean=0.0)
(pid=23257) 2019-10-05 11:32:58,034        INFO sampler.py:521 -- Inputs to
compute_actions():
(pid=23257)
(pid=23257) { 'agent_0': [ { 'data': { 'agent_id': 'agent_0',
(pid=23257)                            'env_id': 0,
(pid=23257)                            'info': {},
(pid=23257)                            'obs': np.ndarray((2,), dtype=float64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                            'prev_reward': 0.0,
(pid=23257)                            'rnn_state': []},
(pid=23257)                  'type': 'PolicyEvalData'},
(pid=23257)                { 'data': { 'agent_id': 'agent_0',
(pid=23257)                            'env_id': 1,
(pid=23257)                            'info': {},
(pid=23257)                            'obs': np.ndarray((2,), dtype=float64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
```

```
(pid=23257)                                 'prev_reward': 0.0,
(pid=23257)                                 'rnn_state': []},
(pid=23257)                   'type': 'PolicyEvalData'},
(pid=23257)                { 'data': { 'agent_id': 'agent_0',
(pid=23257)                            'env_id': 2,
(pid=23257)                            'info': {},
(pid=23257)                            'obs': np.ndarray((2,), dtype=float64,
min=-1.155, max=-1.155, mean=-1.155),
(pid=23257)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                            'prev_reward': 0.0,
(pid=23257)                            'rnn_state': []},
(pid=23257)                   'type': 'PolicyEvalData'},
(pid=23257)                { 'data': { 'agent_id': 'agent_0',
(pid=23257)                            'env_id': 3,
(pid=23257)                            'info': {},
(pid=23257)                            'obs': np.ndarray((2,), dtype=float64,
min=-0.367, max=-0.367, mean=-0.367),
(pid=23257)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                            'prev_reward': 0.0,
(pid=23257)                            'rnn_state': []},
(pid=23257)                   'type': 'PolicyEvalData'},
(pid=23257)                { 'data': { 'agent_id': 'agent_0',
(pid=23257)                            'env_id': 4,
(pid=23257)                            'info': {},
(pid=23257)                            'obs': np.ndarray((2,), dtype=float64,
min=-1.115, max=-1.115, mean=-1.115),
(pid=23257)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                            'prev_reward': 0.0,
(pid=23257)                            'rnn_state': []},
(pid=23257)                   'type': 'PolicyEvalData'},
(pid=23257)                { 'data': { 'agent_id': 'agent_0',
(pid=23257)                            'env_id': 5,
(pid=23257)                            'info': {},
(pid=23257)                            'obs': np.ndarray((2,), dtype=float64,
min=-0.294, max=-0.294, mean=-0.294),
(pid=23257)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                            'prev_reward': 0.0,
(pid=23257)                            'rnn_state': []},
(pid=23257)                   'type': 'PolicyEvalData'},
(pid=23257)                { 'data': { 'agent_id': 'agent_0',
(pid=23257)                            'env_id': 6,
(pid=23257)                            'info': {},
(pid=23257)                            'obs': np.ndarray((2,), dtype=float64,
min=0.777, max=0.777, mean=0.777),
(pid=23257)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                            'prev_reward': 0.0,
(pid=23257)                            'rnn_state': []},
(pid=23257)                   'type': 'PolicyEvalData'},
(pid=23257)                { 'data': { 'agent_id': 'agent_0',
(pid=23257)                            'env_id': 7,
(pid=23257)                            'info': {},
(pid=23257)                            'obs': np.ndarray((2,), dtype=float64,
min=1.338, max=1.338, mean=1.338),
(pid=23257)                            'prev_action': np.ndarray((), dtype=int64,
```

```
min=0.0, max=0.0, mean=0.0),
(pid=23257)                                 'prev_reward': 0.0,
(pid=23257)                                 'rnn_state': []},
(pid=23257)                       'type': 'PolicyEvalData'}],
(pid=23257)    'agent_1': [ { 'data': { 'agent_id': 'agent_1',
(pid=23257)                             'env_id': 0,
(pid=23257)                             'info': {},
(pid=23257)                             'obs': np.ndarray((2,), dtype=float64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                             'prev_reward': 0.0,
(pid=23257)                             'rnn_state': []},
(pid=23257)                       'type': 'PolicyEvalData'},
(pid=23257)                     { 'data': { 'agent_id': 'agent_1',
(pid=23257)                             'env_id': 1,
(pid=23257)                             'info': {},
(pid=23257)                             'obs': np.ndarray((2,), dtype=float64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                             'prev_reward': 0.0,
(pid=23257)                             'rnn_state': []},
(pid=23257)                       'type': 'PolicyEvalData'},
(pid=23257)                     { 'data': { 'agent_id': 'agent_1',
(pid=23257)                             'env_id': 2,
(pid=23257)                             'info': {},
(pid=23257)                             'obs': np.ndarray((2,), dtype=float64,
min=-1.155, max=-1.155, mean=-1.155),
(pid=23257)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                             'prev_reward': 0.0,
(pid=23257)                             'rnn_state': []},
(pid=23257)                       'type': 'PolicyEvalData'},
(pid=23257)                     { 'data': { 'agent_id': 'agent_1',
(pid=23257)                             'env_id': 3,
(pid=23257)                             'info': {},
(pid=23257)                             'obs': np.ndarray((2,), dtype=float64,
min=-0.367, max=-0.367, mean=-0.367),
(pid=23257)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                             'prev_reward': 0.0,
(pid=23257)                             'rnn_state': []},
(pid=23257)                       'type': 'PolicyEvalData'},
(pid=23257)                     { 'data': { 'agent_id': 'agent_1',
(pid=23257)                             'env_id': 4,
(pid=23257)                             'info': {},
(pid=23257)                             'obs': np.ndarray((2,), dtype=float64,
min=-1.115, max=-1.115, mean=-1.115),
(pid=23257)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                             'prev_reward': 0.0,
(pid=23257)                             'rnn_state': []},
(pid=23257)                       'type': 'PolicyEvalData'},
(pid=23257)                     { 'data': { 'agent_id': 'agent_1',
(pid=23257)                             'env_id': 5,
(pid=23257)                             'info': {},
(pid=23257)                             'obs': np.ndarray((2,), dtype=float64,
min=-0.294, max=-0.294, mean=-0.294),
```

```
(pid=23257)                               'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                               'prev_reward': 0.0,
(pid=23257)                               'rnn_state': []},
(pid=23257)                     'type': 'PolicyEvalData'},
(pid=23257)                   { 'data': { 'agent_id': 'agent_1',
(pid=23257)                               'env_id': 6,
(pid=23257)                               'info': {},
(pid=23257)                               'obs': np.ndarray((2,), dtype=float64,
min=0.777, max=0.777, mean=0.777),
(pid=23257)                               'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                               'prev_reward': 0.0,
(pid=23257)                               'rnn_state': []},
(pid=23257)                     'type': 'PolicyEvalData'},
(pid=23257)                   { 'data': { 'agent_id': 'agent_1',
(pid=23257)                               'env_id': 7,
(pid=23257)                               'info': {},
(pid=23257)                               'obs': np.ndarray((2,), dtype=float64,
min=1.338, max=1.338, mean=1.338),
(pid=23257)                               'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                               'prev_reward': 0.0,
(pid=23257)                               'rnn_state': []},
(pid=23257)                     'type': 'PolicyEvalData'}]}
(pid=23257)
(pid=23257) 2019-10-05 11:32:58,034        INFO tf_run_builder.py:92 -- Executing
TF run without tracing. To dump TF timeline traces to disk, set the
TF_TIMELINE_DIR environment variable.
(pid=23257) 2019-10-05 11:32:58,156        INFO sampler.py:548 -- Outputs of
compute_actions():
(pid=23257)
(pid=23257) { 'agent_0': ( np.ndarray((8,), dtype=int64, min=0.0, max=13.0,
mean=4.125),
(pid=23257)                []),
(pid=23257)                { 'q_values': np.ndarray((8, 15), dtype=float32,
min=-0.109, max=6.121, mean=1.271)}),
(pid=23257)   'agent_1': ( np.ndarray((8,), dtype=int64, min=0.0, max=12.0,
mean=7.125),
(pid=23257)                []),
(pid=23257)                { 'q_values': np.ndarray((8, 15), dtype=float32,
min=-1.62, max=1.373, mean=-0.018)})}
(pid=23257)
(pid=23257) 2019-10-05 11:32:58,596        INFO sample_batch_builder.py:161 --
Trajectory fragment after postprocess_trajectory():
(pid=23257)
(pid=23257) { 'agent_0': { 'data': { 'actions': np.ndarray((32,), dtype=int64,
min=2.0, max=14.0, mean=8.625),
(pid=23257)                          'agent_index': np.ndarray((32,), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                          'dones': np.ndarray((32,), dtype=bool,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                          'eps_id': np.ndarray((32,), dtype=int64,
min=651449133.0, max=651449133.0, mean=651449133.0),
(pid=23257)                          'infos': np.ndarray((32,), dtype=object,
head={'delta': 0.2992657094002690 3}),
(pid=23257)                          'new_obs': np.ndarray((32, 2), dtype=float32,
min=-1.613, max=1.621, mean=0.157),
(pid=23257)                          'obs': np.ndarray((32, 2), dtype=float32,
```

```
min=-1.613, max=1.621, mean=0.129),
(pid=23257)                             'prev_actions': np.ndarray((32,),
dtype=int64, min=0.0, max=14.0, mean=8.219),
(pid=23257)                             'prev_rewards': np.ndarray((32,),
dtype=float32, min=0.0, max=0.4, mean=0.271),
(pid=23257)                             'q_values': np.ndarray((32, 15),
dtype=float32, min=-1.492, max=6.059, mean=1.284),
(pid=23257)                             'rewards': np.ndarray((32,), dtype=float32,
min=0.149, max=1.108, mean=0.779),
(pid=23257)                             't': np.ndarray((32,), dtype=int64, min=0.0,
max=31.0, mean=15.5),
(pid=23257)                             'unroll_id': np.ndarray((32,), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                             'weights': np.ndarray((32,), dtype=float32,
min=2.298, max=3.194, mean=2.475)},
(pid=23257)                  'type': 'SampleBatch'},
(pid=23257)   'agent_1': { 'data': { 'actions': np.ndarray((32,), dtype=int64,
min=0.0, max=14.0, mean=7.219),
(pid=23257)                             'agent_index': np.ndarray((32,), dtype=int64,
min=1.0, max=1.0, mean=1.0),
(pid=23257)                             'dones': np.ndarray((32,), dtype=bool,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                             'eps_id': np.ndarray((32,), dtype=int64,
min=651449133.0, max=651449133.0, mean=651449133.0),
(pid=23257)                             'infos': np.ndarray((32,), dtype=object,
head={'delta': 1.0751406043285947}),
(pid=23257)                             'new_obs': np.ndarray((32, 2), dtype=float32,
min=-1.613, max=1.621, mean=0.157),
(pid=23257)                             'obs': np.ndarray((32, 2), dtype=float32,
min=-1.613, max=1.621, mean=0.129),
(pid=23257)                             'prev_actions': np.ndarray((32,),
dtype=int64, min=0.0, max=14.0, mean=7.125),
(pid=23257)                             'prev_rewards': np.ndarray((32,),
dtype=float32, min=0.0, max=0.42, mean=0.295),
(pid=23257)                             'q_values': np.ndarray((32, 15),
dtype=float32, min=-2.096, max=1.617, mean=-0.192),
(pid=23257)                             'rewards': np.ndarray((32,), dtype=float32,
min=0.391, max=1.033, mean=0.872),
(pid=23257)                             't': np.ndarray((32,), dtype=int64, min=0.0,
max=31.0, mean=15.5),
(pid=23257)                             'unroll_id': np.ndarray((32,), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=23257)                             'weights': np.ndarray((32,), dtype=float32,
min=2.303, max=2.57, mean=2.402)},
(pid=23257)                  'type': 'SampleBatch'}}
(pid=23257)
(pid=23257) 2019-10-05 11:32:58,661        INFO rollout_worker.py:485 -- Completed
sample batch:
(pid=23257)
(pid=23257) { 'count': 256,
(pid=23257)   'policy_batches': { 'agent_0': { 'data': { 'actions':
np.ndarray((256,), dtype=int64, min=0.0, max=14.0, mean=7.332),
(pid=23257)                                              'agent_index':
np.ndarray((256,), dtype=int64, min=0.0, max=0.0, mean=0.0),
(pid=23257)                                              'dones':
np.ndarray((256,), dtype=bool, min=0.0, max=0.0, mean=0.0),
(pid=23257)                                              'eps_id':
np.ndarray((256,), dtype=int64, min=52064235.0, max=1699625468.0,
mean=788479707.75),
```

```
(pid=23257)                                                 'infos':
np.ndarray((256,), dtype=object, head={'delta': 0.29926570940026903}),
(pid=23257)                                                 'new_obs':
np.ndarray((256, 2), dtype=float32, min=-1.917, max=1.646, mean=-0.036),
(pid=23257)                                                 'obs': np.ndarray((256,
2), dtype=float32, min=-1.917, max=1.646, mean=-0.058),
(pid=23257)                                                 'prev_actions':
np.ndarray((256,), dtype=int64, min=0.0, max=14.0, mean=7.051),
(pid=23257)                                                 'prev_rewards':
np.ndarray((256,), dtype=float32, min=0.0, max=0.42, mean=0.266),
(pid=23257)                                                 'q_values':
np.ndarray((256, 15), dtype=float32, min=-1.492, max=6.397, mean=1.248),
(pid=23257)                                                 'rewards':
np.ndarray((256,), dtype=float32, min=0.137, max=1.136, mean=0.773),
(pid=23257)                                                 't': np.ndarray((256,),
dtype=int64, min=0.0, max=31.0, mean=15.5),
(pid=23257)                                                 'unroll_id':
np.ndarray((256,), dtype=int64, min=0.0, max=0.0, mean=0.0),
(pid=23257)                                                 'weights':
np.ndarray((256,), dtype=float32, min=2.285, max=3.248, mean=2.459)},
(pid=23257)                                       'type': 'SampleBatch'},
(pid=23257)                           'agent_1': { 'data': { 'actions':
np.ndarray((256,), dtype=int64, min=0.0, max=14.0, mean=6.75),
(pid=23257)                                                 'agent_index':
np.ndarray((256,), dtype=int64, min=1.0, max=1.0, mean=1.0),
(pid=23257)                                                 'dones':
np.ndarray((256,), dtype=bool, min=0.0, max=0.0, mean=0.0),
(pid=23257)                                                 'eps_id':
np.ndarray((256,), dtype=int64, min=52064235.0, max=1699625468.0,
mean=788479707.75),
(pid=23257)                                                 'infos':
np.ndarray((256,), dtype=object, head={'delta': 1.0751406043285947}),
(pid=23257)                                                 'new_obs':
np.ndarray((256, 2), dtype=float32, min=-1.917, max=1.646, mean=-0.036),
(pid=23257)                                                 'obs': np.ndarray((256,
2), dtype=float32, min=-1.917, max=1.646, mean=-0.058),
(pid=23257)                                                 'prev_actions':
np.ndarray((256,), dtype=int64, min=0.0, max=14.0, mean=6.586),
(pid=23257)                                                 'prev_rewards':
np.ndarray((256,), dtype=float32, min=0.0, max=0.42, mean=0.278),
(pid=23257)                                                 'q_values':
np.ndarray((256, 15), dtype=float32, min=-2.101, max=2.057, mean=-0.175),
(pid=23257)                                                 'rewards':
np.ndarray((256,), dtype=float32, min=0.232, max=1.141, mean=0.82),
(pid=23257)                                                 't': np.ndarray((256,),
dtype=int64, min=0.0, max=31.0, mean=15.5),
(pid=23257)                                                 'unroll_id':
np.ndarray((256,), dtype=int64, min=0.0, max=0.0, mean=0.0),
(pid=23257)                                                 'weights':
np.ndarray((256,), dtype=float32, min=2.241, max=2.721, mean=2.407)},
(pid=23257)                                       'type': 'SampleBatch'}},
(pid=23257)   'type': 'MultiAgentBatch'}
(pid=23257)
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorflow/python/framework/dtypes.py:516: FutureWarning: Passing (type, 1) or
'1type' as a synonym of type is deprecated; in a future version of numpy, it will
be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   _np_qint8 = np.dtype([("qint8", np.int8, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
```

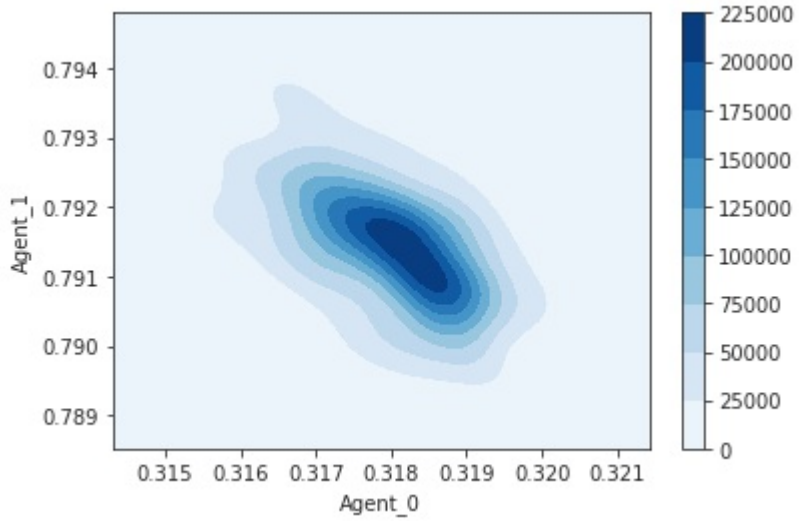
```
tensorflow/python/framework/dtypes.py:517: FutureWarning: Passing (type, 1) or
'1type' as a synonym of type is deprecated; in a future version of numpy, it will
be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   _np_quint8 = np.dtype([("quint8", np.uint8, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorflow/python/framework/dtypes.py:518: FutureWarning: Passing (type, 1) or
'1type' as a synonym of type is deprecated; in a future version of numpy, it will
be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   _np_qint16 = np.dtype([("qint16", np.int16, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorflow/python/framework/dtypes.py:519: FutureWarning: Passing (type, 1) or
'1type' as a synonym of type is deprecated; in a future version of numpy, it will
be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   _np_quint16 = np.dtype([("quint16", np.uint16, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorflow/python/framework/dtypes.py:520: FutureWarning: Passing (type, 1) or
'1type' as a synonym of type is deprecated; in a future version of numpy, it will
be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   _np_qint32 = np.dtype([("qint32", np.int32, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorflow/python/framework/dtypes.py:525: FutureWarning: Passing (type, 1) or
'1type' as a synonym of type is deprecated; in a future version of numpy, it will
be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   np_resource = np.dtype([("resource", np.ubyte, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorboard/compat/tensorflow_stub/dtypes.py:541: FutureWarning: Passing (type, 1)
or '1type' as a synonym of type is deprecated; in a future version of numpy, it
will be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   _np_qint8 = np.dtype([("qint8", np.int8, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorboard/compat/tensorflow_stub/dtypes.py:542: FutureWarning: Passing (type, 1)
or '1type' as a synonym of type is deprecated; in a future version of numpy, it
will be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   _np_quint8 = np.dtype([("quint8", np.uint8, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorboard/compat/tensorflow_stub/dtypes.py:543: FutureWarning: Passing (type, 1)
or '1type' as a synonym of type is deprecated; in a future version of numpy, it
will be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   _np_qint16 = np.dtype([("qint16", np.int16, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorboard/compat/tensorflow_stub/dtypes.py:544: FutureWarning: Passing (type, 1)
or '1type' as a synonym of type is deprecated; in a future version of numpy, it
will be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   _np_quint16 = np.dtype([("quint16", np.uint16, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorboard/compat/tensorflow_stub/dtypes.py:545: FutureWarning: Passing (type, 1)
or '1type' as a synonym of type is deprecated; in a future version of numpy, it
will be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   _np_qint32 = np.dtype([("qint32", np.int32, 1)])
(pid=23269) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
tensorboard/compat/tensorflow_stub/dtypes.py:550: FutureWarning: Passing (type, 1)
or '1type' as a synonym of type is deprecated; in a future version of numpy, it
will be understood as (type, (1,)) / '(1,)type'.
(pid=23269)   np_resource = np.dtype([("resource", np.ubyte, 1)])
(pid=23269) WARNING:tensorflow:From /home/lorenzo/anaconda3/envs/py36/lib/
python3.6/site-packages/tensorflow/python/compat/v2_compat.py:61:
disable_resource_variables (from tensorflow.python.ops.variable_scope) is
deprecated and will be removed in a future version.
(pid=23269) Instructions for updating:
```

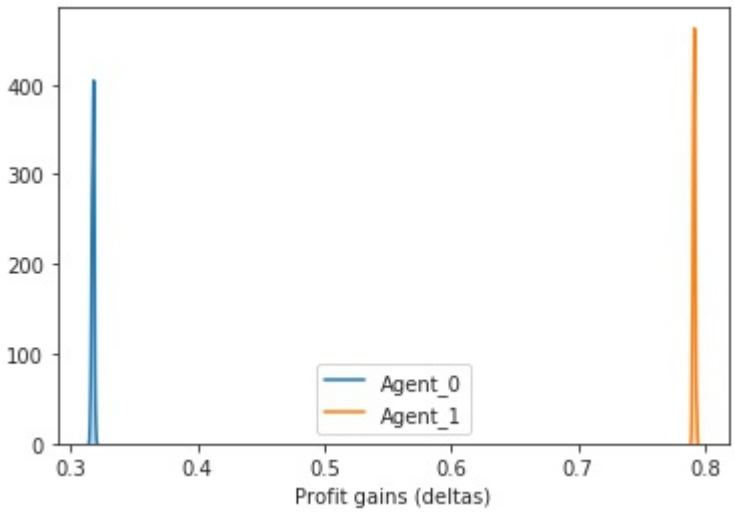
```
(pid=23269) non-resource variables are not supported in the long term
Episode reward 5756.591028063981
Episode 1 of 100
Episode reward 5757.2320136274
Episode 2 of 100
Episode reward 5755.441213696364
Episode 3 of 100
Episode reward 5757.788105240648
Episode 4 of 100
Episode reward 5754.058689571482
Episode 5 of 100
Episode reward 5755.403089063919
Episode 6 of 100
Episode reward 5756.154742480864
Episode 7 of 100
Episode reward 5757.372716448058
Episode 8 of 100
Episode reward 5754.004053263244
Episode 9 of 100
Episode reward 5755.381963273251
Episode 10 of 100
Episode reward 5755.47683240248
Episode 11 of 100
Episode reward 5755.46084621738
Episode 12 of 100
Episode reward 5755.962758699215
Episode 13 of 100
Episode reward 5755.047163117575
Episode 14 of 100
Episode reward 5755.922069007039
Episode 15 of 100
Episode reward 5756.352924833054
Episode 16 of 100
Episode reward 5755.923780946597
Episode 17 of 100
Episode reward 5756.2237614549995
Episode 18 of 100
Episode reward 5755.659961239005
Episode 19 of 100
Episode reward 5755.364273857826
Episode 20 of 100
Episode reward 5756.211308910293
Episode 21 of 100
Episode reward 5754.801403476093
Episode 22 of 100
Episode reward 5756.513470151612
Episode 23 of 100
Episode reward 5757.164349822371
Episode 24 of 100
Episode reward 5755.094592094573
Episode 25 of 100
Episode reward 5755.178837302949
Episode 26 of 100
Episode reward 5757.215989418091
Episode 27 of 100
Episode reward 5755.044183407146
Episode 28 of 100
Episode reward 5755.583273798522
Episode 29 of 100
```

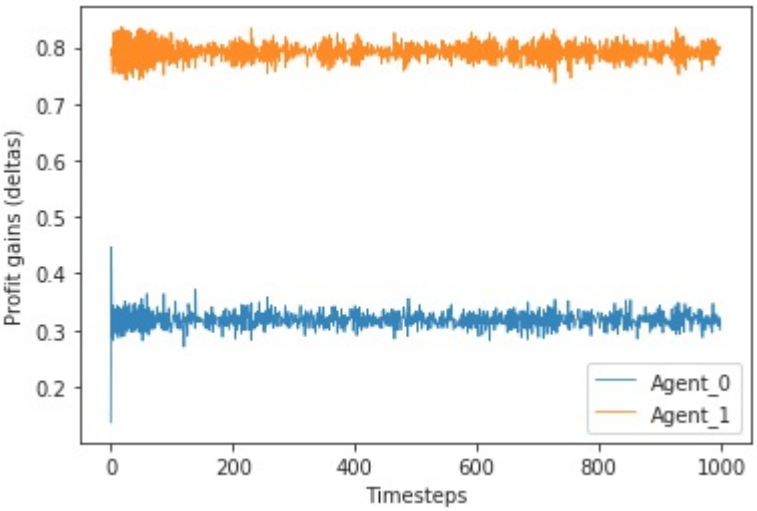
```
Episode reward 5755.597924524774
Episode 30 of 100
Episode reward 5755.561009560341
Episode 31 of 100
Episode reward 5753.722606470515
Episode 32 of 100
Episode reward 5755.369791550584
Episode 33 of 100
Episode reward 5755.578655083336
Episode 34 of 100
Episode reward 5755.657206171462
Episode 35 of 100
Episode reward 5755.305476997377
Episode 36 of 100
Episode reward 5756.113326440645
Episode 37 of 100
Episode reward 5755.599194327526
Episode 38 of 100
Episode reward 5755.775411305665
Episode 39 of 100
Episode reward 5755.627837473114
Episode 40 of 100
Episode reward 5756.825493040384
Episode 41 of 100
Episode reward 5754.375866736508
Episode 42 of 100
Episode reward 5755.80475347983
Episode 43 of 100
Episode reward 5754.549921454246
Episode 44 of 100
Episode reward 5755.058313665336
Episode 45 of 100
Episode reward 5756.252980157725
Episode 46 of 100
Episode reward 5755.069113210044
Episode 47 of 100
Episode reward 5757.051663266328
Episode 48 of 100
Episode reward 5753.874704713228
Episode 49 of 100
Episode reward 5754.055986620063
Episode 50 of 100
Episode reward 5756.240711603016
Episode 51 of 100
Episode reward 5756.502892881303
Episode 52 of 100
Episode reward 5754.515564496166
Episode 53 of 100
Episode reward 5754.40026499055
Episode 54 of 100
Episode reward 5754.548065693165
Episode 55 of 100
Episode reward 5755.9776293580235
Episode 56 of 100
Episode reward 5755.116907920595
Episode 57 of 100
Episode reward 5753.569507893472
Episode 58 of 100
Episode reward 5755.478283156733
```

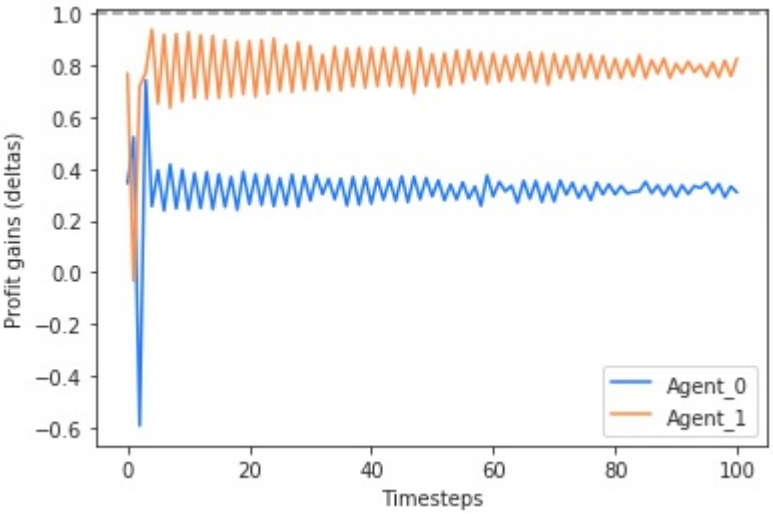
```
Episode 59 of 100
Episode reward 5755.229939630534
Episode 60 of 100
Episode reward 5755.207243787886
Episode 61 of 100
Episode reward 5755.529177583095
Episode 62 of 100
Episode reward 5756.138235868967
Episode 63 of 100
Episode reward 5755.208928334781
Episode 64 of 100
Episode reward 5754.823700171943
Episode 65 of 100
Episode reward 5757.804253749209
Episode 66 of 100
Episode reward 5756.63519773626
Episode 67 of 100
Episode reward 5755.939808093425
Episode 68 of 100
Episode reward 5757.320029568055
Episode 69 of 100
Episode reward 5756.408075394918
Episode 70 of 100
Episode reward 5756.0680448127605
Episode 71 of 100
Episode reward 5754.352695112289
Episode 72 of 100
Episode reward 5756.613508534132
Episode 73 of 100
Episode reward 5756.14882227798
Episode 74 of 100
Episode reward 5755.985934445137
Episode 75 of 100
Episode reward 5756.68530212155
Episode 76 of 100
Episode reward 5754.049034831124
Episode 77 of 100
Episode reward 5756.735449857586
Episode 78 of 100
Episode reward 5753.452870287662
Episode 79 of 100
Episode reward 5755.971840578874
Episode 80 of 100
Episode reward 5756.513361803005
Episode 81 of 100
Episode reward 5755.644411269557
Episode 82 of 100
Episode reward 5755.496198612798
Episode 83 of 100
Episode reward 5756.949814730929
Episode 84 of 100
Episode reward 5755.924044679131
Episode 85 of 100
Episode reward 5755.940863644632
Episode 86 of 100
Episode reward 5755.697733538368
Episode 87 of 100
Episode reward 5756.388179624725
Episode 88 of 100
```

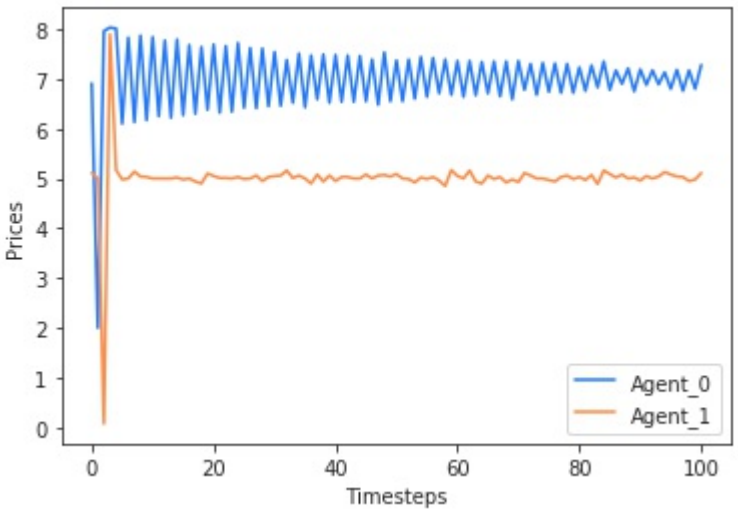
```
Episode reward 5754.692915054598
Episode 89 of 100
Episode reward 5754.6575464529815
Episode 90 of 100
Episode reward 5756.094514529053
Episode 91 of 100
Episode reward 5756.319697409377
Episode 92 of 100
Episode reward 5756.208341356544
Episode 93 of 100
Episode reward 5755.2923337564625
Episode 94 of 100
Episode reward 5752.601745609122
Episode 95 of 100
Episode reward 5754.777558120673
Episode 96 of 100
Episode reward 5755.8477291179925
Episode 97 of 100
Episode reward 5756.371727627003
Episode 98 of 100
Episode reward 5755.723874774537
Episode 99 of 100
Episode reward 5757.0102657515245
Overall deltas mean: 0.5547 and std: 0.2367
Agent0 deltas mean: 0.3179 and std: 0.0009
Agent1 deltas mean: 0.7914 and std: 0.0008
/home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/seaborn/
distributions.py:679: UserWarning: Passing a 2D dataset for a bivariate plot is
deprecated in favor of kdeplot(x, y), and it will cause an error in future
versions. Please update your code.
  warnings.warn(warn_msg, UserWarning)
```




400
300
200
100
0
0.3
0.4
0.5
0.6
0.7
0.8
Agent_0
Agent_1
Profit gains (deltas)


0.8
0.7
0.6
0.5
0.4
0.3
0.2
Profit gains (deltas)
0
200
400
600
800
1000
Timesteps
Agent_0
Agent_1


1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
Profit gains (deltas)
0
20
40
60
80
100
Timesteps
Agent_0
Agent_1




```
Traceback (most recent call last):

  File "<ipython-input-1-b9d5ffeb4c0f>", line 1, in <module>
    runfile('/home/lorenzo/algorithmic-pricing/rollout/rollout.py', args='/home/
lorenzo/algorithmic-pricing/train_results/Azure_ApexDQN_Disc/azure_disc_10_res2/
APEX_MultiAgentFirmsPricing_0_2019-09-21_14-03-453aaxi5sb/checkpoint_940/
checkpoint-940 --run APEX --env env_disc', wdir='/home/lorenzo/algorithmic-
pricing/rollout')

  File "/home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
spyder_kernels/customize/spydercustomize.py", line 827, in runfile
    execfile(filename, namespace)

  File "/home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
spyder_kernels/customize/spydercustomize.py", line 110, in execfile
    exec(compile(f.read(), filename, 'exec'), namespace)

  File "/home/lorenzo/algorithmic-pricing/rollout/rollout.py", line 404, in
<module>
    Deltas_df = pd.DataFrame(d_array)

  File "/home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/pandas/core/
frame.py", line 440, in __init__
    mgr = init_ndarray(data, index, columns, dtype=dtype, copy=copy)

  File "/home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/pandas/core/
internals/construction.py", line 171, in init_ndarray
    values = prep_ndarray(values, copy=copy)

  File "/home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/pandas/core/
internals/construction.py", line 295, in prep_ndarray
    raise ValueError("Must pass 2-d input")

ValueError: Must pass 2-d input


In [2]:

In [2]:
```