`final version:`
`Robotics and Autonomous Systems, 39 (2002), p.223-233, Elsevier`

An N-Player Prisoner's Dilemma in a Robotic Ecosystem

Andreas Birk [a,1] Julie Wiernik [b]

[a] *International University Bremen, Campus Ring 1, 28759 Bremen, Germany*
[b] *Vrije Universiteit Brussel, AI-Lab, Pleinlaan 2, 1050 Brussels, Belgium*

Abstract

The iterated prisoner's dilemma (iPD) in its standard form is a very well known, popular basis for research on cooperation. Here, a significant generalization of the iPD is presented in form of a N-player game, instead of only two players, with continuous degrees of cooperation, instead of binary cooperation or defection. This continuous N-player PD (CN-PD) is motivated and explained by an actual experimental set-up in form of an artificial ecosystem including mobile robots. Furthermore, the novel strategy of so-called justified-snobism (JS) is introduced, which tries to cooperate slightly more than the average of the group of players. Results from sets of experiments with $N = 20$ are presented that indicate that JS is indeed a successful strategy in the CN-PD.

Key words: cooperation, teams, game theory, ecosystem, animat

1 Introduction

Following the "artificial life route to artificial intelligence" [SB93], robot ecosystems allow the investigation of simple animal-like robots or animats [Wil91] to gain insights about basic biological phenomena including the social aspects of intelligences. Multi-robot environments, see e.g. [DTF98] for an overview, are used to conduct experiments on a wide range of topics like evolution of cooperative control [GHB96], language acquisition [MH96], collective intelligence [Mat93] and food chains [WM96], to name just a few.

---

[1] a.birk@iu-bremen.de

Since at least well fifteen years, the iterated prisoner's dilemma is a very popular basis for more theoretically oriented research on cooperation in general and especially for the evolution of cooperation. It made its way into economics, social sciences, biology, and many other scientific fields. For example, the Social Science Citation Index already counted more than 2500 references to Axelrod's ground-breaking work [Axe84] until spring 2000 [Hof00]. But in its basic form the iterated prisoner's dilemma is very limited as it is only a two-player game and as it allows only binary decisions whether to cooperate or to defect. Both aspects have been separately discussed and criticized many times before. For the lack of N-player aspects, see for example [Col82,Har68]. For the in-sufficiency of binary cooperation see for example the work of [RS98], where also a successful strategy for the continuous 2-player PD is presented.

Here, a scenario within a robot ecosystem is used to motivate a generalization of the prisoner's dilemma (PD) to a case with continuous degrees of cooperation and N players (CN-PD). Especially, it is shown how this scenario can be treated in a formal model. In addition, the novel strategy of so-called Justified-Snobism (JS) is presented. JS cooperates slightly more than the average cooperation level of the group of N players if a non-negative pay-off was achieved in the previous iteration, and it cooperates exactly at the previous average cooperation level of the group otherwise. So, JS tries to be slightly more cooperative than the average. This leads to the name for this strategy as the snobbish belief to be "better" than the average of the group is somehow justified for players which use this strategy. In the N-player case of the PD, previous research up to our knowledge failed to come up with evolved cooperation or successful strategies, except for small N. It seems that the combination of the N-player case with degrees of cooperation is beneficial, as it allows to make subtle reactions to the average group behavior, like JS does. This view is supported by results presented in this article for $N = 20$.

The rest of this article is structured as follows. In section 2, a scenario in a robot ecosystem is introduced, which is used to motivate and to explain the basic ideas of the CN-PD. Section 3 presents a formal model of this scenario. In section 4, strategies for iterated games are discussed. Experiments with $N = 20$ and their results are described in the 5th section. Section 6 concludes the article.

2 The Background

In its basic version [Ste94,McF94], the so-called VUB ecosystem (figure 1) consist of

the charging-station where robots can autonomously re-load their batter-

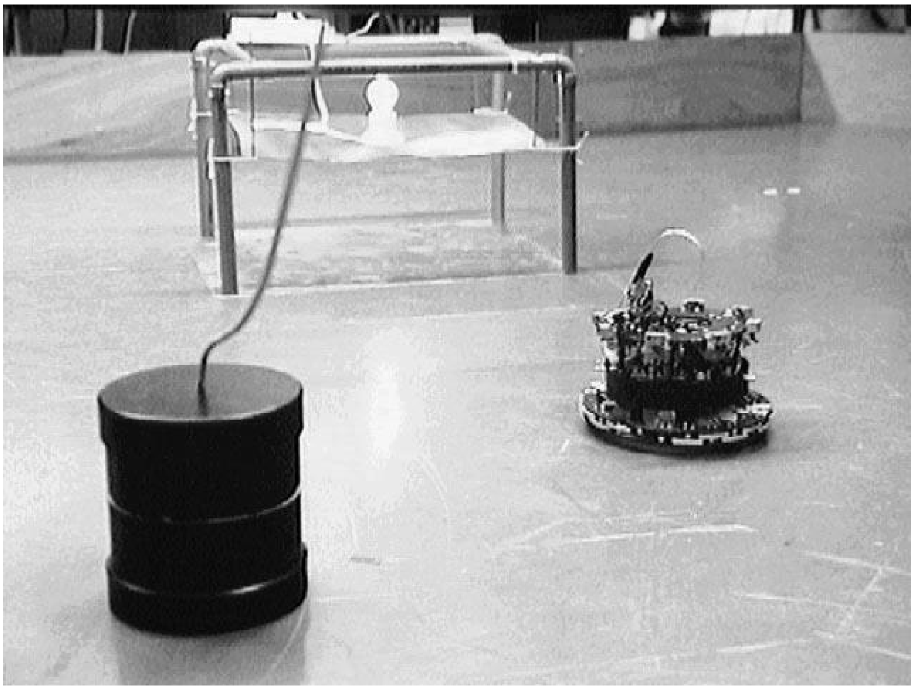

Fig. 1. A partial view of the basic ecosystem with the charging station (center), one of the mobile robots (right) and one of the competitors (left).

ies. The charging-station is equipped with a bright white light on top as a kind of landmark.

simple mobile robots , the so-called "moles" [2] which are equipped with basic behaviors for obstacle avoidance. In addition, they do phototaxis towards the charging station and towards

the competitors , boxes housing lamps connected to the same global energy source as the charging station. The modulated light of these lamps can be dimmed by other inhabitants of the ecosystem by knocking against the boxes. The competitors establish a kind of working task which is paid in energy.

The mechanical design of the different robots in the ecosystem, like e.g. the moles (figure 2), is very simple and imprecise. The mechanical components of most robots are based on toy-construction-kits like e.g. LEGO^{TM} and Fischertechnik^{TM}. But on the electronics side, the robots are equipped with elaborated on-board control-computers and sensors [BKW00,BKW98,Ver96].

An extended version of the ecosystem features an additional inhabitant, the so-called "head" (figure 3). It consists of a camera on a pan-tilt-unit and it has

---

[2] The names for this and of other robot-"species" should not be taken literally. The name "moles" attributes to the fact, that these robots have no on-board vision. We use these names for reasons of convenience, not to imply or suggest any deeper relation with the biological versions.

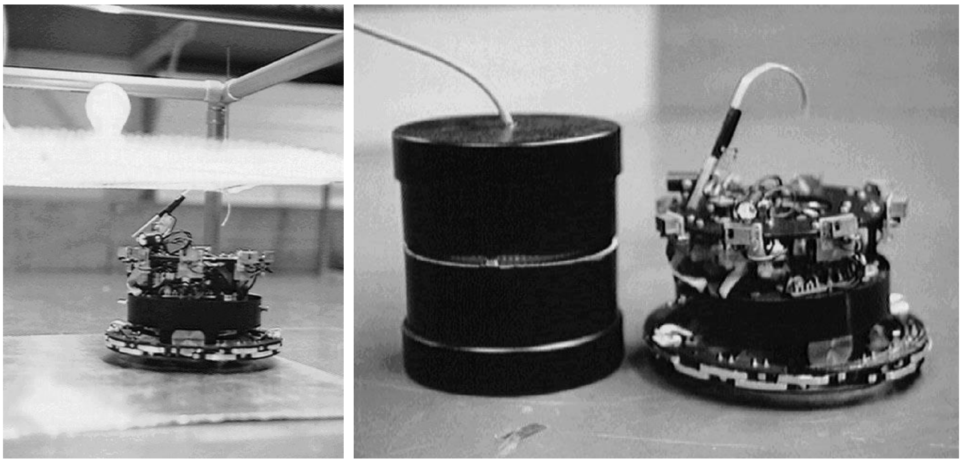

Fig. 2. A so-called mole, a simple autonomous robot which is capable to stay operational in the ecosystem over extended periods in time. It features several basic behaviors. Among them, obstacle avoidance to prevent physical damage, photo-taxis to the charging-station to re-load its batteries (left), and a photo-taxis to the competitors which results in a kind of "fighting"-behavior (right).

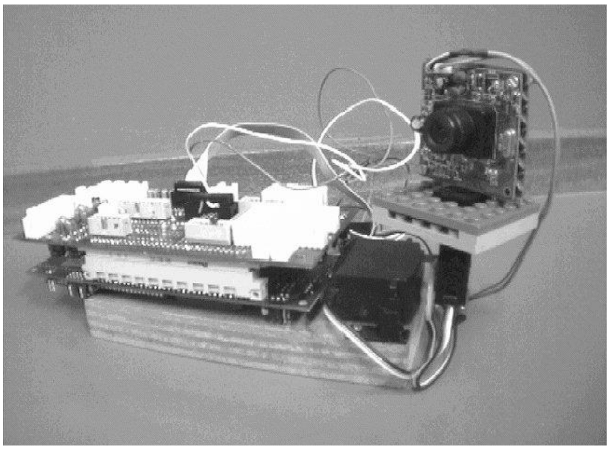

Fig. 3. The "head", an immobile robot with strong vision capabilities. It is immobile and therefore forced to cooperate to receive "food" in form of energy.

quite some vision capabilities. As it is not mobile, the head cannot access the charging-station and it is forced to cooperate. As illustrated in figure 4, the head can track mobile robots and it can perceive so-called pitfalls which are kind of inverse charging stations where the batteries of the mobile robots are partially dis-charged via a resistor. When a mobile robot approaches a pitfall, which he cannot distinguish from a charging station, the head can warn the mobile robot. The mobile robot in exchange can share the benefit of the saved energy with the head.

The head has no other alternative to access food in the form of energy. Therefore, it always issues a warning whenever it perceives a mole close to a pitfall, hoping to receive a reward. The head's perception is influenced by its stamina,

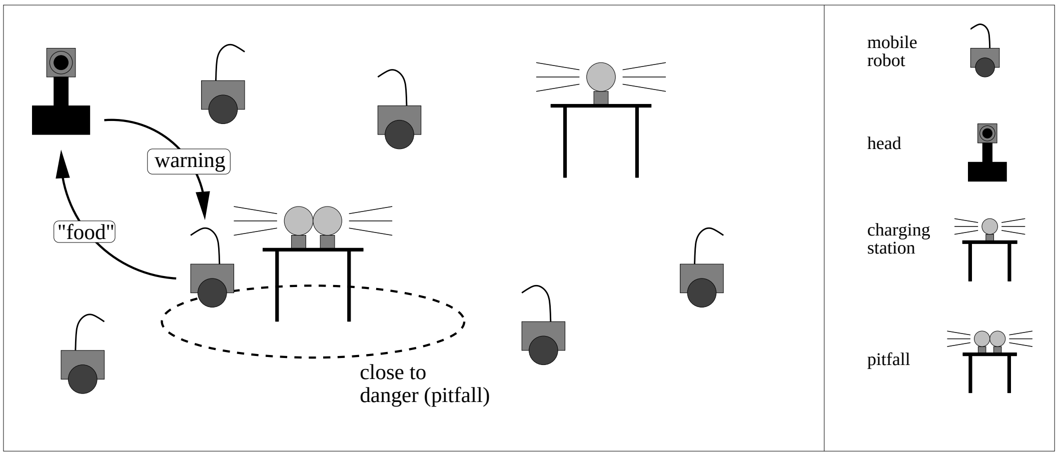

Fig. 4. The extended ecosystem with a head and several moles. Pitfalls in the form of inverse charging-stations can suck energy out of a mole. Unlike moles, a head can distinguish pitfalls and charging-stations, and it can warn a mole when being close to a pitfall. The mole in return feeds a part of its benefit in form of energy to the head.

i.e., the better the head is fed the higher the likelihood that the head perceives a mole when it is close to a pitfall. We assume that

- The head's stamina is directly determined by the sum of the received food.
- The head can not distinguish between moles.

Therefore, a dilemma for the moles arises. Due to the first assumption, it is in the interest of each mole that the head is well fed. Due to the second assumption, there is the temptation to leave the task of actual feeding to others, as the head can not react to the behavior of a single mole and thus punish it when it does not donate energy. In the following section, this dilemma is presented in a formal manner.

3 A Continuous Cooperation N-Player Prisoner's Dilemma

3.1 The iterated Pitfall Game

Let there be N moles and one head. In the rest of this article, we assume that when taking a sufficiently large time-period T, the number of times a mole comes into the dangerous neighborhood of a pitfall is exactly the same for each mole. This means that the stochastics and the asynchronous interactions due to the kind of random walk behavior of moles in the environment are not regarded[3] . The iterated pitfall game proceeds in time steps t, each of length T.

In a time step t, each mole $m_i (1 \leq i \leq N)$ has a gain G_i based on the avoidance of pitfalls due to warnings of the head. As mentioned before, the time period T is sufficiently large that the gain is not stochastically distributed, but it is the same for each mole. So, the gain only depends on the headsight $hs \in [0, 1]$, i.e., the percentage with which the head perceives dangerous situations. Concretely, the gain is the headsight times one hundred energy-units (EU):

$$G_i = hs \cdot 100EU \tag{1}$$

In the beginning of each time step t, each mole m_i invests up to seventy-five energy units to feed the head. This investment I_i is proportional to the continuous cooperation level $co_i \in [0, 1]$ of m_i:

$$I_i = co_i \cdot 75EU \tag{2}$$

As mentioned before, the headsight hs depends on the amount of food the head receives from the moles, i.e., the head is completely fed when it receives the 75 energy units from every mole. Concretely, we define the headsight hs as the averaged sum of cooperation levels in time step t:

$$hs = \sum_{1 \leq i \leq N} co_i / N \tag{3}$$

The pay-off po_i for a mole m_i is the difference between gain and investment:

$$po_i = G_i - I_i = \sum_{1 \leq j \leq N} co_j / N \cdot 100EU - co_i \cdot 75EU \tag{4}$$

A single pitfall game forms an N-player Prisoner's Dilemma with continuous degrees of cooperation:

- In the full cooperation case, i.e., everybody does a full investment, every mole has a pay-off of 25 energy units.
- If all moles defect, i.e., nobody invests, their pay-off is a punishment of 0 energy units.
- Maximum sucking occurs for a mole m_i if it pays the full investment and every other mole does not invest. In this case the pay-off for m_i is $100/N - 75 \leq -25$.

---

[3] This is mainly done for the sake of simplicity, which allows to compare results from simulation with previous work in this field. We also did preliminary experiments including these aspects, which are subject for future publications.

- The maximum temptation pay-off is achieved when one mole m_i invests nothing but all the other moles do full investments. Then the pay-off for m_i is $(N-1)/N \cdot 100 \geq 50$.
- In between these cases, continuous forms of these types of pay-offs occur.

In the following subsection, we discuss the different types of pay-off for different combinations of cooperation levels in a formal manner.

3.2 The Pay-Off Types

Let $\mathcal{C}$ denote the space of possible combinations of cooperation levels, i.e., $\mathcal{C} = [0,1]^N \subset I\!R^N$. Given a group $\mathcal{G}$ of $N \geq 2$ moles as an ordered set $\{m_1, ..., m_N\}$. The state of $\mathcal{G}$ in a single pitfall game can be described as a vector $v_c \in \mathcal{C}$, i.e., the vector of cooperation levels co_i of each mole m_i. The pay-off of a mole m_i is a function $F_p(v_c)$. Let $\bar{co}$ denote the average cooperation level of the group $\mathcal{G}$ in time step t, i.e.:

$$\bar{co} = \sum_{1 \leq i \leq N} co_i / N \tag{5}$$

Due to the nature of the pitfall game, the pay-off for a mole m_i can directly be computed from co_i and $\bar{co}$. Namely, the pay-off function $f_p : [0,1] \times [0,1] \rightarrow I\!R$ is defined as:

$$f_p(co_i, \bar{co}) = co_i \cdot -75EU + \bar{co} \cdot 100EU \tag{6}$$

Now, we can extend the terminology for pay-off values in the standard prisoner's dilemma, with pay-off types for cooperation (C), punishment (P), temptation (T), and sucking (S), as follows:

- Full cooperation as all fully invest: $C_{all} = f_p(1.0, 1.0) = 25.0$
- All punished as nobody invests: $P_{all} = f_p(0.0, 0.0) = 0.0$
- Maximum temptation: $T_{max} = f_p(0.0, \frac{N-1}{N}) \geq 50.0$
- Maximum sucking: $S_{max} = f_p(0.0, \frac{1}{N}) \leq -25.0$

For $co, \bar{co} \neq 0.0, 1.0$, we get the following additional types of pay-offs, the so-called partial temptation, the weak cooperation, the single punishment, and the partial sucking. They are not constants (for a fixed N) like the previous ones, but actual functions in $(co, \bar{co})$. Concretely, they are sub-functions of $f_p(co, \bar{co})$, operating on sub-spaces defined by relations of co in respect to $\bar{co}$ (table 1):

Note that for a fixed average cooperation level $\bar{co}$ and two individual cooperation levels $co' > co''$, it always holds that $f_p(co', \bar{co}) < f_p(co'', \bar{co})$. Therefore

$co < \bar{co}$	$\bar{co} \leq co < 4/3 \cdot \bar{co}$	$co = 4/3 \cdot \bar{co}$	$co > 4/3 \cdot \bar{co}$
$T_{partial}(co, \bar{co})$	$C_{weak}(co, \bar{co})$	$P_{single}(co, \bar{co})$	$S_{partial}(co, \bar{co})$
$\in]0, 100[$	$\in]0, 25[$	$= 0$	$\in]-100, 0[$
partial temptation	weak cooperation	single punish	partial sucking

Table 1
Additional pay-off types in the pitfall game

it holds for an individual player in a single game that:

- The partial temptation pays always better than weak cooperation.
- The partial temptation increases with decreasing individual cooperation.
- The absolute value of partial sucking increases with increasing individual cooperation.

This can also be stated as:

$$\begin{aligned} T_{max} > T_{partial}(.) > C_{all} > C_{weak}(.) > 0.0 \\ P_{single}(.) = P_{all} = 0.0 \\ S_{max} < S_{partial}(.) < 0.0 \end{aligned} \tag{7}$$

The equation 7 illustrates the motivation for the names of the different types of pay-off. The attribute *max* for temptation T and sucking S indicates that these are the maximum absolute values. The *partial* accordingly indicates that these values are only partially reached through the related T or S functions. The attribute *weak* for the cooperation function C relates to the fact that though the player receives a positive pay-off, it is less than in the maximum cooperation case where *all* players fully cooperate. When no player invests, *all* are punished with a Zero pay-off. Whereas in the *single* case, at least the individual player we are looking at gets punished with a Zero pay-off, other players can receive all types of pay-off.

4 Iterations and Strategies

4.1 Continuous 2-Player iPD

[RS98] investigate an iterated 2-player prisoner's dilemma with continuous degrees of cooperation, i.e., where the players can invest a real-numbered value I, corresponding to their degree of cooperation. The investment values of the two players are added, then scaled by a factor $k > 1$, and finally evenly distributed between the two players. So, each player i receives $(I_1 + I_2) \cdot k/2$. The pay-off p_i is accordingly:

$$p_i = \frac{(I_1 + I_2) \cdot k}{2} - I_i \text{ with } i = 1, 2 \tag{8}$$

In the iterated version of this 2-player game, [RS98] investigate following strategies, i.e., schemes for a player to alter its investment $I(t)$ in round t depending on its pay-off $p(t-1)$ and the investment $I'(t-1)$ of the partner in the previous round $t-1$:

Non-altruism (NA) : always completely defect, i.e., $I(t) = 0$

Give-as-good-as-you-get (GGG) : match your partner's last investment, i.e., $I(t) = I'(t-1)$

Short-changer (SC) : subtract a small constant c from your partner's last investment, i.e., $I(t) = I'(t-1) - c$

Raise-the-stakes (RTS) : match your partner's last investment when your previous investment is undercut, increase slightly your own previous investment when it is matched, and increase your own previous investment more than slightly when it is more than matched, i.e.,
- $p(t-1) < 0 : I(t) = I'(t-1)$
- $p(t-1) = 0 : I(t) = I(t-1) + c$
- $p(t-1) > 0 : I(t) = I(t-1) + 2 \cdot c$

Occasional-short-changer (OSC) : a slight variation of RTS, where occasionally the small constant c is subtracted from the RTS-investment

Occasional-cheat (OC) : an other slight variation of RTS, where occasionally nothing is invested

All-or-nothing (AON) : invest nothing when your previous investment is undercut, otherwise invest a constant c', i.e.,
- $p(t-1) \geq 0 : I(t) = c'$
- $p(t-1) < 0 : I(t) = 0$

Anything-will-do (AWD) : always cooperate at a fixed level, i.e., $I(t) = c'$

The main focus of [RS98] is on Raise-The-Stake (RTS). In simulated evolutionary experiments with reproduction rates proportional to pay-offs in pairwise games, they show that

- RTS spreads to fixation when starting with an even mix of strategies
- RTS invades a population of NA, which is also seeded with the other strategies
- RTS resists invasion by the other strategies

They conclude that they show for the first time, up to their knowledge, that "relationships involving increasing investment are robust in the face of cheats and subtle cheats".

As mentioned in the introduction, the N-player prisoner's dilemma is quite different from the 2-player case and it seems to constitute a much harder problem for the evolution of cooperation. We belief that the failure of previous research (up to our knowledge) to come up with a successful strategy in the N-player case is due to the fact that so far no continuous degrees of cooperation were considered for a group of N players.

With continuous degrees of cooperation, it is possible to make slight changes in one's own amount of altruism. Especially, we propose here a strategy based on the *averaged* level of cooperation of the whole group. In the standard binary case of complete cooperation or defection, this value is, though not being really continuous, at least a fractional number. So, it is very difficult to react to subtle tendencies in the group in the standard case; the averaged level of cooperation can take $N+1$ values of the fractions i/N $(0 \leq i \leq N)$, but only two reactions (cooperate or defect) are possible.

At a first glance, readers might object that the average level of cooperation of the whole group is not very suited to base a strategy upon. Naively, every player in the group has to observe the behavior of every other player in the last step to compute the average. For large groups, this seems to be neither feasible nor very intuitive. But this naive view is simply false. Note that the gain G_i of a player as defined in equation 1 depends on exactly the averaged cooperation level. Therefore, this information is presented to each player by default. In the example of the pitfall game, the stamina and the sight of the head fill the role for this task.

Following the hypothesis of the previous subsection that strategies involving slightly increasing investments can be beneficial, we introduce here the strategy of **Justified-Snobism (JS)**, which cooperates slightly more than the average cooperation level of the group in the previous step with non-negative pay-off, and it cooperates exactly at the average cooperation level of the group when a negative pay-off was achieved.

So, JS is at least as cooperative as the average, and sometimes even more cooperative. Thus the motivation for of name of this strategy, as the snobbish belief to be better than the average of the group (in terms of altruism) is somehow justified for players which use this strategy. Given a constant c_{JS} for the additional cooperativeness, JS for a player i is defined as:

Justified-Snobism (JS): (9)

$$po_i(t-1) \geq 0 : co_i(t) = \bar{co}(t-1) + c_{JS}$$
$$po_i(t-1) < 0 : co_i(t) = \bar{co}(t-1)$$

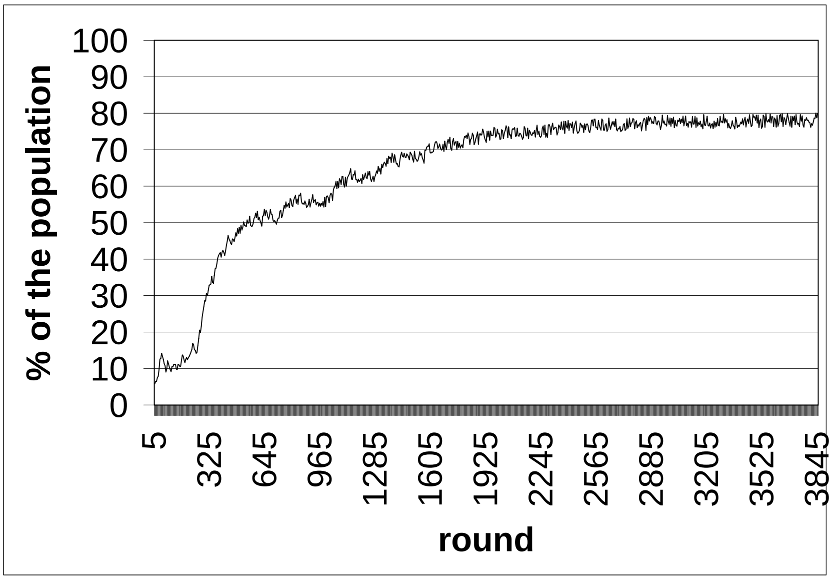

Fig. 5. **Justified Snobism (JS)** spreading in a population which starts with an even mix of strategies.

Furthermore, we define following strategies:

Follow-the-masses (FTM) : match the average cooperation level from the previous round, i.e., $co_i(t) = \bar{co}(t-1)$

Hide-in-the-masses (HIM) : subtract a small constant c from the average cooperation level, i.e., $co_i(t) = \bar{co}(t-1) - c$

Occasional-short-changed-JS (OSC-JS) : a slight variation of JS, where occasionally the small constant c is subtracted from the JS-investment

Occasional-cheating-JS (OC-JS) : an other slight variation of JS, where occasionally nothing is invested

Challenge-the-masses (CTM) : Zero cooperation when the previous average cooperation is below one's one cooperation level, a constant cooperation level c' otherwise, i.e.,

- $co_i(t-1) \geq \bar{co} : co_i(t) = c'$
- $co_i(t-1) < \bar{co} : co_i(t) = 0$

Non-altruism (NA) : always completely defect, i.e., $co_i(t) = 0$

Anything-will-do (AWD) : always cooperate at a fixed level, i.e., $co_i(t) = c'$

5 Experiments and Results

To test Justified-Snobism (JS), several sets of evolutionary experiments in simulations were undertaken, where a population pop of N_{pop} strategies $s_n (1 \leq n \leq N_{pop})$ evolve in iterations of the continuous N-player prisoner's dilemma

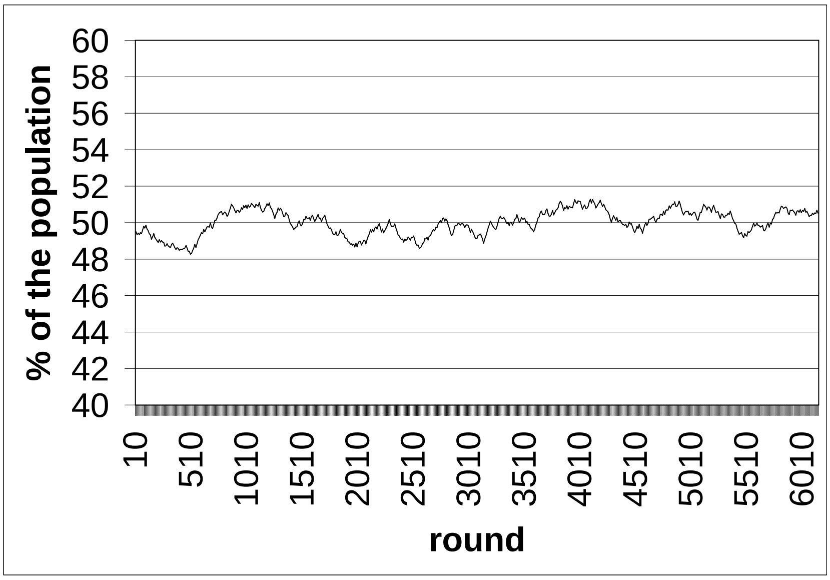

Fig. 6. Resistance of **Justified Snobism (JS)** against invasion of other strategies, which are constantly seeded into the population.

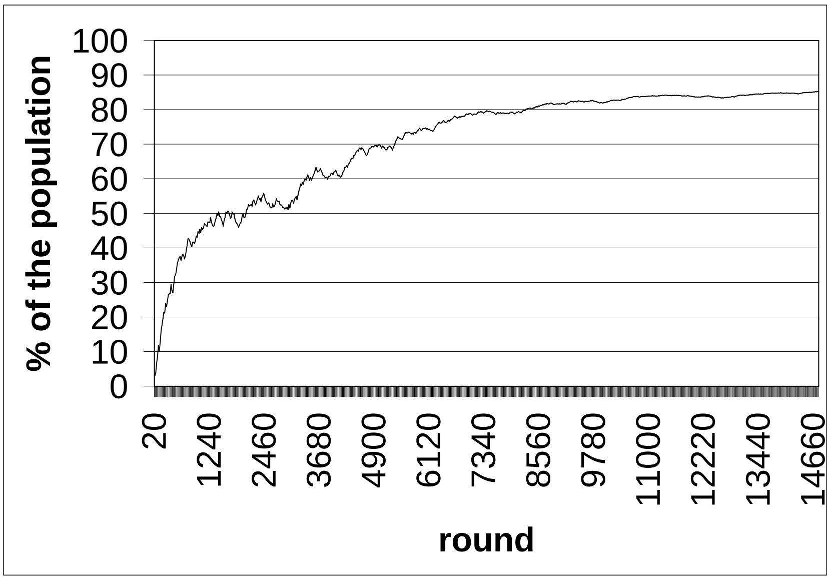

Fig. 7. Invasion of **Justified Snobism (JS)** into a population of Non-Altruistic (NA) strategies, which is also seeded with other strategies.

(CN-PD). The population size of all experiments reported here is set to $N_{pop} = 1000$. The fitness $f(s_n)$ of a strategy s_n is determined by the running average of its pay-offs. The evolution proceeds in rounds r. Given the population $pop(r-1)$ of the previous round, the population $pop(r)$ is determined by

- reproducing each strategy s_n in population $pop(r-1)$ proportionally to its fitness $f(s_n)$
- selecting N_{pop} strategies with roulette-wheel selection for population $pop(r)$

Within a round r, k iterations of the CN-PD are played, with $k = 50$ for the experiments reported here. The two-player strategies presented in subsection 4.1 can participate by selecting a partner, which is fixed for this particular round. If for example strategy s_{724} is GGG and it has the partner s_{64}, then $co_{724}(t) = co_{64}(t-1)$ for $1 < t \leq k$. Each non-trivial strategy[4] is initialized at $t = 0$ with a random value smaller than 1/2. The number of players per dilemma N is set to 20. This is significantly more than in any other previously published research up to our knowledge.

The constant c' for levels of cooperation in the strategies AON, AWD, and CTM is randomly chosen from $]0, 1]$ for each single instance of each strategy at the beginning of each experiment. The small constant c which is added or subtracted to levels of cooperation in RTS, OSC, OC, and SC, is set to 0.01. The constant c_{JS} which is added to levels of cooperation in JS is also set to 0.01.

All experiments done so far indicate that JS is a successful strategy in the iterated continuous N-player prisoner‘s dilemma. This is especially motivated by the following three classes of experiments:

- class 1: JS spreading in a population with an even mix of strategies
- class 2: JS resisting invasion of other strategies
- class 3: JS invading a population of NA, also seeded with other strategies

Results from twenty averaged runs of class 1 experiments are shown in figure 5. The population *pop* at round $r = 0$ and time step $t = 0$ is initialized with an even mix of strategies. In following rounds, JS spreads in the population and becomes dominant.

Figure 6 shows results from twenty averaged runs of class 2 experiments. The population is at the beginning initialized with 70% JS and an even mix of the other strategies for the remaining 30%. In each round, 5% of JS strategies are replaced by randomly chosen other strategies. Nevertheless, the dominant presence of JS in the population does not break down, it even increased in some single experiments.

Last but not least, results from twenty averaged runs of class 2 experiments are shown in figure 7. The initial population consists here of 90% NA and an even mix of the other strategies for the remaining 10%. In these experiments, JS invades NA and becomes dominant.

In addition to these results, it showed that the partner-based strategies seem to be weaker than the average-based strategies in the sense that they tend to

---

[4] Strategies are denoted as trivial when they constantly cooperate at the same level. For example, NA is trivial as it cooperates with $\forall t : co(t) = 0$.

be the first ones to become less present in the population.

6 Conclusion

In this article, a scenario in a robot ecosystem is described in which the origins of cooperation among multiple robots are studied. A so-called "head" can monitor N mobile robots and warn them in dangerous situations caused by so-called pitfalls. The robots in turn "feed" the head with energy. The individual real-valued pay-off for a robot due to warnings depends on the stamina of the head, which in turn depends on the accumulated real-valued energy-donations of all robots. As the head cannot distinguish between robots, it offers its service equally to each robot independent of the robot's donation, causing a dilemma. On the one hand, it is in the interest of each robot that the head is well fed, on the other, this task can be tried to be left to others.

The experimental set-up is formalized as an N-player Prisoner's Dilemma with continuous degrees of cooperation, i.e., real-valued levels of cooperation between Zero and One. This formal model significantly generalizes previous work n the field building on the classic prisoner's dilemma. Our experimental set-up and its formal model are complementary to each other. The formal model allows theoretical work and experiments in simulations, especially in comparison with previous research work on cooperation. The real-world ecosystem allows to investigate aspects which are very difficult and tedious to incorporate into a simulation, especially asynchronous, spatially distributed interactions between robots. The combination of a physical experimental set-up and a thorough theoretical model allows to benefit from both worlds and a cross-fertilization between them. Accordingly, the scenario can be used for a wide body of research. For example, it is also used in experiments on the evolution of trust [Bir00a] and the learning of trust [Bir00b].

In addition to the presentation of the experimental set-up and its formal model, an important basic result from simulations is presented in this article. Namely, it is shown that cooperation can not only emerge in simple two player scenarios, but also in larger groups despite a significant temptation to cheat on others. These results with up to twenty players are based on the novel strategy of Justified-Snobism (JS). JS cooperates slightly more than the average cooperation level of the group of N players if a non-negative pay-off was achieved in the previous iteration, and it cooperates exactly at the previous average cooperation level of the group otherwise. So, JS is at least as cooperative as the average, and it is sometimes even more cooperative than the average. This leads to the name for this strategy as the snobbish belief to be "better" than the average of the group is somehow justified for players which use this strategy.

References

[Axe84] R. Axelrod. *The Evolution of Cooperation.* Basic Books, 1984.

[Bir00a] Andreas Birk. Boosting cooperation by evolving trust. *Applied Artificial Intelligence Journal*, 14(8), September 2000.

[Bir00b] Andreas Birk. Learning to trust. In Falcone, Singh, and Tan, editors, *Deception, Fraud and Trust in Agent Societies*, LNAI. Springer, 2000.

[BKW98] Andreas Birk, Holger Kenn, and Thomas Walle. Robocube: an "universal" "special-purpose" hardware for the robocup small robots league. In *4th International Symposium on Distributed Autonomous Robotic Systems.* Springer, 1998.

[BKW00] Andreas Birk, Holger Kenn, and Thomas Walle. On-board control in the robocup small robots league. *Advanced Robotics Journal*, 14(1):27 – 36, 2000.

[Col82] A.M. Colman. *Game Theory and Experimental Games.* Pergamon Press, 1982.

[DTF98] Alexis Drogoul, Milind Tambe, and Toshio Fukuda, editors. *Collective Robotics.* LNAI 1456. Springer, 1998.

[GHB96] Robert Ghanea-Hercock and David P. Barnes. An evolved fuzzy reactive control system for cooperating autonomous robots. In *From Animals to Animats 4. Proc. of the Fourth International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books, Cambridge, 1996.

[Har68] G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.

[Hof00] Robert Hoffmann. Twenty years on: The evolution of cooperation revisited. *Journal of Artificial Societies and Social Simulation*, 3(2), 2000.

[Mat93] Maja J. Mataric. Designing emergent behaviors: From local interactions to collective intelligence. In *From Animals to Animats 2. Proc. of the Second International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books, Cambridge, 1993.

[McF94] David McFarland. Towards robot cooperation. In Dave Cliff, Philip Husbands, Jean-Arcady Meyer, and Stewart W. Wilson, editors, *From Animals to Animats 3. Proc. of the Third International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books, Cambridge, 1994.

[MH96] Alexandros Moukas and Gillian Hayes. Synthetic robotic language acquisition by observation. In *From Animals to Animats 4. Proc. of the Fourth International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books, Cambridge, 1996.

[RS98] Gilbert Roberts and Thomas N. Sherratt. Development of cooperative relationships through increasing investment. *Nature*, 394 (July):175–179, 1998.

[SB93] L. Steels and R. Brooks, editors. *The "artificial life" route to "artificial intelligence".Building situated embodied agents.* Lawrence Erlbaum Associates, New Haven, 1993.

[Ste94] Luc Steels. A case study in the behavior-oriented design of autonomous agents. In Dave Cliff, Philip Husbands, Jean-Arcady Meyer, and Stewart W. Wilson, editors, *From Animals to Animats 3. Proc. of the Third International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books, Cambridge, 1994.

[Ver96] Dany Vereertbrugghen. Design and implementation of a second generation sensor-motor control unit for mobile robots. Technical Report Thesis, Tweede Licentie Toegepaste Informatica, Vrije Universiteit Brussel, AI-lab, 1996.

[Wil91] Stewart W. Wilson. The animat path to ai. In *From Animals to Animats. Proc. of the First International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books, Cambridge, 1991.

[WM96] Barry Brian Werger and Maja J. Mataric. Robotic "food" chains: Externalization of state and program for minimal-agent foraging. In *From Animals to Animats 4. Proc. of the Fourth International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books, Cambridge, 1996.