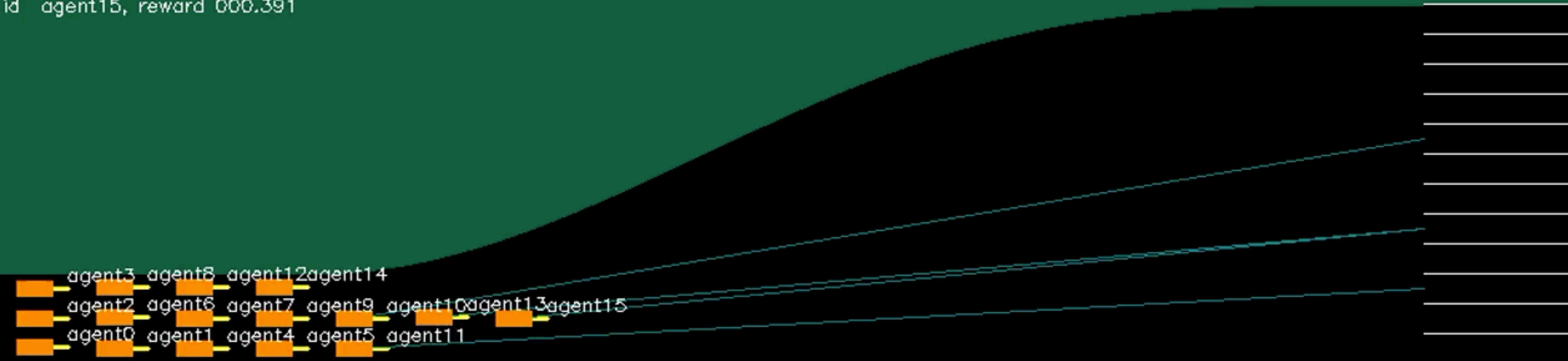




Emergence of “intelligent behaviors”

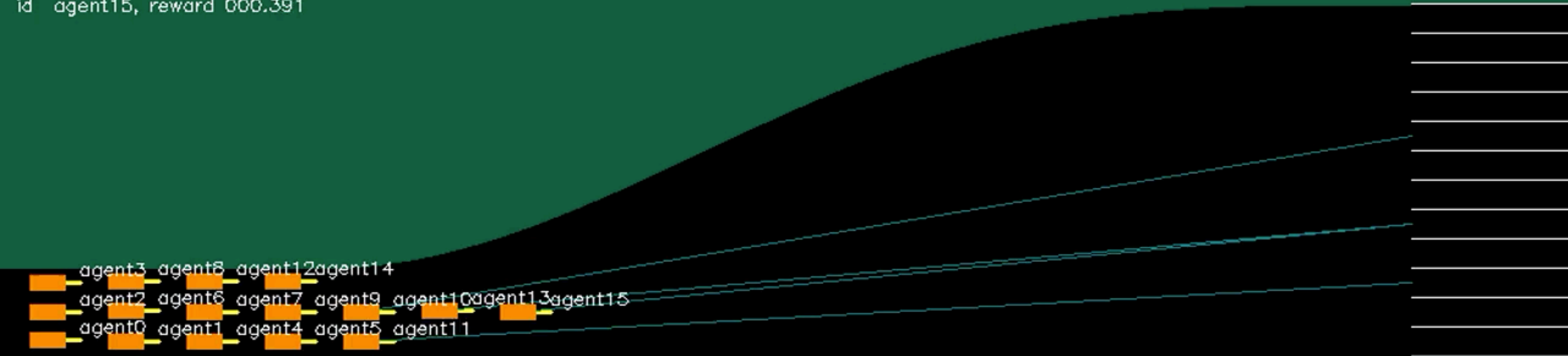
id agent0, reward 000.328
id agent1, reward 000.315
id agent2, reward 000.319
id agent3, reward 000.250
id agent4, reward 000.398
id agent5, reward 000.475
id agent6, reward 000.408
id agent7, reward 000.330
id agent8, reward 000.216
id agent9, reward 000.371
id agent10, reward 000.362
id agent11, reward 000.475
id agent12, reward 000.330
id agent13, reward 000.495
id agent14, reward 000.464
id agent15, reward 000.391



PPO-0409-2 100-iters

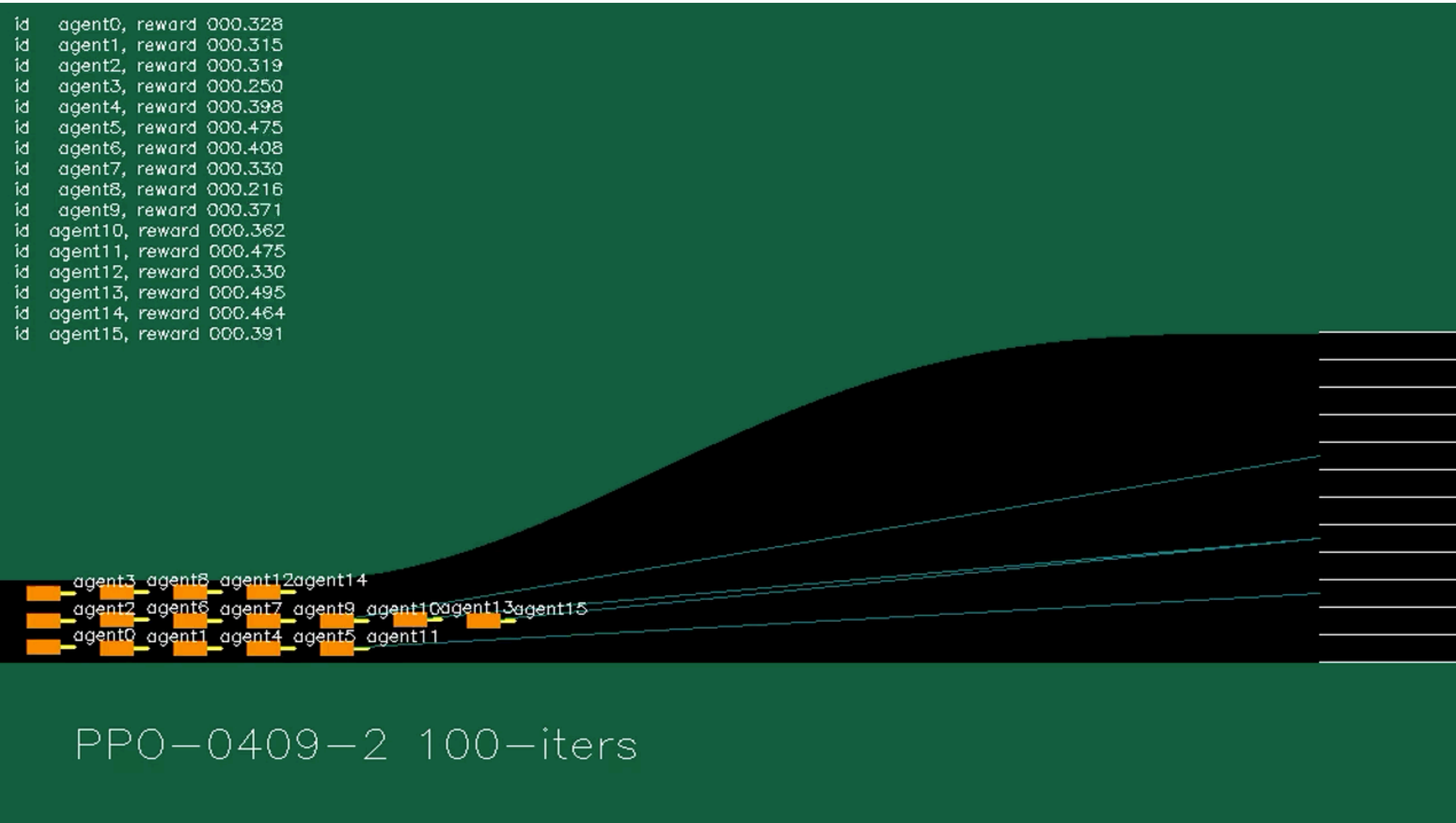
- At 100th iteration
- Better speed control.

```
id agent0, reward 000.328
id agent1, reward 000.315
id agent2, reward 000.319
id agent3, reward 000.250
id agent4, reward 000.398
id agent5, reward 000.475
id agent6, reward 000.408
id agent7, reward 000.330
id agent8, reward 000.216
id agent9, reward 000.371
id agent10, reward 000.362
id agent11, reward 000.475
id agent12, reward 000.330
id agent13, reward 000.495
id agent14, reward 000.464
id agent15, reward 000.391
```



PPO-0409-2 100-iters

Emergence of “intelligent behaviors”



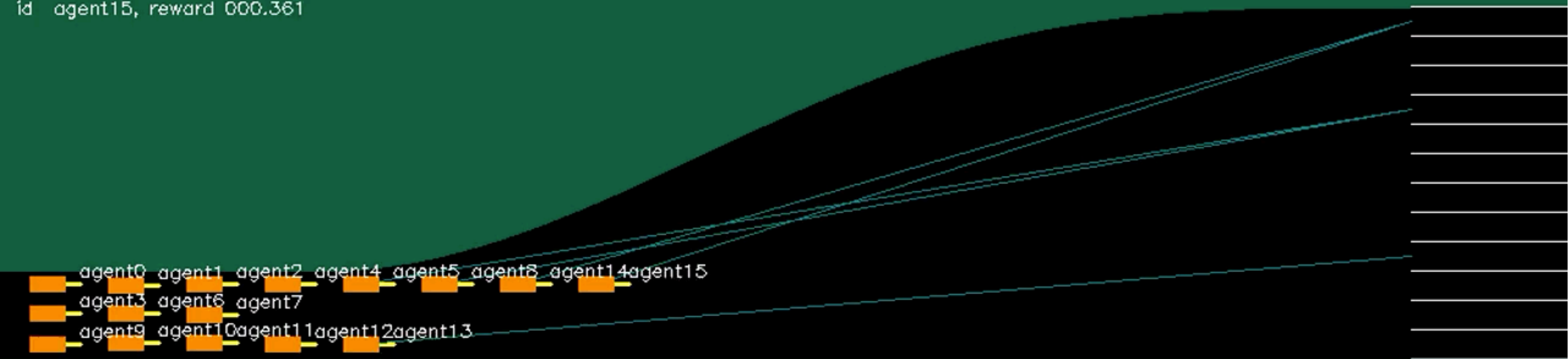
- At 100th iteration
- Better speed control.

Emergence of “intelligent behaviors”

Waiting and observing behaviors emerge! But still easy to collide.

- At 220th iteration

```
id agent0, reward 000.242
id agent1, reward 000.298
id agent2, reward 000.405
id agent3, reward 000.303
id agent4, reward 000.199
id agent5, reward 000.326
id agent6, reward 000.324
id agent7, reward 000.354
id agent8, reward 000.404
id agent9, reward 000.237
id agent10, reward 000.328
id agent11, reward 000.297
id agent12, reward 000.287
id agent13, reward 000.337
id agent14, reward 000.378
id agent15, reward 000.361
```



PPO-0409-2 220-iters

Metric	Value
iteration	220
— performance —	
agent collision	0.029
boundary col. rate	0
gate collision rate	0
pass rate	0.9709
success rate	0.1662
policy loss	-0.00224
value loss	1.47595
episode reward	1643.995
single reward	102.749
episode length	193.753
— configuration —	
num workers	48
env per worker	16
cpu per worker	1
sample batch size	100
train batch size	80000