

The Organisation of Sociality: A Manifesto for a New Science of Multi- Agent Systems

Pietro Panzarasa & Nicholas R. Jennings
Department of Electronics and Computer Science
University of Southampton
Southampton SO17 1BJ, U.K.
{pp; nrj}@ecs.soton.ac.uk

Abstract. In this paper, we pose and motivate a challenge, namely the need for a new science of multi-agent systems. We propose that this new science should be grounded, theoretically on a richer conception of sociality, and methodologically on the extensive use of computational modelling for real-world applications and social simulations. Here, the steps we set forth towards meeting that challenge are mainly theoretical. In this respect, we provide a new model of multi-agent systems that reflects a fully explicated conception of cognition, both at the individual and the collective level. Finally, the mechanisms and principles underpinning the model will be examined with particular emphasis on the contributions provided by contemporary organisation theory.

1. Introduction

This paper is about sociality in multi-agent systems (MASs) - how it is generated and how it evolves over time. We propose to describe, evaluate, and integrate some of the major perspectives on organisation design and behaviour, and to use these perspectives as a basis for a new science of MASs. Our articles of faith are simple. We believe that to engineer MASs that can cope with a wide range of phenomena and the problems traditionally associated with more complex, uncertain, and constantly changing environments, we need to supplement orthodox multi-agent system (MAS) theory with a novel and broader conception of sociality. This need is also motivated by the fact that agents, written by different stakeholders, with different aims and interests, are required to interact in flexible ways, to face the problem of bounded rationality and inter-agent opportunism, to make rational decisions, to adaptively learn from their experience, to deal with inertial non-linear dynamics, and to exhibit forms of pro-active behaviour. Under these conditions, many of the current theoretical tools and principles of MAS research seem to be inadequate. Our response to these deficiencies is placed against the backdrop provided by the broader tradition of organisation theory. By discussing and taking cognisance of what has been learned in organisational research and thought, we explore new directions that can further MAS theory and practice. In particular, we speculate on a range of concepts, principles and analytical tools that can pave the way for a new science of MASs.

Our undertaking is premised on a scrutiny, criticism and re-conceptualisation of the major pillar of MAS research, namely the very notion of MAS. The

traditional concern of much of the MAS community has been with the design of relatively few computational artificial agents that need to be coordinated in order to accomplish a common goal [49]. Obtaining a systematic understanding of other forms of MASs, such as knowledge networks [35], or virtual organisations composed of an increasingly large number of human and artificial agents [7], or large business firms in which economic activity is undertaken through differentiation and coordination of tasks [29], has rarely been a high-priority objective on most researchers' agendas. In an attempt to offer a plausible cure for these inadequacies, we propose a three-step approach. First, we distinguish between the notion of organisation and MAS, and subsume the former under the latter. Any organisation is a MAS in which some form of joint behaviour is carried out through differentiation and coordination of tasks among the constituent members [29, 30, 43]. However, not every MAS is an organisation, for there may be systems in which only cognitive joint processes are undertaken (e.g. knowledge networks). Second, we distinguish between the organisation of a MAS, and a MAS as an organisation. The former refers to the internal structure and coordination of the constituent parts of a MAS, whereas the latter refers to a MAS that relies upon its internal structure to undertake joint behaviour. Finally, building on the above distinctions, we set forth an alternative view of the MAS by extending its traditional semantics so as to capture the richness and complexity of a wide range of possible pluralistic forms for organising sociality. Among other things, our analysis will suggest a notion of MAS that conveys the idea of any form of sociality, formally or informally organised, planned or emergent, that involves either artificial or human autonomous, loosely interdependent, cognitive agents (see section 2).

Building on this conceptual background, we will develop a model of MAS in which mechanisms of joint cognition and behaviour are predicated in terms of cognition and behaviour at the individual level. At the heart of this model is the argument that agenthood and cognition occur not only at the individual level, but also as an emergent phenomenon from the interaction among multiple agents [6]. Sociality is then described in terms of higher-order joint mental attitudes and joint behaviour that build on and transcend the mental attitudes and behaviours of cognitively and socially interconnected individual agents [36, 52]. Finally, the mechanisms and principles underpinning the model will be examined drawing on the contributions given by the leading scholars in organisation theory in the last few decades. In doing so, we will pursue two objectives. First, an attempt will be made to offer a critical assessment of different schools of thought, and integrate them into a synthetic perspective on organisation design and behaviour. Second, we will try to further strengthen our model of MAS by appealing to the analytical tools provided by organisation theories and by using them in an integrated way in order to gain an insight into the fundamental mechanisms through which sociality comes into existence and changes over

time. In pursuing both objectives, we will use the canonical issue of the relation between organisation and environment as a guiding principle. By showing how this issue has been variously addressed by different scholars, we will show not only how the contributions can be brought together into an articulated unified perspective, but also how the divergence of views can be used as a step towards a new conception of the environment and the way in which MASs relate to it. This will pave the way for a set of implications regarding the conditions under which MASs will behave passively, re-actively, or proactively, and the type of cognitive processes on which each of these forms of behaviour rests.

The remainder of the paper is organised as follows. The next section will introduce our model of MAS. In section 3, the model will be discussed by looking for common threads among organisational theories. The discussion will focus on the determinants of sociality (section 3.1), and on the main mechanisms through which sociality changes over time (section 3.2). Section 4 summarises our major findings and provides indications for future research. It is perhaps in setting the future research agenda that the main contribution of this paper lies. As it stands, the paper represents a manifesto for the future landscape of MAS research. Only time will tell whether this vision becomes a reality.

2. A Model of MAS

Science is primarily aimed at finding order in the complexity of nature. To do so, it must partition the object of analysis into various discrete segments, explain their interdependencies, and establish boundaries. Ludwig von Bertalanffy, noting this, drew its logical conclusion [45]. He observed that, if science is primarily concerned with problems of organisation, then the most general science of all is expected to be a science of organisation itself. This is the core idea that underpins what is known as "general system theory" [5, 21]. This theory has inspired our work in two main ways. First, building on the attempts that have been variously made by scholars to develop and apply the systems view to the behavioural and social sciences [21, 23], we will use the open-system perspective for representing and reasoning about MASs. Second, drawing on recent advances in the cybernetics of social systems, we will conduct our analysis by endorsing the point of view of the MAS, here modelled as a cognitive entity that, through feedback and morphogenetic mechanisms, is capable not only of performing maintenance and adaptive behaviour, but also of autonomously re-organising and re-producing its own structure from within [31].

Figure 1 illustrates our model of MAS, its structure and the dynamic interplay of its main components. In this section, we give a brief overview of the model and its articulation, whereas a critical discussion will be carried out in the following section. Against the theoretical background outlined above, the model is premised on the following definition of MAS:

Definition 1. A MAS is a social and cognitive entity, with a relatively identifiable boundary, that functions on a relatively continuous basis through the coordination of loosely interdependent, cognitive and autonomous agents.

Instead of discussing the above definition by looking at each of its components in the order they appear, we will follow an alternative strategy. We will reconstruct the genesis of the definition by adhering to the path through which a MAS comes into existence, shapes its structure, engages in cognitive or behavioural activities, and produces outcomes. The remainder of this section will therefore be articulated as follows. First, the determinants of sociality will be introduced (section 2.1); the individual agent will then be modelled (section 2.2); the other two main components of the system's structure will then be outlined, namely the contextual factors (section 2.3) and the organisational variables (section 2.4); the concept of agenthood will then be extended from the individual to the collective level (section 2.5); finally, the outcomes of the system's processes, and their effects, will be examined (section 2.6).

2.1 Determinants

The model rests on two assumptions. First, the MAS is seen as located in an environment from which it may be constrained and upon which it may impact¹ [52]. Second, the generation and functioning of the MAS is modelled as motivated by a dependence network, namely a set of social dependence relations that hold among agents [17, 49]. Environment and dependence network are inherently intertwined with each other for the following reasons. First, changes in the environment (e.g. availability of resources) bring about changes in the ways the agents depend on one another. Second, the dependence network impacts back onto the environment by affecting the ways in which agents are clustered and organised, which in turn affects the productivity of the resources used and the ways in which these are allocated between the agents. Again, this impacts on the ways in which agents depend on one another, and so on *ad infinitum*.

Environment and dependence network influence the generation and functioning of the MAS by affecting its structure and, via this, its processes. The structure of a MAS refers to the enduring pattern of systematic relationships among the components of the system. The system's processes refer to the dynamic integration of the structural components into a consistent set of cognitive and behavioural joint activities. At the heart of the complex mechanisms that relate the system's structure to its processes lies our conception of *socio-cognitive agenthood* which, in turn, builds on the role

¹ As explained in section 3.2, the strength of the dependence between MAS and environment may have differing degrees, from the case in which the MAS is deterministically "selected" by its environment, through to the case in which the environment is relatively impotent in its effect on the MAS. In particular, the latter case will give us the opportunity to discuss and re-assess the very notion of environment (see section 3.2.2.3).

that cognition plays both at the individual and the joint level.

2.2 The Individual Agent

Agenthood is seen as existing at two interrelated levels: the individual and the collective one. At each level, agenthood is expressed in terms of cognition and behaviour. Let us start our analysis with the individual level. The first component of the centre box in Figure 1 shows the set of individual agents who are the members of the MAS and, therefore, represent its basic structural components. Here, the individual agent is regarded as an autonomous, socio-cognitive entity capable of individual social behaviour [47]. First, being autonomous means being able to operate independently [52]. Second, for an agent to be termed *cognitive* it must be endowed with mental attitudes representing the world and motivating action [36, 39, 41, 52]. Further, for a cognitive agent to be deemed *socio-cognitive* it must not only have an intentional stance towards the environment, but also represent other agents as cognitive agents similarly endowed with mental attitudes for representational and motivational purposes [14]. These properties of individual agenthood can be conceptualised and operationalised by modelling the agent as having a *cognitive structure* and being capable of performing *cognitive processes*. According to mainstream cognitive science, a cognitive structure is a set comprising basic mental attitudes, cognitive principles for modifying these attitudes, operators for adding mental attitudes, the cognitive frame, the language, control procedures, social interaction propensities, and principles and mechanisms for generating inter-agent social behaviour [6]. A key role within the agent's cognitive structure is played by the *mental state*. This is a set of the basic mental attitudes, including beliefs, goals, intentions, preferences, etc. [11, 41]. Building on its cognitive structure, the agent can engage in a variety of cognitive processes, such as conceptualisation, sense-making, categorisation, theoretical and practical reasoning, and so forth [2]. These processes are primarily intended to process, modify or discard the agent's mental attitudes and/or other components of its cognitive structure, or even generate new ones. In this respect, cognitive structure and cognitive processes are related to each other, in that the former provides the cognitive components for undertaking the latter, while the latter affect the former by modifying these components or adding new ones.

Besides being endowed with a cognitive structure and capabilities of undertaking cognitive processes, the agent is also modelled as being capable of *individual social behaviour* aimed at producing effects on other agents. Specifically, an agent's behaviour is social when it is regulated by the agent's cognitive structure, and is oriented towards another agent who, in turn, is regarded as a cognitive agent whose behaviour is regulated by a cognitive structure and a set of cognitive processes [47]. According to this conception, the agent's individual social behaviour rests on its cognitive structure, in that it is controlled by the

principles of social behaviour which, in turn, are affected by the agent's mental attitudes [6]. Therefore, individual social behaviour is indirectly influenced by the agent's cognitive processes because these impact upon the mental attitudes and principles that ultimately trigger and control behaviour. Finally, individual social behaviour impacts back upon the agent's cognitive structure, and via this, on its cognitive processes. In fact, by interacting with others, the agent acquires new information, fulfils its goals, modifies its preferences, etc., and this, in turn, will provide new inputs for the agent's cognitive processes (see also section 2.5).

2.3 Contextual Factors

Both the agent's cognition and social behaviour are constrained by and impact upon two interrelated super-individual components of the system's structure: a set of contextual factors and the organisational variables (shown in the middle of the centre box of Figure 1). The former include five main subsets of situational factors. First, the techniques (i.e., the equipment and processes required to transform inputs into outputs) and technologies (i.e., the set of codified knowledge concerning the techniques adopted) that are used to undertake social activity or simply to manage the social interactions within the system [43]. Second, the institutional variables, namely, the high-level functions pursued by the MAS within a higher-level system (e.g. the economy as a whole), the ways in which the outcomes of its activities are allocated among its members, and the nature of its social externalities (e.g. opportunities for value creation) [38]. Third, the social profile of the system, namely, the nature of the social relations among its members (i.e., their content, finalities, complexity, intensity, specificity, strength, frequency, formalisation and degree of trust, reciprocity and clarity of expectations), the number and types of sub-groups, the set of roles and their relations, the nature and articulation of the communication processes, the social influence patterns that the agents use to govern their interdependencies and make decisions, and the ways in which conflict between agents and/or sub-groups is managed [12, 19]. Fourth, the structural properties of the inter-agent social network, namely its dimension, connectivity, degree centrality, betweenness centrality, structural equivalence, clustering, and the number and forms of cliques [28, 46]. Finally, the set of physical objects and resources that are used within the system.

All the above contextual factors need to be consistent with one another, in that, each of them is constrained by and constrains all the others. Also, they need to be consistent with the set of the agents who are the members of the system. For example, the technology used, by affecting the way in which activities are carried out, is related to what physical resources are used and how they are logistically organised, which in turn may affect the number and identity of the agents that are members of the system (as they control these resources), which affects what social relations and subgroups are formed in the

system, which determines the structural properties of the social network representing the links between each agent, which in turn can impact on how effectively and efficiently the system can undertake its institutional functions, and whether or not social externalities are generated in pursuing social activity, which impacts back onto the choice of the appropriate technologies and techniques, and so on *ad infinitum*. Finally, this causal chain can be re-formulated in the opposite direction; it may also start from the set of the agents or any other contextual factor, and it may relate all its elements in any other different order.

2.4 Organisational Variables

Organisational variables refer to the ways in which social activity (if any) is divided into distinct tasks and then coordination is achieved among these tasks within the system. As it turns out, organisational variables can be analysed into two fundamental components: the organisational structure and the organisational mechanisms. Even though there is by no means complete agreement as to what makes up the term "organisational structure", almost all definitions seem to convey the idea that this notion refers to the "division of labour" into a relatively stable pattern of tasks that are allocated among the members of a system in order to carry out a joint activity [33]. The organisational mechanisms, on the other hand, refer to the ways in which distinct tasks, once distributed among the members of the system, can be coordinated to accomplish social activity [20]. Examples of these mechanisms are: procedures for controlling and evaluating behaviour and results; standardised procedures for collaborative decisions; organisational power; remuneration and the reward system. Whenever any form of social activity is to be carried out, these organisational mechanisms lie at the root of the equilibrium and stability of the system. In fact, they allow, together with other integrative mechanisms within the set of social contextual factors (e.g. communication, social influence), first the components of the organisational structure to combine with one another, second the organisational structure as a whole to dynamically integrate with all the other structural components of the system. Conversely, when no social activity is performed, equilibrium is brought about only by using social integrative mechanisms.

2.5 Joint Cognition and Behaviour

Organisational and social mechanisms can be seen as the key elements for understanding the path leading from the individual level of agenthood to the collective one. As with the individual agent, we have two fundamental components of the collective level of agenthood: joint cognition and joint behaviour. Again, cognition can be analysed into its static and dynamic components. On the one hand, the mental states of the members of the system combine in such a way that a *joint mental state* will ensue [52]. This includes varying joint doxastic, motivational and deontic mental

attitudes, such as mutual beliefs, joint intentions, joint commitments [36]. On the other hand, joint cognitive processes rest on and transcend the complex interplay between the agents' individual cognitive processes. Similarly, joint behavioural processes build on and transcend the agents' individual social behaviours. Against this background, two fundamental questions regarding the internal structure and functioning of collective agenthood are: (i) to what extent does the joint mental state relate to the set of joint processes, and how can the former regulate the evolution of the latter? (ii) are the cognitive and behavioural joint processes related to each other, and if so, what are the mechanisms that govern this relation?

Answering these questions leads us to elucidate the core principles that underpin our conception of socio-cognitive agenthood. In turn, these principles strictly build on and can be explained in terms on the key functions played by the individual agents' cognition and behaviour. As to the first question, we note that joint mental state and joint processes as a whole co-evolve, in that changes in one of them impact upon the other in a two-way direction, through the mediation of the individual's cognition and behaviour. In order to gain an insight into this co-evolution, it is, therefore, necessary to understand the mediating role played by the individual's cognition and behaviour.

First, joint processes find their cognitive roots in the system's joint mental state, in a causal chain involving the agents' cognitive structures, cognitive processes, and individual social behaviours [36]. As to the joint cognitive processes, they are rooted in the joint mental state through the impact that this has on the individual agents' mental states, and these on the agents' cognitive processes. For example, it is a mutual belief, as well as various other forms of shared information, that leads the agents to undertake individual cognitive processes of reasoning that, ultimately, combine into joint cognitive processes of sense-making or knowledge creation. Likewise, joint behavioural processes build on the system's joint mental state through the impact that this has on the agents' mental states and cognitive processes and these, in turn, on the agent's individual social behaviour. For example, it is a joint commitment to achieving some state of affairs that gets the agents' mental states to trigger and control the performance of the cognitive processes of practical reasoning concerning how to attain that state. In turn, the outcome of these processes will provide the agents with the necessary cognitive ingredients for undertaking individual social behaviours that, ultimately, will combine into a joint behavioural process of negotiation aimed at reaching an agreement among the agents about how to attain that state.

Second, joint processes feedback into the system's joint mental state by affecting the individual agents' cognitive structures. In fact, joint processes produce results that directly affect the agents' cognitive structures. As the individual agent engages in joint processes, it may acquire new beliefs, fulfil its goals and intentions, or modify some of them. Experiential

wisdom accumulates as a result of positive and negative reinforcement of prior mental attitudes [6]. Mental attitudes that have led to what are encoded as positive outcomes are reinforced, while those that have led to negative outcomes are modified or discharged [40]. Finally, as a result of their impact upon the agents' mental states, joint processes affect the system's joint mental state as this is cognitively grounded on the individual agents' mental attitudes. For example, failure to successfully undertake negotiation may lead the agents (or some of them) to generate the belief that no agreement can reasonably be made. In this case, the agents will exit negotiation and the joint commitment they maintained with one another will be dropped.

In the light of the above observations, answering the second question - whether and how cognitive and behavioural processes are related to each other - is reasonably straightforward. As with the individual level, these processes are related also at the collective level. However, as opposed to the individual level, joint cognitive processes are not directly linked to joint behavioural processes, but they co-evolve through the mediation of the individual's cognition and behaviour. Thus, changes in either cognitive or behavioural joint processes need to impact on individual agents' cognition and behaviour before they can affect each other and show a consistent path of co-evolution. For example, should a joint commitment to achieving some state via some joint action be formed as a result of a joint cognitive process, this will affect the agents' cognitive structures to the extent that a new joint commitment needs to be consistently integrated within the agents' mental states. In turn, the joint commitment, via affecting the agents' mental states, will generate and control the performance of individual social behaviours that, ultimately, will converge into joint behaviour. Again, co-evolution between joint cognitive and behavioural processes rests on the impact that the outcome of joint behaviour has on the agents' cognitive structures and on the links that these have with individual cognitive processes and, therefore, with the system's joint cognitive ones.

As it stands, our model conveys a conception of joint cognition and behaviour that, even though sympathetic with the idea of allowing some forms of higher-order behaviour and group mind-like constructs to play a role in sociality, nevertheless is not consistent with an anthropomorphic image of the MAS. In fact, the theoretical position we endorsed suggests a notion of MAS in which forms of joint cognition and behaviour are always seen as rooted in the individual agents, namely in their cognitive structures, and their capacity of undertaking cognitive and behavioural processes.

2.6 Outcomes

Before we conclude this section, we need to briefly examine the role of the outcomes and their effects on the system's structure and functioning. We mentioned above that the outcomes may directly impact on the agents' cognitive structures. For example, should the

system generate new knowledge, this will affect the agents' cognitive structures and modify their set of information and beliefs. In turn, new configurations and instantiations of the agents' cognitive structures will impact upon the structure of the dependence network (as changes in beliefs, goals, preferences etc. drive changes in "who depends on whom"). Finally, different ways in which the agents depend on one another will ultimately correspond to different ways in which the system is internally structured and behaves. At an extreme, should the agents fulfil the goals on which their dependence network was grounded, they will exit the system as no further joint activity would be required with respect to these goals. Furthermore, the outcomes may impact on both the contextual factors and the organisational variables. For example, the system's activity can bring about a refinement of the techniques adopted, or it can impact on the type and quantity of resources used. Finally, besides impacting directly on the system, the outcomes may impact onto the environment and, via this, on the system. For example, changes in what resources are available, how scarce they are, how they are allocated among the agents, affect the ways in which different agents depend on one another, and therefore, whether or not a new MAS should be created and with what structure, or an existing one should be differently structured or even abandoned, should it be no longer consistent with the environment and the dependence network.

3. Organisation Theories and MASs

With this section, we begin a critical discussion of the model introduced above. In doing so, we will attempt firstly to look for common threads within organisational research and thought, secondly to show how an integrated organisational perspective can be used as the background for our model, and can help describe and reason about a wide range of forms for organising sociality. In accordance with the internal articulation of the model, we start our discussion with the factors that can be thought of as the determinants of sociality (section 3.1). Section 3.2 will be devoted to a critical evaluation of the environment-MAS relationship. The discussion will shed light on the complexity of this relation, and the ambiguous role played by the environment in its link with the MAS. It will be shown that MASs can be seen as the product of sociality-variation mechanisms reflecting either: (i) the environment's attempts to deterministically "select" the fittest forms of sociality; or (ii) the MASs' attempts to adapt their structure and behaviour to the environment; or (iii) the MASs' attempts to pro-actively shape their structure and behaviour by removing any environmental constraint.

3.1 The Antecedents of Sociality

As shown in Figure 1, what may determine whether or not a MAS is established, how it is structured, and ultimately how it evolves over time is the complex

interplay between the environment and the network of dependencies among agents. The relevance of these factors, and particularly how they relate to each other, will be the subject of the following sections.

3.1.1 Environment

"Just take the universe, subtract from it the subset that represents the organisation, and the remainder is the environment" (see [33], p. 195). We agree with the author of this definition when he adds that, unfortunately, it really isn't that simple. In fact, the issue of defining the environment is a delicate matter that has not only theoretical but also practical implications concerning the governance of pluralistic systems in the real world. First, the notion of environment becomes relevant as long as it is defined with respect to some object. Thus, before a MAS has come into existence, there is no environment as such. Instead, what we can identify is a set of causal factors that can be thought of as the determinants of the generation of a new MAS. Second, given an existing MAS, the problem of defining the environment becomes that of identifying the system's boundaries. Furthermore, building on our definition of MAS as a cognitive entity and the related assumption to conduct our analysis by endorsing the MAS's internal perspective (see section 2), boundaries are what the system considers as such. This raises two related questions. Firstly, how can we, as external observers, identify the factors, if any, that can be regarded as outside the boundaries of a MAS? Secondly, what are the mechanisms, if any, that enable these factors to impact on the MAS?

In order to answer the first question, we will endorse a conception of environment that is comprehensive and articulate enough to account for not only already existing MASs, but also for new ones that have not yet been established.

Definition 2. Before a MAS comes into existence, environment is the set of resources and phenomena that can determine whether or not the system is generated and what its structure and functioning will be. After a MAS has been generated, environment is the set of resources and phenomena that the system, as a cognitive entity, believes are outside its boundaries, and can affect its structure and functioning².

In the above definition, environmental resources may be either agents, physical objects, cognitive properties (e.g. abilities, know-how), techniques, technologies, norms, rules, and institutional values [38], while environmental phenomena may be simple events or actions performed by agents or MASs. The problem here is to understand under what circumstances these factors and phenomena can be thought of as external to

² Note that the definition is consistent with the fact that the MAS may ignore some of the resources and phenomena that can affect and constrain its structure and functioning. However, since the MAS does not take them into account, these factors are not part of the system's environment and, as such, do not represent a potential target at which the system's strategies, policies, activities, etc. may be addressed.

a MAS, thus representing its own environment. However, as might be expected, if we are to endorse the system's viewpoint to define its boundaries, we eventually fall into a vicious circle: as the decision does not rest on us, as observers, everything can be outside and inside the system. Unless we are the agents of the system³, there is no way out of this dilemma, and every attempt to identify what resource is outside the system is bound to fail. One way to circumvent this difficulty, however, is to give up any attempt to define what an environment is by identifying the specific resources and phenomena that do not pertain to a system, and instead try to identify the key properties of these resources, whatever they are, that can affect the nature of the system and how it behaves. To this end, building on Dess and Beard's work [15], we suggest three environmental dimensions: capacity, complexity and volatility. First, the *capacity* of the environment refers to the degree of availability of resources, and the degree of heterogeneity and/or substitution between them. Second, *complexity* refers to the degree of concentration among resources, namely how they are allocated among agents and who has control over what resource. Finally, the *volatility* dimension reflects the degree of instability of the environment, namely how predictable events and actions are over time. Now, having defined the properties of the environment, we turn to the second question above and try to understand how these properties can affect a MAS.

3.1.2 Dependence Network

A fundamental key for understanding how the environment can impact on MASs is to consider how the environment affects the ways in which agents depend on one another. However, before we do this (see section 3.1.3), we need to shed some light on the concept of dependence. As one of the major determinants of social behaviour, social dependence has long been studied by a variety of disciplines, ranging from sociology [4], organisation science [18, 43], to distributed artificial intelligence [17, 49]. In compliance with our objective to offer a synthetic organisational perspective on MASs, social dependence will be here explored building on the major contributions that have been given by the leading scholars in organisational research and thought in the last few decades.

Emerson's studies on dependence provided the spark for the subsequent development of the concept and its incorporation into organisation theory [18]. His general point was that there are different degrees of social dependence. In fact, an agent can be seen as dependent on another (i) in proportion to the former's need for resources or actions which the latter can provide and (ii) in inverse proportion to the availability of other resources or the abilities of other agents to provide the same resources and actions. In terms of a

³ Also in this case, however, defining the environment is not straightforward, as most of the time the agent's and the system's viewpoints do not overlap. In fact, the perspective of the system builds on and transcends the perspective of the constituent agents.

comprehensive theory of organisations, it is straightforward to note how this perspective is relevant for understanding the conditions under which organisations are created, grow, establish relations with other agents or organisations, and sometimes fail. In fact, by focusing on the possibilities of differing degrees of dependence, Emerson's research led other scholars to explore how and to what extent social dependence can be effectively managed, and therefore avoided or sought. Along these lines, Blau generalised Emerson's ideas by analysing the conditions under which agents will be able to maintain their independence in a situation of potential dependence, and the conditions that agents must meet to ensure the dependence of autonomy-seeking agents [4].

The potential of Emerson's studies for the development of a comprehensive theory of organisations has been emphasised by Thompson, who argued that the relations among the agents within an organisation can be represented in terms of their interdependencies [43]. Building on this, he sought to create a classification scheme that was general enough to deal with the range of interdependencies found in complex organizations. His intuition was that each type of interdependency (namely, sequential, pooled, and reciprocal) demands a certain type of coordination that will facilitate organisational effectiveness yet minimise costs. Drawing on Thompson's intuition, Williamson showed how a theory of the organisational boundaries could be developed by linking the effectiveness of various intra- and inter-organisational relations to a taxonomy of interdependencies among agents and economic transactions, under the constraint of the minimisation of coordination costs [50]. Even though Williamson's contribution is primarily restricted to the economic organisation and the economic transaction of goods and services occurring within and between economic organisations, it can nonetheless offer invaluable insights for a theory of MASs that is concerned not only with the economic organisation, but also with other types of pluralistic social systems in which relations and transactions between agents do not necessarily involve economic exchange.

Against this background, it is interesting to understand how the path that led from social dependence to the ontology of the organisation could be further extended so as to derive, from a dependence-based theory of organisations, a more general dependence-based theory of MASs. In this respect, there are three fundamental problems that need to be addressed. Firstly, as noted above, because MASs do not only refer to economic organisations, the notion of dependence needs to be built on a notion of exchange between agents that does not only account for the economic transaction of goods and services, but also for the exchange of information, the emotional exchange (e.g. friendship), and the normative exchange (e.g. exchange of influence, power, norms) [34]. Secondly, the notion of social dependence needs to become comprehensive enough so that two or more agents may happen to be dependent on one another even when no exchange at all (either economic or

social) occurs between them. A key to this refinement is Thompson's notion of *pooled* interdependencies [43], where two or more units contribute separately to a larger unit. In this case, to assume that a MAS is composed of interdependent agents is not necessarily to say that each agent is involved in some form of exchange with each other agent in any direct way. Yet, the agents are interdependent in the sense that unless each acts adequately, the total MAS is jeopardised. Furthermore, by rendering a discrete contribution to the whole, not only can failure of any agent threaten the whole and thus the other agents, but also each agent can be seen as supported by the whole. Thirdly, we need to introduce a further notion of social dependence so that two (or more) agents may happen to be dependent on one another even when between them there are no pooled interdependencies and when no exchange, in any of its forms (i.e. economic, informative, emotional, normative) takes place. This happens when the agents are dependent on one another through the mediation of a third agent. In this respect, we suggest the distinction between *symbiotic* dependencies, where exchange occurs directly between the parties involved, and *competitive* dependencies, where two or more agents depend on the resources/abilities of a third party. Competitive dependencies cannot be equated with pooled interdependencies because the former, as opposed to the latter, do not imply a collaborative effort of the dependent agents toward the achievement of a common goal. Nor do they involve any form of direct exchange between the parties, as the dependent agents are interested simply in the effects that each other's behaviour has on the third party. Therefore, understanding the importance of competitive dependencies is a fundamental key to widening the set of pluralistic systems to be conceptualised against a common theoretical background. For example, competitive dependence is the essence of those mechanisms of organising and governing large numbers of loosely interdependent agents, usually grouped under the common banner of "market". In fact, the firms who refer to the same client are competitively dependent on one another, in the sense that each depends on the resource that the client controls and is able to offer, and the actions it can perform.

3.1.3 Environmental Impact on Social Dependence

Having discussed the notion of social dependence, as a fundamental intermediating link between environment and MAS, we now need to show how the three environmental properties introduced above can affect the set of social dependencies among agents. In the following section, we will explore how dependence, affected by the environment, may impact on the structure of MASs. The first environmental property we mentioned is capacity. The capacity of the environment affects the nature and form of social dependencies in the sense that it determines whether the agents can use their own resources to pursue their goals, or need to lean on others'. Abundant capacity,

for example, and resources extensively available for a variety of purposes, widen the agents' abilities to act in isolation and lower their incentive to organise pluralistic forms of social activity. Conversely, scarcity of resources imposes constraints on the performance of solipsistic activities, and tensions on the functioning of existing MASs⁴. In the same line of argument, but with more emphasis on the distributional dimension, complexity affects dependence, in that it refers to who controls what, thus determining potential social relations among agents. Simple environments, characterised by resources available to everybody, enhance the agents' possibilities to effectively act in isolation, whereas highly complex environments, where resources are dispersed among the agents, restrict these possibilities, and constrain the structure and functioning of existing MASs. Finally, the degree of environmental instability affects social dependence by impacting on the accuracy of the information that the agents maintain about the environment. When there is a high degree of unpredictable change, the environment is highly volatile and dynamic, and the information about it highly uncertain. This makes it difficult for the agents to accurately predict the outcomes of various decision alternatives, and acting in isolation becomes a less attractive way of achieving expected results than alternative forms of joint activity that conversely may help create "negotiated" and more stable environments [50]. Furthermore, highly volatile environments impose tensions on already existing MASs, by making it more difficult to assess and predict the efficiency of their structure and the effectiveness of their behaviour. At the other extreme, when the environment is stable, acting in isolation becomes more effective and attractive, and social activity in existing MASs becomes easier to be efficiently and effectively carried out.

Before we leave the topic, it is interesting to note that the dimension of volatility has long been studied in the organisational literature, particularly in the form of *task uncertainty* [27, 43]. Whereas mainstream decision theory defines uncertain those situations in which one is unable to assign a probability to an event, organisational research has given a more articulate conceptualisation. In the first place, the longer the interval of time between the performance of a task and the generation of its outcomes, the more uncertain the task. In the second place, task uncertainty is related to the agents' cognition. In this respect, the more unstructured, uncodified, inaccurate are the information and know-how of the agent(s) who is to perform a task, the more uncertain is that task. As it turns out, this notion of task uncertainty emphasises the agents' abilities to perform effective actions by focusing on the impact that the environment has on their know-how, and is therefore consistent with our notion of volatility that refers to the impact of the environment on the accuracy of the agents' information

⁴ However, it could be argued that a certain amount of scarcity of resources is necessary in order for a MAS to have the incentive to behave efficiently and to stimulate innovation [26].

concerning "what is the case" and "how to get things done".

3.2 The Environment-MAS Relation: Sociality-Variation Mechanisms

The issue of the environment-MAS relation, in its simplest form, can be presented as the role of the environment *versus* the role of the MAS. It is analogous to nature-nurture debate over formative influences on the development of humans. The more sociologically interesting topic is to determine under what conditions each is most influential, and what questions and explanations are made salient by a study of complex forms of sociality. Every MAS depends on its environment to some degree. However, we cannot ignore that some are much more dependent than others are. One way to understand the key issues involved here is to consider two ideal paradigmatic models of variation in sociality: an *environmental selection* model and a *rational selection* model. The former places heavy reliance on environmental forces and their tendencies to select the fittest forms of sociality, whereas the latter emphasises the active role of agents and MASs in creating and shaping sociality. By presenting the underlying assumptions and characteristics of each mechanism, we will take some steps towards a new integrated perspective. Bypassing the problem of unambiguously defining a "unique" form of sociality that comes into existence and changes over time in a unique specified manner, we will propose a new conception of MASs, here regarded as pluralistic systems that, with respect to their environment, can be either passive, reactive or proactive, depending on what environmental conditions they need to face and how prepared they are to face them.

3.2.1 Environmental Selection: The Case of Emergence

On theoretical grounds, to say that MASs are selected by the environment means rejecting the teleological implications inherent in theories of societal or organisational development based on the role of individuals, and model the source of variation in sociality in terms of any variation-generating mechanism that is built more on unplanned than planned change⁵. In this respect, a great deal of variation is introduced into a MAS through error or random variation, rather than through the rational planning and the choice of alternatives pursued by the system and its members. The reasons why a MAS is generated or an existing MAS is structured and behaves in a particular way may be only dimly understood by the members, even though they will probably attribute the outcomes of the system's

⁵ As explained in Note 2, the MAS may ignore some of the factors that can constrain its structure and behaviour. Therefore, also (some of) these factors can activate natural selection mechanisms that can lead the MAS towards unplanned results. In this section, however, we focus our attention on the selection pursued by those factors that are part of the system's environment (see Definition 2).

activities to their own efforts at rational selection. When the individual agents are relatively powerless to affect the process of generating and structuring sociality, what we need to do in order to understand the sources of variation in sociality is to shift the burden of choice from the individuals' cognition to the working of an impersonal "invisible hand". This invisible hand is the driving force behind the *emergence* of new MASs or new structures and forms of behaviour of existing MASs [1, 16, 22, 38, 49].

The idea of a MAS as an emergent system that is "decided" by the environment without the mediation of the agents' cognition has a two-fold implication. Firstly, a set of agents may happen to become members of a new MAS without having decided to do so. Secondly, an existing MAS may happen to have a structure and to perform behaviours that cannot be totally explained in terms of its own and its members' decisions. In Figure 1, the process of environmental selection has been represented in two related ways. Firstly, through the relation between the dependence network and the MAS. In fact, one of the main reasons that may explain the emergence of MASs is that the agents may not be aware of (some of) their (inter)dependencies. However, being dependent on some other agents may lead them, in trying to pursue their goals, to naturally behave in such a way that new forms of sociality are brought about without being *a priori* represented in their mental states, and the structure of existing forms of organising sociality may change without being the result of rational decisions. Secondly, on a more general level, environmental selection can directly affect the MAS, bypassing the intermediation of the dependence networks. For example, in some cases the environment may be so volatile that any attempt by the agents to predict its future configurations will inevitably fail and so will their attempts to manage the environment and change the system's structure by using tactics and strategies. Or furthermore, the structural conditions of a MAS may be so persistent and inertial, that any decision to change them will inevitably have effects procrastinated in time, and therefore different from what was initially expected.

Within the organisational literature, the late Seventies have seen the growth and development of what certainly stands as the major landmark contribution favouring this extreme environmental-imperative position. This contribution - which has been labelled the natural selection or *population ecology* (PE) view [1, 22] - argues that the environment selects certain types of organisations to survive and others to perish, on the basis of the fit between their structure and environment. Drawing on biology's survival-of-the-fittest doctrine, the PE view claims that organisations that survive have resources and structural dimensions that the casualties do not. Along these lines, shifting the perspective from what had long been the focus in economic and organisational research, it has been argued that it is not the organisation, but the environment, that optimises [22].

As it stands, in its emphasis on the active role of the environment and the limited degree of freedom of the organisation, this position provides contributions as to how to theorise about MASs that need to cope with complex, constantly changing, and uncertain environments. In fact, some of the conditions that, according to the PO view, would favour the activation of mechanisms of natural selection, seem to be particularly suited to describe the challenges that some of today's MASs need to deal with. These conditions are: the high volatility of the environment and the related complexity and uncertainty of the information about it; the dimension of conflicting interests and objectives of the agents involved in most social systems, which makes these systems appear like forms of *integration of the many* rather than holistic super-individual entities, and makes any attempt to internally re-organise their structure rather costly and inconvenient; the costs of innovation; the specialisation of techniques and technologies and the related problem of high costs of information and barriers to entry in most markets; and, finally, the problems of the irrationality of the macro-behaviour of multiple interconnected units, which is particularly true in today's internet-based global communities where a behaviour that is effective when undertaken by a single agent may not turn out to be so when simultaneously undertaken by many related agents.

Even though shifting the focus from the MAS to its environment emphasises the challenges the latter imposes onto the former in terms of inescapable constraints, nonetheless the PE view is limited by the fact that it obscures the MAS's power and ability to partially control the environment in which it acts. This is a major limitation because some of today's MASs still maintain control over their environment, and can insulate themselves against failure. Along these lines, it can be argued that a comprehensive model of MASs needs to be broad enough to account for a variety of forms of organising sociality, ranging from MASs with no control over their internal configuration and deterministically constrained by their environment, through to those that can balance the challenges of the environment with their own abilities to decide how to be structured and behave, and finally MASs that have complete control over their own decisions and behaviour, and can even remove any environmental constraint. In order to account for all these types of MASs, we need to integrate the environmental selection mechanism with new mechanisms of generation of sociality, focused on the MAS's possibilities to rationally select its own structure and behaviour. This will be the topic of the next section.

3.2.2 Rational Selection

Rational selection places reliance on the role of the agents' and the MAS's cognition in managing the challenges posed by the environment. Given the set of environmental constraints, two or more agents may decide whether or not they should join and establish a MAS. Similarly, an existing MAS, through its

members, can decide how to configure its structure and how to behave, given a set of environmental constraints. In both cases, the environment impacts on sociality to the extent that it is cognitively represented in the agents' cognitive structures, and via these, in the system's joint mental state. As shown in Figure 1, rational selection is represented in two related ways. Firstly, through the role of the dependence network. In fact, "who depends on whom" may ultimately determine whether or not two or more agents, aware of their dependencies, decide to establish a new MAS. Secondly, through the impact that the outcomes of the system's processes have on the components of its structure. The system may decide how to be structured, and its decisions and behaviours may lead to outcomes that affect not only the agents, but also the contextual factors and the organisational variables. Besides, by impacting on the agents, the system's outcomes affect their dependencies, and again via these, the system's future configurations.

Rational selection is built on two fundamental processes [24]. The first is *rational planning*, by which expectations about future consequences are used to choose among current alternatives (section 3.2.2.1). The second process is *learning from experience* (section 3.2.2.2). Through learning, feedback from previous experience is used to choose among present alternatives. Both processes are based on the assumption that the environment can be (partially) controlled, in the sense that it provides a set of constraints to which a MAS can adapt its own structure and behaviour. However, besides these processes, there is another one that, while building on both of them, rejects and goes beyond the assumption of a partially controllable environment. This process is *rational proactiveness* (section 3.2.2.3). Either through rational calculation or through learning by trial and error, not only can adaptation to environmental constraints be ensured, but also these constraints can be changed and removed. All of these three processes will now be dealt with in turn.

3.2.2.1 Rational Planning

Rational planning represents a variation-generation mechanism that, building on the individuals' cognitive structures and abilities, is used to explain change in sociality resulting from rational calculation. It inherently rests on a form of cognition that is regarded as a forward-looking form of intelligence premised on an agent's beliefs about the link between the choice of actions and the subsequent impact of those actions on the environment. Typically, agents and MASs can resort to rational planning when: (i) the environment is stable enough to provide a schedule of responses to alternative actions on the part of the agents and the MAS; (ii) the agents (and the MASs) are aware of their own goals, intentions and preferences, and these influence their behaviour; and (iii) the mechanisms through which the agents' cognition and behaviour are transformed into the MAS's cognitive and behavioural processes are known. Under these assumptions, agents

and MASs have a degree of freedom from environmental forces, and changes can be brought about in sociality either by the agents' decisions as to whether and how to generate a new MAS or by MASs' decisions about the effectiveness and efficiency of their structures.

The claim that the system's effectiveness and efficiency are function of the combination between its structure and environment can be regarded as the main tenet of what is known as the *Structural Contingency* (SC) theory of organisations [27]. During the Seventies, this was the theory that represented the dominant paradigm within organisational research. Because all organisations are dependent, in some degree, on their environment, they need to meet the environmental pressures by appropriate structures [20, 27, 43]. Many of the propositions of the SC theory directly draw on March and Simon's cognitive theory of organisations as systems for processing information [29]. Two main concepts are borrowed from the cognitive approach: *differentiation* of an objective into distinct sub-objectives, and *integration* among sub-objectives. Differentiation and integration are then used to help generate indications as to how to build organisational structures with differing degrees of flexibility, or capacity of adaptation, or even capacity to process information and solve problems.

Even though the emphasis has been primarily on the organisation of economic activity, there are many ways in which these indications can be used to enhance MAS theory and practice. Firstly, drawing on the SC claim that there are no universal principles of management, it can be similarly argued that there is no best way to manage or organise a MAS. What structure it should have depends upon the circumstances and the environment. In this view, MAS management practice is entirely *situational*. Secondly, since MAS management is situational, it needs to sharpen its *diagnostic* skills, so it will know the right thing to do at the right time. In this respect, the contingency approach can supply the management of MASs with diagnostic concepts, tools, methods and techniques that are helpful in analysing and solving situational problems. Finally, since there is no value imperative on MAS management other than to diagnose situations correctly, all managerial tasks need to be integrated in order for the MAS to react with rational expediency. Again, the contingency models are useful in this respect as they can be used for a variety of purposes, ranging from strategy formulation, to organisational design and redesign, information decision systems, influence systems and leadership.

Despite its central role in reaffirming the necessity for researchers and practitioners to dwell upon the practical, day-to-day, situational puzzle-solving tasks, the SC theory has been heavily criticised for its paucity of content. Many modern organisation theories arose precisely as a response to that approach, all advocating the possibility of different ways in which the adaptation of the organisation to its environment may occur. Among these theories, in this section we will review some that are built on the

rational planning mechanism; others, centred around the concept of adaptive learning, will be dealt with in next section.

The main line of argument along which some scholars in organisation theory have criticised the SC approach is the following: The adaptation of the organisation to its environment cannot be represented by a nomological law that, to each environmental configuration, assigns only one type of efficient organisational structure. The reason for this is three-fold.

First, there is more than one strategy through which the organisation can act, either in stable or more uncertain environments, and each strategy has its own structural implications. As Child pointed out, the organisational design cannot be seen as the product only of technical demands presented by environment, technology and size [9]. He argued that these variables provide constraints to the set of feasible organisational solutions, but within this set the choice of a particular solution depends on the strategic decision-making carried out by the actors involved in the organisation. Second, the "one-best-way-of-organising" prescription for each type of environment is not entirely justifiable in environments with abundant capacity of resources. In fact, in these situations, either a strictly efficient structure or a structure characterised by "organisational slack"⁶ can be equally adopted, where the latter, even though not justifiable in terms of efficiency criteria, can nonetheless be explained in terms of higher degrees of freedom and power of (some of) the constituent members [13]. Third, the response to an increase in environmental uncertainty cannot be only the adoption of a more flexible organisational structure, as the SC approach seems to suggest, internally more differentiated and based on integrative mechanisms that rely on mutual adjustments among differentiated units [27, 43]. Rather, besides the generation of slack resources that may help adaptation when external pressures become higher and more variable, the organisation may change its boundaries. This is the main line of argument that has been used by two research programmes. On the one hand, the *Markets & Hierarchies* (MH) programme, building on the key behavioural assumptions of actors' bounded rationality and opportunism, shifted the emphasis from the organisational unit to the transaction between units, and addressed the organisational design problem as a transaction cost issue [50, 51]. Against this background, the crucial argument was that unrealised efficiency opportunities always offer an incentive to redefine the boundaries among organisations and/or among internal organisational units until the transaction costs are minimised. On the other hand, the possibility to modify organisational boundaries has been emphasised by the *Resource Dependence* (RD) research programme [37]. This programme rests on the assumption that the problem of adapting the organisational structure to the environment cannot be

explained only in terms of the organisation's need to cope with the environmental uncertainty. Rather, many of the attempts that organisations make to ensure their future survival and growth are primarily related to the problem of managing their dependencies on resources that are controlled by other actors/organisations within the environment. In this perspective, avoiding or weakening resource dependence and/or generating or increasing others' dependence are the driving force behind a number of boundary-spanning strategies, ranging from those impacting on intra-organisational variation (e.g. diversification; horizontal and vertical integration), to those relying upon inter-organisational variation (e.g. joint ventures; interlocking directorates). Finally, besides the emphasis on the organisational boundaries, the RD approach goes beyond the SC "best-fit" principle by admitting that, under conditions of relatively abundant resources, the burden on the organisation becomes weaker and the environmental pressures can be met with more than one efficient boundary-spanning solution.

As with the SC approach, all the above theories provide a number of tools and principles that, once adopted to theorise about and manage MASs, can help take some steps towards a science of MASs that is theoretically more comprehensive and empirically more satisfactory. The first lesson that we learn is that, when the conditions for rational planning are satisfied, the issue of how to best structure a MAS may be ill-founded. Besides the difficulties related to the problem of real-world agents' bounded rationality, there may be more than one efficient way to structure a MAS, and the relative desirability of these ways may depend on the interplay of the (conflicting) interests of the constituent agents. Furthermore, within the set of possible solutions, the MAS may also consider the modification of its boundaries, namely changing the number of its members and, more interestingly, establishing inter-systemic relations with other MASs. Finally, the notion of "slack" is new to MAS research, and as such it may help understand to what extent structural configurations that are *prima facie* inefficient from a cost/benefit viewpoint, can be regarded as the efficient solutions to the two-pronged problem of enabling the achievement of the system's objectives *and* meeting the members' needs to fulfil their own private interests.

3.2.2.2 Adaptive Learning

When the environment is too volatile to be predicted with sufficient accuracy, but at the same time is too stable for a mechanism of environmental selection to make the agents powerless in their cognitive and behavioural processes, variation in sociality cannot take place as a result of a rational calculation of the appropriate means with respect to some ends. In these situations, agents and MASs need to resort to processes of trial and error that, while still built on cognition, nonetheless emphasise a form of intelligence that is premised more on experiential learning than on a forward-looking rational calculation of expectations.

⁶ This is the "disparity between the resources available to the organisation and the payments required to maintain the coalition" ([13] p. 42).

We can imagine a MAS as learning from its experience, modifying its behaviour incrementally on the basis of feedback from the environment. Solutions are tried, their outcomes experienced, and subsequent revisions may occur. Besides the environmental conditions of volatility and ambiguity specified above, there are four main conditions under which a mechanism of learning via trial and error can be more effective than, and sometime the only feasible alternative to, rational planning.

First, agents and MASs may be partially unaware of their goals, intentions and preferences. The ambiguity of mental attitudes is a fundamental factor that, together with environmental volatility, can affect how sociality is organised. Agents and MASs can learn what to expect, what to aim for, how to form their goals, how to commit themselves to them, and how to structure their preferences. Attitudes can be seen as changing over time in response to experience. Second, agents and MASs may be unaware of their abilities; also, they can learn and improve abilities through experience. Learning abilities leads them to be more competent in the activities they have undertaken, and relatively less competent in the activities they have not undertaken. Third, agents' cognition may be only loosely linked to their behaviour. The situation is one that reflects some important inertial dynamics of MASs and their ability to inhibit the modification of individual behaviour on the basis of individual cognition. Fourth, the individual agents may be only partially capable of affecting the system of which they are members, because the mechanisms through which cognition and behaviour at the individual level are related to cognition and behaviour at the collective one may be unknown or ambiguous. When these assumptions are satisfied, the degree of ambiguity and opaqueness is further increased, and the most effective sociality-variation mechanism is to adaptively learn how to change sociality by trying solutions before thinking what solutions to try.

Within organisational research, the most influential background to the development of adaptive experiential learning models can be identified in Simon's work on bounded rationality [40]. Not only are individuals computationally and informationally limited in generating and using cognitive representations of reality, but they are also biased by the tendency to keep such representations also when confronted with empirical evidence that conspires against their accuracy [44]. Similar problems have been addressed with respect to the organisation by scholars with a more cognitively psychological background. Besides the role of time and attention as scarce resources and the role of relatively stable organisational routines [29], one of the most dominant speculations was that the environment is not an objective independent variable, but is what is perceived by the organisation. Should the organisation decide to adapt itself to the environment, this adaptation is intimately biased by the organisation's "choice" of what the environment is, and what are the relevant problems to which attention should be devoted. Along

these lines, Weick's notion of "enacted environment" [48] emphasises the fact that the organisation unintentionally brings about changes in its perceived environment, while trying to adapt itself to environmental modifications that have been caused by its own attempts of adaptation.

Two major models of organisation find their roots within the above conceptual framework of environmental ambiguity, cognitive indeterminism and behavioural adaptation. The first is the *Garbage Can Model of Organisational Choice* proposed by Cohen, March, and Olsen [10]. Here the organisation is viewed as an "organised anarchy" characterised by "a collection of choices looking for problems, issues and feelings looking for decision situations in which they might be aired, solutions looking for issues to which they might be the answer, and decision-makers looking for work" [10]. In this view, problematic preferences, unclear technology, and fluid participation make the organisation's typical choice opportunity appear like "a garbage can into which various kinds of problems and solutions are dumped by participants as they are generated" [10]. The second model to be mentioned here is the "*Behavioral Theory of the Firm*" proposed by Cyert and March [13]. Here, the organisation is modelled as an adaptively rational institution, "constrained by the uncertainty of its environment, the problems of maintaining a viable coalition, and the limitations on its capacity as a system for assembling, storing, and utilizing information" ([13] p. 117). Building on the ideas of bounded rationality, imperfect environmental matching, and unresolved conflict, the authors attempt to formalise a theory of organisational goals, expectations and choice. In doing so, four major relational concepts are introduced as fundamental to an understanding of the decision-making undertaken in a modern, large-scale business organisation. These concepts are: quasi resolution of conflict (i.e., conflict among members' goals is resolved by using local rationality, acceptable-level decision rules, and sequential attention to goals); uncertainty avoidance (i.e., organisations avoid uncertainty by using feedback-react decision procedures and by arranging a negotiated environment); problemistic search (i.e., search is stimulated by a problem and is directed towards finding a solution to that problem); and organisational learning (i.e., organisations exhibit adaptive behaviour over time).

Linking this conceptual framework to MAS research is quite straightforward if we think that many of the implications of both behavioural models discussed above heavily rely upon computer simulations. However, besides this methodological aspect, the research thread here discussed seems to have the potential to affect one of the core conceptual tools adopted in MAS theory and practice, namely the Belief-Desire-Intention logic and agent architecture [52]. If mental attitudes may be ambiguously defined, if agents may not be aware of their intentions and goals, then behaviour cannot always be seen as triggered by the agents' mental states. Behaviour may also come *before* cognition, and therefore action

specifications should be allowed to be instantiated before and independently of the instantiation of the mental attitudes. These, in turn, become a function of the agent's actions. In this view, actions are not rational because they are planned on the basis of mental attitudes that are rational *a priori*; rather, mental attitudes are rational because they are adaptively adjusted to actions that are rational *a posteriori*. Another interesting implication of the adaptive learning framework has to do with the role of randomness in MASs. Not only under the conditions of unclear technology and agents' ambiguous preferences, but also when the number and identity of the agents is problematic and mutable, MASs may look more like "organised anarchies" than structured systems, and their behaviour more like a "garbage can" than an organised process. These conditions, for example, are satisfied by today's internet-based communities where members randomly enter and exit without knowing each other, thus making the communities' size and boundaries extremely uncertain. In such situations, the nature of the MAS's behaviour, the time it takes to be performed, and the problem it solves may depend on a relatively complicated intermeshing of elements, ranging from the mix of choice opportunities available at any one time, the mix of problems that have access to the MAS, and the mix of solutions looking for problems. To some extent, extending the concept of MAS so as to cover not only the traditional notion of integrated system with coordinated tasks, but also the idea of "loosely coupled" systems in which everybody can do everything and where behaviour is loosely coupled to intentions, may have a paradoxical consequence. In fact, loosely coupling phenomena ensure that some degree of foolishness will occur within MASs, no matter how useful for rational coordination and control it may be. Following March's argument [30], it might then be speculated that a deep understanding of the actual behaviour of computational MASs should rest not only on a good technology of distributed intelligence but also on a good technology of "distributed foolishness".

3.2.2.3 Rational Proactiveness

When tension on resources is low enough to enable the agents to control their own dependencies, but at the same time high enough to ensure a differentiation in the agents' control over resources, and furthermore when environmental volatility is too low for natural unplanned selection to take place, but at the same time high enough to ensure a reasonable amount of differentiation in the agents' abilities to cope with uncertainty, then sociality may be generated through mechanisms that heavily rely upon agents' proactiveness. The issue of proactiveness is a delicate one as it rests on a radical reconsideration of the boundaries between MAS and environment. With emergence, the environment is seen as an independent variable that can deterministically affect sociality by making agents' cognition relatively powerless. With rational planning and adaptive learning, the

environment is still regarded as an independent variable that, by providing a set of constraints to agents' cognition, impacts on sociality with varying degrees of intensity. When rationally selected, sociality is therefore the outcome of decisions under environmental constraints. Conversely, proactiveness draws on the idea that the environment is relatively impotent in its effect on sociality. Not only can sociality be rationally selected (either by calculation or by learning), but also this selection is unconstrained: whether and how sociality is generated and changes over time ultimately depends on the agents' and MASs' decisions. To some extent, this line of argument may lead to a re-assessment of the old sociological speculation about the micro-macro link. In fact, being proactive for a MAS means not only deciding and acting under no environmental constraints, but also transforming its own deciding and acting into constraints for others' deciding and acting. Therefore, there is no environment as such for the proactive MAS; rather, the MAS *is* the environment for other MASs.

Within organisational research, the notion that organisational structure and behaviour are imbued with logic, cognition and rationality has historically led to emphasise the role of strategic choice in the exercise of power by the organisation over its internal structure and its relations with the environment. In section 3.2.2.1, we have already mentioned the contribution of Child as to the explanation of decision-making in terms of strategic events built on the "value positions" and preferences of those who hold power [9]. In turn, these studies were prompted by Chandler's idea of the managers' "visible hand" in shaping the structural forms of American industrial enterprises [8]. Here, managers are viewed as acting in a proactive role, and their behaviour is described as intendedly purposeful, hedonistic, and rational. Phenomenologists have provided additional impetus for this view. They suggested that organisations are largely the product of choice and the way individuals construct their social realities [3]. Finally, recent advances in cybernetics and autopoietic systems theory suggest the idea of the organisation as a cognitive self-referential and self-organising system where coordination among its components is designed and implemented from within [32]. In this view, the organisation is autonomously self-referential, in that it subordinates its behaviour to maintaining itself and its internal structure; also, it is autonomously pro-active, in that its boundaries and the type of environment in which it acts entirely depend on how its auto-creation is undertaken.

Within the MAS literature the concept of self-organising MASs has been partially considered by researchers interested in designing the best match among task, environment, structure and performance [42]. However, these studies still conceive the environment in the usual structural-contingent way of a set of constraints to which the MAS needs to adapt. Along these lines, the concept of proactiveness in MAS research is meant to convey the idea of an agent's being able to exhibit goal-directed behaviour [52]. However, this behaviour is still seen as performed by

agents who are constrained by and embodied in some *a priori* environment. Conversely, the studies we discussed in this section suggest the idea that proactive agents may not only exploit, but also create serendipity [52]. Further, for proactive agents, there may be not an *a priori* environment in which they can be modelled as embedded. At most, in such cases, if we still are to maintain a notion of environment, this should be defined only *a posteriori*, after and as a result of the agents' actions, as the set of constraints that these agents, by acting, impose on others' behaviour.

4. Conclusions and Future Work

We started this paper by proposing that there is a need for a new science of MASs. This idea was rooted in the observation that many forms of sociality (e.g. large-scale loosely coupled knowledge communities; virtual organisations; large business corporations) cannot be adequately understood if we limit our analytical tools to those offered by mainstream MAS research. Confronted with this downside, we advocated the possibility of looking for the contributions given by some of the leading scholars in organisation theory, in an attempt to define a new, richer and more complex notion of sociality. Our case for the advantages of this new integrated approach was presented throughout the paper in a number of steps. First, a new conception of MAS has been introduced, distinct from and broader than the definitions traditionally offered by MAS theory. Second, a model centred around this notion of MAS has been developed. Third, the main sociality-variation mechanisms underpinning this model have been discussed in the light of a new integrated perspective that brought together a number of divergent contributions within organisational thought, all prevalently centred around the relation between organisation and environment.

A science is fundamentally an articulated combination of *problems*, *theories* to examine and explain these problems, and *methodologies* to empirically evaluate these theories. A new science of MASs needs to be built upon all these components. In this paper, an attempt has been made to suggest some of the problems that, in our view, should confront the MAS community. Further, some steps towards a new theory of MASs for addressing (some of) these problems have been taken. This theory has been articulated into a model of MAS and a number of implications about what structure and behaviour a MAS is expected to exhibit under different conditions. As such, the theory appeals to a notion of MAS that is consistent with a dual-faceted perspective. First, MAS as an *analytical tool* for representing and reasoning about sociality. Second, MAS as a set of high-level *organisational principles* that can be used by agents for organising a wide range of interrelationships and interactions among them. In this respect, the theory can be used not only for enhancing our understanding of how sociality and its organisation may vary over time, but also as a tool for improving practitioners' ability to design and manage efficient architectures for

computational MASs. Even so, however, the theory, as such, is incomplete. In fact, it is incumbent upon any theoretical attempt at describing, explaining, or predicting new facts, to also provide a set of propositions that can be empirically controlled and tested [25]. This methodological imperative, in turn, calls for the *operationalisation* of the variables through which the propositions of the theory are expressed. These considerations, therefore, suggest the avenues along which our future work will progress.

Finally, on a more epistemological level, the results of our efforts can also be evaluated in terms of what has been done and what is to be done next, towards meeting our initial challenge of a new science of MASs. First, a science needs to be appraised from a theoretical perspective, namely in terms of what type of theories are used to address what problems. This is precisely the arena in which this paper can be placed. Obviously, what is required next is the development of other different (maybe conflicting) theories, that appeal to different scholarly fields in order to provide an increasingly richer and more articulate notion of sociality. Second, a science finds its specificity in the type of methodologies and methods it resorts to. The hallmark of current MAS practice is the use of the computer to engineer systems that can autonomously make good decisions about what to do [52]. If this is the case, a new science of MASs should then be expected to be centred around the use of software engineering for designing and building computational MASs. Besides the benefits they can provide in terms of real-world applications (e.g. in electronic commerce), these systems can also be used as computational models that, once validated, can enhance our understanding of the "real" systems of which they are the analog.

References

1. Aldrich H. E. (1979). *Organizations and Environments*, Englewood Cliffs, N.J., Prentice Hall, Inc.
2. Benjafield J. G. (1992). *Cognition*, Englewood Cliffs, N.J., Prentice Hall, Inc.
3. Berger P. L. and Luckmann T. (1966). *The Social Construction of Reality*, Garden City, N.Y., Doubleday.
4. Blau P. M. (1964). *Exchange and Power in Social Life*, New York, N.Y., John Wiley & Sons, Inc.
5. Boulding K. E. (1956). General systems theory - The skelton of a science, *Academy of Management Journal*, 202-5.
6. Carley K. M. (1989), The value of cognitive foundations for dynamic social theory, *Journal of Mathematical Sociology*, 14(2-3), 171-208.
7. Carley K. M. (1996). A comparison of artificial and human organizations, *Journal of Economic Behavior and Organization*, 896, 1-17.
8. Chandler A. (1962). *Strategy and Structure*, Garden City, N.Y., Anchor Books.
9. Child J. (1972). Organizational structure, environment and performance: The role of strategic choice, *Sociology*, 6, 1-22.
10. Cohen M. D., March J. G. and Olsen J. P. (1972). A garbage can model of organizational choice, *Administrative Science Quarterly*, 17, 1-25.

11. Cohen P. R. and Levesque H. J. (1990). Intention is choice with commitment, *Artificial Intelligence*, 42(3), 213-261.
12. Coleman J. S. (1974). *Power and the Structure of Society*, New York, N.Y., Norton.
13. Cyert R. M. and March J. G. (1992). *A Behavioral Theory of the Firm*, Cambridge, MA, Blackwell Publishers.
14. Dennet D. C. (1987). *The Intentional Stance*, Cambridge, MA, The MIT Press.
15. Dess G. G. and Beard D. W. (1984). Dimensions of organizational task environments, *Administrative Science Quarterly*, 52-73.
16. De Jong K. A. (2001). *Evolutionary Computation. A Unified Approach*, Cambridge, MA, The MIT Press.
17. d'Inverno M. and Luck M. (1996). A formal view of social dependence networks, in Zhang C. and Lukose D. (eds.), *Proceedings of the First Australian Workshop on Distributed Artificial Intelligence*, Lecture Notes in Artificial Intelligence, Vol. 1087, Berlin, Springer Verlag, pp. 115-129.
18. Emerson R. M. (1962). Power-dependence relations, *American Sociological Review*, 27, 31-40.
19. Friedkin, N. E. (1998). *A Structural Theory of Social Influence*, Cambridge, MA, Cambridge University Press.
20. Galbraith J. R. (1973). *Designing Complex Organizations*, Menlo Park, CA, Addison Wesley.
21. Gasser L. (1991). Social conceptions of knowledge and action: DAI foundations and open systems semantics, *Artificial Intelligence*, 47, 107-138.
22. Hannan M. T. and Freeman J. H. (1977). The population ecology of organizations, *American Journal of Sociology*, 82, 929-964.
23. Henderson L. J. (1935). *Pareto's General Sociology*, Cambridge, MA, Harvard University Press.
24. Kurihara S., Aoyagi S. and Onai R. (1997). Adaptive selection or reactive/deliberate planning for the dynamic environment, in Boman M. and Van de Welde W. (eds.), *Multi-Agent Rationality, Proceedings of MAAMAW-97*, Lecture Notes in Artificial Intelligence, Vol. 1237, Berlin, Springer Verlag, pp. 112-127.
25. Lakatos I. (1970). Criticism and the methodology of scientific research programmes, in Lakatos I. and Musgrave A. (eds.), *Criticism and the Growth of Knowledge*, Cambridge, MA, Cambridge University Press.
26. Lawrence P. and Dyer D. (1980). Toward a theory of organizational and industrial adaptation, working paper, Harvard Business School.
27. Lawrence P. R. and Lorsch J. W. (1967). *Organization and Environment: Managing Differentiation and Integration*, Harvard University Graduate School of Business Administration, Boston.
28. Lorrain F. and White H. C. (1971). Structural equivalence of individuals in social networks, *Journal of Mathematical Sociology*, 1, 49-80.
29. March J. G. and Simon H. A. (1958). *Organizations*, New York, N.Y., John Wiley & Sons.
30. March J. G. (1988). *Decisions and Organizations*, New York, N.Y., Basil Blackwell.
31. Maruyama M. (1968). The second cybernetics: Deviation-amplifying mutual casual processes, in Buckley W. (ed.), *Modern Systems Research for the Behavioral Scientist*, Aldine.
32. Maturana H. R. and Varela F. J. (1980). *Autopoiesis and Cognition. The Realization of the Living*, Dordrecht, Holland, Reidel Publishing Company.
33. Miles R. H. (1980). *Macro Organizational Behavior*, Santa Monica, CA, Goodyear Publishing.
34. Mitchell J. C. (1973). Networks, norms and institutions, in Mitchell J. C., *Social Network in Urban Situation*, Manchester, U.K., Manchester University Press.
35. Nonaka I. and Takeuchi H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovations*, New York, N.Y., Oxford University Press.
36. Panzarasa P., Jennings N. R. and Norman T. (2001). Formalising collaborative decision-making and practical reasoning in multi-agent systems, *Journal of Logic and Computation*, forthcoming.
37. Pfeffer J. And Salancik G. R. (1978). *The External Control of Organizations. A Resource Dependence Perspective*, London, U.K., Harper and Row.
38. Powell W. W. and Di Maggio P. J. (eds.). (1991). *The New Institutionalism in Organizational Analysis*, Chicago, The University of Chicago Press.
39. Rao A. S. and Georgeff M. P. (1998), Decision procedures for BDI logics, *Journal of Logic and Computation* 8(3), 293-342.
40. Simon H. A. (1955). A behavioral model of rational choice, *Quarterly Journal of Economics*, 69, 99-118.
41. Singh M. P. (1995). *Multiagent Systems: A Theoretical Framework for Intentions, Know-how, and Communications*, Lecture Notes in Computer Science, Vol. 799, Berlin, Springer Verlag..
42. So Y. And Durfee E. H. (1988). Designing organizations for computational agents, in Prietula M., Carley K. N. and Gasser L. (eds.), *Simulating Organizations. Computational Models of Institutions and Groups*, Menlo Park, CA, AAAI/The MIT Press, pp. 47-64.
43. Thompson J. D. (1967). *Organization in Action*, New York, N.Y., McGraw Hill.
44. Tversky A. and Kahnemann D. (1982). Causal schemas in judgements under uncertainty, in Kahneman D., Slovic P. and Tversky A. (eds.), *Judgement under Uncertainty: Heuristics and Biases*, Cambridge, MA, Cambridge University Press, pp. 117-128.
45. von Bertalanffy L. (1972). The history and status of general systems theory, *Academy of Management Journal*, 407-26.
46. Wasserman S. and Faust K. (1994). *Social Network Analysis: Methods and Applications*, Cambridge, MA, Cambridge University Press.
47. Weber M. (1978), *Economy and Society*, ed. G. Roth and C. Wittich, 2 Vols, Berkeley, CA, University of California Press.
48. Weick K. E. (1969). *The Social Psychology of Organizing*, Reading, MA, Addison-Wesley.
49. Weiss G. (ed). (1999). *Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence*, Cambridge, MA, The MIT Press.
50. Williamson O. E. (1985). *The Economic Institution of Capitalism: Market, Hierarchies and Relational Contracting*, New York, N.Y., Free Press.
51. Williamson O. E. and Ouchi W. C. (1981). The markets and hierarchies program of research: Origins, implications and prospects, in Van de Ven A. H. and Joice W. F. (eds.), *Perspectives on Organization Design and Behavior*, New York, N.Y., John Wiley & Sons, pp. 347-370.
52. Wooldridge M. (2000). *Reasoning About Rational Agents*, Cambridge, MA, The MIT Press.

Figure 1. A Unified Framework for Representing and Reasoning about Multi-Agent Systems.

