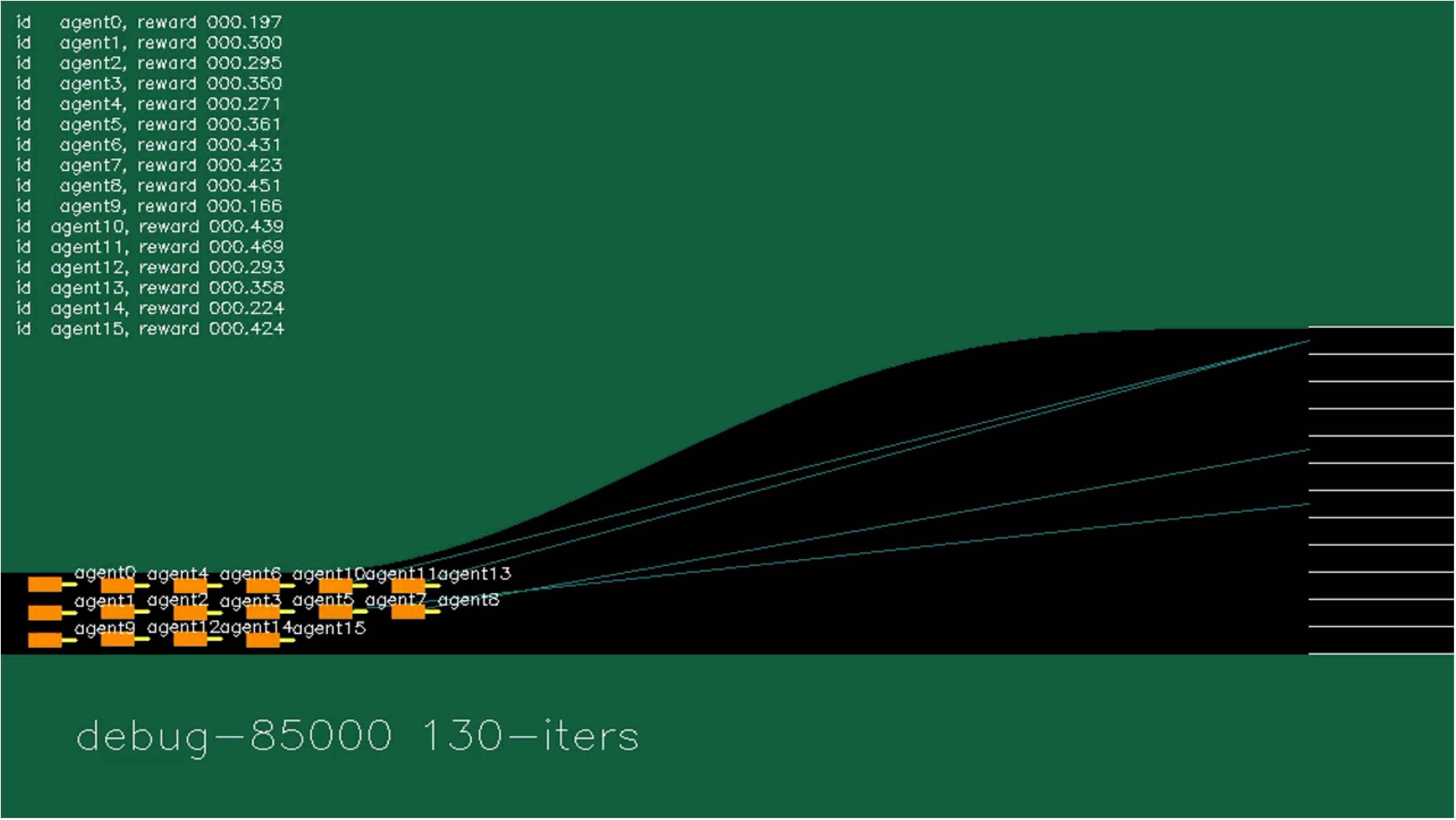




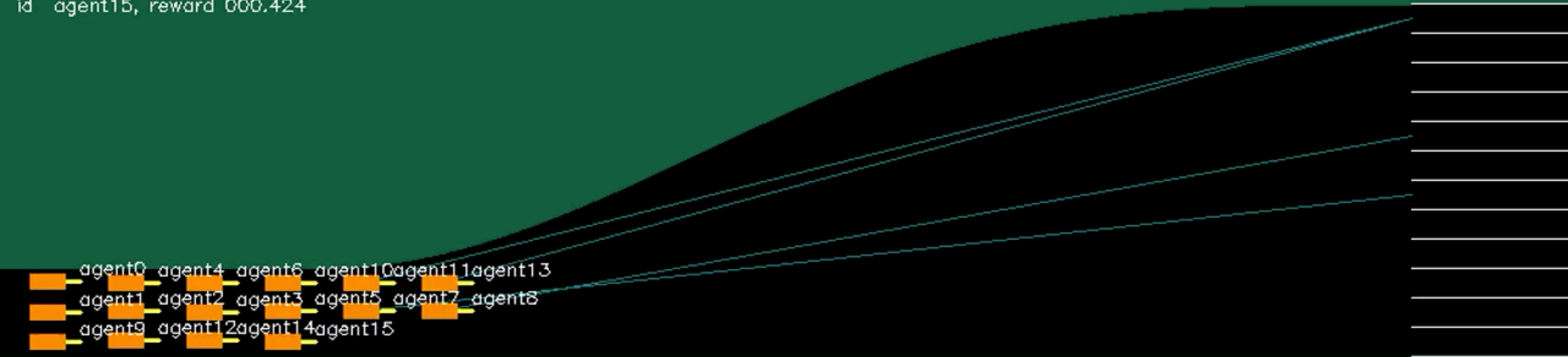
Performance of large batch training



Waiting and serving better! But still easy to collide.

Metric	Value
iteration	132
— performance —	
agent collision	0.0121
boundary col. rate	0
gate collision rate	0
pass rate	0.98796
success rate	0.16492
policy loss	-0.00010
value loss	0.68666
episode reward	1653.808
single reward	103.3630
episode length	194.8668
— configuration —	
num workers	2
env per worker	16
cpu per worker	26
sample batch size	200
train batch size	85000

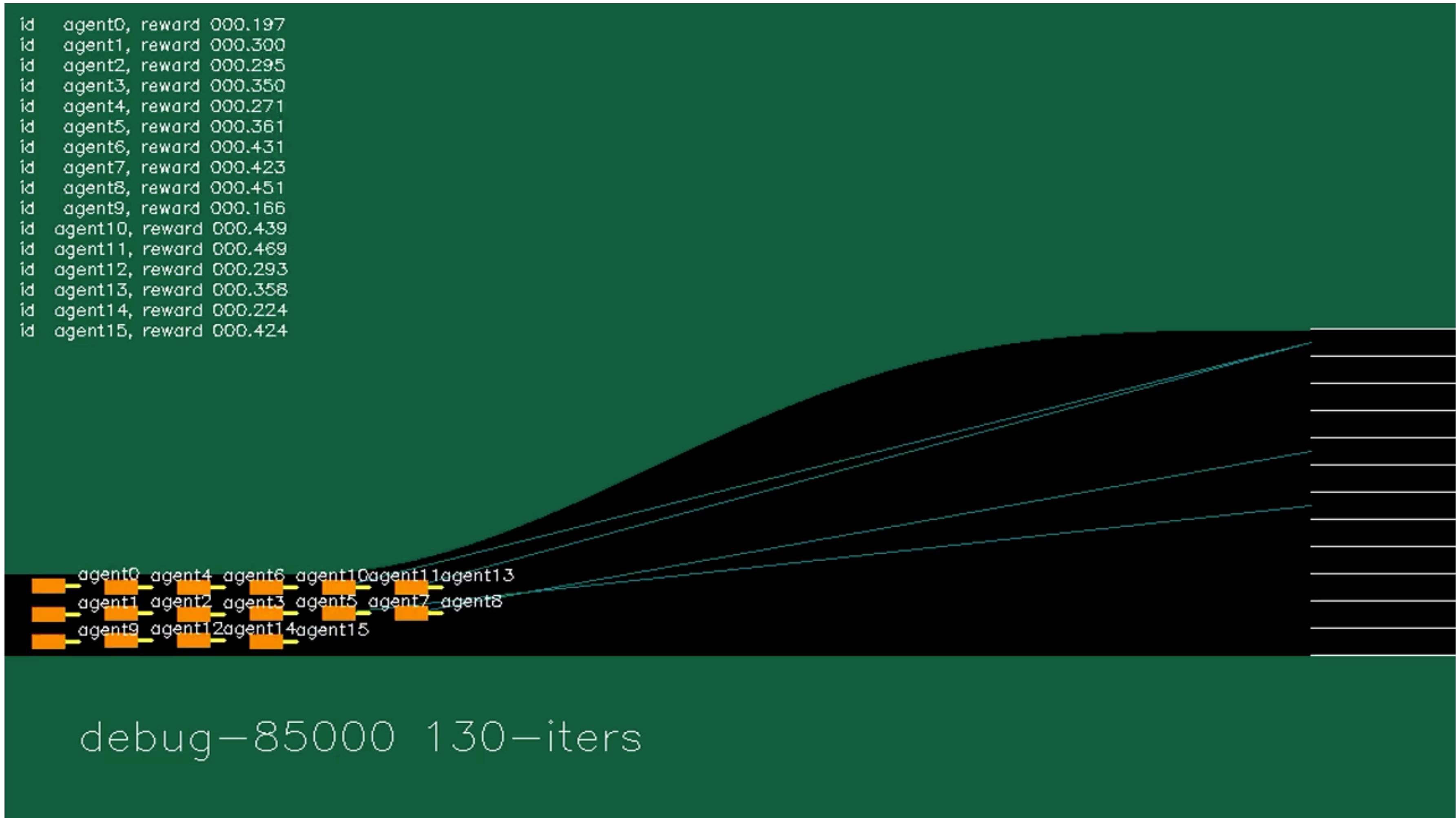
```
id agent0, reward 000.197
id agent1, reward 000.300
id agent2, reward 000.295
id agent3, reward 000.350
id agent4, reward 000.271
id agent5, reward 000.361
id agent6, reward 000.431
id agent7, reward 000.423
id agent8, reward 000.451
id agent9, reward 000.166
id agent10, reward 000.439
id agent11, reward 000.469
id agent12, reward 000.293
id agent13, reward 000.358
id agent14, reward 000.224
id agent15, reward 000.424
```



debug-85000 130-iters

Performance of large batch training

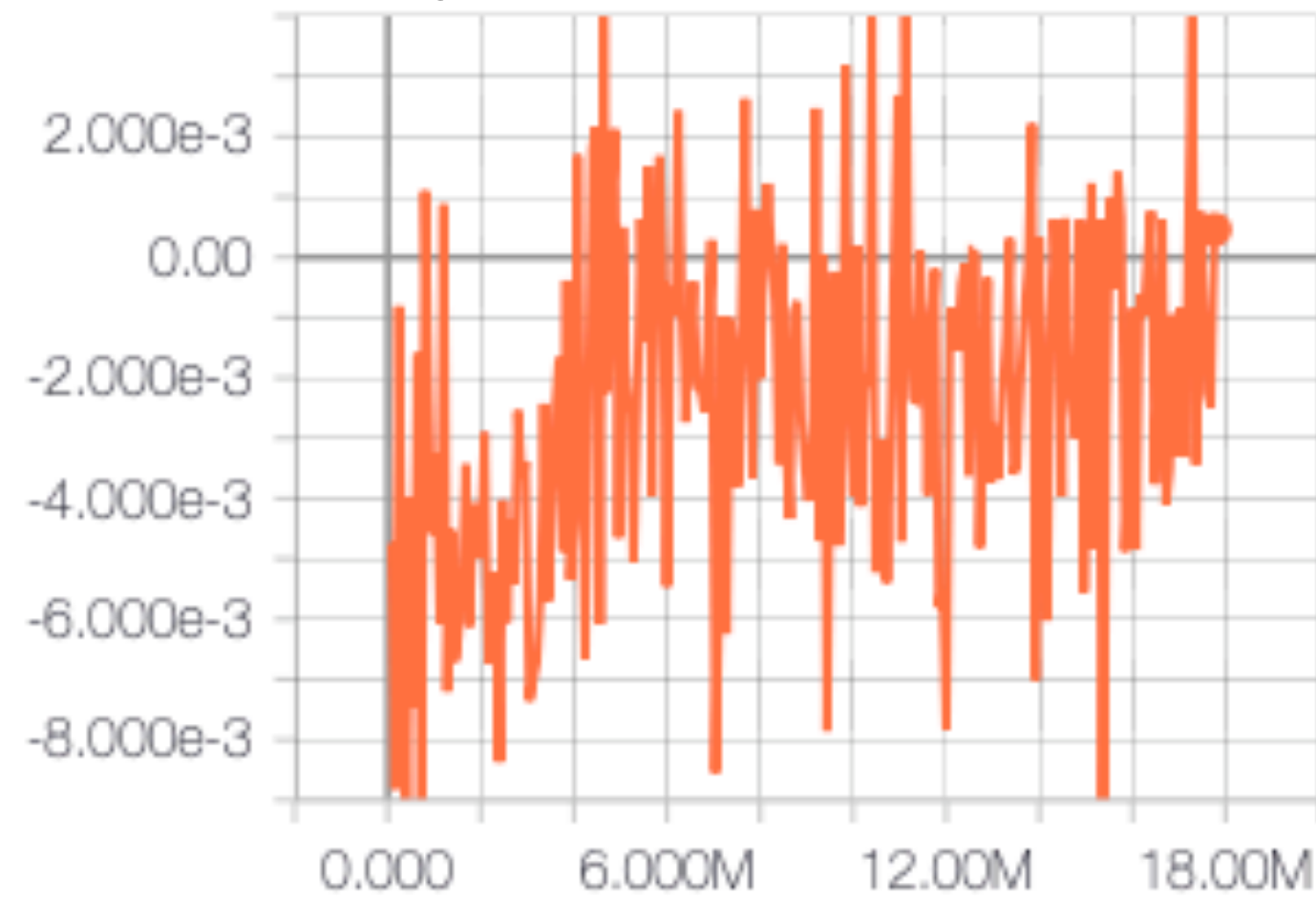
Waiting and observing behaviors emerge! But still easy to collide.



Metric	Value
iteration	132
— performance —	
agent collision	0.0121
boundary col. rate	0
gate collision rate	0
pass rate	0.98796
success rate	0.16492
policy loss	-0.00010
value loss	0.68666
episode reward	1653.808
single reward	103.3630
episode length	194.8668
— configuration —	
num workers	2
env per worker	16
cpu per worker	26
sample batch size	200
train batch size	85000

Discussion of large batch training

Policy Loss



Total Loss



Value Loss



Agent collision rate



Boundary collision rate



Gate collision rate

