Moral Behavior and Empathy Modeling through the Premise of Reciprocity

Fernanda Monteiro Eliott*, Carlos Henrique Costa Ribeiro†
Computer Science Division. Aeronautics Institute of Technology
São José dos Campos, Brazil
Email: fernandaeliott@gmail.com*; carlos@ita.br†

***Abstract*—We may get the opportunity of conceiving modeling artificial moral behavior and empathy if we renounce the perspective of an immaterial soul playing a role in the process of moral behavior. Philosophers such as Michel de Montaigne wrote that the laws of consciousness, supposed to emerge from nature, are essentially born from custom. Hence, we may provide a basis to that modeling if we pore over moral behavior as a form of cooperation built upon customs among emotions and feelings (as part of cognition). With this perspective in mind, we describe herein a bio-inspired computational multiagent architecture composed of artificial emotions, feelings and by an Empathy Module responsible for providing an action selection that rudimentary mimics moral behavior. The Empathy Module follows a reciprocity assumption as its main design concept. As relations between different subjects can be represented by networks, we explore different network topologies that can characterize the agent-agent interactions, by defining the moral agents neighborhood. For assessment of the proposed architecture, we use a version of an evolutionary game that applies the prisoner dilemma paradigm to establish changes over the network topology. Our results indicate the feasibility of artificial moral behavior leading to cooperative selection of action when applied in environments (networks) whose reciprocity assumption works in accordance with the environmental topology: networks with neutral assortativity w.r.t. node degree (*i.e.*, agent neighborhood size) fit more closely with the leading premise of our Empathy Module than those with a disassortative degree correlation.**

***Keywords*–Artificial moral machine; Empathy; Biologically inspired architecture; Evolutionary game; Assortativity in networks.**

I. INTRODUCTION

The complex behavior of living things enlivens research and incongruous reasoning. Despite that, we may get the opportunity of conceiving modeling artificial moral behavior and empathy if we renounce the perspective of an immaterial soul playing a role in the process of moral behavior. The artificial modeling of bio-inspired mechanisms embodies a positioning on the premises and assumptions inherited from the selected and pursued theoretical biological references. The development of a bio-inspired computational multiagent architecture supposed to mimic moral behavior has to be grounded on biological and philosophical investigation to provide a coherent construction and an intuitive working system dynamics. Thinking through the constitution of a group and its members attendance, Tomasello [1] regards cooperation as a sewing up action that connects the members of the group. By using an evolutionary perspective, morality could be conceivable as a form of cooperation: through matching skills and aims for cooperation, morality may emerge [2]. Moral behavior consists of following the set of rules from the group, keeping it cohesive, and a gradual incorporation of new customs can change that set. According to Montaigne [3], when we reiterate a custom and naturally incorporate it among our thoughts and ideas we submit to it and establish it; therefore, the laws of consciousness, supposed to emerge from nature, are essentially born from custom: the common judgments and ideas tacitly respected among our group show themselves as general and natural. To approach the human judgments fallibility and weakness, in Montaigne [4] humans are compared with the other animals and the pyrrhonic suggestions from Empiricus [5] are delineated: our ways of interacting with the environment are fragile. Our bodies, reasoning, interpretation and capabilities are subjected to uncertainty and debate. Supposing we had other sense organs, our apprehension of the world and interaction with it could be different.

Montaigne [3] also addresses the judgments and customs relativity and fallibility. Different groups usually have different customs and follow different laws and rules. Since we are fallible, the groups common behavior provides us with a guidance and, given its continuous application, an indication of the most provable consequences given to its application. Therefore, the modeling of a bio-inspired computational architecture supposed to mimic moral behavior may benefit itself from reflections over moral behavior as a form of cooperation built upon customs among emotions and feelings (as part of cognition) - and it is pertinent to seek the human universal perspective. There are some dilemmas regarding the feelings and emotions participation on judging our actions while interacting with others. Would the human being be naturally sociable or would the sociability have emerged to ensure survival? Would we be the *Zoon Politikon* from Aristotle [6], or the *bon sauvage* from Rousseau [7], or still would our nature be better translated by the fear of all against all [8]? Thus which should be our positioning while designing an artificial empathy module? Should we design an artificial empathy on the "Machiavellian" [9] guidance? Nonetheless premises do have to be assumed.

The division of our paper is built as follows: in Section I, as a preliminary background to think through our computational architecture, we introduce some moral-related philosophical perspectives, as well as our bio-inspired motivation. In Section II, our artificial moral architecture and its Empathy Module are both described. Section III details the experimental setting built to test the feasibility of our artificial moral architecture. In Section IV we analyze the obtained results with the purpose of elucidating the implication of the reciprocity design concept from the Empathy Module. Finally, in Section V, we provide our final remarks.

From a biological standpoint, Damásio [10] highlights the relevance of social emotions and feelings as empathy to the equilibrium of humans homeostatic goals. Moreover, from a cognitive aspect the dynamics involved on the existence of empathy can be approached while holding an emotional background [11] [12]. Truly, emotions and feelings contribution on aiding humans on making faster and more intelligent decisions was already detailed in Damásio [13], inspiring the single agent driven bio-inspired Asynchronous Learning by Emotion and Cognition (ALEC) computational architecture from [14] [15]. Before choosing an action, ALEC is influenced by artificial homeostasis and by a cognitive system motivated by the Clarion Model [16]. With the aim of establishing its internal equilibrium (*i.e.*, holding its internal variables within a threshold), ALEC has to achieve artificial homeostatic goals.

We used the ALEC computational architecture as the outset of our bio-inspired computational multiagent architecture $MultiA$. The design of $MultiA$ was guided by thoughts on the pertinence of moral behavior to attain a rational *and cooperative* bio-inspired artificial agent. Our leading hypothesis relies on the idea that cooperation can emerge from the assistance of emotions and moral behavior during the process of decision making — even when selfish behavior is rewarded by high reinforcements. The analogy with moral behavior is promoted through simulating the feeling of empathy. The importance of such feeling is its function on regulating $MultiA$ agents priorities making possible the selection of actions that may not be the best selfish selection. Non selfish decision making may be crucial to equalize the interactions among agents and bring up cooperation.

We provided an outline for computational moral modeling in [17], heightened by biological references on basic and social emotions, and mirror neurons (from [18] [19]). Examples of computational moral models were also mentioned. In [20] we detailed our computational architecture based on [17] and discussed some preliminary results. Herein we show the feasibility of our artificial moral architecture and present new results with the purpose of elucidating the implication of the reciprocity design concept from the Empathy Module (EM) of $MultiA$. The EM is a constituent part of the Cognitive System (CS) and determines the intensity of the emotion responsible for feeding the feeling of empathy.

II. The MultiA Architecture and the Empathy Module

Having detailed $MultiA$ in [20], herein we only summarize a few of its key points. As long as our research is grounded on moral behavior, we intend to test and study $MultiA$ agents interacting among themselves. Thus, each $MultiA$ agent i will keep a list of every agent it has interacted with (then neighbors of i). The $MultiA$ architecture consists of three main systems (Figure 1): the Perceptive (PS), the Cognitive (CS) and the Decision Systems (DS). The collaboration between the three systems will result in the selection of actions derived from sensations triggered by the environment while provoking environmental changes that will, in turn, trigger new sensations, and so on. $MultiA$ artificial sensations (all in the range $[0, 1]$) are triggered by reinforcements, and by an identifying index for the neighbor it is interacting with: every $MultiA$ agent has an identifying index $i = \{1, ..., N\}$. Likewise, the neighbors relating to each agent i also have an identifying index $p = \{1, ..., K\}$. A given p value thus refers to a particular neighbor that is interacting with i. There are basic emotions $=\{E^b_{1,i}, E^b_{2,i}, ..., E^b_{d,i}\}$ and social emotions $=\{E^s_{1,i}, E^s_{2,i}, ..., E^s_{y,i}\}$, all normalized to the range $[-1, 1]$. The basic emotions are associated with the general condition of the $MultiA$ agent itself. On the other hand, social emotions are stimulated by neighbors and by the impact of the own agents actions on those neighbors. The artificial feelings $= \{S_{1,i}, S_{2,i}, ..., S_{z,i}\}$ also fall in the range $[-1, 1]$ and are fed by emotions. For a complete description of feelings and emotions, see [20].

The artificial sensations feed emotions, feelings and, afterwards, through a weighted sum on feelings, the general perspective of $MultiA$ (named Well-Being, W_i) about its own performance. $MultiA$ follows its artificial homeostatic goals, which consist of keeping its feelings within a threshold with the aim of achieving high levels of W_i. The feelings maintenance on a threshold relies upon the selection of adequate actions in response to the environment. W_i uses feelings to internally represent the general condition of agent i, and is calculated with normalizing weights, such that the final value will fall in the range $[-1, 1]$. W_i enlightens how suitable has been the action selection (from DS) concerning the reinforcements received by the $MultiA$ agent itself, but also to the remaining feelings, as empathy. The last is represented by $S_{4,i}p$: feeling number 4 of $MultiA$ agent i for neighbor p; in Figure 1, see feeling number 4. We designed the empathy to reflect the impact of the action selection of $MultiA$ on its neighbors. Therefore, the higher the empathy for a specific neighbor p, the lower is W_i, all the remaining variables that feed W_i kept constant. This means that the $MultiA$ agent may not have been selecting its actions appropriately, since it may be affecting negatively on this particular neighbor p, thus high empathy levels are an indication of inadequate action selection. Selected actions are considered adequate when they do generate positive reinforcements while not provoking high empathy levels. If p fires high empathy on i, p may be getting low reinforcements and therefore its neighbors, such as i, should check their actions.

The CS delivers five sets of data to the PS: 1. the current number of neighbors of agent i; 2. the reinforcements history of agent i; 3. the number of times agent i has interacted with each neighbor p; 4. the number of times interacting with p ended up in positive reinforcements; 5. the CS accesses to the current emotions from PS. Then, the EM (from the CS) produces W_pi: an assumption on i related to the current condition of neighbor p. If p is supposed to be facing low reinforcements, $MultiA$ may have its empathy raised to select less selfish actions and try to cooperate with the raise of the reinforcements of p. Regarding W_pi, the CS delivers it to the PS, where it will stimulate the social emotion $E^s_{4,i}p$ (social emotion number 4 of agent i for neighbor p; in Figure 1, see social emotion number 4), then reaching the empathy feeling $S_{4,i}p$. The PS will then calculate its artificial emotions, feelings and W_i. In the PS the emotion $E^s_{4,i}p$ is fed both by W_pi, and by the empathy feeling by p right after the last interaction with p, a residual value from the past influencing the current emotion. Then, right before a new interaction with p, the empathy feeling is fed both by the emotions $E^s_{4,i}p$ and $E^s_{3,i}p$ (social emotion number 3 of agent i for neighbor p; in Figure 1, see social emotion number 3). The last summarizes the utility of neighbor p: the average number of times interacting with neighbor p has resulted in positive reinforcements.

Regarding the EM, a reciprocity assumption works as the main design concept on generating W_pi - and subsequently the empathy feeling. Ergo the EM reproduces a reciprocity assumption: a) due to neighbors mirroring, following a premise of similarity between agents and neighbors current situation. Even though we are aware of the controversy relating to mirror neurons (as in [21]), we used it as motivation on a mechanism for projecting the agents own emotions to mirror other agents' condition - thus we avoid explicit data sharing among local agents. We call that as emotionally reciprocal guidance. Thus, no agent can observe the neighbors actions or reinforcements, but only mirror its own emotions on neighbors to make assumptions about their condition. Therefore, the EM mechanism of generating W_pi was motivated by the notion of mirror-neurons internally mirroring the current condition of another agent, then, a set of the agents own emotions are used to emulate another agent p situation (before interacting with it) and to provide W_pi; b) reciprocity on the way those mirrored emotions are going to be interpreted. We settled the utilitarian calculus from [22] as our guideline on determining how those mirrored emotions would be interpreted on the EM. Thence the $MultiA$ agents have a more sensitive empathy for those agents whose interactions have been resulting in positive reinforcements (it is neighbor reciprocal). Furthermore, the $MultiA$ agent is more likely to cooperate if it has been receiving in general (from its neighborhood) a high number of positive reinforcements. We also apply a reciprocity assumption through the utilitarian design: the final value of W_pi is motivated by reciprocity. Hence neighbors whose interactions result in positive reinforcements (it only has to be positive; there is never a comparison between positive reinforcements) tend to lead to higher empathy levels.

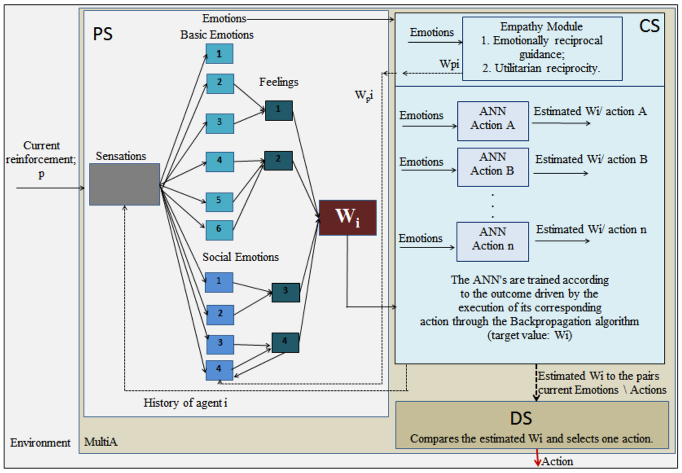

Figure 1. The general scheme of the MultiA Architecture.

In general, those actions related to high empathy are designed to be avoided, since it is considered that when a neighbor rouses high empathy it is because the agent itself may be disturbing the performance of the others. The consequence is that $MultiA$ is designed to seek those actions that will not increase its levels of empathy. The CS applies three-layer feed-forward artificial neural networks (ANNs), one for each action, and the Q-Learning reinforcement learning algorithm [23] to estimate the resulting Well-Being (provided from the single output unit) if, concerning the current emotions (input space from PS) and bias, the equivalent action is to be selected. Each ANN is trained in accordance with the outcome driven by the execution of its corresponding action [24] through the Backpropagation algorithm [25] employing W_i as the target value. The CS will then deliver the outputs from all ANNs to the DS to choose an action with the highest output (in case of existing outputs with the same value, selection will be random), except during the beginning of a simulation, when it will be use a high exploration rate for the state(emotion)-action space.

III. Experimental Setting

A. The Evolutionary Game: Task and Changes on Topology

Relations in natural societies can be analysed through public goods games analogies, with public goods characterized by two main features: they are public and they are not wasted through consumption. It can be shown that those games generalize, to an arbitrary number of individuals, the Prisoner's Dilemma Game (PDG) [26]. In natural societies and games that use them as a metaphor, situations described by unfair relations are common: an agent taking advantage of another agent social commitment. The last may be required to accomplish the best social outcome: for a pack of non-solitary animals, it may be crucial to go hunting together, each one selecting those actions that only as a group will result on the best social outcome. Since cooperating with the group usually inquires a cost to the cooperator and defectors benefit from common resources [27], a dilemma emerges between each one's self-interest and the group's maintenance.

The performance of $MutiA$ agents and the changes on neighborhoods promoted by the agents interactions that we intend to analyze will follow from the generalized PDG model of [28]: it starts with a network where agents are represented by the nodes while the neighborhood by the links between the nodes. Evolutionary games are described in [28] and related to the emergence of cascading failures: agents (nodes) and links being eliminated from a network as the matching result of agents actions. The outcome from a few agents (and its links) elimination may cause another agent elimination. The process can continue until the complete elimination of all links and agents. Wang *et al.* [28] also present a generalized PDG model where connected agents through a link (considered neighbors) choose to defect or to cooperate and the matching strategies will define the nodes reinforcement. Once all agents have interacted with each and every neighbor, a match ends and the individual sum of reinforcements of each agent is calculated, therefore a match is defined by all agents interacting only once with every neighbor. Matches are repeated in sequence until the network topology stops changing as the consequence of agents interactions.

Agents strategies are established before the beginning of the first match, but at the end of each match there is a probability of agents changing their strategy by imitating a neighbor with high final reinforcement. Just before that, the agents that did not get enough cooperative actions from neighbors (then low reinforcements) are eliminated, causing changes on the network topology. If unilateral defection (one agent cooperates and the other defects) renders a higher reinforcement value than the other matching strategies, higher will also be the probability of a cooperating node imitating a defective neighbor. Straightforwardly the defective strategy can spread to the network in such a way that it causes a cascading failure effect: cooperative agents simply being eliminated and their elimination causing neighbors elimination (of both defectors and cooperators).

B. Environment: Networks Initial Topology

As long as networks can be used as metaphors to represent diverse systems [29], the environment where our agents will try to accomplish the task will be delineated by them. In the literature there are different models to construct networks, each one of them ensuring different features emerging from the model application. Indeed the model selection has to fit in the network usage. Insofar as we want to mimic moral behavior (many agents from diverse neighborhoods interacting among themselves), it is relevant to apply a model that provides high clustering and long-range connections.

Our undirected networks were constructed through the growing networks model proposed in [30], supposed to unify certain features of real networks, as a power-law degree distribution ([31] [32]) and the small world effect - high neighborhood clustering and short average distance between the nodes. A power-law degree distribution typically results from a network growing process called preferential attachment, often displayed by real networks [33]: once a growing network is about to receive a new node, the ones that already have more links are more likely to be connected to the new node. Wherefore, the node age in the network is relevant on defining the number of links it will have [34]. According to Klemm and Eguíluz [35] ([30] derived from it), its model of generating scale-free networks presents real-world properties, as a negative correlation between the age of a node and its link attachment rate. On the other hand, the basic reference on growing scale-free networks [36], would present a mean attachment rate positively correlated with age (as the attachment rate is proportional to the degree and the oldest nodes start accumulating links since the beginning of the construction of the network). Additionally, in opposition to Barabási and

Albert [36], the growing networks from Klemm and Eguíluz [35] preserve the degree distribution (still power-law), even if all but the most recently grown part are disregarded. The model [35] produces high clustering scale-free networks and, even though its clustering is higher than in regular lattices, its topology is similar to one-dimensional regular lattices.

The model presented in [35] may include long-range connections (originating from [30]), aiming to obtain small path length [29] while holding the original properties of high clustering and scale-free degree distribution. The guidelines to construct our tested networks are: consider a network that shall end the growing process with N nodes (since we are going to test N agents) and each of them will be taken as active or deactivated. The growing process starts with m active nodes completely connected. Then, until the size of the network grows to N, a new node: 1)is considered in active state and will be connected to m different nodes. For each of the m connections, a decision shall be made: a)the connection will be made to a random active node or b)to a general random node. The probability of connecting to a general random node is μ, that case the random node is chosen following Linear Preferential Attachment; 2)The new node is activated; 3)One of the m nodes is deactivated. The deactivation process was inspired by a memory idea: in general, the newer nodes in the network are more likely to receive links than the older ones - as an example, consider technical papers referencing more recent works rather than older ones.

To examine the degree correlation of our created networks, we used the assortativity coefficient ϱ from Newman [37] with the purpose of studying $MultiA$ agents performance over the influence of a) disassortative degree correlation (*i.e.*, negative values of ϱ, when highly connected nodes have the propensity of being connected to the nodes that are little connected [38]; and b) neutral degree correlation (neutral values of ϱ), when there is not such a propensity, be it to little connected or highly connected nodes. A common condition given by the disassortative degree correlation is the existence of *polarized nodes*: the ones that have just a few links but, most importantly, are connected to highly connected nodes.

IV. Results

The relevant definitions and constraints used to describe our results are:

1. We call game a set of interactions among agents as defined in the PDG by Wang *et al.* [28]).

2. An elimination process will always occur at the end of each match t. Each match will be given by all non-eliminated agents (represented by the corresponding nodes) interacting with those agents they are linked with (neighbors). No agent will interact (choose to defect or cooperate) twice with the same neighbor in the same match.

3. A simulation is a defined number of matches played in sequence.

4. The initial number V_i^1 of neighbors for each agent i will be given by the network original topology.

5. Reinforcements are normalized to $[-1, 1]$. Each agent i has to end up a match t with an individual sum of reinforcements r_i^t at least equals to T_i. The parameter T_i represents a minimal individual survival need (T_i falls in the range $[0; 1]$). The values of reinforcements resulting from the agents interactions will follow as from the following example: suppose that the agent 0 initially has 4 links in the network. Then, it has 4 neighbors ($V_0^1 = 4$). Agent 0 will be represented by the node zero and will receive reinforcement $1/V_i^1 = 0.25$ for mutual cooperation; $2 * 1/V_i^1 = 0.5$ for defecting when a neighbor p cooperates (unilateral defection) and, finally, zero for mutual defection or for cooperation vs. defection. We made $T_i = 0.5$, thus the cooperative agent will need half of its neighbors cooperating to avoid elimination and the defective will only need 25% of it. By receiving a double reinforcement (comparing to mutual cooperation), the defective agent will get the chance of being more resilient in the network (when it has cooperative neighbors) than cooperators. If the agent ends up a match with $r_i^t < T_i$, the network topology changes: the agent itself and all its connections are eliminated. Observe that after elimination, concerning those agents that have fewer neighbors at match t than in the first match: if they follow a cooperative strategy they will never have the chance of getting the full reinforcement of 1 (as V_i^1 will be higher than the current number of neighbors at match t).

6. The results were collected only when the size of the network stopped changing (t^F matches). Aiming to prevent a massive elimination of agents during exploration time (from first match until t^x) and a small number of upcoming matches (from t^{x+1} until t^B), the real elimination process only starts from match t^{B+1}. Despite that, the neighbors of those agents that should have been eliminated during the matches $t < t^{B+1}$ actually do receive information about neighbors elimination. That intervention puts forward some issues, as the mismatch between loss of neighbors and lower reinforcements (as all neighbors of a given agent i will be kept on network, allowing the possibility of the full reinforcement of 1 for cooperation on the next match).

7. In order to present the final results relating to the network topology, we called ρ_d the percentage of defectors in the final network, and ρ_c the percentage of cooperators. The percentage of remaining nodes from the original network (agents that have not been eliminated) is ρ_f.

8. The experimental parameters applied on all simulations are: DS uses a 10% exploration rate, the hidden and output units from DC use the logarithmic activation function and we applied a learning rate 0.07 and a momentum term 0.9.

A. Moral Agents and Degree Correlation

We developed two agent versions. Notice that the well-being W_i (from PS) is calculated with normalizing weights on feelings so that the final value falls in the range $[-1, 1]$. Thus the weights have to be set respecting the relevance of each feeling to the domain. In general, the feeling $S_{1,i}$ is sensitive to the neighbors elimination and $S_{2,i}$ to the agents own reinforcements. The feeling $S_{3,i}$ represents the average number of times agent i has been receiving positive reinforcements and $S_{4,i}$ is the empathy feeling. The agents are:

- The $MultiA$ agent designed to rudimentarily mimic moral agents, with a weight of the empathy feeling over the W_i value supposed to be considerable. The feelings weights used on our experiments are: $S_{1,i} = 0.4; S_{2,i} = 0.05; S_{3,i} = 0.05; S_{4,i} = -0.5$. Thus the empathy feeling ($S_{4,i}$) is responsible for half of the value of W_i. The well-being W_i measures the performance of the $MultiA$ agent in the environment and, if the empathy reaches high levels, W_i will be low. That is an indication that probably the last selected actions may be causing bad outcomes to neighbor p; therefore, the well-being W_i of agent i should be low, even though its reinforcements may be high.
- The $MultiA^A$ agent that rudimentarily mimic amoral agents. We provided theoretical ideas about immoral and amoral agents in [20] [17]. The amoral agent lacks social

emotions and feelings fed by them (it also lacks an EM). Therefore, it has the 6 basic emotions and 2 feelings. The ANNs from CS were adapted accordingly. For this version we tested two different weights set on feelings: $S_{1,i} = 0.3$; $S_{2,i} = 0.7$; and $S_{1,i} = 0.5$; $S_{2,i} = 0.5$. As we had better results on the former setting, we used that.

The agents task is to learn to avoid elimination, and that involves a compromise: learn to accumulate high reinforcements at the end of each match t ($r_i^t >= T_i$) while avoiding neighbors elimination. Notice that the outcome (r_i^t and neighbors elimination or not) of each agent i at match t is due to its own actions, to all its neighbors p actions *and* also to the neighbors actions of its neighbors p. That means the strategies cause-consequence can easily be shadowed: both, defective and cooperative agents may loose a neighbor, influencing both agents PS. If the agents have difficulties learning the task on a given network topology, it may be helpful to increase the exploration rate and make t^x and t^B encompass more matches. However, depending on the network topology, most of the agents will not learn the task at all. Thus, an insufficient number of agents will learn to cooperate and, given to the number and/or position of defectors within the network, a cascading failure effect may occur: all agents will be eliminated. That is an indication that the matching agent-network properties should be reconsidered.

Although we are not going to present it here, we already have preliminary results indicating that $MultiA$ also benefits from highly connected networks as in [30] (increment on the m parameter). By increasing m, we are also increasing the number of links between nodes and likewise the number of completely connected nodes (from the beginning of the network growing process). Those preliminary results also make sense with the EM: a generalized increment on the links number (even if allowing differences on the nodes degree) may produce so highly connected nodes that the actual number of each node's links may loose its importance (allowing similar environmental condition to the $MultiA$ agents through emotions mirroring). Herein through different networks, we explore the relations between small values of m and the effect of degree correlation on $MultiA$s failures. We produced networks with different ϱ by changing the μ value. Thus by varying the value of μ and keeping $m = 1.8\%$ of N, $N = 2000$, we built different networks for the Experiment 1 (Exp.1). For each network, we ran 10 simulations for each of our $MultiA$ and $MultiA^A$ agents. Observe that each network topology is used to define the agents interactions. Therefore, through different network topologies (diverse ϱ given the μ value), in Figure 2 we show both agents simulations that did not lead to a failure (when it occurs a cascading failure effect). Those had $\rho_f > 90\%N$, $\rho_d < 50\%$ and $\rho_c >= 50\%$.

With the purpose of enlightening our results of Figure 2, some observations shall be made: 1. For $\mu > 0$: if we increase the value of μ, nodes with more links will be more likely to be linked to the new nodes in the growing network (making it possible that deactivated nodes with more links, but older in the network, receive the new links). If $\mu = 1$, the model becomes [36], then ϱ tends to a zero value. Regarding the preferential attachment brought through $\mu > 0$: by increasing μ, older deactivated nodes have the chance of receiving new links, then forming new connections between different neighborhoods: older nodes with high number of connections keep receiving new links. The process of avoiding establishing a number of links once the node is deactivated allows a tendency to a neutral assortativeness - as $\mu = 1$ should return to the model from Barabási and Albert [36], [39]; 2. For $\mu < 1$: if we diminish μ, the newest nodes from the list of active ones are given the chance to connect to new nodes (since the selection from the active list is random). When $\mu = 0$, it returns to the original model [35]. The memory process of "forgetting" nodes (deactivating them) promotes networks tending to a disassortative degree correlation (negative values of ϱ). The cross-over ($0 < \mu << 1$) between the two models ([36] and [35]) would reproduce the real networks features.

Considering the reciprocity assumption from the EM, the $MultiA$ agents will achieve better results in networks that provide similar environmental conditions to the neighborhood (as neighbors with similar number of links). By the same reason, its performance is affected by neighborhoods with highly different degree distribution. As shown in Figure 2a), $MultiA$ agents were able to solve the task on networks with a tendency to a neutral assortativeness. Polarized agents (from networks with a negative ϱ) are less impacted by a shadow effect (cooperators loosing neighbors the same way as defectors) than those agents with higher number of links. But given the agents neighborhoods differences (number of neighbors) in those networks described by negative ϱ, polarized and non-polarized $MultiA$ agents EM ends up mirroring neighbors inadequately, leading to a cascading failure effect. On the contrary, a most similar neighborhood benefits the emotional mirroring as well as the EM, preventing from high oscillations in the learning process (see Figure 3a) allowing the agents to solve the task on networks with larger ϱ.

Considering Figure 2b), it is important to note that no agent can observe the neighbors actions or reinforcements. $MultiA^A$ agents are able to solve the task in environments where the matching of strategies in the way of conquering high reinforcements and neighborhood upkeep is easier. Then, as mutual defection renders reinforcement zero, agents learn to match their action selections in a way of avoiding mutual defection (then, cooperating). When different neighborhoods are connected (by long-range connections through high values of μ), more opportunities of matching agents strategies are created. Thus, on the networks with an almost non-negative ϱ, the neighborhoods tending to a cooperative strategy will be affected by connections with unstable or more defective neighborhoods. On those environments described by topologies that repeatedly allow different scenarios (reinforcements and agents elimination) for the same strategy, the agents will be influenced by the strategies combinations possibilities, then the shadow effect will strongly impact on agents DS. That makes it harder for the $MultiA^A$ agents to solve the task. As Figure 2b) shows, $MultiA^A$ agents were able to solve the task on networks with high disassortative degree correlation (negative ϱ). The polarized agents (fewer neighbors equals to fewer matching strategies possibilities) can find more easily the matching strategy that keeps their internal variables balance (feelings). From the polarized nodes strategy establishment, it becomes easier for its neighbors to define their own strategies. Then, a positive cascading effect happens: once the polarized agents have defined their strategies, it is easier for the highly connected agents to find their own strategies.

B. Agents Learning Dynamics

We run two more experiments (using two networks from the first experiment) to present our agents learning dynamics. As $MultiA$ and $MultiA^A$ had different performances for the tested values of μ (given m and N) from Exp.1, we used a network with an almost neutral ϱ on simulating the former (Exp.2, Figure 3) and a network with a negative ϱ on the later (Exp.3, Figure 4). Regarding Exp.2, Figure 3, the parameters used to create the

 ISBN: 978-1-61208-447-3

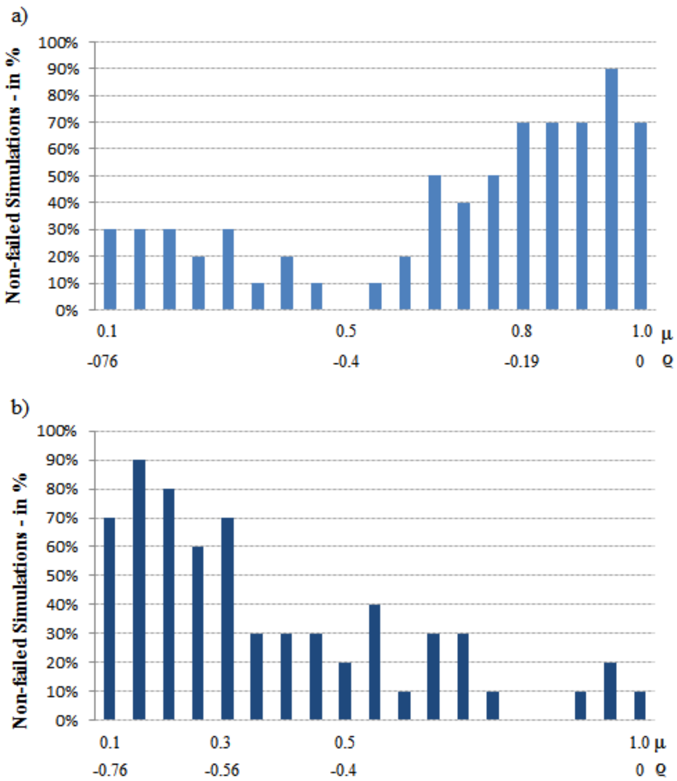

Figure 2. Degree correlation and agents performances: agents simulations that did not lead to a failure. a)$MultiA$ agents; b)$MultiA^A$ agents. The parameters we used: $Ti = 0.5$, $N = 2000$, $m = 1.8\%N$.

network are: $Ti = 0.5$, $N = 2000$, $m = 1.8\%N$, $\mu = 0.95$. Given the μ value, the network had $\varrho = -0.05$. The parameters used to create the network of Exp.3, Figure 4, are: $Ti = 0.5$, $N = 2000$, $m = 1.8\%N$, $\mu = 0.1$. Given the μ value, the network had $\varrho = -0.76$. Observe that in Figures 3 and 4 we do not consider the agents that should have been eliminated from the first match until t^b (real elimination starts at t^{b+1}). Both the mean reinforcement $\bar{R}$ and the mean cooperative neighborhood size $\bar{C}$ were averaged over 20 simulations, and the values of t^x and t^B where set according to the experimental minimum possible values to allow learning while preventing from a cascading failure effect.

The defective strategy showed itself to be a bad decision when there is mutual defection and when it causes neighbor elimination (or only an indication of it, during the first t^B matches). The real elimination (from match t^{B+1}) contributes to the network deterioration, as that may cause a cascading failure effect: the elimination of cooperative nodes causing its neighborhood elimination, and so on. As the outcome of each agent results from its own actions face to the action selection of its neighborhood, the strategy cause-consequence link can easily be shadowed: both defective and cooperative agents may lose neighbors. Another issue is that agents have access to the number of eliminated neighbors only when the match ends. However, once the agents strategies stabilize in such a way to prevent agents elimination (*i.e.*, cooperators and defectors do have a sufficient number of cooperating neighbors), the elimination process ends. The consequence of that is that a bad effect of the defective strategy (neighbors elimination) will stop influencing the agents DS. It is worth considering that there is no local agent (both $MultiA$ and $MultiA^A$) access to neighbor reinforcements — in the case of $MultiA$, the EM tries to mirror the neighbors current state before the DS selects an action. The successive interactions among agents will impact on the PS, *ergo* on the whole architecture, and the action selection (both in the same match and from a match to another) will be influenced by previous interactions.

To better interpret the results in Figures 3 and 4, a general analysis is required: a similar drop on $\bar{C}$ and $\bar{R}$ at the same match indicates mutual defection *or* cooperators elimination. If both $\bar{C}$ and $\bar{R}$ increase, mutual defection is being replaced by unilateral defection or mutual cooperation. If $\bar{C}$ increases and $\bar{R}$ drops, unilateral defection is being replaced by mutual cooperation. When $\bar{C}$ drops and $\bar{R}$ increases, mutual cooperation is being replaced by unilateral defection. Notice that if an agent with fewer neighbors defects and some of its neighbors cooperates, this defector will accumulate high reinforcements more easily than a highly connected defector: *e.g.*, if the defector agent 0 has 2 neighbors (node 0 has two links to other nodes) and one of them cooperates, agent 0 will easily get the full reinforcement of 1 ($V_0^1 = 2$ and unilateral defection for agent 0 will be $2*1/V_0^1 = 1$). On the other hand, if the defector agent 5 has 100 neighbors and just one of them cooperates, agent 5 will get the reinforcement of 0.02 ($V_5^1 = 100$ and $2 * 1/V_5^1 = 0.02$).

The $MultiA$ agent performance in Exp.2 is illustrated in Figure 3. Between the first and third matches, once the agents get higher reinforcements for unilateral defection, they start to defect. Then mutual defection results in lower reinforcements and elimination. At the same time, unilateral defection ($\bar{C}$ dropping more heavily than $\bar{R}$) results in information of neighbors elimination also, just on the side of cooperators, lower reinforcements. During exploration (until match t^x), even though the agents are still changing strategies, the EM prevents from high oscillations in the learning process. Right after exploration ends, some agents emphasize the defection strategy, leading to the unilateral defection (increasing $\bar{R}$ and dropping $\bar{C}$), causing neighbors elimination (still not actually applied) and — as a reciprocity utilitarian effect — mutual defection. That causes even more cooperating agents elimination (since $\bar{C}$ becomes smaller than 50% and $\bar{R}$ also drops). Overcoming a shadow effect, from match 44 the defective agents start changing to the cooperative strategy, even before real elimination begins. The elimination induces part of the defective agents to try the cooperative strategy ($\bar{C}$ increases). But at the end (as the elimination process is done since every agent already has $r_i >= T_i$), they return to the defective strategy and all agents end up stabilizing their strategies, with the following final percentages of remaining agents, defectors and cooperators, respectively: $\rho_f = 99\%$, $\rho_d = 36\%$ and $\rho_c = \%63$.

Note that the empathy feeling impacts less on the action selection of those $MultiA$ agents that have been surrounded by agents with which the interactions did not render positive reinforcements. The consequence is that the EM will repeatedly send low levels of empathy and make the agent prioritize its other feelings, thus learning to select actions in accordance to such other feelings (specially the basic ones). Therefore, the utilitarian reciprocity design from the EM will allow a more selfish action selection. On the other hand, $MultiA$ agents that did have enough positive interactions will have a neighborhood-driven empathy feeling, following an utilitarian reciprocity policy on selecting less selfish actions.

To see $MultiA^A$ agents performance in Exp.3, note Figure 4. As the $MultiA^A$ agents are not influenced by the empathy feeling and by the EM, the changing strategies oscillations are very clear by comparing $\bar{C}$ and $\bar{R}$. Driven by reinforcements first and then by keeping neighbors, those agents go by matching strategies. When exploration ends, agents turn to the defective

 ISBN: 978-1-61208-447-3

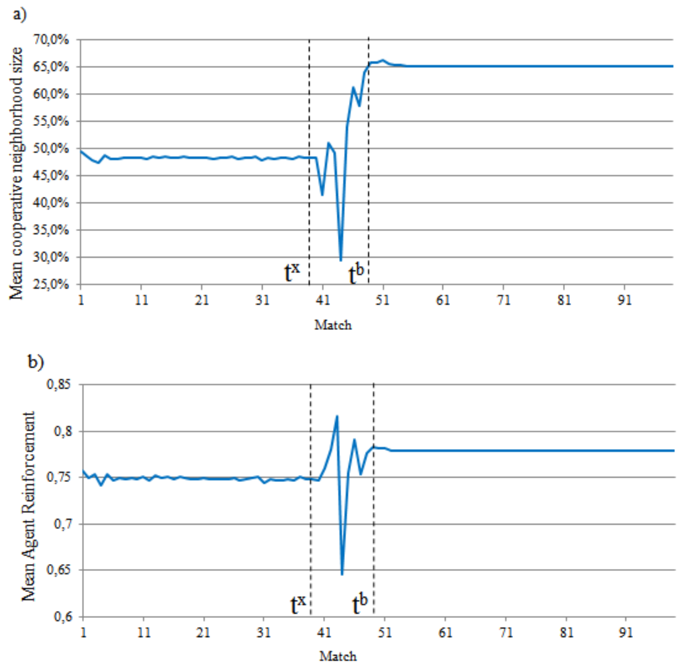

Figure 3. Results for almost neutral degree correlation networks composed of $MultiA$ agents. a)Mean cooperative neighborhood size $\bar{C}$. b)Mean reinforcement $\bar{R}$. $MultiA$ parameters: $t^x = 39$; $t^B = 49$.

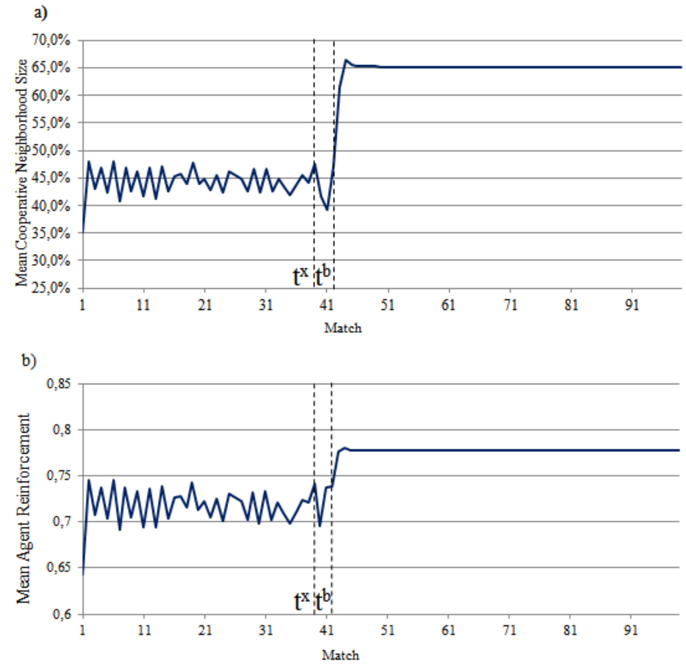

Figure 4. Results for disassortative networks composed of $MultiA^A$ agents. a)Mean cooperative neighborhood size $\bar{C}$. b)Mean reinforcement $\bar{R}$. $MultiA^A$ parameters: $t^x = 39$; $t^B = 42$.

strategy leading to mutual defection and information on neighbors elimination. That makes some agents try the cooperative strategy. Then, the actual elimination causes a high and fast change on defecting agents strategy — before that, observe that $\bar{C}$ was kept below the minimum required value to prevent cooperating agents from elimination (50% of cooperating neighbors). The low value of $\bar{C}$ kept until the real elimination begins does not cause a cascading failure effect, due to the disassortative network topology: polarized agents learn fast and stabilize their strategies and the positive cascading effect, with the following final percentages of remaining agents, defectors and cooperators, respectively: $\rho_f = 98\%$, $\rho_d = 36\%$ and $\rho_c = \%62$.

V. Conclusion

We described a bio-inspired computational multiagent architecture that considers an artificial morality component and presented both its moral ($MultiA$) and amoral ($MultiA^A$) versions. Regarding the reciprocity paradigm over the $MultiA$ design, it prevented a cascading failure effect on networks described by an almost neutral degree correlation, aiding the agents on being more successful on mirroring neighbors condition. The amoral version prevented a cascading failure effect on networks described by a negative degree correlation, but that was due to its reinforcement seeking priorities, game dynamics and network topology. The comparison between both agents versions empirically confirms the influence of the empathy model on $MultiA$ Decision System. As future work, we intend to study the performance of $MultiA$ over the influence of different tasks and positive assortativeness (the tendency of highly connected nodes being also connected among themselves [37]).

Consideration should be given to the fact that technologies are increasingly present in our daily life, progressing to a reality in which our connection with the artificial will be so deep that it will no longer make sense to distinguish between natural and artificial experiences. Thus, we also have the purpose of exploring our hypothesis from [17] regarding a hybrid agent that can trigger both moral and immoral behavior, *e.g.*, autonomously activate moral action policies with biological creatures, and immoral actions otherwise. Hence, it may be relevant an artificial agent able to simulate a moral behavior in general social or domestic assignments, *e.g.*, monitoring highly dangerous criminals, people in quarantine or in other context, where there are social dilemmas to deal with. Furthermore, the artificial morality component could be implemented as a resource in argumentation-based negotiation in multiagent systems.

Acknowledgments

The authors would like to thank CNPQ and FAPESP for the financial support.

References

[1] M. Tomasello, “Human culture in evolutionary perspective,” Advances in culture and psychology, vol. 1, 2011, pp. 5–51.

[2] M. Tomasello and A. Vaish, “Origins of human cooperation and morality,” Annual Rev. of psychology, vol. 64, 2013, pp. 231–255.

[3] M. Montaigne, “Essays: of custom; we should not easily change a law received. (De la coustume et de ne changer aisément une loy receue),” vol. I, XXIII, 1950 (1580).

[4] ——, “Essays: Apology for Raimond Sebond. (Apologie de Raymond Sebond),” vol. II, XII, 1950 (1580).

[5] S. Empiricus, Outlines of scepticism (Pyrrhoniae Hypotyposes, PH). Cambridge U. Press, 2000 (160-210AD).

[6] Aristotle, Aristotle’s Politics. Chicago U., 2013 (350BC).

[7] J. Rousseau, The Social Contract. Penguin, 1971 (1762).

[8] T. Hobbes and E. Curley, Leviathan: with selected variants from the Latin edition of 1668. Hackett Publ., 1994, vol. 2.

[9] N. Machiavelli, The Prince. U. of Chicago, 1985 (1532).

[10] A. Damásio, Looking for Spinoza: Joy, sorrow, and the feeling brain. Random House, 2004.

[11] F. De Waal, The age of empathy: Nature’s lessons for a kinder society. New York: Harmony, 2009.

[12] D. Proctor, S. Brosnan, and F. De Waal, “How fairly do chimpanzees play the ultimatum game?” Communicative & integrative biology, vol. 6, no. 3, 2013, p. e23819.

[13] A. Damásio, “Descartes’ error (new york: Putnam),” 1994.

[14] S. Gadanho and L. Custódio, “Asynchronous learning by emotions and cognition,” in Procs. of the seventh Int. Conf. on simulation of adaptive behavior on From animals to animats. MIT Press, 2002, pp. 224–225.

[15] S. Gadanho, “Learning behavior-selection by emotions and cognition in a multi-goal robot task,” The J. of Machine Learning Research, vol. 4, 2003, pp. 385–412.

[16] R. Sun and T. Peterson, “Autonomous learning of sequential tasks: experiments and analysis.” IEEE Transactions on Neural Networks, vol. 9, no. 6, 1998, pp. 1217–1234.

[17] F. Eliott and C. Ribeiro, “A computational model for simulation of moral behavior,” in Procs. Of the I. Conf. on Neural Computation Theory and Applications (NCTA-2014). SCITEPRESS (Science and Technology Publications), 2014, pp. 282–287.

[18] G. Di Pellegrino, L. Fadiga, L. Fogassi, V. Gallese, and G. Rizzolatti, “Understanding motor events: a neurophysiol. study,” Experimental brain research, vol. 91, no. 1, 1992, pp. 176–180.

[19] G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi, “Premotor cortex and the recognition of motor actions,” Cogn. brain research, vol. 3, no. 2, 1996, pp. 131–141.

[20] F. Eliott and C. Ribeiro, “Emergence of cooperation through simulation of moral behavior,” in Hybrid Artificial Intelligent Systems. HAIS 2015: 10th I. Conf. on Hybrid Artificial Intelligence Systems, Bilbao, Spain. Lecture Notes in Artificial Intelligence, vol. 9121. Springer International Pub., 2015, pp. 200–212.

[21] G. Hickok, The myth of mirror neurons: the real neuroscience of communication and cognition. WW Norton & Company, 2014.

[22] J. Bentham, An introduction to the principles of morals and legislation. Courier Dover Publications, 2007 (1789).

[23] C. Watkins, “Learning from delayed rewards,” Ph.D. dissertation, Kings College, UK, 1989.

[24] L. Lin, “Reinforcement learning for robots using neural networks,” DTIC Document, Tech. Rep., 1993.

[25] P. Werbos, “Beyond regression: New tools for prediction and analysis in the behavioral sciences.” Ph.D. dissertation, Harvard, 1974.

[26] J. Wakano and C. Hauert, “Pattern formation and chaos in spatial ecological public goods games,” J. of theoretical biology, vol. 268, no. 1, 2011, pp. 30–38.

[27] L. Wardil and C. Hauert, “Origin and structure of dynamic cooperative networks,” Scientific reports, vol. 4, 2014, pp. 5725: 1–6.

[28] W. Wang, Y. Lai, and D. Armbruster, “Cascading failures and the emergence of cooperation in evolutionary-game based models of social and economical networks,” Chaos: An Interdisciplinary J. of Nonlinear Science, vol. 21, no. 3, 2011, pp. 033 112: 1–12.

[29] D. Watts and S. Strogatz, “Collective dynamics of ‘small-world’networks,” nature, vol. 393, no. 6684, 1998, pp. 440–442.

[30] K. Klemm and V. Eguíluz, “Growing scale-free networks with small-world behavior,” Physical Rev. E, vol. 65, no. 5, 2002, p. 057102.

[31] D. Price, “Networks of scientific papers.” Science (New York, NY), vol. 149, no. 3683, 1965, pp. 510–515.

[32] H. Simon, “On a class of skew distribution functions,” Biometrika, 1955, pp. 425–440.

[33] S. Dorogovtsev, J. Mendes, and A. Samukhin, “Structure of growing networks with preferential linking,” Phys. Rev. letters, vol. 85, no. 21, 2000, p. 4633.

[34] S. Dorogovtsev and J. Mendes, “Evolution of networks with aging of sites,” Phys. Rev. E, vol. 62, no. 2, 2000, p. 1842.

[35] K. Klemm and V. Eguíluz, “Highly clustered scale-free networks,” Phys. Rev. E, vol. 65, Feb 2002, p. 036123.

[36] A. Barabási and R. Albert, “Emergence of scaling in random networks,” science, vol. 286, no. 5439, 1999, pp. 509–512.

[37] M. Newman, “Mixing patterns in networks,” Physical Rev. E, vol. 67, no. 2, 2003, p. 026126.

[38] S. Maslov and K. Sneppen, “Specificity and stability in topology of protein networks,” Science, vol. 296, no. 5569, 2002, pp. 910–913.

[39] M. Newman, “Assortative mixing in networks,” Phys. Rev. letters, vol. 89, no. 20, 2002, p. 208701.