



is a great application for your next meta-learning algorithm!

We found *high quality* bots with *diverse strategies*, and put them in a deep-learning-friendly sandbox (aka the Hanabi Learning Environment)

First player	Simplebot	Valuebot	Holmesbot	Outer	Iggi	Piers	Rainbow	Van-Den-Bergh
Simplebot	16.8	15.7	12.8	0.0	4.1	1.1	1.7	0.2
Valuebot	15.2	18.0	17.6	0.0	3.8	1.3	2.0	0.0
Holmesbot	11.2	18.3	14.7	0.0	1.4	0.6	0.5	0.0
Outer	0.0	0.0	0.0	14.1	1.0	4.1	6.2	9.0
Iggi	3.9	3.8	1.8	1.8	16.2	11.8	2.7	6.0
Piers	1.8	0.5	0.2	7.2	10.3	15.9	5.5	9.4
Rainbow	0.3	2.0	0.3	6.6	4.0	5.6	18.1	2.6
Van-Den-Bergh	0.0	0.2	0.0	10.6	4.7	8.3	2.7	10.5

github.com/aronсар/hoad

Here they are playing 2p games with each other. **Cooperation** is easy only when both players are on the same page!

How did we do it?

Original bots play games in native

env



Recreate* games in HLE



Train MLP to imitate original bots



These MLPs can now play with one another!

Agent	Original Self Play Score	Imitator Self Play Score	Imitator Accuracy
Simplebot	16.9	16.8	99.7
Valuebot	19.8	18.0	92.0
Holmesbot	20.8	14.7	90.3
Outer	14.5	14.1	66.7
Iggi	17.0	16.2	90.9
Piers	17.3	15.9	85.8
Rainbow	18.5	18.1	77.5
Van-Den-Bergh	14.0	10.5	81.2

*We observe only the starting deck and player actions in the native env, and replay the game in HLE, thus we save time by not having to translate the game representation from the source language to Python