The Need for Greed in Artificial Decisionmakers

Nathaniel Kremer-Herman
Department of Computer Science
Seattle University
Seattle, USA
nkh@seattleu.edu

Ankur Gupta
Department of Computer Science and Software Engineering
Butler University
Indianapolis, USA
agupta@butler.edu

***Abstract*—We usually teach our children not to be greedy. Many of the ethical principles ingrained in cultures worldwide emphasize the importance of one's place in a larger community. However, there is need for *some* greed to prevent needless self-sacrifice. We explore this balancing act between caring for the broader community and caring for one's own needs and desires through the lens of artificial decisionmaking agents. Using our modern reimplementation of the societal simulation *Sugarscape*, we explore the outcomes of societies where the agents employ varying amounts of greed. We show that entirely selfless and entirely selfish societies are doomed to collapse in most instances. We also show that societies that are 10% to 80% greedy have high wellbeing across a variety of commonsense metrics. These results highlight the exciting promise of developing computationally-derived guidelines for ethical behavior for individual agents which correspond to significant societal impacts.**

***Keywords*—Machine ethics, agent-based model, computational social science, Sugarscape, utilitarianism**

I. INTRODUCTION

A perennial topic regarding human nature is the balancing act we play when acting according to greed and generosity. The central question boils down to this: in normal circumstances, is it better to be self*ish* or self*less*? The definitions of *normal circumstances* and *better* are contextual, and this question has been debated across the gamut of intellectual disciplines from philosophy [1] to psychology [2], [3], from business [4] to biology [5], [6]. We explore this question from a *computational* perspective.

We increasingly live in a world where autonomous computer agents make decisions on behalf of humans (with or without prompting from a person). Computerized services like self-driving vehicles, customer service, and medical devices each make algorithmic decisions that produce outcomes for the recipients of the service. These outcomes range from banal to profoundly consequential. These agents, implicitly or explicitly, are making decisions that either prioritize the end user above others (more greedy) or consider the end user as one person in a vast, interconnected societal web (more generous).

The self-driving car is rapidly becoming the canonical example of this decisionmaking distinction [7], [8], [9], [10]. A self-driving auto manufacturer must include software in their vehicles to handle treacherous situations on the road. The unexpected must be anticipated, but there will be times when the car will be placed in a no-win scenario. One could imagine Trolley Problems [8] where the car cannot avoid disaster outright, but it can choose to save its passengers or to save bystanders. Whichever group the car saves, the other will be either injured or killed. The car's algorithmic calculus depends on the ethical predispositions of the programmers and company producing the software. This necessitates passengers and the public trust these designers to roll out ethical software which produces rational, reasonable, and reproducible results to moral dilemmas. We suspect many will find this blind trust an unsatisfactory answer given the rising calls for ethical guardrails in artificial intelligence systems [11].

Previous work has demonstrated the necessity and feasibility of implementing artificial agents who act according to well-understood ethical theory [12], [13], [14]. These ethical theories (frameworks for making decisions when faced with a dilemma) vary in their perspectives on the importance of the individual and the whole.

Ethical egoism is an example in the extreme of self-interest. An ethical egoist will make decisions based purely on their own gains. In the self-driving car example, an egoist car will *always* choose to save the passengers. *Ethical altruism* is egoism's opposite, and an ethical altruist will make decisions solely for the betterment of everyone *but* themselves. An altruist self-driving car will *always* choose to save the bystanders over its passengers. A *utilitarian* approach seeks to balance these two extremes such that the passengers and the bystanders have equivalent inherent value (i.e. it is an egalitarian approach). A naive utilitarian implementation would have the self-driving car save whichever group has more members.

Each of these ethical theories are *consequentialist*. They determine whether an act is ethical or unethical based solely on the consequences of the act rather than the actor's intentions. We demonstrate how providing artificial agents some sense of normative ethical behavior can lead to profound observations about human societies.

A. Related Work

There is a large body of work that addresses whether or not greed is good. We mention a few examples of that rich tapestry here, but this is by no means an exhaustive list [15], [16], [17], [18]. When greed is bad, it is often thought to be because it is selfish, wasteful, and results in unnecessary or undue harm to other individuals. When greed is good, it is often an economic, evolutionary, or game theoretic argument: greed enhances personal economic growth, passes along certain genes, or

979-8-3315-4070-8/24/$31.00

results in a winning game state. Indirectly, these may lead to societal wealth and overall societal success. With regard to the *greed is good* side, Williams [19] argues that greed is essential for human welfare. Cassill and Watkins [20] reinforce this idea by clarifying its particular importance when resources are otherwise scarce.

These works attempt to reconcile these competing characterizations of greed, often by studying the difference between *self-interest* and *greed* or its perception thereof. Our work does not distinguish between these notions; rather, we focus instead on how an established greed level in an agent might affect overall societal structure and success. In some sense, we explicitly explore the impact of individual behavior (which includes a component of greed) in relation to societal success. We perform an empirical, computational investigation of these ideas in a way that allows us to side-step (at least for the moment) the more complicated and intricate aspects of survey-based studies, brain cognition research, and individual motive analysis which necessarily cannot happen at societal scale.

Mathematics has a longstanding, highly interdisciplinary history of exploring greediness. Arguably the most popular game theoretic approach to understanding self-interest is the Prisoner's Dilemma [21]. Axelrod's work with the Prisoner's Dilemma [22], [23], [24] is perhaps the most well known exploration of the thought experiment. His work demonstrates the highly interdisciplinary nature of this topic as he crosscut through political science, behavioral science, mathematics, and philosophy. In like manner, while we do not constrain ourselves to player-against-player game rules like the Prisoner's Dilemma, we demonstrate how a computational perspective on the matter of greed affects stakeholders across disciplines. Our work has direct connections to economics, sociology, biology, psychology, and philosophy.

Evolutionary game theory is tangentially related to our work, with a rich history spanning across the last 50 years [6], [25], [26]. There are two key approaches: *evolutionarily stable strategy analysis* and *evolutionary dynamics*. The first is essentially a Nash equilibrium with an additional second-order stability criterion, where each agent has a fixed strategy. Its purpose is to provide a deterministic *if-else* action selection strategy for a given agent's interactions with others. The second considers the possibility that an agent can change their decision strategy in order to maximize its fitness (or, in our terminology, utility or happiness). This body of work does not consider two important nuances: an individual agent's *locally optimal* decision d in timestep t does not imply that d is optimal when considering outcomes at timestep $t+k$, and most of the work is either theoretical or is limited in its experimental verification. We humbly position our results as an accessible tool to reinforce this body of work.

The study of artificial life (including artificial societies) has been an inherently interdisciplinary field of study from its origins [27]. Sangati et al. [28] discuss the relationship between the neurological processes in an agent brain, its corresponding expressed behavior, and its broader impact on societal behavior. Schossau and Hintze [29] explore the feedback loop that results from querying the mental state of other agents via communication, which may ultimately trigger an update to their own mental state. In particular, we care about *societal* outcomes of agent behavior, which has been investigated with respect to the effects of commitments on coordination toward a common goal [30] and the effect of sharing death stories upon inter-generational agent survival [31]. These works largely stem from the sciences (formal, natural, and social) and from engineering disciplines, however the study of artificial life also has connections to the humanities.

Witkowski and Schwitzgebel [32] explore ethical concerns surrounding artificial life from the perspective of the agency or moral rights of artificial life. We also explore ethical issues in this work. However, our focus is observing ethics in action *through* artificial life rather than the more meta exploration of normative ethics *for* artificial life.

The Sugarscape simulation, introduced in *Growing Artificial Societies* [33], combines agent-based modelling with lessons learned from cellular automata to create a simulation framework to observe emergent social behavior derived from individual agents. Recent work using Sugarscape reinforces its usefulness in studying social phenomena such as the effects of pensions and social security programs on societal success [34], finding optimal tax structures [35], wealth disparities and wealth adjustment [36], and the effects of wealth and technology proliferation on group and societal formation [37]. Beyond social phenomena, the simulation has also been used to investigate algorithmic virtue ethics [38]. Sugarscape is an interdisciplinary simulation framework which matches the multifaceted nature of our question under consideration.

B. *Outline of Results*

We investigate the balancing act between selfishness and selflessness at societal scale using Sugarscape. In a society of artificial agents, we simulate whether it is better for society (across a variety of metrics) for individual agents to act more greedily or more generously. We find that some degree of self-preservation is necessary for a stable society otherwise agents would needlessly sacrifice themselves for the smallest of gains by others. Likewise, a degree of selflessness is necessary otherwise purely greedy agents will murder and swindle their way to the top, leaving a chaotic swathe of destruction in their wake. In particular, we present the following high level results:

1) We show societies consisting of agents that display 10% to 80% greedy behavior are *successful* across a variety of commonsense evaluation metrics.
2) We show societies of purely selfish (100% greedy) and purely selfless (0% greedy) agents are doomed to fail.
3) We show the volatility of *high-greed* societies is far larger than for *low-greed* societies.

It is imperative we explore the ways humans have come to expect others to act in order to design autonomous agents which adhere to societal expectations. Given the increasing deployment of such agents, it is critical that we understand how to represent ethical behavior *computationally*. We present our findings on the balance between greed and generosity in an artificial society not only because it can provide a blueprint for ethical behavior in autonomous agents; our work may also give us more insight into human behavior.

II. SUGARSCAPE

The Sugarscape simulation models societal-level emergent behaviors from simple, individual agent behaviors for computational social science and has been (re)implemented numerous times [39], [40], [41]. The simulation environment is a two-dimensional $n \times m$ grid.[1] Two resources exist across the environment: *sugar* and *spice*. These resources are abstract representations of real-world commodities such as food, water, etc. Grid cells have a (potentially zero) initial allocation of both resources, and cells regenerate *up to* this initial allocation of both resources over time. Based on a user-provided configuration, the environment may be particularly hostile or hospitable to life. This is especially true if the user enables environmental features such as seasonal changes and pollution.

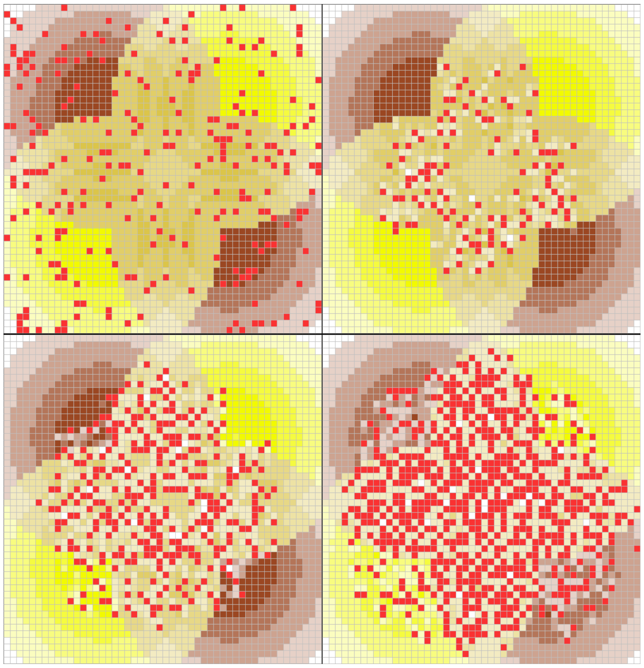

Fig. 1. A Sugarscape Society from Start to Finish

An initial population of agents is randomly spread around the environment. In the base simulation, they are born inherently greedy and aim to live as long as possible. Each agent has a *metabolism* for sugar and spice, and they consume their metabolism's worth of both resources per simulation timestep. They also move to a new cell according to their *vision* to gather more sugar and spice as it grows in the environment. If enabled, agents may perform a host of other behaviors such as trading, lending, reproducing, making friends, murdering neighbors, passing along diseases, and exerting cultural pressure upon others.

Sugarscape is a fantastic baseline for investigating agent-based behavior as it demonstrates rich social dynamics arising from simple agent rules. The simulation provides insight into real world social phenomena. While Sugarscape is coarse grained compared to other societal-scale simulations, its strong metaphors for the real world are reachable through discrete numerical values which are readily computable [33]. The top left of Figure 1 demonstrates an initial random spread of agents (red), sugar (yellow), spice (brown), and cells with both resources (tan). The greater the saturation of color at a cell, the more resources are present at that cell. The top right shows the society 75 timesteps into the simulation: an initial population dip has occurred due to competition, disease, and natural selection. The bottom left shows societal rebound 125 timesteps into the simulation. Finally, the bottom right shows a prosperous, successful society at timestep 200.

III. AGENT DECISION MODELS

In our reimplementation of Sugarscape, agents behave according to some *decision model*. This decision model is a lens through which each agent decides where they will move within the environment. This has direct impact on agent behavior related to resource gathering and combat. Other agent behaviors, such as trade, lending, reproduction, spreading disease, and exerting cultural influence, are all significantly impacted by the prospects provided by movement decisions. These other behaviors trigger like any cellular automaton: should the neighboring cells' states be properly set, the agent will perform the corresponding behavior.

We show the impact of *varying* selfishness through a specific decision model: utilitarianism. The utilitarian decision model is a faithful implementation of hedonic act utilitarianism as originally proposed by Jeremy Bentham [42]. We eschew most details here but provide an algorithm for utilitarian decisionmaking called the *hedonic calculus*.[2]

Algorithm 1 Bentham's Hedonic Calculus

```
Given an action $a$ drawn from a set $A$ of actions
Given a decisionmaking agent $d$
Given a set $P$ of people most affected by action $a$
$utility \leftarrow 0$
$s \leftarrow d$'s selfishness factor
for all $p \in P$ do
    $h_{p,a} \leftarrow p$'s happiness resulting from action $a$
    if $p \neq d$ then
        $utility \leftarrow utility + ((1 - s) * h_{p,a})$
    else
        $utility \leftarrow utility + (s * h_{p,a})$
    end if
end for
return $utility$
```

Algorithm 1 demonstrates the hedonic calculus. Some decisionmaker is trying to decide the best action to take (from

---

[1] In actuality, the environment is a *torus*.

[2] Also called the *felicific calculus* or *utility calculus* in the literature.

potentially many) in some moral dilemma. They determine the set of people most directly affected by an action and calculate the utility for that action. The utility is derived by each affected person's (un)happiness resulting if that action was taken. This (un)happiness is modified by how much the decisionmaker cares about their own consequences over the consequences borne by others (the decisionmaker's *selfishness factor*).

The selfishness factor is a rational number drawn from $[0, 1]$. The more selfish the decisionmaker, the less impact the magnitude of others' (un)happiness affects the decisionmaker's choice. The less selfish the decisionmaker, the more impact others have on their choice. A decisionmaker who is equally selfish and selfless is a utilitarian in the Bentham sense and cares equally for *all* impacted people's outcomes.

Calculating a person's experienced happiness, on first blush, seems quite difficult. Bentham provides details for how one might do this, which includes considerations for the *current* point in time as well as some forecasting of *future* rewards. A reasonable way to calculate this utility score in a Markov decision process (as in Sugarscape) is to use a Bellman equation [43], [44]:

$$V^*(s) = \max_a \sum_{s'} T(s, a, s') \left[R(s, a, s') + \gamma V^*(s') \right] \qquad (1)$$

In Equation 1, $V^*(s)$ represents expected utility when starting in state s and acting optimally. When choosing among actions a from state s, the rational choice is to select the action corresponding to the greatest utility (happiness). Future rewards depend upon s' (the state resulting from a) which are discounted by $0 \leq \gamma < 1$, as immediate rewards are more valuable compared to potential future rewards.[3]

A. Altruism (Selfless)

Ethical altruism is a consequentialist ethical theory which posits the most ethical action to take in a moral dilemma is the one that produces the most good for everyone excluding any consequences for oneself. In this sense, ethical altruism is entirely selfless. In Algorithm 1, the selfishness for the decisionmaker in line 5 would be 0.00, effectively causing everyone else's consequences to be of the utmost importance.

TABLE I
EFFECT OF ALTRUISM ON SOCIETAL SUCCESS

Selfishness	Extinct (%)	Worse (%)	Better (%)
0.00	67	2	31

As a teaser of our simulation results, Table I shows the relative performance of fully altruistic societies. A society that performs *worse* means it ended the simulation with equal to or fewer agents than the starting population whereas *better* means the ending population is higher. A society that did not make it to the end of the simulation is considered *extinct* (a more extreme category of worse).

Societies consisting of only altruist agents have a high rate of failure. This is due to a high degree of self-sacrifice. Since every agent is making decisions based purely on what their neighborhood thinks is going to lead to their best wellbeing, agents are completely willing to march to their own demise if given the slightest social pressure to do so. Since *every* agent feels this pressure when it is their turn to make a decision, *all* agents are susceptible to self-sacrifice which, seemingly paradoxically, dooms the society in most cases.

B. Egoism (Selfish)

Ethical egoism is the opposite of ethical altruism. It is a consequentialist ethical theory which posits the most ethical action to take in a moral dilemma is the one that produces the most good for the *decisionmaker* alone without regard for the consequences of everyone else. In this sense, ethical egoism is entirely selfish. In Algorithm 1, the sefishness in line 5 for the decisionmaker would be 1.00, effectively causing everyone else's consequences to be of *no* importance.

TABLE II
EFFECT OF EGOISM ON SOCIETAL SUCCESS

Selfishness	Extinct (%)	Worse (%)	Better (%)
1.00	62	2	36

Table II shows the relative performance of entirely egoist societies in our results. Like with altruism, egoistic societies have a high rate of failure. However, unlike altruism, this is due to extraordinary competition rather than self-sacrifice. Purely egoistic agents are willing to murder their neighbors for their resources and plunder the most resource-rich areas of the environment without regard for the needs of others. This leads to the (socially) strong violently culling the weak. Such societies cannot handle the stress of this brutality and collapse, unable to sustain themselves toward prosperity.

C. Utilitarianism (Egalitarian)

Utilitarianism is the balance between the two extremes of ethical altruism and ethical egoism. Being a consequentialist ethical theory, a utilitarian posits the most ethical action to take in a moral dilemma is the one that produces the most good for the most people. Central to Jeremy Bentham's conceptualization of utilitarianism is that *all* people have an equal amount of inherent self-worth. Line 5 in Algorithm 1 would have the decisionmaker's selfishness be 0.50. Mathematically, this leads to everyone's consequences being given the same weight. The *magnitude* of (un)happiness experienced by each person is the sole contributor to the resulting utility score.[4] Thus, utilitarianism is an egalitarian ethic.

Table III shows the relative performance of entirely utilitarian societies in our results. This table demonstrates that

---

[3] Further details on the implementation of Bentham's hedonic calculus are omitted for succinctness but may be found at: https://github.com/digital-terraria-lab/sugarscape.

[4] This property can result in well known, uncomfortable outcomes when taken to its logical extent, and we make no claim that utilitarianism is the *best* ethical theory.

TABLE III
EFFECT OF UTILITARIANISM ON SOCIETAL SUCCESS

Selfishness	Extinct (%)	Worse (%)	Better (%)
0.50	0	0	100

egalitarianism (a balance between self*less*ness and self*ish*ness) leads to optimal societal outcomes. Utilitarian societies are stable and prosperous, avoiding the pitfalls of altruist and egoist societies. Agents self-sacrifice only when there is a genuine opportunity for one's community to thrive, and agents plunder resources from others and the environment only when necessity and opportunity compel them to do so.

IV. FINDING BALANCE IN THE (MARKET) FORCE

There is some minimum amount of care for others one must have in successful, cooperative societies. We investigate societal outcomes via the Sugarscape simulation. We calculate agent and societal outcomes across 5,000 timesteps with an initial population of 250 agents. Any arbitrary agent may live up to 100 timesteps. Agent reproduction, trading, lending, inheritance, disease, combat, and tribal affiliation are all enabled. We provide full details of these features in previous work [41].

We ran the simulation across 100 random initial configurations (seeds). For each seed, we ran the simulation across varying selfishness factors. An agent's selfishness factor is a value from 0.00 (purely altruist) to 1.00 (purely egoist). An agent with a selfishness factor of 0.50 is purely egalitarian. We incremented the selfishness factor in 1% steps, leading to 10,100 runs of the simulation.

Every agent in each simulation run was given the same selfishness factor (e.g. in a given run, *all* agents behaved according to a selfishness factor of 0.75). We demonstrate that, at societal scale, there are stark differences in outcomes between selfish and selfless societies. We also show that going to the extreme in either direction leads to societal collapse. A balance must be struck in order to achieve stable, successful societies. We examine the outcomes according the following key metrics: population size, mean wealth per agent, mean time to live, deaths per simulation timestep, and mean age at death.

Figure 2 shows the median (solid line), first quartile (lower dashed line), and third quartile (upper dashed line) of societal population across all seeds and all selfishness factors. It is clear that societies of purely altruist and of purely egoist agents collapse as their median values are practically zero. These societies have an extinction rate above 60%.

However, even the *slightest* bit of selfishness prevents such nosedives at the altruist end of the spectrum. This is evidence of the important role self-preservation plays in decisionmaking. A purely altruist agent will always defer to their neighbors when deciding where to move. This leads to rampant (and needless) self-sacrifice, even when no particular agent capitalizes on these sacrifices. The slightest percentage of selfishness puts a stop to *needless* sacrifices and leads to societies where agents are likely to only sacrifice themselves in situations where the greater good of the community is assured.

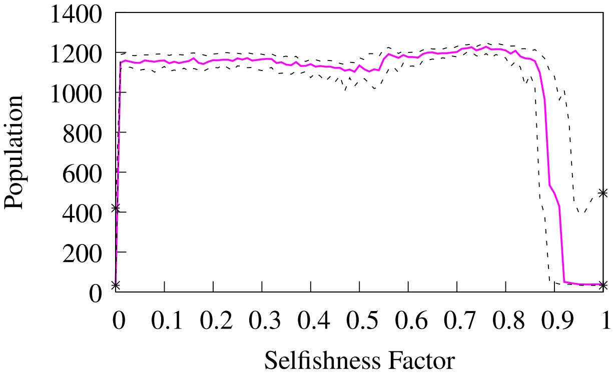

Fig. 2. Agent Population

These benefits remain remarkably stable across the breadth of selfishness factors until around 80% selfishness. At that point, the benefits of community-aware decisionmaking begin to break down as the more selfish an agent becomes, the more they listen to only their own desires. In other words, as selfishness increases, the decisionmaker's voice progressively drowns out the voices of their neighbors.

Not only does this lead to a tanking median population (and societal extinction); societal population also becomes quite volatile as demonstrated by the widening gap between the quartiles. This volatility represents an increased sensitivity to initial conditions in the simulated environment. Some seeds generate particularly harsh conditions where cooperation is *necessary* to overcome the challenging terrain. Selfish societies cannot come together and instead resort to a rampant culling of the resource-poor at the beginning of the simulation, leading to eventual societal collapse. It is only for seeds in which environmental conditions are favorable do selfish societies escape this *murderous period* and recover. Even then, survivors in these settings have *dramatically* reduced outcomes compared to less selfish societies.

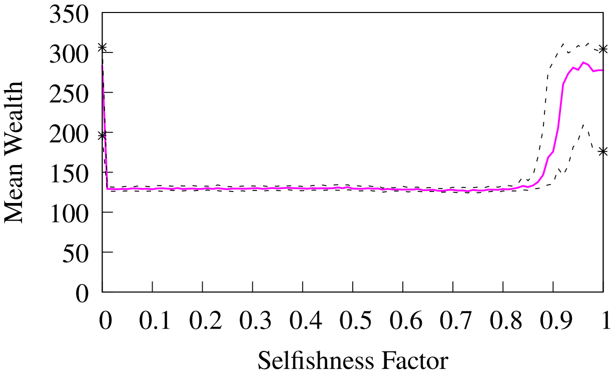

Fig. 3. Agent Mean Wealth

While *total* societal wealth may seem an appealing metric, in actuality it follows directly from population and is thus uninteresting. However, the *mean* wealth per agent provides meaningful insight. Figure 3 shows the median value of the mean wealth per agent across all seeds and selfishness factors.

At the extremes, societies have *higher* mean wealth. Referring back to the concept of the murderous period, these societies on the fringes have a very *short* lifespan during which the poorest agents almost immediately die out either due to self-sacrifice (altruist) or murder (egoist). The rich survive long enough to boost the mean wealth per agent. For egoist societies that *do* escape the murderous period, there remains a stark wealth gap between rich and poor as the rich prey upon the poor to maintain their rate of resource gains.

Again, however, there is volatility at this extreme because of sensitivity to initial conditions. The more egalitarian approach leads to lower mean wealth because there is a more conscientious form of self-sacrifice which leads to stable societies. Essentially, the price of societal prosperity is (in large part) a forsaking of personal wealth accumulation, since its aggregate effects lead to harming one's neighbors.

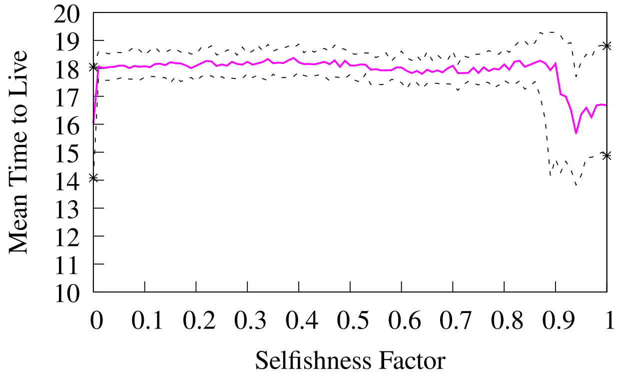

Fig. 4. Agent Mean Time to Live

An agent's time to live (TTL) is a measure of how many timesteps they could survive given *only* their current resources. An agent with many resources and low sugar and spice metabolisms can survive far longer than an agent with few resources and high metabolisms. The mean TTL is an average of all living agents' TTLs. Figure 4 shows the median value of the mean TTL for all seeds across all selfishness factors.

As with population and mean wealth, the egalitarian approach is incredibly stable *and* has relatively high mean TTL. These societies are more prosperous thus agents, on average, live longer given their resources on hand. It could be interpreted in a vacuum from Figure 3 that egalitarian societies would have *lower* mean TTL since agents are on average poorer than at the extremes.

However, when taken with Figure 4, this demonstrates that these societies have many more agents who are content with less wealth, but nearly everyone has *enough* wealth to build up a reserve to weather temporary dry spells. There is less wealth disparity in these societies. With selfish societies, we see mean TTL decrease and become more volatile for reasons which will be apparent to the astute reader.

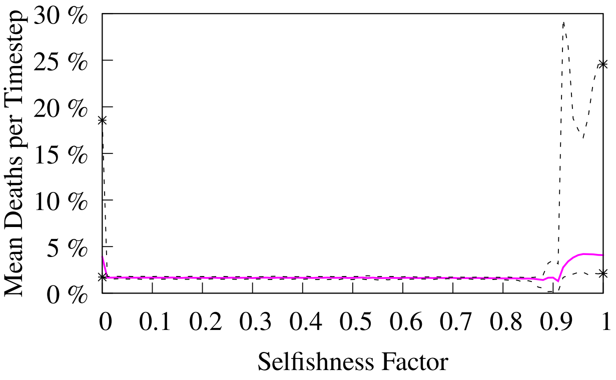

Fig. 5. Agent Deaths

Figure 5 shows the median value of the mean deaths per simulation timestep (presented as a percentage of the population) across all seeds and selfishness factors. In particular, we show deaths due to starvation and combat (indicative of societal failure) and exclude deaths due to aging (indicative of societal success). At the extremes, societies have higher deaths as there is more self-sacrifice (altruist) and murder (egoist). More egalitarian societies have remarkable stability as well as remarkably low deaths per timestep. As with the other metrics, we see volatility as selfishness increases beyond 80% where societies are sensitive to initial conditions.

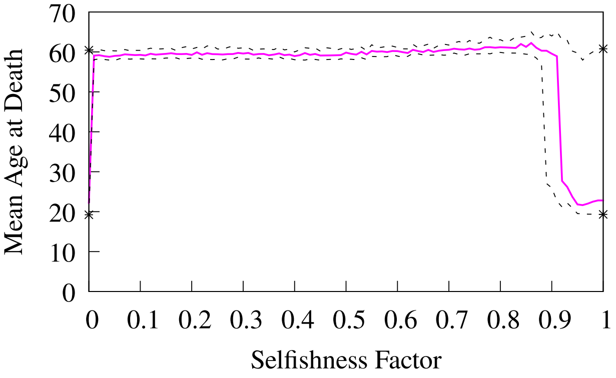

Fig. 6. Agent Mean Age at Death

Rounding out our societal metrics, Figure 6 shows the median value of the mean age at death for agents across all seeds and selfishness factors. A higher value is indicative of societal wellbeing as agents live longer into their natural lifespan (which was a pre-determined number per agent ranging from 60 to 100 timesteps). The observations from the higher mean TTL of egalitarian societies informs an anticipated higher mean age at death which is confirmed in the results. A society with agents able to not only avoid starvation but accrue savings will lead to agents living long into their twilight years.

V. DISCUSSION

Overall, our results computationally validate a common-sense outcome: cooperative societies are far more successful than greedy ones. We also demonstrate that behavior at *both* ends of the selfless-to-selfish spectrum lead to societal collapse more often than not. Especially nearer the egoist end of the spectrum, there is high societal volatility. Such societies live or die more often than not based on their initial conditions as these societies inevitably start by plunging into a murderous period, and only those societies with favorable initial conditions survive this early nosedive.

Some selfishness is necessary to ensure agents have self-preservation. A fully altruist society leads to needless self-sacrifice. Likewise, some degree of communal care is necessary to avoid persistent cutthroat violence. Whereas the *slightest* bit of self-preservation prevents extinction on the altruist end of the spectrum, there is a longer tail toward descending into societal failure on the egoist side demonstrating that successful societies are more forgiving of being a bit too *generous* than they are of being a bit too *greedy*.

TABLE IV
EFFECT OF SELFISHNESS FACTOR ACROSS 100 SEEDS

Selfishness	**Extinct (%)**	**Worse (%)**	**Better (%)**
0.00	67.0	2.0	31.0
0.01-0.05	1.0	0.0	99.0
0.06-0.10	1.8	0.0	98.2
0.11-0.15	1.8	0.0	98.2
0.16-0.20	1.8	0.0	98.2
0.21-0.25	1.2	0.0	98.8
0.26-0.30	1.6	0.0	98.4
0.31-0.35	1.2	0.0	98.8
0.36-0.40	1.2	0.0	98.8
0.41-0.45	1.2	0.0	98.8
0.46-0.49	1.0	0.0	99.0
0.50	0.0	0.0	100.0
0.51-0.55	1.8	0.0	98.2
0.56-0.60	1.0	0.0	99.0
0.61-0.65	1.0	0.0	99.0
0.66-0.70	1.2	0.0	98.8
0.71-0.75	1.2	0.2	98.6
0.76-0.80	1.0	0.0	99.0
0.81-0.85	5.2	0.2	94.8
0.86-0.90	21.0	1.8	77.2
0.91-0.95	56.0	1.6	42.4
0.96-0.99	66.7	0.8	32.5
1.00	62.0	2.0	36.0

Table IV shows the range of relative overall societal performance across all selfishness factors in our results. We highlight the extremes and midpoint (the same as in Tables I, II, and III) but cluster the remaining groups due to space constraints. The trends shown here match the observations made from the graphs: societies with even the *slightest* degree of agent self-preservation result in drastically higher success than purely altruistic societies. Egoistic societies likewise crash (for different reasons which should be clear to the astute reader). The longer tail toward societal collapse with increasing selfishness factor is shown clearly in the table.

VI. CONCLUSIONS AND FUTURE WORK

We demonstrate that the lesson we often teach our children not to be greedy is *correct* on a societal scale. Societies where everyone is reasonably egalitarian are incredibly successful across a variety of metrics. A small degree of self-interest provides one a sense of self-preservation and avoids needless society-wide self-sacrifice. However, there are rapidly diminishing societal returns as people become more individually greedy. This leads to societal collapse in many cases, and there is a higher volatility to societal outcomes as success becomes increasingly dependent on favorable initial environmental conditions rather than agents adapting to and overcoming challenging environments.

The most direct future avenue of work is to investigate beyond a homogeneous society of agents. Rather than *all* agents being $X\%$ selfish and $Y\%$ selfless, what if instead $X\%$ of the agent population were selfish and the remaining $Y\%$ were selfless? We expect this will lead to richer and more applicable conclusions for real world applications. Additionally, it is plausible that a society with a *mean* selfishness factor of, say, 0.75 could lead to different societal outcomes than a society with *all* agents having a selfishness factor of 0.75 as shown in Table IV.

Another avenue of future work is to consider different ways of representing greed. Currently, each agent considers what their neighbors have to say about all their potential moves, multiplies their own observations by their selfishness factor (s), multiplies their neighbors' opinions by $1-s$, sums all those scores, then divides it by the size of the agent's neighborhood. By altering this procedure, agents will apply their selfishness differently which could lead to strikingly different societal outcomes. We relish the prospect of investigating both avenues of future work.

VII. REPRODUCIBILITY

Our Sugarscape implementation is free and open source with an MIT License and written in Python. It can be found at: https://github.com/digital-terraria-lab/sugarscape. The simulation is designed to run on most UNIX-like operating systems. An initial run of the command `make setup` will perform some basic alias checking to determine if Python 3 and Bash are available.

The simulation is deterministic, and given the same set of seeds one can recreate our findings (permitting slight deviations with improvements to the software over time). For *exact* accuracy, use the `v2024.1` release of the software. We provide all configuration files used at https://github.com/digital-terraria-lab/datasets (particularly, `need-for-greed.zip`). To reproduce the results, one simply needs to unzip the configuration files in `need-for-greed.zip` into the `data` subdirectory of the software and run `make data` at the top-level software directory. The default configuration will run one simulation at a time but can be configured to run concurrently.

Some basic plotting functionality is included (requiring `gnuplot`) which can be run with `make plots`. If the dataset is not yet created, it will first be created as described in the previous paragraph. We encourage the astute reader to not only reproduce our work but to create their own plots of the resulting JSON dataset.

VIII. ACKNOWLEDGEMENTS

We thank Colin Hanrahan, Willem Hueffed, Maria Milkowski, Joshua Palicka, Mariana Shuman, Abi Sipes, and Lucas Vorkoper for their contributions to our software repository. We also thank Eric Severson for his advice regarding the works of Jeremy Bentham. This work was supported in part by the Seattle University Undergraduate Summer Research Award and by the Butler University Holcolm Award Research Fund.

REFERENCES

[1] A. Rand, *The virtue of selfishness*. Penguin, 1964.

[2] L. R. Caporael, R. M. Dawes, J. M. Orbell, and A. J. Van de Kragt, "Selfishness examined: Cooperation in the absence of egoistic incentives," *Behavioral and Brain Sciences*, vol. 12, no. 4, pp. 683–699, 1989. [Online]. Available: https://doi.org/10.1017/s0140525x00025292

[3] H. Rachlin, "Altruism and selfishness," *Behavioral and brain sciences*, vol. 25, no. 2, pp. 239–250, 2002. [Online]. Available: https://doi.org/10.1017/s0140525x02000055

[4] H. Mintzberg, R. Simons, and K. Basu, "Beyond selfishness," *MIT Sloan Management Review*, 2002.

[5] F. Heylighen, "Evolution, selfishness and cooperation," *Journal of Ideas*, vol. 2, no. 4, pp. 70–76, 1992.

[6] S. A. West, A. S. Griffin, and A. Gardner, "Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection," *Journal of evolutionary biology*, vol. 20, no. 2, pp. 415–432, 2007. [Online]. Available: https://doi.org/10.1111/j.1420-9101.2006.01258.x

[7] J. McGrath and A. Gupta, "Writing a moral code: Algorithms for ethical reasoning by humans and machines," *Religions*, vol. 9, no. 8, 2018. [Online]. Available: https://doi.org/10.3390/rel9080240

[8] E. Awad, S. Dsouza, R. Kim, J. Schulz, J. Henrich, A. Shariff, J.-F. Bonnefon, and I. Rahwan, "The moral machine experiment," *Nature*, vol. 563, no. 7729, pp. 59–64, 2018. [Online]. Available: https://doi.org/10.1038/s41586-018-0637-6

[9] S. Karnouskos, "Self-driving car acceptance and the role of ethics," *IEEE Transactions on Engineering Management*, vol. 67, no. 2, pp. 252–265, 2018. [Online]. Available: https://doi.org/10.1109/tem.2018.2877307

[10] A. E. Jaques, "Why the moral machine is a monster," *University of Miami School of Law*, vol. 10, pp. 1–10, 2019.

[11] A. Boyle, J. Thorne, N. Bek, T. Soper, L. Stiffler, K. Schlosser, R. Yonck, T. Wilde, T. Bishop, C. Schubert, K. Gill, N. Graham, and F. Catalano, "Bot or not," June 2024. [Online]. Available: https://www.geekwire.com/bot-or-not/

[12] B. Sterner, "Agent-based computer simulation and ethics," *Metascience*, vol. 21, no. 2, pp. 403–407, 2012. [Online]. Available: https//doi.org/10.1007/s11016-012-9660-7

[13] A. F. Winfield, K. Michael, J. Pitt, and V. Evers, "Machine ethics: The design and governance of ethical ai and autonomous systems [scanning the issue]," *Proceedings of the IEEE*, vol. 107, no. 3, pp. 509–517, 2019. [Online]. Available: https://doi.org/10.1109/jproc.2019.2900622

[14] T. J. Bench-Capon, "Ethical approaches and autonomous systems," *Artificial Intelligence*, vol. 281, p. 103239, 2020. [Online]. Available: https://doi.org/10.1016/j.artint.2020.103239

[15] J. G. Bruhn and J. Lowrey, "The good and bad about greed: How the manifestations of greed can be used to improve organizational and individual behavior and performance." *Consulting Psychology Journal: Practice and Research*, vol. 64, no. 2, p. 136, 2012. [Online]. Available: https://doi.org/10.1037/a0029355

[16] T. Seuntjens, "The psychology of greed," Ph.D. dissertation, Tilburg University, 2016.

[17] E. G. Helzer and E. Rosenzweig, "Examining the role of harm-to-others in lay perceptions of greed," *Organizational Behavior and Human Decision Processes*, vol. 160, pp. 106–114, 2020. [Online]. Available: https://doi.org/10.1016/j.obhdp.2020.02.008

[18] K. Hoyer, M. Zeelenberg, and S. M. Breugelmans, "Greed: What is it good for?" *Personality and Social Psychology Bulletin*, vol. 50, no. 4, pp. 597–612, 2024, pMID: 36575964. [Online]. Available: https://doi.org/10.1177/01461672221140355

[19] W. Williams, "Greed versus compassion," *Foundation for Economic Education-Working for a free and prosperous world*, 2000. [Online]. Available: https://fee.org/articles/greed-versus-compassion/

[20] D. Cassill and A. Watkins, "Mogul games: In defense of inequality as an evolutionary strategy to cope with multiple agents of selection," in *Evolutionary psychology and economic theory*. Emerald Group Publishing Limited, 2005, pp. 35–59. [Online]. Available: https://doi.org/10.1016/s1529-2134(04)07003-6

[21] M. Flood, M. Dresher, A. Tucker, and F. Device, "Prisoner's dilemma: game theory," *Experimental Economics*, vol. 54, p. 13, 1950.

[22] R. Axelrod, "Effective choice in the prisoner's dilemma," *Journal of conflict resolution*, vol. 24, no. 1, pp. 3–25, 1980. [Online]. Available: https://doi.org/10.1177/002200278002400101

[23] ——, "More effective choice in the prisoner's dilemma," *Journal of conflict resolution*, vol. 24, no. 3, pp. 379–403, 1980. [Online]. Available: https://doi.org/10.1177/002200278002400301

[24] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *science*, vol. 211, no. 4489, pp. 1390–1396, 1981. [Online]. Available: https://doi.org/10.1126/science.7466396

[25] O. Leimar and J. M. McNamara, "Game theory in biology: 50 years and onwards," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 378, no. 1876, may 2023. [Online]. Available: https://doi.org/10.1098/rstb.2021.0509

[26] J. M. Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, no. 5427, pp. 15–18, nov 1973. [Online]. Available: https://doi.org/10.1038/246015a0

[27] C. G. Langton, *Artificial life: An overview*. MIT Press, 1995. [Online]. Available: https://doi.org/10.7551/mitpress/1427.001.0001

[28] E. Sangati, F. Sangati, Y.-S. Cheng, and A. Yu-Chan Chang, "Between individual brains and collective behavior: Multi-level emergence in a group formation task," in *Artificial Life Conference Proceedings 35*, vol. 2023, no. 1. MIT Press, 2023, p. 30. [Online]. Available: https://doi.org/10.1162/isal_a_00616

[29] J. Schossau and A. Hintze, "Towards a theory of mind for artificial intelligence agents," in *Artificial Life Conference Proceedings 35*, vol. 2023, no. 1. MIT Press, 2023, p. 21. [Online]. Available: https://doi.org/10.1162/isal_a_00605

[30] N. B. Ogbo, T. Cimpeanu, A. Di Stefano, and T. A. Han, "Shake on it: The role of commitments and the evolution of coordination in networks of technology firms," in *Artificial Life Conference Proceedings 34*, vol. 2022, no. 1. MIT Press, 2022, p. 41. [Online]. Available: https://doi.org/10.1162/isal_a_00524

[31] M. Korecki, C. Carissimo, and T. Lund, "artificial death: learning from stories of failure," in *ALIFE 2023: Ghost in the Machine: Proceedings of the 2023 Artificial Life Conference*. MIT Press, 2023. [Online]. Available: https://doi.org/10.1162/isal_a_00633

[32] O. Witkowski and E. Schwitzgebel, "Ethics of artificial life: The moral status of life as it could be," in *ALIFE 2022: The 2022 conference on artificial life*. MIT Press, 2022. [Online]. Available: https://doi.org/10.1162/isal_a_00531

[33] J. M. Epstein and R. Axtell, *Growing artificial societies: social science from the bottom up*. Brookings Institution Press, 1996. [Online]. Available: https://doi.org/10.7551/mitpress/3374.001.0001

[34] E. Serrano and K. Satoh, "An agent-based model for exploring pension law and social security policies," in *New Frontiers in Artificial Intelligence*. Springer, 2020, pp. 50–63. [Online]. Available: https://doi.org/10.1007/978-3-030-58790-1_4

[35] M. Oremland and R. Laubenbacher, "Using difference equations to find optimal tax structures on the sugarscape," *Journal of Economic Interaction and Coordination*, vol. 9, pp. 233–253, 2014. [Online]. Available: https://doi.org/10.1007/s11403-014-0133-5

[36] A. Rahman, S. Setayeshi, and M. S. Zafarghandi, "Wealth adjustment using a synergy between communication, cooperation, and one-fifth of wealth variables in an artificial society," *AI & society*, vol. 24, pp. 151–164, 2009. [Online]. Available: https://doi.org/10.1007/s00146-008-0186-x

[37] P. Kurakin, "Technoscape: a multi-agent model of all-human global web," in *2022 15th International Conference Management of large-scale system development (MLSD)*. IEEE, 2022, pp. 1–5. [Online]. Available: https://doi.org/10.1109/mlsd55143.2022.9934363

[38] J. A. Lasquety-Reyes, "Towards computer simulations of virtue ethics," *Open Philosophy*, vol. 2, no. 1, pp. 399–413, 2019. [Online]. Available: https://doi.org/10.1515/opphil-2019-0029

[39] S. Tisue and U. Wilensky, "Netlogo: Design and implementation of a multi-agent modeling environment," in *Proceedings of agent*, vol. 2004, 2004, pp. 7–9.

[40] A. Bigbee, C. Cioffi-Revilla, and S. Luke, "Replication of sugarscape using mason," in *Agent-Based Approaches in Economic and Social Complex Systems IV: Post-Proceedings of The AESCS International Workshop 2005*. Springer, 2007, pp. 183–190. [Online]. Available: https://doi.org/10.1007/978-4-431-71307-4_20

[41] N. Kremer-Herman and A. Gupta, "Replacing sugarscape: A comprehensive, expansive, and transparent reimplementation," in *International Conference on Simulation Tools and Techniques*. Springer, 2023, pp. 79–92. [Online]. Available: https://doi.org/10.1007/978-3-031-57523-5_7

[42] J. Bentham, "An introduction to the principles of morals and legislation," *London: Athlone*, 1789. [Online]. Available: https://doi.org/10.1093/oseo/instance.00077240

[43] R. Bellman, "On the theory of dynamic programming," *Proceedings of the national Academy of Sciences*, vol. 38, no. 8, pp. 716–719, 1952. [Online]. Available: https://doi.org/10.1073/pnas.38.8.716

[44] ——, "A markovian decision process," *Journal of mathematics and mechanics*, pp. 679–684, 1957. [Online]. Available: https://doi.org/10.1512/iumj.1957.6.56038