C Illustrations

We include several figures – tables and illustrations:

- Figure 2 contains an illustration of the abstract model.
- Figure 3 and Figure 4 contains graphs regarding the location based simulations. In particular, Figure 3 illustrates the average number of iterations required for different parameter combinations to reach a majority coalition. This is depicted as a function of the total number of agents, assuming convergence within $10,000$ iterations. To gain a deeper understanding of how different parameter combinations influence the convergence, we conducted a detailed analysis as well. In Figure 4, we present the mean of the average distances between each ideal point of an agent and the coalition points, specifically those that have reached a majority.

  Note that,

  - The linear regression model for predicting the average number of iterations resulted in an adjusted R^2 of 0.683. Coefficients provide insights into the linear connection. However, for the linear regression model on mean distances, the adjusted R^2 is 0.177, reflecting a noisier and less explainable relationship, as observed in Figure 5, 7, and 8 (in the supplementary material).
  - The logistic regression model demonstrates $F1$ score 0.89, reflecting a strong ability of prediction.
- Figure 5 contains a graph regarding the text based simulation.
- We illustrate in Table 2 the example given in the main text.
- We illustrate another example to show the capabilities of our model with fixed parameters outlined in Figure 7 and Table 2. This example also involves 10 ideal sentences of agents but this time deals with ways to fight for rights of minorities (of maximum 15 words).
- Figures 6, 8, 9, and 10 shows raw statistical analysis.
- Figure 11 shows the different agent distributions used in the experiments.

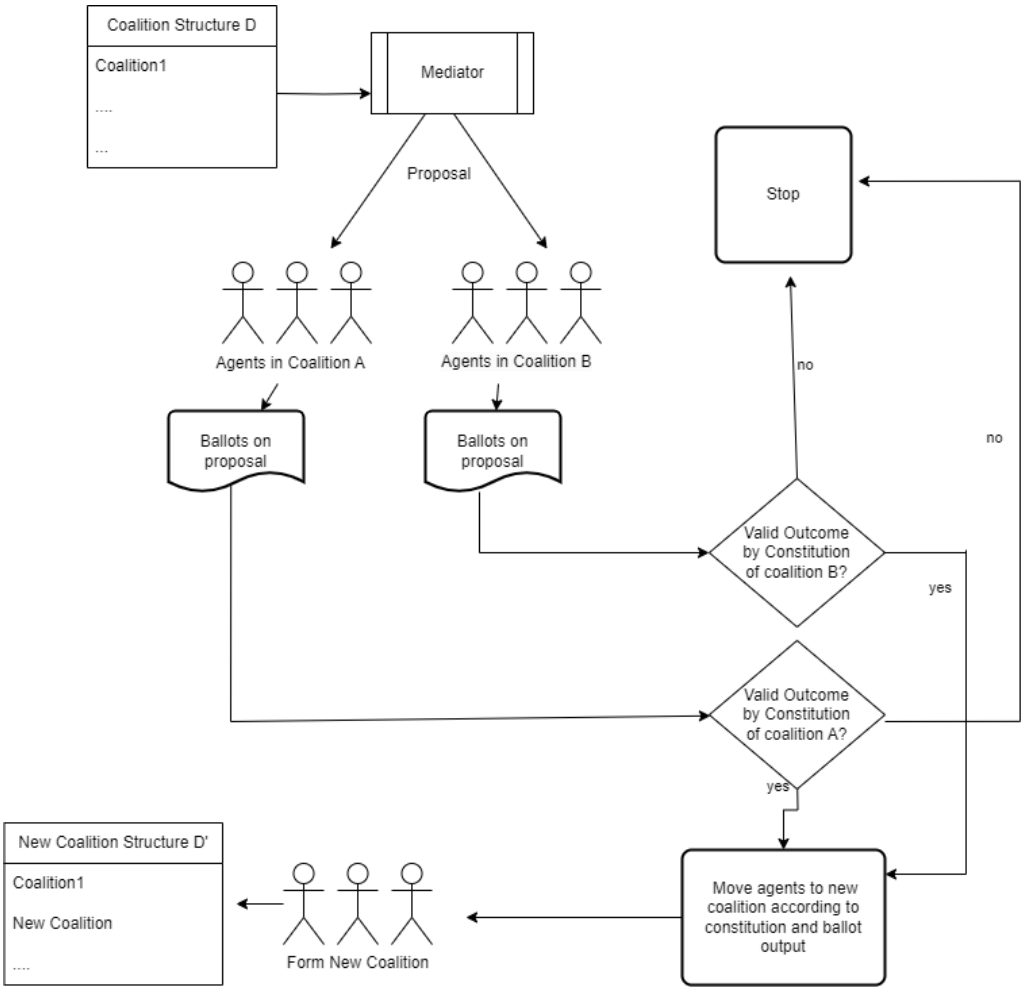

Figure 2: Illustration of the Abstract Model.

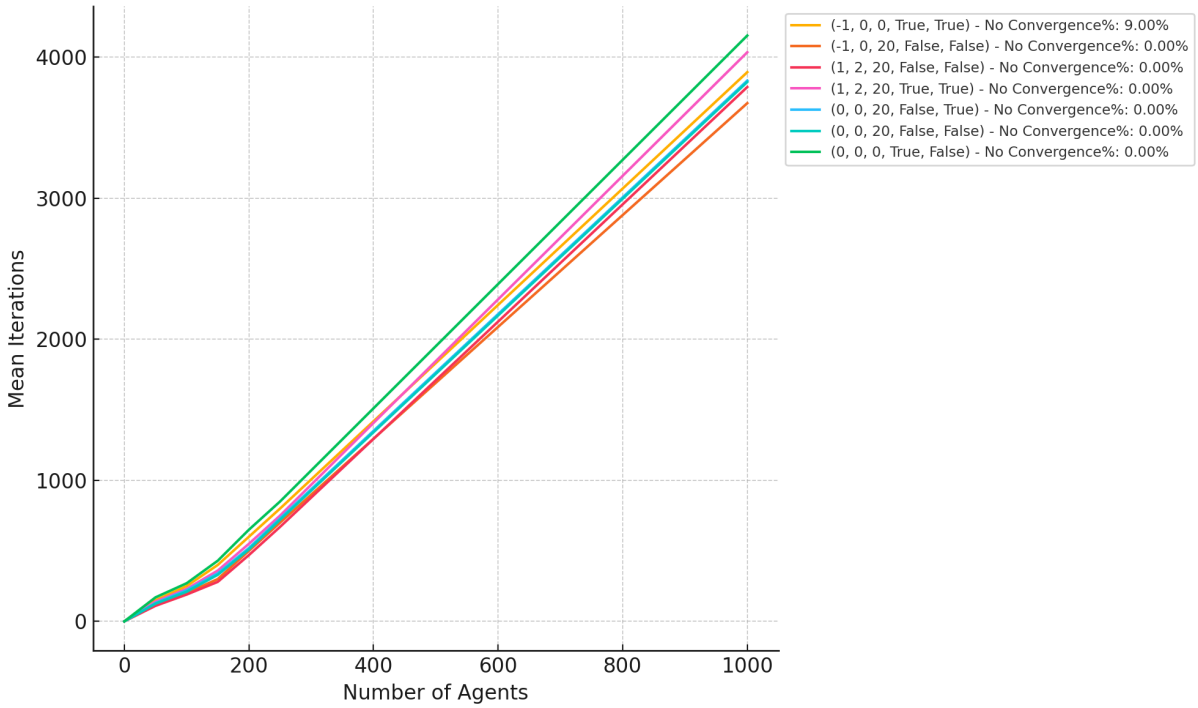

Figure 3: Location Use case- Speed of convergence (i.e., average number of iterations) as a function of the number n of agents; each of the line corresponds to different combination of the parameters α, g, σ, coalition discipline, I.

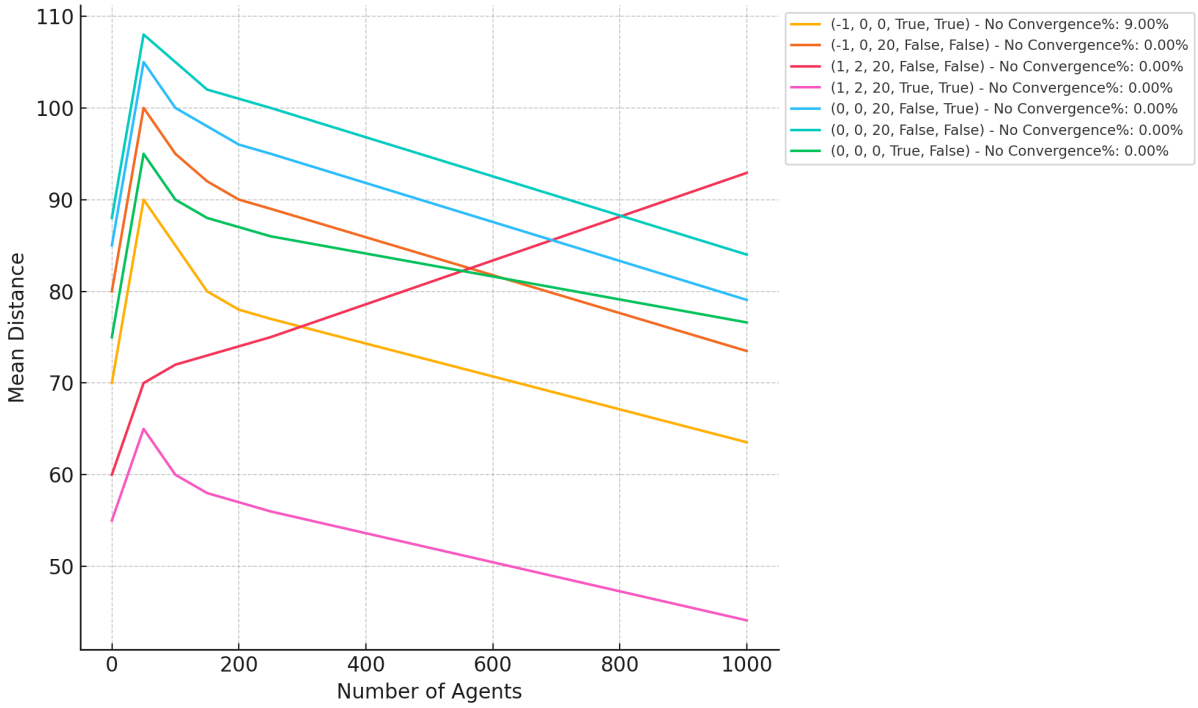

Figure 4: Location Use case- A noisy outcome: number of agents VS average distance chosen combinations of α, g, σ, coalition discipline, I.

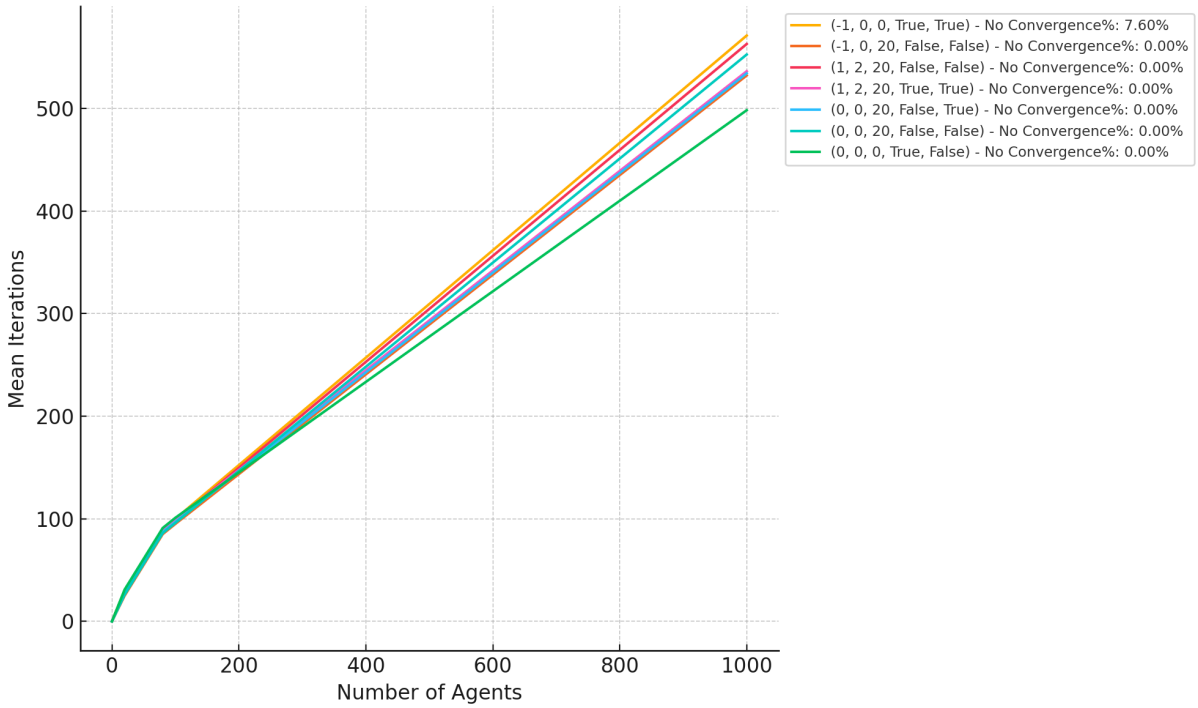

Figure 5: Text use case: Number of agents as a function of the average number of iterations and chosen combinations of α, σ, C, I.

R-squared:	0.685		
Model:	OLS	**Adj. R-squared:**	0.683
Method:	Least Squares	**F-statistic:**	345.0
Prob (F-statistic):	6.76e-235		
Time:	13:26:49	**Log-Likelihood:**	-3965.6
No. Observations:	960	**AIC:**	7945.
Df Residuals:	953	**BIC:**	7979.
Df Model:	6		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
intercept	-6.6379	1.364	-4.866	0.000	-9.315	-3.961
n	1.5814	0.066	24.030	0.000	1.452	1.711
n·α	0.1633	0.022	7.485	0.000	0.120	0.206
n·peaks	-0.0339	0.016	-2.146	0.032	-0.065	-0.003
n·C	0.5116	0.057	9.051	0.000	0.401	0.622
peaks·C	-1.6322	0.572	-2.856	0.004	-2.754	-0.510
sigma·C	0.4137	0.058	7.085	0.000	0.299	0.528

Figure 6: Linear Regression for Average Iterations Results- Location use case.

Iter. 1	Iter. 2	Iter. 3	Iter. 4	Iter. 5	Iter. 6	Iter. 7
0: "Educate yourself on the issues facing minority communities"		1,0: Amplify the voices of marginalized groups				
1: "Use your privilege to amplify marginalized voices"		1,0: Amplify the voices of marginalized groups				
2: Advocate for policies that promote equality						
3: "Join or support organizations advancing minority rights"					3,7,8,4: "Engage in advocacy and support minority rights"	6,3,7,8,4: "Participate in protests to show solidarity"
4: "Engage friends and family in conversations"	8,4: "Engage friends and family while standing in solidarity"		7,8,4: "Support businesses for marginalized empowerment"		3,7,8,4: "Support advocacy and minority rights"	6,3,7,8,4: "Participate in protests to show solidarity"
5: "Challenge discriminatory behavior"				5,9: "Confront discrimination and learn intersectional issues"		
6: "Vote for political candidates who prioritize minority rights"						6,3,7,8,4: "Participate in protests to show solidarity"
7: "Support minority-owned businesses"			7,8,4: "Support businesses for marginalized empowerment"		3,7,8,4: "Support advocacy and minority rights"	6,3,7,8,4: "Participate in protests to show solidarity"
8: "Stand in solidarity during protests"	8,4: "Engage friends and stand in solidarity"		7,8,4: "Support businesses for marginalized empowerment"		3,7,8,4: "Support advocacy and minority rights"	6,3,7,8,4: "Participate in protests to show solidarity"
9: "Learn about intersectional issues"				5,9: "Confront discrimination and learn intersectional issues"		

Table 2: Matrix representation of a coalition formation process fighting for the rights of minorities. Each column represents an iteration, and each row corresponds to an agent. Initially, each agent is in its singleton coalition. The table shows the compromise sentences accepted by pairs of coalitions as the process evolves.

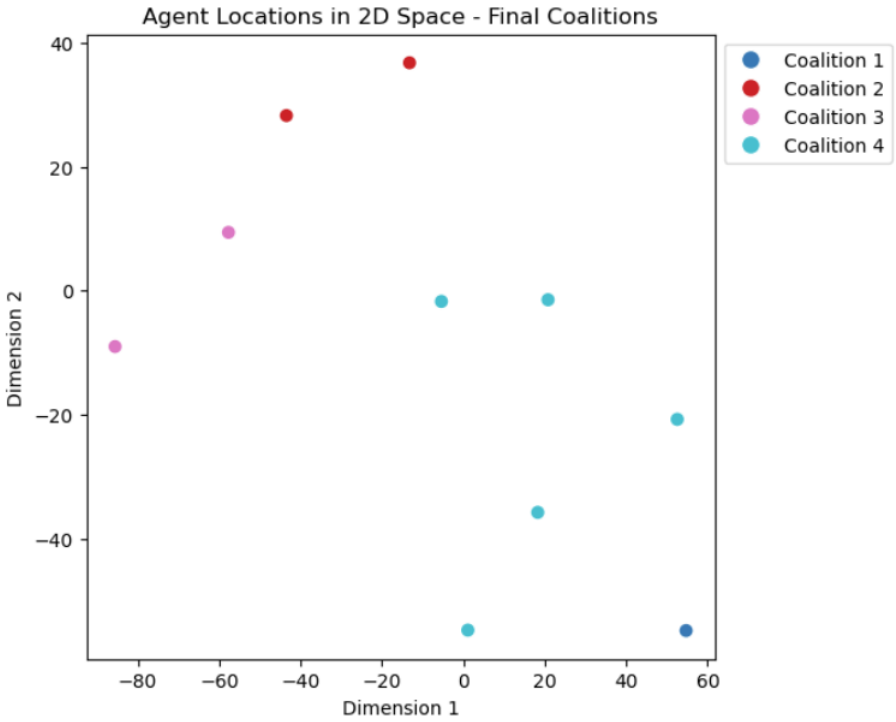

Figure 7: Coalition formation result- "Fighting for Rights of Minorities", for $n = 20, C = \text{False}, \alpha = 0, \sigma = 0, I = \text{True}$.

Model:	Logit	**Df Residuals:**	762
F1 Score:	0.89	**Pseudo R-squ.:**	0.5432
Log-Likelihood:	-212.55		
converged:	True	**LL-Null:**	-465.36
LLR p-value:	4.928e-107		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
n	-0.0425	0.009	-4.785	0.000	-0.060	-0.025
α	0.3311	0.153	2.163	0.031	0.031	0.631
peaks	-0.2525	0.080	-3.140	0.002	-0.410	-0.095
sigma	-0.1620	0.015	-11.152	0.000	-0.190	-0.134
C	5.3229	0.404	13.170	0.000	4.531	6.115
I	-1.2725	0.253	-5.037	0.000	-1.768	-0.777

Figure 8: Logistic Regression for the probability of encountering non-convergent results-Location use case.

R-squared:	0.180				
Model:	OLS		**Adj. R-squared:**		0.177
Method:	Least Squares		**F-statistic:**		69.75
Prob (F-statistic):	8.53e-41				
Log-Likelihood:	-3986.9				
No. Observations:	960		**AIC:**		7982.
Df Residuals:	956		**BIC:**		8001.
Df Model:	3				
	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
const	53.4245	1.317	40.552	0.000	50.839	56.010
n	0.4051	0.045	9.096	0.000	0.318	0.493
α	2.8802	0.610	4.723	0.000	1.683	4.077
C	-10.1666	0.996	-10.209	0.000	-12.121	-8.212

Figure 9: Linear Regression for Average Distance.

R-squared:	0.999				
Model:	OLS		**Adj. R-squared:**		0.999
Method:	Least Squares		**F-statistic:**		2.404e+04
Prob (F-statistic):	4.35e-89				
Time:	11:41:12		**Log-Likelihood:**		-109.19
No. Observations:	64		**AIC:**		224.4
Df Residuals:	61		**BIC:**		230.9
Df Model:	2				
	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
const	-4.0914	0.288	-14.188	0.000	-4.668	-3.515
n	1.0078	0.005	218.869	0.000	0.999	1.017
alpha·C	0.9930	0.308	3.224	0.002	0.377	1.609

Figure 10: Linear Regression for Average Iterations Results- Text use case.

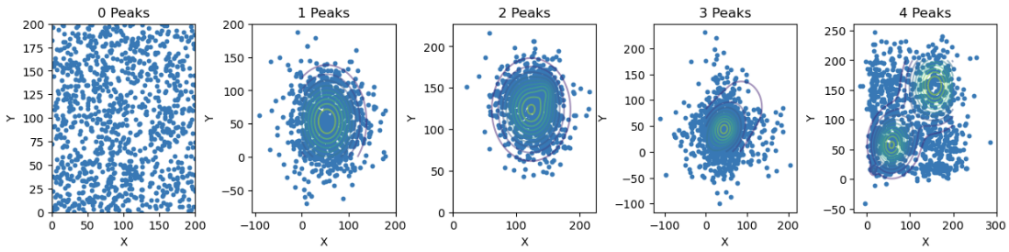

Figure 11: 1000 agents sampled from a GMM with a mixture of 0,1,2,3,4 Gaussians over a Euclidean 2D grid over (0,200) square, using weights distributed by Dirichlet distribution. The color of the contour lines represents different density levels in the mixture model; for 0 peaks, the density is constant across the grid.