Emotions in collective risk dilemmas using fuzzy linguistic rules

Manuel Chica
Dpt. of Computer Science and AI,
University of Granada
manuelchica@ugr.es

Oscar Cordon
Dpt. of Computer Science and AI,
Universidad de Granada
ocordon@decsai.ugr.es

Hussein Abbass
University of New South Wales,
Canberra
h.abbass@unsw.edu.au

Abstract

The goal of this paper is to propose how to model emotions in climate change policies through evolutionary game theory (i.e., collective risk dilemmas) using fuzzy linguistic rules. We will first study evolutionary game theory models to explore climate change policies and decisions to enhance cooperation between policy players. Using computational agent-based simulations, we will build a framework where we incorporate and extend evolutionary game models with players' emotions in the game dynamics. These emotions, modeled by fuzzy linguistic rules, are a way for players to reconsider their strategies when playing the game. Results show how these rules representing emotions can better control defection in the collective-risk dilemmas while modeling a more realistic way of dealing with players' features.

Keywords: Collective risk dilemmas, fuzzy linguistic rules, evolutionary games, agent-based modeling, emotions.

1. Introduction

From COVID-19 restrictions (Chica et al. (2021, 2022)) to climate change decisions (Vasconcelos et al. (2013)), averting catastrophic events often requires individuals to cooperate for the collective good (Ostrom (2010)). This is challenging in social dilemmas related to climate change, as there is a need for a contribution, tempting free riders to take advantage of the efforts of others (Levin (2014)). Here, the key element is balancing selfish interests and common good (Wang et al. (2020)). An example of mitigating climate change is the decision to reduce carbon emissions, which is one of the most important applications of climate dilemmas (Milinski et al. (2011), Pacheco et al. (2014), Tavoni et al. (2011), and Vasconcelos et al. (2014)). One way to facilitate cooperation is to have strong local institutions, which are important for successful governance in public good games (Smirnov (2019)).

The role of local institutions to punish free-riders has been a useful tool for promoting cooperation (Vasconcelos et al. (2013)). In the latter work, the study of a collective-risk dilemma (CRD) (Milinski et al. (2008) and Santos and Pacheco (2011)) demonstrated the importance of local and global institutions to increase the achievement of agreements of public groups. Specifically, CRDs are multiplayer public good games in which every player can contribute some amount to avoid a certain risk of failure. However, institutions are not enough to avoid free riders in climate dilemmas and more interventions are needed (Vasconcelos et al. (2013)).

Emotion emerges along with individual interactions, and numerous experimental studies have shown that emotion plays an important role in individual decision-making. Including emotions in evolutionary games has been a topic of interest in the area. One of the first studies was in Szolnoki et al. (2011). The authors observed how imitating emotions such as goodwill and envy from the most successful players, instead of pure strategies in social dilemmas, was able to resolve social dilemmas in structured populations. In Szolnoki et al. (2013a), authors considered a model in which sympathy and envy are the two emotions that determine the strategy of each player in any given interaction. In the same study, they discovered how players show high sympathy and high envy in lattices and regular random graphs, while heterogeneous networks, such as scale-free distributions, lead to low sympathy and envy. The study of Chen et al. (2021) assumed that

URI: https://hdl.handle.net/10125/111632
978-0-9981331-9-5

emotion can influence one's own willingness to change strategies and quantified emotions in a cumulative way. They also divided players in the network into two types: competitive individuals and non-competitive individuals. Finally, another relevant previous work is done in Chen et al. (2025) where there is a diffusion of emotions. These emotions significantly impact group cooperation in non-linear ways, by also having interesting effects of emotional values, cumulative emotions, diffusion parameters on the evolution of group behavior.

In the case of this paper, our research proposal is to study how players' emotions influence the output and behavior of CRDs. We will first study evolutionary game theory models for exploring climate change policies and decisions to enhance cooperation between policy players (e.g., countries in international organizations such as UN climate meetings (Cole (2015))). Later, we will build a simulation framework in which we can incorporate and extend CRD to incorporate emotions into the model. In climate change management, emotions have an impact and represent a threat when making decisions (Reser and Swim (2011)) and this also applies to the context of cooperating for global climate change policies.

We will model these emotions in a simulation framework using fuzzy linguistic rules in a knowledge base. A fuzzy modeling library is used to extend the traditional evolutionary game theory models for the tragedy of commons in climate change using fuzzy linguistic rules. Some attempts to link emotions with tragedy of commons and evolutionary games were done before but without a focused climate change application (Greenwood et al. (2017) and Szolnoki et al. (2013b)). Thanks to this additional module, modelers and stakeholders will be able to add emotions to the decision making process of the evolutionary game and depict differences and impact on a set of diverse climate change policies (Reser and Swim (2011)).

Each CRD player has a fuzzy rule-based system to model their emotions. These emotions are automatically used, and their parameters are updated at every time step of the simulation. The goal of these rules is to decide whether to trigger or not the traditional update rule (normally the Fermi function) to copy others' strategies. In a nutshell, if the emotions of the players overpass a given threshold, the players are supposed to change their state. If not, if their emotions are not intense, they will keep their current strategies and will not change them. In the experiments, we will run simulations to determine the main differences in the output of the groups playing the evolutionary game and to see the output implications when setting fuzzy emotions in the players' reasoning.

In summary, we can state the final outcome of this research line as presenting a more complete evolutionary game model based on CRD to better represent how to manage climate change policies and to understand how emotions play a role in these decisions. The structure of the paper is as follows. First, the description of the model in Section 2.1 and the inclusion of fuzzy emotions in Section 3. Later, the results and analysis of the experiments are presented in Section 4. Finally, the concluding remarks in Section 5.

2. Evolutionary model description

In this section, we will review the description of the CRD model, the payoffs of the game, and update rule.

2.1. Collective risk-dilemma

This threshold public good game dilemma, called the collective risk dilemma (CRD), consists of a finite set of agents Z (players) distributed in groups of size N. Each player i chooses a strategy s from three possibilities at every time step ($s(i) = \{C, D, P\}$) Vasconcelos et al. (2013): being a cooperator (strategy C), a defector (strategy D), or a punisher (strategy P). The agents' strategies are initially assigned at random.

Each player is involved in only one group and all players start with a benefit b. k_C, k_D, and k_P are the numbers of players adopting the strategy C, D, and P, in a group, respectively. Players with strategy C and P provide a fraction $c \in (0, 1)$ at each step of their benefit b for the public good. D players do not provide anything and therefore their benefit remains as b. If, in a group, there are not enough C or P players (that is, at least a threshold of n_{pg} players of the group size N), there is a probability of risk r for the group players to lose all their benefits (that is, collective disaster).

In order to limit the number of free riders or defectors (players with strategy D) in a group, P players pay a tax called π_t to create and maintain the local institution of the group. If the group is able to keep a minimum number of P players who contribute with taxes (a minimum ratio given by parameter n_p to keep at least $\pi_t n_p$), the defectors of the group are then penalized with a fine π_f.

2.2. Payoffs definition and group metrics

Taking into account the players in a group, we can determine the net wealth of individual agents based on their payoffs, and according to the strategy adopted by them and the other members of the group. The payoff Π_C of a cooperator, the payoff Π_D of a defector, and the payoff Π_P of a punisher can be obtained as follows:

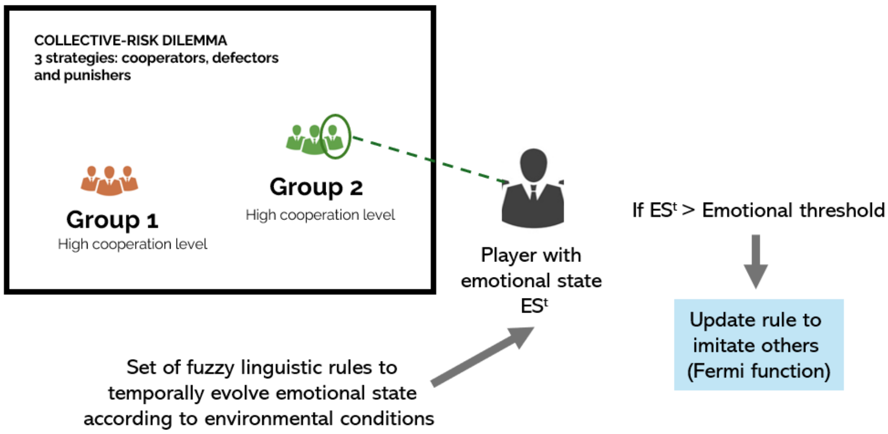

Figure 1. Scheme of the proposal having emotions into players of a CRD.

$$\begin{aligned}
&\Pi_C = -cb + b\Theta(k_C + k_P - n_{pg}) + \\
&(1-r)b[1 - \Theta(k_C + k_P - n_{pg})], \\
&\Pi_D = \Pi_C + cb - \Delta, \\
&\Pi_P = \Pi_C - \pi_t,
\end{aligned} \tag{1}$$

where $\Theta(x)$ is a Heaviside function that returns 1 if $x \geq 0$ and 0 if $x < 0$. Δ is the punishment function for defectors by the institution of the group, equal to 0 if the number of punishers in the group (k_P) is lower than $n_p N$: $\Delta = \pi_f \Theta(k_P - n_p N)$.

From the local decisions of the players in their groups, we obtain two important public good measures or metrics of the whole population to observe the cooperation level in a global term. The first is the prevalence of the institution η_I which is the ratio of groups that have a minimum number of players with strategy P who contribute with a tax π_t to preserve the public good. The second measure is the achievement of the group η_G, which measures the ratio of groups in the population with enough contributors C and P to the public good n_{pg}.

2.3. Update rule and evolutionary dynamics

After playing the CRD at each time-step and calculating their payoffs, the agents have an opportunity to update their strategies according to their fitness. The fitness of an agent i is calculated as the average pay-off of the players in the population who played the same strategy as i in the previous round $t-1$.

The update of the strategy follows an evolutionary procedure based on imitation that can be interpreted as information exchange within a social learning process (Nowak et al. (2010)). To be more precise, at the time step t, a player i evaluates its previous fitness results in $t-1$ and decides whether to imitate the strategy of another player j' using an evolutionary update rule. These rules of imitative nature are widely used in the literature and represent a situation where bounded rationality or lack of information forces players to copy (imitate) other strategies (Schlag (1998)).

We use the Fermi function, a stochastic pairwise comparison rule, where players can make mistakes during the imitation process (i.e., imitate players with a lower fitness value). Agent i with strategy X adopts strategy Y of agent j with a probability given by the following equation:

$$prob_i^t j = \frac{1}{1 + e^{-\beta(f_Y - f_X)}}, \tag{2}$$

where the intensity of the selection parameter β is equal to 0.5. f_X, f_Y are the fitness values for both strategies X and Y in step $t-1$.

Furthermore, players can always change their strategies by adopting a random strategy following a probability mutation mechanism μ.

Table 1. Fuzzy rules for emotional changes in Cs

	k_D/N is low	k_D/N is high
r is low	Positive	Negative
r is high	Neutral	Very negative

3. Emotions through fuzzy modeling

The traditional CRD described in the previous section is extended with this fuzzy model. The goal is

Table 2. Fuzzy rules for emotional changes in Ds

	k_P/N is low	k_P/N is high
r is low	Positive	Negative
r is high	Neutral	Very negative

to incorporate emotions into all players, independently of their strategy. These emotions will trigger the udpate rule mechanism. Therefore, the update rule (i.e., Fermi's update rule described in Section 2.3) is only triggered when a player feels emotionally uncomfortable. Please note that the mutation process is independent of the emotions' levels. Figure 1 shows an illustrative diagram with the main components of our proposal.

Mathematically, each player has an emotional state variable $ES_i^t \in \{0, ..., 10\}$ evolving over time and an emotional threshold $\overline{ES}_i$ that does not change over time. These two variables have integer values and control the emotional state of each player. When $ES_i^t > \overline{ES}_i$, the player uses the update rule to mimic others according to the rewards of their strategy in $t+1$. After applying the update rule, even if the strategy remains the same, the emotional state ES_i is reset to 0.

ES_i^0 is set to 0 while $\overline{ES}_i$ is set at random when initializing the simulation. The initialization of these thresholds can be uniformly distributed in the range or initialized using a Gaussian distribution. ES_i is increasing or decreasing overtime depending on the circumstances of the group of the player i. These emotional changes δ_E can be of four types: positive ($\delta_E = -1$), neutral ($\delta_E = 0$), negative ($\delta_E = 1$), and veryNegative ($\delta_E = 2$). The updates for each of the four emotions are listed below:

- Positive: $ES_i^t = \max\{ES_i^{t-1} - 1, 0\}$.
- Neutral: $ES_i^t = ES_i^{t-1}$.
- Negative: $ES_i^t = \min\{ES_i^{t-1} + 1, 10\}$.
- VeryNegative: $ES_i^t = \min\{ES_i^{t-1} + 2, 10\}$.

At each step t, the emotional state of the players ES_i is updated following a set of fuzzy linguistic rules in a TSK rule base with a $\min$ method for the AND operator, which will determine the increase or decrease in their emotions depending on the group and external conditions such as risk probability r or the number of other playing strategies.

We define 4 fuzzy input variables, one for risk r, k_D/N, and k_P/N in the range [0,1] with two low and high linguistic labels (linear fuzzy variables); and one for $n_p - k_P/N$, with two low (*Z shape*) and high (*S shape*). The fuzzy system also has a fuzzy output variable Δ_E with four labels (positive, neutral, negative, and veryNegative). Tables 1, 2, and 3 show these dynamics that link emotions with the variables of the groups. Finally, Table 4 presents the fuzzy linguistic rules, specific to each of the three strategies.

4. Experiments and results

In this section, we first establish the conditions and model specifications for the experiments (CRD model, fuzzy rules, and agent-based simulation). Later, we analyze the results and observe the implications of using fuzzy rules in a CRD.

4.1. Experimental setup

The results were obtained by simulating an agent-based model of the CRD evolutionary game. The implementation was done in Java and using the JFML library for fuzzy modeling. The model was run for 50 MC realizations in a population of $1,000$ agents for $1,000$ time steps of simulation. The initialization of the thresholds for fuzzy rules was performed using a Gaussian distribution with a standard deviation of 0.5 and two mean values 3 and 8. The mutation probability μ of the model was set to 0.01.

Regarding the parameters of the CRD, we establish an equal distribution of strategies in the initial population ($k_C \approx k_D \approx k_P$). The penalty tax is 0.03, the penalty fine is 0.3, and the thresholds for contributors and coordination are 0.25 and 0.75, respectively. The cost of cooperating is $c = 0.1$. We set two models setting; one with all the players playing a single group and another setting having groups of 4 players.

4.2. Analysis of the results

The goal of the experiments is to compare the dynamics without the use of emotions (i.e., the traditional CRD where update rules are triggered for all the time steps for all the players). This is equivalent to setting the minimum emotional threshold value of 0. The first results of the model without taking into account emotions and fuzzy linguistic rules (that is, the same model as in Vasconcelos et al. (2013) but using an agent-based implementation). The plots of Figures 2 and 3 show the dynamics of traditional CRD with groups of 4 players and one single group ($1,000$ players), respectively. Cooperation is almost inexistent when low-risk values are set, as already shown in the literature. When playing in a single group, defection if full independently from the risk level. These results are the baseline for comparing those using fuzzy emotions.

Figures 4 and 5 show the CRD results when fuzzy

Table 3. Fuzzy rules for emotional changes in Ps

	k_D/N is low	k_D/N is high
$n_p \leq \frac{k_P}{N}$ is enough ($n_p - \frac{k_P}{N}$ is low)	Positive	Negative
$n_p > \frac{k_P}{N}$ is not enough ($n_p - \frac{k_P}{N}$ is high)	Neutral	Very negative

Table 4. Fuzzy knowledge base of rules for the three strategies of the game

`Fuzzy knowledge base for agents playing strategy` C
IF r is low AND k_D/N is low, THEN Δ_E is positive
IF r is low AND k_D/N is high, THEN Δ_E is negative
IF r is high AND k_D/N is low, THEN Δ_E is neutral
IF r is high AND k_D/N is high, THEN Δ_E is veryNegative

`Fuzzy knowledge base for agents playing strategy` D
IF r is low AND k_P/N is low, THEN Δ_E is positive
IF r is low AND k_P/N is high, THEN Δ_E is negative
IF r is high AND k_P/N is low, THEN Δ_E is neutral
IF r is high AND k_P/N is high, THEN Δ_E is veryNegative

`Fuzzy knowledge base for agents playing strategy` P
IF $(n_p - k_P/N)$ is low AND k_D/N is low, THEN Δ_E is positive
IF $(n_p - k_P/N)$ is low AND k_D/N is high, THEN Δ_E is negative
IF $(n_p - k_P/N)$ is high AND k_D/N is low, THEN Δ_E is neutral
IF $(n_p - k_P/N)$ is high AND k_D/N is high, THEN Δ_E is veryNegative

rules are used in groups of 4 players. The first figure, Figure 4, is for the initialization of low threshold values (Gaussian distribution of 3 as a mean) while the second figure, Figure 5, is for high threshold values (mean of 8 for the Gaussian distribution). We can see that in both cases the results are quite similar. In addition, there are differences with respect to traditional CRD as changes when risk r is decreasing are less significant. Then, we can see that it is easier to get better results in terms of cooperation when we have emotions in the game.

In Figures 6 and 7 we have the same sensitivity analysis on risk r as before but when using a single group of 1,000 players instead of groups of 4 players. In this case, more differences are observed when the risk increases (cooperation is easier). Interestingly, punishers are the most prevalent strategy in the population.

We can see that by including emotions, we can get better results, as we have these thresholds to stop imitation until we reach the emotional thresholds. Another interesting insight of these results is to show how we control the evolutionary dynamics of the game. The point here is that, in classic game theory, the only way to change the dynamics of the game is to change the utility function. With emotions, interventions that affect emotions offer another set of tools to control dynamics. The point here is that emotions are driven by perception. So, changing how people perceive things could change their emotions, which would then change the dynamics.

5. Concluding remarks

In this study, we have used CRD as an evolutionary game simulated using agent-based models together with fuzzy linguistic rules. These rules are associated with the players' reasoning when playing in groups, and model the emotions of the players, linked to the level of risk disaster and the strategies of the other members of the group.

Different risk levels modeled by the parameter r were tested to compare the system without using emotions in the reasoning of the players and the use of emotions by the players before applying their update rules. The fuzzy interface to model emotions showed that can better control defection in the game for low values of the risk. In addition, it smooths the strategy levels when modifying the risk parameter, as it acts as a firewall before launching the update dynamics of the players.

Several future extensions of this model with linguistic rules can be followed. First, we can use historical decisions by the agents to change the emotions' response or the next update rule output. Additionally, linguistic modifiers can be used to

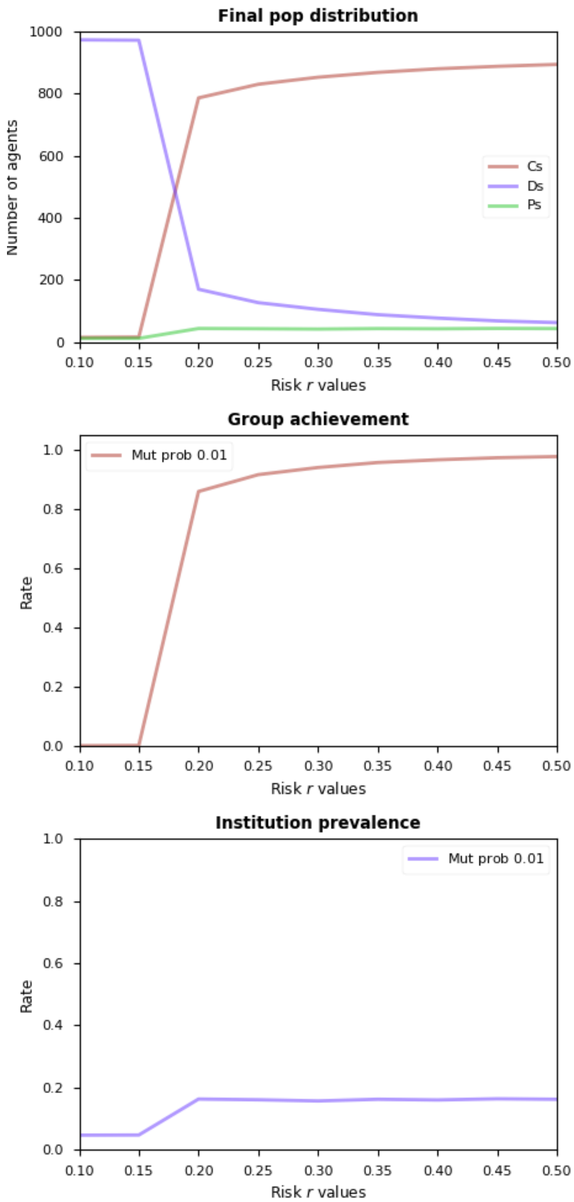

Figure 2. Cooperation metrics when changing risk parameter r in groups of four players.

increase/decrease the output of the rules together with a broader experimental setup via sensitivity analysis and wider ranges of the parameters. Finally, automatic machine learning methods such as genetic fuzzy systems can be used to learn the rules and threshold values of the model.

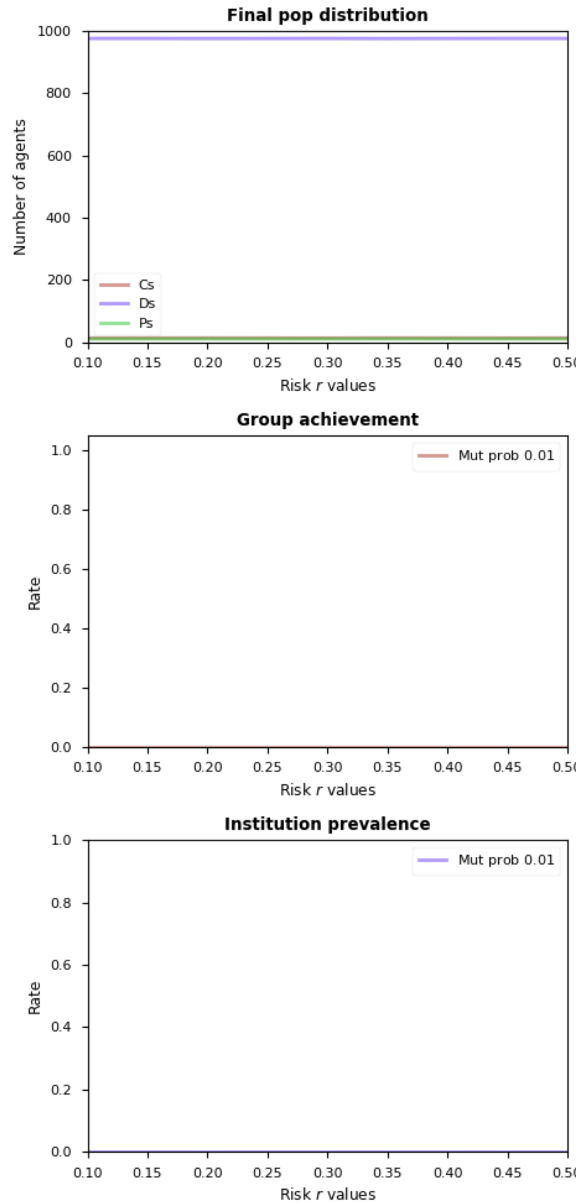

Figure 3. Cooperation metrics when changing risk parameter r with a single group (whole population).

Acknowledgments

This work was supported by the Andalusian Government under grant Emergia (EMERGIA21_00139) and by grant CNS2024-154265 funded by MCIN/AEI/10.13039/501100011033 and, as appropriate, by "ESF Investing in your future" or by "European Union NextGenerationEU/PRTR".

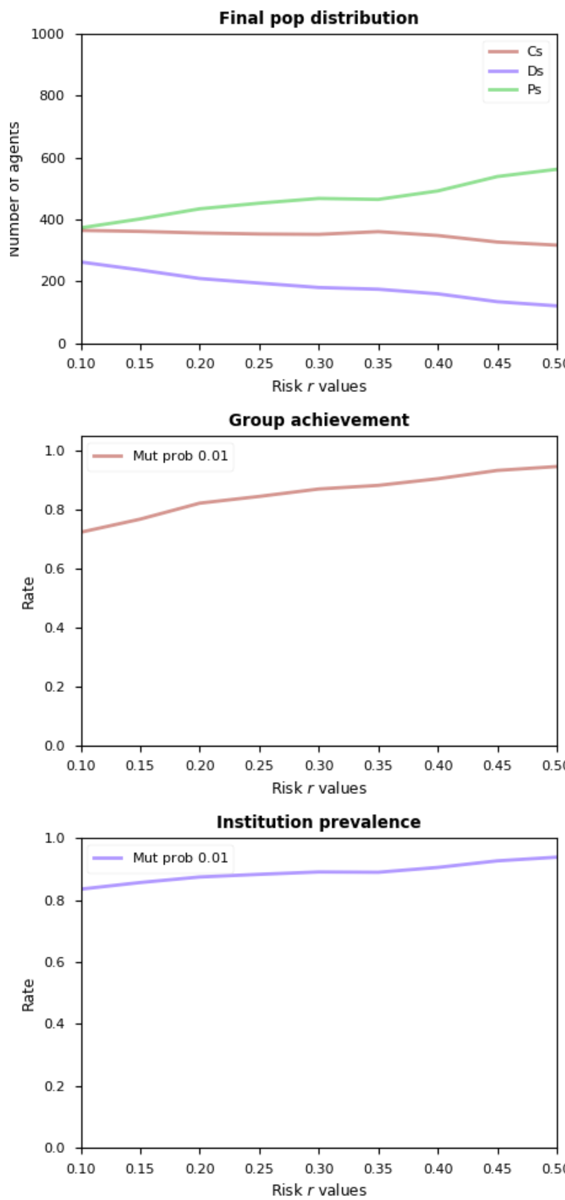

Figure 4. Cooperation metrics using CRD (groups of four players) with fuzzy linguistic rules using low thresholds for different risk parameter r.

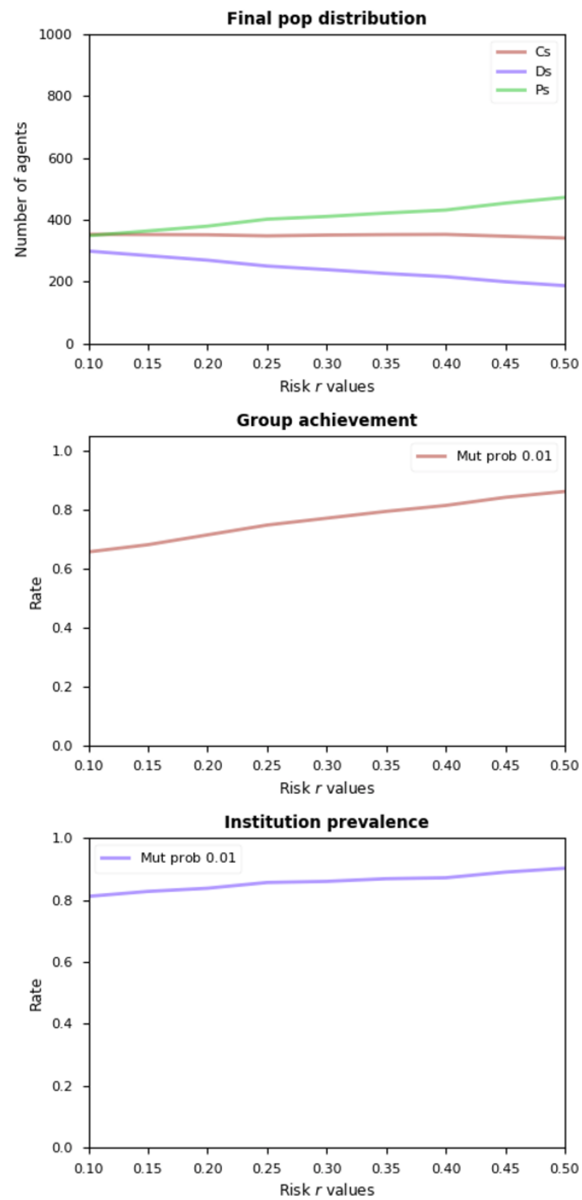

Figure 5. Cooperation metrics using CRD (groups of four players) with fuzzy linguistic rules using high thresholds for different risk parameter r.

References

Chen, W., Wang, J., & Chen, Y. (2025). The effect of emotional diffusion on the evolution of cooperation in multiplex networks. *Chaos, Solitons & Fractals*, *199*, 116860.

Chen, W., Wang, J., Yu, F., He, J., Xu, W., & Wang, R. (2021). Effects of emotion on the evolution of cooperation in a spatial prisoner's dilemma game. *Applied Mathematics and Computation*, *411*, 126497.

Chica, M., Hernández, J. M., & Bulchand-Gidumal, J. (2021). A collective risk dilemma for tourism restrictions under the covid-19 context. *Scientific Reports*, *11*(1), 1–12.

Chica, M., Hernández, J. M., & Santos, F. C. (2022). Cooperation dynamics under pandemic risks and heterogeneous economic interdependence. *Chaos, Solitons & Fractals*, *155*, 111655.

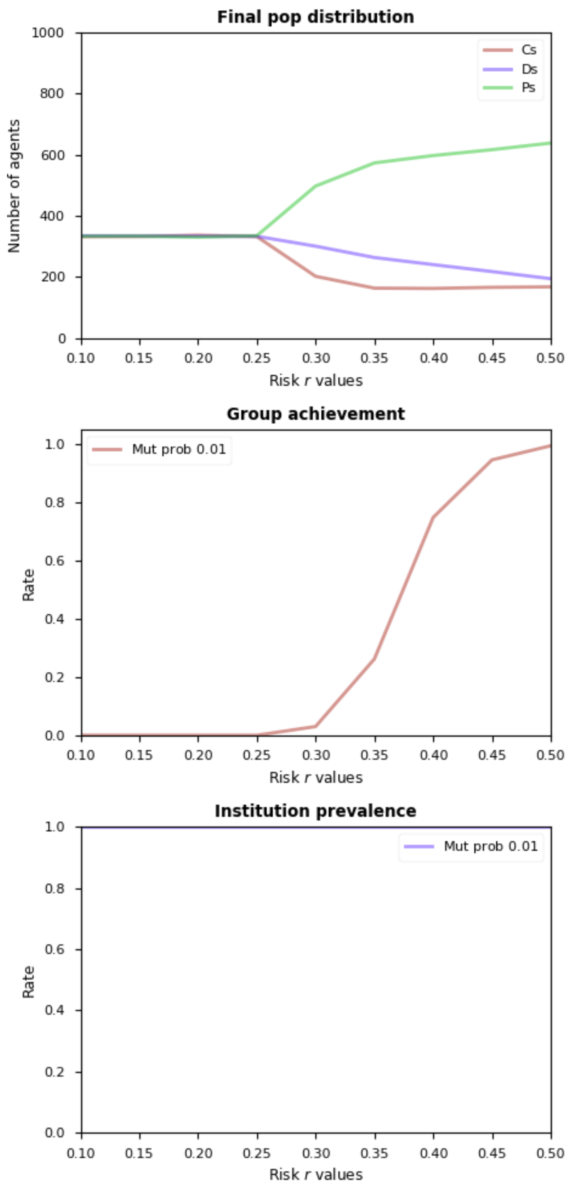

Figure 6. Cooperation metrics using CRD (in one single group) with fuzzy linguistic rules using low thresholds for different risk parameter r.

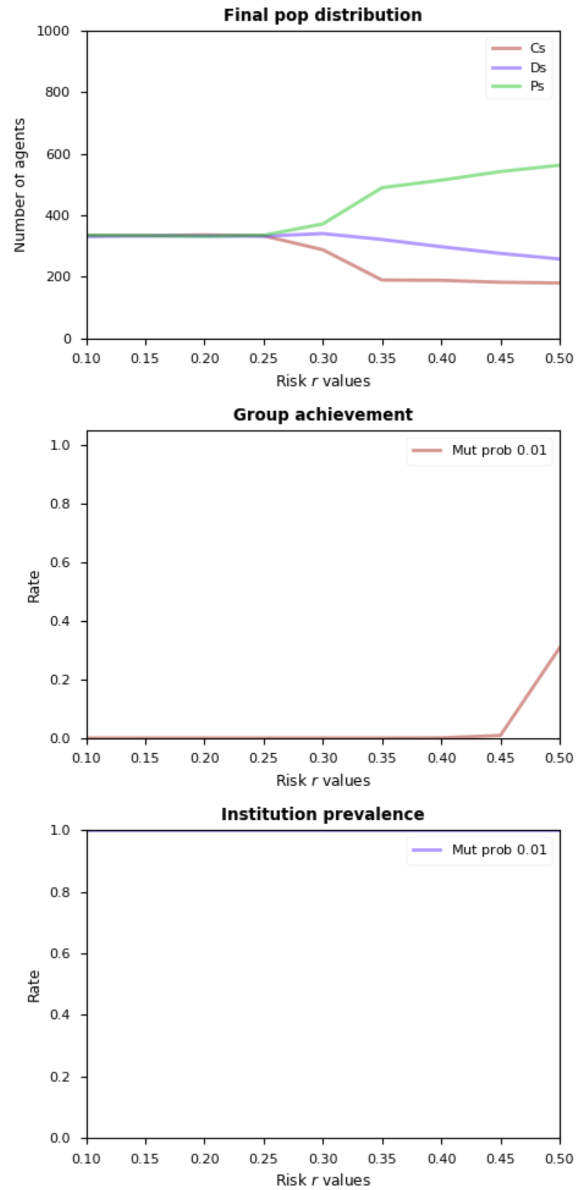

Figure 7. Cooperation metrics using CRD (in one single group) with fuzzy linguistic rules using high thresholds for different risk parameter r.

Cole, D. H. (2015). Advantages of a polycentric approach to climate change policy. *Nature Climate Change*, *5*(2), 114.

Greenwood, G., Abbass, H. A., & Petraki, E. (2017). Emotion, trustworthiness and altruistic punishment in a tragedy of the commons social dilemma. *Australasian Conference on Artificial Life and Computational Intelligence*, 12–24.

Levin, S. A. (2014). Public goods in relation to competition, cooperation, and spite. *Proceedings of the National Academy of Sciences*, *111*(Supplement 3), 10838–10845.

Milinski, M., Röhl, T., & Marotzke, J. (2011). Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. *Climatic change*, *109*(3), 807–814.

Milinski, M., Sommerfeld, R. D., Krambeck, H. J., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proc Natl Acad Sci USA*, *105*(7), 2291–2294.

Nowak, M. A., Tarnita, C. E., & Antal, T. (2010). Evolutionary dynamics in structured populations. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *365*(1537), 19–30.

Ostrom, E. (2010). Polycentric systems for coping with collective action and global environmental change. *Global Environmental Change*, *20*(4), 550–557.

Pacheco, J. M., Vasconcelos, V. V., & Santos, F. C. (2014). Climate change governance, cooperation and self-organization. *Physics of Life Reviews*, *11*(4), 573–586.

Reser, J. P., & Swim, J. K. (2011). Adapting to and coping with the threat and impacts of climate change. *American Psychologist*, *66*(4), 277.

Santos, F. C., & Pacheco, J. M. (2011). Risk of collective failure provides an escape from the tragedy of the commons. *Proceedings of the National Academy of Sciences*, *108*(26), 10421–10425.

Schlag, K. H. (1998). Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, *78*(1), 130–156.

Smirnov, O. (2019). Collective risk social dilemma and the consequences of the us withdrawal from international climate negotiations. *Journal of Theoretical Politics*, *31*(4), 660–676.

Szolnoki, A., Xie, N.-G., Wang, C., & Perc, M. (2011). Imitating emotions instead of strategies in spatial games elevates social welfare. *Europhysics Letters*, *96*(3), 38002.

Szolnoki, A., Xie, N.-G., Ye, Y., & Perc, M. (2013a). Evolution of emotions on networks leads to the evolution of cooperation in social dilemmas. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, *87*(4), 042805.

Szolnoki, A., Xie, N.-G., Ye, Y., & Perc, M. (2013b). Evolution of emotions on networks leads to the evolution of cooperation in social dilemmas. *Physical Review E*, *87*(4), 042805.

Tavoni, A., Dannenberg, A., Kallis, G., & Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proc Natl Acad Sci USA*, *108*(29), 11825–11829.

Vasconcelos, V. V., Santos, F. C., & Pacheco, J. M. (2013). A bottom-up institutional approach to cooperative governance of risky commons. *Nature Climate Change*, *3*(9), 797.

Vasconcelos, V. V., Santos, F. C., Pacheco, J. M., & Levin, S. A. (2014). Climate policies under wealth inequality. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(6), 2212–2216.

Wang, Z., Jusup, M., Guo, H., Shi, L., Geček, S., Anand, M., Perc, M., Bauch, C. T., Kurths, J., Boccaletti, S., et al. (2020). Communicating sentiment and outlook reverses inaction against collective risks. *Proceedings of the National Academy of Sciences*, *117*(30), 17650–17655.