

Results Round 1

Each agent played 100 games with 4 instances of itself and 3 instances of a non-negotiating Baseline Agent. We counted the average number of Supply Centers conquered per agent instance per game. We say an agent 'passes Round 1' if it scores a statistically significant higher average than the Baseline agent.

Only CoalitionBot and Gunma were able to outperform the Baseline agent. However, the results of Gunma were not significant. Therefore, CoalitionBot is the only agent that managed to pass Round 1.

	Total	Average	Std. Error	Result
CoalitionBot	2211	5.528	0.123	PASS
Baseline Agent	1189	3.963	0.149	

Table 1: CoalitionBot defeated the Baseline Agent with a p-value of $1.9 \cdot 10^{-8}$.

	Total	Average	Std. Error	Result
Gunma	1980	4.95	0.195	fail
Baseline Agent	1420	4.733	0.205	

Table 2: Gunma defeated the Baseline Agent, but with a p-value of 0.47, so this is not significant.

	Total	Average	Std. Error	Result
Baseline Agent	1479	4.93	0.209	
M@sterMind	1921	4.803	0.175	fail

	Total	Average	Std. Error	Result
Baseline Agent	1632	5.44	0.242	
GamlBot	1768	4.42	0.174	fail

Results Round 2

In round 2 we played 100 games with 1 instance of each agent together with 3 instances of the Baseline Agent. We see that when negotiating with other agents the CoalitionBot is not able to significantly outperform the other agents.

	Total	Average	Std. Error
Gunma	569	5.69	0.3
GamlBot	531	5.31	0.334
CoalitionBot	494	4.94	0.289
Baseline Agent	1362	4.54	0.168
M@sterMind	444	4.44	0.29

Conclusions

In order to win the challenge an agent had to pass round 1, as well as beat the other agents in round 2. None of the agents managed to do this, so the Challenge ends with no winner.

Although CoalitionBot was able to outperform the non-negotiating agent, it only performs well when negotiating with other instances of itself. When it comes to negotiating with other agents, the best agent seems to be Gunma, but only by a small difference.

Roughly speaking we can interpret this as CoalitionBot being the best at *making* proposals, while Gunma being the best at evaluating and *accepting* incoming proposals.

Finally, it is worth noting there was a 5th submission, called DDAgent2, but this agent was too slow to participate. Most of the time it was not able to submit its orders before the deadline making it pointless to even complete a single game.