

Electronic Social Capital for Self-Organising Multi-Agent Systems

PATRICIO E. PETRUZZI, JEREMY PITT, and DÍDAC BUSQUETS, Imperial College London

It is a recurring requirement in open systems, such as networks, distributed systems and socio-technical systems, that a group of agents must coordinate their behaviour for the common good. In those systems – where agents are heterogeneous – unexpected behaviour can occur due to errors or malice. Agents whose practices free-ride the system can be accepted to a certain level; however, not only do they put the stability of the system at risk, but they also compromise the agents that behave according to the system’s rules.

In social systems, it has been observed that *social capital* is an attribute of individuals that enhances their ability to solve collective action problems. Sociologists have studied collective action through human societies and observed that social capital plays an important role in maintaining communities through time as well as in simplifying the decision making in them. In this work, we explore the use of Electronic Social Capital for optimising self-organised collective action.

We developed a context-independent Electronic Social Capital framework to test this hypothesis. The framework comprises a set of handlers that capture events from the system and update three different forms of social capital: trustworthiness, networks and institutions. Later, a set of metrics are generated by the forms of social capital and used for decision-making. The framework was tested in different scenarios such as 2-player games, n -player games and public goods games. The experimental results show that social capital optimises the outcomes (in terms of long-term satisfaction and utility), reduces the complexity of decision-making and scales with the size of the population.

This work proposes an alternative solution using Electronic Social Capital to represent and reason with qualitative, instead of traditional quantitative, values. This solution could be embedded into socio-technical systems to incentivise collective action without commodifying the resources or actions in the system.

Additional Key Words and Phrases: social capital, nested self-organising systems, electronic institutions, resource allocation, common-pool resource

ACM Reference format:

Patricio E. Petruzzi, Jeremy Pitt, and Didac Busquets. 2016. Electronic Social Capital for Self-Organising Multi-Agent Systems. 1, 1, Article 1 (January 2016), 25 pages.

DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Open systems, such as networks, distributed systems and socio-technical systems, often face the challenge that a group of participants must cooperate or coordinate their behaviour for the *common good*. Usually, these system participants are referred to as *agents*, and form opportunistic alliances

This work is partially supported by the UK EPSRC Grand Challenge Project Autonomous Power System.

This article is an extended synthesis of Petruzzi et al. [36–39].

Authors’ address: P.E. Petruzzi, J. Pitt and D. Busquets, Department of Electrical and Electronic Engineering, Imperial College London, SW7 2BT, UK; emails: {p.petruzzi12, j.pitt, d.busquets-font}@imperial.ac.uk.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM. XXXX-XXXX/2016/1-ART1 \$15.00

DOI: 10.1145/nnnnnnn.nnnnnnn

to manage *resources* and perform *collective actions* that would otherwise be impossible or very costly to achieve individually. These systems also lack a central authority, which makes them dependant on self-organising mechanisms to succeed.

When agents of these systems share a resource, they must agree on how the appropriation of it is managed. Generally, the objective of allocating resources is to find a distribution of resources that is considered ‘fair’ and sustainable over time. The nature of the resource as *common* and the lack of a central authority to control it can encourage agents to free-ride the system by not contributing to it. This can be either the maintenance or the management of resources, depending on system’s characteristics. A traditional approach is to use mechanism design [25] to model the situation as a strategic game and then use game theory, which allows the agents to analyse the game they are confronted with and decide on the best course of action or strategy to play.

Alternatively, Elinor Ostrom analysed how human societies are able to create institutions for the management and government of *resources* [32]. This allowed them to avoid the “tragedy of the commons” [22] (depletion of resources) predicted by a game-theoretic analysis. She defined institutions as “a set of working rules that are used to determine who is eligible to make decisions in some arena, what actions are allowed or constrained, ... [and] contain prescriptions that forbid, permit or require some action or outcome” [32, p. 51]. The “working rules” specify procedures for operational, collective and constitutional choices.

In later work [33], *social capital* has been analysed and defined as “an attribute of individuals that enhances their ability to solve collective action problems”. Researchers also proposed that *social capital* created in one institution can enable other institutions to succeed, when both institutions are codependent [32, pp.133–136].

An example of two codependent institutions is an irrigation system. In this example, the system is actually composed of at least two different common resources: the water and the channel. Each of these resources represents a collective action problem to their users. Users must share the maintenance cost to keep the channel functional, which represents a provisioning problem. On the other hand, the water use represents an appropriation problem that is associated with water allocation among users and the tools used to monitor the compliance with water rights. Despite the difficulties of managing different codependent institutions, Ostrom’s analysis of communities across the world shows that some communities have actually managed to avoid the resource depletion.

In this work, we explore if the use of Self-Organising Electronic Institution (SOEI) with Electronic Social Capital (ESC) can enhance the ability to solve collective action problems in Multi-Agent System (MAS). We developed a context-independent ESC framework to test this hypothesis. Furthermore, we test whether ESC created in one SOEI can enable another SOEI to succeed, when both these institutions are codependent, as shown in the previous example.

This is the fundamental contribution of this paper. It defines a social capital framework that not only defines data structures for representing attributes of agents, but also defines processes for updating, evaluating and decision-making with social capital. Also, it presents a set of experiments results which show that social capital optimises outcomes, reduces the complexity of decision-making and scales with the size of the population in different collective action situations. And it provides an implementation of the framework which can be reused for further experiments.

Therefore, this paper is structured as follows. Section 2 analyses cooperation in MAS and introduces social capital. The electronic social capital framework is detailed in Section 3 and an implementation is presented. Section 4, 5, 6 describe the experiments performed in 2-player games, n -player games and public goods games. Section 7 presents a real world application scenario. After discussing related work in Section 8 and further work in Section 9, we summarise and conclude in Section 10.

2 SOCIAL CAPITAL FOR COLLECTIVE ACTION

In this section, we discuss a set of common characteristics shared between different types of open MAS that are subject to the problem of resource allocation, such as manufacturing and scheduling, public transport, grid computing or network routing. We analyse cooperation in MAS from a game-theoretical point of view and contrasted with Ostrom's work on cooperation in human societies. To conclude, we present the notion of social capital, which can play a key role in promoting cooperation among institutions, and some examples of basic forms of social capital used in computer systems are described.

2.1 Resource allocation in open multi-agent systems

The allocation of resources is relevant to a wide range of domains, such as manufacturing [5] and scheduling [10], public transport [6], grid computing [18] and network routing [15]. All of these examples share the following common characteristics [7]:

- *Resources* refers to items that are being distributed among the participating agents of the system, and can be classified into *divisible* and *indivisible*, depending on their nature.
- *Allocation* refers to a particular distribution of resources between the agents.
- Agents can have *preferences* regarding the allocations they receive. In some cases, agents can communicate their preferences by sending demands to other participants.
- The *allocation method* defines the procedures required to perform an allocation. Usually, the method can be divided into centralised or distributed.
- The *objective* of resource allocation is to find an allocation that is either feasible or optimal. The feasibility refers to finding a suitable allocation that solves the problem. The optimality, on the other hand, relates to finding an allocation that maximises utility for the allocator.

In some scenarios, a combination of different objectives might be also possible. The main objective could include searching for an optimal solution amongst a set of feasible ones.

2.2 Cooperation: Game theory view

An agent involved in an interaction with other agents faces the problem of deciding whether to cooperate (e.g. provision resources or consume the allocated amount) with the other agents or not. A traditional approach is to model the situation as a strategic game and use game theory, which allows the agents to analyse the game they are confronted with and decide on the best action or strategy to play. This usually involves computing the Nash equilibrium [30], with which the agent is guaranteed that it cannot make itself better off by unilaterally changing its strategy.

While game theory has been long applied to many problems, it relies on strong assumptions, such as: *static games*, in the sense that games are one-shot interactions instead of repeated over time; *rational agents* with infinite computational power to compute the best strategy; and *complete information* about the other agents' available strategies and possible payoffs. Moreover, game theory imposes other constraints and assumptions that might not be appropriate to model realistic scenarios [1]. For example, it does not usually consider the possibility of refusing to engage in an interaction, even if the agent believes that it is more beneficial to do so.

The Logic of Collective Action [31] analysed the premise of rational individuals. Although, when individuals within a group are considered *rational* and *self-interested*, and the group's and individuals' interests are aligned, it would be logical for them to act together to pursue a common objective, this is not always the case. Individuals will always have incentives to free-ride the system, but this could be counteracted by only providing the benefits to active participants.

Moreover, Elinor Ostrom analysed how human societies were able to create institutions for the management and government of resources [32], which allowed them to avoid the "tragedy of the

commons” [22] (depletion of the resources) predicted by a game-theoretical analysis. Furthermore, “What makes these models so dangerous – when they are used metaphorically as the foundations for policy – is that the constraints that are assumed to be fixed for the purpose of analysis are taken on faith as being fixed in empirical settings... [I’d] rather address the question of how to enhance the capabilities of those involved to change the [constraints] to lead to outcomes other than remorseless tragedies” [32, pp.6–7].

In the next section we introduce the notion of social capital, which can be used to “enhance the capabilities of those involved...” and facilitate the management of resources to avoid the resource depletion.

2.3 Social capital

Capital comprises multiple forms, such as physical, human or social. Physical capital is entirely tangible and created by modifying materials to assist production (e.g. computers, railroads or factories). Human capital is less tangible, created by knowledge and skills obtained by an individual [45]. The theory of social capital gained importance through the integration of classical sociological theory with the description of an intangible form of capital.

Through the concept of social capital, researchers have tried to propose a synthesis between the value contained in communitarian approaches and the individualism professed by the *rational choice theory*. Social capital is therefore defined as a set of intangible collective resources that an individual or a group of individuals holds and “comes about through changes in the relations among them that facilitate action” [9].

Other relevant social capital definitions include: “The features of social organization, such as networks, norms and trust, that facilitate coordination and cooperation for mutual benefit” [43]; “The aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition...” [3]; and “An attribute of individuals that enhances their ability to solve collective action problems” [33].

Although it appears that there is still no agreement regarding on a single concrete and formal definition of ‘social capital’, researchers seem to agree on the qualitative value of connections and/or relations among individuals. These developed relations combined with institutionalised rules that are commonly agreed upon can facilitate the coordination (and cooperation) when performing collective actions.

Ostrom and Ahn observed that social capital has multiple forms, of which they identified three:

- *Trustworthiness*: This is distinct from trust and is related to reputation (i.e. a shared understanding of someone’s willingness to honour agreements and commitments).
- *Networks*: These include strong and weak ties, and identify channels through which people communicate, as well as other social relations.
- *Institutions*: These are identified as sets of conventional rules by which people voluntarily, and mutually agree to, regulate their behaviour.

They also suggested that *trust* itself was the ‘glue’ that enabled these various forms of social capital to be leveraged for solving collective action problems (see Figure 1) – for example, the sustainability of a Common Pool Resource (CPR). Social capital generates ‘reliance’ trust; where reliance trust can be seen as a complexity-reducing decision-making shortcut that helps resolve collective action problems.

Inspired by an analysis of *trust* as comprising a belief component and an expectation component [26], a trust decision can be modelled reasoning about three components: the belief that there is a rule; the expectation that someone else’s behaviour will conform to that rule; and the expectation that a third party will punish any behaviour that does not conform to that rule.

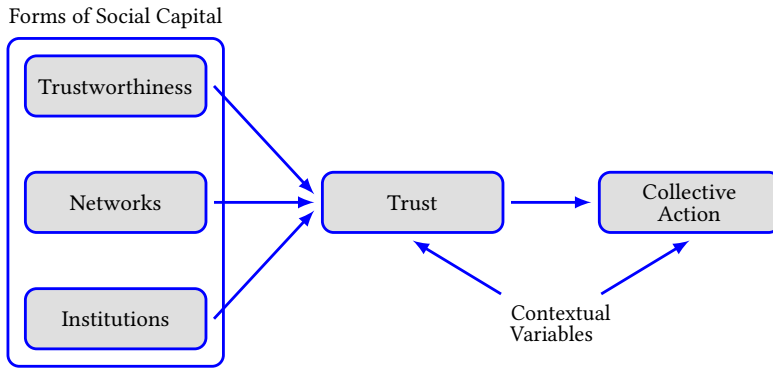


Fig. 1. Ostrom and Ahn's social capital model.

Social capital has also been studied in the context of MAS. For example, it was stated that “there is a big interest in literature about ‘social capital’ and its powerful effects on the wellbeing of both societies and individuals, often it is not clear enough what is it the object under analysis” [14]. In that paper, they proposed ‘trust’ as the capital of agents.

[41] identified several examples of computer systems which represent and reason with social capital in computational form, including forgiveness in e-commerce, legitimate claims for ‘fair’ resource allocation in open networks, and demand-side self-organisation in SmartGrids. The description of those examples follows.

Human society has evolved a standard mechanism for recovering from error in general: *forgiveness*. From the psychological literature, four positive motivations for forgiveness can be identified, comprising twelve constituent signals. This has been formalised in a computational model of forgiveness. The critical aspect of this forgiveness model is that some of the constituent signals, for example ‘prior beneficial relationship’ are an indication of some form of social capital.

Similarly, a theory of distributive justice based on legitimate claims [44] was used in a computational model for self-organised, ‘fair’ resource allocation. In this model the representation of some of the claims – notably the claims according to efforts and sacrifices, and according to socially-useful services – provide a ranking based on quantitative representation of a form of social capital.

In another resource allocation experiment, agents requested time slots for resource access, receive an allocation and are allowed to exchange those allocations among themselves. During each exchange, agent keep a count of “favours received” or “favours done” depending on who benefits from that exchange. The favours are a form of social capital and it has been shown that a system with favours can outperform one without (in terms of the percentage of time slot preferences which are satisfied).

While these examples demonstrate the potential benefit of some form of social capital in computer-assisted or computer-mediated decision-making, they are generally implicit, disjoint and concerned with informing individual-rather than community-based action. What we propose here is a unified computational framework for representing and reasoning about social capital directly targeted to resolving collective action problems.

3 ELECTRONIC SOCIAL CAPITAL FRAMEWORK

We propose a formal framework to represent and reason about (an electronic version of) social capital. In this framework, events enhance or diminish the different forms of social capital, and a decision-making module uses the information from the forms of social capital to decide either to cooperate or not with another agent. Figure 2 shows a schematic view of the framework. A detailed specification of the generic framework can be found in [38].

We have implemented three forms of social capital: trustworthiness, networks and institutions. Each form has an internal update function which process the events, data structures to store the information and outputs a set of metrics. The trustworthiness form of social capital keeps the reputation information received from different sources. The network form stores information about all the actions that the other agents perform. This form of social capital can subjectively analyse the value of a relationship between two different agents among them. The third form of social capital, institutions, uses information on institutional actions performed in the system. A detailed description of each form of social capital is presented later in this section.

We have developed our framework on top of Presage2 [27]. Presage2 is a Java platform for developing discrete time driven animation and simulations of collective adaptive systems. It provides services to simulate large, heterogeneous populations, multiple different networks, inter-agent communication, the physical environment, event recognition, data logging and visualisation.

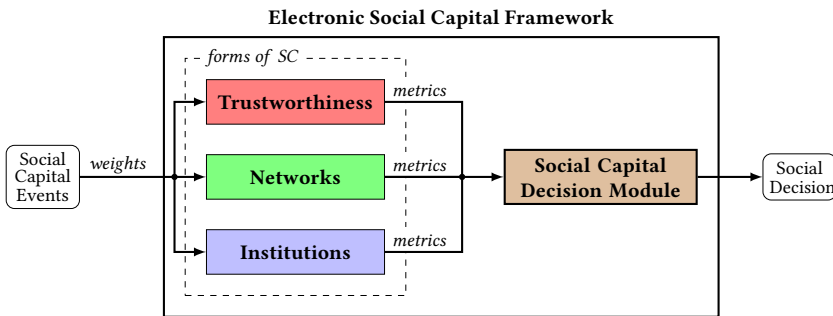


Fig. 2. ESC framework

3.1 Social capital events

The framework comprises a set of context-independent events with an associated weight. When an agent perceives an action performed in the environment, the agent is responsible for translating the action into the appropriate social capital event. In the current version of the framework, the weights are fixed in the range of $(0, 1]$ and assigned by perceived significance (i.e. the weight of being expelled from an institution is higher than the weight for being sanctioned). Furthermore, the weights specify to which forms of social capital are bound. An event can be bound to more than one form of social capital. For example, a sanction performed by an institution is connected to the forms 'networks' and 'institutions'. This is, mainly because, at the networks form the social capital of the agent sanctioned is diminished, while at institutions form, the social capital of the sanctioning institution is increased.

Figure 3 shows the UML diagram of the main classes implemented to model the events. First, the abstract class *SocialEvent* must be extended to define a new event. Depending on which forms of social capital the event affects, it must also implement the appropriated interface. This solution simplifies the assignment of the events to different forms of social capital. This diagram shows only

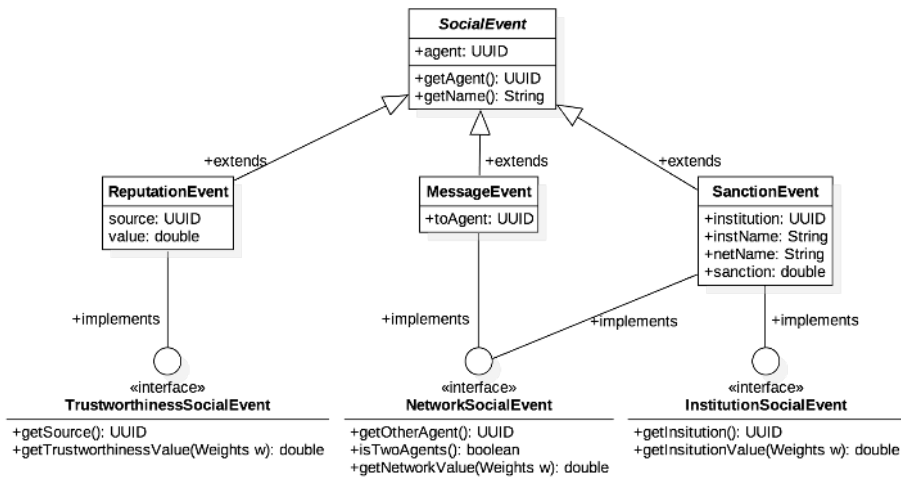


Fig. 3. Social events UML class diagram

an example with three social events. More events can be added to the framework by extending and implementing these classes.

3.2 Forms of social capital

3.2.1 Trustworthiness. This form of social capital collects reputation events. These events have a reputation source, an ‘agent identification’ and a reputation value. The update function normalises the reputation value to a value between zero and one and stores it. The data structure chosen to store this information is a *tuple*, which keeps the source, agent identification and normalised reputation value and updates them after each event is received. When a new reputation event is received, the previous value is overwritten, as the old reputation data is already included in the new one by the reputation sources. This form of social capital produces an output metric for each reputation source included in the system.

3.2.2 Networks. The networks form of social capital receives events regarding agents’ interactions in the system. Examples of these events are whether an agent has cooperated or not, messages sent or received, and all the institutional actions performed (from an agent’s perspective, not the institution’s). Events have a specific weight based on their relevance. Sending a message, successfully cooperating or being banned from an institution do not have the same importance. Later, the event with its weight is used by the update function to update the value at the data structure. The data structure chosen to store the data is a graph, where each vertex represents an agent. The edges – with a value from zero to one assigned – represent the social capital between the agents. It is important to remark that this form of social capital updates information on the agents’ actions even when the event is collected by an agent that does not participate in that action. For example, if two agents interact in some context where a third agent witnesses that action and its outcome, the *networks* form of social capital at the third agent will update the third agent’s graph (the edge between the two agents) with the appropriate value for the outcome of that action. The output metrics of this form of social capital are the values of the edges of a given agent.

3.2.3 Institutions. This form of social capital uses events related to the institutions, such as rules followed, sanctions being applied, agents joining or leaving, etc. A social capital value (between zero and one) is stored for each institution; this value represents the aggregated value of all the

events involving a particular institution. The output metrics of this form are the social capital values of the institutions that an agent is participating.

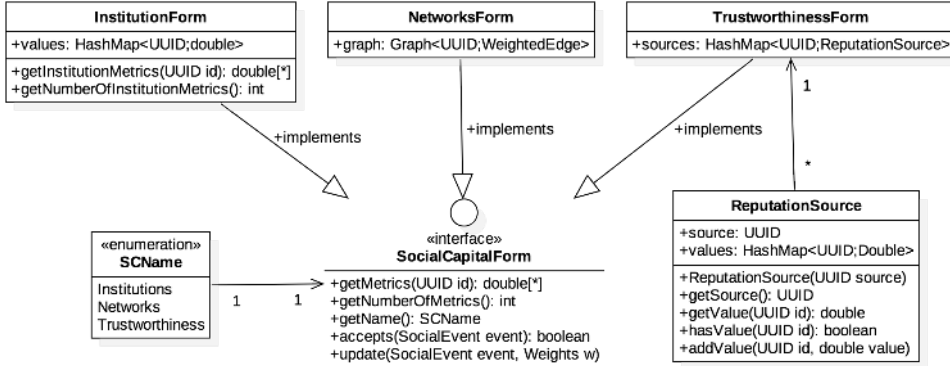


Fig. 4. Forms of social capital UML class diagram

Figure 4 shows the UML diagram of the main classes implemented to model the forms of social capital. The interface *SocialCapitalForm* defines all the methods that any form of social capital must implement, allowing the addition of forms in the future. The methods *accept* and *update* implement the update functions. Furthermore, the method *getMetrics* provides the social capital metrics. One class for each form of social capital is also shown, where the data structures are defined, and the appropriate methods to handle them. These methods are not shown in the diagram for clarity. Table 1 shows a simplification of the data types used in each form of social capital.

Table 1. Data types used in the different forms of social capital

Form of social capital	Data types
Trustworthiness	HashMap < UUID, HashMap < UUID, Double > >
Networks	Graph: - Vertex < UUID > - Edges < UUID, UUID, Double >
Institutions	HashMap < UUID, Double >

3.3 Update of the different forms of social capital

When a social capital event is feed into the framework, the data structures at each form are updated according to the weight w of the event. The networks and institutions forms of social capital use the following equation to reinforce their social capital at the *update* method:

$$sc^{i,t+1} = \begin{cases} sc^{i,t} + w \cdot (1 - sc^{i,t}) & \text{if enhances cooperation} \\ sc^{i,t} \cdot (1 - w) & \text{otherwise} \end{cases}$$

In this equations, $sc^{i,t+1}$ is the social capital for an agent or institution i (depending on which form is updating) and w is the weight assigned to the event used as a learning factor. The first expression is used to reinforce the social capital for the agent or institution i if its actions enhanced cooperation. Conversely, we apply the second expression for actions that diminish cooperation.

The trustworthiness form of social capital does not use this function, because the new reputation event contains the updated value of the agents' reputations. In this case, the new reputation value replaces the old one stored within the appropriated source.

Example. When a sanction event with a weight $w = 0.00005$ and a ID $id = 1$, is fed to the institutions form of social capital, the current $sc^{id,t}$ is retrieved from the HashMap using the id . This value is set by default at 0.5. The new value $sc^{id,t+1}$ is calculated applying the previous formula which results in $sc^{id,t+1} = 0.50025$. This value is stored into the HashMap and will be used as one of the metrics provided by the institution id in the social capital decision module.

3.4 Social capital decision module

Given an agent, the generic model combines n social capital indicators from the forms of social capital to make a social decision, i.e. cooperate. In different contexts, a distinct set of indicators might be enabled. Formally, the cooperation decision uses the social capital $SC(a)$ associated with an agent a as:

$$SC(a) = \sum_{i=1}^n w_i v_i$$

Here, w_i is the weight of each social capital indicator (normalised such that they add up to 1) and v_i is the social capital value of each indicator. This value allows the agent to decide whether to cooperate or not with another agent.

Similarly, to compute the social capital $SC(i)$ associated with a given institution i , the social decision module uses:

$$SC(i) = institution(i)$$

where $institution(i)$ is the social capital indicator associated to i by the institutions form of social capital. The value of social capital of i then allows the agent to make decisions about the institution – for instance, whether to join it, leave it, recommend it to another agent, etc.

Example. The decision module with the three forms of social capital described before might get three indicators: one from *Trustworthiness* – if only one reputation source is included, one from *Networks* – from the edge connecting the agent, and one from *Institutions* – if the agent participates in only one institution. Using the this implementation of the decision module, the three indicators received will be combined by a simple average, and the value used to decide to cooperate or not. An agent will cooperate when the average is equal to or greater than 0.5.

3.5 Brief note on complexity

The ESC framework implementation uses *online* learning through event handlers that update the social capital attributes. The metrics used in this implementation are either constant or linear. Furthermore, the decision-making process combines the values of different metrics using a linear function. Therefore, the complexity of social decision-making in the social capital model (this instantiation) is linear; and in general is linear in the number of attributes and independent of the number of players and their strategies.

It has been suggested that even less is known about computing equilibria for n -player games than for the special case of 2-player games [34]. They, however, show that there is a polynomial-time algorithm for finding a Nash equilibrium in a certain type of n -player k -strategy game; but no such algorithm is known for general games, even when $k = 2$. Given the complexity results for our framework, we suggest that social capital is a way of side-stepping the intractability of computing equilibria in n -player collective action situations.

4 EXPERIMENTS IN 2-PLAYER GAMES

A common requirement of distributed multi-agent systems is for the agents themselves to negotiate pairwise agreements on performing a joint action. If the computational resources expended in negotiating an optimal solution exceed the marginal benefits gained from that negotiation, then it would be more expedient and efficient to use the *memory* of past interactions to short-cut the complexity of decision-making in joint or collective actions of this kind. In the irrigation system example in [section 1](#), the collaboration between the provision institution and the appropriation institution can be modelled as a 2-player game; where a successful cooperation can benefit both institutions.

In this section, the Electronic Social Capital framework is used to decide whether to cooperate or defect in a strategic game. An experimental testbed was implemented, in which a population of agents repeatedly played pairwise games, and used the social capital framework as the basis for their action-selection.

4.1 Testbed: Cooperation game

The Cooperation Game (CG) is a strategic game where a population of agents is repeatedly randomly paired to play a game against each other. At each round, each player has a randomly designated opponent and a two-player strategic game to play. Once paired, players must choose either to *cooperate*, *defect* or *refuse to play*. Then, the payoff matrices are applied and the agents receive or lose points depending on what they have played. If one of the players refuses to play, the game is cancelled and agents do not receive or lose any points. A global count of points is kept for all the players, and it is used to evaluate their performance over the time. Another feature of the CG is institutions, player can participate in different institutions which came with a set of benefits and/or obligations while playing the games. A more detailed description of the CG can be found in [36].

4.1.1 Step-by-step algorithm. Algorithm 1 shows the step-by-step procedure of the CG. At the beginning of each round, random pairs of agents are generated and a specific pairwise game is selected for each of the pairs. We called this a Match. Then, players select the action they want to play in that scenario (lines 9-11). At the end of the round, the players' actions are grouped with their matches and the outcome is calculated (line 13). Here, players will gain or lose points based on their actions and the game's payoff matrix (line 14). Institutions apply their sanctions based on the match results (line 15); agents that violate the rules of the institutions they are members of get sanctioned. Lastly, new institutions are created and the members of the institutions are updated, all based on agents' requests during that round. This process is repeated for every round until the end of the simulation.

4.1.2 Social players. Social players are agents who participate in the cooperation game and have some form of social capital included in their decision-making. Algorithm 1 describes their behaviour in a round of the simulation. When a round starts, players receive a random and limited number of results for the other players' matches (i.e. each agent will receive a different set of results). We use this to model agent observation, communication with other agents, or any form of publishing results that will deficiently spread the match results to other agents. Social agents update their social capital with this information (lines 5-7). Subsequently, they again update their social capital with all the information received by the institutions they are members of (lines 8-10). All institutions send information about who joined, left, was sanctioned, was rewarded or was expelled. In addition, if an institution called a vote to accept a new member, the vote is sent at this point. Next, the player has a probability q of creating an institution (line 11). The configuration of the institution is randomly generated. Two institutions with the same characteristics are not allowed.

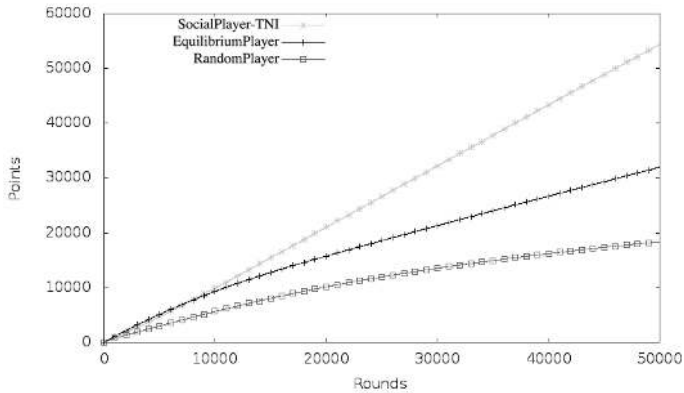
<pre> 1: $A \leftarrow$ set of n agents 2: $G \leftarrow$ set of m games 3: $t \leftarrow 0$ 4: repeat 5: $p \leftarrow$ <i>generate_random_pairs</i>(A, G) 6: for each pair $(i, j, g) \in p$ do 7: create match $m_{ij}(g)$ 8: end for 9: for each agent $i \in A$ do 10: play $action_i$ 11: end for 12: for each pair $(i, j, g) \in p$ do 13: $mr_{ij} \leftarrow$ <i>match_result</i>($g, action_i, action_j$) 14: <i>update_points</i>(mr_{ij}) 15: <i>institutional_sanctions</i>(mr_{ij}) 16: end for 17: create new institutions 18: update institution membership 19: $t \leftarrow t + 1$ 20: until $t == T_{lim}$ </pre>	<pre> 1: $E \leftarrow$ set of n institutional events 2: $R \leftarrow$ set of m other match results 3: $I \leftarrow$ set of o institutions joined by the agent 4: $J \leftarrow$ set of p invitations to join institutions 5: for each match result $mr \in R$ do 6: create social interaction si_{mr} 7: end for 8: for each event $e \in E$ do 9: create social interaction si_e 10: end for 11: create new institution, with probability q 12: join random institution, with probability r 13: leave random institution, with probability s 14: for each invitation $j \in J$ do 15: if <i>accept</i>(j) then 16: join j 17: end if 18: end for 19: for each institution $i \in I$ do 20: if not <i>cooperate</i>(i) then 21: leave i 22: end if 23: end for 24: $opp_t \leftarrow$ <i>current_match_opponent</i> 25: if <i>cooperate</i>(opp_t) then 26: play cooperate 27: else 28: play <i>refuse_to_play</i> 29: end if 30: if opp_{t-1} played cooperate then 31: send invitation to my institution 32: create social interaction si_{coop} 33: end if 34: if opp_{t-1} played not_cooperate then 35: create social interaction si_{not_coop} 36: end if </pre>
--	--

ALGORITHM 1: Cooperation game (left) and Social player (right)

With a probability r and with a probability s , social players will join or leave a random institution respectively (lines 12-13). In these cases, a Boltzmann distribution of the institutions based on their social capital value is used to choose one (when choosing which one to leave the value is inverted using $1 - value_{sc}$). The next step is to process the invitations to join institutions. In order to decide whether to accept the invitation or not, players check the social capital of the institution and of the player who sent the invitation. If these values are greater than a certain threshold, the invitation is accepted (lines 14-18). Following this, the agent checks the social capital of each institution it is a member of. If the value is lower than a threshold, the agent will leave the institution (lines 19-23). Afterwards, the current opponent is retrieved and the action to play is chosen according to this opponent's social capital. If the opponent's social capital value is lower than a threshold, social players refuse to play against this opponent (lines 25-29). Lastly, they receive the information about the last match's results and they update their social capital accordingly (lines 30-36).

4.1.3 *Other players.* The CG also includes other kinds of players. *Probabilistic* players also use social capital but, even if the social capital advises them to cooperate, they play defect in a pre-defined percentage of rounds. *Equilibrium* players choose their actions based on the Nash equilibrium at each game. They defect at prisoner's dilemma and coordination games, and cooperate at full and partial convergence. Lastly, *Random* players choose to *cooperate*, *defect* or *refuse to play* according to random selection.

4.2 Experiments



Agent Type	Score	%	Agent Type	Score	%
SocialPlayer-TNI	54404	100	EquilibriumPlayer	31991	58
SocialPlayer-NI	53118	97	Probabilistic-T-0.25	30663	56
SocialPlayer-TI	52715	96	Probabilistic-N-0.25	30237	55
SocialPlayer-TN	47871	87	Probabilistic-TNI-0.25	28811	52
SocialPlayer-N	47754	87	Probabilistic-TNI-0.75	26595	48
SocialPlayer-T	46393	85	Probabilistic-TNI-0.5	25453	46
Probabilistic-TNI-0.1	45735	84	RandomPlayer	18381	33
SocialPlayer-I	45216	83	Probabilistic-I-0.25	10927	20

Fig. 5. Points obtained for players at prisoner's dilemma

Our first experiment was run a simulation where agents always play prisoner's dilemma. The results are shown in Figure 5. The plot shows the evolution of the average points achieved by different type of agents (for the sake of clarity, only the type obtaining the most points is shown, as well as the *equilibrium* and *random* agents). On the right-hand side of the figure, the average of the final score for each type of agent in the last round of the simulations is shown, as well as the percentage of points w.r.t the best score. It can be seen that, at the beginning *equilibrium* players outperform the rest. Due to the lack of interaction between agents, the forms of social capital do not have any information as yet, and *social* and *probabilistic* players get defected by *equilibrium* players. As the game evolves, each form of social capital starts collecting information and feeding their decision modules. After this, players with any form of social capital start refusing to play against the *equilibrium* players. At around round 8000, the *social* player with trustworthiness, networks and institutions forms of social capital outperforms the *equilibrium* player. *Random* players underperform all other players, but one. A huge difference can be seen between *social* players using any form of social capital and the rest of the players.

5 EXPERIMENTS IN N -PLAYER GAMES

In the irrigation system example in [section 1](#), the collaboration between members of a provision institution or an appropriation institution can be modelled as a n -player game; where a successful cooperation can benefit all the participating members.

In this section, further experiments using the ESC framework are presented. Firstly, a new testbed for n -player games is described, including the customisations regarding the institutions and reputation sources used. Moreover, the different players implemented for this scenario are specified; four new experiments were carried out with these players. To conclude, some related work is presented and compared to the experiments performed.

5.1 Testbed: Unscrupulous diner's dilemma game

A social dilemma is a situation where one individual could benefit from acting selfishly unless everyone acts selfishly, in which case the whole group loses [21]. In a scenario where a group of friends are dining in a restaurant with an unexpressed agreement to equally divide the bill, one person could order an expensive meal and enjoy a magnificent dinner at a bargain price. If everyone at the dinner does so, however, they will all end with an astronomical bill to pay. This situation is called the Unscrupulous Diner's Dilemma (UDD)[21]. A more detailed description of the UDD game can be found in [38].

5.1.1 Step-by-step algorithm. Algorithm 2 shows a step-by-step procedure of the game. At the beginning of each round, agents must choose in which group dinner they will participate. Then, agents select the meal they want to order in that group (lines 8-10). At the end of the round, each bill is calculated (line 12). Here, players will spend money and gain joy based on their actions and the game's payoff matrix (lines 13-14). Institutions apply their sanctions based on the bill (line 15); agents that violate the rules of the institutions they are members of get sanctioned. Finally, the members of the institutions are updated, all based on the agents' requests during that round. This process is repeated for every round until the end of the simulation.

5.1.2 Social players. Social players are agents who participate in the game and have some form of social capital included in their decision-making. Algorithm 2 describes their behaviour in a round of the simulation. When a round starts, players receive the bill for the last meal they participated in and they update their social capital with this information (line 5). This information is also partially reported to the system's reputation. Subsequently, they again update their social capital with all the information received by the institutions they are member of (lines 8-10). All institutions send information about who joined, left, was sanctioned or was expelled. In addition, if an institution called a vote to accept a new member or expel a current member, the vote is sent at this point. Next, the player has a probability r of joining a random institution (line 11). In this cases, a Boltzmann distribution of the institutions based on their social capital value is used to choose one (when choosing which one to leave, the value is inverted using $1 - value_{sc}$). Following this, the agent checks the social capital of each institution of which it is a member. If the value is lower than a threshold, the agent will leave the institution (lines 12-26). After this, the current group dinner is chosen by selecting the one organised by the institution (it is member of) with the higher social capital. Finally, the agent combines the social capital of the participants at the dinner by calculating the average value. If the value is higher than a threshold, the agent orders the inexpensive meal. If not, the expensive meal is chosen.

5.1.3 Other players. The UDD game also includes other kinds of players. *Dominant* players always order an expensive meal. This strategy is strictly dominant and the unique Nash equilibrium. *Random* players choose an *expensive* or *inexpensive* meal by random selection.

<pre> 1: $A \leftarrow$ set of n agents 2: $D \leftarrow$ set of d dinner_groups 3: $t \leftarrow 0$ 4: repeat 5: for each agent $i \in A$ do 6: choose $dinner_group_i$ 7: end for 8: for each agent $i \in A$ do 9: order $meal_i$ 10: end for 11: for each $d \in D$ do 12: $bill_{id} \leftarrow create_bill(d)$ 13: $update_spent(bill_{id})$ 14: $update_joy(bill_{id})$ 15: $institutional_sanctions(bill_{id})$ 16: end for 17: update institution membership 18: $t \leftarrow t + 1$ 19: until $t == T_{lim}$ </pre>	<pre> 1: $E \leftarrow$ sequence of events 2: $O \leftarrow$ set of o orders from last bill 3: $I \leftarrow$ set of i institutions joined by the agent 4: for each order $o \in O$ do 5: create social capital event sc_o 6: report reputation, with probability q 7: end for 8: for each event $e \in E$ do 9: create social capital event sc_e 10: end for 11: join random institution, with probability r 12: for each institution $i \in I$ do 13: if not $cooperate(i)$ then 14: leave i 15: end if 16: end for 17: $gd_t \leftarrow choose_group_dinner$ 18: if $cooperate(gd_t)$ then 19: order inexpensive meal 20: else 21: order expensive meal 22: end if </pre>
---	--

ALGORITHM 2: Unscrupulous diner's dilemma game (left) and social player (right).

5.2 Experiments

Experiment 1: Comparative evaluation. For the first experiment, the simulation was populated with 50 agents of each type, for a total of 150 agents. The average values of 50 simulations were used for the results in this section.

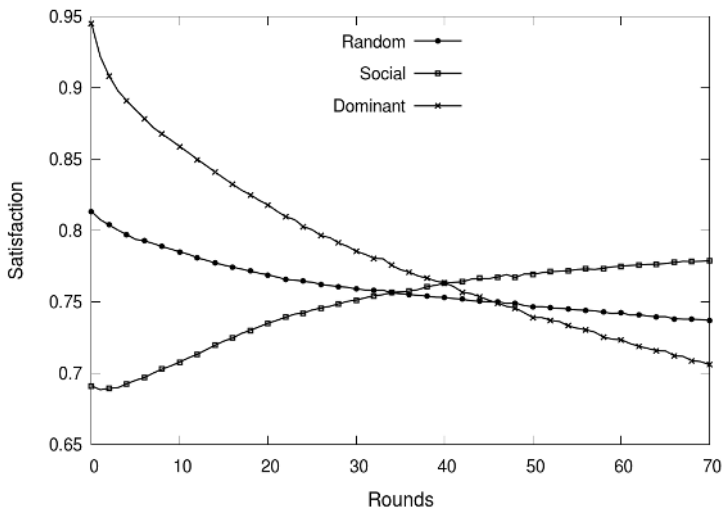


Fig. 6. Players' satisfaction

With this setup, the performance of social capital was investigated in a heterogeneous population of agents, with an equal number of each type. The results are shown in Figure 6.

The plot shows the evolution of the average satisfaction achieved by the different type of agents. At the beginning, the *dominant* players outperform the rest. Due to the lack of interaction between the agents, the forms of social capital do not have any information, and the *social* players share the cost of the expensive meals ordered by the *dominant* and *random* players.

As the game evolves, each form of social capital starts collecting information. Then, *social* players start choosing to group with other *social* players in institutions where sanctions are applied and where repeated violators are expelled. At around round 42, the *social* players outperform the *dominant* players. An important difference can be seen between *social* players using social capital and the rest of the players.

Experiment 2: Scaling with size of group. For the next experiment, simulations were run with different number of players but with the same proportion of *social*, *dominant* and *random* players. With this setup, the effect of scale – in terms of size of the population – on the performance of social capital was investigated. Figure 7 shows the break points when *social* players outperform (in terms of satisfaction) *dominant* and *random* players. These results show that, as the system scales, the rounds to achieve the breakpoint increase less than linearly.

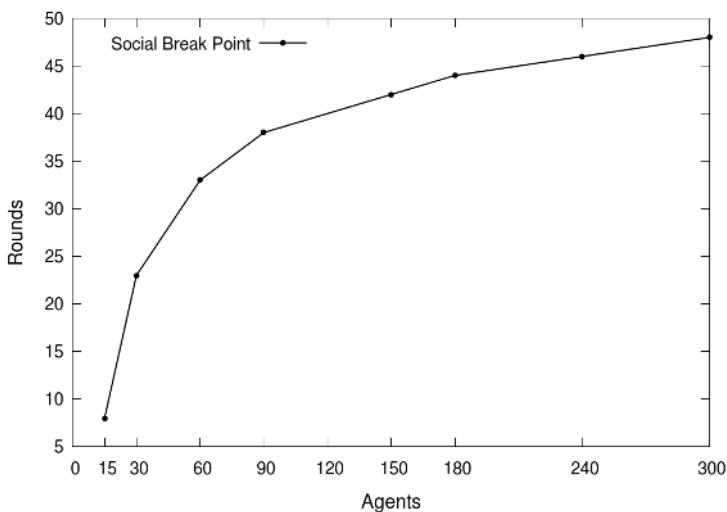


Fig. 7. Social players' break points

6 EXPERIMENTS IN PUBLIC GOODS GAMES

In this section, the role and nature of social capital using agents with learning capabilities in the context of Ostrom's institutional principles is investigated. In particular, the eighth principle: multiple layers of nested enterprises. Both inter-agent and inter-institutional interactions are then situated in an extension of a Public Goods Game (PGG) for an electricity scenario. These type of games are commonly used to study the effects of free-riding in the system, which reduces the positive benefits of cooperation in otherwise unregulated situations.

6.1 Testbed: Electricity public goods game

In a PGG, a set of players must choose how many resources they will provide for a common pool. Once provisioned, the total resources are multiplied by an incremental factor and distributed among the participants. When all the players collaborate and offer resource provisions to the system, they all benefit by receiving more resources. However, a selfish player is tempted to free-ride the system by collecting the distributed resources without provisioning, which reduces the benefits and promotes free-riding among the other players. If this situation is not managed, it will provoke a vicious cycle that leads to the failure of the system. A more detailed analysis of conditional cooperation on PGG can be found in [17].

A variation of a PGG for an Electricity Public Goods Game (EPGG) is proposed. The game consists of a set of n players that possesses a quantity of electricity $g_i \in [0, 1]$, and requires another quantity of electricity $r_i \in [0, 1]$ (which are individually and randomly assigned). Each player decides to contribute a part c_i to the common pool or to store it s_i for the future. The amount of energy contributed and stored must not be greater than the amount available. When the energy is stored, a discount factor α is applied and the available energy at the next time step will be increased with the stored energy, simulating the costs of energy storage infrastructures such as batteries or capacitors. Alternatively, if players contribute to the common pool, another discount factor β is applied, which represents the network infrastructure costs. A more detailed description of the EPGG game can be found in [39].

The EPGG game shares some similarities with the example of the irrigation system described in [section 1](#). In irrigation systems, top-enders must agree with bottom-enders how water is appropriated without depleting it at the top. Similarly, in the EPGG there is a resource surplus at some locations, emulating the top-enders, but this situation is only temporary. Depending on the energy generation conditions, the surplus will indistinctly change between locations; making each location's long-term efficiency codependent on the others.

6.1.1 Step-by-step algorithm. The EPGG game was time-driven implemented. At the beginning of each round, players have a quantity of electricity available and another quantity needed. Common Pool Resource (CPR) must be created to start accepting contributions, and the institutional rules of the CPR must be defined. Players can join and leave any CPR at the beginning of each round. This allows players to leave when they believe that the rules of the current CPR are not favourable and, to create a new one when all CPRs available are disadvantageous. Each player must inform how much will provision and/or it will demand to all the CPRs it is associated with. With this information, CPRs can calculate their surplus or need for electricity and perform the same provision and/or demand to the upper level of CPRs. The decision on how much to provision and demand between nested CPRs is done by one of the players acting as a CPR leader. Following this, by a top-down hierarchy, each CPR allocates the player's demands using the allocation method chosen when the CPR was created. Allocations from top-level CPRs are included as available resources in the lower levels and reallocated recursively until the bottom level. With the allocation results, CPRs update their internal state processing events and sanctions, updating reputations and calling for votes if needed. A description of the CPRs's features follows in the next section. Finally, each player will consume from the grid the required electricity if the electricity allocated from the CPRs and saved is not enough to cover its needs. [Algorithm 3](#) shows a detailed step-by-step procedure of the game.

6.1.2 Social players. [Algorithm 3](#) describes the behaviour of *Social* players in a round of the game. These players implement the social capital framework with a new decision module. At the beginning of each round, *social* players provide feedback about the previous decisions done. The new decision-making module is explained in detail in the the following sections. Later, *social*

<pre> 1: $A \leftarrow$ set of n agents 2: $C \leftarrow$ set of c CPRs 3: $max_lvl \leftarrow$ maximum nested levels 4: $t \leftarrow 0$ 5: repeat 6: for each agent $i \in A$ do 7: calculate g_i {resource available} 8: calculate r_i {resource need} 9: end for 10: for each agent $i \in A$ do 11: manage CPR membership 12: end for 13: for each agent $i \in A$ do 14: $request_provision(i)$ 15: $request_demand(i)$ 16: end for 17: $lvl \leftarrow 0$ 18: repeat 19: for each CPR $c \in C$ where $level = lvl$ do 20: $request_provision(c)$ 21: $request_demand(c)$ 22: end for 23: $lvl \leftarrow lvl + 1$ 24: until $lvl == max_lvl$ 25: repeat 26: for each CPR $c \in C$ where $level = lvl$ do 27: $a_c \leftarrow allocate(method)$ 28: end for 29: $lvl \leftarrow lvl - 1$ 30: until $lvl == 0$ 31: for each $c \in C$ do 32: update reputation 33: update events 34: institutional sanctions 35: process votes 36: end for 37: for each agent $i \in A$ do 38: $request_consumption(i)$ 39: end for 40: $t \leftarrow t + 1$ 41: until $t == T_{lim}$ </pre>	<pre> 1: $E \leftarrow$ sequence of e events 2: $C \leftarrow$ set of c CPR available 3: $J \leftarrow$ set of c CPR joined by the agent 4: $dm \leftarrow$ decision-making 5: $feedback(dm)$ 6: for each event $e \in E$ do 7: update social capital sc_e 8: if $e = vote$ then 9: $vote(dm(v))$ 10: end if 11: end for 12: if $size(J) \neq max$ then 13: $rc \leftarrow random(C)$ 14: if $dm(rc)$ then 15: $join(rc)$ 16: end if 17: end if 18: for each CPR $j \in J$ do 19: if not $dm(j)$ then 20: leave j 21: end if 22: end for 23: $g_i \leftarrow$ resource available 24: $r_i \leftarrow$ resource need 25: for each CPR $c \in J$ do 26: $provision(g_i/\gamma)$ 27: $demand(d_i/\gamma)$ 28: end for 29: $e_i \leftarrow$ calculate total available energy 30: $consume(r_i - e_i)$ </pre>	<p>Repeat for each CPR leading</p>
--	--	--

ALGORITHM 3: Electricity public goods game (left) and social player (right).

players update their social capital with the relevant events from the CPRs. The possible events are: a member joined, left, was sanctioned, or was expelled, what allocations were performed, reputation updates or calls for voting. In the case of voting, the actions are also performed. Management of the CPR membership is done after all the events have been processed. Players will join a new CPR based on the *social capital decision* for that CPR (only when the maximum numbers of CPRs defined by *max* is not reached; this is limited to 12 in this implementation). The same occurs when players must decide if they stay or leave. The following action is to *provision* to and *demand* from to the

common pools. Players will calculate the difference between g and r to decide if they will provision to or demand from the pool (surplus electricity or need for electricity). A factor γ is defined to distribute the provision or demand proportional to all the CPRs the player is participating in. Now all the actions previously explained are repeated for all the CPRs the player is leading, because each CPR becomes a player in the upper level and the leader acts on its behalf. Resource allocation follows, and players receive the electricity assigned to them. Finally, players calculate the difference between the total available electricity and the amount they need; the difference is consumed from the grid.

6.1.3 Other players. The EPGG also includes other kinds of players. *Free-rider* players always save the energy not consumed and participate in all the CPRs without provisioning. Their objective is to reduce the higher ‘cost’ of storing resources by free-riding. They will vote ‘yes’ in any vote, as more players in the CPR will facilitate free-riding and expelling other free-riders allows them to receive more resources. *Random* players participate in the EPGG by deciding all their actions through random selection. This includes joining or leaving CPRs, any vote decision and how much to demand or provision.

6.2 Machine learning for contextualised decision-making

In previous experiments with the framework, social capital indicators were combined using the arithmetic mean to generate a unique output value $o \in [0, 1]$. It has been shown that, when multiple forms of social capital are present, distinct indicators might produce mixed results; using only a few forms of social capital could outperform using more forms in some contexts. Therefore, we propose an alternative approach for decision-making based on machine learning.

Most of the decisions using the framework are *binary*. For example, in the EPGG game, players decide if they will contribute to the common pool, when to join or leave a CPR or to vote on different aspects of CPR management. Essentially, these are *yes/no* decisions. The method chosen is a nonlinear classifier implemented using a Support Vector Machine (SVM) [2]. We used the SVM provided by the machine learning library Encog [23] with the radial basis function as kernel.

To be able to train the SVM without any prior data, a dataset with all the discretised values for the metrics’ inputs and the arithmetic mean was defined. The values were discretised to two decimals and a threshold of 0.5 was defined to split the output values into two categories. This initial configuration enables the decision-making module to initially work as in the previous versions. Every time a decision is taken, the module will save the output (indicators and output) for later learning; this can be adjusted by a learning rate discarding some of the decisions and was initially set to 50%. These decisions will be added to the training set based on the feedback received regarding the outcomes of the decision. If the decision was ‘ok’, it will be included in the training set as it is; in the contrary, it will be included switching the category.

The re-training of the SVM is triggered by a threshold defined by a percentage of values changed – in the size of the training set – mainly to avoid the computing costs associated with a considerably big dataset. By default the SVM is retrained when 1% of the dataset is changed.

6.3 Experiments

In all the following experiments, the cost of saving energy was set at 50% ($\alpha = 0.5$) and the cost of providing to the common pool was set at 10% ($\beta = 0.9$). Those values were chosen based on the assumption that the cost of storage is higher than the cost of the network ($\alpha > \beta$) (see Section 6.1). The values were considered suitable to model an actual electricity market scenario and provide enough ‘difference’ to analyse the results.

Experiment 1: Multiple forms of social capital. For the first experiment, the simulation was populated with *random*, *free-rider* and *social* players. *Social* players using the three forms of social capital and combinations of only two were included. The form *institutions* is present in all of the *social* players, as it is essential to manage and evaluate the CPRs. Thirty players of each type participate in an EPGG game with one location and one level of CPRs. The average values of 10 simulations have been used for the results of these experiments.

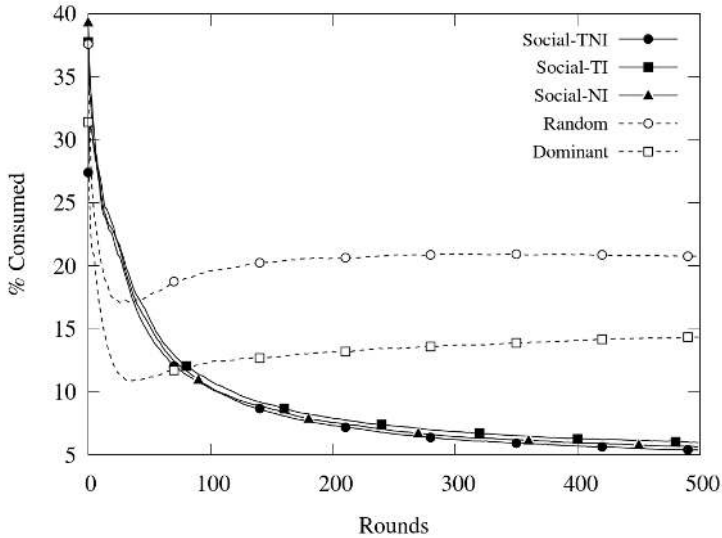


Fig. 8. Consumption using multiple forms of social capital

Figure 8 shows the average electricity consumed from the grid for each type of players. At the beginning, *free-rider* players outperform the rest by free-riding the system. *Social* players have no social capital information and start exploring the different institutional settings. As the game evolves, social capital starts 'learning' which institutions are beneficial and agents start providing to the right common pool. By their interactions in these institutions, they also vote to expel free-riders, maximising their own allocations of resources. *Random* players also benefit by some free-riding (behaving half of the time like *free-rider* players) which is also diminished after a few rounds. Having more forms of social capital is moderately beneficial in this scenario. Players using three forms outperform the others by reducing the institutions that allow free-riding at a faster speed. After 500 rounds, the results are close to the ideal scenarios presented in the previous section. The *free-rider* players slightly benefit by the initial situation, but it shows that is not sustainable as the simulation runs.

Experiment 2: Nested CPRs. For the next experiment, we ran the simulation with 3 locations using *social*, *free-rider* and *random* players and with I, II and III levels of nested CPRs. With this setup, we investigated how beneficial it was to add the structure in terms of reducing the consumption from the main grid.

Figure 9 shows the players' consumption. *Random* players are not included as no changes were manifested. The creation of a nested structure of CPRs allows the three locations to balance their energy surplus and needs, which benefits all the players. However, this gain is not equally

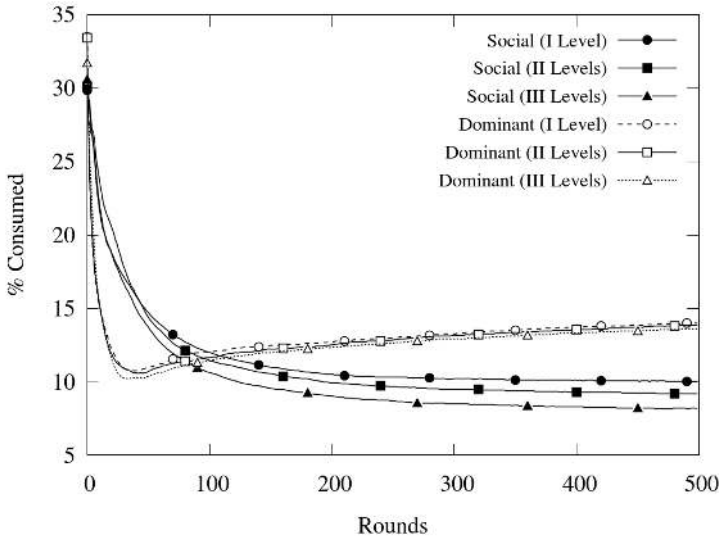


Fig. 9. Consumption for I, II and III levels of nested CPRs

distributed among all the players. *Free-rider* players decrease their consumption marginally, whereas *social* players considerably reduce their energy consumption. Table 2 compares the relative energy consumed at round 50 and 500 for the I, II and III nested levels. The results of *social* players and *free-rider* players at round 50 with one nested level are used as baseline for comparison purposes. As shown in the table, at round 50, *social* players reduce their consumption by 7% when three layers are added. In contrast, the benefit for the *free-rider* players is only 3%. These results are more prominent at round 500, where the reduction is 20% (0.55/0.68) for the first and 2.5% (1.24/1.27) for the second.

Table 2. Electricity consumption evolution in EPGG

Player Type	Round	I-Level	II-Levels	III-Levels
Social	50	1	0.98	0.93
	500	0.68	0.62	0.55
Free-rider	50	1	0.99	0.97
	500	1.27	1.26	1.24

7 APPLICATION SCENARIO

The previous proposed scenario can be grounded in a potential real world application for energy production and consumption self-organisation. A real scenario where home users will be able to produce their own electricity using renewable generators (solar, wind, etc), appliances getting smarter (Internet of Things) and electric cars gaining territory is not in a distant future.

The renewable energy sources often can not be controlled – i.e. solar panels will generate electricity when there is sun and wind turbines will only work with enough wind, but appliances

and cars need electricity based on users' requirements. Possible solutions to this problem include the installation of expensive batteries at home to store the surplus for later use or feeding the electricity surplus into the grid (usually at a low price) and later consume the necessary electricity from the grid (usually at a high price). By implementing a system using Social Capital for collective action (in this case electricity production and consumption balancing), houses could automatically arrange the use of the produced energy. The surplus of one house could provide electricity to another house, and later when the need and surplus are reversed, the first house can get the electricity that requires. By using Social Capital, the system can be decentralised (without a central authority) and resilient, avoiding free-riders that consume the electricity without providing any in the future.

Social Capital appears to be also important in other areas of socio-technical systems where collective action is required, such as sharing economy, peer production, common shared spaces and flat-sharing.

8 RELATED WORK

Cooperative behaviour can be achieved through repeated interactions and the agents' desire to avoid retaliation threats in the case of non-cooperative behaviour [29]. In this case, however, cooperation does not arise from a willingness to do so, but rather from the objective of reducing the probability of being punished.

Norms can also lead to cooperative behaviour. Norms can be learnt from repeated social interactions using (e.g. reinforcement learning) [46] or through aggregation techniques (e.g. ensemble methods) [47]. After a sequence of interactions, agents might learn that cooperation is beneficial. This can then be explicitly stated as a norm, or implicitly internalised by the agents as a social convention. In the case of explicit norms, the compliance to them could be incorporated in the update of the social capital related to institutions.

Given a definition of trust as a willingness to expose oneself to risk, then there is a question: what is to be done when the trust decision is wrong. The common answer in the trust literature is to trash the reputation, which is why the two are found conjoined, as in "trust and reputation". But what we see here is that the reputation (or trustworthiness) is just one element of the social capital attributes which is an input to the trust decision itself, and a wrong decision can update all the different forms of social capital. In addition, *forgiveness* can be construed (in this framework) as another social decision-making process using the social capital indicators as its input.

The notion of trust is widely studied in the field of multi-agent systems [40]. However, most of those works define trust as a function or a value, which typically computes the probability of a beneficial outcome. The use of trust in this context refers to it as the 'glue' that allows the different forms of social capital to be combined together. Therefore, the notion of trust that is most compatible with that presented here is which presents different situations in which it could (objectively) said that *A trusts B* [26]. It then identifies two common features of these situations; firstly, *A* has a belief that there is a rule and, secondly, that *A* has an expectation that *B*'s behaviour will conform to that rule. This is what is seen in this framework and the notion that 'trust' is the glue between social capital and (successful) collective action. The belief that there is a rule is captured by the social capital attribute of institutions, and the expectation of conformance is captured, in effect, by the output of the corresponding decision-making process.

In the fields of self-organising, multi-agent and legal systems, there are many notations (and often associated tools) for describing organisations and institutions. This includes MOISE (an organisational model for MAS) [24], OMACS (organisational model for adaptive computational systems) [11], LAO (logic and organisations) [13], LGI (law governed interaction) [28] and electronic

institutions [19], amongst others. However, to the best of our knowledge, none of these works address the issue of multiple institutions whose interactions create social capital and reasoned with as part of the decision-making processes of the component entities of the institutions (or the entities acting on behalf of (i.e. empowered by) the institutions that they represent).

The issue of multiple interacting institutions has been addressed [8, 35], but these works do not consider the concepts addressed here: norm-governed institutions, Ostrom's institutional design principles, and most significantly, inter-institutional social capital. However, recent work on teams, team structures and team coordination [16] could offer some useful insights into *structuration* (the duality of agency and structure, in that structures are made up of agents, and agents have memory of structures [20]). This could prove highly relevant to the formal conception of nested enterprises and entities, because this memory should perhaps includes elements of the reputational and relational economies which are, in fact, the essence of social capital.

9 FURTHER WORK

In the foundational work on the eight institutional design principles [32], Ostrom did not much highlight the role of social capital. However, a multi-institutional case study was examined in some detail ([32, pp.133–136]), involving water basins in California. It is implicit in her analysis that, in addition to all eight features of successful CPR management systems, elements of 'social capital' were also in play between the various public enterprises that sustained the water basins.

The scenario discussed in this work is intended as an abstraction of the decentralised Community Energy System idea [42], which has also been analysed as a "polycentric set" of different actors in [12]. However, even in abstract terms, it demonstrates two of those elements of social capital evident in the water basin scenario: an interplay of relations at multiple inter-institutional layers; and the learning of effective relationships by individual entities acting on behalf of, or as the representative of, the institutions. What is missing, perhaps, is the interaction between multiple heterogeneous institutions in a "polycentric" set. This missing element, though, points the way towards further research in three directions: the relationship between social capital and the institutional design principles, in particular principle P7; the relationship between social capital and polycentric governance; and the role of social capital in what have been called *holonic institutions* [12].

In the first direction, Ostrom's institutional design principle P7 states that there should be a "minimal recognition of the right to self-organise". Recent work [12] has tried to formalise this principle in terms of empowerment and entitlement relations between *nested* enterprises as a trade-off between the rights and powers of the 'inner' and 'outer' institutions. Thus the principle acts as a kind of 'mediating constraint' between institutions that serves to limit excess of autonomy on the one hand and excessive interference (from the outside) on the other. By contrast, we believe that social capital acts as a kind of 'mediating enabler' that reinforces the trust between institutions that enables successful collective action at the institutional layers. Therefore, an intriguing next step is to converge these experiments in inter-institutional social capital with investigations into the formalisation of Principle P7.

A successful convergence of these ideas would enable further research into the second direction, and an analysis of a polycentric set of institutions. In the experiments described in this paper, the institutions at each of the nested layers are essentially homogenous, especially in terms of their intended goals. We also need to set up and analyse a set of nested institutions where there are different (and even conflicting) aims, ownership models, and participation strategies. This will allow modelling of scenarios where polycentricity is clearly an important feature, such as the

irrigation systems studied by Ostrom, but also smart grids or community energy systems, and other large-scale infrastructures.

Finally, this study of social capital and polycentric governance would contribute to the third direction of research, namely that of holonic institutions. The idea of holonic institutions is to converge the benefits of voluntary regulation according to mutually agreed rules (as found in institutions) with the multi-scale, multi-criteria optimisation offered by holonics. The critical question here is the extent to which social capital can contribute to innovation in, of and between institutions. This is, arguably, essential to any development in (for example) Smart Cities, where the building of such a city from scratch is effectively impossible, and needs to proceed from a potential melange of pre-existing organisations, institutions and other vested interests. We believe that social capital is not just the precondition to successful collective action, but also the precondition for successful collective (institutional) innovation.

10 SUMMARY AND CONCLUSIONS

We presented the social capital framework that defines data structures for representing attributes of agents, the processes for updating, evaluating and decision-making with electronic social capital, and providing for effective action selection across multiple individual and institutional collective action situations. The implementation of the framework on top of Presage2 was described and it is distributed as a module of the simulation platform under GPL license.

Furthermore, the experiments in 2-player games show that the use of social capital clearly benefits the agents individually, as well as the system as a whole. One of the main effects of using social capital is the facilitation to achieve win-win situations, where two agents involved in a pairwise interaction benefit from behaving cooperatively. It has also been observed that social capital acts as a catalyst for self-organisation: agents decide with whom to interact according to their social capital, and they also use the social capital information to join or leave institutions. Moreover, the experiments in n -player games show that social capital: optimises the outcomes (in terms of long-term satisfaction and utility) of collective decision-making in competitive environments, compared to alternative simplistic strategies; reduces the complexity of that decision-making; and scales with the size of the population (contra [31]), because its complexity is independent of population size.

Throughout the experiments, it was observed that social capital facilitates the self-organisation of 'like-minded' individuals into groups, and incentivises those individuals, who may initially be disinclined to conform to the group norms. In conclusion, it is this self-organisation into groups that may be the most significant outcome of these experiments, especially if we understand the organisation of these groups in terms of *communities*.

It has been argued that the value of communities is that they can resolve certain types of collective action problems, or reap the benefits from other forms of collective action, that are otherwise resistant to purely market-based or (top-down) policy-based solutions [4]. This is because communities can leverage relational information, as provided by social capital, in ways that market-based economies, if they are solely reliant on transactional information, cannot.

It is hence predicted that self-organised community systems will be of increasing importance as a mechanism for solving collective action problems in the digital society and achieving satisfactory outcomes for citizens. Electronic forms of social capital will be an essential feature of such systems, along with self-governance and collective attention [12].

At last, we reported new experimental results which show the significance of inter-institutional social capital in the self-organisation of sustainable structures of nested enterprises, which: facilitates transmission of prosocial behaviour from 'lower' levels to 'higher'; beneficially affects the

development and maintenance of the nested structure; and assists the members' transition amongst the structure.

Those last experiments showed how the aggregation of agents into CPR-institutions (that are embedded in a system of nested enterprises) promotes prosocial behaviour which can be propagated throughout the system by inter-institutional social capital. We would argue that the (formal) representation of social capital is a critical aspect to effective coordination in open multi-agent systems, especially where issues of scalability demand that the system organises itself into a layered structure of nested enterprises.

In conclusion, this work proposes an alternative solution using Electronic Social Capital to represent and reason with qualitative, instead of traditional quantitative, values. This solution could be embedded into socio-technical systems to incentivise collective action without commodifying the resources or actions in the system.

11 ACKNOWLEDGEMENTS

This work is partially supported by the UK EPSRC Grand Challenge Project Autonomous Power System. The Electronic Social Capital Framework has been a work in progress developed in [36–39]. Experiments in section 4 have been synthesised from [36], experiments in section 5 have been synthesised from [38], and experiments in section 6 have been synthesised from [39]. This article is an extended synthesis of Petruzzi et al. [36–39].

REFERENCES

- [1] Ken Binmore. *Natural Justice*. Oxford University Press, 2005.
- [2] Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, COLT '92, pages 144–152, New York, NY, USA, 1992. ACM.
- [3] Pierre Bourdieu and Loïc J.D Wacquant. An invitation to reflexive sociology. *Chicago University Press*, 1992.
- [4] Samuel Bowles and Herbert Gintis. Social capital and community governance. *The Economic Journal*, 112(483): F419–F436, 2002.
- [5] Hendrik Van Brussel, Jo Wyns, Paul Valckenaers, Luc Bongaerts, and Patrick Peeters. Reference architecture for holonic manufacturing systems: {PROSA}. *Computers in Industry*, 37(3):255 – 274, 1998.
- [6] Estelle Cantillon and Martin Pesendorfer. Chapter 22 auctioning bus routes: The london experience, 2004.
- [7] Yann Chevaleyre, Paul E Dunne, Ulle Endriss, Jérôme Lang, Michel Lemaitre, Nicolas Maudet, Julian Padget, Steve Phelps, Juan A Rodriguez-Aguilar, and Paulo Sousa. Issues in multiagent resource allocation. *Informatica*, 30(1), 2006.
- [8] Owen Cliffe, Marina De Vos, and Julian Padget. Specifying and reasoning about multiple institutions. In *Coordination, Organizations, Institutions, and Norms in Agent Systems II*, number 4386 in LNCS, pages 67–85, 2006.
- [9] James S. Coleman. Social capital in the creation of human capital. *American Journal of Sociology*, 94:S95–S120, 1988.
- [10] National Research Council. *Visionary Manufacturing Challenges for 2020*. The National Academies Press, Washington, DC, 1998.
- [11] Scott A. DeLoach. OMACS: A framework for adaptive, complex systems. In V. Dignum, editor, *Multi-Agent Systems: Semantics and Dynamics of Organisational Models*. IGI Global, 2009.
- [12] Ada Diaconescu and Jeremy Pitt. Holonic institutions for multi-scale polycentric self-governance. In *COIN'14 Proceedings*, number 9372 in LNAI. Springer, 2015.
- [13] Virginia Dignum and Frank Dignum. A logic of agent organizations. *Logic Journal of the IGPL*, 20(1):283–216, 2012.
- [14] Rino Falcone and Cristiano Castelfranchi. Trust and relational capital. *Comput. Math. Organ. Theory*, 17(2):179–195, 2011.
- [15] Rainer Feldmann, Martin Gairing, Thomas Lücking, Burkhard Monien, and Manuel Rode. *Selfish Routing in Non-cooperative Networks: A Survey*, pages 21–45. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.
- [16] Mariana Ramos Franco, Gustavo A. L. Campos, and Jaime Sim ao Sichman. An empirical approach for relating environmental patterns with agent team compositions. In *COIN@AAMAS2016 Pre-proceedings*, 2016.
- [17] Simon Gächter. Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications. In *Psychology and Economics: A Promising New Cross-Disciplinary Field*, pages 19–50. MIT Press, Cambridge, MA, 2007.
- [18] Aram Galstyan, Karl Czajkowski, and Kristina Lerman. Resource allocation in the grid with learning agents. *Journal of Grid Computing*, 3(1):91–100, 2005.

- [19] Andrés García-Camino, Pablo Noriega, and Juan Antonio Rodríguez-Aguilar. Implementing norms in electronic institutions. In *Proceedings of the Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 667–673. ACM Press, 2005.
- [20] Anthony Giddens. *The Constitution of Society*. University of California Press, 1984.
- [21] Natalie S. Glance and Bernardo A. Huberman. The dynamics of social dilemmas. *Scientific American*, 270(3):76–81, 1994.
- [22] Garrett Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.
- [23] Jeff Heaton. Encog: Library of interchangeable machine learning models for java and c. *Journal of Machine Learning Research*, 16:1243–1247, 2015.
- [24] Jomi Hübner, Jaime Sichman, and Olivier Boissier. Developing organised multiagent systems using the MOISE⁺ model. *IJAOS*, 1(3/4):370–395, 2007.
- [25] Leonid Hurwicz. The design of mechanisms for resource allocation. *The American Economic Review*, 63(2):1–30, 1973.
- [26] Andrew J. I. Jones. On the concept of trust. *Decision Support Systems*, 33(3):225–232, 2002.
- [27] Samuel Macbeth, Didac Busquets, and Jeremy Pitt. Principled operationalization of social systems using presage2. In D. Gianni, A. D’Ambrogio, and A. Tolk, editors, *Modeling and Simulation-Based Systems Engineering Handbook*. CRC Press, 2014.
- [28] Naftaly Minsky. *Law-Governed Interaction (LGI): A Distributed Coordination and Control Mechanism*. Rutgers University, 2005.
- [29] Roger B. Myerson. *Game theory - Analysis of Conflict*. Harvard University Press, 1997.
- [30] John F. Nash. Equilibrium points in n -person games. *Proc. of the National Academy of Sciences*, 36:48–49, 1950.
- [31] Mancur Olson. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Number v.124 in Harvard economic studies. Harvard University Press, 1965.
- [32] Elinor Ostrom. *Governing the Commons*. Cambridge University Press, 1990.
- [33] Elinor Ostrom and TK. Ahn. *Foundations of Social Capital*. Edward Elgar Publishing, 2003.
- [34] Christos H. Papadimitriou and Tim Roughgarden. Computing equilibria in multi-player games. In *Proc. 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 82–91, 2005.
- [35] Jigar Patel, W. T. Luke Teacy, Nicholas R. Jennings, Michael Luck, Stuart Chalmers, Nir Oren, Timothy J. Norman, Alun Preece, Peter M. D. Gray, Gareth Shercliff, Patrick J. Stockreisser, Jianhua Shao, W. Alex Gray, Nick J. Fiddian, and Simon Thompson. Agent-based virtual organisations for the grid. *Multiagent and Grid Systems*, 1(4):237–249, 2005.
- [36] Patricio E. Petruzzi, Didac Busquets, and Jeremy Pitt. Experiments with social capital in multi-agent systems. In *17th International Conference on Principles and Practice of Multi-Agent Systems*, pages 18–33, 2014.
- [37] Patricio E. Petruzzi, Didac Busquets, and Jeremy Pitt. Social capital as a complexity reduction mechanism for decision making in large scale open systems. In *8th IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, pages 145–150, 2014.
- [38] Patricio E. Petruzzi, Didac Busquets, and Jeremy Pitt. A generic social capital framework for optimising self-organised collective action. In *9th IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, pages 21–30, 2015.
- [39] Patricio E. Petruzzi, Jeremy Pitt, and Didac Busquets. Inter-institutional social capital for self-organising ‘nested enterprises’. In *10th IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, pages 90–99, 2016.
- [40] Isaac Pinyol and Jordi Sabater-Mir. Computational trust and reputation models for open multi-agent systems: a review. *Artificial Intelligence Review*, 40(1):1–25, 2013.
- [41] Jeremy Pitt and Andrzej Nowak. The reinvention of social capital for socio-technical systems. *IEEE Technol. Soc. Mag.*, 33(1):27–33, 2014.
- [42] Jeremy Pitt, Didac Busquets, Aikaterini Bourazeri, and Patricio E. Petruzzi. Collective intelligence and algorithmic governance of socio-technical systems. In Daniele Miorandi, Vincenzo Maltese, Michael Rovatsos, Anton Nijholt, and James Stewart, editors, *Social Collective Intelligence: Combining the Powers of Humans and Machines to Build a Smarter Society*, pages 31–50. Springer Verlag, 2014.
- [43] Robert D. Putnam. The prosperous community: Social capital and public life. *The American Prospect*, pages 35–42, 1993.
- [44] Nicholas Rescher. *Distributive Justice*. Bobbs-Merrill, 1966.
- [45] Theodore W. Schultz. Investment in human capital. *The American economic review*, 51(1):1–17, 1961.
- [46] Toshiharu Sugawara. Emergence and stability of social conventions in conflict situations. In *22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 371–378, 2011.
- [47] Chao Yu, Minjie Zhang, Fenghui Ren, and Xudong Luo. Emergence of social norms through collective learning in networked agent societies. In *12th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 475–482, 2013.