# Interaction-based Group Identity Detection
# via Reinforcement Learning and Artificial Evolution

Corrado Grappiolo, Julian Togelius
Center for Computer Games Research
IT University of Copenhagen, Denmark
{cogr, juto}@itu.dk

Georgios N. Yannakakis
Institute of Digital Games
University of Malta, Malta
georgios.yannakakis@um.edu.mt

## ABSTRACT

We present a computational framework capable of inferring the existence of group identities, built upon social networks of reciprocal friendship, in Complex Adaptive Artificial Societies (CAAS) by solely observing the flow of interactions occurring among the agents. Our modelling framework infers the group identities by following two steps: first, it aims to learn the ongoing levels of cooperation among the agents and, second, it applies evolutionary computation, based on the learned cooperation values, to partition the agents into groups and assign group identities to the agents.

Experimental investigations, based on CAAS of agents who interact with each other by means of the Ultimatum (or Bargain) Social Dilemma Game, show that a cooperation learning phase, based on Reinforcement Learning, can provide highly promising results for minimising the mismatch between the existing and the inferred group identities. The proposed method appears to be robust independently of the size and the ongoing social dynamics of the societies.

## Categories and Subject Descriptors

I.2.m [**Computing Methodologies**]: Artificial Intelligence—*Miscellaneous*

## General Terms

Algorithms

## Keywords

Group Identity Detection; Adaptive Artificial Societies; Ultimatum Game; Reinforcement Learning; Evolutionary Computation.

## 1. INTRODUCTION

A Complex Adaptive System (CAS) [18] is a set of interconnected parts or units which interact repeatedly with

each other and are equipped with self-organisation properties which allow them to form global patterns. These patterns hold a higher level of complexity than the one inside each single unit. Moreover, some CAS have a recurrent structure, meaning that the global patterns influence the behaviour and the self-organisation properties of the parts. As a consequence, the understanding of how the global structures *emerge* from the interactions of the units cannot be obtained by merely summing the properties of each single part [18]. In fact, within CAS research, the formation of global patterns is also referred to as *Emergence of Complexity*; this clearly states that the modelling of the transition from the local perspective — i.e. the units and their interactions — to the global perspective — i.e. the *emergent* complex patterns — is far from being trivial, due to many factors including the number of units composing the CAS, the volume of interactions, the units' own behavioural mechanisms (which could constitute a CAS itself), and the noise present in the environment.
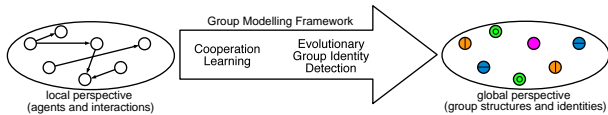
As an example of a CAS, consider human societies [2]: the interactions among the individuals together with the individuals' inner characteristics lead to social phenomena such as friendship [31] and culture [27]. These, in turn, clearly have an impact on how the individuals interact with each other. For instance, they may lead to the spontaneous formation of groups. These groups may lead the formation of group identities — i.e. the ability of the individuals to identify themselves as belonging to a group [4]. Ultimately, such group identities have an influence on the level of cooperation or altruism [11] and collaboration (i.e. reciprocal altruism [18]) of the interactions.

The research presented in this paper aims to define a computational framework — hereafter called Group Modelling (GM) Framework — capable of inferring the presence of group structures, built upon social networks of reciprocal friendship, and assign relative group identities, to Complex Adaptive Artificial Societies (CAAS). The agents forming the CAAS manifest relational-based social preferences [24] when interacting with each other by means of the Ultimatum (or Bargain) Social Dilemma Game [15]. The friendship social networks are generated at a global level and are subsequently communicated to the agents, which *react* to them and adapt their behaviours. In other words, the CAS' phenomena under investigation define the mechanism by which the global structures influence the agents.

Our modelling approach leverages on the concept that individuals belonging to the same group (*in-group*) tend to be more cooperative with each other than when they interact

**Figure 1: The structure of our Group Modelling Framework.**

with individuals belonging to other groups (*out-group*) [4]. The framework is composed of two interconnected modules, as depicted in Figure 1: the first module, called Cooperation Learning Module (CL), is in charge of learning the ongoing levels of cooperation existing among the artificial agents; the second module, called Group Identity Detection Module (GID), partitions the agents into community structure-based groups [19], based on the learnt cooperation values, and assigns a group identity to each one of the identified structures. The main focus of this paper is the investigation of appropriate and efficient techniques to be used for the CL module, being its computation the fundamental feature for the correct computation of the GID module. In particular, we consider three different learning rules for the CL module: (1) Reinforcement Learning's (RL) $\alpha$-constant Monte Carlo ($\alpha$-cMC) update rule, (2) Temporal-Difference (TD) update rule [28], and (3) pheromone evaporation update rule used in Ant Colony Optimisation [5]. For the GID module we rely on a Evolutionary Algorithm (EA) following the principles introduced in [9].

In this paper we conduct experimental studies on two society sizes, i.e. 20 and 50 agents. The results obtained showcase that $\alpha$-cMC, combined with the stochastic search GID module, lead to the highest performing outcomes (i.e. lowest misclassification error between the inferred and the *true* group identities) across the two population sizes and against the ongoing dynamics of the adaptive agents.

To the best of the authors' knowledge, there has not been any attempt to infer the existence of group structures in Complex Adaptive Artificial Societies (CAAS) of agents capable of manifesting relational-based social preferences. The main finding of this paper is the demonstration that RL is a sound technique for cooperation learning. Moreover, we also highlight that its combination with the EA-based GID module — that is a framework which had already proven its efficacy with respect to the detection of group identities in Complex Static Artificial Societies [9] — also offers reliable methods for the detection of group identities in CAAS.

Our modelling framework has several possible real-life applications. It can be used as a data mining tool not only for the detection of community structures in given networks — i.e. the sole application of GM's GID module — [10]; it can also be used to infer underlying complex social networks of users, in complex interactive environments e.g. multiplayer collaborative games [30], indirectly and in real-time by means of gameplay data analysis; as a consequence, our framework can easily be used within — and as a further extension of — Experience-Driven Procedural Content Generation [33], for the creation of adaptive multiplayer games [34]. Furthermore, providing that the interaction protocol is kept the same across human-based (*in-vivo*), and agent-base scenarios (*in-vitro*), our framework can be used to compare the two emergent global patterns in order to further increase the level of realism of the in-vitro simulations, and hence to contribute to the research field of evolution of cooperation [12].

## 2. RELATED WORK

There are a number of studies investigating groups of agents and their behaviours in artificial societies. Nowak et al. [25], among others, focus on the evolution of collaboration by evolving the policies of artificial agents. The approach, however, neglects the impact of collaboration on group formation. Similar research was conducted by Hammond and Axelrod [13], though it was focused on the evolution of ethnocentrism. Among the studies on collective behaviour, Lerman and Galstyan [20] create mathematical models (through differential equations) of the collective behaviour of simple multi-agent systems, such as social insects. Their approach differs from ours in that they aim to build a generic model of an agent, based on observations, and then devise a mathematical model of it. Martinez et al. [23] investigated the use of rule-learning algorithms to predict group behaviours in artificial societies. Their method is based on historical data and even though we share a common goal (i.e. modelling of group dynamics), that study does not aim to model group identities in real time. The use of social networks in Agent-Based Modelling is not novel [14]. For instance, Xianyu [32] utilises spatial networks in order to regulate the interactions among the agents and allow them to manifest inequity aversion-based social preferences [15]. Similar approaches to the use of complex networks can also be seen, among others, in the work of Liu et al. [21]. Our work differs in that our agents are spatially close to each other and interact with any of the other agents in the society; moreover, we use the social network to represent the agent's social relationships, i.e. friendship. Evolutionary algorithms (EAs) have been used for the detection of community structures, such as in the studies on Pizzuti [26] and Tasgin [29], and Liu et al.[22], who propose multi-objective evolutionary computation for the detection of overlapping community structures. Nevertheless, their work did not examine directed weighted networks. Farmer and Fotheringham [6], on the other hand, aim to identify community structures in such complex networks. While we are inspired by that study, we use principles of evolutionary computation (instead of spectral partitioning) for identifying group structures.

## 3. INTERACTION PROTOCOL

Our choice to use the Ultimatum (or Bargain) Social Dilemma Game (UG) for our research is motivated by its extensive application in many akin studies, among which we remark those which put an emphasis on social preferences (see [7, 8, 15, 24, 32] among others). In the simplest UG version [15], an individual, called the provider $p$, receives an endowment $e$. $p$ has the duty to bargain $e$ with another individual, called the receiver, $r$, by making an offer $0 \leq o \leq e$. $r$ has the possibility to either accept the offer (in which case $r$ will gain $o$ and $p$ will gain $e - o$) or dismiss it (in which case both parts will gain zero).

In our study, the UG is implemented as follows: within a society $\mathcal{S} = \{a_0, a_1, \ldots, a_{n-1}\}$ of $n$ agents, for each iteration (Episode) $t$, each agent $a_i$ interacts with the remaining $n-1$ agents by playing two rounds of the UG: one in which $a_i$ plays the role of the provider agent, and one in which it plays the role of the receiver agent. We will hereafter refer

to the offer made by $a_i$ to $a_j$ as $o_{i,j}$, whilst the decision made by $a_j$ with respect to $o_{i,j}$ will be $d_{j,i} = 1$ for "accept" and $d_{j,i} = 0$ for "dismiss".

## 4. GLOBAL SOCIETAL PATTERNS

The complex global pattern under investigation is the existence of group structures, based on the notion of reciprocal friendship, perceived by the adaptive agents of the society at any Episode $t$ of our experiments. We define $\mathcal{FN}_{rec}$ as the society's well-structured reciprocal friendship network:

$$\mathcal{FN}_{rec} = (\mathcal{S}, \mathcal{F}_{rec}) \tag{1}$$

where the $n$ agents of $\mathcal{S}$ are the nodes of the network, and $\mathcal{F}_{rec} = \{(a_i, a_j), (a_j, a_k), \ldots\}$ represents the set of reciprocal friendship edges (or links) among the agents. For each experiment, $\mathcal{F}_{rec}$ is initialised by containing only one friendship link between two agents $a_i$ and $a_j$ uniformly sampled within $\mathcal{S}$. The evolution of friendship is implemented as follows: at regular intervals of $T$ Episodes, a new edge is added to $\mathcal{F}_{rec}$ by applying the well-established Barabási-Albert algorithm for the creation of scale-free networks [1]. When all agents have at least one friendship link, the process ends.

As soon as a new edge is added to $\mathcal{F}_{rec}$, the two involved agents are notified, their behaviour is affected (i.e. they start behaving more altruistically with each other), and the social perception towards each other moves towards friendship (see Section 5). However, the agent's adaptation is not immediate, meaning that at each episode $t$, within $\mathcal{S}$, there exists the following *actual* friendship network:

$$FN_{direct} = (\mathcal{S}, F_{direct})$$

which differs from $\mathcal{FN}_{rec}$ by $F_{direct} = \{(a_i, a_j), (a_j, a_k), \ldots\}$ which is the set of directed friendship links among the agents in $\mathcal{S}$. For instance, $(a_i, a_j) \in F_{direct}$ represents the perception of friendship of agent $a_i$ towards $a_j$ only. $FN_{direct}$ is built at the end of each episode $t$ by querying each agent about their social perception of friendship towards each other agent in the society. The *actual* reciprocal friendship matrix, $FN_{rec}$, approximation of $\mathcal{FN}_{rec}$, is obtained from $FN_{direct}$ by extracting the $(a_i, a_j)$ and $(a_j, a_i)$ edges from $F_{direct}$):

$$FN_{rec}(\mathcal{S}, F_{rec}) \subseteq FN_{direct} \tag{2}$$

Subsequently, $FN_{rec}$ is partitioned into community structures by means of the well-established Clauset-Newman-Moore algorithm [3]. The number of detected communities, together with the number of agents with no friendship links, will constitute the existing number of group structures. The agents belonging to the same community structure will be assigned to the same group identity, whilst the agents with no friends will have their own unique group identity. As an illustrative example, Figure 4(a) depicts the scale-free $FN_{rec}$ network of $n = 50$ agents and the 49 edges obtained at the end of one experimental run. The agents' group identities are represented by different colours summing up to six group structures in that example.

## 5. SOCIALLY-DRIVEN AGENT POLICIES

We define a socially-driven agent $a_i \in \mathcal{S}$ which is capable of manifesting different levels of altruistic behaviours —

when playing the provider role in the UG [24] — depending on which other agent $a_j \in \mathcal{S}$ it interacts with.

Each agent considers each other agent as either a stranger (S), acquaintance (A), or friend (F). This is implemented by allowing each agent to keep the following vector of social perceptions:

$$\vec{s}_i = (s_{i,0}, \ldots s_{i,i-1}, s_{i,i+1}, \ldots s_{i,n-1}) \tag{3}$$

where $s_{i,j} \in [0, 1]$ represents $a_i$'s social perception towards $a_j$. The classification of the perceptions into S, A and F is done as follows:

$$F_{direct}(a_i, a_j) = \begin{cases} S & \text{if } 0 \leq s_{i,j} < \tau_{i,A} \\ A & \text{if } \tau_{i,A} \leq s_{i,j} < \tau_{i,F} \\ F & \text{if } \tau_{i,F} \leq s_{i,j} \leq 1 \end{cases} \tag{4}$$

where $\tau_{i,A}$ and $\tau_{i,F}$, $0 < \tau_A < \tau_{i,F} < 1$, represent, respectively $a_i$'s acquaintanceship and friendship thresholds. Each agent's $\vec{s}_i$ vector is initialised with values uniformly sampled within the $[0, \tau_{i,A}]$ interval. Only the two agents $a_i$ and $a_j$, which are chosen to initialise $\mathcal{F}_{rec}$, have the values for $s_{i,j}$ and $s_{j,i}$ initialised by uniform sampling within the $[\tau_{i,F}, 1]$ and $[\tau_{j,F}, 1]$ intervals, respectively.

### 5.1 Offer and Decision Making Policy

Each agent $a_i$ maintains a parameter, $\hat{o}_i \in [0, 1]$, which represents the (normalised) maximum portion of $e$ it is willing to share with any other agents in $\mathcal{S}$. At each Episode $t$, and for each agent $a_j$, $a_i$ will formulate the offer $o_{i,j}$ as follows:

$$o_{i,j} = \begin{cases} e\ \hat{o}_i\ s_{i,j} + \delta_{i,o} & \text{if } u \leq 0.5 \\ e\ \hat{o}_i\ s_{i,j} - \delta_{i,o} & \text{otherwise} \end{cases} \tag{5}$$

where $\delta_{i,o}$ is an offset representing the noise in the environment and $u$ is uniformly sampled. Agent $a_i$ will decide whether to accept or dismiss $a_j$'s offer as follows:

$$d_{i,j} = \begin{cases} 1 & \text{if } o_{j,i} \geq e\ s_{i,j} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

### 5.2 Adaptation Mechanism

Eq. (5) states that, on average, the higher $s_{i,j}$ is, the higher $o_{i,j}$ will be. As long as $a_i$ has no friendship link, no agent adaptation occurs. However, when $a_i$ is notified to become friend with $a_j$ (i.e. $(a_i, a_j) \in \mathcal{F}_{rec}$) starting from episode $t + 1$, and for each remaining episode, agent $a_i$ will adapt its social perception towards $a_j$ as follows:

$$s_{i,j} \leftarrow \min(s_{i,j} + \delta_{i,s}, 1) \tag{7}$$

where $\delta_{i,s}$ is an offset regulating the adaptation step. At the beginning of each experimental run, each agent generates its values for $params = \{\hat{o}, \delta_o, \delta_s, \tau_A, \tau_F\}$ by uniformly sampling them within their own intervals $[x_{min}, x_{max}]$, $0 < x_{min} < x_{max} < 1, \forall x \in params$. Clearly, depending on the values $\delta_{i,s}$, $\delta_{j,s}$, $\tau_{i,F}$ and $\tau_{j,F}$, the agents will move towards the perception of friendship in a more-or-less delayed manner with respect to the representation kept by $\mathcal{FN}_{rec}$. Moreover, the different $\hat{o}_i$ boundaries will allow for the generation of different levels of altruism within the society.

# 6. GROUP MODELLING FRAMEWORK

As already mentioned, the Group Modelling (GM) framework proposed is composed of two pipelined modules: the first one, called Cooperation Learning Module (CL), aims to learn the ongoing levels of cooperation existing among the agents by leveraging on the concept that individuals belonging to the same group (*in-group*) tend to cooperate more with each other than when they interact with individuals belonging to other groups (*out-group*) [4]; the second module, called Group Identity Detection Module (GID), aims to partition the agents in order to have high in-group and low out-group cooperation [4, 6, 19].

This study focuses on evaluating three different techniques for CL, two of them based on reinforcement learning [28] (see Subsection 6.1) and one inspired by the concept of pheromone evaporation adopted in Ant Colony Optimisation (ACO) [5] (see Subsection 6.2). The result of CL is a cooperation matrix, $\mathcal{C} = n \times n$, where $\mathcal{C}(i,j)$ represents the cooperation value of agent $a_i$ towards agent $a_j$ updated at the end of episode $t$. Subsequently, $\mathcal{C}$ is processed by the GID module, which utilises an evolutionary algorithm in order to partition the society into community structures and assign group identities to the agents [9] (see Subsection 6.3).

## 6.1 Cooperation Learning via Reinforcement Learning

The assumption made by this approach is that the agents have the Markov property [28], that is, they perform *actions* based on their current internal *state*, in order to reach an envisaged *goal* state. Therefore, CL aims to learn the Society's *true state* — represented by the set of the $n$ vectors $\{\vec{s}_i\}$; see Eq. (3) — by processing the *actions* of the agents (see Eq. (5) and Eq. (6)) which represent $\mathcal{S}$'s *observations* about the *true state*. CL assumes that the *goal state* of the society is to have as many friendship links as possible. Assuming that friendship influences altruism or cooperation [11, 24], CL's reward function is defined so that observed altruistic behaviours would lead to high reward values.

Two are the reward functions used and examined in this study, namely the probabilistic Agent Preference ($\mathcal{I}_{\mathrm{AP}}$) and the Entropy-Agent Preference ($\mathcal{I}_{\mathrm{HAP}}$) Interaction Classifier functions [11], hereby described. Let

$$O_i = \{o_{i,0}, \ldots o_{i,i-1}, o_{i,i+1}, \ldots o_{i,n-1}\} \qquad (8)$$

define the whole set of offers made by $a_i$ at Episode $t$; $p_{\mathrm{AP}}(a_i, a_j)$ calculates the probability that an agent $a_i$ is *in-group* with $a_j$ as follows:

$$p_{\mathrm{AP}}(a_i, a_j) = \frac{o_{i,j} - \min(O_i)}{\max(O_i) - \min(O_i)} \qquad (9)$$

whilst $p_{HAP}(a_i, a_j)$ is obtained as the linear combination of $p_{AP}$ and the normalised entropy [11] of the offers:

$$p_{HAP}(a_i, a_j) = -\frac{p_{AP}(a_i, a_j)}{ln(n-1)} \sum_{j=1, j \neq i}^{n} \frac{o_{i,j}}{O_i} ln\left(\frac{o_{i,j}}{O_i}\right) \qquad (10)$$

The two reward functions return either *in-group* ($\mathcal{I} = 1$) or *out-group* ($\mathcal{I} = 0$) as follows:

$$\mathcal{I}_{\mathrm{AP}}(i,j) = \begin{cases} 1 & \text{if } u \leq p_{\mathrm{AP}}(a_i, a_j) \\ 0 & \text{otherwise} \end{cases} \qquad (11)$$

$$\mathcal{I}_{\mathrm{HAP}}(i,j) = \begin{cases} 1 & \text{if } u \leq p_{\mathrm{HAP}}(i,j) \\ 0 & \text{otherwise} \end{cases} \qquad (12)$$

where $u$ is sampled uniformly. We consider two learning rules in this study: the $\alpha$-constant Monte Carlo ($\alpha$-cMC) update rule for non-stationary environments [28]:

$$\mathcal{C}(i,j) \leftarrow \mathcal{C}(i,j) + \alpha\left[\mathcal{I}_{\mathrm{AP}}(i,j) - \mathcal{C}(i,j)\right] \qquad (13)$$

where $\alpha$ is a parameter regulating the learning step, and the Temporal-Difference learning's TD(0) rule [28]:

$$\mathcal{C}(i,j) \leftarrow \mathcal{C}(i,j) + \beta\left[\mathcal{I}_{\mathrm{AP}}(i,j) + \gamma\,\mathcal{C}(j,i) - \mathcal{C}(i,j)\right] \qquad (14)$$

where $\beta$ is a parameter regulating the learning step and $\gamma$ is the discount factor giving more/less importance of the the reciprocal cooperation value, $\mathcal{C}(j,i)$, for the update of $\mathcal{C}(i,j)$.

## 6.2 Cooperation Learning via Pheromone Evaporation

Assuming that each agent represents an *ant*, and that the offers made represent the *pheromone* deposited along the edge $(i,j)$ of a complete weighted network, the ACO-based update rule for cooperation learning is defined as follows:

$$\mathcal{C}_{i,j} \leftarrow (1 - \rho)\,\mathcal{C}_{i,j} + \rho\,o_{i,j} \qquad (15)$$

where $\rho$ is the pheromone evaporation coefficient. This update rule, essentially, aims to learn the ongoing offers made by each agent in the society.

## 6.3 Evolutionary Group Identity Detection

Matrix $\mathcal{C}$ can be interpreted as a complete, directed and weighted network, $\mathcal{CN} = (\mathcal{S}, \mathcal{C})$. The goal of the GID module is the partitioning of $\mathcal{CN}$ into community structures [19]. However, the task of optimal network partitioning into community structures is known to be computationally hard (NP-complete) over the set of all graphs of a given size [6, 19]; for these reasons, we rely upon the benefits of genetic search for approximating a solution [9].

The EA we implemented considers a population of $m$ chromosomes; each chromosome $x_i$ has $n$ genes, each representing the group identity $c_k$ of agent $a_k \in \mathcal{S}$. The EA aims to maximise the following modularity measure [6, 9]:

$$f(x) = \frac{1}{w} \sum_{i,j} \left[\mathcal{C}(i,j) - \frac{w_i^{\mathrm{in}}\,w_j^{\mathrm{out}}}{w}\right] \delta(c_i, c_j) \qquad (16)$$

where $w = \sum_{i,j} \mathcal{C}(i,j)$ is the total sum of the up-to-date $\mathcal{C}$ values, $w_i^{\mathrm{in}}$ is the in-degree of vertex $i$/agent $a_i$, $w_j^{\mathrm{out}}$ is the out-degree of vertex $j$/agent $a_j$, $c_i$ is the label of the community/group identity to which vertex $i$/agent $a_i$ is assigned, and $\delta(c_i, c_j)$ is the Kronecker delta function, for which $\delta(c_i, c_j) = 1$ if $c_i = c_j$ and $\delta(c_i, c_j) = 0$ otherwise.

The possible gene values are taken from an alphabet $G$, which holds the labels of possible group structures. The set of different $c \in G$ symbols in $x_i$, $\hat{G}_i \subseteq G$, represents the group structures detected by $x_i$. At the beginning of each experimental setup, $G$ is initialised as $G = \{0, 1\}$ and the chromosomes are initialised with gene values sampled uniformly within $G$. In other words, our GM framework initially assumes the existence of two groups within $\mathcal{S}$.

At the end of episode $t$, once $\mathcal{C}$ has been updated, the EA is reinitialised by following these five steps: (1) the EA recalculates the fitness values of its genotypes, and sorts the population by descending fitness, since the updates of $\mathcal{C}$ performed by the $CL$ module might have generated changes in the chromosome ranking of the EA; (2) only the chromosomes with the highest possible fitness, $\{x_0, \ldots x_j\}$ remain in the population, the remaining $m - (j + 1)$ chromosomes are discarded; (3) the EA calculates the average number of group structures detected by the individuals who survived the previous step, $\mu_{\hat{G}} = (\sum_{i=0}^{j} \hat{G}_i)/(j + 1)$; (4) the EA initialises $G$ as $G = \{0, \ldots \mu_{\hat{G}}/2\}$; (5) the EA fills its population by generating $m - (j + 1)$ chromosomes which are uniformly sampled within $G = \{0, \ldots \mu_{\hat{G}}/2\}$. The fitness of the $m - (j - 1)$ chromosomes is calculated and the whole population of $m$ chromosomes is sorted by decreasing fitness.

At this point, the evolutionary process starts. Each evolutionary generation includes the following key steps: (1) the first half of the population remains in the current population, the second half is replaced by $m/2$ offspring; (2) each new offspring $x_h$ is generated as follows: first, two chromosomes ($x_i$ and $x_j$) are chosen via rank selection and are recombined via uniform crossover; then, mutation is applied to each gene with a mutation probability $p_{mut}$. The gene value under mutation is replaced by one of the symbols in $G$ which are picked via a uniform random distribution; (3) the fitness of the $m/2$ offspring is calculated.

At the end of each evolutionary generation, the average fitness value of the first fittest half of the population is calculated. If the average is not improving for $l$ consecutive generations (local limit), $G$ is updated by inserting a new symbol. Alternatively, if the average is not improving for $g$ consecutive generations (global limit), the EA terminates and the chromosome with the highest fitness, $x_{\text{GID}}$, is the one chosen by GID to represent the group structures and identities of $\mathcal{S}$.

# 7. RESULTS

In order to examine which of the CL three modules, combined with the evolutionary GID module, would better infer the existence of *emergent* group structures of reciprocal friendship we have conducted experiments on two scenarios: 20 and 50 agent societies. Each scenario was iterated for 30 experimental runs; the three GM frameworks were evaluated against the same experimental setup, in order to grasp a better, more generic understanding of their performance, and to limit random fluctuations introduced by the interaction classifiers and the stochastic nature of the EA.

The performance measure we consider in this study is the normalised mismatch error $nme(C_{\text{RF}}, x_{\text{GID}})$ [9], between the community structures of $FN_{\text{rec}}$ (see Sec. 4), and the inferred group identities with the highest obtained fitness function, $x_{\text{GID}}$ (see Subsec. 6.3), calculated as follows:

$$nme(C_{\text{RF}}, x_{\text{GID}}) = \frac{n - h(C_{\text{RF}}, x_{\text{GID}})}{n}$$

where $h(C_{\text{RF}}, x_{\text{GID}})$ is the maximum assignment score obtained by running Kuhn's Hungarian algorithm [16].

The societies were instantiated by using the following parameters: $\hat{o}_{min} = 0.3$, $\hat{o}_{max} = 0.5$, $\delta_{o,min} = 0.01$, $\delta_{o,max} = 0.03$, $\delta_{s,min} = 0.02$, $\delta_{s,max} = 0.05$, $\tau_{A,min} = 0.5$, $\tau_{A,max} = 0.6$, $\tau_{F,min} = 0.5$, $\tau_{F,max} = 0.7$; $T = 5$ and 150 episodes

were considered for $n = 20$, whilst $T = 12$ and 700 episodes were considered for $n = 50$. These parameters allow for the instantiation of societies of heterogeneous agents, which are capable of manifesting different levels of relational-based social preferences [24], yet by avoiding the generation of overly chaotic dynamics. The following parameters were used for the three CL modules: $\alpha = 0.1$, $\beta = 0.25$, $\gamma = 0.25$, $\rho = 0.3$. Finally, all frameworks shared the same parameters for the GID Module: $m = 50$, $g = 50$, $l = 5$ and $p_{\text{mut}} = 0.8$. These parameters were retrieved after a preliminary — though not exhaustive — fine tuning phase of the framework modules.

## 7.1 20-Agent Societies

Figure 2(a) depicts the average performance, and relative standard deviations, of the three GM frameworks based on $\alpha$-constant Monte Carlo ($\alpha$-cMC; see Eq. 13), Temporal Difference (TD0, see Eq. 14) and pheromone evaporation update rules (Pheromone; see Eq. 15) for the 20-agent society scenario, across 30 iterations. The grey vertical bar at $t = 119$ indicates the Episode beyond which no new reciprocal friendship links are perceived by the agents. Figure 2(b) depicts, in detail, the lowest scored $nme$, for the three frameworks, across those 30 iterations.

As it can be clearly understood, a CL module based on RL manages to better infer the existing global structures, as $TD0$ reaches a score of $nme = 0$ for 57% of the iterations and $\alpha$-cMC scores $nme = 0$ for 50% of the runs, whereas Pheromone obtains the same performance only 20% of the runs. Remarkably, TD0 manages to score the smallest average $nme$ across all experiments: 51%, 41% and 8% of the Episodes for TD0, $\alpha$-cMC and Pheromone, respectively; moreover, TD0 registers the lowest average $nme$ from the early stages up to $t = 109$, that is, when the agents' social perceptions are nearly consolidated; then, from there until the second to last Episode, $\alpha$-cMC registers the lowest average $nme$.

With respect to this study, hence, it appears that the combination of $\mathcal{I}_{\text{HAP}}$ with TD0 is a more suitable technique for cooperation/collaboration learning during the *emergence* of global patterns, though $\alpha$-cMC seems to better perform when the global dynamics slow down and tend to reach an equilibrium.

## 7.2 50-Agent Societies

Figure 3(a) depicts the average performance, and relative standard deviations, of the three GM frameworks, for the 50-agent society scenario, across 30 iterations; the grey vertical bar at $t = 602$ indicates the Episode beyond which no new reciprocal friendship links are perceived by the agents. In addition, Figure 3(b) depicts the lowest scored $nme$ for the three frameworks across the 30 iterations.

Although none of the three framework manages to reach $nme = 0$, $\alpha$-cMC and TD0 outperform Pheromone in terms of minimum $nme$ scored (see Figure 3(b)). Pheromone's best score is $nme = 0.26$ (i.e. 13 misclassified agents) whilst both $\alpha$-cMC and TD0 score $nme = 0.12$ (i.e. 6 misclassified agents). Moreover, for 57% and 40% of the runs, $\alpha$-cMC and TD0, respectively, obtain a minimum $nme < 0.26$.

The evidence that both RL-based frameworks perform better than Pheromone throughout the whole experiment duration is observed in Figure 3(a). On average, Pheromone never manages to perform better than neither $\alpha$-cMC nor TD0. Moreover, we observe that $\alpha$-cMC performs better
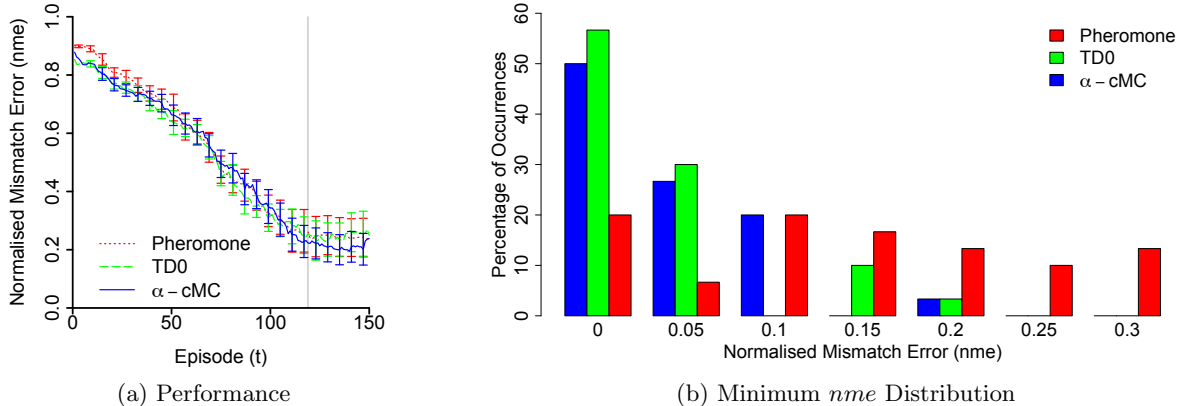
| (a) Performance | (b) Minimum $nme$ Distribution |

**Figure 2: 20-Agent Societies**

than TD0 for 71% of the experiment's episodes. Furthermore, $\alpha$-cMC consolidates its role of best framework starting from Episode 450, i.e. when the number of reciprocal friendship links is still far from reaching an equilibrium.

With respect to this study we conclude that the combination of $\mathcal{I}_{\mathrm{AP}}$ with $\alpha$-constant Monte Carlo update rule is a more suitable technique for cooperation learning during both the *emergence* and consolidation of global patterns.

### 7.2.1 Analysis via Network Visualisation

Despite the fact that $\alpha$-cMC has proven to be the best GM framework, it never reached $nme = 0$ for the 50-agent scenario. Moreover, in both 20 and 50-agent scenarios, the framework seems to converge, on average, to a suboptimal $nme$ of about 0.20 and 0.36 respectively. Arguably, two might be the possible reasons: either (1) the GID module cannot explore well the search space of group identity combinations, or (2) CL provides misleading information, within $\mathcal{C}$, about the existing friendship links.

In order to identify what caused the performance drop in the 50-agent scenario, we focus on the experimental run for which $\alpha$-cMC scored the lowest misclassification error, $nme = 0.2$, at the end of the simulation (i.e. Episode 700). Within this experimental run, the emphasis is given to $t = 677$ and $t = 678$, i.e. the Episodes in which $\alpha$-cMC scores its lowest misclassification error $nme = 0.16$.

Figure 4(a) depicts a graphical representation of $FN_{\mathrm{rec}}$ at $t \in [599, 700]$; the network's consolidated *true* group identities are represented by the network's node colours. Similarly, Figures 4(b) and 4(c) depict the inferred group identities, detected by $\alpha$-cMC, at $t = 677$ and $t = 678$. At $t = 677$, the $\alpha$-cMC misclassifies five single agents — i.e. $a_{20}$, $a_{27}$, $a_{34}$, $a_{44}$ and $a_{47}$ — and a *subgroup* composed of agents $a_{25}$, $a_{39}$ and $a_{40}$. At $t = 678$, instead, the GM changes the group identities of only $a_{34}$ and $a_{47}$. At $t = 677$, GID's EA algorithm executed 66 generations, whilst at $t = 678$ the total number of generations was 123. The fittest member of the population increases its fitness from $f(x) = 0.0877$ at $t = 677$ to $f(x) = 0.0897$ at $t = 678$. Remarkably, at the end of the computation for both Episodes, the first half of EA's genetic population, i.e. the candidate parents, scored the same highest $f(x)$ values meaning that, the population converged prematurely, the uniform crossover operation be-

came meaningless and the only way to explore the search space was through the mutation operation. This suggests that the GM framework would need further modifications for increasing and maintain genetic diversity (such as fitness sharing and niching [10]) if we desire lower $nme$ values in the more complex agent scenarios. Nevertheless, the modularity measure Eq. (16), based on the learned $\mathcal{C}$ matrix at $t = 677$ and $t = 678$, with respect to the *true* group identities, returns $f(x) = 0.0771$ and $f(x) = 0.0781$, respectively; in other words, GID manages to detect group identities which have a higher modularity measure than the *true* ones.

## 8. DISCUSSION AND FUTURE WORK

The combination of the Agent Preference Interaction Classifier ($\mathcal{I}_{\mathrm{AP}}$) with $\alpha$-constant Monte Carlo ($\alpha$cMC) update rule for the cooperation learning module, together with the EA approach adopted by the group identity detection module, allows for the construction of a GM framework which managed to reach a zero misclassification error for 20-agent societies. Although it could not reach optimal performance in the experiments with the 50-agent societies, the GM framework provided the best misclassification error — among the different framework configurations we have considered in this study — of *emergent* and consolidated group structures and group identities.

In general, a RL-based update rule appears to be a key component for the indirect inference of group structures by solely observing the flow of interactions among the agents. In fact, the Pheromone evaporation technique, which merely attempts to keep track of the ongoing offers made by the agents, provided the worst results for both 20 and 50-agent societies. The rationale behind the success of $\alpha$-cMC, and in part TD0, is their ability to perform the transition from the raw interactions to the two *in/out-group* classes. This was achieved via the one-to-many interaction classifiers $\mathcal{I}_{\mathrm{AP}}$ and $\mathcal{I}_{\mathrm{HAP}}$ [11]. These classifiers measure altruism as a more general property of the agents, which cannot be successfully estimated by isolating each single agent interaction.

Although the considerations made in subsection 7.2.1 may suggest that CL could potentially lead to a misleading representation of the ongoing dynamics occurring within the society, we rather argue that CL provides an augmented rep-
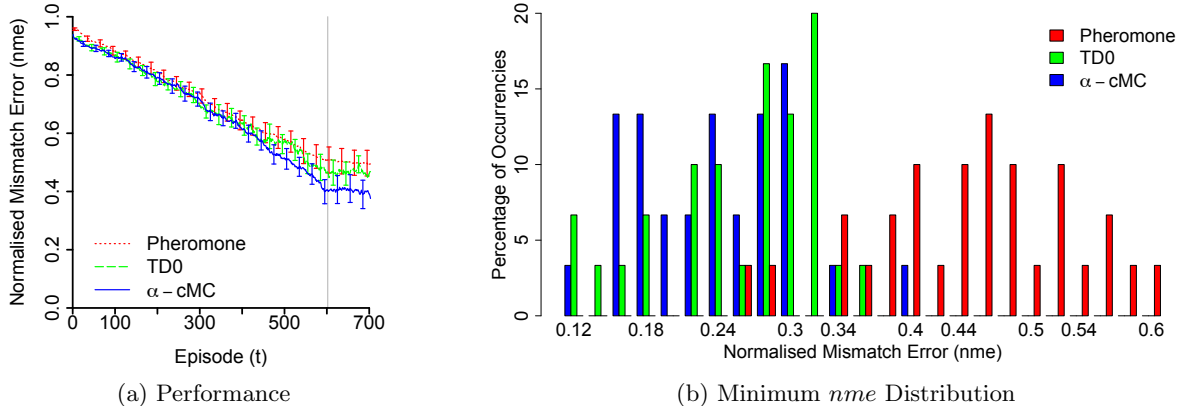
(a) Performance

(b) Minimum $nme$ Distribution

**Figure 3: 50-Agent Societies**



(a) Consolidated *True* Group Identities.

(b) Inferred Group Identities at $t = 677$.
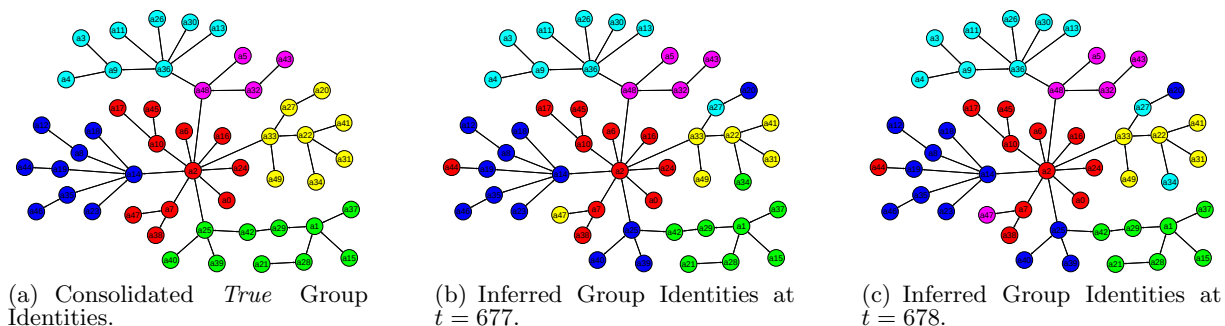
(c) Inferred Group Identities at $t = 678$.

**Figure 4: Differences in Group Identity Assignment between the *True* and the inferred community structures.**

resentation of $FN_{\text{rec}}$, by transforming the dichotomy "friend-vs-non friend" into a continuous one. This opens to a plethora of improvement possibilities, two of which are proposed here: (1) a new module could be embedded between CL and GID; its purpose would be to refine the $\mathcal{C}$ matrix by e.g. identifying possible *in/out-group* thresholds and then apply some form of saturation of the $\mathcal{C}(i, j)$ values; (2) given that the $\mathcal{C}(i, j)$ values range within the $[0, 1]$ interval, GM could be enriched by a new interactive component, giving the possibility to the agents to self-report, at any time, their own perceptions of friendship, and thus directly affect $\mathcal{C}$.

Although the EA algorithm used for the GID module managed to reach a zero error misclassification for the 20-agent scenario, it showed partial inefficacy for the 50-agent scenario; the reason is most likely due to the fact that it seems incapable of maintaining a sufficient genetic diversity in its chromosomes. Therefore, we are confident that our EA approach, being generic in terms of genetic representations and operations, and being capable of automatically augmenting its alphabet size and terminate its computation, is a robust technique against the scaling of the society size. Ongoing work aimed to overcome the genetic diversity drawback are being conducted; more specifically, we are currently focusing on the use of automatic niche shifting [10].

The choice of using the Ultimatum Game as interaction protocol opens for immediate applications of our GM framework to new experimental scenarios based on both artificial societies and on humans; the interactions protocols

can either be canonical Social Dilemmas, e.g. the Dictator Game [7], or complex scenarios such as collaborative multiplayer games [30]. In accordance with the latter, ongoing work are focused on the application of our framework to the analysis of the Glitch gameplay data [17], in order to detect and further analyse eventual discrepancies between inferred groups and the game's social networks of friendship among the players. Finally, the GM's fundamental characteristic of indirect modelling can be combined with Procedural Content Generation (PCG) to extend the field of Experience-Driven PCG to adaptive multiplayer games [33].

Beyond future application of our modelling framework, we will also embark a thorough fine-tuning phase of the GM's parameters; more specifically, we are considering the use of meta-EAs for the artificial evolution of genetic populations of GM frameworks.

## 9. CONCLUSIONS

We proposed a computational framework capable of inferring the existence of group structures based on social networks of reciprocal friendship. This framework assigns relative group identities to complex adaptive artificial societies of social agents which interact with each other through the ultimatum (or bargain) social dilemma game. In contrast to some other approaches, our modelling framework performs the inference based only on the flow of interactions occurring within the society. On that basis, it combines a

reinforcement learning-based cooperation learning module (CL) and an evolutionary algorithm-based group identity detection module (GID). Experimental setups conducted on adaptive societies of two different sizes showed the effectiveness of our framework in minimising the misclassification error between the *true* and inferred group identities.

## 10. ACKNOWLEDGMENTS

## 11. REFERENCES

[1] A. Barabási and R. Albert. Emergence of Scaling in Random Networks. *Science*, 286(5439):509–512, 1999.

[2] W. Buckley. *Society-A Complex Adaptive System: Essays in Social Theory*, volume 9. Routledge, 1998.

[3] A. Clauset, M. Newman, and C. Moore. Finding Community Structure in Very Large Networks. *Physical Review E*, 70(6):066111, 2004.

[4] R. M. Dawes and D. M. Messick. Social Dilemmas. *International Journal of Psychology*, 2(35):111–116, 2000.

[5] M. Dorigo and K. Socha. An Introduction to Ant Colony Optimization. *Handbook of Approximation Algorithms and Metaheuristics*, pages 26–1, 2006.

[6] C. Farmer and A. Fotheringham. Network-based Functional Regions. *Environment and Planning A*, 43(11):2723–2741, 2011.

[7] E. Fehr and K. M. Schmidt. A Theory Of Fairness, Competition, And Cooperation. *The Quarterly Journal of Economics*, 114(3):817–868, 1999.

[8] R. Forsythe. Fairness in Simple Bargaining Experiments. *Games and Economic Behavior*, 6(3):347–369, 1994.

[9] C. Grappiolo, J. Togelius, and G. N. Yannakakis. Artificial Evolution for the Detection of Group Identities in Complex Artificial Societies. In *Proc. of IEEE SSCI ALife*, pages 126–133, 2013.

[10] C. Grappiolo, J. Togelius, and G. N. Yannakakis. Shifting Niches for Community Structure Detection. In *Proc. of CEC, (to appear)*, 2013.

[11] C. Grappiolo and G. N. Yannakakis. Towards Detecting Group Identities in Complex Artificial Societies. In *Proc. of SAB*, pages 421–430, 2012.

[12] G. W. Greenwood and D. Ashlock. Evolutionary Games and the Study of Cooperation: Why Has So Little Progress Been Made? In *Proc. of WCCI*, pages 680–687, 2012.

[13] R. A. Hammond and R. Axelrod. The Evolution of Ethnocentrism. *Journal of Conflict Resolution*, 50(6):926–936, 2006.

[14] M. Janssen and E. Ostrom. Empirically Based, Agent-based Models. *Ecology and Society*, 11(2):37, 2006.

[15] J. H. Kagel, C. Kim, and D. Moser. Fairness in Ultimatum Games with Asymmetric Information and Asymmetric Payoffs. *Games and Economic Behavior*, 13(1):100–110, 1996.

[16] H. Kuhn. The Hungarian Method for the Assignment Problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, 2006.

[17] P. M. Landwehr. A collection of economic and social data from glitch, a massively multiplayer online game. 2013.

[18] S. J. Lansing. Complex Adaptive Systems. *Annual Review of Anthropology*, 32:183–204, 2003.

[19] E. A. Leicht and M. E. J. Newman. Community Structure in Directed Networks. *Physical Review Letters*, 100(11):118703, 2008.

[20] K. Lerman and A. Galstyan. Automatically Modeling Group Behavior of Simple Agents. In *Proc. of AAMAS*, 2004.

[21] J. Liu, H. Abbass, W. Zhong, and D. Green. Local-Global Interaction and the Emergence of Scale-free Networks with Community Structures. *Artificial Life*, 17(4):263–279, 2011.

[22] J. Liu, W. Zhong, H. Abbass, and D. Green. Separated and overlapping community detection in complex networks using multiobjective Evolutionary Algorithms. In *Proc. of CEC*, pages 1–7, 2010.

[23] V. Martinez, G. I. Simari, A. Sliva, and V. Subrahmanian. CONVEX: Similarity-Based Algorithms for Forecasting Group Behavior. *IEEE Intelligent Systems*, 23:51–57, 2008.

[24] F. Marzo, B. J. Grosz, and A. Pfeffer. Social preferences in Relational Contexts. In *Proc. of Collective Intentionality*, 2005.

[25] M. A. Nowak, C. E. Tarnita, and T. Antal. Evolutionary Dynamics in Structured Populations. *Philosophical Transactions of the Royal Society*, 365(1537):19–30, 2010.

[26] C. Pizzuti. GA-Net: A genetic Algorithm for Community Detection in Social Networks. *Parallel Problem Solving from Nature*, 5199:1081–1090, 2008.

[27] H. Situngkir. On Selfish Memes: Culture as Complex Adaptive System. *Journal of Social Complexity*, 2(1):20–32, 2004.

[28] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction*. Cambridge University Press, 1998.

[29] M. Tasgin, A. Herdagdelen, and H. Bingol. Community Detection in Complex Networks Using Genetic Algorithms. *arXiv:0711.0491*, 2007.

[30] Y. Vogiazou and M. Eisenstadt. Designing Multiplayer Games to Facilitate Emergent Social Behaviours Online. *Interactive Technology and Smart Education*, 2:113–126, 2005.

[31] E. Wolf. Kinship, Friendship, and Patron-client Relations in Complex Societies. *The Social Anthropology of Complex Societies*, pages 1–22, 1966.

[32] B. Xianyu. Social Preference, Incomplete Information, and the Evolution of Ultimatum Game in the Small World Networks: An Agent-Based Approach. *Journal of Artificial Societies and Social Simulation*, 13(2):7, 2010.

[33] G. N. Yannakakis and J. Togelius. Experience-Driven Procedural Content Generation. *IEEE Transactions on Affective Computing*, 2:147–161, 2011.

[34] G. N. Yannakakis, J. Togelius, R. Khaled, A. Jhala, K. Karpouzis, A. Paiva, and A. Vasalou. Siren: Towards Adaptive Serious Games for Teaching Conflict Resolution. In *Proc. of ECGBL*, pages 412–417, 2010.