

The Axiomatisation of Socio-Economic Principles for Self-Organising Systems

Jeremy Pitt and Julia Schaumeier
Department of Electrical & Electronic Engineering
Imperial College London
Exhibition Road, London, SW7 2BT UK
Email: {j.pitt, j.schaumeier09}@imperial.ac.uk

Alexander Artikis
Institute of Informatics & Telecommunications
National Centre for Scientific Research “Demokritos”
Athens 15310, Greece
Email: a.artikis@iit.demokritos.gr

Abstract—We are interested in engineering for open, embedded and resource-constrained systems, which have applications in ad hoc, sensor and opportunistic networks. In such systems, there is decentralised control, competition for resources and an expectation of both intentional and unintentional errors. The ‘optimal’ distribution of resources is then less important than the ‘robustness’ or ‘survivability’ of the distribution mechanism, based on collective decision-making and tolerance of unintentional errors. We therefore seek to model resource allocation in the network as a common pool resource management problem, and apply a formal characterisation of Ostrom’s socio-economic principles for building enduring institutions. This paper presents a complete axiomatisation in the Event Calculus of six of Ostrom’s eight principles, describes a preliminary testbed for experimenting with the axiomatisation, and considers the work from a methodological perspective of sociologically-inspired computing for self-organising systems.

Keywords—Self-Organisation, Socio-Economics, Norms.

I. INTRODUCTION

Open embedded systems consist of heterogeneous components of unknown provenance that are coordinating their behaviour in the context of an environment which may be perturbed by outside events. Such systems arise in a new class of wireless network, for example mobile ad hoc, opportunistic, sensor and vehicular networks; in service-oriented systems like virtual organisations and cloud computing applications; and increasingly in demand-side infrastructure management, for water, energy, and so on.

All these applications share a number of features. Primarily, decision-making is too fast, frequent and complicated for operator intervention and/or provision: therefore the system has to be able to operate autonomously. Being open, there is no central controller, no common goal and no common knowledge: therefore collective decisions must be made in the face of both uncertainty and possibly conflicting opinions and requirements. Openness also implies the system must operate in expectation of error, non-compliance to the specification and other sub-ideal behaviour, including both intentional and unintentional violations; but the system components cannot expect any level of cooperation, i.e. that appropriate action will be taken to recover from errors or sub-ideal states. Finally, the systems is resource-constrained,

and the components are required to share and appropriate resources in order to satisfy individual goals.

The ‘optimal’ distribution of resources is then less important than the ‘robustness’ or ‘survivability’ of the distribution mechanism, based on collective decision-making and tolerance of unintentional errors. Accordingly, we examine three propositions: firstly (**p1**), that open, embedded and resource-constrained systems can be considered from the perspective of institutions for management of common pool resources (CPR); secondly (**p2**), that socio-economic principles for enduring institutions can be considered from the perspective of norm-governed systems, and can be axiomatised using action languages used in Artificial Intelligence for reasoning about action, agency and norms; and thirdly (**p3**), that such an axiomatisation can be used as an executable specification for systematic experiments to test whether these principles are necessary and sufficient conditions for *enduring* institutions.

This paper is organised as follows. In the next section, we review the background work against which these propositions are to be tested, namely the work on CPR management of Ostrom [1], the approach to norm-governed systems specification [2], and action languages [3]. In Section III, we give an abstract specification of an institutional resource allocation system that will be used to illustrate the axiomatisation given in Section IV. Section V considers the work from a methodological perspective, and describes preliminary work in developing an experimental testbed for investigating the axiomatisation with respect to ‘endurance’. After discussing related and further work in Section VI, we summarise in Section VII with some comments on sociologically-inspired computing for self-organising systems.

II. BACKGROUND

A. Governing the Commons

Ostrom [1] argued that management of common pool resources (CPR) need not lead to a ‘tragedy of the commons’ as predicted by game theory, and that there was an alternative to privatisation or centralised control of the resource. She observed that in many cases, for example in Spain, Switzerland, Japan and the US, communities were able to manage their own affairs by defining *institutions* to govern their commons. We define an ‘institution’ as a set of working

rules used to determine who is eligible to make decisions in what area, what actions are allowed or constrained, what aggregation rules are used, and so on.

However, Ostrom also observed that there were occasions when the institutions were *enduring*, and others where they were not. Accordingly, eight design principles were identified for *self*-management of common pool resources (CPR) to endure [1, p. 90]. These were:

- 1) Clearly defined boundaries: those who have rights or entitlement to appropriate resources from the CPR are clearly defined, as are its boundaries;
- 2) Congruence between appropriation and provision rules and the state of the prevailing local environment;
- 3) Collective choice arrangements: in particular, those affected by the operational rules participate in the selection and modification of those rules;
- 4) Monitoring, of both state conditions and appropriator behaviour, is by appointed agencies, who are either accountable to the resource appropriators or are appropriators themselves;
- 5) A flexible scale of graduated sanctions for resource appropriators who violate communal rules;
- 6) Access to fast, cheap conflict resolution mechanisms;
- 7) Existence of and control over their own institutions is not challenged by external authorities; and
- 8) Systems of systems: layered or encapsulated CPRs, with local CPRs at the base level.

B. Norm-Governed Systems

The study of legal, social and organisational systems has often been formalised in terms of norm-governed systems. We maintain the standard and long established distinction between physical capability, institutionalised power, and permission (see e.g. [4] for illustrations of this distinction). Accordingly, a specification of a norm-governed system expresses five aspects of social constraint: (i) physical capabilities; (ii) institutionalised powers; (iii) permissions, prohibitions and obligations of the agents; (iv) sanctions and enforcement policies that deal with the performance of prohibited actions and non-compliance with obligations; and (v) designated roles of empowered agents.

The first aspect of a specification of social constraints concerns the externally observable physical capabilities of a society's members.

The term institutional, or 'institutionalised', power refers to the characteristic feature of an institution, whereby designated agents, often acting in specific roles, are empowered to create or modify facts of special significance in that institution – *institutional facts* [5] – usually by performing a specified kind of act, in certain cases a speech act.

The next aspect of a specification of social constraints provides the definitions of permitted, prohibited and obligatory actions. These definitions are application-specific. In some cases, we might want to associate institutional powers

with permissions. In some societies, for example, an agent is permitted to perform an action if that agent is empowered to perform that action. According to this definition, an agent is always permitted to exercise its institutional powers. In other societies the relationship is stronger: an agent is permitted to perform an action *if and only if* it is empowered to perform that action. In general, however, there is no standard, fixed relationship between powers and permissions.

Determining what actions are permitted, prohibited or obligatory enables the classification of the behaviour of individual agents and the society as a whole into categories such as 'social' or 'anti-social', 'acceptable' or 'unacceptable', and so on. The behaviour of an agent might be considered 'anti-social' or 'unacceptable' if that agent performs certain forbidden actions or does not comply with its obligations.

Social constraints also express the sanctions and enforcement policies that deal with 'anti-social' or 'unacceptable' behaviour. We are concerned with two issues: firstly, when is an agent sanctioned, and secondly, what is the penalty that the agent faces, in the case that it does get sanctioned.

Finally, we associate a social role with a set of *preconditions* that agents must satisfy in order to be eligible to occupy that role, and a set of *constraints* that govern the behaviour of the agents once they occupy the role (cf. [6]). In general, an agent may be assigned a role if the following two criteria are met: firstly, that the agent satisfies the role preconditions, and secondly that the assignment of the role to the agent does not violate the *role-assignment constraints*.

C. Action Languages: the Event Calculus (EC)

To specify the axiomatisation of Ostrom's socio-economic principles of enduring institutions in terms of the concepts of a norm-governed system, we use a language that enables representing and reasoning about action, agency, social constraints and change. There are various alternative languages; we use the Event Calculus (EC) [3] for clarity of exposition and for use in executable specification.

The EC is a logic formalism for representing and reasoning about actions or events and their effects. The EC is based on a many-sorted first-order predicate calculus. For the version used here, the underlying model of time is linear; we will use non-negative integer time-points (although this is not an EC restriction). We do not assume that time is discrete but we do impose a relative/partial ordering for events: for non-negative integers, $<$ is sufficient.

An *action description* in EC includes axioms that define: the action occurrences, with the use of happensAt predicates; the effects of actions, with the use of initiates and terminates predicates; and the values of the fluents, with the use of initially and holdsAt predicates. Table I summarises the main EC predicates. Variables, that start with an upper-case letter, are assumed to be universally quantified unless otherwise indicated. Predicates, function symbols and constants start with a lower-case letter.

Table I: Main Predicates of the Event Calculus.

Predicate	Meaning
Act happensAt T	Action Act occurs at time T
initially $F = V$	The value of fluent F is V at time 0
$F = V$ holdsAt T	The value of fluent F is V at time T
Act initiates $F = V$ at T	The occurrence of action Act at time T initiates a period of time for which the value of fluent F is V
Act terminates $F = V$ at T	The occurrence of action Act at time T terminates a period of time for which the value of fluent F is V

Where F is a *fluent*, which is a property that is allowed to have different values at different points in time, the term $F = V$ denotes that fluent F has value V . Boolean fluents are a special case in which the possible values are *true* and *false*. Informally, $F = V$ holds at a particular time-point if $F = V$ has been *initiated* by an action at some earlier time-point, and not *terminated* by another action in the meantime.

Events initiate and terminate a period of time during which a fluent holds a value continuously. Events occur at specific times (when they *happen*). A set of events, each with a given time, is called a *narrative*.

The utility of the EC comes from being able to reason with narratives. Therefore the final part of an EC specification is the domain-independent ‘engine’ which computes what fluents hold, i.e. have the value *true* in the case of boolean fluents, or what value a fluent takes, for each multi-valued fluent. This can be used to compute a ‘state’ of the specification, which changes over time, and includes the roles, powers, permissions and obligations of agents.

III. RESOURCE ALLOCATION SYSTEMS

Consider an abstract specification which can be instantiated for many different types of open, embedded system which require some partition of a divisible good. For example, consider a water management system with a resource (a reservoir of water) and a set of appropriators (agents) who draw water from the reservoir.

This can be formulated as a resource allocation system defined at time t by $\langle \mathcal{A}, P, m \rangle_t$, where \mathcal{A} is the set of appropriators (agents); P is the pooled resources (a divisible good); and m is the resource allocation, where at each time t , m_t is a mapping from members of \mathcal{A} to a fraction of P , $m_t : \mathcal{A} \mapsto [0, P]$ such that $\sum_{a \in \mathcal{A}} m_t(a) \leq P$. There are various ways of determining m_t , for example by auctions [7], or cake-cutting algorithms [8].

Now let \mathcal{I} be an institutional resource allocation system defined at time t by:

$$\mathcal{I}_t = \langle \mathcal{A}, \epsilon, L, m \rangle_t$$

where (omitting the subscript t when obvious from context):

- \mathcal{A} is the set of agents;

- ϵ is the environment, a pair $\langle Bf, Ff \rangle$ with Bf the set of ‘brute’ facts whose values are determined by the physical state, including the resource(s) P to be allocated, and Ff a set of fluents, or ‘institutional’ facts, with values V determined by conventional state;
- L , is the resource allocation ‘legislature’, the set of rules by which agents are allocated resources; and
- m is a partial function $\mathcal{A} \mapsto [0, P]$ which specifies the amount of resources allocated to each agent a in \mathcal{A} .

Clearly \mathcal{I} subsumes $\langle \mathcal{A}, P, m \rangle$. Each system determines m_t for each t , but \mathcal{I} will use the rules in L and the state of ϵ (at time t) to determine m_t , rather than an auction (say).

Following Ostrom [1, pp. 52-53], the rules in L can be divided into three types – *OC*, *SC* and *CC* – where *OC* = *operational choice* rules, concerned with appropriation, monitoring and enforcement; *SC* = (*social*) *collective choice* rules, concerned with determining the operational rules, adjudication, etc.; and *CC* = *constitutional choice* rules, concerned with eligibility and determining the social collective choice rules. Then suppose that we have two types of method, *raMethod* and *wdMethod*, where *raMethod* is the type of resource allocation method, e.g. random, queue, ration and priority; and *wdMethod* is the type of winner determination method, e.g. plurality, runoff, borda, etc.

\mathcal{I} operates in time slices. During a time slice an agent a will try to appropriate resources as a fraction of P , by making a demand $d(a)$, receiving an allocation $m(a)$, and making an appropriation. Let $v_a^1(ra)$ be the preference of each agent a for every resource allocation method $ra \in raMethod$, and let $v_a^2(wd)$ be the preference of each agent a for every winner determination method $wd \in wdMethod$. Assuming the winner determination method for the constitutional choice rule is fixed, i.e. some constant $k \in wdMethod$ (although it need not be, cf. [9]), then a constitutional choice rule $ccr \in CC$ maps a set of expressed preferences to a winner determination method according to k ; a social collective choice rule $scr \in SC$ maps another set of expressed preferences to a resource allocation method according to this winner determination method; and an operational choice rule $ocr \in OC$ maps a set of demands ($d : \mathcal{A} \mapsto [0, P]$) to a set of allocations ($m : \mathcal{A} \mapsto [0, P]$) according to this resource allocation method:

$$\begin{aligned} ccr &: \{v_a^2(\cdot)\}_{a \in \mathcal{A}} \times k \rightarrow wdMethod \\ scr &: \{v_a^1(\cdot)\}_{a \in \mathcal{A}} \times wdMethod \rightarrow raMethod \\ ocr &: (\mathcal{A} \mapsto [0, P]) \times raMethod \rightarrow (\mathcal{A} \mapsto [0, P]) \end{aligned}$$

A *valid* allocation satisfies the constraints firstly, that $\sum_{a \in \mathcal{A}} m(a) \leq P$, and secondly that for all $a \in \mathcal{A}$, $m(a) \leq d(a)$. It is a violation of the rules of \mathcal{I} to compute an invalid allocation or for an agent to appropriate more resources than it is allocated.

This validates the proposition **p1** that open, embedded and resource-constrained systems can be considered from

Table II: Fluents for the Axiomatisation.

Fluent	ValueRange(type)
$role_of(A, I)$	$\{head, member, monitor, \dots\}$
$res_alloc_meth(I)$	$\{ration, queue, \dots\}$
$win_det_meth(I)$	$\{plurality, runoff, \dots\}$
$offences(A, I)$	<i>integer</i>
$sanction_level(A, I)$	<i>integer</i>
$pow(Agent, Action)$	<i>boolean</i>
$per(Agent, Action)$	<i>boolean</i>
$obl(Agent, Action)$	<i>boolean</i>

the perspective of institutions for management of common pool resources. In the next section, we examine in detail the proposition **p2** that the institutional policies in L can encapsulate Ostrom's principles for enduring institutions, if considered from the perspective of norm-governed systems and axiomatised using an action language.

IV. AXIOMATISING OSTROM'S PRINCIPLES

In this section, we will consider in detail the formalisation of six of the eight socio-economic principles defined by Ostrom [1], as they were reviewed in Section II-A, using the concepts identified in Section II-B, and in terms of Event Calculus axioms, as described in Section II-C.

Some of the fluents F in whose values we are interested are shown in Table II. These fluents record the roles that agents occupy. The basic role is *member*, and there are other roles with special significance, such as *head* and *monitor*. We assume that these latter roles require membership and are exclusive, but different institutions can have different rules relating to role assignment. Non-members have no role in the institution, i.e. *null*.

Two fluents record the number of rule violations *offences* and the sanctions imposed *sanction_level*. The multi-valued fluents *res_alloc_meth* determines which resource allocation method is selected, and *win_det_meth* the winner determination method to select it. The final three fluents record the (institutionalised) powers, permissions and obligations of each agent. Further fluents are used for defining other data structures (queues and lists) but are not listed here.

A. Principle 1: Clearly Defined Boundaries

Principle 1 states that those who have rights or entitlement to appropriate resources from the CPR are clearly defined, as are its boundaries.

There are three aspects to axiomatising this principle: firstly, separating those who have rights and entitlements from those who do not; secondly, expressing precisely what those rights and entitlements are; and thirdly defining its boundaries. As the last of these is to do with physical constraints, we will not consider it further here.

The first issue can be dealt with using role-based access control (e.g. [10]) and defining a role-assignment protocol, in order to distinguish between those agents in \mathcal{A} that are

members of the institution and those which are not. We define a fluent $role_of(A, I) = member$ which holds (is true) if A is a member of I and does not hold otherwise.

An agent can apply for membership to an institution I if it is a non-member and it qualifies for consideration for membership. This qualifying check is domain-specific, and can be as a simple (i.e. no constraints) or as complicated as necessary. It is not affected by any of the actions that can be performed by agents, and is therefore a rigid designator.

$$\begin{aligned}
 apply(A, I) & \text{ initiates} \\
 applied(A, I) & = true \quad \text{at } T \leftarrow \\
 role_of(A, I) & = null \quad \text{holdsAt } T \wedge \\
 qualifies(A, I, member) &
 \end{aligned}$$

An agent A can be included as a member if agent C performs the designated action (*include*), and C is empowered to do so if A applied to I , A was approved, and agent C was indeed the empowered agent, by virtue of occupying the role *head* in institution I .

$$\begin{aligned}
 include(C, A, I) & \text{ initiates} \\
 role_of(A, I) & = member \quad \text{at } T \leftarrow \\
 pow(C, include(C, A, I)) & = true \quad \text{holdsAt } T \\
 pow(C, include(C, A, I)) & = true \quad \text{holdsAt } T \leftarrow \\
 applied(A, I) & = true \quad \text{holdsAt } T \wedge \\
 approved(A) & = true \quad \text{holdsAt } T \wedge \\
 role_of(C, I) & = head \quad \text{holdsAt } T
 \end{aligned}$$

A mechanism for exclusion complements that for inclusion; the issue of sanctions is discussed later under Principle 5.

$$\begin{aligned}
 exclude(C, A, I) & \text{ initiates} \\
 role_of(A, I) & = null \quad \text{at } T \leftarrow \\
 pow(C, exclude(C, A, I)) & = true \quad \text{holdsAt } T \\
 pow(C, exclude(C, A, I)) & = true \quad \text{holdsAt } T \leftarrow \\
 sanction_level(A, I) & = 2 \quad \text{holdsAt } T \wedge \\
 role_of(C, I) & = head \quad \text{holdsAt } T
 \end{aligned}$$

The second issue, that of formally characterising the rights and entitlements, will be discussed when we consider Principle 3, after the axiomatisation of Principle 2.

B. Principle 2: Congruence

Principle 2 states that there should be congruence between appropriation and provision rules and the state of the prevailing local environment.

This requires axioms for valid demands; axioms for the power of the *head* agent to grant allocations which are dependent on the state of the local environment; and axioms concerning the rights and entitlements of the agents.

Agents make a demand R for resources, where R is some fraction of the pooled resources P . To make a *valid* demand

in I , an agent must be empowered, and it is empowered if it is a member of I , it has not made demand in this time slice, and it has not been sanctioned (see Principle 5). This enforces the ‘boundary’ conditions from Principle 1, as any non-member or excluded member cannot make valid demands (their demand actions are ‘noise’). A valid demand also adds a demand to the demand queue fluent.

$demand(A, R, I)$ initiates
 $demand_q(I) = R$ at $T \leftarrow$
 $\mathbf{pow}(A, demand(A, R, I)) = true$ holdsAt T
 $demand(A, R, I)$ initiates
 $demand_q(I) = Q ++ [(A, R)]$ at $T \leftarrow$
 $demand_q(I) = Q$ holdsAt $T \wedge$
 $\mathbf{pow}(A, demand(A, R, I)) = true$ holdsAt T
 $\mathbf{pow}(A, demand(A, R, I)) = true$ holdsAt $T \leftarrow$
 $role_of(A, I) = member$ holdsAt $T \wedge$
 $demand_q(I) = 0$ holdsAt $T \wedge$
 $sanction_level(A, I) = 0$ holdsAt T

Now recall that access to resources in our example depends on resource availability. There were four levels, random, queue, ration, and priority. These determine the conditions on the power of the *head* to allocate resources.

$\mathbf{pow}(C, allocate(C, A, R, I)) = true$ holdsAt $T \leftarrow$
 $demand_q(I) = R$ holdsAt $T \wedge$
 $demand_q(I) = [(A, R) \mid Rest]$ holdsAt $T \wedge$
 $role_of(C, I) = head$ holdsAt $T \wedge$
 $res_alloc_meth(I) = queue$ holdsAt T
 $\mathbf{pow}(C, allocate(C, A, R', I)) = true$ holdsAt $T \leftarrow$
 $demand_q(I) = R$ holdsAt $T \wedge$
 $demand_q(I) = [(A, R) \mid Rest]$ holdsAt $T \wedge$
 $role_of(C, I) = head$ holdsAt $T \wedge$
 $res_alloc_meth(I) = ration(R'')$ holdsAt $T \wedge$
 $((R > R'' \wedge R' = R'') \vee (R \leq R'' \wedge R' = R))$

The last line says that either the agent demanded more than the ration, in which case all it gets is the ration; or it demanded less than (or equal to) the ration, in which case it gets what it demanded. The axioms for the other resource allocation methods are similar and are omitted.

Allocation is closely associated with the issue of rights and entitlements. It has been argued [10] that in access control and resource allocation situations of the type being analysed here, where there may be both ‘valid’ and ‘invalid’ demands, the notions of permission and prohibition are insufficient, and a notion of *entitlement* is required.

As in [11], for an agent that is entitled to be allocated resources, there is a corresponding obligation on another

agent – the one occupying the role of *head* – to grant valid demands, as determined by the allocation method, e.g.:

$\mathbf{obl}(C, allocate(C, A, R, I)) = true$ holdsAt $T \leftarrow$
 $demand_q(I) = R$ holdsAt $T \wedge$
 $demand_q(I) = [(A, R) \mid Rest]$ holdsAt $T \wedge$
 $role_of(C, I) = head$ holdsAt $T \wedge$
 $res_alloc_meth(I) = queue$ holdsAt T

and similarly for the other appropriation rules.

C. Principle 3: Collective Choice Arrangements

Principle 3 concerns collective choice arrangements: in particular, that the agents affected by the operational rules participate in the selection and modification of those rules.

In [12], we investigated the interleaving of rules of social order (i.e. a norm-governed system), rules of social exchange (e.g. opinion formation), and rules of social computational choice to balance the choice of security policy against the available energy in an ad hoc network. This work showed how brute facts, such as the energy level, and institutional facts such as the security level, could be correlated by using processes of opinion formation and collective choice expressed in EC axioms. This allowed collective self-determination of the security level by a vote, based on the opinions communicated through a social network.

We illustrate and exemplify the relevance of this work for the axiomatisation of Principle 3 by showing axioms for participation, selection of a rule affecting the members, and modification of a rule by the affected members.

For participation, we need to ensure that the members of institution I are empowered to vote. This is given by:

$\mathbf{pow}(A, vote(A, X, M, I)) = true$ holdsAt $T \leftarrow$
 $status(M, I) = open$ holdsAt $T \wedge$
 $role_of(A, I) = member$ holdsAt T

This states that agent A has the power to vote on issue M in institution I if two conditions are satisfied. Firstly, that the status of the issue is *open*, i.e. an appropriately empowered agent in I has called for a vote (opened a ballot) on M , which set the fluent $status(M, I)$ to *open*; and no appropriately empowered agent in I has closed the ballot, i.e. has set $status(M, I)$ to *closed*. Secondly, the agent must have the role of *member* in I . Note that what X denotes, either yes/no, a number, a candidate list, etc., depends on the content of M (see below).

Designated actions, i.e. votes by member agents, can be specified to establish the necessary institutional facts:

$vote(A, X, M, I)$ initiates
 $votes_cast(M, I) = [X \mid Votelist]$ at $T \leftarrow$
 $votes_cast(M, I) = Votelist$ holdsAt $T \wedge$
 $\mathbf{pow}(A, vote(A, X, M, I)) = true$ holdsAt T

For the selection of a rule affecting the participants in I , we can for example arrange for a vote on the appropriation rule by calling for a vote. Suppose that in institution i the agent in the role of *head* is c , the current allocation method is *queue*, the winner determination method for choosing the resource allocation method is *plurality*, and there are agents a, b and c . Then suppose we had the following narrative:

```
open_ballot(c, res_alloc_meth, i) happensAt 1
vote(a, ration, res_alloc_meth, i) happensAt 2
vote(b, ration, res_alloc_meth, i) happensAt 3
close_ballot(c, res_alloc_meth, i) happensAt 4
```

We also have the following axiom:

```
obl(C, declare(C, W, M, I)) = true holdsAt T ←
role_of(C, I) = head holdsAt T ∧
status(M, I) = closed holdsAt T ∧
votes_cast(M, I) = Votelist holdsAt T ∧
win_det_meth(M, I) = WDM holdsAt T ∧
winner_determination(WDM, Votelist, W)
```

Then the next event in the narrative should be:

```
declare(c, ration, res_alloc_meth, i) happensAt 5
```

which changes the fluent via the axiom:

```
declare(C, W, M, I) initiates
M(I) = W holdsAt T ←
pow(C, declare(C, W, M, I)) = true holdsAt T
```

so that the appropriation rule in I is now *ration*.

For the modification of a rule affecting the participants in the institution, note that we can specify exactly the same process, but with the issue M being the winner determination rule for the appropriation rule. For more on this style of hierarchical dynamic specification, see [9].

D. Principle 4: Monitoring

Principle 4 is concerned with ensuring that monitoring, of both state conditions and appropriator behaviour, is by appointed agencies, who are either accountable to the resource appropriators or are appropriators themselves.

In one sense, this principle is simply axiomatised by introducing a new role, *monitor*, to which the head agent is empowered to assign to a member agent.

```
assign(C, B, I) initiates
role_of(B, I) = monitor at T ←
pow(C, assign(C, B, I)) = true holdsAt T
pow(C, assign(C, B, I)) = true holdsAt T ←
role_of(B, I) = member holdsAt T ∧
role_of(C, I) = head holdsAt T
```

Appointment to the *monitor* role is associated with *obligations* to sample the state of the environment, observe appropriations, and report this information to the head. The former information is used to trigger a change to the appropriation rule congruent to the state of the environment (Principle 2) using the collective choice protocols (Principle 3). The observation of appropriations is used to ensure the rules are being followed; the report of a misappropriation can lead to a sanction (Principle 5) and a dispute (Principle 6). Note that the role of *monitor* empowers one agent to report another:

```
pow(B, report(B, A, R, I)) = true holdsAt T ←
role_of(B, I) = monitor holdsAt T ∧
role_of(A, I) = member holdsAt T
```

This principle is deeply connected with event recognition and opinion formation, and their interleaving [12].

E. Principle 5: Graduated Sanctions

Principle 5 states that there should be a flexible scale of graduated sanctions for resource appropriators who violate communal rules. For example, for a first offence the sanction level is increased to 1 and the power to demand is withdrawn; for a second offence the sanction level is increased to 2 and the agent may be excluded from the institution.

For example, consider the following narrative, with member agent a , head agent c , and the *ration* appropriation rule in force, and that $r > r'$:

```
demand(a, r, i) happensAt 14
allocate(c, a, r', i) happensAt 15
appropriate(a, r, i) happensAt 16
report(b, a, r, i) happensAt 17
```

Agent a has violated the communal rule by appropriating resources to which it was not entitled, and is reported by a monitor agent b (see Principle 4).

We can add an axiom that counts rule violation offences:

```
report(B, A, R, I) initiates
offences(A, I) = O1 at T ←
pow(A, report(B, A, R, I)) = true holdsAt T ∧
offences(A, I) = O holdsAt T ∧ O1 = O + 1 ∧
res_alloc_meth(I) = ration(R') holdsAt T ∧
allocated(A, R', I) = true holdsAt T ∧ R > R'
```

and empower the head agent to sanction offences:

```
sanction(C, A, S, I) initiates
sanction_level(A, I) = S at T ←
pow(C, sanction(C, A, S, I)) = true holdsAt T
pow(C, sanction(C, A, S, I)) = true holdsAt T ←
role_of(C, I) = head holdsAt T ∧
offences(A, I) = S holdsAt T
```

If an agent A is sanctioned at level 1 for a first offence, then it is not empowered to *demand*. The head agent is empowered to ‘reset’ the sanction level $\text{sanction_level}(A, I) = 0$, so that A once again has its power, but the number of offences does not decrease (i.e. $\text{offences}(A, I) = 1$ still holds). If agent A violates the appropriation rule again, and is sanctioned a second time, the head agent is empowered to exclude agent A because $\text{sanction_level}(A, I) = 2$ (as specified in Principle 1).

Graduated sanctions interleave closely with the conflict resolution mechanisms of Principle 6, which can help treat intentional and unintentional violations differently.

F. Principle 6: Conflict Resolution

Principle 6 states that the institution should provide rapid access to low-cost conflict resolution mechanisms. Alternative Dispute Resolution (ADR) has numerous benefits as an alternative to litigation, including lower cost, shorter time, and damage limitation. It can preserve and even strengthen relationships among the parties [14].

The axiomatisation of ADR protocols is therefore a key element of providing low-cost, rapid conflict resolution mechanisms for self-governing commons. Here we will briefly discuss a simple appeals procedure, and then consider a more refined approach.

1) ‘Simple’ Appeals Procedure: From the specification of the previous two principles, once the monitor has reported an agent, its number of offences is incremented. Given a certain number of offences, the head is empowered to apply a sanction. However, an appeal against a sanction can be made by the sanctioned agent:

$$\begin{aligned} \text{appeal}(A, S, I) & \text{ initiates} \\ & \text{appealed}(A, S, I) = \text{true} \quad \text{at } T \quad \leftarrow \\ \text{pow}(A, \text{appeal}(A, S, I)) &= \text{true} \quad \text{holdsAt } T \\ \text{pow}(A, \text{appeal}(A, S, I)) &= \text{true} \quad \text{holdsAt } T \quad \leftarrow \\ \text{role_of}(A, I) &= \text{member} \quad \text{holdsAt } T \quad \wedge \\ \text{sanction_level}(A, I) &= S \quad \text{holdsAt } T \end{aligned}$$

The head agent can uphold the appeal, which removes the sanction, and decrements the offence count.

$$\begin{aligned} \text{uphold}(C, A, S, I) & \text{ initiates} \\ \text{sanction_level}(A, I) &= S1 \quad \text{at } T \quad \leftarrow \\ \text{pow}(C, \text{uphold}(C, A, S, I)) &= \text{true} \quad \text{holdsAt } T \quad \wedge \\ \text{sanction_level}(A, I) &= S \quad \text{holdsAt } T \quad \wedge \\ S1 &= S - 1 \\ \text{uphold}(C, A, S, I) & \text{ initiates} \\ \text{offences}(A, I) &= O1 \quad \text{at } T \quad \leftarrow \\ \text{pow}(C, \text{uphold}(C, A, S, I)) &= \text{true} \quad \text{holdsAt } T \quad \wedge \\ \text{offences}(A, I) &= O \quad \text{holdsAt } T \quad \wedge \\ O1 &= O - 1 \end{aligned}$$

$$\begin{aligned} \text{pow}(C, \text{uphold}(C, A, S, I)) &= \text{true} \quad \text{holdsAt } T \quad \leftarrow \\ \text{role_of}(C, I) &= \text{head} \quad \text{holdsAt } T \quad \wedge \\ \text{appealed}(A, S, I) &= \text{true} \quad \text{holdsAt } T \end{aligned}$$

This simple appeals procedure allows the head agent to apply some form of ‘common sense’ reasoning to the application of the graduated sanctions. This allows for tolerance of unintentional violations through the application of *forgiveness*, an essential psychological construct complementary to trust in establishing long-lasting social relations [13].

Note that we are here assuming that the monitor is a completely reliable observer and reporter. However, a more substantive appeals procedure would take into account that the monitor may incorrectly report an agent’s appropriation, so it would need to give grounds for an appeal, and that might require a more complex protocol.

2) ‘Complex’ Dispute Resolution: A more complex alternative dispute resolution protocol is presented in [14]. This protocol has three main stages: initiating a dispute, selecting a dispute resolution method, and resolving a dispute. The initiation phase assigns the roles to the two agents in I involved in the dispute, as the litigants involved in a dispute, and secondly initialises the starting values for the second phase, which is the selection of the ADR method.

In the second phase, the protocol tracks through a fluent *adr_method* which ADR method is to be used to resolve the dispute. The value of this fluent can be either *null* (no method), negotiation, arbitration, or mediation; and the protocol proceeds by one litigant proposing a method, which may be agreed or rejected by the other. However, we stipulate that it is not permitted to reject arbitration as a proposed method, to reflect the legal principle that ‘everyone’ should have a right to a jury trial (which is effectively what the arbitration method supports), if that is what one of the parties desires.

Once an ADR method is selected, the dispute is resolved in the third phase using another protocol, which can also be specified in the EC. Negotiation can use the contract-net protocol, mediation can follow an argumentation protocol, and arbitration can use a protocol for jury trials.

V. PERSPECTIVES ON THE AXIOMATISATION

The previous section has shown that Ostrom’s socio-economic principles for enduring institutions can be axiomatised using languages from Artificial Intelligence to reason about action, agency and norms. In this section, we briefly consider the significance of this result from a methodological and an experimental perspective. The methodological perspective situates this work in the context of *sociologically-inspired computing*. The experimental perspective considers the relevance of this work for validating proposition **p3**.

A methodology for sociologically-inspired computing is illustrated in Figure 1. We start from observed phenomena, for example a (human) social, legal or organisational system.

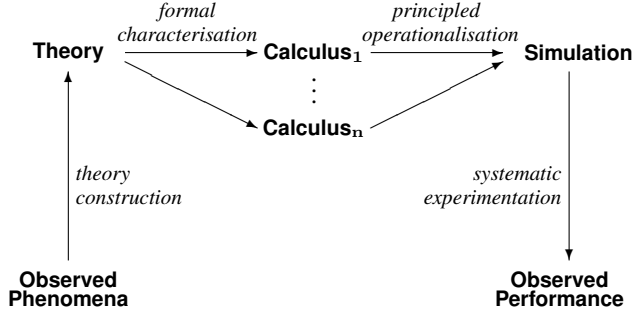


Figure 1: Method for Sociologically-Inspired Computing

The process of *theory construction* creates what we call a pre-formal ‘theory’, usually specified in a natural language. Ostrom [1] comes into this category, as it is an evidence-based theory of enduring institutions but without formalism. The process of *formal characterisation* represents such theories in a calculus of some kind, where by calculus we mean any system of calculation or computation based on symbolic representation and manipulation. (This representation can be at different levels of abstraction depending on the intended role of the formal characterisation: expressive capacity for conceptual clarity, computational tractability for proving properties, etc.) The process of formal characterisation in the previous section has represented the pre-formal theory into the formal action language of the Event Calculus. The step of *principled operationalisation* embeds such formal representations in simulations which include detailed implementation of individual agents.

In this final step, we are building a multi-agent CPR testbed for *systematic experimentation* to validate **p3**. The testbed is designed to be configurable so that we can observe the behaviour of the system with different principles active. The principles which are the subjects of these experiments are *congruence* (Principle 2) and *monitoring* (Principle 4), as axiomatised in the legislature, and investigate their impact on creating *enduring* institutions. (The remaining principles will be included in the testbed in future work.)

Consider again the definition of an institutional resource allocation system of Section III, $\mathcal{I}_t = \langle \mathcal{A}, \epsilon, L, m \rangle_t$. We define a *run* of \mathcal{I} as a sequence of environment states ϵ_t , with $t = 0, \dots, n$. This $n \in \mathbb{N}$ denotes the *lifespan* of the institution and is reached when $P_t < 0$ or $\mathcal{A}_t = \emptyset$. Each state transition is labelled by a set of actions, which may be endogenous or exogenous events. One exogenous event ensures the value of P in ϵ_{t+1} is replenished by the replenishment rate P_{rep} , which is added to the value of P in ϵ_t , less the demands, allocations and appropriations as determined by endogenous actions of the agents according to the legislature L .

A run requires first defining the agent population with

specific characteristics of agents, such as the amount of resource they tend to request, and their ‘propensity for non-compliance’, a probability that they will appropriate more than their allocation.

The basic algorithm and operation of the testbed is described in Table III. (Note that the events of the EC narrative associated with each step are shown with variables, but there are several events at each step and each event would have different instantiations for the variables.) A run then starts with a (default) role assignment, and the system cycles through the following steps, if the corresponding principles are selected: declaration of the resource allocation method, demand and allocation; appropriation; monitoring, i.e. reporting, and possible exclusion.

In every environment state the reservoir starts with a certain amount of water, at most up to a limit P_{max} . Member agents can demand a share R , from the water in the reservoir. Afterwards, the head of the institution allocates the resource among all agents, as described in Section IV, according to the operational resource allocation method. The agents then appropriate their resources and may intentionally take more than they are allocated. If monitoring is activated, there is a probability that their non-compliance will be observed and an exclusion might follow.

The system then advances to the next environment state and the resource is replenished. This repeats, until the replenishment did not restore sufficient resources (it is unsustainable) or there are no agents left in the system. From an external perspective, the aim of ‘the system’ is to maximise the lifespan n .

Table III: Algorithm for CPR testbed.

initially $role_of(A, I) = member$	# role assignment
if {Principle 2} then:	
initially $role_of(C, I) = head$	
if {Principle 4} then:	
initially $role_of(B, I) = monitor$	
$t \leftarrow 0, P \leftarrow P_{max}$	# full resources
while $P \geq 0 \ \&\& \ \exists A \in I \ \&\& \ t < t_{max}$ do:	
if {Principle 2} then:	# congruence
declare $(C, W, raMethod, I)$ happensAt ...	
demand (A, R, I) happensAt ...	# request
allocate (C, A, R', I) happensAt ...	# allocation
appropriate (A, R'', I) happensAt ...	# appropriation
if {Principle 4} then:	
report (B, A, R'', I) happensAt ...	# monitoring
exclude (C, A, I) happensAt ...	# sanction
$t \leftarrow t + 1$	
$P \leftarrow \min(P_{max}, P + P_{rep} - \sum_A R'' - P_{mon})$	# replenish

Figure 2 shows eight runs of \mathcal{I} with a population \mathcal{A} of 100 agents, where different principles have been selected. For simulation purposes, the maximal lifespan n is constrained by $t_{max} = 500$. The refill rates are the same for each run but vary every 50 time steps between high (h), moderate (m) or low (l), furthermore $P_{max} = 10000$ and $R_a \approx 50$,

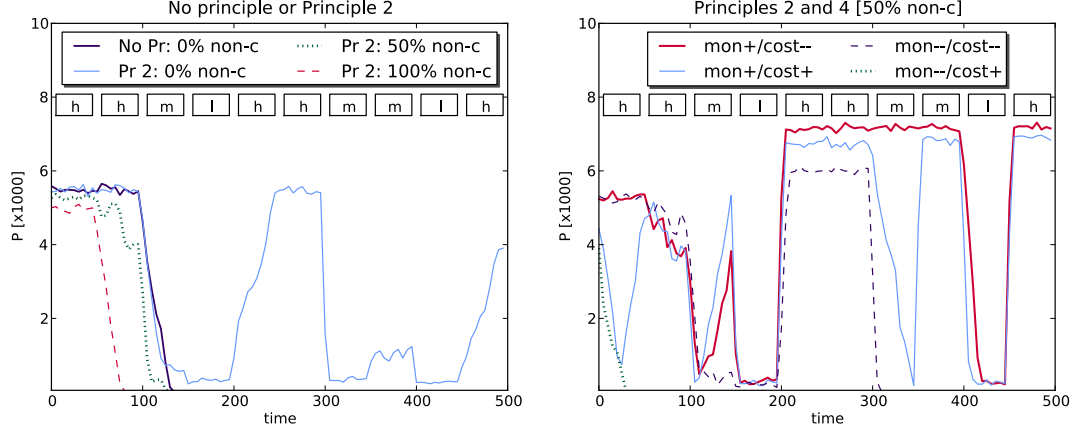


Figure 2: Lifespan of \mathcal{I} using selected principles and varying refill rates.

for all $a \in \mathcal{A}$. There are two resource allocation policies: *queue*, where resources are allocated to the first m agents in a queue, provided $\sum_{a=1}^m R_a \leq P_t$; and *ration*, where each agent is allocated $\min(R_a, P_t/|\mathcal{A}|)$. An agent is compliant if it observes this allocation, and non-compliant if, with some probability, it appropriates more than it was allocated.

On the left hand side, there is one run where no principle has been selected and all agents are compliant, and three runs where Principle 2 has been selected with a different proportion of non-compliant agents (0%, 50% and 100%). The graphs show that with no principle activated, even when all agents comply, the system goes bankrupt with a relatively short lifespan. With Principle 2 and 0% non-compliance, the system responds to variations in P_{rep} with a congruent appropriation and therefore endures until t_{max} . For the case with 50% and 100% non-compliant members, the system goes bankrupt again. Thus in an open system with sub-ideal behaviour Principle 2 alone is not enough to ensure an enduring institution.

On the right hand side of Figure 2, Principles 2 and 4 are activated and the proportion of non-compliant members is 50%. The four runs use different settings for monitoring success (high/low: $+/-$) and cost (expensive/cheap: $+/-$). This cost, P_{mon} , is met from the resources. When monitoring success is low but the cost is high, the lifespan is shorter than with Principle 4 inactive. As the cost decreases, the lifespan increases considerably, but the low monitoring success still leads to bankruptcy. For the other two cases, the monitoring success increases. Eventually, enough non-compliant members are excluded and the situation is similar to Principle 2 but with fewer agents and (approximately) 0% non-compliance. The remaining members manage to preserve the resource and absorb even high monitoring costs. Thus the agents do not bankrupt their system.

We conclude that exhaustive monitoring (i.e. the costs are too high with respect to its success), can do harm to the system, but if the rules are designed such that the cost is

commensurate with the effort, it leads to a more stable and enduring institution. This connects to Ostrom’s observation [1, p. 96] that the cost of monitoring is dependent on other conditions, and so the members should be able to modify the monitoring rules as per Principle 3.

Finally, the results support the supposition that with every additional principle activated, the institutional lifespan increases. Thus, we conjecture that using all the principles will ultimately lead to enduring *electronic* institutions too, but this remains for further work.

VI. RELATED AND FURTHER WORK

Traditionally, the role of a software engineer has been to apply some methodology to implement a ‘closed’ system which satisfies a set of functional and non-functional requirements. Our problem is to engineer ‘open’ systems where the primary non-functional requirement, that the system should endure, is an *emergent* property, and is a side-effect of the *interaction* of components rather than being the goal of any of those components. More generally, unplanned emergent behaviour exhibited by complex socio-technical systems cannot easily be handled by top-down design methods. Thus the approach proposed here has much in common with other new design methods, for example *design for emergence* [15], for systems that adapt and evolve, and where the design method specifically targets ‘self- \ast ’ properties.

Our aim has been to leverage Ostrom’s work for *agent-based software engineering*, but there is also related research from the perspective of *agent-based modelling*. This reveals many additional parameters to consider in developing experiments to test the emergent property of endurance. For example, [16] investigate whether or not people are prepared to invest their own resources in endogenous rule change, e.g. from open access to private property. We will also have to design experiments which consider the ‘cost’ of rule changes, the costs of monitoring and dispute resolution, and the impact this has on ‘endurance’. In addition, Ostrom’s

original analysis has been extended to introduce more than 30 factors which influence endurance [17], and we may need to enrich our model with these additional parameters.

Although many related works on institutional action and institutionalised power parameterise their formal accounts with respect to an institution, for ‘simplicity’ or ‘expediency’ it is assumed that there is just one institution. However, the key feature of Principle 8 is that there are layered or encapsulated CPRs, or multiple CPRs operating in the same space. This is why we have included the parameter *I* in the fluents and actions of our EC specification, as a placeholder for further work on systems of systems of CPR. We also plan to implement an ‘asynchronous’ version of the testbed based on Ostrom’s notion of a *decision arena* and an efficient EC dialect [18]. All resource allocation decisions would take place in one decision arena, all dispute resolution procedures would take place in another, and so on. This would allow the application of operational choice rules, collective choice rules, etc. to overlap and interleave, rather than all being resolved within one time slice.

VII. SUMMARY AND CONCLUSIONS

The main contribution of this paper is the first logical axiomatisation in an action language of Ostrom’s socio-economic principles for common pool resource management using self-organising institutions. It has shown how the three types of institutional choice rules, at constitutional, collective and operational levels, can be given a common treatment in the same formalism. Moreover, the computational basis of the formalisation provides the foundations for implementing a testbed to examine the proposition that Ostrom’s principles are necessary and sufficient conditions for creating *enduring* institutions. A successful outcome of these experiments will offer a proof-of-concept for innovative design and specification of self-organising systems for a range of open, embedded, and resource-constrained systems; or the experiments will determine the constraints, conditions and perhaps other factors [17] that need to be considered for these principles to work.

However, we believe that the work reported here has laid the foundations to address further challenges, for example in the automation of enduring institutions for cloud computing, the development of sustainable institutions for smarter infrastructure management using socio-technical systems, and a deeper investigation into the development of institutions that are not only self-organising, but are also *self-aware*.

ACKNOWLEDGMENTS

We would particularly like to thank the anonymous reviewers for their valuable feedback and useful suggestions.

REFERENCES

[1] E. Ostrom, *Governing the Commons*. CUP, 1990.

[2] A. Artikis, M. Sergot, and J. Pitt, “Specifying norm-governed computational societies,” *ACM Transactions on Computational Logic*, vol. 10, no. 1, pp. 1–42, 2009.

[3] R. Kowalski and M. Sergot, “A logic-based calculus of events,” *New Generation Computing*, vol. 4, pp. 67–95, 1986.

[4] A. Jones and M. Sergot, “A formal characterisation of institutionalised power,” *Journal of the IGPL*, vol. 4, no. 3, pp. 429–445, 1996.

[5] J. Searle, *Speech Acts*. CUP, 1969.

[6] I. Pörn, *Action Theory and Social Science: Some Formal Models*, ser. Synthese Library. D. Reidel, 1977, vol. 120.

[7] I. Kremer and K. Nyborg, “Divisible-good auctions: The role of allocation rules,” *RAND Journal of Economics*, vol. 35, no. 1, pp. 147–159, 2004.

[8] S. Brams and A. Taylor, *Fair Division: From Cake-Cutting to Dispute Resolution*. CUP, 1996.

[9] A. Artikis, “Evaluating dynamic protocols for open agent systems,” in *Proc. of Int. Conf. on Autonomous Agents and Multi-Agent Systems (AAMAS)*. ACM, 2009, pp. 97–104.

[10] B. S. Firozabadi and M. Sergot, “Contractual access control,” in *Proc. 10th Int. Workshop on Security Protocols (2002)*, ser. LNCS, B. Christianson, B. Crispo, J. Malcolm, and M. Roe, Eds., vol. 2845. Springer, 2004, pp. 96–102.

[11] J. Pitt, L. Kamara, M. Sergot, and A. Artikis, “Voting in multi-agent systems,” *Computer Journal*, vol. 49, no. 2, pp. 156–170, 2006.

[12] J. Pitt, D. Ramirez-Cano, M. Draief, and A. Artikis, “Interleaving multi-agent systems and social networks for organized adaptation,” *Computational and Mathematical Organization Theory*, 2011.

[13] A. Vasalou, A. Hopfensitz, and J. Pitt, “In praise of forgiveness: Ways for repairing trust breakdowns in one-off online interactions,” *Int. J. of Human-Computer Studies*, vol. 66, no. 6, pp. 466–480, 2008.

[14] J. Pitt, D. Ramirez-Cano, L. Kamara, and B. Neville, “Alternative dispute resolution in virtual organizations,” in *Proc. ESAW’07*, ser. LNCS, A. Artikis, G. O’Hare, K. Stathis, and G. Vouros, Eds., vol. 4995, 2007, pp. 72–89.

[15] M. Uliuru, “Evolving the ‘DNA blueprint’ of eNetwork middleware to control resilient and efficient cyber-physical ecosystems,” in *2nd Int. Conf. Bionetics*, 2007, pp. 41–47.

[16] M. Janssen, R. Goldstone, F. Menczer, and E. Ostrom, “Effect of rule choice in dynamic interactive spatial commons,” *Int. Journal of the Commons*, vol. 2, no. 2, pp. 288–311, 2008.

[17] A. Agrawal, “Common property institutions and sustainable governance of resources,” *World Development*, vol. 29, no. 10, pp. 1623–1648, 2001.

[18] A. Artikis, M. Sergot, and G. Paliouras, “A logic programming approach to activity recognition,” in *Proc. of the 2nd ACM Int. Workshop on Events in Multimedia*, 2010, pp. 3–8.