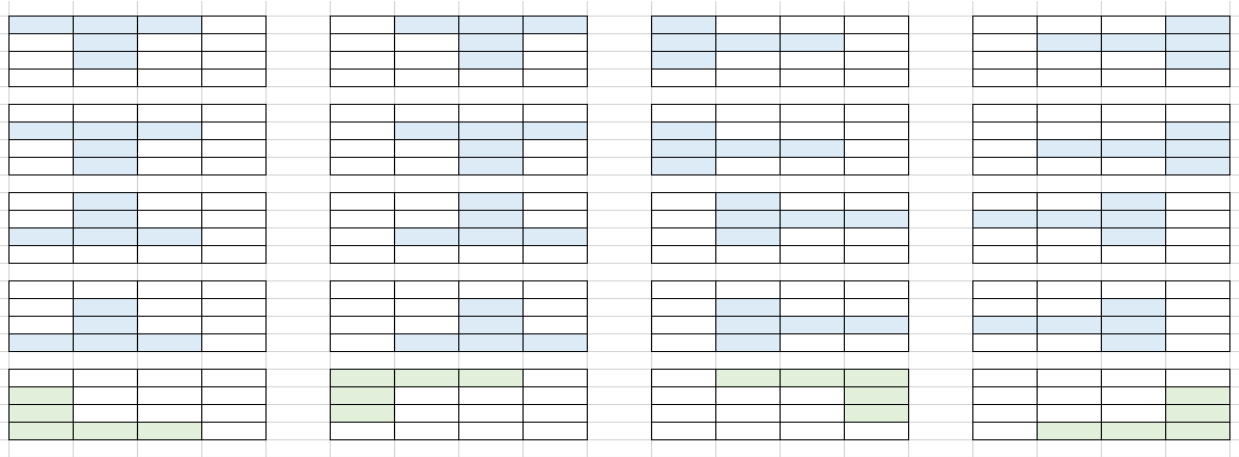# TCG-Project2-2584 Report

Student: 翁玉芯 ID: m093875

Part I : Network Design

In this project, I'd tried many kinds of network and finally found that the 20x5-tuple network has good performance than others, I use the patterns as the picture below, it is the combination of 16 T-shape patterns and 4 L-shape patterns.

Part II : The Used Method

The method I use is backward TD learning, it learn by updating the afterstate value, it's same as the method TA had used during the demo in the lecture.

First, update the last afterstate S'_{T-1}, the adjustment will be **alpha * (0-V(S'_{T-1}))**, where V(S'_t)= the sum of all corresponding feature weights. Next, from t=T-2 to 0, the adjustment will equal to **alpha * (r_{t+1} + V(S'_{t+1}) − V(S'_t))** .

Part III : The Training Process

I train it with alpha = 0.0125 for 60000 games, then I found it is going to converge so I adjust the alpha value to 0.01 for the following 10000 episodes, then 10000 games with alpha=0.008 and 10000 games with alpha=0.005, next I'd tried to train it with alpha=0.003 for 10000 games and found that when the training episodes reach 9000(total training episode = 99000), it has better performance than 10000 (total training episode = 100000), so I stop it at 9000, then keep training it with alpha=0.001 for8000 games. After **112000 episodes** of training the **win rate = 98.9%** and the **average score = 217022**

Part IV : The Training Result

The result is as the 2 figures below, the left one shows the growth of win rate during the training, and the right one shows the growth of the average score. There's also detailed statistic is below, max tile

means the number of tiles that are larger than 2584 and max tile % means the percentage of the largest tile
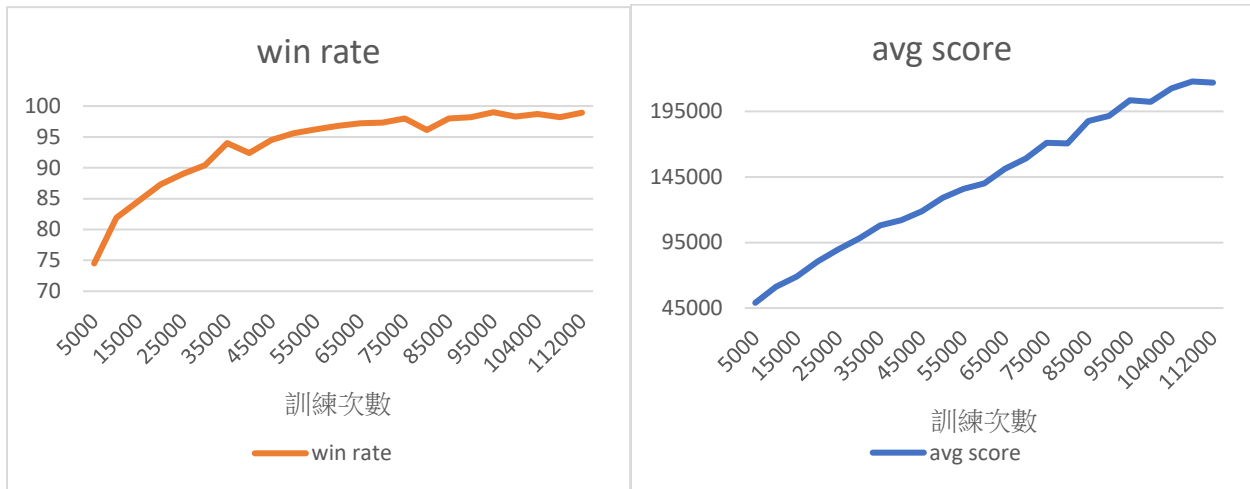
train	alpha	win rate	max tile	max tile %
5000	0.0125	74.5%	+2	1.5
10000	0.0125	81.9%	+2	8.9
15000	0.0125	84.6%	+3	0.9
20000	0.0125	87.3%	+4	0.1
25000	0.0125	89%	+4	0.1
30000	0.0125	90.4%	+3	8.5
35000	0.0125	94%	+4	0.5
40000	0.0125	92.4%	+4	1.3
45000	0.0125	94.5%	+4	1
50000	0.0125	95.6%	+4	2.2
55000	0.0125	96.2%	+4	2
60000	0.0125	96.8%	+4	3.5
65000	0.0100	97.2%	+4	6.2
70000	0.0100	97.3%	+4	6.7
75000	0.0080	98%	+5	0.1
80000	0.0080	96.1%	+4	11.1
85000	0.0050	98%	+4	16
90000	0.0050	98.2%	+4	14.7
95000	0.0030	99%	+5	0.4
99000	0.0030	98.3%	+5	0.8
104000	0.0010	98.7%	+5	0.7
109000	0.0010	98.2%	+5	1.3
112000	0.0010	98.9%	+5	1.2