

Systems of Exchange Values as Tools for Multi-Agent Organizations

Graçaliz P. Dimuro¹ A. C. Rocha Costa^{1,2}
Luiz A. M. Palazzo¹

¹ Escola de Informática
Universidade Católica de Pelotas - UCPel
R. Félix da Cunha 412, 96010-000 Pelotas, RS, BRAZIL

²Programa de Pós-graduação em Computação
Universidade Federal do Rio Grande do Sul - UFRGS
Av. Bento Av. Bento Gonçalves 9500, Bloco IV, 91501-970 Porto Alegre, RS, BRAZIL
{liz; rocha; lpalazzo}@atlas.ucpel.tche.br

Abstract

This paper introduces systems of exchange values as tools for the organization of multi-agent systems. Systems of exchange values are defined on the basis of the theory of social exchanges, developed by Piaget and Homans. A model of social organization is proposed, where social relations are construed as social exchanges and exchange values are put into use in the support of the continuity of the performance of social exchanges. The dynamics of social organizations is formulated in terms of the regulation of exchanges of values, so that social equilibrium is connected to the continuity of the interactions. The concept of supervisor of social equilibrium is introduced as a centralized mechanism for solving the problem of the equilibrium of the organization. The equilibrium supervisor solves such problem making use of a qualitative Markov Decision Process that uses numerical intervals for the representation of exchange values.

Keywords: Social Exchanges, Exchange Values, Exchange Values-Based Social Organization, Social Equilibrium, Equilibrium Supervisor, Qualitative Markov Decision Process

1 Introduction

The exchange values approach to social interactions [31, 19], in which an interaction is seen as an exchange of services among a small group of agents, along with the corresponding evaluation of (assignment of values to) such services, was conceived as a methodological stance where organizations are taken

as collections of small groups of agents.

Each such group is seen as having its own internal interaction dynamics, and the collection of all groups may be recursively construed as a collection of small groups of small groups, etc., thus picturing the organization as a hierarchy of levels of structurally similar dynamical systems, all centered around the issue of exchanging values between their components.

The overall balance of exchange values involved in the operation of a system can be analyzed for its state in terms of equilibrium (beneficial for all agents) or disequilibrium (beneficial to a particular agent or group of agents, or beneficial to nobody).

It is usually attributed to George Caspar Homans the initial development of the exchange values approach to social organizations [19, 20]. In connection to that, the exchange values approach is sometimes criticized as too narrow in scope, on the basis of its supposedly inherent behavioristic limitations, due to Homans' explicit adoption of Burrhus Skinner's behavioristic psychology as an ancillary explanatory theory of individual human behavior.

Nevertheless, the exchange values approach to the analysis of social interaction has nothing inherently behavioristic in it. Well before Homans introduced his ideas in [19], Jean Piaget proposed a constructive, non-behavioristic, theory of social exchanges, where exchange values play a central role [30, 31].

Moreover, Piaget's theory of exchange values goes much further than Homan's theory, both in what concerns the scope of application of exchange values (going as far as showing their role in the origination of laws, moral rules and organizational norms) and in what concerns the degree of formalization of the ideas.

The main contribution of the exchange values approach to social exchanges, in comparison to the classical, quantitative utility functions based approach, is the introduction of the possibility of taking care of the subjective values, of qualitative nature, with which everyone judges the daily exchanges he has (*good, bad, better than, worst than*, etc.), which usually cannot be faithfully represented quantitatively, due to the lack of neat objective conditions for their measurement.

In this paper, we build on our previous work regarding the application of Piaget’s approach to the analysis of interactions in multi-agent systems [34, 35, 14] and cooperative environments [8, 10], to consider the way systems of exchange values can be taken as useful tools for helping to solve problems concerning the organization of multi-agent systems.¹ Also, we explain our understanding of the connection between Piaget’s and Homans’ approaches to exchange values, showing the role that Homan’s theory can play in the broader analytical framework set up by Piaget.

The paper is organized as follows. In Sect. 2, we summarize the sociological basis of the work, as encompassed by Piaget’s theory of exchange values (Sect. 2.1), by Homans’ theory of social behavior (Sect. 2.2), and by the way we think both theories fit together (Sect. 2.3).

In Sect. 3, we present in an axiomatic way our notion of social organization: in Sect. 3.2 we state how social organizations are construed by social functions, social roles, exchange values, and social rules. In Sect. 3.4, we express the dynamics of social organizations, in terms of regulation of exchanges of values, so that social rules and social equilibrium can be connected to that dynamics. Section 3.5 introduces the notion of supervisors of social equilibrium.

In Sect. 4 we bring the main concrete contribution of the paper, namely, a general model of a supervisor of social equilibrium. Section 4.1 introduces our way of representing qualitative exchange values by intervals of a numeric scale. Section 4.2 explains the use of such intervals to the modelling of social exchanges between two agents, and Sect. 4.3 generalizes that modelling to a matrix-like notation capable of representing social exchanges between all agents of an organization. Section 4.4 shows how the exchange values equilibration problem of an organization can be solved by a qualitative interval Markov Decision Process [37], which uses the interval representation for qualitative values

¹Values have been extensively used in the MAS area, through value-based and market-oriented decision, and value-based social theory, see e.g., [2, 36, 25]. The merits and interest in using Piaget’s notion of exchange values is discussed in [34, 35], showing its complementarity to the dependence theory approach.

introduced earlier.

Section 5 brings the analyzes of the model: a theoretical analysis (Sect. 5.1), regarding the reachability of the equilibrium state, considering the case in which all agents follow the recommendations of the equilibrium supervisor; and a comparative analysis (Sect. 5.2) of simulations of (unsupervised and supervised) exchanges processes, considering different degrees of obedience to the supervisor. Section 6 is the Conclusion.

2 Sociological Bases

2.1 Piaget’s Theory of Exchange Values

As a cognitive psychologist, Piaget gave much less attention in his work to social and affective aspects of human behavior, than to its cognitive aspects. Nevertheless, he had the opportunity to express in various places his ideas about those subjects (e.g., in [30, 31]).

2.1.1 Interactions as Social Exchanges

In what concerns human society, Piaget adopts a relational approach, such that the structure of a society is defined as a relational structure where the relationships among the individuals are established by social exchanges among them. Thus, *interactions* are understood as exchanges of *services* among individuals, involving not only the realization of services by some individuals on behalf of others, but also the valuation of such services, from various points of view, by every individual involved in them.

A service performed by an individual, however, is not a simple action or interference on the action of somebody else. To be counted as a service, an action performed by an individual has to be understood by all individuals involved in the action as an intentional action, directed toward some other individual, thus allowing its evaluation as beneficial or prejudicial to the latter. The evaluation of a service by an individual (either the server of the service or its client) is done on the basis of a *scale* of so-called *exchange values*, which are of a qualitative nature, since such values express subjective evaluations ².

² A scale of exchange values is a set of qualitative, ordinal values, which can be compared for their magnitudes (less than, equal, greater than), but cannot be algebraically operated in an unrestricted way, as fully quantitative values can. E.g., the values can be added or subtracted: if $a < b$ and $a < c$ then $a < b + c$. However, the differences between values cannot be compared: if $a < b$ and $c < d$ it is not possible to decide which of $b - a < d - c$ or $b - a = d - c$, or else $b - a > d - c$, is true.

2.1.2 Exchange Values and Social Equilibrium

Exchange values give rise to a *qualitative economy of social exchanges*, where individuals acquire *credits* for services they have performed, and *debts* to others for services the others have performed to them. The *balances* of exchange values allow individuals to observe the state of equilibrium of the social exchanges (even between just two individuals) and to react according to such state (e.g., trying to enforce equilibrium, to overcome high debts, to keep their status as privileged beneficiaries of the exchanges etc.).

Qualitative exchange values encompass *economic values* as a particular kind of (quantitative) values, and are seen as the cornerstones of social rules. *Social rules* of the many different varieties (formal or informal; moral, economical or juridical, including organizational norms) can often be understood as means put to operate in an effort to guarantee that the overall balance of exchange values are kept in certain equilibrated (or disequilibrated, i.e., favorable to some individuals or groups of individuals) states, so individuals are kept motivated (or, enforced) to continue their participation in those exchanges.

A crucial aspect for the right understanding of Piaget's approach is the clear differentiation between the notions of *social equilibrium* and *social order* (social stability). By social equilibrium it should be understood a kind of equality, or equity, in the distribution of exchange values among the agents participating in the exchange. By social order, or social stability, it should be understood the temporal continuity of the established set of social exchanges. The two concepts are orthogonal, in the sense that any stable or unstable exchange can be either in equilibrium or in disequilibrium, and any equilibrated or disequilibrated exchange can be either stable or unstable. On the other hand, they are not independent, in the sense that each may impact the other ³.

The main technical contribution of this paper is the concept of an *equilibrium supervisor*, a system component (possibly an agent) that, at each moment, is able to recommend that particular exchanges among agents be performed so that a subset of the agents of the system (possibly, the whole set of agents) be kept in equilibrium (or, disequilibrium). That is, so that all agents in that set benefit from the exchanges in equal (or, differentiated) terms. In other words, equilibrium supervisors embody social rules designed to keep a set of agents in certain (equilibrated or disequilibrated) states of exchange values.

³But this mutual dependence was not fully exploited by Piaget, who concentrated on stable ("static") societies.

2.1.3 Material and Virtual Exchange Values

Social exchanges are classified by Piaget into two broad categories: *immediate exchanges* and *deferred exchanges*. In immediate exchanges, individuals exchange services in an immediate way, service for service, so that the evaluations of such services can be done immediately, as the services are being performed, allowing each individual to immediately regulate the quality and quantity of the service it performs for the other (as when two people exchange material goods, negotiating the quantities in which each good will participate in the exchange).

Two kinds of values are associated to such services, corresponding to the *investment (cost)* necessary to perform them, and to the *satisfaction* they may cause to the client. Such values are called *material exchange values*.

Deferred exchanges involve a separation in time between the stages of an exchange of services, and give rise to so-called *virtual exchange values*, encompassing *credits* and *debts*: after an individual performed a service for the other, the first is entitled a credit for the service performed, and the other is entitled a debit for receiving the service, and is supposed to pay that debit in the future ⁴.

2.1.4 Social Rules

Virtual values have the weakness that they tend to vanish as time passes: in a very distant future, one may not feel obliged to perform a service to somebody, in return for a service he received before, precisely because that fact, having happened in a too distant past, may have lost its importance for the current situation.

Thus, besides serving the purpose of governing individual reactions to the state of the overall balance of exchange values, social rules also play the essential role of a means to avoid the vanishing of virtual values, by associating such values to behaviors that aim their conservation, and that are to be mandatorily performed. Virtual values have their most neat application to the regulation of exchanges between two individuals, but are easily extended to cover relations involving more individuals.

Social rules regulating the preservation of such values are either *moral rules*, historically created by the group, or *private contracts* spontaneously established between the individuals involved in the exchanges.

⁴The term *virtual value* refers precisely to the fact that they represent services that are yet to be performed, in the future.

2.1.5 The Structure of Social Exchanges

A *social exchange* between two agents, α and β , is performed involving two types of stages, illustrated by the schemes in Fig. 1. In stages of type $I_{\alpha\beta}$, the agent α realizes a service for β . The exchange values involved in this type of exchange stage are the following:

- $r_{I_{\alpha\beta}}$ is the value of the *investment* done by α for the realization of the service for β ;
- $s_{I_{\beta\alpha}}$ is the value of β 's *satisfaction* due to the receiving of the service done by α ;
- $t_{I_{\beta\alpha}}$ is the value of β 's *debt*, the debt it acquired to α for its satisfaction with the service done by α ;
- $v_{I_{\alpha\beta}}$ is the value of the *credit* that α acquires from β for having realized the service for β .

Investment values are always *negative*, while the other values may be either positive or negative.

In stages of the type $II_{\alpha\beta}$, α asks β the payment for the service he did previously for β , and the values related with this exchange stage – $v_{II_{\alpha\beta}}$, $t_{II_{\beta\alpha}}$, $r_{II_{\beta\alpha}}$ and $s_{II_{\alpha\beta}}$ – have similar meaning. $r_{I_{\alpha\beta}}$, $s_{I_{\beta\alpha}}$, $r_{II_{\beta\alpha}}$ and $s_{II_{\alpha\beta}}$ are called *material values*. $t_{I_{\beta\alpha}}$, $v_{I_{\alpha\beta}}$, $t_{II_{\beta\alpha}}$ and $v_{II_{\alpha\beta}}$ are the *virtual values*. The order in which the exchange stages may occur is not necessarily $I_{\alpha\beta}$ – $II_{\alpha\beta}$.

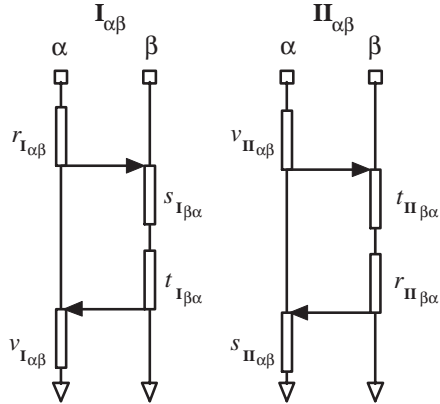


Figure 1: Stages of social exchanges.

Piaget's modelling of social exchanges has an algebraic flavor, aiming at the formalization of algebraic laws for the operations involved in those exchanges, laws that serve as the bases for formalization of the rules that determine the equilibrium of exchanges:

Rule $I_{\alpha\beta}$: $(r_{I_{\alpha\beta}} = s_{I_{\beta\alpha}}) \wedge (s_{I_{\beta\alpha}} = t_{I_{\beta\alpha}}) \wedge (t_{I_{\beta\alpha}} = v_{I_{\alpha\beta}})$

Rule $II_{\alpha\beta}$:

$$(v_{II_{\alpha\beta}} = t_{II_{\beta\alpha}}) \wedge (t_{II_{\beta\alpha}} = r_{II_{\beta\alpha}}) \wedge (r_{II_{\beta\alpha}} = s_{II_{\alpha\beta}})$$

Rule $I_{\alpha\beta}II_{\beta\alpha}$: $v_{I_{\alpha\beta}} = v_{II_{\alpha\beta}}$.

Rule $I_{\alpha\beta}$ states the conditions for the internal equilibrium of stage $I_{\alpha\beta}$, implying that the investment made by α in the performance of the service for β equals the credit that β assigns to α , that is

$$\text{Rule } I_{\alpha\beta} \Rightarrow r_{I_{\alpha\beta}} = v_{I_{\alpha\beta}}.$$

Rule $II_{\alpha\beta}$ states the conditions for the internal equilibrium of stage $II_{\alpha\beta}$, implying that the credit charged by α on β equals the satisfaction α gets from the return service performed by β , that is

$$\text{Rule } II_{\alpha\beta} \Rightarrow v_{II_{\alpha\beta}} = s_{II_{\alpha\beta}}.$$

Rule $I_{\alpha\beta}II_{\alpha\beta}$ states the conditions for the external equilibrium between the two stages, $I_{\alpha\beta}$ and $II_{\alpha\beta}$, implying that the initial investment made by α equals the final satisfaction it gets from the interaction with β , that is

$$\text{Rule } I_{\alpha\beta}II_{\alpha\beta} \Rightarrow r_{I_{\alpha\beta}} = s_{II_{\alpha\beta}}.$$

The equilibrium rules play a central role in Piaget's explanation of the dynamics of the social organization and in the identification of situations of disequilibrium (including several kinds of social crises).

The development of certain moral rules, for instance, is associated to the equilibrium rules, for such moral rules are developed precisely for the sake of guaranteeing the validity of the equilibrium rules, through the enforcement of certain behaviors that can compensate behaviors which are prone to produce (or, that may have produced) a situation of disequilibrium.

The main role of the equilibrium supervisor introduced in this paper is precisely that of indicating what behaviors should be performed at each moment, by the agents, so as to compensate extant deviations from the state of equilibrium of exchange values.

2.2 Homans' Theory of Elementary Social Behaviors

Homans approached the subject of social exchanges (that he called *elementary social behaviors*) from a different point of view [20]: he was interested in explaining why each agent behave the way he does, in such exchanges. Being a sociologist, Homans borrowed from Skinner the theory of *operant conditioning*, as a means to explain why men continue to behave in certain ways (or, change behaviors), in certain situations.

Homans looked at the exchanges values when he looked for a sufficient stimulus for continued (or, discontinued) social behavior, and found it in the concept of *profit*, defined as $profit = benefit - cost$, where *benefit* and *cost* are defined almost exactly as in Piaget's

theory. Profit, in this qualitative sense, is seen as the element that can play the role of stimulus: bigger profit means stronger stimulus to continue the current behavior, while smaller profit means stronger stimulus to discontinue the current behavior.

Of course, search for profit maximization was known, to Homans, not to lead necessarily to the best overall results for the partners of an interaction, especially when looked from the point of view of social equilibrium, as Pareto had shown much earlier⁵. So, Homans had to extend Skinner's conceptions of human behavior with notions that are essentially non-behavioristic (such as the notion of *personal integrity before a group* [19]) in order to produce consistent explanations.

Homans' proposal gave the starting point for the formalization we present below, where the process of *social control* that regulates an interaction (which in Homans' theory results from the combination of the various individual behaviors involved in the interaction) is reduced to a Markov Decision Problem [37], to be solved by a central equilibrium supervisor.

2.3 The Role That Homans' Theory May Have in Piaget's Framework

The striking similarity between the basic concepts of the two theories, by Homans and Piaget, regarding the idea that a society is based on an organization where the relationships between individuals are conceived as valued exchanges, puts the question about the possible ways the two theories can be made closer to each other.

From our point of view, the way the two theories can be combined is based on the following observations. On the one hand, in the theory of the psychologist Piaget, the psychological aspects of social interactions were kept out of the formalization of social exchanges, to allow for a qualitative algebraic form. The price Piaget paid for that is that the theory had to give up producing a description of the decision processes that the individuals adopt while interacting, that is, to give up explaining the individuals' behaviors.

On the other hand, in the theory of the sociologist Homans, the psychological aspect of social interactions was brought to the foreground, in order to produce a description of the decision processes adopted by the interacting individuals, and an explanation of their behaviors. The behavioristic basis of the adopted psychological explanation, however, forced two consequences. First, the exchange values that could be

⁵Homans and Piaget were both sympathetic readers of Pareto, see [22] and [30].

taken into account were restricted to those of a quantitative nature, so that the notion of *profit* could be introduced, as a measurable and comparable difference between *benefits* and *costs*. Second, the psychological theory itself had to be extended in non-behavioristic directions, in order to accommodate subjective phenomena not directly observable, but essential for the explanation of moral behaviors.

The very constructions of the two approaches, thus, already indicate the way they may be combined: Piaget's theory can be used to formalize an explanation of social exchanges where the agents are understood algebraically, on the basis of the operations they use to handle the exchange values, while Homans' theory can be adapted to deal with the piagetian algebra of qualitative values, in order to formalize the internal decision processes that agents follow when deciding on the keeping or changing of their social behaviors.

In this paper, we present a first step in that direction. We don't show agents making decisions on the basis of their own idiosyncratic criteria, but we show a simplified version of the problem, where all agents are assumed to uniformly follow the same set of criteria when deciding what exchanges to propose to another agent, but where each agent is allowed to answer to that proposal in its own manner.

Halving the complexity of the problem in this way, we were able to perform the first, provisory stage in the formalization of the proposed integration of the two sociological theories that were taken as our bases of our work. The next step in the evolution of this synthesis, accounting for the agents' individual decision making with respect to exchange values, is briefly considered in the Conclusion of the paper.

3 Social Organizations

In this section, we build on previous work on the dynamics of multi-agent organizations [4, 5, 6, 7, 3, 13] in order to coordinate the notions of organization and exchange values, so that systems of exchange values can become useful tools for the organization of multi-agent systems.

The central concept introduced here is that of *dynamics of exchange values* in social organizations.

3.1 The Notion of Social Organization

The notion of organization is not univocal. There are at least two main senses in which it is used, namely, a *functional* sense and a *structural* sense. In the functional sense, organization is one of the *functional in-*

variants that characterize all forms of autonomous dynamic systems [28]. In the structural sense, the organization of an autonomous dynamic system is the relational structure that allows the system's components to interact with each other. In this paper, we use the term *organization* in the latter, structural sense.

In its simplest form, the *organization* of a society S , at a given time t , is conceived as a structure $O_S^t = (A_S^t, E_S^t)$ where A_S^t is the set of agents of the society at time t , and $E_S^t \subseteq A_S^t \times A_S^t$ is the set of social exchanges that are happening at that time⁶. In the following, we will only consider the *static* (synchronic) case, where the organization of a society is not changing while the society is functioning. So, we will let implicit the time index t . Also, since we will consider one single society, we will let implicit the index S .

3.2 Social Functions, Social Roles, Exchange Values, and Social Rules

Societies with the simple form defined above should have a big problem to keep their existence in time: the only permanent element they have are the agents, since exchanges are usually finite, and vanish suddenly. Agents, in such situations, are thus allowed to interact with any other agents in the society, at any time, as they are pleased, or simply not to interact at all. Organization, in such societies, is no more than a set of momentary exchanges. There is no way organization can become a *structural invariant* of the society, in a society with such a simple form.

Many other permanent elements are required, in the organization of a society, if the society is to be able to keep its organization reasonably stable during a certain period. We identify four main notions allowing for such *invariant organization*, namely, *social functions*, *social roles*, *exchange values*, and *social rules*:

- *social functions* are the *services* that agents (or, sets of agents), perform for other agents (or, sets of agents) in the society, and that justify the very existence of the society: agents that are self-sufficient need not live in societies; only agents that have the need that others perform certain services for them care to live together with others;
- *social roles* are the relational elements that establish the links between the agents, the social functions they perform, and the behavior they should

have in order to perform such functions in an adequate way; social roles are the gluing elements of an organization;

- *exchange values* are the means by which the services that an agent performs are evaluated by the members of the society (including the agent itself and the other agents involved in the exchange), so that a resulting *balance of exchange values* can be used by social rules (see below) to compensate behaviors that deviate the society from the desired kind of balance (equilibrated, disequibrated);
- *social rules* are the means by which agents (given the social functions that are to be performed in the society, and the way the social roles were assigned to the agents) are obliged to behave in certain ways, and forbidden to behave in some other ways.

Implicit in the structural conception of organization is, thus, a dynamical notion of *social equilibrium* (or, *disequilibrium*), which – in this paper – we restrict to the notion of *equilibrium* (or, *disequilibrium*) of *social exchanges*, as defined by Piaget.

3.3 Definition of Social Organization

For the purposes of this paper, the notion of social organization can be defined as:

Definition 1 *The organization of a society is a structure $O = (A, F, Ro, E, BV, Ru)$ where: A is the set of agents; F is the set of services (functions) that agents, and sets of agents, should provide for each other; Ro is the set of social roles that agents may be assigned to; E is the set of social exchanges that agents may perform between them; BV is the set of balances of exchange values that supports the various ways agents may evaluate social exchanges; Ru is the set of social rules that regulate the agents' behaviors.*⁷

To the main elements of the organization of a society, given in the definition above, a few more *complementary elements* should be added, so that the dynamics of the organization can be explained:

- the set $IBeh$ of all possible *individual behaviors* of all agents of the society, so that to each agent corresponds the set of individual behaviors that it is capable of realizing, given by $IB : A \rightarrow IBeh$;

⁶We follow [11] in considering the dependence relations as one of the main reasons for the establishment of social exchanges.

⁷We let undetermined the details of the sets in Def. 1, in order to have an *abstract* notion of social organization, which may be instantiated in various ways, in various applications.

- the way the set E of all possible social exchanges are related to the subset of agents that are capable of realizing them together, given by $Cap : \wp(A) \rightarrow E$;
- the way each social function is *implemented* by a set of agents in the form of a social exchange, given by $I : F \times \wp(A) \rightarrow E$;
- the way each social role determines the individual behavior of the agent to which it is assigned, regarding the *performance* of a social function, given by the function $P : F \times Ro \rightarrow IBeh$;
- the way each agent evaluates the performance of an exchange, given by the function $Ev : E \times A \rightarrow BV$
- the way each social rule determines the *permitted*, *obligatory* and *forbidden* behaviors of agents in a social exchange, according to the balance of exchange values assigned to the social exchange, regarding the performance of a social function, given by $Ru : F \times A \times E \rightarrow \wp(IBeh \times \{\mathbf{p}, \mathbf{o}, \mathbf{f}\})$.

The detailed exploration of the connections among the main and complementary elements of a social organization, as defined above, are out of the scope of this paper. We concentrate just on the connections that allow the understanding of our notion of *dynamics of exchange values*, leading to the concept of *exchange value-based social control*, and its provisory centralized form, the *supervisor of social equilibrium*.

3.4 The Dynamics of Exchange Values

Although Piaget, in his sociological works, did not develop his notion of dynamics of an organization, both in his works on the Epistemology of Biology [28] and in his psychological works on the equilibration of cognitive structures [29], that notion is well exposed. We summarize here that *piagetian cybernetics*, concentrating just in the concepts that directly apply to our problem.

In any dynamical system where a notion of *equilibrium* can be defined, two related concepts immediately apply, namely, the concepts of *deviation* and *compensation*. Deviation is any action that may happen in a system and lead it to *disequilibrium*, that is, away from equilibrium. Compensation is any action that may happen in a system, when it is in disequilibrium, and lead it back to equilibrium.

Regulation is the process of determining which compensation should be performed, at a given moment, to

compensate a deviation, when the system is in disequilibrium. In Homan's terms [19], Piaget's process of regulation is a *social control* process that aims to keep the society stable in a state of equilibrium.

Piaget applies such ideas both to the *synchronic regulation* of a system's functioning (where the structure of the system is kept essentially unchanged) and to the *diachronic regulation* of the development of the system's structural organization, when the system is in a process of development. To the first, he assigned the name *minor equilibration*, while the latter he called *major equilibration*. In our case, we focus on the synchronic functioning of a social system, so that only minor equilibrations will be considered.

We consider social systems where the interactions are seen as exchanges of services, and whose equilibrium is defined on the basis of balances of values associated to such exchanges. In such systems, deviations are actions (performances of services) whose evaluations lead the system to a state where the balances of exchange values are such that at least one of the equilibrium rules $I_{\alpha\beta}$, $II_{\alpha\beta}$ and $I_{\alpha\beta}II_{\alpha\beta}$ is not satisfied. A compensation is a performance of a service that may lead the system back to a state where the balances of exchange values are such that those rules are satisfied.

Social rules specify a mechanism of social control by stipulating that, for each state of disequilibrium, the kind of action that should be performed in order to re-establish the equilibrium of the system. Two kinds of such actions are possible: *punishment* and *reciprocation*. Punishment is an action by which an individual suffers some loss in order to be reinforced towards not repeating the deviation action again.

Reciprocation is an action by which an individual is forced to perform a service for the other, in order to compensate him for some service the latter had previously performed for him. Reciprocation is the fundamental operation of compensation, in Piaget's model of regulation of social exchanges.

Of course, everything that has been said here in connection to social control processes that aim to stabilize the society or organization in a state of equilibrium, also applies to social control processes that aim to keep the society or organization stable in a state of disequilibrium⁸.

3.5 Supervisors of Social Equilibrium

In general, such *exchange values-based mechanism of social control* may be put to operate in two main

⁸See [12] for a discussion of the role of social control processes in artificial societies.

ways. On the one hand, social rules may be enforced by *authorities*, which have the capacity to push the agents of the society to follow such rules.

On the other hand, social rules may be *internalized* by the agents, so that agents follow such rules because they are incorporated into the agents' behaviors.

Typical of the social rules enforced by authorities are the *juridical* rules of a society (*laws, statutes, organizational norms*), dealing with rights and duties. Typical of the internalized rules are the *moral* rules, dealing with permissions and obligations.

In the following, as a preparatory step to a future study of decentralized social control mechanisms based on social rules internalized in agents, we introduce a centralized version of such mechanism. We consider the notion of *supervisor of social equilibrium*, a component of the society (possibly an agent) that is able to determine, at each time, the set of compensation actions that may be performed in order to bring the social system back to the equilibrium (or, disequilibrium), regarding the balances of exchange values, and that may recommend some of such actions to the agents of the system, in order to get that equilibrium (or, disequilibrium).

Obviously, supervisors of social equilibrium do not implement moral rules, because moral rules, in the sense defined above, can only be implemented inside the agents themselves, not in a component of the society that is external to them.

So, supervisors of social equilibrium implement juridical rules (laws, norms). However, they do so in a way that they are not *law enforcers*, because we don't require that agents follow the recommendations: agents are allowed to autonomously decide if they are to follow, or not, any given recommendation.

4 A Model Supervisor of Social Equilibrium

In this section, we introduce a formal model for a supervisor of social equilibrium, which is able to implement a regulation process for the equilibration (or, disequilibration) of exchange values.

The regulation process is embedded in the equilibrium supervisor in the form of a *recommendation policy*, which determines for each kind of exchange values state an appropriate compensation action.

However, we don't consider here the connection between the recommendation policy that implements the regulation process, and the expression of such regulation process in the form of juridical rules. That is,

everything works in the model as if the juridical rules were pre-compiled into its recommendation policy, by way of the reward function of the Markov Decision Process solved by the equilibrium supervisor.

Further work is necessary on this compilation process both to understand the connection between the language of juridical laws and its translation in terms of exchange values, and to allow for a dynamic compilation process, capable of supporting dynamic changes in juridical laws, in order to model organizations supporting evolutive juridical systems.

4.1 Using Interval Mathematics for Representing Social Values

Interval Mathematics is a mathematical theory introduced in the 1960's by Moore [26] that aims at the automatic and rigorous control of the errors that arise in numerical computations. Any real number $x \in \mathbb{R}$ is represented by a real interval $X = [x_1, x_2]$, with $x_1, x_2 \in \mathbb{R}$, such that $x_1 \leq x \leq x_2$. x_1 and x_2 denote, respectively, the left and right endpoints of a real interval X . The set of real intervals is denoted by \mathbb{IR} .

The arithmetical operations $*_{\mathbb{IR}} \in \{+, -, \times \div\}$ are defined on \mathbb{IR} as $X *_{\mathbb{IR}} Y = \{x * y \mid x \in X, y \in Y\}$ ⁹, and they can be explicitly calculated by [27]:

$$\text{Addition: } X + Y = [x_1 + y_1, x_2 + y_2]$$

$$\text{Subtraction: } X - Y = [x_1 - y_2, x_2 - y_1]$$

$$\text{Product: } X \times Y = [\min \rho, \max \rho],$$

$$\text{with } \rho = \{x_1 y_1, x_1 y_2, x_2 y_1, x_2 y_2\}$$

$$\text{Quotient: } X \div Y = [\min \sigma, \max \sigma],$$

$$\text{with } \sigma = \left\{ \frac{x_1}{y_1}, \frac{x_1}{y_2}, \frac{x_2}{y_1}, \frac{x_2}{y_2} \right\} \text{ and } 0 \notin Y.$$

The interval arithmetic operations satisfy the *monotonic inclusion property*, for all $X', X'', Y', Y'' \in \mathbb{IR}$ [1, 27]: $X' \subseteq X'', Y' \subseteq Y'' \Rightarrow X' * Y' \subseteq X'' * Y''$. This property plays an important role in the usage of Interval Mathematics to compute with real numbers that are uncertain for some reason (e.g., if they are obtained by a measuring instrument with limited resolution) or that are not exactly representable in a floating point system (e.g., the numbers $\pi, 0.1, \frac{1}{3}$).

A machine interval has floating point numbers as endpoints and outward roundings are used to guarantee that the resulting output interval of any interval computation process contains the actual result [17]. Besides that, the range of the output interval is the

⁹Whenever it can be understood from the context, we shall not use the notation \mathbb{IR} to distinguish interval operations.

indicative of the maximum error that may have occurred in the whole process.

Interval Mathematics has also been applied to represent other kinds of uncertainty rather than numerical uncertainty [16], with applications in Artificial Intelligence, Soft Computing etc.

In this paper, intervals are used to capture the qualitative nature of Piaget's concept of *scale of exchange values* [31]. The chosen representation is a compromise between a purely qualitative and a purely quantitative representation. It makes the representation mathematically operational, and the decision process computationally viable, without being unfaithful to Piaget's approach.

Let $\mathbb{IR}_L = \{[x_1, x_2] \mid -L \leq x_1 \leq x_2 \leq L\}$ be the set of real intervals bounded by $L \in \mathbb{R}$ ($L > 0$) and let $\mathcal{IR}_L = (\mathbb{IR}_L, +, \Theta, \sim, \approx)$ be a *scale of interval exchange values*, where:

- $+$: $\mathbb{IR}_L \times \mathbb{IR}_L \rightarrow \mathbb{IR}_L$ is the L -bounded addition operation defined by

$$X + Y = [\max\{x_1 + y_1, -L\}, \min\{x_2 + y_2, L\}].$$

- A *null value* is any $X \in \mathbb{IR}_L$ such that $\text{mid}(X) = 0$, where $\text{mid}(X) = \frac{x_1 + x_2}{2}$ is the mid point of X . The set of null values is denoted by Θ . $0 = [0, 0]$ is the *absolute null value*.

- A *quasi-symmetric value* of $X \in \mathbb{IR}_L$ is any interval $X' \in \mathbb{IR}_L$ such that $X + X' \in \Theta$. The set of quasi-symmetric values of X is denoted by \tilde{X} .

- \approx is the *qualitative equivalence relation* defined by $X \approx Y \Leftrightarrow \exists Y' \in \tilde{Y} : X + Y' \in \Theta$.

$\mu\tilde{X} \in \tilde{X}$ is said to be the *least quasi-symmetric value* of X if whenever there exists $S \in \tilde{X}$ it holds that $d(\mu\tilde{X}) \leq d(S)$, where $d(X) = x_2 - x_1$ is the diameter of X . For all $X \in \mathbb{IR}_L$, it follows that:

Proposition 1 (i) $\tilde{X} = \{-[\text{mid}(X) - k, \text{mid}(X) + k] \mid k \in \mathbb{R} \wedge k \geq 0\}$; (ii) $\mu\tilde{X} = -[\text{mid}(X), \text{mid}(X)]$.

Proof. It holds that $\text{mid}(X + [-(\text{mid}(X) + k), -(\text{mid}(X) - k)]) = \text{mid}([\frac{x_1 - x_2 - 2k}{2}, \frac{x_2 - x_1 + 2k}{2}]) = 0$. Consider $S \in \tilde{X}$ such that $\text{mid}(S) \neq \text{mid}(X)$. For $k_1 \neq k_2 \in \mathbb{R}$, it follows that $\text{mid}(X + S) = \text{mid}(X + [-(\text{mid}(X) + k_1), -(\text{mid}(X) - k_2)]) = \text{mid}([\frac{x_1 - x_2 - 2k_1}{2}, \frac{x_2 - x_1 + 2k_2}{2}]) = \frac{k_2 - k_1}{2} \neq 0$, which is a contradiction. It follows that any quasi-symmetric value of X is of the form $[-(\text{mid}(X) + k), -(\text{mid}(X) - k)]$ and, when $k = 0$, $\mu\tilde{X} = [-\text{mid}(X), -\text{mid}(X)]$ is obtained. \square

In practical applications, due to the rounding errors that arise in any numerical computation, it is usually not possible to verify whether or not a computed interval is precisely the absolute null value 0 [1, 17, 18, 27]. Then, we shall introduce the concept of *absolute ϵ -null value* $0_\epsilon = [-\epsilon, +\epsilon]$, with $\epsilon \in \mathbb{R}$ ($\epsilon \geq 0$) being a given tolerance. In this case, an ϵ -null value is any $X \in \mathbb{IR}_L$ such that $\text{mid}(X) \in 0_\epsilon$. The set of ϵ -null values is denoted by Θ_ϵ . The related set of ϵ -quasi-symmetric values for an interval value $X \in \mathbb{IR}_L$ is denoted by \tilde{X}_ϵ , and, from Prop. 1, it follows that the least ϵ -quasi-symmetric value is given by

$$\mu\tilde{X}_\epsilon = [-(\text{mid}(X) + \epsilon), -(\text{mid}(X) - \epsilon)]. \quad (1)$$

The *qualitative equivalence relation* (module ϵ) is then defined by $X \approx_\epsilon Y \Leftrightarrow \exists Y' \in \tilde{Y}_\epsilon : X + Y' \in \Theta_\epsilon$.

4.2 The Modelling of Social Exchanges

Let T be a set of discrete instants of time. Let α and β be any two agents. A *qualitative interval exchange-value system* for modelling the exchanges from α to β is a structure $\mathbf{IR}_{\alpha\beta} = (\mathcal{IR}_L; r_{I_{\alpha\beta}}, r_{II_{\beta\alpha}}, s_{I_{\beta\alpha}}, s_{II_{\alpha\beta}}, t_{I_{\beta\alpha}}, t_{II_{\beta\alpha}}, v_{I_{\alpha\beta}}, v_{II_{\alpha\beta}})$, where

$$r_{I_{\alpha\beta}}, r_{II_{\beta\alpha}} : T \rightarrow \mathbb{IR}_L, s_{II_{\alpha\beta}}, s_{I_{\beta\alpha}} : T \rightarrow \mathbb{IR}_L, \quad (2)$$

$$t_{I_{\beta\alpha}}, t_{II_{\beta\alpha}} : T \rightarrow \mathbb{IR}_L, v_{I_{\alpha\beta}}, v_{II_{\alpha\beta}} : T \rightarrow \mathbb{IR}_L \quad (3)$$

are partial functions, called *exchange-value functions*, that evaluate, at each time instant $t \in T$, the investment, satisfaction, debt and credit values¹⁰, respectively, involved in the exchange. The symbol \perp denotes an undefined exchange value. In the following, we use the notation $r_{I_{\alpha\beta}}(t) = r_{I_{\alpha\beta}}^t$, $r_{II_{\beta\alpha}}(t) = r_{II_{\beta\alpha}}^t$, $s_{II_{\alpha\beta}}(t) = s_{II_{\alpha\beta}}^t$, $s_{I_{\beta\alpha}}(t) = s_{I_{\beta\alpha}}^t$, $t_{I_{\beta\alpha}}(t) = t_{I_{\beta\alpha}}^t$, $t_{II_{\beta\alpha}}(t) = t_{II_{\beta\alpha}}^t$, $v_{I_{\alpha\beta}}(t) = v_{I_{\alpha\beta}}^t$ and $v_{II_{\alpha\beta}}(t) = v_{II_{\alpha\beta}}^t$.

For the exchange-value functions given in (2) and (3), at a given time instant t , the following constraints must be satisfied for every pair of agents α and β :

$$r_{I_{\alpha\beta}}^t = \perp \Rightarrow s_{I_{\beta\alpha}}^t = t_{I_{\beta\alpha}}^t = v_{I_{\alpha\beta}}^t = \perp, \quad (4)$$

$$v_{II_{\alpha\beta}}^t = \perp \Rightarrow t_{II_{\beta\alpha}}^t = r_{II_{\beta\alpha}}^t = s_{II_{\alpha\beta}}^t = \perp, \quad (5)$$

$$r_{I_{\alpha\beta}}^t \neq \perp \Rightarrow v_{II_{\alpha\beta}}^t = \perp, \quad (6)$$

where:

- $r_{I_{\alpha\beta}}^t = \perp$ denotes that the agent α did not perform a service for the agent β at time t , and, therefore, all the other corresponding exchange values in the stage I resulted undefined;

¹⁰The values are undefined if no service is done at all at a given moment $t \in T$.

- $v_{\text{II}\alpha\beta}^t = \perp$ denotes that the agent α , at time t , did not charge the credit for a service previously done for the agent β , and, therefore, all the other corresponding exchange values in the stage II resulted undefined.

The implication (6) means that, according to the structure of social exchanges (Fig. 1), it is not possible for an agent α to perform a service for β and, at the same time t , to charge him a credit. From (6) it follows that it is also required that $v_{\text{II}\alpha\beta}^t \neq \perp \Rightarrow r_{\text{I}\alpha\beta}^t = \perp$.

A *configuration of exchange values* for any pair of agents α and β at a time instant t is specified by one of the tuples of well defined exchange values:

$$(r_{\text{I}\alpha\beta}^t, s_{\text{I}\beta\alpha}^t, t_{\text{I}\beta\alpha}^t, v_{\text{I}\alpha\beta}^t), (r_{\text{I}\beta\alpha}^t, s_{\text{I}\alpha\beta}^t, t_{\text{I}\alpha\beta}^t, v_{\text{I}\beta\alpha}^t), \\ (v_{\text{II}\alpha\beta}^t, t_{\text{II}\beta\alpha}^t, r_{\text{II}\beta\alpha}^t, s_{\text{II}\alpha\beta}^t), (v_{\text{II}\beta\alpha}^t, t_{\text{II}\alpha\beta}^t, r_{\text{II}\alpha\beta}^t, s_{\text{II}\beta\alpha}^t).$$

A *social exchange process* between two agents α and β of a multi-agent system, occurring during the time instants $T = t_1, \dots, t_n$, is any finite sequence of such configurations of exchange values e^{t_1}, \dots, e^{t_n} . Each element of this sequence is called a *stage* of the exchange process, which may be a stage of type I or type II.

The *exchange balance* of stages of type I of a social exchange process between any pair of agents α and β that has occurred during a time interval T is a tuple

$$b_{\text{I}\{\alpha,\beta\}}^T = \left(r_{\text{I}\alpha\beta}^T, r_{\text{I}\beta\alpha}^T, s_{\text{I}\alpha\beta}^T, s_{\text{I}\beta\alpha}^T \mid t_{\text{I}\alpha\beta}^T, t_{\text{I}\beta\alpha}^T, v_{\text{I}\alpha\beta}^T, v_{\text{I}\beta\alpha}^T \right), \quad (7)$$

where, for $k = r, s, t, v$,

$$k_{\text{I}\alpha\beta}^T = \sum_{t \in T} k_{\text{I}\alpha\beta}^t \quad \text{and} \quad k_{\text{I}\beta\alpha}^T = \sum_{t \in T} k_{\text{I}\beta\alpha}^t,$$

for all $k_{\text{I}\alpha\beta}^t \neq \perp$ and $k_{\text{I}\beta\alpha}^t \neq \perp$. The exchange balance of stages of type II, denoted by $b_{\text{II}\{\alpha,\beta\}}^T$, is defined analogously. The *general exchange balance* is then given by

$$b_{\{\alpha,\beta\}}^T = b_{\text{I}\{\alpha,\beta\}}^T + b_{\text{II}\{\alpha,\beta\}}^T.$$

The *material results* $\mathbf{m}_{\alpha\beta}$ and $\mathbf{m}_{\beta\alpha}$ of a social exchange process that happens between agents α and β , during the interval T , according to the points of view of α and β , respectively, are given by the sum of the respective material values involved in the process:

$$\mathbf{m}_{\alpha\beta}^T = r_{\text{I}\alpha\beta}^T + s_{\text{II}\alpha\beta}^T + r_{\text{II}\alpha\beta}^T + s_{\text{I}\alpha\beta}^T, \quad (8)$$

$$\mathbf{m}_{\beta\alpha}^T = r_{\text{I}\beta\alpha}^T + s_{\text{II}\beta\alpha}^T + r_{\text{II}\beta\alpha}^T + s_{\text{I}\beta\alpha}^T. \quad (9)$$

Analogously, the *virtual results* $\mathbf{v}_{\alpha\beta}$ and $\mathbf{v}_{\beta\alpha}$ are given by:

$$\mathbf{v}_{\alpha\beta}^T = t_{\text{I}\alpha\beta}^T + v_{\text{II}\alpha\beta}^T + t_{\text{II}\alpha\beta}^T + v_{\text{I}\alpha\beta}^T, \quad (10)$$

$$\mathbf{v}_{\beta\alpha}^T = t_{\text{I}\beta\alpha}^T + v_{\text{II}\beta\alpha}^T + t_{\text{II}\beta\alpha}^T + v_{\text{I}\beta\alpha}^T. \quad (11)$$

The *general results* take into account all kinds of exchanges values, and is obtained by:

$$\mathbf{g}_{\alpha\beta}^T = \mathbf{m}_{\alpha\beta}^T + \mathbf{v}_{\alpha\beta}^T, \quad \mathbf{g}_{\beta\alpha}^T = \mathbf{m}_{\beta\alpha}^T + \mathbf{v}_{\beta\alpha}^T. \quad (12)$$

A social exchange process between a pair of agents α and β is said to be in *equilibrium* (with tolerance $\epsilon \geq 0$) if

$$\mathbf{g}_{\alpha\beta}^T \in \Theta_\epsilon \quad \text{and} \quad \mathbf{g}_{\beta\alpha}^T \in \Theta_\epsilon.$$

The *material equilibrium* is achieved when

$$\mathbf{m}_{\alpha\beta}^T \in \Theta_\epsilon \quad \text{and} \quad \mathbf{m}_{\beta\alpha}^T \in \Theta_\epsilon.$$

4.3 Modelling Social Exchanges Involving Multiple Agents

In this section, a matrix-like notation is introduced to make possible the generalization of the results concerning the social exchanges between two agents, presented in Sect. 4.2, for the case of an organization composed by m agents. An $m \times m$ interval \star -matrix $[x_{ij}]^\star$ is defined as the interval $m \times m$ matrix $[x_{ij}]$ where $x_{ij} = \star$ whenever $i = j$, that is:

$$[x_{ij}]^\star = \begin{pmatrix} \star & x_{12} & \cdots & x_{1m} \\ x_{21} & \star & \cdots & x_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ x_{m1} & x_{m2} & \cdots & \star \end{pmatrix}$$

A 2×2 \star -matrix is denoted schematically by an ordered pair

$$(x_{12}, x_{21})^\star = \begin{pmatrix} \star & x_{12} \\ x_{21} & \star \end{pmatrix} \quad (13)$$

For two interval $m \times m$ \star -matrices $X = [x_{ij}]^\star$ and $Y = [y_{ij}]^\star$, we define:

- The *addition* operation: $X + Y = [x_{ij} + y_{ij}]^\star$.
- Any \star -matrix $[n_{ij}]^\star$ such that $n_{ij} \in \Theta_\epsilon$ is an ϵ -null \star -matrix. The set of such \star -matrices is denoted by N_ϵ .
- The *qualitative equivalence relation* (module ϵ): $X \approx_\epsilon Y \Leftrightarrow (\forall i \neq j) x_{ij} \approx_\epsilon y_{ij}$.

In a multi-agent system composed by m agents, the exchange values determined by the functions defined in (2) and (3) can be represented by the eight $m \times m \times \#T$ interval \star -matrices

$$K_{\text{I}} = \left[k_{\text{I}\alpha\beta}^t \right]^\star \quad \text{and} \quad K_{\text{II}} = \left[k_{\text{II}\alpha\beta}^t \right]^\star,$$

called *investment* ($K = R$), *satisfaction* ($K = S$), *debt* ($K = T$) and *credit* ($K = V$) *matrices* for the social

exchange stages I and II, respectively, that occurred between each two agents α and β in T .¹¹

For each $t' \in T$, $K_I^{t'} = [k_{\alpha\beta}^{t'}]^*$ and $K_{II}^{t'} = [k_{II\alpha\beta}^{t'}]^*$ are $m \times m$ interval \star -matrices. For a time sequence $T = t_1, \dots, t_n$, the four $m \times m$ matrices of *global* investment, satisfaction, debt and credit in T are given by

$$K_I^T = \sum_{t=t_1}^{t_n} K_I^t \text{ and } K_{II}^T = \sum_{t=t_1}^{t_n} K_{II}^t,$$

for $K = R, S, T, V$.

The exchange balance of stages of type I, given in (7), is represented by a tuple of \star -matrices:

$$B_I^T = (R_I^T, S_I^T | T_I^T, V_I^T).$$

Analogously, the exchange balance of stages of type II is represented as a tuple B_{II}^T . Therefore, the general exchange balance can be represented by the tuple:

$$B^T = (R^T, S^T | T^T, V^T),$$

where, for $K = R, S, T, V$, $K^T = K_I^T + K_{II}^T$.

The material, virtual and general results of social exchange processes in a multi-agent system, given in (8–12), are then evaluated by

$$\mathbf{M}^T = R^T + S^T, \mathbf{V}^T = T^T + V^T \text{ and } \mathbf{G}^T = M^T + V^T,$$

respectively.

Thus, a multi-agent system is said to be in equilibrium if

$$\mathbf{G}^T \in N_\epsilon.$$

It is in material equilibrium when

$$\mathbf{M}^T \in N_\epsilon.$$

If a social exchange process between any pair of agents α and β , which occurred in T , is not in material equilibrium according α 's point of view, that is, $\mathbf{m}_{\alpha\beta}^T \notin \Theta_\epsilon$, then the least ϵ -quasi-symmetric of $\mathbf{m}_{\alpha\beta}^T$, given in (1), is said to be the *compensation* value from α 's point of view. The compensation value for $\mathbf{m}_{\alpha\beta}^T \in \Theta_\epsilon$ is the interval $[0, 0]$. A *compensation \star -matrix* for a \star -matrix of material results \mathbf{M}^T is the \star -matrix \mathbf{M}' whose each entry is the compensation value of the corresponding entry of \mathbf{M}^T . It follows that $\mathbf{M}^T + \mathbf{M}' \in N_\epsilon$.

¹¹To extend this representation to *interactions between two groups A and B of agents* one should take into account the sum total of the interactions of every agent of A with every agent of B, e.g., $K_I = [k_{IAB}^t]^*$, where $k_{IAB}^t = \sum_{\alpha \in A, \beta \in B} k_{I\alpha\beta}^t$.

4.4 Solving the Equilibration Problem Using a QI-MDP

We conceive, in the context of a social exchange process in a multi-agent system, a special agent, called *equilibrium supervisor*, which analyzes the exchange processes between each pair of agents and makes suggestions of exchanges to each two agents in order to keep the material results of exchanges in equilibrium. The equilibrium supervisor also takes into account the virtual results of the exchanges in order to decide which type of exchange stage he shall suggest for the two agents.

To achieve that purpose, the equilibrium supervisor models the exchanges between each pair of agents as simultaneous Markov Decision Processes (MDP) [37], where the *states* of the models represent “possible material results of the overall exchanges” and the *optimal policies* represent “sequences of *actions* that the equilibrium supervisor recommends that the interacting agents execute”.

In the following, the Qualitative Interval Markov Decision Process (QI-MDP) is introduced. An initial formulation of this model, considering a society of just two agents, was presented in [14].

4.4.1 The Basis of a QI-MDP for a Multi-Agent System

Consider an admissible tolerance $\epsilon \geq 0$, a bound $L \in \mathbb{R}$ ($L > 0$) for the set of real intervals $\mathbb{I}\mathbb{R}_L$ and $n \in \mathbb{N}$ ($n > 0$). Let $\hat{E} = \{E^{-n}, \dots, E^n\}$ be the set of $2n + 1$ equivalence classes of intervals $X \in \mathbb{I}\mathbb{R}_L$, defined, for $i = -n, \dots, n$, as:

$$E^i = \begin{aligned} & \{X \mid i \frac{L}{n} \leq \text{mid}(X) < (i+1) \frac{L}{n}\} && \text{if } -n \leq i < -1 \\ & \{X \mid -\frac{L}{n} \leq \text{mid}(X) < -\epsilon\} && \text{if } i = -1 \\ & \{X \mid -\epsilon \leq \text{mid}(X) \leq +\epsilon\} && \text{if } i = 0 \\ & \{X \mid \epsilon < \text{mid}(X) \leq \frac{L}{n}\} && \text{if } i = 1 \\ & \{X \mid (i-1) \frac{L}{n} < \text{mid}(X) \leq i \frac{L}{n}\} && \text{if } 1 < i \leq n. \end{aligned} \quad (14)$$

The classes $E^i \in \hat{E}$ are the supervisor representations of classes of unfavorable ($i < 0$), equilibrated ($i = 0$) and favorable ($i > 0$) material results of exchange social processes. Whenever it is understood from the context, we shall denote by E^- (or E^+) any class $E_{i < 0}^i$ (or $E_{i > 0}^i$).

The *accuracy* of the equilibrium supervisor is given by $\kappa_n = \frac{L}{n}$. The range of the midpoints of the intervals that belong to a class E^i is called the *representative* of the class E^i . Whenever it is clear from the context, we shall identify a class E^i with its representative.

The states of the QI-MDP model are \star -matrices $E = [E_{\alpha\beta}^i]^\star$, where each entry $E_{\alpha\beta}^i \in \hat{E}$ is the class representing the material results of the social exchange process between α and β , from the point of view of the agent α . For the analysis of the equilibrium, we shall consider each pair of co-related classes of material results $(E_{\alpha\beta}^i, E_{\beta\alpha}^j)$.

The \star -matrix $[E_{\alpha\beta}^0]^\star$ is the *terminal* state, representing that the system is in equilibrium. However, in some applications, it may be considered just a subset B of agents for the analysis of the equilibrium, even when all agents are involved in the exchanges processes. In this case, the terminal state is such that $E_{\alpha\beta}^i = E_{\beta\alpha}^j = E^0$, for all $\alpha, \beta \in B$.

The *actions* considered in the model are state transitions

$$[E_{\alpha\beta}^i]^\star \xrightarrow{[A_{\alpha\beta}^i]^\star} [E_{\alpha\beta}^{i'}]^\star,$$

with $E_{\alpha\beta}^i, E_{\alpha\beta}^{i'} \in \hat{E}$, where $[A_{\alpha\beta}^i]^\star$ is an interval \star -matrix operator such that $\text{mid}(E_{\alpha\beta}^i + A_{\alpha\beta}^i) \in E_{\alpha\beta}^{i'}$.

$A_{\alpha\beta}^i$ is an *interval action* that should be one of the following types:

- a *compensation interval* C^i , which is the least quasi-symmetric of a class representative E^i ;
- a *go-forward- k -step interval* F_k^i , which is an interval that transforms a class E^i into E^{i+k} , with $i \neq L$ and $i+k \neq 0$;
- a *go-backward- k -step interval* B_{-k}^i , which is an interval that transforms a class E^i into E^{i-k} , with $i \neq -L$ and $i-k \neq 0$.

The set of compensation intervals, denoted by \mathcal{C} , is shown in Table 1. The set \mathcal{F} of go-forward intervals and their respective effects are partially presented in Table 2. The set of go-backward intervals, denoted by \mathcal{B} , can be specified analogously.

Table 1: Specification of compensation intervals

State	Compensation Interval $C^i \in \mathcal{C}$
$E_{-n \leq i < -1}^i$	$[-\frac{(2i+1)L}{2n} - \epsilon, -\frac{(2i+1)L}{2n} + \epsilon]$
E^{-1}	$[\frac{1}{2}(\frac{L}{n} + \epsilon) - \epsilon, \frac{1}{2}(\frac{L}{n} + \epsilon) + \epsilon]$
E^0	$[0, 0]$
E^1	$[-\frac{1}{2}(\frac{L}{n} + \epsilon) - \epsilon, -\frac{1}{2}(\frac{L}{n} + \epsilon) + \epsilon]$
$E_{1 < i \leq n}^i$	$[\frac{(1-2i)L}{2n} - \epsilon, \frac{(1-2i)L}{2n} + \epsilon]$

For example, consider the case of a society with just two agents α and β . For the pair of classes of material

results (using the notation given in (13))

$$(E_{\alpha\beta}^i, E_{\beta\alpha}^j)_{-n \leq i < -1, 1 < j \leq n}^\star \equiv \left(\left[i\frac{L}{n}, (i+1)\frac{L}{n} \right], \left[(j-1)\frac{L}{n}, j\frac{L}{n} \right] \right)^\star,$$

it follows that the *compensation-compensation* action $(C^i, C^j)^\star$ and the *go-backward $_3$ -go-forward $_{+2}$* $(B_{-3}^i, F_{+2}^j)^\star$ action are specified by

$$C^i = \left[-\frac{2i+1}{2}\frac{L}{n} - \epsilon, -\frac{2i+1}{2}\frac{L}{n} + \epsilon \right], \quad (15)$$

$$C^j = \left[\frac{(1-2j)L}{2n} - \epsilon, \frac{(1-2j)L}{2n} + \epsilon \right], \quad (16)$$

$$B_{-3}^i = \left[-3\frac{L}{n} - \epsilon, -3\frac{L}{n} + \epsilon \right], \quad (17)$$

$$F_{+2}^j = \left[2\frac{L}{n} - \epsilon, 2\frac{L}{n} + \epsilon \right], \quad (18)$$

respectively, resulting in the following state transitions, with $-n \leq i < -1$ and $1 < j \leq n$:

$$(E_{\alpha\beta}^i, E_{\beta\alpha}^j)^\star \xrightarrow{(C^i, C^j)^\star} (E_{\alpha\beta}^0, E_{\beta\alpha}^0)^\star$$

$$(E_{\alpha\beta}^i, E_{\beta\alpha}^j)^\star \xrightarrow{(B_{-3}^i, F_{+2}^j)^\star} (E_{\alpha\beta}^{(i-3)}, E_{\beta\alpha}^{(j+2)})^\star.$$

The equilibrium supervisor has to find, for each state $[E^i]^\star$, the action that shall achieve the terminal state or, at least, another state from where the terminal state can be achieved, with the least number of steps and least final *value uncertainty*¹². We observe that the choice of such actions are also regulated by the rules of the social exchanges, and, therefore, there are some state transitions that are not allowed.

Based on an optimal policy, the equilibrium supervisor may be asked to recommend that the agents act optimally. An *optimal exchange recommendation* consists of a function that gives, for each actual material result (represented by a state of the model), a partially defined exchange stage that shall restore or establish the material equilibrium or, at least, give conditions that it be achieved in a least number of steps with least value uncertainty. This partial definition shall be completed by the analysis of the virtual results, which allows the specification of which particular types of exchange stages (I or II) should be considered.

Although the interacting agents acknowledge the optimal recommendations from the equilibrium supervisor, they are autonomous in the sense that *they may*

¹²By value uncertainty of a state $[E^i]^\star$ we mean the diameters of the intervals E^i .

Table 2: Specification of some go-forward intervals and their respective effects

State	Go-forward interval $F_{+k}^i \in \mathcal{F}$	Effect
$E_{-n \leq i < -2}^i$	$[k\frac{L}{n} - \epsilon, k\frac{L}{n} + \epsilon]_{0 < k \leq -i-2}$	$E^i \mapsto E_{i < i+k \leq -2}^{i+k}$
E^{-2}	$[\frac{L}{n} - \epsilon, \frac{L}{n} + \epsilon]$	$E^{-2} \mapsto E^{-1}$
$E_{-n \leq i < -1}^i$	$[k\frac{L}{n} - \epsilon, k\frac{L}{n} + \epsilon]_{1-i \leq k \leq n-i-1}$	$E^i \mapsto E_{1 < i+k \leq n}^{i+k}$
$E_{-n \leq i < -1}^i$	$[-i\frac{L}{n} + \epsilon, -i\frac{L}{n} + \epsilon]$	$E^i \mapsto E^1$
E_{-1}^{-1}	$[\frac{L}{n} + \epsilon, \frac{L}{n} + \epsilon]$	$E^{-1} \mapsto E^1$
E^{-1}	$[k\frac{L}{n} - \epsilon, k\frac{L}{n} + 2\epsilon]_{2 \leq k \leq n}$	$E^{-1} \mapsto E_{1 < k-1 \leq n}^{k-1}$
E^0	$[2\epsilon, \frac{L}{n}]$	$E^0 \mapsto E^1$
E^0	$[k\frac{L}{n}, (k+1)\frac{L}{n}]_{0 < k \leq n-1}$	$E^0 \mapsto E_{1 < k+1 \leq n}^{k+1}$
E^1	$[k\frac{L}{n} - 2\epsilon, k\frac{L}{n} + \epsilon]_{0 < k \leq n-1}$	$E^1 \mapsto E_{1 < k+1 \leq n}^{k+1}$
$E_{1 < i \leq n}^i$	$[k\frac{L}{n} - \epsilon, k\frac{L}{n} + \epsilon]_{0 < k \leq n-i}$	$E^i \mapsto E_{i < i+k \leq n}^{i+k}$

not follow the recommendations exactly. The agents may have different personalities, interests, needs etc., which may lead them not to always follow the recommendations. This means that there is a probability that the system achieves another state different from the suggested by the supervisor and, therefore, there may be a great deal of uncertainty about the effects of the agents actions.

Even if the agents follow a recommendation exactly, we will show that the effect may not be the expected by the supervisor, since it depends on the ratio $\frac{\kappa_n}{\epsilon}$, where $\kappa_n = \frac{L}{n}$ is the equilibrium supervisor accuracy and $\epsilon \in \mathbb{R}$ ($0 \leq \epsilon < \kappa_n$) is the admissible tolerance. On the other hand, in this paper, we assume that there is never any uncertainty about the current state of the system, that is, the equilibrium supervisor always has access to the current configuration of exchange values and has complete and perfect abilities to evaluate the current material balance.

Definition 2 A Qualitative Interval Markov Decision Process (QI-MDP), for keeping in equilibrium the social exchanges in a multi-agent systems of m agents, is a tuple $\langle E, A, F, R \rangle_{\epsilon}^{L, n}$, where:

- The set of the states of the model is the set of $m \times m$ star-matrices $\mathbf{E} = \{ [E_{\alpha\beta}^i]^\star \mid E_{\alpha\beta}^i \in \hat{E} \}$ of classes of material results as specified in (14).
- The set of the actions of the model is the set of $m \times m$ star-matrices $\mathbf{A} = \{ [A_{\alpha\beta}^i]^\star \mid A_{\alpha\beta}^i \in \mathcal{C} \cup \mathcal{F} \cup \mathcal{B} \}$, of compensation $C^i \in \mathcal{C}$, go-forward $F_{+k}^i \in \mathcal{F}$ and go-backward $B_{-k}^i \in \mathcal{B}$ intervals.
- $\mathbf{F} : \mathbf{E} \times \mathbf{A} \rightarrow \Pi(\mathbf{E})$ is the state-transition function, which gives for each state and each action, a probability distribution over the set of states;

- $\mathbf{R} : (\mathbf{E} \times \mathbf{A}) \rightarrow \mathbb{R}$ is the reward function, giving the expected immediate reward gained by choosing an action $[A_{\alpha\beta}^i]^\star$ when the current state is $[E_{\alpha\beta}^i]^\star$.

In this model, the next state and the expected reward depend only on the previous state and the action taken, satisfying the so-called *Markov property*.

4.4.2 The Optimal Policy and the Reward Function

The reward function plays an important role when the equilibrium supervisor is choosing the action that will generate a recommendation of agents interaction, in each state. The supervisor aims to maximize the utility of sequences of actions, evaluated according to the reward function, thus trying to bring the system back to equilibrium (or keep it there) with the least possible number of state transitions.

A sample reward function $\mathbf{R} : (\mathbf{E} \times \mathbf{A}) \rightarrow \mathbb{R}$ that conforms to the idea of supporting a recommendation function that is able to direct pairs of agents into social equilibrium is partially sketched in Table 3, using the notation given in (13). This particular sample function illustrates various requirements that should be satisfied by any reward function of the model.

Observe, for instance, that if the current state is of the type $(E^-, E^+)^\star$, then the best action to be chosen is the *compensation-compensation* action $(C, C)^\star$, which results in a state transition $(E^-, E^+)^\star \mapsto (E^0, E^0)^\star$. Any other choice should make the agents either take a long way to the equilibrium or get away from it.

On the other hand, if the current state is of type $(E^-, E^-)^\star$, then a *compensation-compensation* action $(C, C)^\star$ would generate a recommendation of agent

exchanges of *satisfaction-satisfaction* type, which is impossible according to the model of social interactions [33], since it is impossible for an agent to get a satisfaction value from no service at all. The reward function R should state that $(C, C)^*$ is a very bad action to be chosen in such situation.

Any optimal policy $\pi^* : \mathbf{E} \rightarrow \mathbf{A}$ solving the social equilibrium problem should satisfy the set of requirements expressed by the schema partially sketched in Table 4 (for a pair of agents). Notice that it is a non deterministic policy.

The *optimal value recommendation* associated to an optimal policy π^* is a \star -matrix operator ρ_{π^*} that gives, for each state $[E_{\alpha\beta}^i]^*$ and optimal action $\pi^* [E_{\alpha\beta}^i]^* = [A_{\alpha\beta}^i]^*$, recommendations of exchange stages, possibly partially undefined, consisting of \star -matrices whose elements in symmetric positions are either $(r_{\alpha\beta}, A_{\alpha\beta}^i)$ and $(s_{\beta\alpha}, A_{\beta\alpha}^j)$, or $(s_{\alpha\beta}, A_{\alpha\beta}^i)$ and $(r_{\beta\alpha}, A_{\beta\alpha}^j)$, where $(r_{\lambda\delta}, W)$ means the performance, by the agent λ , of a service with investment value $W < 0$, and $(s_{\delta\lambda}, W')$ means δ 's satisfaction with interval value W' . The optimal value recommendation ρ_{π^*} , corresponding to the the optimal policy shown in Table 4, is partially sketched in Table 5.

Finally, the equilibrium supervisor has to decide which types of exchange stages (I or II) should be recommended. This is done by the analysis of the virtual results from the points of view of each pair of agents α (given in (10)) and β (given in (11)):

- If $\mathbf{v}_{\alpha\beta} > 0$, then α is able to charge β the credit for services previously done. Thus, an exchange stage of type $\text{II}_{\alpha\beta}$ should be recommended.
- If $\mathbf{v}_{\beta\alpha} > 0$, then it is the case that the agent β can charge α the credit for services previously done, indicating that an exchange stage of type $\text{II}_{\beta\alpha}$ should then be recommended.
- If $\mathbf{v}_{\alpha\beta} \leq 0$, then the agent α does not have any credit to charge α . Therefore, the service done by β must be spontaneous. In this case, an exchange stage of type $\text{I}_{\beta\alpha}$ should then be recommended.
- If $\mathbf{v}_{\beta\alpha} \leq 0$, then the agent β does not have any credit to charge β , resulting that an exchange stage of type $\text{I}_{\alpha\beta}$ should then be chosen.

Table 6 shows the criteria used by the equilibrium supervisor to reason about the possible *stage recommendations*, based on virtual results, according to the discussion presented in the paragraph above. Observe

that these alternatives are not mutually exclusive. The final decision of which type of exchange stage shall be executed is let to the agents to decide.

Table 6: Criteria for deciding on types of exchange stages to be recommended

Virtual Result	Type of Exchange Stage	Label
$\mathbf{v}_{\alpha\beta} > 0$	$\text{II}_{\alpha\beta}$	T^1
$\mathbf{v}_{\beta\alpha} > 0$	$\text{II}_{\beta\alpha}$	T^2
$\mathbf{v}_{\alpha\beta} \leq 0$	$\text{I}_{\beta\alpha}$	T^3
$\mathbf{v}_{\beta\alpha} \leq 0$	$\text{I}_{\alpha\beta}$	T^4

The stage effects of the recommendations (Table 6) are sketched in the simplified state transition diagram shown in Fig. 2. The dot lines represent alternative paths to the equilibrium state that were not considered as optimal recommendations since they are considered impossible according to the structure of the exchange stages. The symbol \star of the \star -matrices was omitted.

5 Analysis of the Model

For the analysis of the model, we consider a multi-agent organization whose agents are classified by the degree in which they follow the recommendations given by the equilibrium supervisor. The *obedient agent* always follows the recommendations; the *disobedient agent* may not follow the recommendations.

5.1 Theoretical Results

The analysis concerns the reachability of the terminal state and the complexity of the regulation process in terms of number of steps that are necessary to achieve the equilibrium.

If no agent follows the recommendations of the equilibrium supervisor, the latter is unable to regulate the system. Its effectiveness increases with the number of agents that follow its recommendations.

Thus, for the theoretical analysis, we shall consider only the case in which the *agents always follow the recommendations given by the equilibrium supervisor*. We show that, even in this favorable case, the decision process is a non-trivial one, due the qualitative nature of exchange values and to the restrictions imposed by the definition of exchange, that always requires a service (with a definite cost) to be done in any exchange stage. However, we show that under some conditions, it is always possible to have the system equilibrated in at most four steps.

Table 3: Partial schema of the reward function R for the case of two interacting agents

R	(C, C)	$(0_\epsilon, C)$	$(C, 0_\epsilon)$	(B_{-1}, F_{+1})	(B_{-3}, F_{+3})	(F_{+1}, B_{-1})	(C, B_{-1})	(B_{-3}, C)	(F_{+1}, C)
(E^-, E^+)	30	20	-30	-5	-10	3	20	20	20
(E^+, E^-)	30	-30	20	3	5	-5	20	18	-30
(E^+, E^+)	30	20	20	0	0	0	18	20	20
(E^0, E^+)	28	30	0	25	20	25	5	20	25
(E^+, E^0)	28	0	30	25	20	25	20	10	-5
(E^0, E^-)	-30	-30	0	30	13	-20	-5	0	-30
(E^-, E^0)	-30	0	-30	-20	20	30	0	-10	0
(E^-, E^-)	-30	-30	-30	30	0	30	28	26	-30

The symbol \star in the notation of the \star -matrices was omitted.

 Table 4: Partial schema of the optimal policy π^* for the case of two interacting agents

State	Action	State	Action	State	Action
$(E^i, E^j)_{-n \leq i < -1}^{1 < j \leq n}$	(C^i, C^j)	$(E^i, E^j)_{1 < i, j \leq n}$	(C^i, C^j)	$(E^i, E^0)_{-n \leq i < -1}$	$(F_{+(-i+1)}^i, B_{-1}^0)$
$(E^i, E^1)_{-n \leq i < -1}$	(C^i, B_{-1}^1)	$(E^1, E^j)_{1 < j \leq n}$	(B_{-1}^1, C^j)	(E^{-1}, E^0)	(F_{+1}^{-1}, B_{-1}^0)
$(E^{-1}, E^j)_{1 < j \leq n}$	(F_{+1}^{-1}, C^j)	$(E^0, E^j)_{1 < j \leq n}$	$(0_\epsilon, C^j)$	$(E^j, E^1)_{1 < j \leq n}$	(C^j, B_{-1}^1)
(E^{-1}, E^1)	(F_{+1}^{-1}, B_{-1}^1)	(E^0, E^1)	$(0_\epsilon, B_{-1}^1)$	$(E^i, E^{-1})_{-n \leq i < -1}$	$(F_{+(-i+1)}^i, B_{-1}^{-1})$
$(E^j, E^i)_{-n \leq i < -1}^{1 < j \leq n}$	(C^j, C^i)	$(E^j, E^0)_{1 < j \leq n}$	$(C^j, 0_\epsilon)$	$(E^{-1}, E^j)_{-n \leq j < -1}$	(F_{+2}^{-1}, B_{-1}^j)
$(E^1, E^i)_{-n \leq i < -1}$	(B_{-1}^1, C^i)	(E^1, E^0)	$(B_{-1}^1, 0_\epsilon)$	(E^{-1}, E^{-1})	$(F_{+2}^{-1}, B_{-1}^{-1})$
$(E^j, E^{-1})_{1 < j \leq n}$	(C^j, F_{+1}^{-1})	$(E^0, E^i)_{-n \leq i < -1}$	$(B_{-1}^0, F_{+(-i+1)}^0)$	$(E^i, E^j)_{-n \leq i, j < -1}$	$(B_{-1}^j, F_{+(-i+1)}^i)$ or
(E^1, E^{-1})	(B_{-1}^1, F_{+1}^{-1})	(E^0, E^{-1})	(B_{-1}^0, F_{+1}^{-1})		$(F_{+(-i+1)}^i, B_{-1}^j)$

The symbol \star in the notation of the \star -matrices was omitted.

Consider a multi-agent system composed by just two obedient agents α and β and use the notation given in (13). Let $M^\tau = (\mathbf{m}_{\alpha\beta}^\tau, \mathbf{m}_{\beta\alpha}^\tau)^\star$ be the material results, at step τ , of a social exchange process performed by the agents α and β . For a given tolerance $0 < \epsilon < \kappa_n = \frac{L}{n}$, where κ_n is the equilibrium supervisor accuracy, the following results hold:

Proposition 2 *If $\mathbf{m}_{\alpha\beta}^0 \in E^{-1}$ and $\mathbf{m}_{\beta\alpha}^0 \in E^1$, then the system achieves the equilibrium in one step if and only if $1 < \frac{\kappa_n}{\epsilon} \leq 3$.*

Proof. (\Rightarrow) Considering that the system is at the state $(E^{-1}, E^1)^\star$, then, for β 's material result, it holds that $\epsilon < \text{mid}(\mathbf{m}_{\beta\alpha}^0) \leq \frac{L}{n}$, and the optimal recommendation (Table 5, $R11$) is based on the optimal action $C^1 = [-\frac{1}{2}(\frac{L}{n} + \epsilon), -\frac{1}{2}(\frac{L}{n} + \epsilon)]$. It follows that:

$$\begin{aligned} & \epsilon - \frac{1}{2} \left(\frac{L}{n} + \epsilon \right) \\ & < \text{mid}(\mathbf{m}_{\beta\alpha}^0) - \frac{1}{2} \left(\frac{L}{n} + \epsilon \right) \leq \frac{L}{n} - \frac{1}{2} \left(\frac{L}{n} + \epsilon \right) \\ & \Rightarrow \frac{1}{2} \left(-\frac{L}{n} + \epsilon \right) < \text{mid}(\mathbf{m}_{\beta\alpha}^0) \leq \frac{1}{2} \left(\frac{L}{n} - \epsilon \right) \\ & \Rightarrow \frac{1}{2} (-h\epsilon + \epsilon) < \text{mid}(\mathbf{m}_{\beta\alpha}^0) \leq \frac{1}{2} (h\epsilon - \epsilon), \end{aligned}$$

where $\frac{L}{n} = h\epsilon$, with $h > 1$. If the system achieves the equilibrium in the step 1, then it holds that $\frac{1}{2}(h\epsilon - \epsilon) \leq \epsilon$. It follows that $1 < h \leq 3$, and therefore, $1 < \frac{\kappa_n}{\epsilon} \leq 3$, since $\kappa_n = \frac{L}{n}$. The proof for α 's material result is analogous. The proof of (\Leftarrow) is almost straightforward. \square

Proposition 3 *(i) If $\mathbf{m}_{\alpha\beta}^0 \in E^i$, with $1 < i \leq n$, then it is possible to get $\mathbf{m}_{\alpha\beta}^\tau \in E_\alpha^0$ in at most $\tau = 2$ steps if and only if $1 < \frac{\kappa_n}{\epsilon} \leq 3$; (ii) If $\mathbf{m}_{\beta\alpha}^0 \in E^i$, with $-n \leq i < -1$, then it is possible to get $\mathbf{m}_{\beta\alpha}^\tau \in E_\beta^0$ in at most $\tau = 2$ steps if and only if $1 < \frac{\kappa_n}{\epsilon} \leq 3$.*

Proof. (i) (\Rightarrow) Considering that $(i-1)\frac{L}{n} \leq \text{mid}(\mathbf{m}_{\alpha\beta}^0) < i\frac{L}{n}$ and that the optimal recommendation (Table 5, $R3$), thinking about each agent individually, is based on the optimal action $C = \left[\frac{(1-2i)L}{2n}, \frac{(1-2i)L}{2n} \right]$, it follows that:

$$\begin{aligned} & (i-1)\frac{L}{n} + \frac{(1-2i)L}{2n} < \text{mid}(\mathbf{m}_{\beta\alpha}^0) + \frac{(1-2i)L}{2n} \\ & \leq i\frac{L}{n} + \frac{(1-2i)L}{2n} \Rightarrow -\frac{1L}{2n} < \text{mid}(\mathbf{m}_{\beta\alpha}^0) \leq \frac{1L}{2n}. \end{aligned}$$

Table 5: Partial schema of the optimal value recommendation ρ_{π^*}

State	Optimal policy	Value Recommendation	Label
$(E^i, E^j)_{-n \leq i < -1, 1 < j \leq n}$	$(C^i > 0, C^j < 0)$	$((r_{\beta\alpha}, C^j), (s_{\alpha\beta}, C^i))$	R_1
$(E^j, E^i)_{-n \leq i < -1, 1 < j \leq n}$	$(C^j < 0, C^i > 0)$	$((r_{\alpha\beta}, C^j), (s_{\beta\alpha}, C^i))$	R_2
$(E^i, E^j)_{1 < i, j \leq n}$	$(C^i < 0, C^j < 0)$	$((r_{\alpha\beta}, C^i), (s_{\beta\alpha}, C^j))$ or $((r_{\beta\alpha}, C^j), (s_{\alpha\beta}, C^i))$	R_3 R_4
$(E^0, E^j)_{1 < j \leq n}$	$(0_\epsilon, C^j < 0)$	$((r_{\beta\alpha}, C^j), (s_{\alpha\beta}, 0_\epsilon))$	R_5
$(E^j, E^0)_{1 < j \leq n}$	$(C^j < 0, 0_\epsilon)$	$((r_{\alpha\beta}, C^j), (s_{\beta\alpha}, 0_\epsilon))$	R_6
$(E^0, E^i)_{-n \leq i < -1}$	$(B_{-1}^0 < 0, F_{+(-i+1)}^i > 0)$	$((r_{\alpha\beta}, B_{-1}^0), (s_{\beta\alpha}, F_{+(-i+1)}^i))$	R_7
$(E^i, E^0)_{-n \leq i < -1}$	$(F_{+(-i+1)}^i > 0, B_{-1}^0 < 0)$	$((r_{\beta\alpha}, B_{-1}^0), (s_{\alpha\beta}, F_{+(-i+1)}^i))$	R_8
$(E^{-1}, E^j)_{1 < j \leq n}$	$(F_{+1}^{-1} > 0, C^j < 0)$	$((r_{\beta\alpha}, C^j), (s_{\alpha\beta}, F_{+1}^{-1}))$	R_9
$(E^1, E^i)_{-n \leq i < -1}$	$(B_{-1}^1 < 0, C^i > 0)$	$((r_{\alpha\beta}, B_{-1}^1), (s_{\beta\alpha}, C^i))$	R_{10}
(E^{-1}, E^1)	$(F_{+1}^{-1} > 0, B_{-1}^1 < 0)$	$((r_{\beta\alpha}, B_{-1}^1), (s_{\alpha\beta}, F_{+1}^{-1}))$	R_{11}
(E^1, E^{-1})	$(B_{-1}^1 < 0, F_{+1}^{-1} > 0)$	$((r_{\alpha\beta}, B_{-1}^1), (s_{\beta\alpha}, F_{+1}^{-1}))$	R_{12}
$(E^i, E^1)_{-n \leq i < -1}$	$(C^i > 0, B_{-1}^1 < 0)$	$((r_{\beta\alpha}, B_{-1}^1), (s_{\alpha\beta}, C^i))$	R_{13}
(E^{-1}, E^0)	$(F_{+1}^{-1} > 0, B_{-1}^0 < 0)$	$((r_{\beta\alpha}, B_{-1}^0), (s_{\alpha\beta}, F_{+1}^{-1}))$	R_{14}
(E^0, E^{-1})	$(B_{-1}^0 < 0, F_{+1}^{-1} > 0)$	$((r_{\alpha\beta}, B_{-1}^0), (s_{\beta\alpha}, F_{+1}^{-1}))$	R_{15}
$(E^i, E^j)_{-n \leq i, j < -1}$	$(F_{+(-i+1)}^i > 0, B_{-1}^j < 0)$ or $(B_{-1}^j < 0, F_{+(-i+1)}^i > 0)$	$((r_{\beta\alpha}, B_{-1}^j), (s_{\alpha\beta}, F_{+(-i+1)}^i))$ or $((r_{\alpha\beta}, B_{-1}^j), (s_{\beta\alpha}, F_{+(-i+1)}^i))$	R_{16} R_{17}

The symbol \star in the notation of the \star -matrices was omitted.

Therefore, it holds that $\mathbf{m}_{\beta\alpha}^1 \in E_\alpha^1$. From Prop. 2, it follows that with one more step we can get the desired result. The proofs of (i)(\Leftarrow) and (ii) are analogous. \square

The following result is almost immediate:

Proposition 4 *If $\mathbf{m}_{\alpha\beta}^0 \in E^i$, $1 < i \leq n$, is such that $\frac{2i+1}{2} \frac{L}{n} - \epsilon \leq \text{mid}(\mathbf{m}_{\alpha\beta}^0) \leq \frac{2i+1}{2} \frac{L}{n} + \epsilon$, then $\mathbf{m}_{\alpha\beta}^1 \in E_\alpha^0$.*

From Prop. 3, it follows that an individual transition from a material result that belongs to a class E^i , with $1 < i \leq n$ or $-n \leq i < -1$, to the equilibrium can be done in at most two steps ($E^i \mapsto E^1$ (or $E^{-1}) \mapsto E^0$). However, in any interaction between two agents, combined transitions departing from a state $(E^i, E^j)^\star$ or $(E^j, E^i)^\star$, with $1 < i \leq n$ and $-n \leq j < -1$, may result in a state different from $(E^1, E^{-1})^\star$, $(E^{-1}, E^1)^\star$ or $(E^0, E^0)^\star$. We may have, for example, $(E^{-1}, E^0)^\star$, and, in this case, it will not be possible to get the equilibrium in one more step, since any *compensation* or *go-forward* action for α is not allowed without a corresponding β 's service. The solution is then to have a transition to $(E^1, E^{-1})^\star$ and then, finally, to reach $(E^0, E^0)^\star$. Thus, the overall process takes three steps.

The worst case is when the interaction presents material results that belong to the state $(E^i, E^j)^\star$, with $-n \leq i, j < -1$, since two simultaneous positive compensation actions (that would require a recommendation of satisfaction values for the two agents without

any service at all) are not allowed. In this case, the optimal recommendation (Table 5) leads the agents to get the material equilibrium in at most four steps, by one of the following transitions:

$$\begin{aligned}
(E^i, E^j)_{-n \leq i, j < -1}^\star &\xrightarrow{R_{16}} (E^1, E^j)_{-n \leq j < -1}^\star \\
&\xrightarrow{R_{10}} (E^0, E^{-1})^\star \xrightarrow{R_{15}} (E^{-1}, E^1)^\star \xrightarrow{R_{11}} (E^0, E^0)^\star, \\
(E^i, E^j)_{-n \leq i, j < -1}^\star &\xrightarrow{R_{17}} (E^j, E^1)_{-n \leq j < -1}^\star \\
&\xrightarrow{R_{13}} (E^{-1}, E^0)^\star \xrightarrow{R_{14}} (E^1, E^{-1})^\star \xrightarrow{R_{12}} (E^0, E^0)^\star.
\end{aligned}$$

5.2 Simulation Results

Simulations of supervised social exchange processes were developed in the Python programming language, generating two types of reports: tables with the configurations of exchange values and material results at each time $t \in T = 0, \dots, 1000$, and graphics showing the trajectory of the mid points of the material results of the exchanges between two agents of a (simulated) system. The material and virtual values that agents could use at each exchange stage were set to vary in the range $-100 \dots +100$. A tolerance of $[-\epsilon, \epsilon] = [-25, 25]$ was adopted for the definition of the equilibrium point.

First, we considered simulations of social exchanges processes in organizations where the supervisor was inactive, Fig. 3, having exchange values bound to

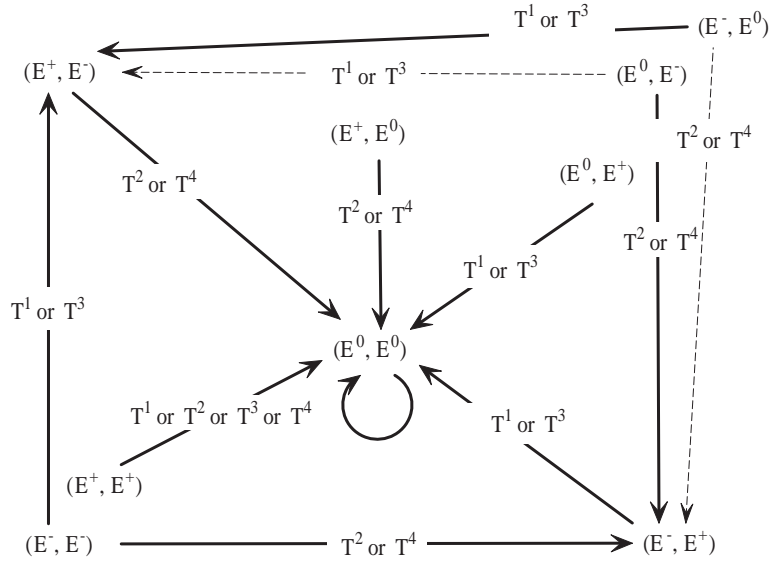


Figure 2: Effects of stage and optimal value recommendations

$L = 2400$). In those simulations, the social exchanges were totally random along all the time interval T . Then we awakened up the supervisor and let it make recommendations, which the agents followed or not, according to their degree of obedience to the supervisor.

In successive experiments, we increased the percentage of agent obedience to the supervisor, generating five different simulations: obedience during 1% of the time (Fig. 4, exchange values bound to $L = 1400$), obedience during 25% of the time (Fig. 5, exchange values bound to $L = 700$), obedience during 50% of the time (Fig. 6, exchange values bound to $L = 350$), obedience during 75% of the time (Fig. 7, exchange values to bound $L = 185$) and obedience during 100% of the time (Fig. 8, exchange values bound to $L = 105$).

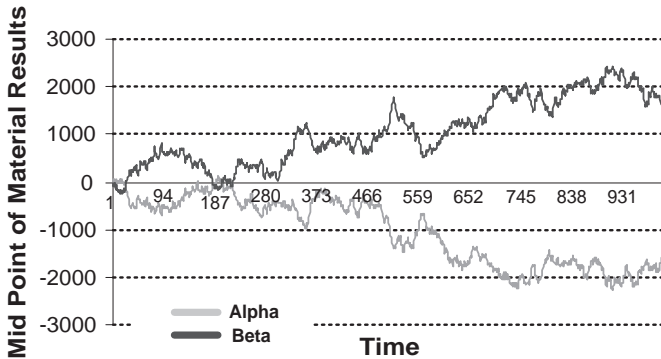


Figure 3: Unsupervised social exchange process

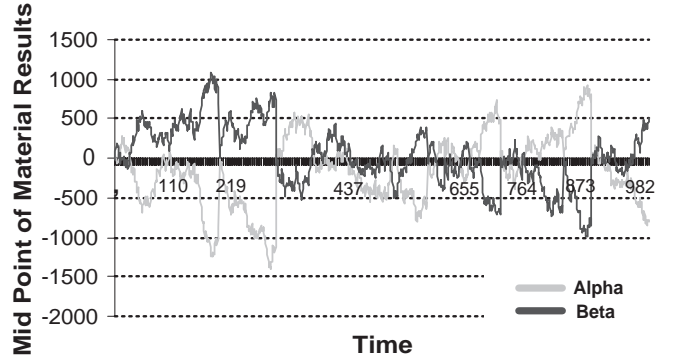


Figure 4: Agent obedience in 1% of the time

In the unsupervised exchange processes shown in Fig. 3, agents α and β present behaviors such that β profited from the interaction much more than α , which was forced to keep its material results at a negative level for the most part of the experiment. Considering the given tolerance, one finds that the system starts in equilibrium, but was unable to keep it.

Figure 4 shows a 1% supervised exchange processes. The figure shows that just such level of supervision is enough to make the two agents alternate their kinds of behaviors during the exchanges, thus avoiding that one of them profits from the interaction at the expense of the other. Also, the system was able to achieve the equilibrium in various opportunities (e.g., at $t = 196$ and at $t = 893$), but was kept yet in disequilibrium almost all of the time.

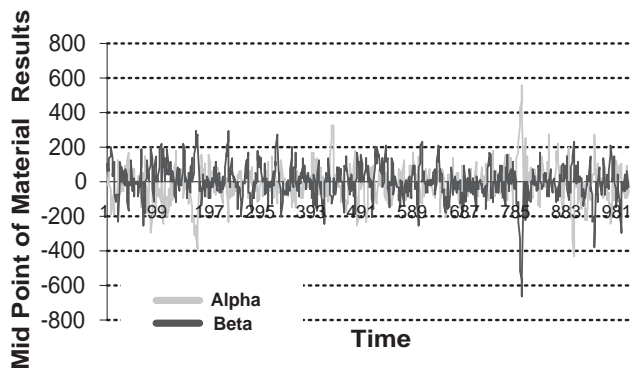


Figure 5: Agent obedience in 25% of the time

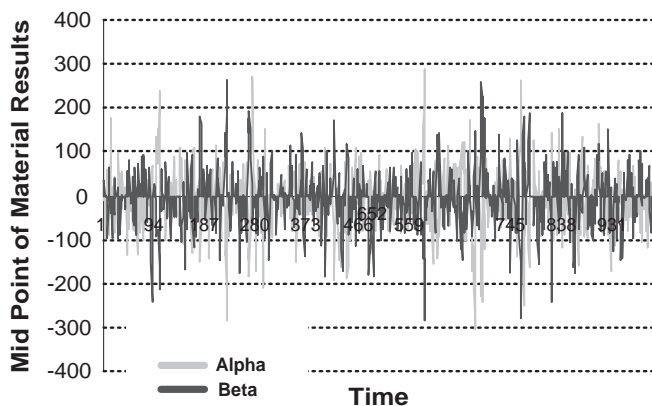


Figure 6: Agent obedience in 50% of the time

Figures 5, 6, 7 and 8 shows the increasing effect of the supervisor recommendations, according the increasing level of the agents' obedience to it. In particular, one can see that the range of deviations of the material results from the equilibrium was progressively reduced as the agents progressively adhered to the supervisor's recommendations.

The simple simulations that we produced seem to confirm well the theoretical predictions that we could derive from the supervisor model in section 5.1.

6 Conclusion

This paper introduced the QI-MDP version of the Markov Decision Process. The combination of interval-based modelling and qualitative approach to the comparison of values of the model made it well suited for solving the problem of keeping social exchanges in equilibrium.

From the point of view of Piaget's theory of social

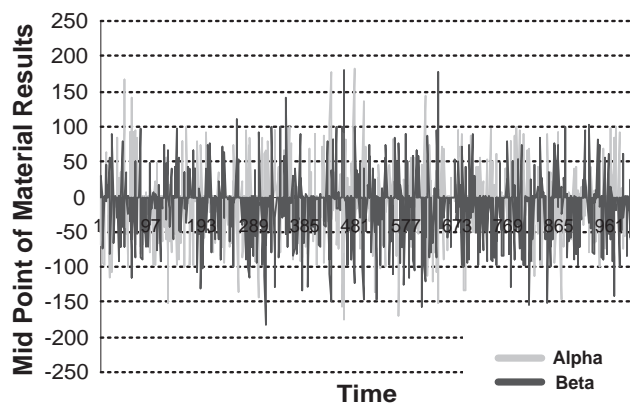


Figure 7: Agent obedience in 75% of the time

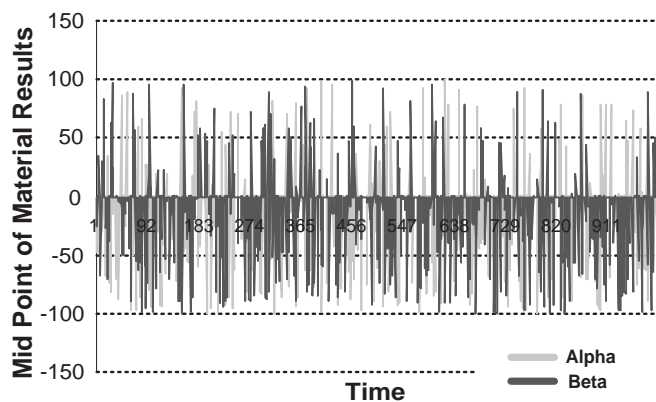


Figure 8: Agent obedience in 100% of the time

interactions, it means a sound way of making practical use of the INRC group of social exchanges that structure the social interactions and defines its equilibrium problem [9, 10].

The QI-MDP model is general enough to be applied to other problems, besides the problem of keeping social interactions in equilibrium. It can be adapted to model situations in which the social interactions should be kept stable, but in disequilibrium. This can be done by choosing a non null terminal state for the supervisor.

The model can also be applied to equilibrium problems of other kinds of systems, besides systems of social exchanges, if such systems have one single terminal (equilibrated, disequilibrated) state.

Regarding the notion of *social control* introduced in [23], and explored in [12] in connection to multi-agent systems, the social rules concerned with exchange values introduce the notion of *exchange values-*

based social control, first analyzed in [19]. Thus, the present paper can be seen as a preliminary step in the computational formalization of such notion.

Immediate future work will be concerned with the case of an equilibrium supervisor that is not able to determine the material balance of social exchange processes with complete reliability (i.e., it is not allowed to know all the exchange values of the two agents). In this case, a partially observable Markov decision process (POMDP) shall be considered (see, p.ex., [15]), since the equilibrium supervisor shall be able to make external observations (also probabilistic) to help him to decide about the recommendations.

Further future works will deal with (i) the internalization of the model decision process introduced in the paper, in each agent of the organization, so that the mechanism of exchange values-based social control that it supports can be performed in a decentralized way, and (ii) further exploration of the role of exchange values in dependence-based agent interactions.

Acknowledgements

This work was partially supported by CNPq and FAPERGS.

The authors thank the anonymous referees for the valuable suggestions, specially the requirement to distinguish between equilibrium and order (stability).

References

- [1] G. Alefeld and J. Herzberger. Introduction to Interval Computations. Academic Press, New York, 1983.
- [2] L. Antunes and H. Coelho. Decisions Based Upon Multiple Values: the BVG Agent Architecture. In *Proceedings of 9th Portuguese Conference on Artificial Intelligence, EPIA'99*, Évora, 1999. P. Barahona and J.J. Alferes (Eds.), Springer Lecture Notes in Artificial Intelligence, Vol. 1695, pages 297–311, 1999.
- [3] R.H. Bordini, A.C.R. Costa, J.F. Hübner, and R.M. Viccari. Linguistic Support for Agent Migration. In V. Lesser and L. Gasser (Eds.), *Proceedings of the 1st International Conference on Multi-Agent Systems*, San Francisco, page 441, 1995.
- [4] A.C. Rocha Costa, J.M.V. Castilho, and D. Claudio. Functional Processes and Functional Roles in Societies of Computing Agents. In *Proceedings of Brazilian Symposium on Artificial Intelligence, SBIA 93*, Porto Alegre, pages 267–276, 1993.
- [5] A.C.R. Costa, J.F. Hübner, and R.H. Bordini. On Entering an Open Society. In *Proceedings of Brazilian Symposium on Artificial Intelligence, SBIA 94*, Fortaleza, pages 535–546, 1994.
- [6] A.C.R. Costa, J.M.V. Castilho, and D.M. Claudio. Toward a Constructive Notion Of Functionality. *Cybernetics and Systems*, 6(4):443–480, 1995.
- [7] A.C.R. Costa and Y. Demazeau. Toward a Formal Model of Multi-agent Systems with Dynamic Organization. In J.W. Perram and J.-P. Müller (Eds.), *Proceedings of 2nd. International Conference on Multi-agent Systems, ICMAS'96*, Kyoto, page 431, 1996.
- [8] A.C.R. Costa. The Piagetian Theory of Social Exchanges and its Application to Learning Environment. *Informática na Educação: Teoria e Prática*, 6(2):77–900, 2003. (in Portuguese)
- [9] A.C.R. Costa and G.P. Dimuro. About INRC Groups of Social Exchange Stages: a Formal Critical Analysis of Piaget's Social Exchange Model. Technical Report, ESIN/UCPel, Pelotas, 2003. (available at <http://gmc.ucpel.tche.br/valores>, Mar. 2005, in Portuguese)
- [10] A.C.R. Costa and G.P. Dimuro. The Case for Using Exchange Values in the Modelling of Collaborative Learning Interactions. In J. Mostow and P. Tedesco (Eds.), *Proceedings of the 7th International Conference on Intelligent Tutoring Systems/2nd International Workshop on Designing Computational Models of Collaborative Learning Interaction*, Maceió, pages 19–24, 2004.
- [11] C. Castelfranchi, M. Miceli, and A. Cesta. Dependence Relations among Autonomous Agents. In E. Werner and Y. Demazeau (Eds.), *Decentralized A.I.-3*, pages 215–227. Elsevier, Amsterdam, 1992.
- [12] C. Castelfranchi. Engineering Social Order. In A. Omicini, R. Tolksdorf, and F. Zambonelli (Eds.), *Engineering Societies in the Agents World*, pages 1–18. Springer, Berlin, 2000.
- [13] Y. Demazeau and A.C.R. Costa. Populations and Organizations in Open Multi-Agent Systems. In *Proceedings of 1st National Symposium on Parallel and Distributed AI*, Hyderabad, 1996.

- [14] G.P. Dimuro and A.C.R. Costa. Interval-based Markov Decision Processes for Regulating Interactions Between Two Agents in Multi-Agent Systems. In: *Extended Abstracts of the Workshop on State-of-the-Art in Scientific Computing/Mini-Symposium on Interval Methods*, pages 51–58, Lyngby, 2004. (the full paper will appear in Springer Lecture Notes in Computer Science)
- [15] L.P. Kaelbling, M.L. Littman, and A.R. Cassandra. Planning and Acting in Partially Observable Stochastic Domains. *Artificial Intelligence*, 101(1):99–134, 1998.
- [16] R.B. Keafort and V. Kreinovich (Eds). Applications Of Interval Computations. Kluwer, Boston, 1996.
- [17] U. Kulisch and W.L. Miranker (Eds). A New Approach to Scientific Computation. Academic Press, New York, 1983.
- [18] U. Kulisch. Advanced Arithmetic for the Digital Computer, Design of Arithmetic Units. *Electronic Notes in Theoretical Computer Science*, 24, 1999.
- [19] G.C. Homans. Social Behavior as Exchange. *American Journal of Sociology*, 63(1958):597-606. Reprinted in [21].
- [20] G.C. Homans. Social Behavior – Its Elementary Forms. Harcourt, Brace & World, New York, 1961.
- [21] G.C. Homans. Sentiments and Activities. The Free Press of Glencoe, New York, 1962.
- [22] G.C. Homans and C.P. Curtis, Jr. An Introduction to Pareto – His Sociology. Alfred A. Knopf, New York, 1934.
- [23] G.C. Homans. The Human Group. Harcourt, Brace & World, New York, 1950.
- [24] W.S. Lovejoy. A Survey of Algorithmic Methods for Partially Observable Markov Decision Processes. *Annals of Operations Research*, 28(1):47–65, 1991.
- [25] M. Miceli and C. Castelfranchi. The role of evaluation in cognition and social interaction. In K. Dautenhahn (Ed.), *Human cognition and agent technology*, pages 225–262. John Benjamins, Amsterdam, 2000.
- [26] R.E. Moore. Interval Analysis. Prentice-Hall, Englewood Cliffs, 1966.
- [27] R.E. Moore. Methods and Applications of Interval Analysis. SIAM, Philadelphia, 1979.
- [28] J. Piaget. Biology and knowledge : an essay on the relations between organic regulations and cognitive processes. The University of Chicago Press, Chicago, 1971.
- [29] J. Piaget. The development of thought : equilibration of cognitive structures. The Viking Press, New York, 1977.
- [30] J. Piaget. L’Explication en Sociologie. In J. Piaget (Ed.), *Introduction à la Épistémologie Génétique – Tome III: La Pensée Biologique, La Pensée Psychologique et la Pensée Sociologique*. Presses Universitaire de France, Paris, 1950. Reprinted in [33].
- [31] J. Piaget. Essay sur la Théorie des Valeurs Qualitatifs en Sociologie Statique (“synchronique”). In [32].
- [32] J. Piaget. Études Sociologiques. Librairie Droz, Paris, 1965. English translation in [33].
- [33] J. Piaget. Sociological Studies. Routledge, London, 1995.
- [34] M.R. Rodrigues, A.C.R. Costa, and R. Bordini. A System of Exchange Values to Support Social Interactions in Artificial Societies. In *Proceeding of the Second International Conference on Autonomous Agents and Multiagents Systems, AAMAS 2003*, Melbourne, pages 81–88, 2003.
- [35] M.R. Rodrigues and A.C.R. Costa. Using Qualitative Exchange Values to Improve the Modelling of Social Interactions. In D. Hales, B. Edmonds, E. Norling, and J. Rouchier (Eds.), *Proceedings of 4th Workshop on Agent Based Simulations*, Melbourne, 2003. Springer Lecture Notes in Computer Science Series, Vol. 2927, pages 57–72, 2003.
- [36] W.E. Walsh and M.P. Wellman. A Market Protocol for Distributed Task Allocation. In *Proceedings of the Third International Conference on Multiagent Systems*, Paris, pages 325–332, 1998.
- [37] D.J. White. Markov Decision Processes. Wiley, New York, 2002.