```
(pid=25726)         'agent_1': np.ndarray((2,), dtype=float16, min=1.236,
max=1.236, mean=1.236)},
(pid=25726)    2: { 'agent_0': np.ndarray((2,), dtype=float16, min=1.556,
max=1.556, mean=1.556),
(pid=25726)         'agent_1': np.ndarray((2,), dtype=float16, min=1.556,
max=1.556, mean=1.556)},
(pid=25726)    3: { 'agent_0': np.ndarray((2,), dtype=float16, min=1.567,
max=1.567, mean=1.567),
(pid=25726)         'agent_1': np.ndarray((2,), dtype=float16, min=1.567,
max=1.567, mean=1.567)},
(pid=25726)    4: { 'agent_0': np.ndarray((2,), dtype=float16, min=1.633,
max=1.633, mean=1.633),
(pid=25726)         'agent_1': np.ndarray((2,), dtype=float16, min=1.633,
max=1.633, mean=1.633)},
(pid=25726)    5: { 'agent_0': np.ndarray((2,), dtype=float16, min=1.9, max=1.9,
mean=1.9),
(pid=25726)         'agent_1': np.ndarray((2,), dtype=float16, min=1.9, max=1.9,
mean=1.9)},
(pid=25726)    6: { 'agent_0': np.ndarray((2,), dtype=float16, min=1.571,
max=1.571, mean=1.571),
(pid=25726)         'agent_1': np.ndarray((2,), dtype=float16, min=1.571,
max=1.571, mean=1.571)},
(pid=25726)    7: { 'agent_0': np.ndarray((2,), dtype=float16, min=1.619,
max=1.619, mean=1.619),
(pid=25726)         'agent_1': np.ndarray((2,), dtype=float16, min=1.619,
max=1.619, mean=1.619)}}
(pid=25726) 2019-09-21 14:06:30,646        INFO sampler.py:305 -- Info return from
env: { 0: {'agent_0': {}, 'agent_1': {}},
(pid=25726)    1: {'agent_0': {}, 'agent_1': {}},
(pid=25726)    2: {'agent_0': {}, 'agent_1': {}},
(pid=25726)    3: {'agent_0': {}, 'agent_1': {}},
(pid=25726)    4: {'agent_0': {}, 'agent_1': {}},
(pid=25726)    5: {'agent_0': {}, 'agent_1': {}},
(pid=25726)    6: {'agent_0': {}, 'agent_1': {}},
(pid=25726)    7: {'agent_0': {}, 'agent_1': {}}}
(pid=25726) 2019-09-21 14:06:30,646        INFO sampler.py:403 -- Preprocessed
obs: np.ndarray((2,), dtype=float16, min=1.339, max=1.339, mean=1.339)
(pid=25726) 2019-09-21 14:06:30,647        INFO sampler.py:407 -- Filtered obs:
np.ndarray((2,), dtype=float64, min=0.0, max=0.0, mean=0.0)
(pid=25726) 2019-09-21 14:06:30,655        INFO sampler.py:521 -- Inputs to
compute_actions():
(pid=25726)
(pid=25726) { 'agent_0': [ { 'data': { 'agent_id': 'agent_0',
(pid=25726)                            'env_id': 0,
(pid=25726)                            'info': {},
(pid=25726)                            'obs': np.ndarray((2,), dtype=float64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                            'prev_reward': 0.0,
(pid=25726)                            'rnn_state': []},
(pid=25726)                  'type': 'PolicyEvalData'},
(pid=25726)                { 'data': { 'agent_id': 'agent_0',
(pid=25726)                            'env_id': 1,
(pid=25726)                            'info': {},
(pid=25726)                            'obs': np.ndarray((2,), dtype=float64,
min=-0.707, max=-0.707, mean=-0.707),
(pid=25726)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
```

```
(pid=25726)                                 'prev_reward': 0.0,
(pid=25726)                                 'rnn_state': []},
(pid=25726)                   'type': 'PolicyEvalData'},
(pid=25726)                { 'data': { 'agent_id': 'agent_0',
(pid=25726)                            'env_id': 2,
(pid=25726)                            'info': {},
(pid=25726)                            'obs': np.ndarray((2,), dtype=float64,
min=1.096, max=1.096, mean=1.096),
(pid=25726)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                            'prev_reward': 0.0,
(pid=25726)                            'rnn_state': []},
(pid=25726)                   'type': 'PolicyEvalData'},
(pid=25726)                { 'data': { 'agent_id': 'agent_0',
(pid=25726)                            'env_id': 3,
(pid=25726)                            'info': {},
(pid=25726)                            'obs': np.ndarray((2,), dtype=float64,
min=0.873, max=0.873, mean=0.873),
(pid=25726)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                            'prev_reward': 0.0,
(pid=25726)                            'rnn_state': []},
(pid=25726)                   'type': 'PolicyEvalData'},
(pid=25726)                { 'data': { 'agent_id': 'agent_0',
(pid=25726)                            'env_id': 4,
(pid=25726)                            'info': {},
(pid=25726)                            'obs': np.ndarray((2,), dtype=float64,
min=0.982, max=0.982, mean=0.982),
(pid=25726)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                            'prev_reward': 0.0,
(pid=25726)                            'rnn_state': []},
(pid=25726)                   'type': 'PolicyEvalData'},
(pid=25726)                { 'data': { 'agent_id': 'agent_0',
(pid=25726)                            'env_id': 5,
(pid=25726)                            'info': {},
(pid=25726)                            'obs': np.ndarray((2,), dtype=float64,
min=1.551, max=1.551, mean=1.551),
(pid=25726)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                            'prev_reward': 0.0,
(pid=25726)                            'rnn_state': []},
(pid=25726)                   'type': 'PolicyEvalData'},
(pid=25726)                { 'data': { 'agent_id': 'agent_0',
(pid=25726)                            'env_id': 6,
(pid=25726)                            'info': {},
(pid=25726)                            'obs': np.ndarray((2,), dtype=float64,
min=0.131, max=0.131, mean=0.131),
(pid=25726)                            'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                            'prev_reward': 0.0,
(pid=25726)                            'rnn_state': []},
(pid=25726)                   'type': 'PolicyEvalData'},
(pid=25726)                { 'data': { 'agent_id': 'agent_0',
(pid=25726)                            'env_id': 7,
(pid=25726)                            'info': {},
(pid=25726)                            'obs': np.ndarray((2,), dtype=float64,
min=0.333, max=0.333, mean=0.333),
(pid=25726)                            'prev_action': np.ndarray((), dtype=int64,
```

```
min=0.0, max=0.0, mean=0.0),
(pid=25726)                                 'prev_reward': 0.0,
(pid=25726)                                 'rnn_state': []},
(pid=25726)                       'type': 'PolicyEvalData'}],
(pid=25726)    'agent_1': [ { 'data': { 'agent_id': 'agent_1',
(pid=25726)                             'env_id': 0,
(pid=25726)                             'info': {},
(pid=25726)                             'obs': np.ndarray((2,), dtype=float64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'prev_reward': 0.0,
(pid=25726)                             'rnn_state': []},
(pid=25726)                       'type': 'PolicyEvalData'},
(pid=25726)                     { 'data': { 'agent_id': 'agent_1',
(pid=25726)                             'env_id': 1,
(pid=25726)                             'info': {},
(pid=25726)                             'obs': np.ndarray((2,), dtype=float64,
min=-0.707, max=-0.707, mean=-0.707),
(pid=25726)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'prev_reward': 0.0,
(pid=25726)                             'rnn_state': []},
(pid=25726)                       'type': 'PolicyEvalData'},
(pid=25726)                     { 'data': { 'agent_id': 'agent_1',
(pid=25726)                             'env_id': 2,
(pid=25726)                             'info': {},
(pid=25726)                             'obs': np.ndarray((2,), dtype=float64,
min=1.096, max=1.096, mean=1.096),
(pid=25726)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'prev_reward': 0.0,
(pid=25726)                             'rnn_state': []},
(pid=25726)                       'type': 'PolicyEvalData'},
(pid=25726)                     { 'data': { 'agent_id': 'agent_1',
(pid=25726)                             'env_id': 3,
(pid=25726)                             'info': {},
(pid=25726)                             'obs': np.ndarray((2,), dtype=float64,
min=0.873, max=0.873, mean=0.873),
(pid=25726)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'prev_reward': 0.0,
(pid=25726)                             'rnn_state': []},
(pid=25726)                       'type': 'PolicyEvalData'},
(pid=25726)                     { 'data': { 'agent_id': 'agent_1',
(pid=25726)                             'env_id': 4,
(pid=25726)                             'info': {},
(pid=25726)                             'obs': np.ndarray((2,), dtype=float64,
min=0.982, max=0.982, mean=0.982),
(pid=25726)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'prev_reward': 0.0,
(pid=25726)                             'rnn_state': []},
(pid=25726)                       'type': 'PolicyEvalData'},
(pid=25726)                     { 'data': { 'agent_id': 'agent_1',
(pid=25726)                             'env_id': 5,
(pid=25726)                             'info': {},
(pid=25726)                             'obs': np.ndarray((2,), dtype=float64,
min=1.551, max=1.551, mean=1.551),
```

```
(pid=25726)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'prev_reward': 0.0,
(pid=25726)                             'rnn_state': []},
(pid=25726)                   'type': 'PolicyEvalData'},
(pid=25726)                 { 'data': { 'agent_id': 'agent_1',
(pid=25726)                             'env_id': 6,
(pid=25726)                             'info': {},
(pid=25726)                             'obs': np.ndarray((2,), dtype=float64,
min=0.131, max=0.131, mean=0.131),
(pid=25726)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'prev_reward': 0.0,
(pid=25726)                             'rnn_state': []},
(pid=25726)                   'type': 'PolicyEvalData'},
(pid=25726)                 { 'data': { 'agent_id': 'agent_1',
(pid=25726)                             'env_id': 7,
(pid=25726)                             'info': {},
(pid=25726)                             'obs': np.ndarray((2,), dtype=float64,
min=0.333, max=0.333, mean=0.333),
(pid=25726)                             'prev_action': np.ndarray((), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'prev_reward': 0.0,
(pid=25726)                             'rnn_state': []},
(pid=25726)                   'type': 'PolicyEvalData'}]}
(pid=25726)
(pid=25726) 2019-09-21 14:06:30,656        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=25726) 2019-09-21 14:06:30,781        INFO sampler.py:548 -- Outputs of
compute_actions():
(pid=25726)
(pid=25726) { 'agent_0': ( np.ndarray((8,), dtype=int64, min=1.0, max=4.0,
mean=1.75),
(pid=25726)                 [],
(pid=25726)                 { 'q_values': np.ndarray((8, 5), dtype=float32,
min=-0.149, max=3.984, mean=1.522)}),
(pid=25726)   'agent_1': ( np.ndarray((8,), dtype=int64, min=0.0, max=3.0,
mean=1.375),
(pid=25726)                 [],
(pid=25726)                 { 'q_values': np.ndarray((8, 5), dtype=float32,
min=-2.941, max=1.322, mean=-0.176)})}
(pid=25726)
(pid=25726) 2019-09-21 14:06:31,153        INFO sample_batch_builder.py:161 --
Trajectory fragment after postprocess_trajectory():
(pid=25726)
(pid=25726) { 'agent_0': { 'data': { 'actions': np.ndarray((32,), dtype=int64,
min=0.0, max=4.0, mean=1.438),
(pid=25726)                          'agent_index': np.ndarray((32,), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                          'dones': np.ndarray((32,), dtype=bool,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                          'eps_id': np.ndarray((32,), dtype=int64,
min=144169260.0, max=144169260.0, mean=144169260.0),
(pid=25726)                          'infos': np.ndarray((32,), dtype=object,
head={'delta': -0.711568527438792}),
(pid=25726)                          'new_obs': np.ndarray((32, 2), dtype=float32,
min=-1.521, max=1.532, mean=-0.256),
(pid=25726)                          'obs': np.ndarray((32, 2), dtype=float32,
```

```
min=-1.521, max=1.532, mean=-0.257),
(pid=25726)                             'prev_actions': np.ndarray((32,),
dtype=int64, min=0.0, max=4.0, mean=1.406),
(pid=25726)                             'prev_rewards': np.ndarray((32,),
dtype=float32, min=0.0, max=0.303, mean=0.22),
(pid=25726)                             'q_values': np.ndarray((32, 5),
dtype=float32, min=-0.306, max=2.788, mean=0.637),
(pid=25726)                             'rewards': np.ndarray((32,), dtype=float32,
min=0.213, max=0.864, mean=0.647),
(pid=25726)                             't': np.ndarray((32,), dtype=int64, min=0.0,
max=31.0, mean=15.5),
(pid=25726)                             'unroll_id': np.ndarray((32,), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'weights': np.ndarray((32,), dtype=float32,
min=2.303, max=2.665, mean=2.42)},
(pid=25726)                  'type': 'SampleBatch'},
(pid=25726)   'agent_1': { 'data': { 'actions': np.ndarray((32,), dtype=int64,
min=0.0, max=4.0, mean=2.094),
(pid=25726)                             'agent_index': np.ndarray((32,), dtype=int64,
min=1.0, max=1.0, mean=1.0),
(pid=25726)                             'dones': np.ndarray((32,), dtype=bool,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'eps_id': np.ndarray((32,), dtype=int64,
min=144169260.0, max=144169260.0, mean=144169260.0),
(pid=25726)                             'infos': np.ndarray((32,), dtype=object,
head={'delta': -0.6810216555029001}),
(pid=25726)                             'new_obs': np.ndarray((32, 2), dtype=float32,
min=-1.521, max=1.532, mean=-0.256),
(pid=25726)                             'obs': np.ndarray((32, 2), dtype=float32,
min=-1.521, max=1.532, mean=-0.257),
(pid=25726)                             'prev_actions': np.ndarray((32,),
dtype=int64, min=0.0, max=4.0, mean=2.062),
(pid=25726)                             'prev_rewards': np.ndarray((32,),
dtype=float32, min=0.0, max=0.204, mean=0.128),
(pid=25726)                             'q_values': np.ndarray((32, 5),
dtype=float32, min=-3.572, max=0.059, mean=-2.037),
(pid=25726)                             'rewards': np.ndarray((32,), dtype=float32,
min=0.168, max=0.52, mean=0.376),
(pid=25726)                             't': np.ndarray((32,), dtype=int64, min=0.0,
max=31.0, mean=15.5),
(pid=25726)                             'unroll_id': np.ndarray((32,), dtype=int64,
min=0.0, max=0.0, mean=0.0),
(pid=25726)                             'weights': np.ndarray((32,), dtype=float32,
min=2.289, max=2.539, mean=2.428)},
(pid=25726)                  'type': 'SampleBatch'}}
(pid=25726)
(pid=25726) 2019-09-21 14:06:31,207        INFO rollout_worker.py:485 -- Completed
sample batch:
(pid=25726)
(pid=25726) { 'count': 256,
(pid=25726)   'policy_batches': { 'agent_0': { 'data': { 'actions':
np.ndarray((256,), dtype=int64, min=0.0, max=4.0, mean=1.859),
(pid=25726)                                              'agent_index':
np.ndarray((256,), dtype=int64, min=0.0, max=0.0, mean=0.0),
(pid=25726)                                              'dones':
np.ndarray((256,), dtype=bool, min=0.0, max=0.0, mean=0.0),
(pid=25726)                                              'eps_id':
np.ndarray((256,), dtype=int64, min=144169260.0, max=1483390642.0,
mean=717353850.125),
```

```
(pid=25726)                                                 'infos':
np.ndarray((256,), dtype=object, head={'delta': -0.711568527438792}),
(pid=25726)                                                 'new_obs':
np.ndarray((256, 2), dtype=float32, min=-1.989, max=2.683, mean=-0.059),
(pid=25726)                                                 'obs': np.ndarray((256,
2), dtype=float32, min=-1.989, max=2.683, mean=-0.03),
(pid=25726)                                                 'prev_actions':
np.ndarray((256,), dtype=int64, min=0.0, max=4.0, mean=1.816),
(pid=25726)                                                 'prev_rewards':
np.ndarray((256,), dtype=float32, min=0.0, max=0.437, mean=0.206),
(pid=25726)                                                 'q_values':
np.ndarray((256, 5), dtype=float32, min=-0.416, max=4.217, mean=1.321),
(pid=25726)                                                 'rewards':
np.ndarray((256,), dtype=float32, min=0.108, max=1.277, mean=0.599),
(pid=25726)                                                 't': np.ndarray((256,),
dtype=int64, min=0.0, max=31.0, mean=15.5),
(pid=25726)                                                 'unroll_id':
np.ndarray((256,), dtype=int64, min=0.0, max=0.0, mean=0.0),
(pid=25726)                                                 'weights':
np.ndarray((256,), dtype=float32, min=2.279, max=2.793, mean=2.45)},
(pid=25726)                                       'type': 'SampleBatch'},
(pid=25726)                           'agent_1': { 'data': { 'actions':
np.ndarray((256,), dtype=int64, min=0.0, max=4.0, mean=1.801),
(pid=25726)                                                 'agent_index':
np.ndarray((256,), dtype=int64, min=1.0, max=1.0, mean=1.0),
(pid=25726)                                                 'dones':
np.ndarray((256,), dtype=bool, min=0.0, max=0.0, mean=0.0),
(pid=25726)                                                 'eps_id':
np.ndarray((256,), dtype=int64, min=144169260.0, max=1483390642.0,
mean=717353850.125),
(pid=25726)                                                 'infos':
np.ndarray((256,), dtype=object, head={'delta': -0.6810216555029001}),
(pid=25726)                                                 'new_obs':
np.ndarray((256, 2), dtype=float32, min=-1.989, max=2.683, mean=-0.059),
(pid=25726)                                                 'obs': np.ndarray((256,
2), dtype=float32, min=-1.989, max=2.683, mean=-0.03),
(pid=25726)                                                 'prev_actions':
np.ndarray((256,), dtype=int64, min=0.0, max=4.0, mean=1.75),
(pid=25726)                                                 'prev_rewards':
np.ndarray((256,), dtype=float32, min=0.0, max=0.433, mean=0.204),
(pid=25726)                                                 'q_values':
np.ndarray((256, 5), dtype=float32, min=-3.98, max=1.501, mean=-0.99),
(pid=25726)                                                 'rewards':
np.ndarray((256,), dtype=float32, min=0.147, max=1.242, mean=0.597),
(pid=25726)                                                 't': np.ndarray((256,),
dtype=int64, min=0.0, max=31.0, mean=15.5),
(pid=25726)                                                 'unroll_id':
np.ndarray((256,), dtype=int64, min=0.0, max=0.0, mean=0.0),
(pid=25726)                                                 'weights':
np.ndarray((256,), dtype=float32, min=2.247, max=2.753, mean=2.405)},
(pid=25726)                                       'type': 'SampleBatch'}},
(pid=25726)   'type': 'MultiAgentBatch'}
(pid=25726)
(pid=25738) /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=25738)   _np_qint8 = np.dtype([("qint8", np.int8, 1)])
(pid=25738) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
```

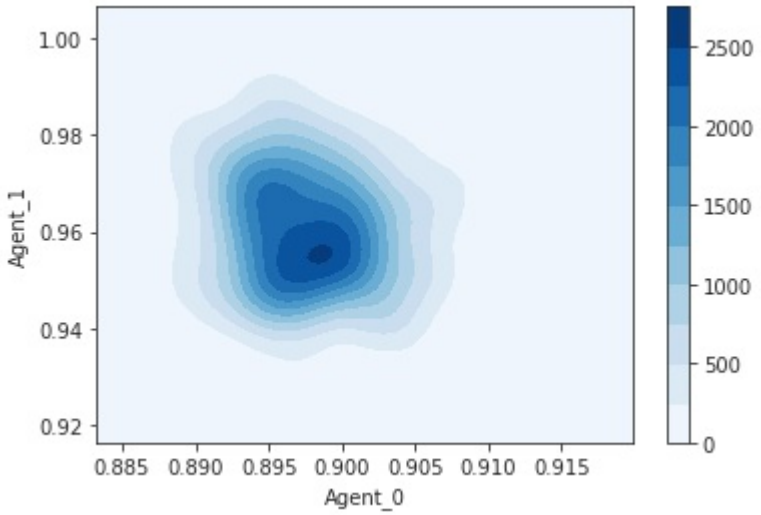
```
tensorflow/python/framework/dtypes.py:517: FutureWarning: Passing (type, 1) or
'1type' as a synonym of type is deprecated; in a future version of numpy, it will
be understood as (type, (1,)) / '(1,)type'.
(pid=25738)   _np_quint8 = np.dtype([("quint8", np.uint8, 1)])
(pid=25738) /home/lorenzo/anaconda3/envs/py36/lib/python3.6/site-packages/
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=25738)   _np_qint16 = np.dtype([("qint16", np.int16, 1)])
(pid=25738) /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=25738)   _np_quint16 = np.dtype([("quint16", np.uint16, 1)])
(pid=25738) /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=25738)   _np_qint32 = np.dtype([("qint32", np.int32, 1)])
(pid=25738) /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=25738)   np_resource = np.dtype([("resource", np.ubyte, 1)])
(pid=25738) /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=25738)   _np_qint8 = np.dtype([("qint8", np.int8, 1)])
(pid=25738) /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=25738)   _np_quint8 = np.dtype([("quint8", np.uint8, 1)])
(pid=25738) /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=25738)   _np_qint16 = np.dtype([("qint16", np.int16, 1)])
(pid=25738) /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=25738)   _np_quint16 = np.dtype([("quint16", np.uint16, 1)])
(pid=25738) /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=25738)   _np_qint32 = np.dtype([("qint32", np.int32, 1)])
(pid=25738) /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=25738)   np_resource = np.dtype([("resource", np.ubyte, 1)])
(pid=25738) 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=25738) Instructions for updating:
```

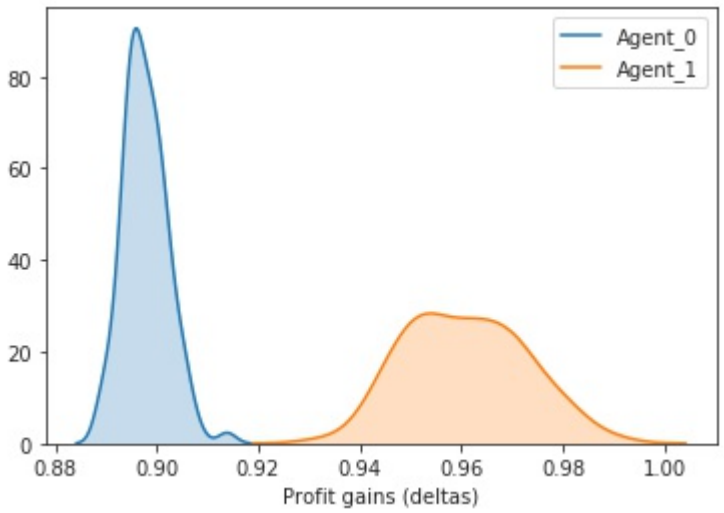
```
(pid=25738) non-resource variables are not supported in the long term
Episode reward 6614.881827699062
Episode 1 of 100
Episode reward 6620.462570431692
Episode 2 of 100
Episode reward 6590.501715317872
Episode 3 of 100
Episode reward 6580.738807639121
Episode 4 of 100
Episode reward 6591.12933930372
Episode 5 of 100
Episode reward 6572.3488013542055
Episode 6 of 100
Episode reward 6609.518205745466
Episode 7 of 100
Episode reward 6601.799505522896
Episode 8 of 100
Episode reward 6588.519891702732
Episode 9 of 100
Episode reward 6589.778403299667
Episode 10 of 100
Episode reward 6609.448091844422
Episode 11 of 100
Episode reward 6603.581135126863
Episode 12 of 100
Episode reward 6576.48077530799
Episode 13 of 100
Episode reward 6601.506293136692
Episode 14 of 100
Episode reward 6583.394376343002
Episode 15 of 100
Episode reward 6583.269398948098
Episode 16 of 100
Episode reward 6599.330890750429
Episode 17 of 100
Episode reward 6599.687441887042
Episode 18 of 100
Episode reward 6585.749083413137
Episode 19 of 100
Episode reward 6572.266980590549
Episode 20 of 100
Episode reward 6614.269498388876
Episode 21 of 100
Episode reward 6567.580933694507
Episode 22 of 100
Episode reward 6588.067393797141
Episode 23 of 100
Episode reward 6559.240823448945
Episode 24 of 100
Episode reward 6581.388512345514
Episode 25 of 100
Episode reward 6592.515971114938
Episode 26 of 100
Episode reward 6598.397298788607
Episode 27 of 100
Episode reward 6576.70018601171
Episode 28 of 100
Episode reward 6604.782012055611
Episode 29 of 100
```

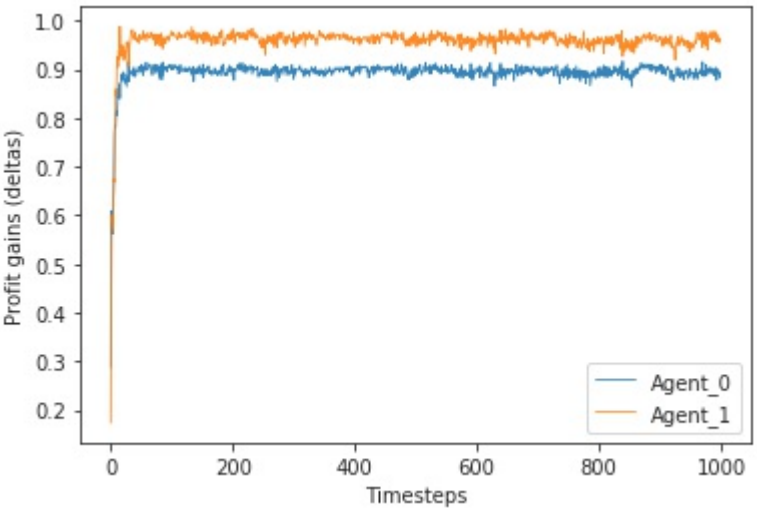
```
Episode reward 6581.0760668010125
Episode 30 of 100
Episode reward 6604.130518849176
Episode 31 of 100
Episode reward 6579.119991468802
Episode 32 of 100
Episode reward 6571.828822839452
Episode 33 of 100
Episode reward 6581.468566602975
Episode 34 of 100
Episode reward 6612.904445005635
Episode 35 of 100
Episode reward 6596.87864883645
Episode 36 of 100
Episode reward 6606.3647906117985
Episode 37 of 100
Episode reward 6585.695427084948
Episode 38 of 100
Episode reward 6593.011986308304
Episode 39 of 100
Episode reward 6597.525104510016
Episode 40 of 100
Episode reward 6616.964331630053
Episode 41 of 100
Episode reward 6601.255692572118
Episode 42 of 100
Episode reward 6578.1687894460665
Episode 43 of 100
Episode reward 6602.354309562849
Episode 44 of 100
Episode reward 6586.502545793765
Episode 45 of 100
Episode reward 6592.560287309553
Episode 46 of 100
Episode reward 6580.759535201601
Episode 47 of 100
Episode reward 6570.580607496362
Episode 48 of 100
Episode reward 6596.841786936744
Episode 49 of 100
Episode reward 6597.7395863739675
Episode 50 of 100
Episode reward 6590.019530849823
Episode 51 of 100
Episode reward 6608.894286156483
Episode 52 of 100
Episode reward 6578.924896436074
Episode 53 of 100
Episode reward 6597.6823281302495
Episode 54 of 100
Episode reward 6591.868300156559
Episode 55 of 100
Episode reward 6587.538506563926
Episode 56 of 100
Episode reward 6585.885813839768
Episode 57 of 100
Episode reward 6586.768988836012
Episode 58 of 100
Episode reward 6587.470722848166
```

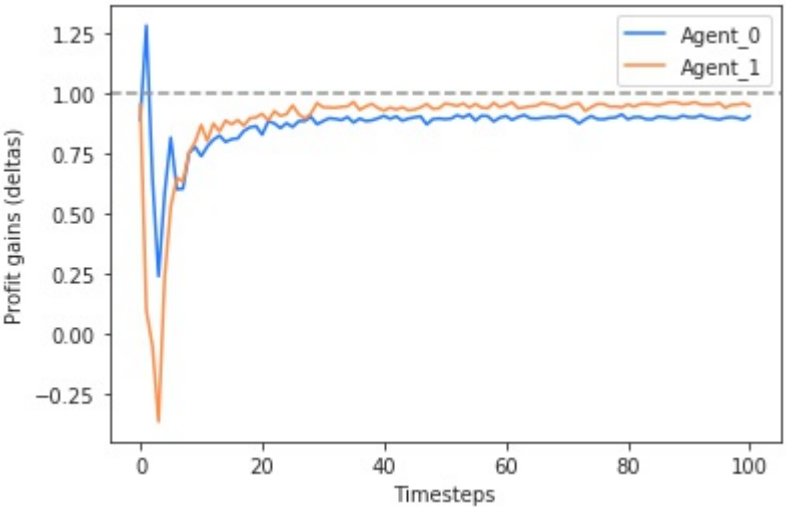
```
Episode 59 of 100
Episode reward 6599.681957263204
Episode 60 of 100
Episode reward 6593.041223636895
Episode 61 of 100
Episode reward 6575.761594329423
Episode 62 of 100
Episode reward 6573.268485063934
Episode 63 of 100
Episode reward 6588.176870647673
Episode 64 of 100
Episode reward 6587.172929994304
Episode 65 of 100
Episode reward 6575.279346525543
Episode 66 of 100
Episode reward 6581.928259360288
Episode 67 of 100
Episode reward 6574.443449235184
Episode 68 of 100
Episode reward 6621.836908915284
Episode 69 of 100
Episode reward 6576.26340222693
Episode 70 of 100
Episode reward 6584.054407205489
Episode 71 of 100
Episode reward 6576.363073895057
Episode 72 of 100
Episode reward 6585.043301701248
Episode 73 of 100
Episode reward 6591.3660672213255
Episode 74 of 100
Episode reward 6595.491297225957
Episode 75 of 100
Episode reward 6616.248843852168
Episode 76 of 100
Episode reward 6590.160857811141
Episode 77 of 100
Episode reward 6585.270497459148
Episode 78 of 100
Episode reward 6587.8352177350525
Episode 79 of 100
Episode reward 6605.3214837568885
Episode 80 of 100
Episode reward 6573.544304894614
Episode 81 of 100
Episode reward 6594.990719967783
Episode 82 of 100
Episode reward 6569.877843154768
Episode 83 of 100
Episode reward 6603.74305734762
Episode 84 of 100
Episode reward 6595.956548308386
Episode 85 of 100
Episode reward 6609.418460382369
Episode 86 of 100
Episode reward 6591.666712350678
Episode 87 of 100
Episode reward 6597.79583332376
Episode 88 of 100
```

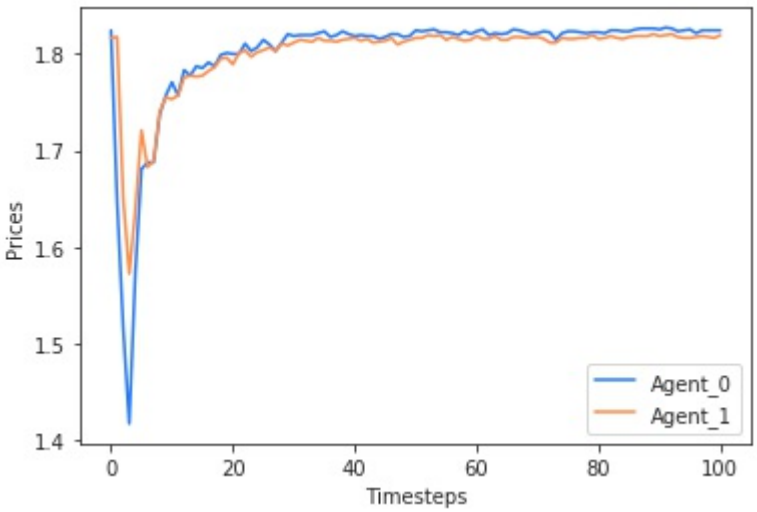
```
Episode reward 6606.283466402426
Episode 89 of 100
Episode reward 6581.4651193660175
Episode 90 of 100
Episode reward 6566.770993038112
Episode 91 of 100
Episode reward 6597.387838662651
Episode 92 of 100
Episode reward 6608.02968897967
Episode 93 of 100
Episode reward 6597.9137898296485
Episode 94 of 100
Episode reward 6618.7067636329175
Episode 95 of 100
Episode reward 6581.153495908784
Episode 96 of 100
Episode reward 6603.682262470751
Episode 97 of 100
Episode reward 6595.0251567505675
Episode 98 of 100
Episode reward 6605.89287614964
Episode 99 of 100
Episode reward 6598.4291817508565
Overall deltas mean: 0.9293 and std: 0.0327
Agent0 deltas mean: 0.8978 and std: 0.0043
Agent1 deltas mean: 0.9608 and std: 0.0116
/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)
```




Agent_0
Agent_1
80
60
40
20
0
0.88
0.90
0.92
0.94
0.96
0.98
1.00
Profit gains (deltas)


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
Profit gains (deltas)
Agent_0
Agent_1
0
200
400
600
800
1000
Timesteps


1.25
1.00
0.75
0.50
0.25
0.00
−0.25
Profit gains (deltas)
Agent_0
Agent_1
0
20
40
60
80
100
Timesteps




```
Traceback (most recent call last):

  File "<ipython-input-1-c2828c5a10e8>", line 1, in <module>
    runfile('/home/lorenzo/algorithmic-pricing/rollout/rollout.py', args='/home/
lorenzo/algorithmic-pricing/train_results/Azure_ApexDQN_Cont/
azure06_cont_DQN_res2/
APEX_MultiAgentFirmsPricingContinuous_0_2019-09-06_10-17-13df1x7oyx/
checkpoint_850/checkpoint-850 --run APEX --env env_cont', 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]:
```