GDC | March 21-25, 2022
San Francisco, CA

Reinforcement Learning for Efficient Cars and Tracks Design in Racing Games

Minggao Wei
Senior AI Engineer
NetEase Games AI Lab

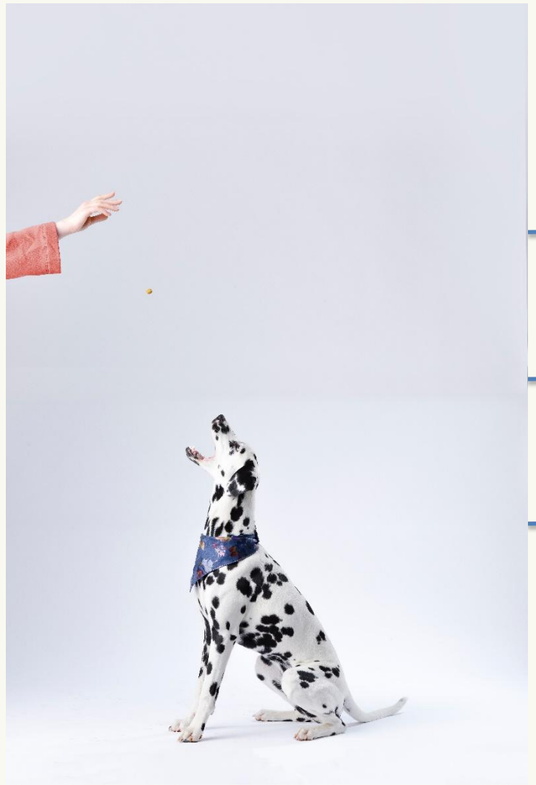
#GDC22

- Established in 2017
- Apply AI technology to Games
- Research Interests: Vision, NLP, RL and Speech Processing

Outline

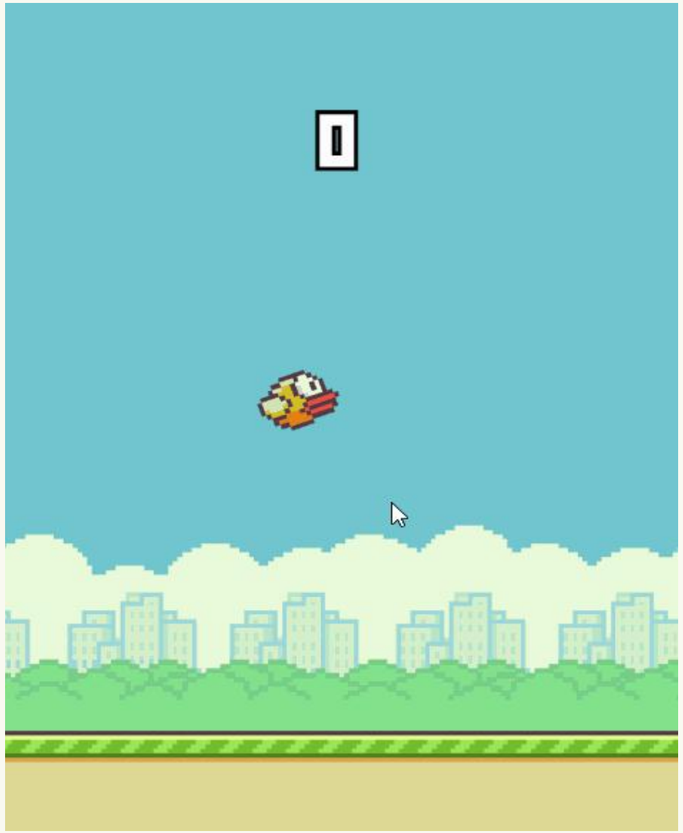
- Brief Introduction to Reinforcement Learning
- The Problem Space and Challenges
- The Solution and Results
- Takeaways

Brief Introduction to RL



1. hear “down” and sit down
2. receive food or not
3. correct its action

Brief Introduction to RL



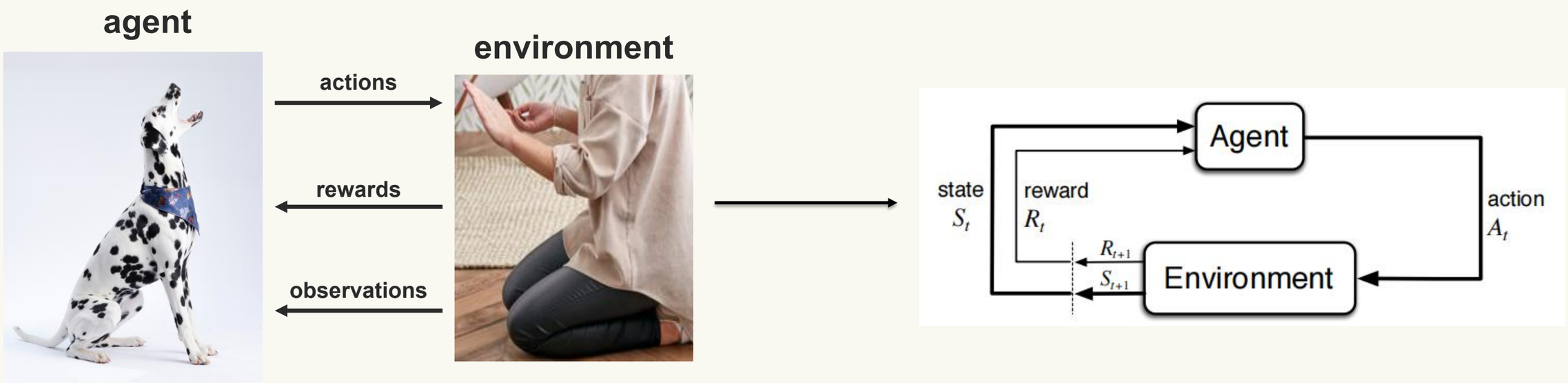
Flappy Bird

→ 1. touch the screen

→ 2. the game ends or not

→ 3. correct its action

Brief Introduction to RL



Brief Introduction to RL

The basic elements in RL

- environment（env）
- agent
- reward
- action
- policy
- observation

The process of RL

while true:

1) the agent collects data by interacting with env
2) the agent corrects its policy according to the collected data

Brief Introduction to RL

1. Three interfaces of the environment are necessary.
   - observation = env.reset()
   - observation = env.step(action)
   - terminal = env.terminal()

2. Design the reward and observation.

3. Choose the RL algorithm:
   - DQN, SAC and PPO, etc.

Brief Introduction to RL

Simple training process:

```
policy = DQN()
env = Env()
data_buffer = []
observation = env.reset()
while True:
    action = policy.choose_action(observation)
    next_observation, r = env.step(action)
    terminal = env.is_terminal()
    if terminal == True:
        next_observation = env.reset()
    data_buffer.append([observation, action, r, terminal, next_observation])
    observation = next_observation
```

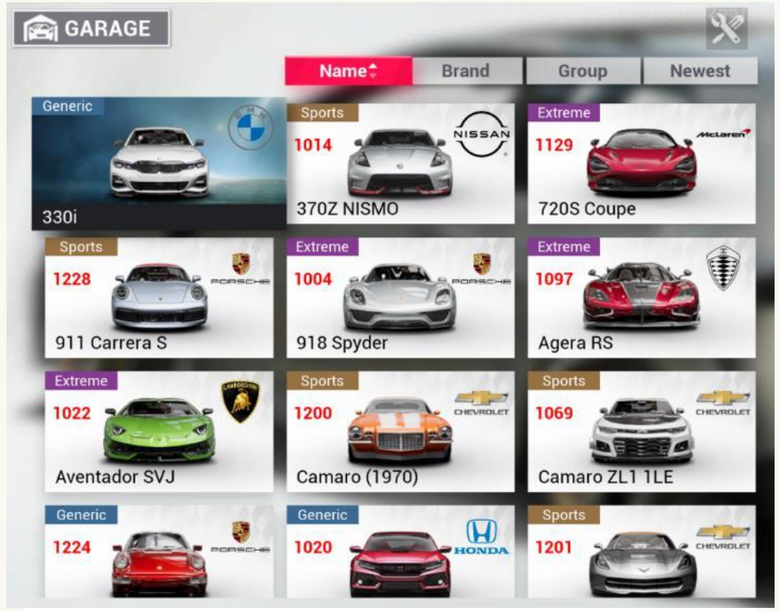
The Problem Space and Challenges

RACING
MASTER

Racing Master is a racing game developed by NetEase Games, emphasizing a realistic driving experience.

The Problem Space and Challenges

Racing car and track diversity



Hundreds of racing cars in Racing Master

Curve drift

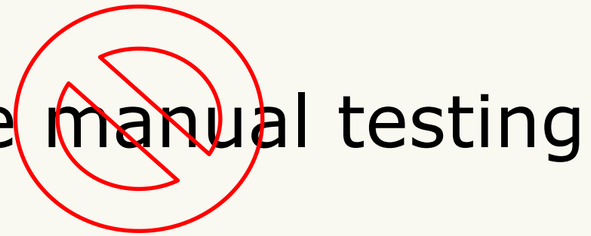
Tracks in different styles in Racing Master

RACING MASTER

The Problem Space and Challenges

Challenges for testing:

1. High-level players can tell the real handling differences between racing cars and the drift areas, as well as the necessity for designed curves.
2. The growing diversity of racing cars and tracks remarkably increases the testing workload.



RACING MASTER

The Problem Space and Challenges

It’s time to try reinforcement learning!

The Problem Space and Challenges

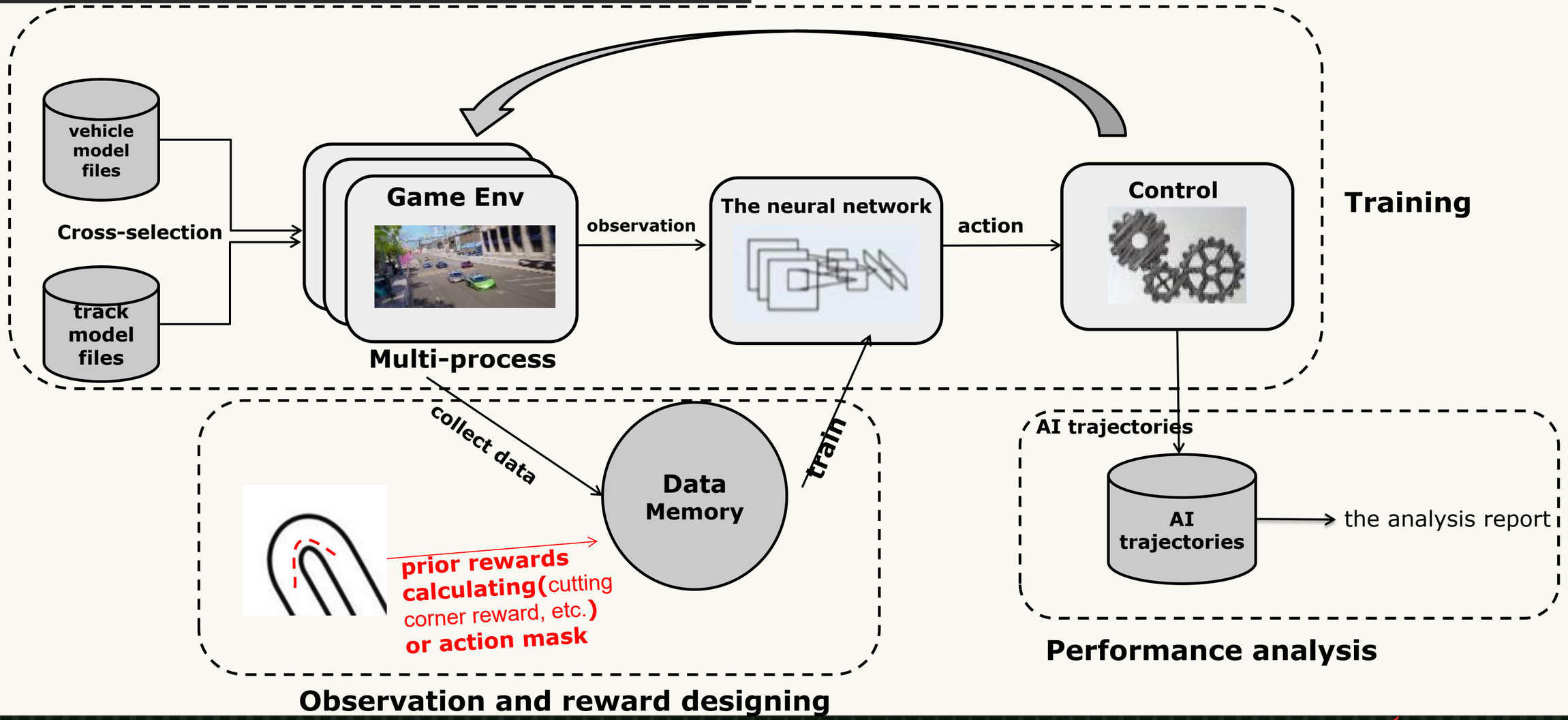
The characteristics of the system

1. For verifying racing car diversity, the system can achieve a super-human performance of 50 cars for each track in less than an hour.
2. For testing each curve’s drift areas, the system presents them along with their necessity.

The Solution and Results

- The Architecture of RL
- State Representation and Reward Engineering
- Details and Tips for Training and Modeling
- Results

The Architecture of RL in Racing Master



State Representation of Cars and Tracks

- Difficulties
  - Complete vehicle attributes
  - Complex road information

State Representation of Cars and Tracks

- Solutions:
  - Complete vehicle attributes to distinguish vehicle differences

Maximum power	Maximum torque	Mass	Size	Max Wheel Radius
1	1	1	3	1

State Representation of Cars and Tracks

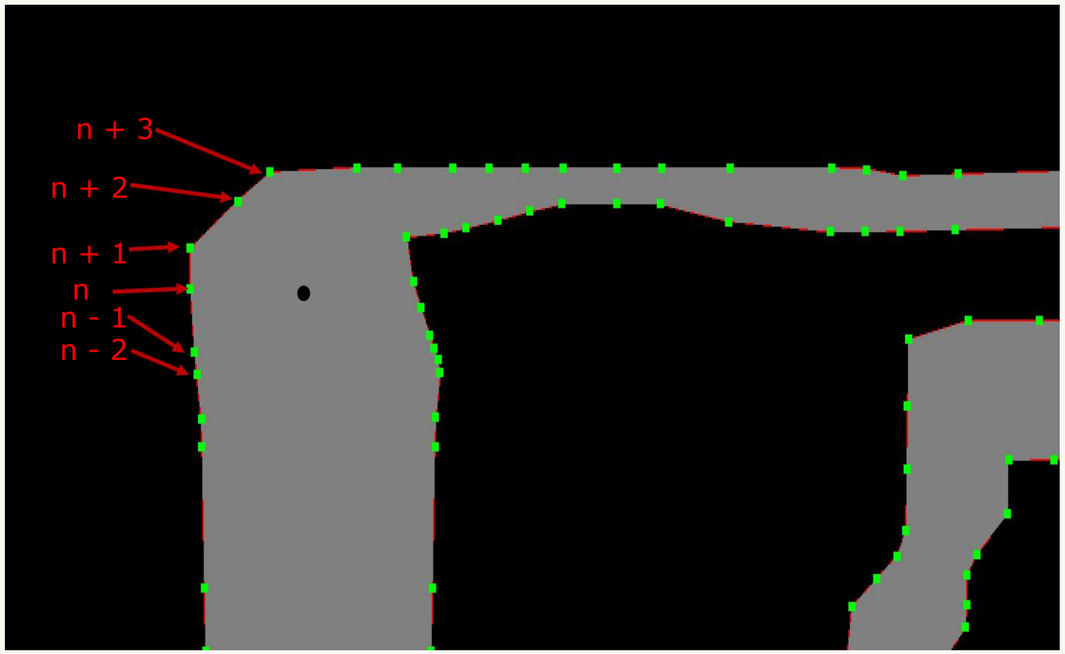
- Solutions:
  - Dynamic driving information of the vehicle

Gear	Rpm	Throttle	Handbrake		Speed
11 discrete	1	2 discrete	2 discrete	...	1

State Representation of Cars and Tracks

- Solutions:
  - Environmental perception information

Sensor features	Track features
19	30 * 8



Track Information

Reward Engineering and Action Mask

- Difficulties
  - High-level, collision-free and as fast as possible
  - Anthropomorphic, such as cutting the corner

Reward Engineering and Action Mask

- Solutions:

$$r = \alpha * r_{collision} + \beta * r_{speed} + \gamma * r_{cutting\ corner} + r_{tick} + r_{terminal}$$

a small negative constant every step

— the longer the trajectory, the smaller the total reward

Reward Engineering and Action Mask

- Difficulties:
  - Anthropomorphic handling

Anthropomorphic handling:
1.Each action keeps 100ms at least 2.The switch interval between the LEFT action and the RIGHT action cannot be less than 500ms 3.The switch interval between throttle and hand brake cannot be less than 500ms

Reward Engineering and Action Mask

- Solutions
  - Action mask

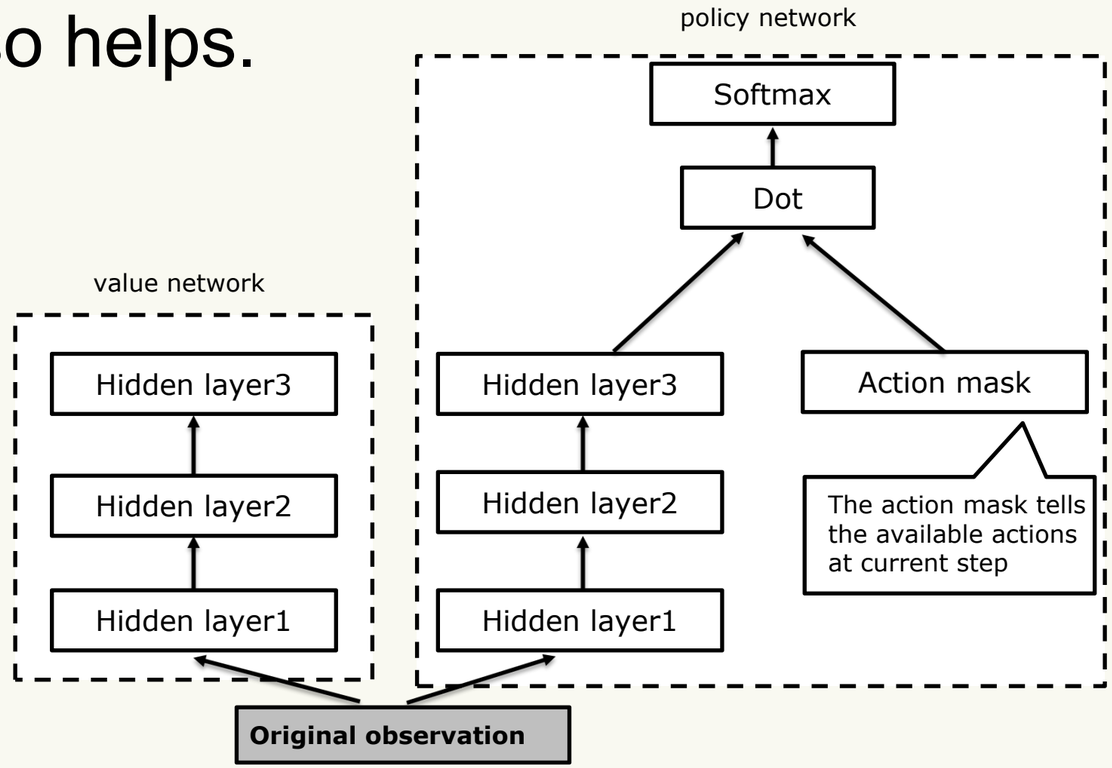
    Given the action tick is 100ms and the last step is LEFT, then the RIGHT action is not available at the current step.

Details and Tips for Training and Modeling

- Difficulties
  - The model should converge fast.
  - The training algorithm and framework should be efficient.

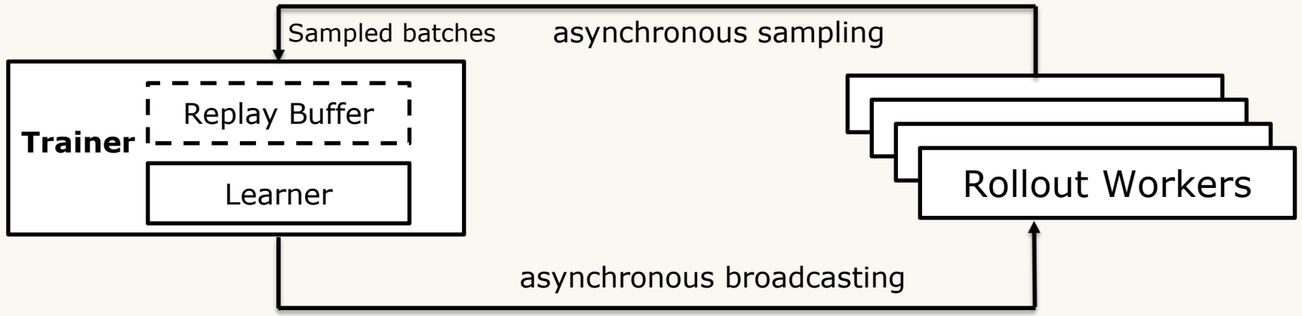
Details and Tips for Training and Modeling

- Solutions:
  - Independent value estimation network.
  - Action mask also helps.



Details and Tips for Training and Modeling

- Solutions
  - Asynchronous Proximal Policy Optimization (APPO)

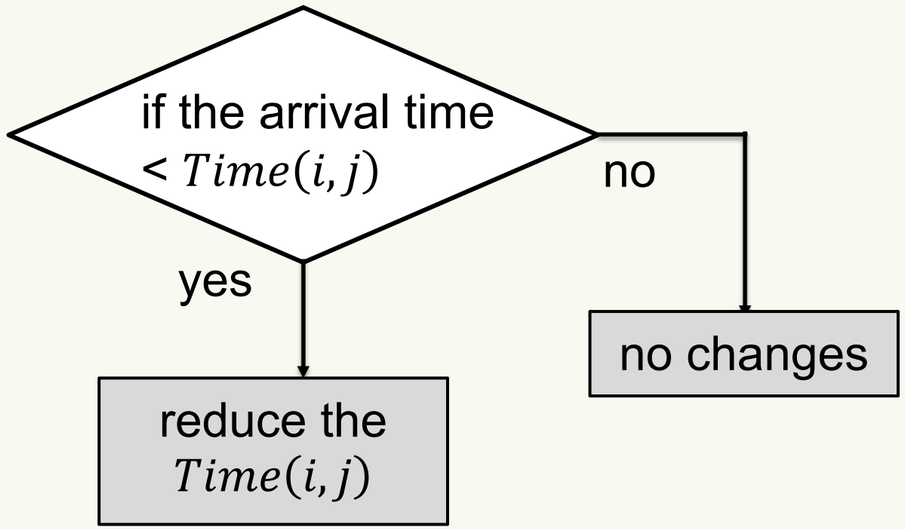



```
Worker Process:
    while True:
        1. model = GetNewModel()
        2. data = CollectData()
        3. SendData_ToMemory()

Learner Process:
    while True:
        1. data = SampleData_FromMemory()
        2. model = Train(data)
        3. SaveModel()
```

Details and Tips for Training and Modeling

- Tips for accelerating training
  - Dynamic terminal reward;

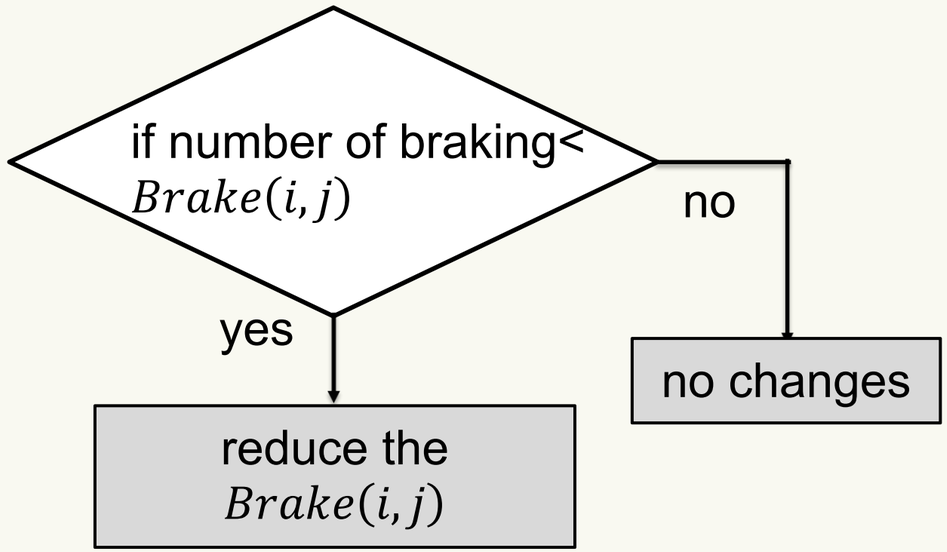


$$r_{termial} = \begin{cases} 1 \; if \; the \; arrival \; time < Time(i,j) \\ 0 \; if \; the \; arrival \; time \geq Time(i,j) \end{cases}$$

$Time$ is a global variable, which saves the running time thresholds of all cars and tracks. $Time(i,j)$ denotes the current running time threshold of the i-th car in the j-th track.

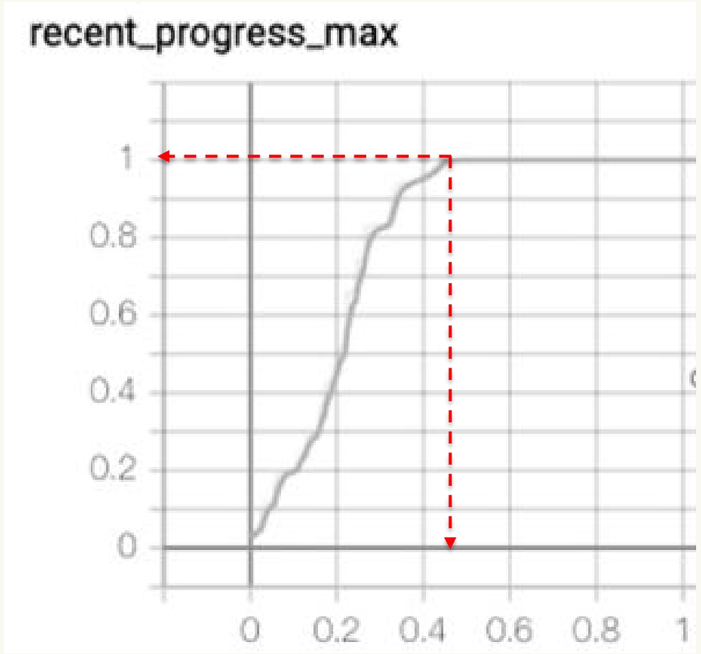
Details and Tips for Training and Modeling

- Tips for accelerating training
  - Dynamic action mask;



$Brake$ is a global variable, which saves the limits of the numbers of braking times of all tracks. $Brake(i,\ j)$ denotes the current limit of the number of braking times of the i-th curve in the j-th track.

The Outputs of the AI System



Training Curve. The X-axis denotes the training time. The Y-axis represents the ratio of cars that have finished the progress.

The Outputs of the AI System

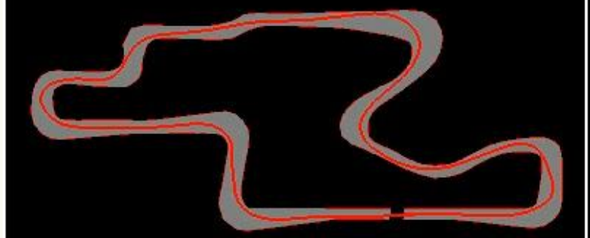
1. The optimal racing trace
   - The designers can figure out any car’s trajectory which they want to inspect.


route:SanF_route_0, car:1219
route:SanF_route_0, car:1001
route:SanF_route_0, car:1010
route:SanF_route_1, car:1219
route:SanF_route_1, car:1001
route:SanF_route_1, car:1010
route:Guangzhou_route_2, car:1219
route:Guangzhou_route_2, car:1001
route:Guangzhou_route_2, car:1010
route:chicago_route_3, car:1219
route:chicago_route_3, car:1001
route:chicago_route_3, car:1010
route:chicago_route_4, car:1219
route:chicago_route_4, car:1001
route:chicago_route_4, car:1010
route:chicago_route_5, car:1219
route:chicago_route_5, car:1001
route:chicago_route_5, car:1010
route:chicago_route_5, car:1097
route:chicago_route_5, car:1007
route:chicago_route_5, car:1024
Optimal path

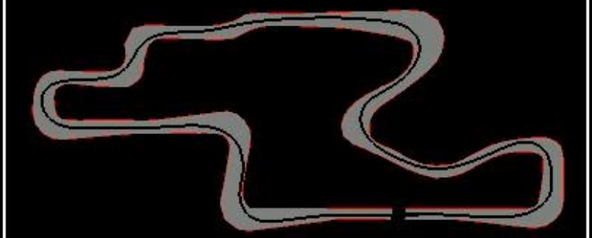
The Outputs of the AI System

1. The optimal racing trace
   - The dynamic running details can also be animated.

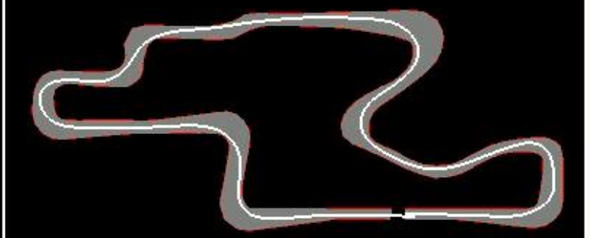
route:chicago_route_5, car:1219
speed:160.73455810546875
speedLimit:280
power:480
torque:571.917
throttle:1.0, brake:0.0, direction:-1.0, handbrake:0.0



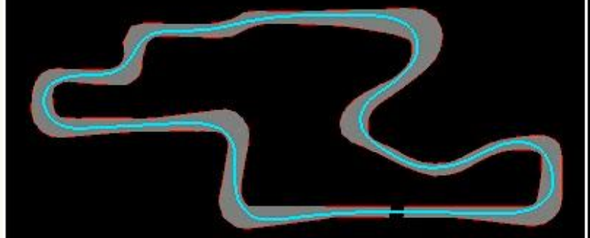
route:chicago_route_5, car:1001
speed:208.52442932128906
speedLimit:370
power:820
torque:819.98
throttle:1.0, brake:0.0, direction:0.0, handbrake:0.0



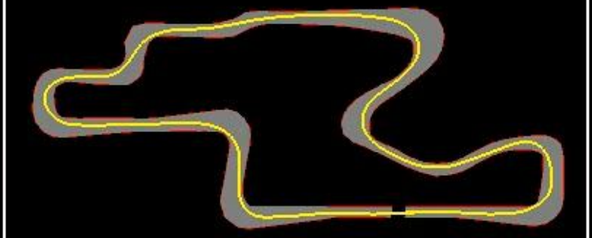
route:chicago_route_5, car:1010
speed:176.31761169433594
speedLimit:308
power:585
torque:656.06
throttle:1.0, brake:0.0, direction:1.0, handbrake:0.0



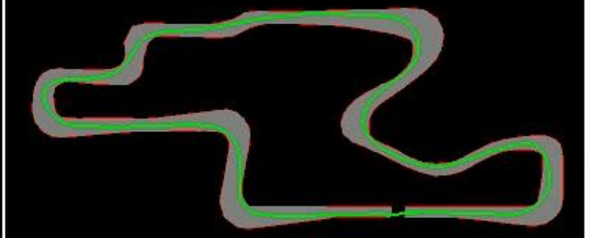
route:chicago_route_5, car:1097
speed:178.00424194335938
speedLimit:370
power:1160.0
torque:1249.74
throttle:1.0, brake:0.0, direction:-1.0, handbrake:0.0



route:chicago_route_5, car:1003
speed:172.66525268554688
speedLimit:310
power:800.0
torque:673.72
throttle:1.0, brake:0.0, direction:0.0, handbrake:0.0



route:chicago_route_5, car:1024
speed:129.43624877929688
speedLimit:233
power:181.0
torque:187.7
throttle:1.0, brake:0.0, direction:0.0, handbrake:0.0



The Outputs of the AI System

2. The best time report for all cars

Car id

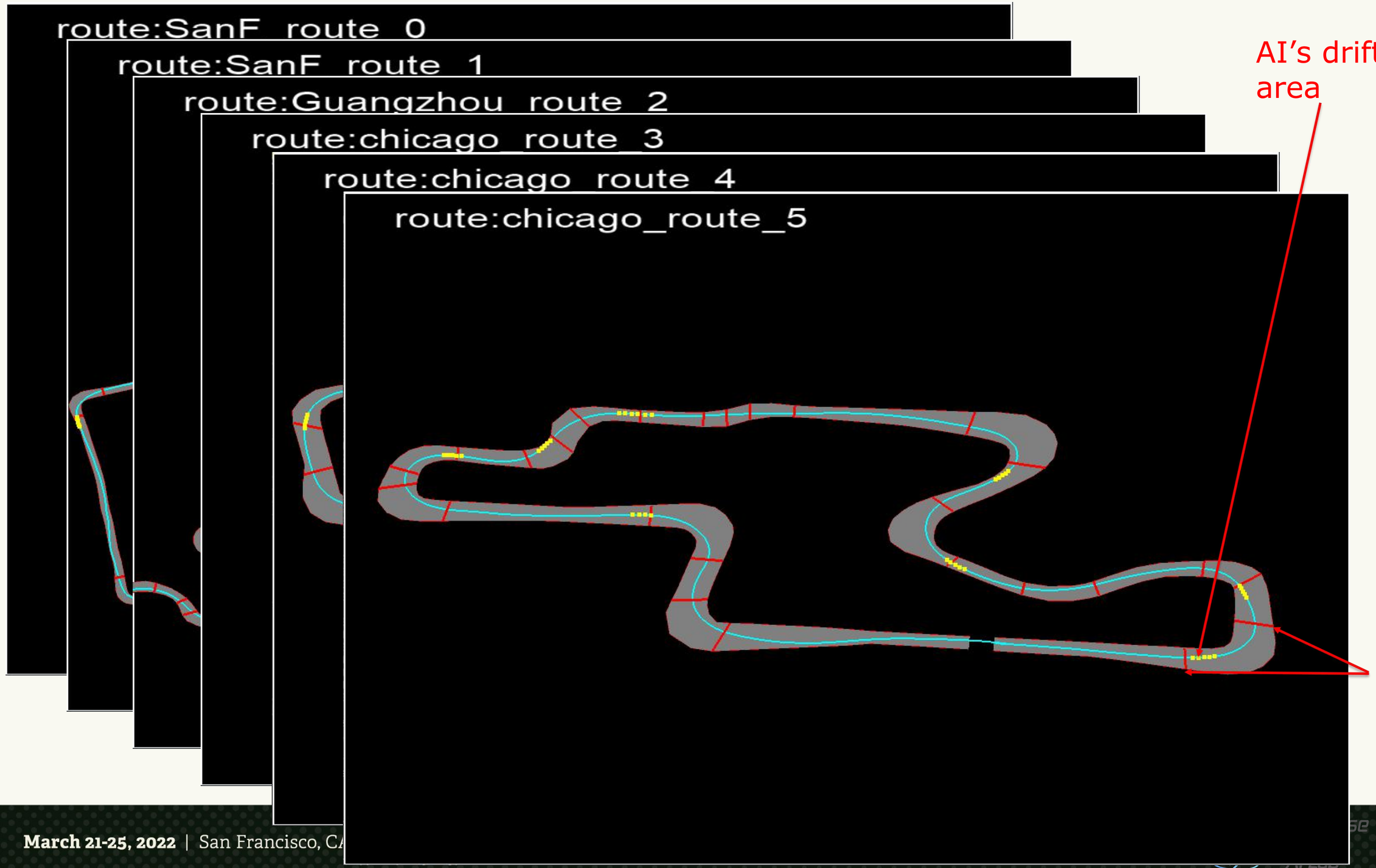
Track id →

	SanF_route_0	SanF_route_1	SanF_route_2	SanF_route_3	SanF_route_4	SanF_route_5
1097	99	97	75	45	54	36
1219	122	120	91	53	65	44
1224	117	115	88	53	64	42
1017	106	103	78	47	57	38
1004	104	103	79	47	57	39
1001	102	101	77	46	57	38
1209	104	102	77	46	56	38
1022	106	106	80	48	58	39
1002	105	105	81	48	58	39
1069	108	106	81	49	60	40
1007	110	108	83	50	60	41
1033	114	112	87	52	63	42
1012	113	107	93	57	67	45
1010	107	106	81	50	59	39
1027	109	108	83	51	60	40
1223	107	107	81	49	59	39
1228	110	108	82	50	60	40
1003	98	97	75	45	54	37
1005	107	104	79	49	58	39
1006	112	111	86	52	62	41
1068	104	101	77	47	56	38
1016	110	109	83	50	61	41
1011	105	105	80	49	57	39
1014	114	111	86	51	62	41
1032	117	117	89	54	65	42
1008	117	115	87	53	63	42
1158	116	116	88	55	64	42
1170	115	114	87	52	63	42
1025	114	114	88	52	63	42
1015	121	118	91	54	68	44
1126	116	114	88	54	64	42
1039	119	118	90	55	65	43
1024	138	135	105	66	76	51

	chicago_route_0	chicago_route_1	chicago_route_2	chicago_route_3	chicago_route_4	chicago_route_5
1097	99	86	55	33	41	45
1219	127	104	67	40	49	54
1224	119	100	64	39	49	52
1017	106	91	58	33	44	47
1004	108	93	59	35	44	48
1001	103	90	58	35	44	48
1209	105	91	58	35	43	47
1022	114	94	60	36	46	51
1002	110	95	60	35	45	50
1069	107	94	61	36	46	50
1007	118	97	62	37	47	51
1033	113	103	66	40	49	54
1012	136	111	70	40	53	56
1010	105	94	61	37	46	51
1027	116	96	62	37	46	51
1223	113	95	61	35	45	49
1228	110	95	62	37	47	50
1003	96	86	56	33	42	46
1005	114	93	59	35	44	48
1006	113	100	63	38	49	52
1068	103	90	58	34	43	47
1016	113	95	61	37	47	51
1011	114	93	59	35	44	48
1014	126	98	63	36	46	52
1032	126	102	65	39	49	53
1008	127	101	64	39	48	52
1158	126	104	66	40	50	54
1170	122	103	65	38	48	53
1025	123	102	65	38	49	54
1015	136	110	71	41	51	58
1126	127	103	65	39	48	53
1039	127	102	66	39	50	54
1024	152	122	78	47	58	63

The Outputs of the AI System

3. Each curve’s drift areas and their necessity


route:SanF_route_0
route:SanF_route_1
route:Guangzhou_route_2
route:chicago_route_3
route:chicago_route_4
route:chicago_route_5
AI's drift area
drift area designed by developer

Analysis for the Cars and the Curves

	SanF_route_0	SanF_route_1	SanF_route_2	SanF_route_3	SanF_route_4	SanF_route_5
1097	99	97	75	45	54	36
1219	122	120	91	53	65	44
1224	117	115	88	53	64	42
1017	106	103	78	47	57	38
1004	104	103	79	47	57	39
1001	102	101	77	46	57	38
1209	104	102	77	46	56	38
1022	106	106	80	48	58	39
1002	105	105	81	48	58	39
1069	108	106	81	49	60	40
1007	110	108	83	50	60	41
1033	114	112	87	52	63	42
1012	113	107	93	57	67	45
1010	107	106	81	50	59	39
1027	109	108	83	51	60	40
1223	107	107	81	49	59	39
1228	110	108	82	50	60	40
1003	98	97	75	45	54	37
1005	107	104	79	49	58	39
1006	112	111	86	52	62	41
1068	104	101	77	47	56	38
1016	110	109	83	50	61	41
1011	105	105	80	49	57	39
1014	114	111	86	51	62	41
1032	117	117	89	54	65	42
1008	117	115	87	53	63	42
1158	116	116	88	55	64	42
1170	115	114	87	52	63	42
1025	114	114	88	52	63	42
1015	121	118	91	54	68	44
1126	116	114	88	54	64	42
1039	119	118	90	55	65	43
1024	138	135	105	66	76	51

A

B

C

	chicago_route_0	chicago_route_1	chicago_route_2	chicago_route_3	chicago_route_4	chicago_route_5
1097	99	86	55	33	41	45
1219	127	104	67	40	49	54
1224	119	100	64	39	49	52
1017	106	91	58	33	44	47
1004	108	93	59	35	44	48
1001	103	90	58	35	44	48
1209	105	91	58	35	43	47
1022	114	94	60	36	46	51
1002	110	95	60	35	45	50
1069	107	94	61	36	46	50
1007	118	97	62	37	47	51
1033	113	103	66	40	49	54
1012	136	111	70	40	53	56
1010	105	94	61	37	46	51
1027	116	96	62	37	46	51
1223	113	95	61	35	45	49
1228	110	95	62	37	47	50
1003	96	86	56	33	42	46
1005	114	93	59	35	44	48
1006	113	100	63	38	49	52
1068	103	90	58	34	43	47
1016	113	95	61	37	47	51
1011	114	93	59	35	44	48
1014	126	98	63	36	46	52
1032	126	102	65	39	49	53
1008	127	101	64	39	48	52
1158	126	104	66	40	50	54
1170	122	103	65	38	48	53
1025	123	102	65	38	49	54
1015	136	110	71	41	51	58
1126	127	103	65	39	48	53
1039	127	102	66	39	50	54
1024	152	122	78	47	58	63

	Harbour_route_0	Harbour_route_1
1097	26	71
1219	31	86
1224	30	80
1017	28	74
1004	28	74
1001	29	73
1209	27	73
1022	28	77
1002	28	76
1069	29	76
1007	30	79
1033	32	83
1012	32	87
1010	29	77
1027	30	78
1223	28	77
1228	29	77
1003	27	71
1005	28	76
1006	30	79
1068	27	74
1016	30	79
1011	28	75
1014	30	80
1032	30	81
1008	30	80
1158	30	82
1170	30	82
1025	30	81
1015	33	89
1126	30	82
1039	31	82
1024	36	96

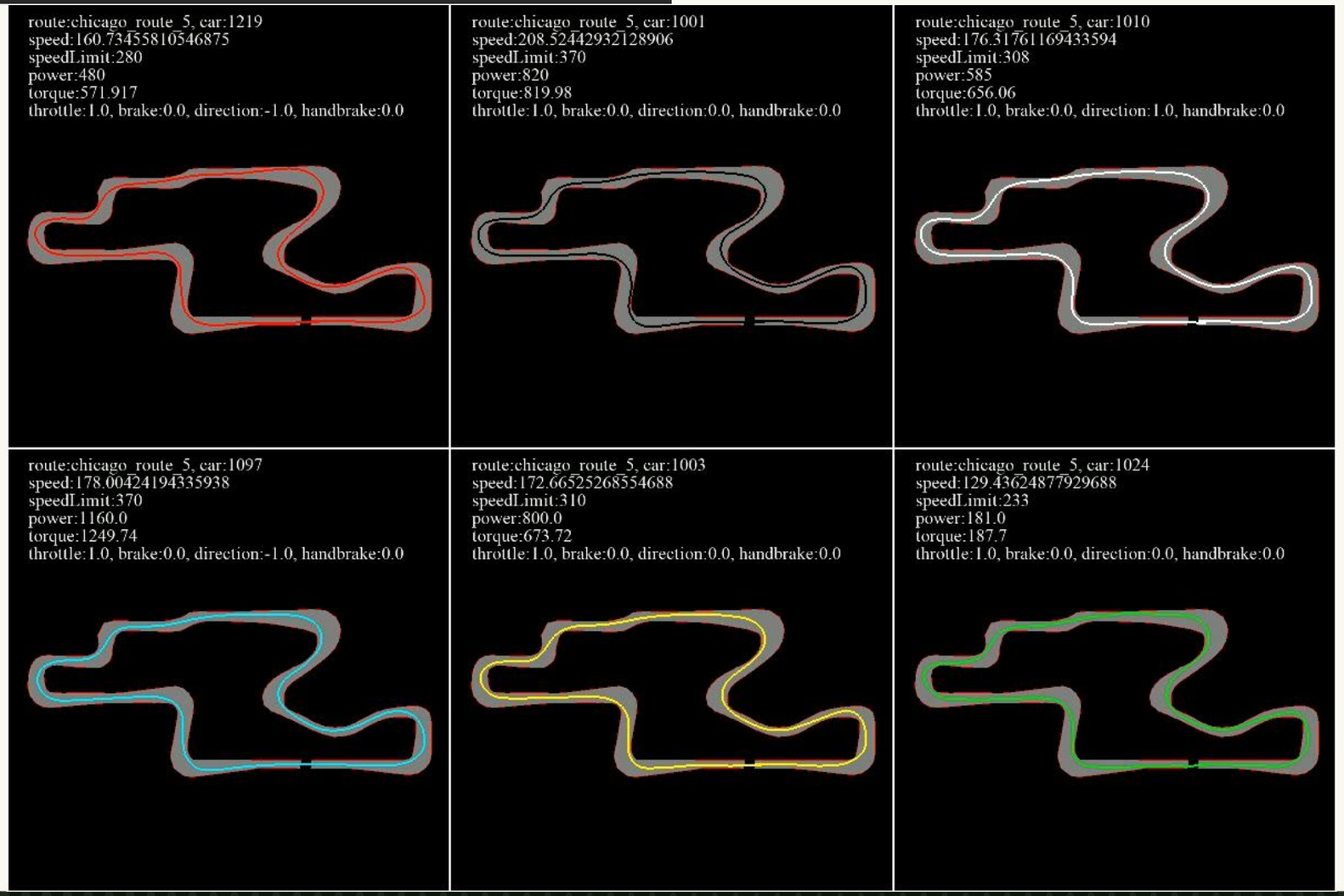
	tokyo_route_0
1097	165
1219	208
1224	202
1017	190
1004	182
1001	181
1209	180
1022	188
1002	192
1069	189
1007	189
1033	203
1012	218
1010	184
1027	201
1223	185
1228	193
1003	177
1005	183
1006	188
1068	177
1016	194
1011	199
1014	201
1032	206
1008	203
1158	202
1170	205
1025	211
1015	218
1126	209
1039	208
1024	262

	Guangzhou_route_0	Guangzhou_route_1
1097	98	115
1219	117	126
1224	106	123
1017	101	119
1004	111	120
1001	102	112
1209	104	115
1022	105	123
1002	105	124
1069	114	124
1007	106	124
1033	109	127
1012	125	137
1010	112	122
1027	105	121
1223	112	121
1228	106	117
1003	96	113
1005	103	120
1006	103	123
1068	105	116
1016	106	126
1011	100	120
1014	115	136
1032	107	128
1008	110	124
1158	108	131
1170	111	124
1025	115	132
1015	121	134
1126	115	127
1039	114	127
1024	126	146

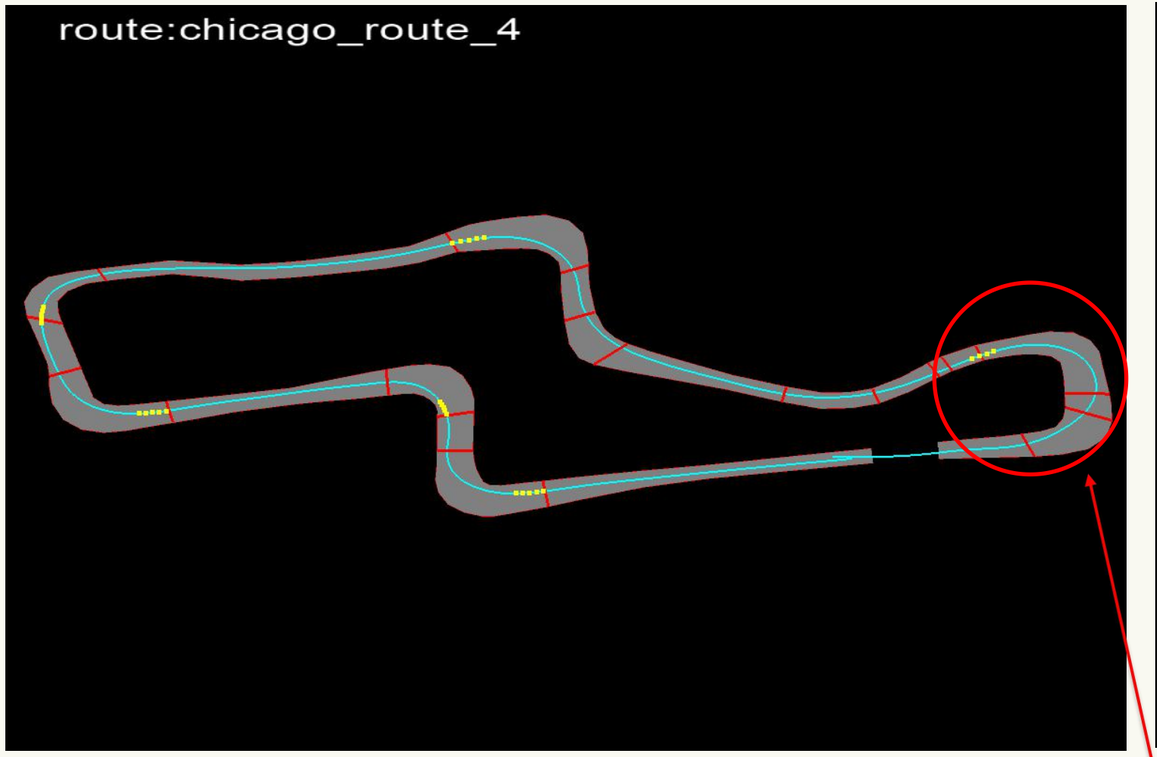
The better equipped cars perform better on all tracks, such as 1097, 1006 and 1024.

	Maximum power	Maximum torque	Speed Limit
1097	1160	1249.74	370
1006	520.0	565.47	288
1024	181	187.7	233

Analysis for the Cars and the Curves

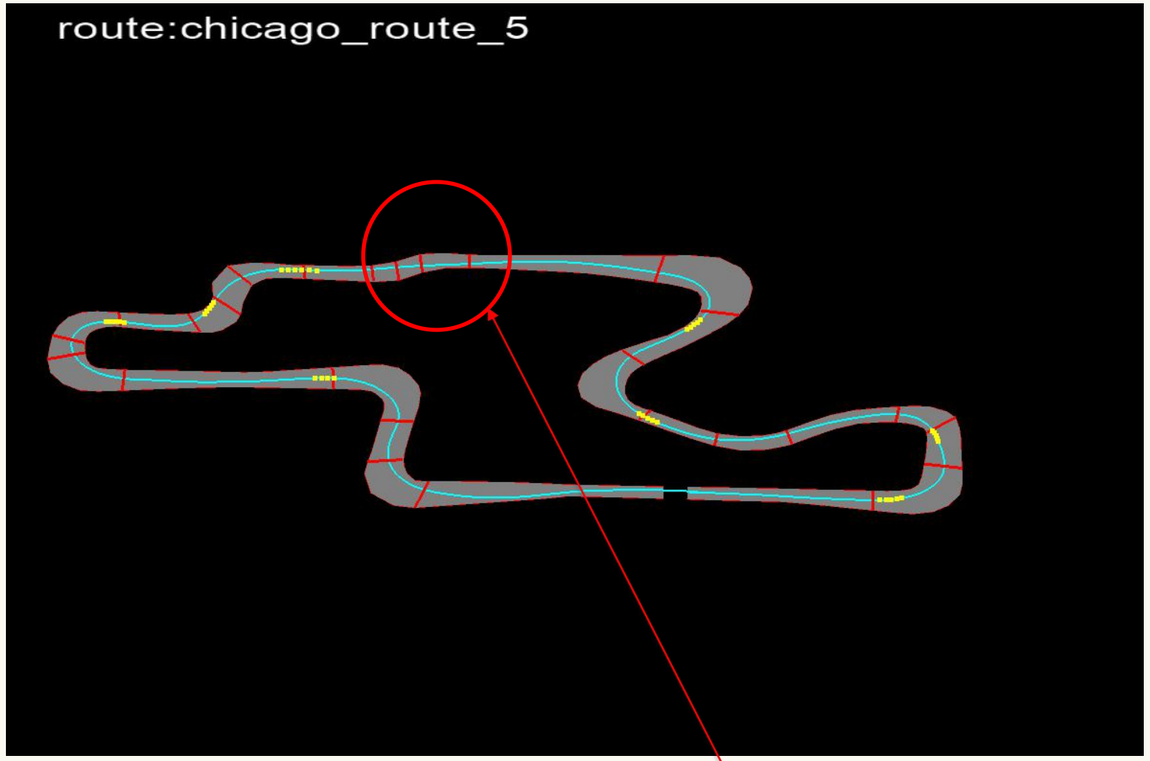


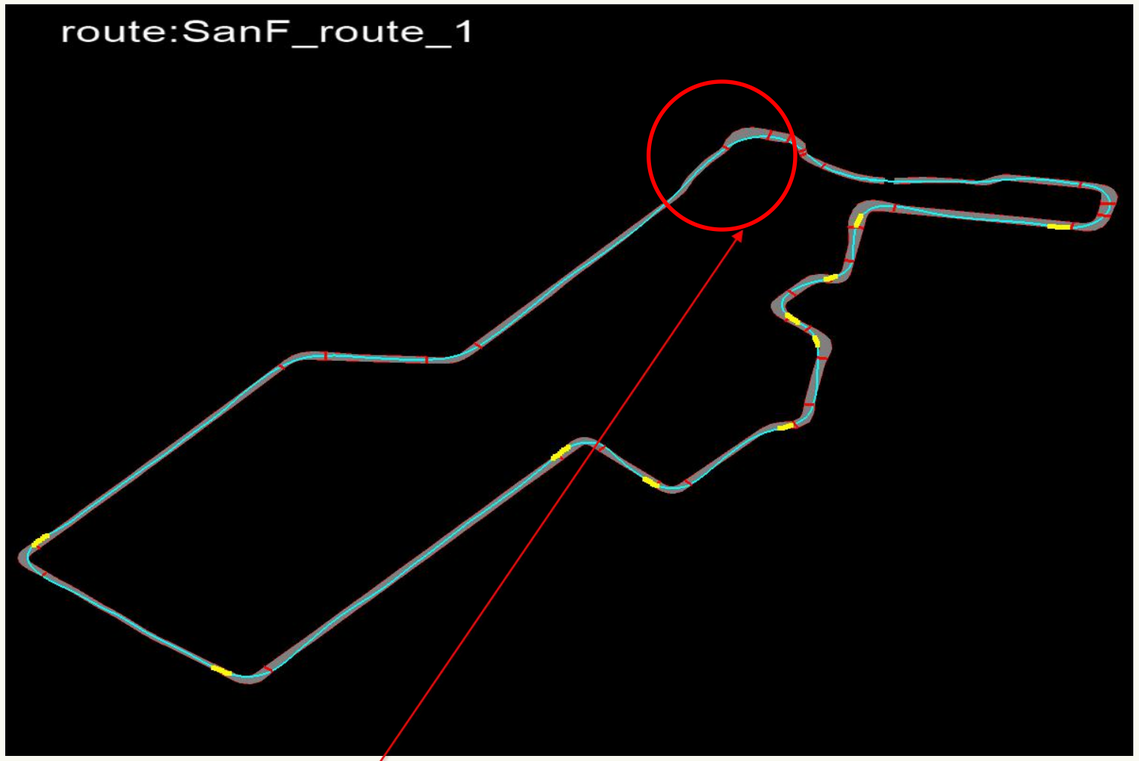
Analysis for the Cars and the Curves



Sometimes passing through the corner is easy with only one handbrake needed, because the width of these U turns are set too wide or the straight acceleration track ahead is too short.

Analysis for the Cars and the Curves





The curve may be unnecessary, or the shape of the curve may be sharper.

The Secondary Product

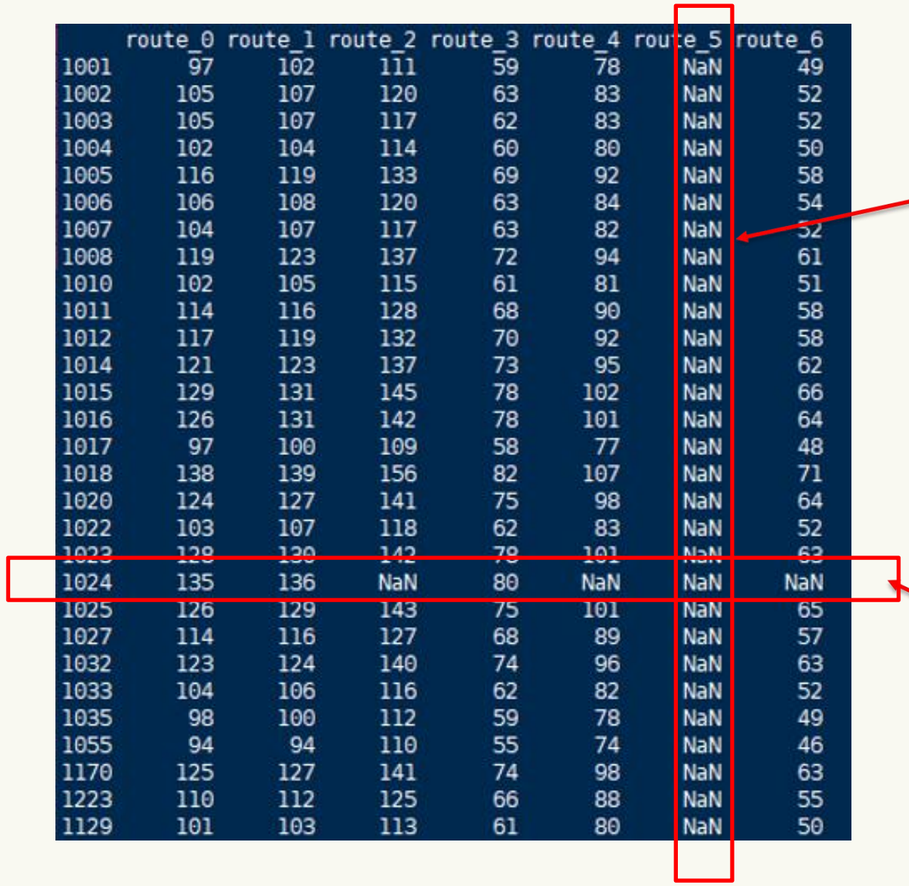
- Help the players to experience the game

Player in San Francisco

The AI’s trajectories help the players to run faster. The green arrows represent the normal acceleration area, and the yellow arrows represent the handbrake area.

The Secondary Product

- Help the developers to debug



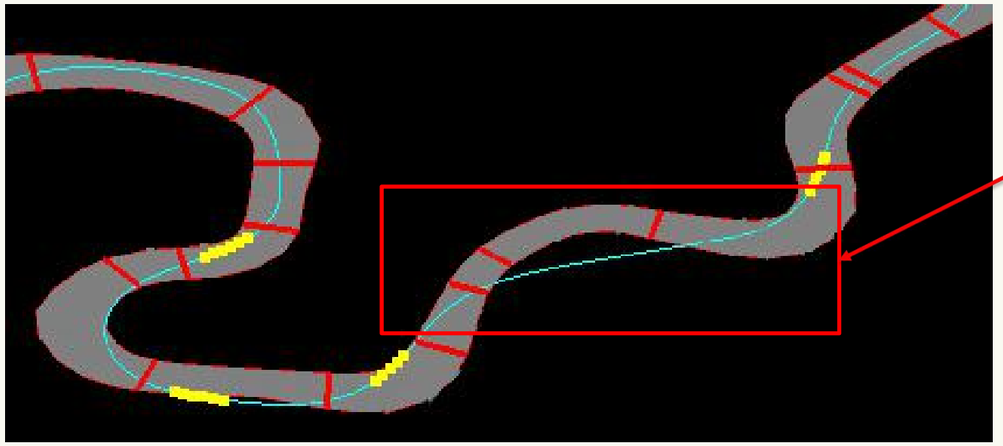
	route_0	route_1	route_2	route_3	route_4	route_5	route_6
1001	97	102	111	59	78	NaN	49
1002	105	107	120	63	83	NaN	52
1003	105	107	117	62	83	NaN	52
1004	102	104	114	60	80	NaN	50
1005	116	119	133	69	92	NaN	58
1006	106	108	120	63	84	NaN	54
1007	104	107	117	63	82	NaN	52
1008	119	123	137	72	94	NaN	61
1010	102	105	115	61	81	NaN	51
1011	114	116	128	68	90	NaN	58
1012	117	119	132	70	92	NaN	58
1014	121	123	137	73	95	NaN	62
1015	129	131	145	78	102	NaN	66
1016	126	131	142	78	101	NaN	64
1017	97	100	109	58	77	NaN	48
1018	138	139	156	82	107	NaN	71
1020	124	127	141	75	98	NaN	64
1022	103	107	118	62	83	NaN	52
1023	128	130	142	78	101	NaN	63
1024	135	136	NaN	80	NaN	NaN	NaN
1025	126	129	143	75	101	NaN	65
1027	114	116	127	68	89	NaN	57
1032	123	124	140	74	96	NaN	63
1033	104	106	116	62	82	NaN	52
1035	98	100	112	59	78	NaN	49
1055	94	94	110	55	74	NaN	46
1170	125	127	141	74	98	NaN	63
1223	110	112	125	66	88	NaN	55
1129	101	103	113	61	80	NaN	50

The air wall was in the middle of route_5 and prevented all traffic from passing.

The 1024 car failed to arrive on time because the max speed was set small.

The Secondary Product

- Help the developers to debug



Due to obstacles around the curve did not come into effect in the Guangzhou track, the car could go straight through the house.

Takeaways

- Anybody can try RL: it’s not difficult.
- RL helps game design: not only racing games.

Acknowledgement

- Key Colleagues:
  - Zhongyue Huang
  - Guangjun Zeng
  - Yue Lin
  - Li Zhao
  - Guan Yang
  - Ruimin Shen
  - Weijun Hong
  - Haicheng Chen
  - Xin Wen

GDC × NetEase Games
Passion of gamers

March 21-25,2022
San Francisco,CA

THANKS FOR WATCHING

China: https://hr.game.163.com/recruit.html

Overseas: https://www.neteasegames.com/careers

Twitter

Facebook

Youtube

#GDC22
#NetEase Games