ABM Paper:
Uninformed Individuals Promote Democratic Consensus in Animal Groups

Final Report for the Seminar
“Computational Modelling of Social Systems”

submitted by
Gina-Maria Unger
[deleted]

at the

Universität Konstanz

Faculty of Law, Economics and Politics

Department of Politics and Public Administration

Examiner: Professor Doctor David Garcia

Konstanz, 2023

Table of Content

List of Figures

1. Motivation for Modelling Collective-Decision Making Processes

When it comes to reaching consensus in collective decision-making, it is possible that in the absence of formalised democratic processes, a numerically smaller but strongly opinionated minority overpowers a numerically bigger majority. For instance, the modelling results of Galam and Jacobs (2007) suggest that if a minority is more inflexible in changing its opinion than a majority, it is more likely, that the decision outcome is in favour of the minority. Furthermore, it is generally assumed that the minority's influence on the process outcome is even stronger, when uninformed individuals, namely individuals that "lack a preference or are uninformed about the […] the collective decision" (Couzin et al, 2011), are a part of the deciding group. This assumption is made, as uninformed individuals are viewed as more vulnerable to manipulation by an opinionated minority and thus might facilitate a minority-preferred decision outcome. Given this frame the work of Couzin et al. (2011) addresses two research questions:

1. *Whether and if so, under which conditions can a self-interested and strongly opinionated minority exert its influence on consensus decision-making?*
2. *What role do uninformed individuals play in this process?*

Although their work was published in 2011 these questions are still of great importance considering that some of the most influential political institutions, such as the Council of the European Union, decide on consensus-based rules.

To get an overview of Couzin et al. (2011) work, it will be discussed in the following sections. More precisely, the following second section presents the spatial model the authors relied on to model consensus-decision making processes in the presence of an opinionated minority and uninformed individuals. Further, this section shortly outlines the model's explicit and implicit assumptions. In the third section, the main results of this modelling process will be stated and compared to a limited extent to own results. The fourth section compares other models discussed in the lecture to the spatial model. Lastly, the final fifth section contains a general reflection on Couzin et al. (2011) agent-based modelling approach, laying the focus on the quality of the modelling approach itself and the model's plausibility.

2. Couzin et al.'s Model Formalization

When it comes to the modelling collective-decision making processes, Couzin et al. (2011) rely on three different types of models. More precisely, their main models are a spatial model and a simplified version thereof. These models are used to model the collective consensus decision-making context, study underlying mechanisms, and give possible answers to the research questions. Furthermore, since a spatial context might not be very realistic, Couzin et al. (2011) also implement an adaptive network and convention model to check if the spatial model's results can be generalized. In the subsequent, the full spatial model will be described.

<u>Spatial Model</u>

Couzin et al. (2011) choose to stick to a slightly customized spatial model from Couzin et al. (2005) that was developed for studying collective dynamics in animal groups. This model allows modelling the process of consensus decision-making in groups for a binary choice in the presence of two or more subgroups – here the majority, minority, and uninformed. The binary choice, as well as the group individual's current opinion, preferred choice, and strength of choice preference, get represented as points and vectors in a vector space. Further, in this space, individuals can also interact, hence influence each other's current status.

Describing the model in a more mathematical manner, one can denote the size of the whole group with N, then N_1, N_2, and N_3 state the sizes of the subgroups – namely the majority, the minority, and the uninformed. Each individual i gets represented at time step t by its position $c_i(t)$ in the vector space and a unit direction vector $\hat{v}_l(t)$, that gives its motion intentions. An individual modifies its motion in the next time step t according to its social and goal desires. When it comes to its social desires, two zones, which are centered around the focal individual, are of importance. The repulsion zone with a radius of α and at the traction zone with a radius of ρ, where α must be smaller than ρ. Concretely, if other individuals j are in individual $i's$ repulsion zone, it moves away from them according to the rule:

$$s_i(t+\Delta t) = -\sum_{j\neq i} \frac{c_j(t)-c_i(t)}{\|c_j(t)-c_i(t)\|}$$

In simple words, this could be expressed as: "The individual i tries to move in the opposite direction as of most of the individuals j in its repulsion zone". Otherwise, if no other individuals j are in individual $i's$ repulsion zone, but some are in its attraction zone, it aligns its motion with them consistent with the rule:

$$s_i(t+\Delta t) = \sum_{j\neq i} \frac{c_j(t)-c_i(t)}{\|c_j(t)-c_i(t)\|} + \sum_j \hat{v}_j$$

That means the individual i takes the current position of the other individuals j as well as their motion direction into account for its next movement. Besides this social desire, individuals might also have some goal desires, namely, that their preferred outcome gets selected or equivalent to reach their preferred target location. Consequently, each individual i has a preferred direction of travel $\hat{g}_l$, which is a continuously changing unit vector pointing at its preferred spatial target. Furthermore, the individual possesses a weighting term ω_i indicating its preference strength for its preferred outcome. Thus, if ω_i is close to 0, the respective individual i has no desire to move in its preferred direction or is uninformed. If ω_i approaches 1, the individual i weights its social and goal desires equally. If ω_i exceeds 1, the individual i weights its goal desires more strongly. Ergo, ω_i can be used to set the strength of subgroup opinion strengths, what specifically allows to model situations with strongly opinionated minorities. Thus, ω_1 and ω_2 get defined as the group weighting terms for the subgroups N_1 and N_2. Combining everything together, the individual $i's$ social and goal desires for the next time step t can be denoted as:

$$d_i(t+\Delta t) = \frac{s_i(t+\Delta t)}{\|s_i(t+\Delta t)\|} + \omega_i \hat{g}_l$$

One can clearly see that if the individual's goal desires are greater than its social desires the second part of the formula outweighs the first. In addition, by further normalizing this formula, the individual $i's$ desired motion direction for the next time step t can be retrieved as:

$$\hat{d}_i(t+\Delta t) = \frac{d_i(t+\Delta t)}{\|d_i(t+\Delta t)\|}$$

Further it is assumed that the individual motion $\hat{d}_i(t+\Delta t)$ is also a subject to random influences. Thus, $\hat{d}_i(t+\Delta t)$ gets rotated by a fixed angle θ.

When it comes to the specific model parameters and settings Couzin et al. (2011) placed the majority target at $(100, -50)$ and minority target at $(100, 50)$. At the simulation start all individuals were initialized with random starting positions and orientations within a box 2 by 2 units that was centered at $(0, 0)$. After a transition period of 200 time steps, which is approximately the time needed until a group gets stable, the group got recentered at (0,0) and was allowed to approach the targets. If the group centroid was within 4 units of a target, this target choice was recorded as the result of the respective trial. Overall, 20000 replications were made for each combination of the parameters N_1, N_2, N_3, ω_1 and ω_2. The remaining fixed parameter values were selected to be $\alpha = 1$, $\rho = 6$, and $\theta = 2$ for all trials.

Summarizing all assumptions made in the modelling process, the following picture emerges:

1. A binary choice, which a group must consensually decide on, can be represented as two target points in a vector space.
2. A group individual, its current position, preferred choice, and choice preference strength, can be defined by scalars, points, or vectors in the vector space.
3. Each individual is a member of one subgroup, of which at least two exist and that have either a preferred decision outcome they try to reach or no such aspirations. The first two subgroups prefer each one of the targets in the spatial decision space and are of different sizes. The third subgroup, if existent, is not focused on any target.
4. Individuals avoid collisions with other individuals in the vector space.
5. Individuals are attracted towards each other and can probabilistically align their direction of travel with near neighbors.

However, when taking a closer look, one must notice that these assumptions are not exhaustive. Since it is assumed that individuals have social desires that manifest by them avoiding collisions with others and show attraction towards others, one must further assume that individuals know the position and intended motion direction of other individuals. This implicit assumption is quite problematic, as such all-encompassing information statuses of individuals are rather unrealistic when it comes to real world consensus-decision making.

3. Spatial Model Results

Summarizing the results of Couzin et al. (2011) modelling process can be reduced to stating their three main observations, which will be outlined in the subsequent. Firstly, the simulations indeed showed that self-interested and strongly opinionated minorities can exert their influence on group movement decisions, when their preference strength for a decision outcome exceeds the majority's counterpart. For instance, when having a glance at the *Figure 1*, one can see that for an increasing minority preference strength and a constant majority preference strength, the majority's target gets less reached:

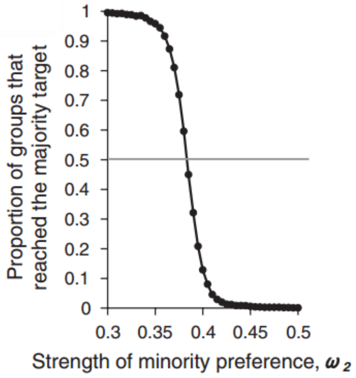

Figure 1: Proportion of groups that reached the majority target depending on the strength of the minority preference ω_2 for $\omega_1 = 0.3$, $N_1 = 6$, and $N_2 = 5$

Secondly, uninformed individuals can weaken the minority's influence and hence indirectly assist the majority. Precisely, uninformed individuals tend to adopt the opinion of the local majority, what results in settings with a small number of uninformed individuals in an advantage for the majority and in settings with an increasing number in a close-to-equal probability of selecting either outcome. Viewing *Figure 2* one can notice these distributional variations:

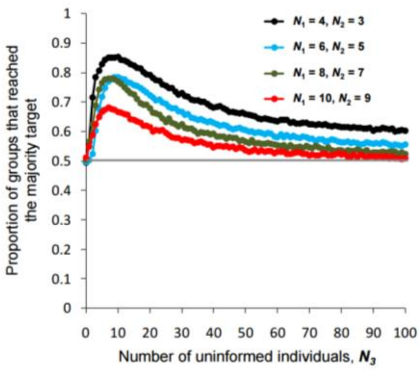

Figure 2: Proportion of groups that reached the majority target depending on the number of uninformed individuals N_3 for a range of N_1 and N_2 values

Thirdly, the last finding of importance is that if the number of uninformed individuals increases to much, relatively to the number of individuals in majority group, a consensus decision becomes more unlikely and the group gets unstable. Specifically, that behavior occurs as the tendency of the uninformed individuals to adopt the opinion of the local neighborhood works not as a system stabilizing factor if the neighborhood mostly consists of uninformed individuals too. Hence, the benefit of the uninformed individuals for the majority is discounting.

To complement these results, one of the objectives for this report was to reproduce them to some extent. However, since the implementation of the spatial model and the run of simulations are quite complicated, this objective could just be fulfilled to a very limited extent. More precise, three major obstacles arose due the result reproduction process:

1. *The coding of the model components and inherent model dynamics*
2. *The lack of specific model information*
3. *The lack of computational power*

The first obstacle could be solved to the extent that all described model components and inherent dynamics got encoded. When it comes to the parameter settings, the same settings were used as selected by Couzin et al. (2011), except that the number of replications was reduced to 100 due the lack of computational power. Nevertheless, although the given model information is extensive, for the matters of a model implementation some aspects are still missing or not precise enough. For instance, Couzin et al. (2011) never stated if uninformed individuals have a preferred direction of travel $\hat{g}_l$. They could be either initialized without one or get assigned to a preferred target by random, thus having a preferred direction of travel $\hat{g}_l$. Since their respective weighting term ω_3 is defined as 0 this differentiation is not of great importance as $\hat{g}_l$ would drop out of the formula either way, but as this missing of information is not an exception, it is problematic factor when it comes to implementation matters.

Due these limitations the reproduction of the results was limited to the first finding that was: “self-interested and strongly opinionated minorities can exert their influence on group movement decisions”. Hence, trying to reproduce *Figure 1* resulted into the following:

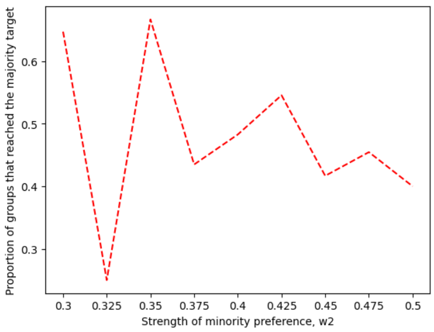

Figure 3: Reproduced *Figure 1* for variable strengths of the minority preference $\omega_2 = \{0.3, 0.325, 0.35, 0.375, 0.4, 0.425, 0.45, 0.475, 0.5\}$, $\omega_1 = 0.3$, $N_1 = 6$, and $N_2 = 5$

Unfortunately *Figure 3* does not compare well to *Figure 1*, both distributions differ greatly in their shape even when considering the fact that *Figure 3* looks more angular as it was produced with less data points. Contrasting both figures the reproduced results show a fluctuating trend for the effect of the strength of the minority preference ω_2 on the majority target approaches rather than a decreasing. Most likely one of the stated obstacles or a mix thereof led to this mismatch in simulation results. Nevertheless, one must note that *Figure 3* shows, although having a fluctuating nature, a general downward trend what can be interpreted in favor of Couzin et al. (2011) results. Having also an additional look at the number of groups splits for $\omega_2 = \{0.3, 0.325, 0.35, 0.375, 0.4, 0.425, 0.45, 0.475, 0.5\}$ per 100 replications the assumption hardens that something must have gone wrong due the model reimplementation process, as the numbers of $splits = \{83, 88, 85, 77, 71, 78, 76, 78, 85\}$ are very high. One explanation for this result and the high splitting numbers could be that the maximal number of time steps for the group to approach the targets was set too low (fixed to 1000), unfortunately this limit must be set due the missing of more powerful computational resources.

4. Comparison to Other Lecture Models

Comparing Couzin et al. (2011) spatial model to other models that were discussed in the lecture, the model seems to have the most similarities with the linear and non-linear voter models for modelling opinion dynamics. For instance, having a glance at the voter models, one can see that the spatial model's definition of an individual i over its position $c_i(t)$ at time step t compares quite well to those of an agent i being defined over its opinion θ_i at a given time step t. Nevertheless, the spatial model allows for the modelling of situations with multiple subgroups, while both voter models are limited to settings with two subgroups. Otherwise, when taking the number of outcomes an individual or agent can choose from into account, all models are restricted to a binary choice, which is in the case of the spatial model implemented over two spatial targets and for the voter model as the agent's choices 0 and 1.

Furthermore, in all three models a possible interaction of the individual or agent with the local neighborhood gets implemented, that can shape its current position or opinion. In the spatial model this process is implemented by the repulsion and attraction rule, and in the voter models by taking the frequency of agents with opposite opinions in the neighborhood $f_{\mathrm{i}}^{1-\theta_i}$ into account. Both, the spatial and the non-linear voter model also incorporate some sort of preference strength for a specific outcome that can reduce the chance to exhibit an opposing opinion than the initially preferred one. In the spatial model this is enabled by the individuals i weighting term ω_i, while in the non-linear voter model the response function to the frequency of other opinions in the local neighbourhood $\kappa(f)$ can be adjusted in a respective manner. Contrary, the linear voter does not allow for the incorporation of such a differing behaviour. Further, compared to the other models, the non-linear voter model is the only model that can take the past into account by incorporating a memory effect.

Lastly, when looking at the model outcomes the spatial model leads, depending on the selection of ω_i and N_3, either to a reach of consensus or to the coexistence of opinions. Otherwise, in the linear voter model consensus always appears as outcome, while the non-linear voter model also facilitates consensus and group splits as outcomes. In sum one can say that the spatial model is of great similarity to the voter model, and specifically of similarity to the non-linear voter model.

Besides the voter models some results of the adaptive network model can be compared in their behavioural dynamic to Granovetter's threshold model. Glancing at the following figures:

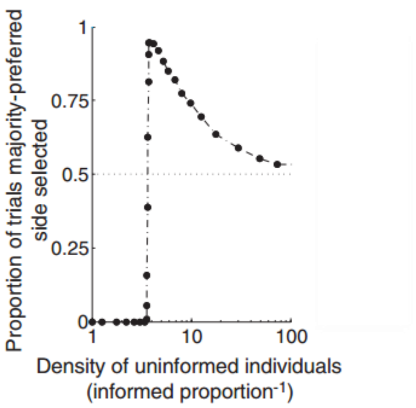

Figure 4: Proportion of trials the majority-preferred side gets selected in the network based on the density of uninformed individuals

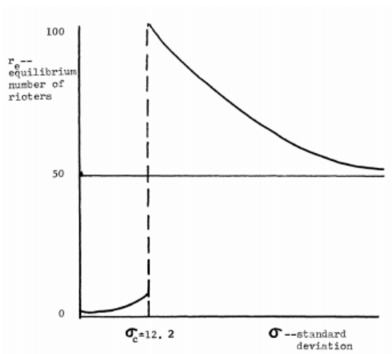

Figure 5: Equilibrium number of rioters r_e based on the standard deviation σ of the distribution of thresholds to join collective action

One can see that the proportion of trials in which the majority-preferred side is selected based on the density of uninformed individuals follows a nearly identical distribution as the equilibrium number r_e based on the standard deviation σ of the distribution of thresholds. Both distributions show a sharp increase at a critical value where a phase transition happens, either from a minority to majority win or from less than 50% to more than 50% of possible rioters joining.

5. General Critical Reflection of the Model

Sorting Couzin et al. (2011) modelling approach of consensus decision-making in the context of general agent-based modelling approaches, one can state that the authors try not to explain an empirical observation, but rather test hypotheses in an abstracted system to enable the testing of more specific hypotheses in the real world.

In general, Couzin et al. (2011) modelling approach is of good quality even by nowadays standards. Firstly, the model allows drawing causal conclusions, as the individuals can have different current positions, motion directions, and preferred choices, which can also partly change over time. Hence, for each individual a counterfactual behaviour exists, which permits that the modelling outcome can vary and is not ad hoc. Secondly, the specific modelling outcome or more interestingly the outcomes of multiple simulations are easily measurable, e.g. as the proportion the majority target has been reached. Thirdly, the design of the modelling approach is also quantifiable to a limited extent in the real world. Meanwhile their studies Couzin et al. (2011) conducted also a repeated experiment with fish to test the hypotheses that “uninformed individuals should inhibit the influence of a strongly opinionated minority and return the control to the numerical majority”. The experimental results indicate that uninformed fish help the majority to maintain the control and approach their preferred food target, even when the minority fish group was strongly opinionated and favored an other target. Lastly, the only point of criticism when it comes to the quality of the modelling approach is that the utilized spatial model is not minimal. However, Couzin et al. (2011) corrected for this by implementing a simplified model version that still leads to the same results.

Contrary to Couzin et al. (2011) well performance when it comes the quality of their modelling approach its plausibility it a bit dubious. First of all, the choice to stick to a spatial context is suboptimal and does not compare well to real world consensus-decision settings. Still, the authors control to some extent for this critic point by implementing two non-spatial models – the adaptive network and convention model. Nevertheless, some additionally made model assumptions, besides the already mentioned implicit assumption that individuals have an all-encompassing information status, remain unhandled and problematic. For instance, the assumption that individuals do not want to have the same position or collied with other individuals in the decision space is unrealistic. Moreover, the existence of clearly identifiable, well separable and opposing decision choices does not reflect most real-world decision choices, which are often quite similar in nature and differ just in few aspects. Lastly, as also already referred in the fourth section, the inability of the agents to remember the past sounds unlikely. All in all, when it comes to the model plausibility there might be some room for improvement.

6. Bibliography

Couzin, I. D., Ioannou, C. C., Demirel, G., Gross, T., Torney, C. J., Hartnett, A., ... & Leonard, N. E. (2011). Uninformed individuals promote democratic consensus in animal groups. *science, 334*(6062), 1578-1580.

Couzin, I. D., Krause, J., Franks, N. R., & Levin, S. A. (2005). Effective leadership and decision-making in animal groups on the move. *Nature, 433*(7025), 513-516.

Galam, S., & Jacobs, F. (2007). *The role of inflexible minorities in the breaking of democratic opinion dynamics. Physica A: Statistical Mechanics and its Applications, 381,* 366-376.