# Social Motives and Social Contracts in Cooperative Survival Games*

Matthew Scott[1], Mathieu Dubied[1,2], and Jeremy Pitt[1]

[1] Imperial College London, London SW7 2BX, United Kingdom
{mss2518,md1721,jpitt}@ic.ac.uk
[2] ETH Zürich, 8092 Zürich, Switzerland

**Abstract.** Cooperative survival games are a sub-class of resource competition games wherein self-interest appears to be the rational choice in the short-term, but if every 'player' always acts out of self-interest, extinction is guaranteed in the long-term. The situation is dramatised in the film *The Platform* (*El Hoyo*); in this paper, we implement a self-organising multi-agent system that approximately recreates the cooperative survival game depicted in this film. In a series of experiments, we investigate how communication, a pre-existing tendency to sociality (characterised by social motives) and a capacity for social construction (characterised by social contracts) enables a collective of random individuals to establish a stable institution that increases their overall life expectancy. The experimental results provide some insight into how a pro-social personality and the ability to bootstrap institutions enable a random collective to find a solution to what is effectively an iterated game of Rawl's *Veil of Ignorance*.

**Keywords:** Multi-Agent System · Governance · Collective Action.

## 1 Introduction

Cooperative survival games are a sub-class of iterative resource competition games wherein self-interest appears to be the rational choice in the short-term, but if every 'player' always acts out of self-interest, elimination or extinction is inevitable in the long-term. The players need to maintain a critical mass that can gather sufficient resources to survive this iteration to ensure that here are sufficient players to survive the next iteration. Dropping below a certain threshold means that "if one is lost, all are lost".

Cooperative survival games are a popular form of entertainment in low- or zero-stakes entertainment, as seen in board games (e.g., Ravine) and computer games (e.g., "Don't Starve' "Rust," "Minecraft"), and have been analysed extensively in anthropological studies of collective behaviour in extreme environmental conditions [3][10]. Addressing anthropogenic climate change can be seen as a high-stakes cooperative survival game on a planetary scale with nation states as the players.

* Supported by Imperial College London and the Swiss Study Foundation

Ostrom has shown how collectives have solved the common-pool resource management (CPR) problem by using self-governing institutions [13], i.e. sets of mutable, mutually-agreed conventional rules which the members voluntarily regulate their behaviour. Considering a cooperative survival games as a form of extreme, high-stakes CPR problem where any one individual maximising self-interest or free-riding is an existential hazard to all, this paper addresses the question of how to bootstrap the formation of such an institution from a starting position of complete ignorance. In this initial situation, the players have no knowledge of the other players, and there are no rules, no social network, and no external authority. The players only have their personal psychological characteristics (which we call social *social motives*) and an ability for the social construction [2] of social contracts (which we call *treaties*).

Accordingly, this paper is structured as follows. In Section 2, we first present a scenario, which is based on the film *The Platform* (*El Hoyo*), and related work that provides the background to the multi-agent simulator developed in Section 3, and the social motives for agents specified in Section 4. Section 5 presents the experimental results which show how communication, a pre-existing tendency to sociality (characterised by social motives) and a capacity for social construction (characterised by social contracts or treaties) enables a collective of random individuals to establish a stable institution that increases their overall life expectancy. Finally, Section 6 concludes with some observations on how pro-social behaviour and the ability to bootstrap institutions enable a collective to find a psychologically and sociologically plausible solution to what is effectively an iterated game of Rawl's *Veil of Ignorance* [15].

## 2 Scenario and Related Work

For this paper, we consider the social dilemma presented the 2019 film 'The Platform' (*El Hoyo*).This film envisions a tower consisting of $N$ floors with a pair of prisoners on each floor.

A platform laden with food descends through a central shaft in the tower, stopping at consecutive floors. The prisoners are allowed to eat as much as they want while the platform has stopped on their floor, but cannot save food "for later". Obviously it is advantageous to be on a low-numbered floor to have first access to the food on the platform; however there is a 'reshuffle' after $D$ days, with all the agents are randomly re-assigned to new floors, and with no knowledge of which floor they will be re-assigned.

It has been shown that by taking an approach inspired by moral philosophy there are solutions to the *social contract design problem* [5]. This means that, for any non-cooperative game, it is theoretically feasible to define a social contract which produces a modified game that optimises for a moral imperative. Ostrom's work, as previously mentioned [13], provides empirical evidence that it is practically possible for groups of people to resolve collective action situations through the social construction of self-governing institutions. Effectively this is identifying the institutions, understood as a set of rules, as the social contract and sustainability of the common-pool resource as the moral imperative.

The question addressed in this paper is under what conditions is it practically possible for groups of *agents* to resolve a collective action situation, specifically that posed by *The Platform* scenario. In this scenario, we presume that the motivation for creating a social contract comes from an abstraction of the psychological concept of *social motives* [16][12], which Folmer describes as "the psychological processes that drive people's thinking, feeling and behavior in interactions with other people." Social motives are further identified as a potential source of conflict, with Folmer also claiming that "the actions that are dictated by one individual's motives are incompatible with, or even harmful to, the interests of others." In other words, the social contract must not only solve the social dilemma, but must also resolve any residual tension between potentially conflicting social motives.

Although, without loss of generality, we make some modifications to the scenario from the film – for example, we assume one prisoner per floor rather than a pair (although that is only required for dramatic effect), no movement between floors, and communication allowed between adjacent floors – we are assuming strict constraints of no prior knowledge, no pre-existing social network and no external authority, with the additional complications of a dynamic population and periodic floor re-assignment. The challenge is then to determine whether, despite the combination of limited communication, social motives and the propensity for social construction, the agents can 'find' a social contract which is a solution to the current formulation of the game and perpetuates across subsequent re-formulations.

## 3 Simulator Design

To simulate *The Platform*, we implement a self-organising, multi-agent system. This system consists of a set of agents, stored inside a 'tower' data structure which acts a server, handling agent interactions and containing the setup parameters for the simualtion.

External to the basic representation of agents in the tower, we further represent the infrastructure of the simulator by modelling the agents' health, global utility, and treaties.

### 3.1 Health Modelling

All agents have a health value that exists on a continuous spectrum with three additional discrete levels of *criticalLevel*, *weakLevel* and *maxHP*. An agent is considered to have critical health if it falls between the *criticalLevel*, the minimum possible health, and *weakLevel*, the cutoff for the critical region. An agent process is terminated if they remain in this region for $N$ days, equal to *maxDaysCritical*.

An agent's health is updated through two mechanisms: agents eating food (appropriating resources), which causes a positive change, and the cost of living, which causes a negative change.

Mathematically, the mapping between food intake and health is parameterised as follows:

$$newHP = currentHP + w(1 - e^{\frac{-foodTaken}{\tau}}) \tag{1}$$

with $\tau$ offered as a tuning parameter to either increase or decrease the magnitude of health change from one unit of food and $w$ a variable to represent the width of the gap between the *weakLevel* and *maxHP*. This function is chosen similarly to a step response function to replicate 'diminishing returns' and prevent rapid changes in health. An agent in the critical region has a slightly different update function:

$$newHP = currentHP + \min\left\{HPReqCToW, w(1 - e^{\frac{-foodTaken}{\tau}})\right\} \tag{2}$$

to ensure that a critical agent must first transition to weak, before applying Equation (1). Hence, *HPReqCToW* represents the change in health required to transition from the critical region to the weak level.

An agent's health will decay according to the equation:

$$newHP = currentHP - [b + s(currentHP - WeakLevel)] \tag{3}$$

and is subsequently bounded to the range [*criticalLevel*, *maxHP*]. We note that critical agents affected differently by health decay. If an agent is unable to achieve *HPReqCToW*, they will be reset to the *criticalLevel*. Conversely, if they do appropriate this food, they will be reset to the *weakLevel*.

### 3.2 Global Utility

To assess the performance of the agents in the tower as a group, we investigate their *social welfare*, based on each agent's individual utility [14].

In this scenario, each agent $i \in \{1, \ldots, N\}$ carries out four actions at each iteration $t \in \{1, \ldots, \infty\}$: it first determines the resources it has on the platform ($g_i$), then its need for resources ($q_i$). After this, it receives an allocation of resources ($r_i$) from the treaties it has formed and finally makes an appropriation of resources ($r_i'$). Since agents are programmed to be *honest*, we assert that $r_i' = r_i$.

The need for resources $q_i$, looks to reward agents who take food only when necessary. Hence:

$$q_i = \frac{numberDaysInCriticalState}{maxDaysInCriticalState} \tag{4}$$

The total resources accrued at the end of an iteration, $R_i$, is then defined as:

$$R_i = r_i' + (g_i - p_i) \tag{5}$$

which gives the utility per agent:

$$u_i = \begin{cases} \alpha_i q_i + \beta_i(R_i - q_i) & \text{if } R_i \geq q_i \\ \alpha_i R_i - \gamma_i(q_i - R_i) & \text{else} \end{cases} \tag{6}$$

where $\alpha_i$, $\beta_i$ and $\gamma_i$ are tuning parameters that follow the rule $\alpha_i > \gamma_i > \beta_i$. In our work, we use the values $\alpha_i = \alpha = 0.2$, $\beta_i = \beta = 0.1$, and $\gamma_i = \gamma = 0.18$.

Finally, we use (6) to compute an average global utility, which corresponds to the social welfare $SW$ divided by the number of agents:

$$U = \frac{\sum_i^N u_i}{N} = \frac{SW}{N} \tag{7}$$

### 3.3 Treaties

To successfully handle treaties, an agent must be able to propose, evaluate, and propagate treaties. In addition, we enforce the agents act *honestly*, and therefore comply with the treaties to which they agree.

Treaties are codified as data structures with three main parts: a condition on the HP, a request concerning the amount of food and a duration. One such example of a treaty is: "if *currentHP* $\geq 60$, take $\leq 5$ food for 5 days.". They serve as an extension of message passing, wherein a treaty is proposed verbally either 1 floor above or below the floor of the proposer. Such proposals happen *asynchronously* in the tower and are implemented with concurrent channels, meaning that all agents can send treaties simultaneously. When a treaty is proposed, it enters the receiver's 'inbox' to be processed.

An agent may compile a treaty with *newTreaty*$(p_1, p_2, \ldots p_n)$, which packages $n$ arguments into a data structure to be subsequently sent as a proposal to an agent. Upon agreeing with a treaty, both agents involved will place this data into their respective *activeTreaties* arrays. The array is then iterated through to provide constraints on the agents' consumption.

## 4 Agent Design

All agents $i \in A$ are implemented as a data structure with parameterisation to participate in the various communication methods $c \in C$, resulting in a set of interactions defined by $\mathrm{I} = < A, C >$. Each agent inherits from the *baseAgent* structure and also contains the fields contained in Table 1. We note only the most relevant fields for quantifying the agent have been included.

| **Parameter** | **Range** |
|---|---|
| BaseBehaviour | $N \in [0,10]$ |
| Stubbornness | $N \in [0,1]$ |
| MaxBehaviourSwing | $N \in [0,10]$ |
| ParamWeights | { $HPW$:$N$, $FW$:$N$ } |
| FloorDiscount | $N \in [0,1]$ |
| MaxBehaviour | $N = 10$ |

| **Parameter** | **Range** |
|---|---|
| Config | *config*{} |
| CurrBehaviour | $N$ |
| MaxFloorGuess | $N$ |
| AverageFoodIntake | $N$ |
| ShortTermMemory | $[N]$ |
| LongTermMemory | $[N]$ |
| ActiveTreaties | [*Treaty*] |

Table 1: The *Config* (left) and *Agent* (right) data structures.

### 4.1 Social Motives

**Social Motives Spectrum** The agent's behaviour revolves around the concept of social motives [12], which Folmer defines as "the psychological processes that drive people's thinking, feeling and behavior in interactions with other people" [16]. This in turn leads to a "mixed-motive" setting [17] in the tower. From this concept, we abstract 4 distinct social motives:

*Altruist*: The disinterested and selfless concern for the well-being of others. An altruist then acts in a way that purely benefits others, even if it means harming themselves.

*Collectivist*: The practice or principle of giving a group priority over each individual in it. A collectivist then acts in a way that benefits the group, themselves included, over purely the individual.

*Selfish*: Being concerned excessively or exclusively with oneself. A selfish agent will act in a way to satisfy themselves, but not necessarily with the intent to harm the other agents.

*Narcissist*: An excessive interest or admiration of oneself. A narcissistic agent will act in a way that not only benefits themselves, but also hinders the collective.

For this implementation, we assert that all agents' social motives can be defined on a spectrum, with one end corresponding to pure altruism, and the other to pure narcissism, which we codify as a continuous value between 0.0 and 10.0 respectively. Fig. 1 illustrates the spectrum of social motives.

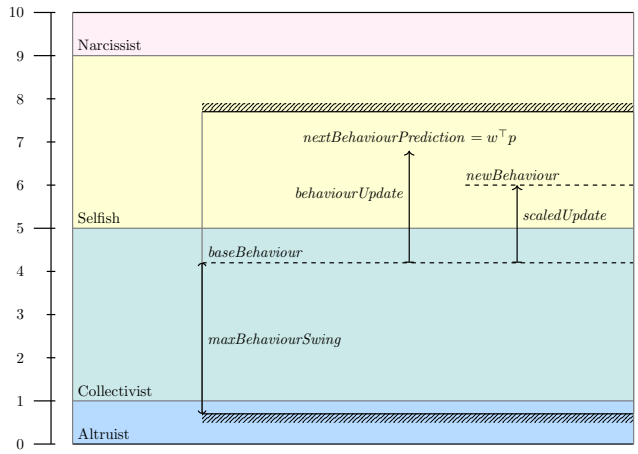



Fig. 1: Illustration of a change in social motive.

**Changing Social Motives** This paper proposes that it is both limiting and unrealistic for an agent to express 1 social motive for its entire lifespan. For this reason, agents are able to dynamically update their initially assigned social motive to reflect the duality of "nature vs nurture" [9]: an agent's genotype does not necessarily match the agent's phenotype.

To codify this idea, we use a 'predictor' that calculates a *behaviourUpdate* from the feature transformations of the 1) current health of the agent (8) and 2) floor that the agent is located on (9). These feature transformations map their respective features to a range [0,1], with poorer performances (low health, low floor) tending towards 1 to represent a skew towards narcissistic behaviour:

$$hpScore = 1 - \frac{currentHP}{maxHP} \tag{8}$$

Agents forecast the *maxFloor* by keeping track of the lowest floor they have visited. The lower down the floor, the faster Equation (9) tends to 1. This is to have the agents tend towards narcissism faster as they reach lower floors. We take $\lambda$ as the *floorDiscount* variable from Table 1.

$$floorScore = \frac{e^{\frac{\lambda \cdot currentFloor}{maxFloor}}}{e^{\lambda}} \tag{9}$$

The predictor then weights these feature transformations to yield a value in the range [0,10]:

$$\begin{gathered} p = [currentHPUtility, currentFloorUtility], \quad w = [HPW, FW] \\ nextBehaviourPrediction = w^{\top} p \end{gathered} \tag{10}$$

and we construct a vector illustrating the change in social motive as:

$$behaviourUpdate = nextBehaviourPrediction - currentBehaviour \tag{11}$$

This paper further asserts that agents are unlikely to rapidly change their social motive, instead requiring multiple similar experiences to alter their phenotype. We hence offer a concept of *stubbornness*, which limits the vectorial change in *behaviourUpdate*:

$$scaledUpdate = behaviourUpdate \cdot (1 - stubbornness) \tag{12}$$

$$newBehaviour = currentBehaviour + scaledUpdate \tag{13}$$

With the new social motive defined as the movement from the current behaviour using the *scaledUpdate* vector. Finally, we propose that a genotypically altruistic agent, say, is unlikely to make a severe transition in personality to full narcissism. This is solved by introducing a *maxBehaviourSwing*, which bounds the total change in social motive that an agent can experience.

Agents are also able to dynamically update the weights in Equation (10) in order to make more permanent shifts towards narcissism if one of the parameters is constantly evaluated poorly. If the agent's health is below 20, increase $HPW$ by 0.05 and decrease $FW$ by 0.05. Alternatievly, If the agent's average food intake is less than 1 per turn, decrease $HPW$ by 0.1 and increase $FW$ by 0.1. After this update, we ensure that the weights remain in the range [0,1].

### 4.2 Food Consumption

Resources are conditionally appropriated depending on both the social motive and environmental factors such as commitments to messages and treaties. The baseline behaviours exhibited by the different social motives are as follows:

*Altruist*: An altruistic agent always takes 0 food, as it is only concerned for the well-being of others with a total disregard for itself.

*Collectivist*: A collectivist agent consume the food required to survive, and consumes no food when not in danger of dying. To codify this, agents randomly choose a day in the range [1, *maxDaysCritical*] and take food once they have remained at critical health for this period. This has the effect of staggering when collectivists are able to take food, to prevent the entire tower simultaneously depleting resources.

*Selfish*: A selfish agent always aims to stay at the *healthyLevel*. This means that it will always appropriate the food required to reach this point.

*Narcissist*: A narcissistic agent takes maximum amount of food consumable, since it is purely be concerned for its own well-being whilst sabotaging the others.

### 4.3 Handling Treaties

**Evaluating Treaties** It is through the agents interacting with one another that a social network is formed. Agents use techniques from forecasting, utility theory and the link weights in the social network to handle the acceptance or rejection of treaties. The results of these decisions then update the link weights.

Given that treaties do not have any immediate effect, but instead influence the future consumption of an agent, agents forecast to assess the present value of a treaty. This is codified by using two separate arrays corresponding to long-term and short-term memory and storing the amount of food received each day (Fig. 2), with the short-term memory reset after each reshuffle. The reason for having two memory types is to allow agents to separately look at the current reshuffle period and total experience in the tower, which aligns with the core assumption in cognitive psychology that there are separate systems for long- and short-term memory [11].

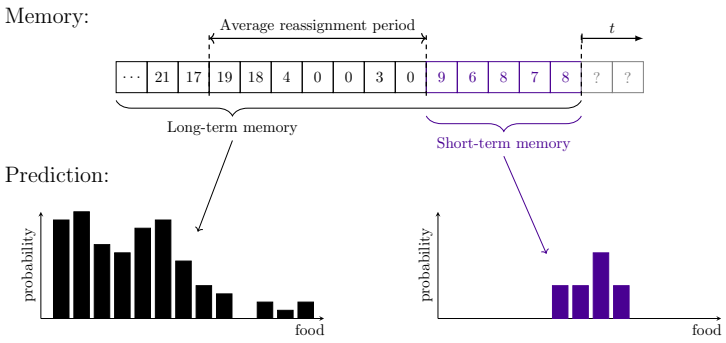



Fig. 2: Illustration of different agent memory types

Since the reshuffle period is unknown to agents, they forecast this information by averaging over all previous reshuffling periods.

Agents must also contrast the effect that a treaty will have on the *future* food intake to assess if it is beneficial or not. Since the satisfaction of gaining or losing wealth is non-linear [7], utility functions can account for this by mapping the monetary value of a good or service to an individual's preference [6].

Therefore, an agent calculates the expected utility both with and without a treaty and subsequently maximising its estimated future benefit. The utility of gaining an uncertain amount of food represented by the discrete random variable $x$ (based on past experience), is computed with:

$$x = [p_1 : x_1, p_2 : x_2; ... \; p_n : x_n] \tag{14}$$

where an agent expects to get amount $x_i$ of food with probability $p_i$, such that:

$$E[U(x)] = p_1 \times U(x_1) + p_2 \times U(x_2) + ... + p_n \times U(x_n) \tag{15}$$

Prospect theory [8] is a well-established model of how a change in value is perceived or, alternatively, how much utility is gained or lost from a change in value. This model comprises four main principles:

*Greediness*: Agents are generally greedy, meaning that more of a resource is at least beneficial. Utility functions are hence generally increasing.

*Diminishing sensitivity*: Marginal returns are strictly decreasing, thus the greater the personal wealth of an agent, the less they value the resource.

*Risk aversion*: Agents generally try to avoid risk. With risk aversion, the amount of food the agent perceives as equivalent to a random distribution (its *certainty equivalent* $C$) is hence less than its mean.

$$U(C) = E[U(x)] < U(E[x]) \tag{16}$$

*Loss aversion*: Losing some amount of food is generally perceived as worse than gaining that same amount. Agents hence weight loss higher than gain

Using these concepts, we identify a *gain (g)* and *cost (c)* associated with each unit of food received *(x)*, as well as the *risk aversion (r)* to define the utility of receiving a unit of food. The amount of food that the collectivist and selfish agents would need to consume in order to maximise their utility varies depending on the current health level. The peak of its total utility function thus needs to be able to vary too. We account for this by introducing a scaling factor $a$ as:

$$a = \frac{1}{z}\left(\frac{cr}{g}\right)^{\frac{r}{1-r}} \tag{17}$$

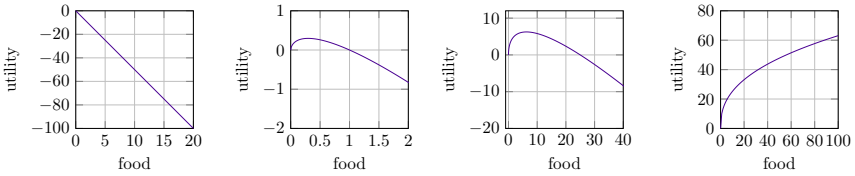
yielding:

$$U(x) = g(ax)^{\frac{1}{r}} - cax \tag{18}$$

with $z$ being the desired food intake, falling at the maximum of this function.

The utility calculation for each different social motive has been parameterised according to three insights: 1.) the more selfish an agent is, the greedier it is, 2.)

the more an agent cares for the greater good, the greater its social cost associated with consumption and 3.) more narcissistic people are generally less risk-averse [4]. The resulting utility functions are shown in Fig. 3.




(a) Altruist $g = 0, r = 1, c = 5$ (b) Collectivist $g = 2, r = 1.5, c = 2$ (c) Selfish $g = 5, r = 2, c = 1$ (d) Narcissist $g = 10, r = 2.5, c = 0$

Fig. 3: Different utility functions used to rate treaties according to the social motive of the agents.

Agents also use the proportion of estimated days before the next reshuffle period in order to weight how much they should focus on the short term. To optimise survivability, agents ignore the expected long-term utility when their health is on a critical level.

Let $b_{short}$ and $b_{long}$ be the estimated short and long term benefit of a treaty, respectively. Also, let the estimated days remaining on the current level be given by $d_{current}$ and the duration of a treaty by $d_{treaty}$. The total benefit, $b_{tot}$ is then:

$$b_{tot} = \frac{d_{current}}{d_{treaty}} \times b_{short} + (1 - \frac{d_{current}}{d_{treaty}}) \times b_{long} \tag{19}$$

Overall, the algorithm that agents follow when considering treaties is summarised as follows:

1. Check if the trust in the *proposing* agent is high enough
2. Check if the treaty is compatible with treaties that the agent signed already
3. Calculate the expected short- and long-term utility according to Equation (15)
4. Amplify the utility if it is negative to simulate loss-aversion.
5. Calculate the utility of the food it can feasibly take under the treaty
6. Compute the estimated benefits of signing the treaty as $U(\text{sign}) - U(\text{don't sign})$.
7. Choose to focus on the long- or short-term benefit according to Equation (19)
8. Sign the treaty if its overall benefit is positive

**Proposing and Propagating Treaties** Altruist agents wish to sacrifice themselves by taking 0 food and narcissist agents wish to sacrifice others by taking all the food. This means that these agent types will never sign treaties, as it goes against their strategy. The collectivist and selfish agents are therefore the two social motives that propose treaties.

Once a treaty has been accepted or rejected, it is possible for the agent to re-propose the same treaty to its neighbour. Logically, the best possible strategy is to propagate one single treaty throughout the tower and have all agents behave uniformly. Narcissist agents act to avoid this, hoping for the downfall of the collective and hence refuse to propagate treaties. All other agents, however, propagate the treaty five floors above and below.

## 5 Experimental Results and Discussion

In this section, we use the simulator and agent designs introduced in Section 3 and Section 4 to assess the performance of the studied system.

We divide the simulations into 4 groups characterised by having different initialisation parameters. Table 2 summarises the simulations parameters. The percentages of each social motive correspond to the initial distribution. If not explicitly mentioned, we run experiments using 100 agents, with 100 food initially on the platform for 60 days and with a reshuffle period of 30 days. Our simulations results are given as the average over 30 repeated simulations.

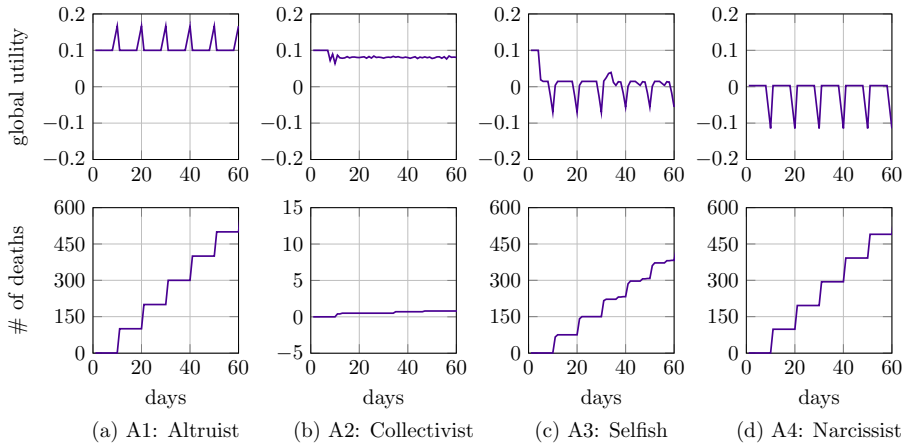
| | A1 | A2 | A3 | A4 | B1 | B2 | C1 | C2 | C3 | D1 | D2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| % Altruist | 100 | 0 | 0 | 0 | 10 | 0 | 10 | 10 | 10 | 100 | 100 |
| % Collectivist | 0 | 100 | 0 | 0 | 40 | 80 | 40 | 40 | 40 | 0 | 0 |
| % Selfish | 0 | 0 | 100 | 0 | 40 | 20 | 40 | 40 | 40 | 0 | 0 |
| % Narcissist | 0 | 0 | 0 | 100 | 10 | 0 | 10 | 10 | 10 | 0 | 0 |
| Stubbornness | - | - | - | - | - | - | 0.2 | 0.2 | 0.2 | 0.8 | 0.8 |
| MaxBehaviourSwing | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 8 | 8 | 6 | 6 |
| Communication | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | Y |

Table 2: Summary of the experiments.

In addition, cases C3 and D2 have additional specific features. In comparison to C1, C3 introduces more restrictive treaties. This is also the case for D2.

### 5.1 Simulation A

The first set of simulations we analyse are simulations that include agents that all have the same social motive. Moreover, these agents do not have the ability to change their social motive. The simulations results are shown in Fig. 4.




(a) A1: Altruist (b) A2: Collectivist (c) A3: Selfish (d) A4: Narcissist

Fig. 4: Simulation results for a group of agents with uniform fixed social motive.

We observe that a system containing purely altruists (Fig. 4 (a)) effectively self-destructs, since by acting purely selflessly, these agents never take any resources. As the agents all die at the same time and are replaced by a new group of altruists, we see a step pattern in the number of deaths over time.

Similar to the altruist agents, the narcissists have a large number of deaths among them every 10 days (Fig. 4 (d)). This is due to the agents on the upper floors of the tower taking all of the food, leaving none for the agents below.

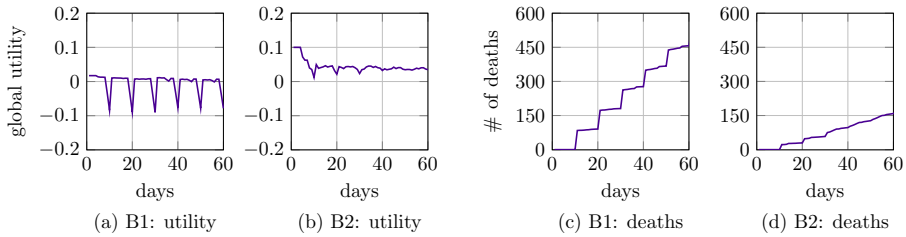
The main difference between the collectivist and the narcissist agents can be seen in their corresponding global utility. The patterns can be explained by (6), which yields positive values only for A1, but leads to negative spikes for A4.

As a compromise between the two systems, a system including only selfish agents present a lower number of deaths and a better global utility than A4 (Fig. 4 (c))

Finally, the collectivists instantaneously achieve a stable society in which (almost) none of the agents die (Fig. 4 (b)). We also note a uniformly positive curve for global utility over time that is smoother than for the other social motives. This reflects the increased social cohesion between the agents and identifies the almost perfect allocation of resources, leading to no wasted utility.

### 5.2 Simulation B

Having assessed groups of agents of each social motive individually in Section 5.1, we increase the complexity of the system by having agents with different *fixed* social motives in the tower. The inability for these agents to change their social motive with time leads to the simulation results shown in Fig. 5.




(a) B1: utility (b) B2: utility (c) B1: deaths (d) B2: deaths

Fig. 5: Simulation results for a group of agents with different fixed social motives.

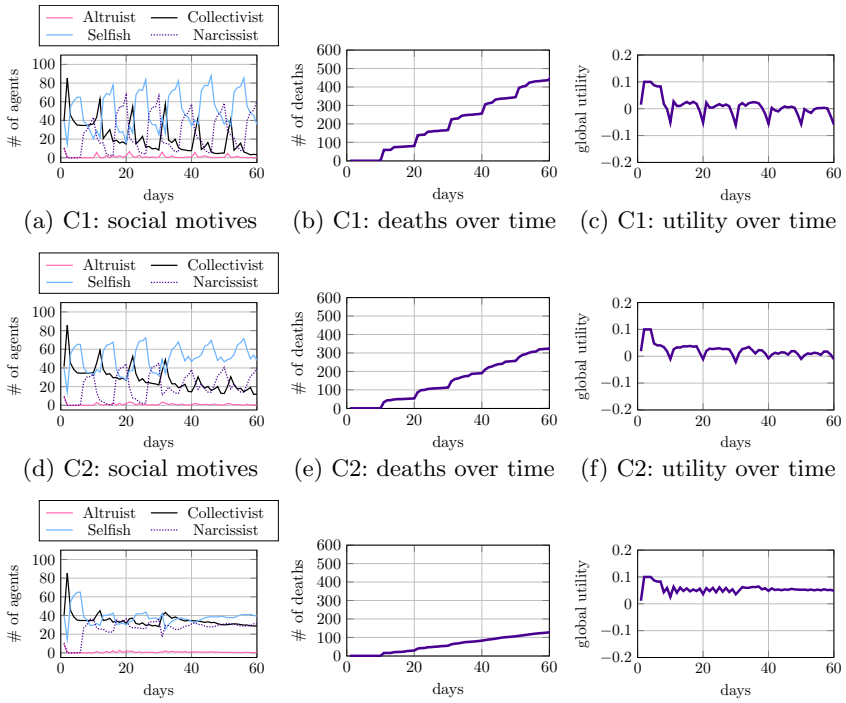
Through comparing B1 to B2, we see that the system comprising a larger amount of collectivists (B2) outperforms the system with a comparatively smaller amount of collectivists (B1). This is to be expected, as the more collectivist agents there are, the more similar to Fig. 4 (b) the system will be.

A second result illustrated by this simulation is that the action of introducing treaties (Section 4.3) is not always relevant. The collectivist agents accept the collectivist and selfish treaty (the collectivist one being more restrictive), but the selfish agents *only* accept the selfish treaty. This way, the two agent types are following their natural strategy concerning food intake (Section 4.2). Knowing this, the system shows similar results with and without treaties, hence we only show the results where communication is allowed.

### 5.3 Simulation C

This set of simulations builds on top of the framework set by simulation B, instead investigating the behaviour of a system comprising different distributions of *fluid* social motives. We utilise different levels of communication and treaties to contrast the results using the treaties introduced in Section 4.3 (Fig. 6 (a-c)). We simulate the system under two other configurations: without considering any form of communication (Fig. 6 (d-f)), and by using a different, more restrictive treaty (Fig. 6 (g-i)).

This treaty restricts the amount of food its members can take when their health drops below the *weakLevel*: "if *currentHP* $<$ *weakLevel*, take $\leq 2$ food."




(a) C1: social motives (b) C1: deaths over time (c) C1: utility over time

(d) C2: social motives (e) C2: deaths over time (f) C2: utility over time

(g) C3: social motives (h) C3: deaths over time (i) C3: utility over time

Fig. 6: Simulation results for case C. C1 and C3 include communication, but C2 does not. C3 includes a more restrictive treaty.

The overarching comment to draw from this set of results is the impact of specific treaties on the global utility. Although thought to improve the global utility, treaties might have a negative effect on it: the results C1 using the collectivist treaty as introduced Section 4.3 are worst than the ones obtained without communication (C2).

As the agents' health falls, their social motives tend to change toward narcissist. Instead of following the natural decision of this social motive, the agents have to follow the treaties they signed. The moment their health falls below

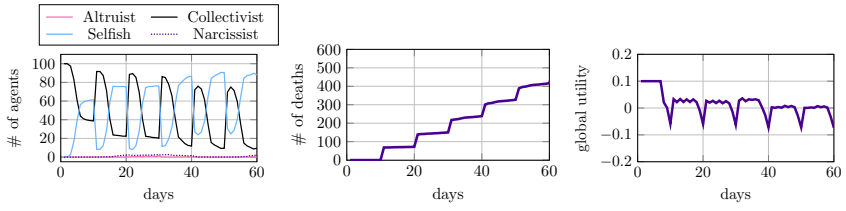
the *weakLevel*, these treaties no longer applies and they will follow their natural food intake rule defined in Section 4.2. However, this leads to a lot of wasted resources at this critical health level. Notably, each food intake greater than 2 will not offer additional utility to agents whose health falls below the *weakLevel*: any food intake greater than or equal to 2 upgrades the agents health to the *weakLevel*. The waste of common pool resources can also be visualised in Fig. 6 (c), where the global utility becomes strongly negative every 10 days

This waste of common resources induced by agents following the collectivist treaty is arguably due to a poor treaty design. To contrast these results, we can propose a different, more effective treaty. Simulation C3 introduced a treaty that applies when the agents HP is below the *weakLevel*. As can be seen in Fig. 6 (h) and (i), this treaty allows for better performance of the system.

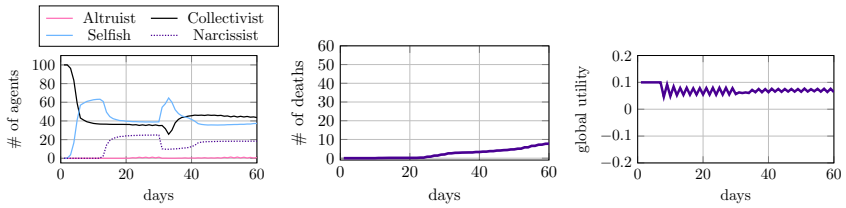
### 5.4 Simulation D

In these experiments, we initialise the tower's population with collectivist agents only, but with the possibility for them to change their social motive over time.

The goal of these experiments is to evaluate if a society comprised solely of collectivists is able to remain stable over time. In addition, we investigate the effect of treaties on such a system. The simulation results are shown in Fig. 7.




(a) D1: social motives (b) D1: deaths over time (c) D1: utility over time




(d) D2: social motives (e) D2: deaths over time (f) D2: utility over time

Fig. 7: Simulation results different treaties acceptances.

The results of D1 are similar to the ones of C1 in terms of (high) number of deaths and (low) global utility. The replacement of terminated agents by collectivist agents leads to an oscillatory behaviour between 2 quasi-stable states, with convergence to both a high concentration of collectivists and selfish agents in inverse proportions. We hence deem this experiment as 2-phase polystable [1].

Using the more restrictive treaty introduced in Section 5.3 on a system initially composed solely of collectivist agents leads to an impressive performance

(Fig. 7 (e) and (f)). In addition, the use of this treaty also allows for a stable distribution of the social motives across the tower (Fig. 7 (d)). This stability can also be seen in Fig. 6 (g). Despite the presence of selfish (and even narcissistic) agents in the tower, they all follow the rule dictated by the treaties they signed whilst being collectivist.

In addition, we can also see the effect of the reshuffle period on the social motives distribution in Fig. 7 (d). The reshuffle period is 30 days in this case and we see a global shift toward collectivism at that moment.

## 6 Summary and Conclusions

### 6.1 Summary

Our first set of experiments shows the natural strategies taken by agents of different (fixed) social motives, and therefore gives us a baseline (A). The collectivist strategy is by far the one achieving the highest global utility. Consequently, the more collectivist agents in the tower, the higher the global utility (B).

However, the natural tendency of agents in an economy of scarcity is to make a transition towards the narcissistic end of the spectrum. This leads to a higher overall distribution of selfish agents, and therefore a higher number of deaths and lower global utility (C, D).

To counteract this fact, it is possible to design social contracts in the form of treaties between the agents. Treaties serve as a stabilising self-organising mechanism, with appropriately constrictive treaties (C3, D2) even allowing for the integration of narcissists into the population, despite their natural tendency to destabilise a system. Treaties may also change a polystable system into a purely stable system, when sufficiently strong as to enforce a collectivist mindset. Oscillatory distributions of social motives can be brought to a static distribution using this mechanism (D1, D2). However, designing treaties that lead to a high global utility is not a trivial task. Agents using poorly designed treaties might even perform worst than agents only following their natural strategy without using any sort of communication (C1, C2).

### 6.2 Future Work

Our future work would focus around adapting the ways in which we model the agents' changes in social motives. One such way is to make agents tend towards altruism, rather than narcissism, when faced with adversarial conditions. This could be interpreted as an understanding of the agent's environment and the long-term improvement of the individual utility through a short-term sacrifice, thus bringing the system back to an equilibrium.

Furthermore, we might imagine a randomly distributed assignment of behavioural weights (Table 1) across different agents. This would illustrate how different agents react to their condition, from which the concept of agent personality could be derived. For example, some agents may encounter a comfortable situation (high HP, high floor) and take advantage of it by acting selfishly, while

another agent may encounter the same situation and take the opportunity to make a positive impact for their fellow agents below by acting altruistically.

Finally, we want to analyse the effects of a larger number of treaties on global utility. The choice of treaties which lead to an increase in the global utility is not straightforward. Since treaties are expressed in a generic way, it may be possible to tune the treaty parameters to find optimal treaties in a given scenario.

### 6.3 Conclusion

In conclusion, we observe that the scenario demands that the prisoners in the tower are effectively faced with an iterated version of Rawls' Veil of Ignorance: they have to decide repeatedly what sort of society they would prefer if they did not know what position they would occupy in such a society. This work shows that even with limited communication and a population with diverse social motives, the ability to construct social contracts leads to a stable society which perpetuates across generations, arguably showing that there is some psychological and sociological plausibility to Rawls' theory, although there is still work to be done on establishing whether or not, even if our agents establish a stable and self-perpetuating social contract, it is the 'best' social contract.

## References

1. Ashby, W.: Design for a Brain; The Origin of Adaptive Behavior - Scholar's Choice Edition, chap. 4-5. Creative Media Partners, LLC (2015)
2. Berger, P., Luckmann, T.: The Social Construction of Reality. Penguin Books (1966)
3. Briggs, J.: Never in Anger: Portrait of an Eskimo Family. Harvard Uni. Press (1970)
4. Campbell, W.K., Goodie, A.S., Foster, J.D.: Narcissism, confidence, and risk attitude. Journal of behavioral decision making **17**(4), 297–311 (2004)
5. Davoust, A., Rovatsos, M.: Social contracts for non-cooperative games. In: Third AAAI/ACM Conf. on Artificial Intelligence, Ethics, and Society. pp. 43–49 (2020)
6. Fishburn, P.C.: Utility theory for decision making. Tech. rep., Research analysis corp McLean VA (1970)
7. Fishburn, P.C.: Nonlinear preference and utility theory, p. 2. Johns Hopkins University Press (1988)
8. Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk, pp. 99–127. World Scientific (2013)
9. Kong, A., Thorleifsson, G.: The nature of nurture: Effects of parental genotypes. Science **359**(6374), 424–428 (2018)
10. Norberg-Hodge, H.: Ancient futures: learning from Ladakh. Sierra Club Books, San Francisco (1991)
11. Norris, D.: Short-term memory and long-term memory are still different. Psychological bulletin **143**(9), 992 (2017)
12. Oppenheimer, O.: The origin of social motives. Educational Theory **4**(2), 95–104 (1954)
13. Ostrom, E.: Governing the Commons. Cambridge (1990)
14. Pitt, J.: Self-Organising Multi-Agent Systems, pp. 149–151. World Scientific (2021)
15. Rawls, J., Kelly, E.: Justice as Fairness: A Restatement. Belknap Press (2001)
16. Reinders Folmer, C.: Social Motives, pp. 886–890. SAGE (06 2016)
17. Schelling, T.C.: The strategy of conflict prospectus for a reorientation of game theory. The Journal of Conflict Resolution **2**(3), 203–264 (1958)