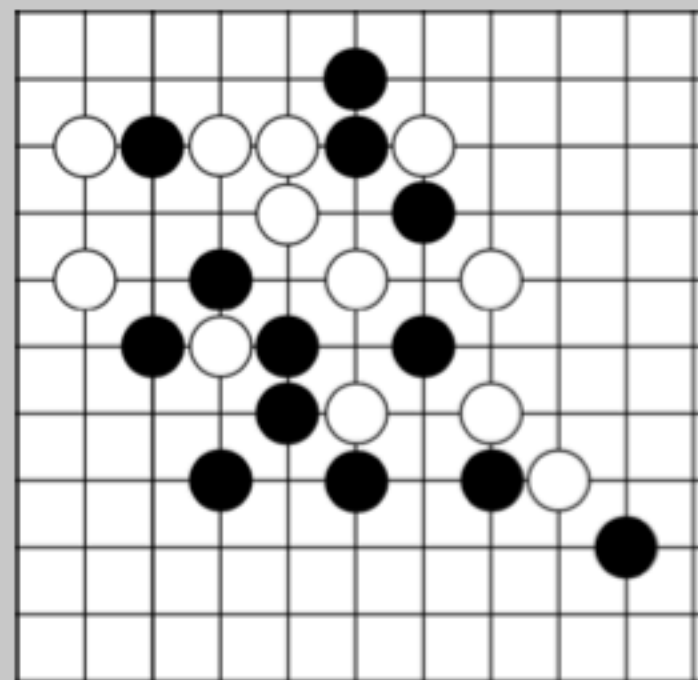


# Self-Play

2	1	0	0	1	2	0	2	0
0	1	2	1	2	0	2	0	2
2	2	1	1	0	2	0	2	2
0	1	0	1	1	0	0	1	2
2	0	1	0	1	2	2	1	1
2	2	1	2	0	1	2	0	1
1	2	1	0	0	2	2	2	1
0	0	0	0	1	2	1	1	1
1	0	1	2	0	0	1	0	1



Go board engine

Agent #1

Agent #2

Game records

Experiences

Game records

DRL

[0.0.1.0.0.0.0.0.0.0.]
[0.0.0.0.0.0.0.0.0.0.]
[0.0.0.0.0.0.0.0.0.0.]
[0.0.0.0.0.0.0.0.0.0.]
[0.1.0.0.0.0.0.0.0.0.]
[0.0.0.0.0.0.0.0.1.0.]
[0.0.0.0.0.0.0.0.0.0.]
[0.0.0.0.0.0.0.0.0.0.]

ResNet

policy:  $\pi_{\theta}(s,a)$

value:  $v_w(s)$

Loss

policy and value targets

2	1	0	0	1	2	0	2	0
0	1	2	1	2	0	2	0	2
2	2	1	1	0	2	0	2	2
0	1	0	1	1	0	0	1	2
2	0	1	0	1	2	2	1	1
2	2	1	2	2	0	2	0	1
1	2	1	0	0	2	2	2	1
0	0	0	0	1	2	1	1	1
1	0	1	2	0	0	1	0	1