A Conceptual Framework for Self-Organising MAS

Andrea Omicini*, Alessandro Ricci*, Mirko Viroli*, Cristiano Castelfranchi†, Luca Tummolini†
*DEIS, Alma Mater Studiorum, Università di Bologna
via Venezia 52, 47023 Cesena, Italy
Email: andrea.omicini@unibo.it, mirko.viroli@unibo.it, aricci@deis.unibo.it
†Institute of Cognitive Sciences and Technologies, CNR
viale Marx 15, 00137 Roma, Italy
Email: c.castelfranchi@istc.cnr.it, tummoli@ip.rm.cnr.it

***Abstract*— In this seminal paper, we sketch a general conceptual framework for self-organising systems (SOSs) that encompasses both stigmergy and MAS coordination, and potentially promotes models of self-organisation for MASs where interaction between cognitive agents is mediated by the environment, by means of artifacts provided by the agent infrastructure. Along this line, we first introduce the notions of *Behavioural Implicit Communication* (BIC) as a generalisation of stigmergy, and of *shared environment* (s-env) as a MAS environment promoting forms of observation-based coordination (such as BIC-based ones) that exploit cognitive capabilities of intelligent agents to achieve MAS self-organisation.**

I. Introduction

Self-organisation is typically associated to natural systems, where global coherent behaviour emerges from a multiplicity of local interactions between non-intelligent system components, in absence of global centralised control. For instance, physical systems like molecules of magnetic materials, biological systems like cytoskeletal filaments in cytoplasm of eukaryotic cells [1], social systems like insect societies [2], all exhibit forms of local interaction between very simple system components that result in higher-level forms of organisation, which can be reduced neither to the individual component's behaviour, nor to explicit external control or constraints over system's evolution. Self-organisation is also found in (human) social systems, where it emerges from non-directed local interactions between humans [3]. Robustness, fault-tolerance and adaptability to changes are typical features of those sorts of self-organising systems (SOSs henceforth) that computer scientists and engineers are nowadays trying to capture and bring to computational systems.

By definition, SOSs are those systems that exhibit some forms of global order (organisation, structure, architecture, ...), or direction, that emerge as the result of apparently non-ordered, non-directed local behaviour. Correspondingly, fundamental definitory features of SOSs are the lack of centralised control, and locality of interaction between components.

The very fact that natural SOSs often exhibit global "intelligent" (in a very broad sense) behaviours in spite of their non-intelligent individual components (magnetic particles, cytoskeletal filaments, ants) has led a good deal of the SOS research in computer science to focus on SOSs based of very simple software components. This is the case, for instance, of most of the literature on ant-based systems, trying to capture the principle of self-organisation by mostly focusing on the patterns of interaction between ant-like components, rather than on their inner structure and functioning, as in the case of stigmergy coordination [4].

This has changed in the last few years, with Multi-Agent Systems (MASs henceforth) taking momentum in the SOS field [5]. There, the most typical model for local interaction between components (agents) is based on direct communication: according to [6], self-organising MASs are typically driven by social interaction (communication, negotiation, coordination) among autonomous entities. This is the case, for instance, of the AMAS theory [7], where self-organisation depends on the ability of the agents to be locally "cooperative" – based on their ability to subjectively interpret interactions with other agents and the environment. Also, this corresponds to well-known patterns of self-organisation in human organisations [3].

On the other hand, when interaction among agents is mediated (so indirect, as opposed to direct interaction) by the environment, it typically happens that cognitive abilities of agents are not adequately exploited to the aim of self-organisation. According to [8, page 316], there is

> "a fundamental flaw in many studies of self-organisation: the assumption that the subunits of a self-organised system are dumb"

This is the case, for instance, of stigmergy [9] and swarm intelligence [10] applied to MAS coordination, where no use of agent cognitive capabilities is assumed to achieve self-organisation.

Given such premises, in this seminal paper we assume as our conceptual target those forms of self-organisation which are based on mediated interaction through the environment (á la stigmergy), but where intelligence of components plays a relevant role. So, we first demystify the apparent dichotomy between stigmergy coordination and social communication, showing a larger range of options: interaction between cognitive agents is not always reducible to communication, communication is not always explicit, and stigmergy (once properly defined [11]) does not exhaust the whole range of interaction through the environment. This is achieved by adopting the theory of *Behavioural Implicit Communication* (BIC), which models a wide range of social behaviours, and works as a critical decentralised coordination mechanism which is mainly

M. Baldoni, F. De Paoli, A. Martelli, and A. Omicini (Eds.): *WOA 2004 – Dagli Oggetti agli Agenti, Sistemi Complessi e Agenti Razionali*, Torino, Italy, November 29th – December 1st 2004. Pitagora Editrice Bologna. ISBN 88-371-1533-4. http://woa04.unito.it

responsible for social order in human societies [11]. Such a mechanism is shared with animal societies, where it takes the form of stigmergy (which can then be thought as a BIC sub-category), and in the context of MAS provides a more comprehensive theory for self-organisation based on local interactions mediated by the environment that also covers cognitive agents.

Then, we focus on the environmental properties that enable BIC, and devise out the notion of *shared environment* (s-env) as a MAS environment promoting forms of observation-based coordination (such as BIC-based ones) that exploit cognitive capabilities of intelligent agents to achieve MAS self-organisation. In particular, the environment should support observability of agent's behaviour, and enable awareness of observation, through suitably-designed MAS infrastructures. Along this line, a formal model for MAS encompassing both BIC and s-env is introduced, that works as a model for MAS infrastructures enabling and promoting advanced forms of self-organisation for MAS based on cognitive agents, where agents interact through suitable abstractions provided by the infrastructure.

Some meaningful examples are finally discussed, that show how forms of self-organisation can emerge in MASs based on cognitive agents by exploiting the observability features provided by shared environments, focusing in particular on the BIC approach.

II. Self-Organisation through Behavioural Implicit Communication

A. Interaction, Communication, Observation

In this section we briefly introduce various kind of interaction which can be found in complex systems, remarking in particular the relevance of indirect interaction and implicit communication – based on observation and awareness – as far as coordination and self-organisation activities are concerned.

Forms of indirect interaction are pervasive in complex systems, in particular in systemic contexts where systems take the form of structured societies with an explicit organisation, with some cooperative activities enacted for achieving systemic goals. In such contexts, in order to scale with activity complexity, sorts of *mediating artifacts* are shared and exploited to enable and ease interaction among the components. Mediating artifacts of different kind can be identified easily in human society, designed and exploited to support coordination in social activities, and in particular in the context of cooperative work: examples are blackboards, form sheets, but also protocols and norms. Mediation is well focused by some theories such as Activity Theory [12] and Distributed Cognition, [13] adopted in the context of CSCW and HCI, exploring how to shape the environment in terms of mediating artifacts in order to better support cooperative work among individuals. Stigmergy is another well-known form of indirect interaction, exploiting directly the environment as mediating artifact: individuals interact by exploiting shared environmental structures and mechanisms to store and sense kind of signs (such as pheromones in the case of ant-based systems), and processes transforming them (such as evaporation/aggregation of pheromones) [2].

With respect to interaction, communication adds *intentionality*. A famous claim of the Palo Alto psychotherapy school says that "any behaviour is communication" [14]: more generally, we consider communication as any process involving an intentional transfer of information from an agent X (sender) to an agent Y (receiver), where X is *aimed at* informing Y. Agent X's behaviour has the goal or the function of informing agent Y. Agent X is executing a certain action "in order" to have other agents receiving a message and updating their beliefs or epistemic state. Communication is an intentional or functional notion in the sense that it is always goal oriented such that a behaviour is selected also for its communicative effect[1]. In the context of cognitive MAS – composed by intelligent agents – explicit types of (high level) communication are typically adopted for supporting coordination and self-organisation, mainly exploiting common semantics and ontologies.

However, in complex societies explicit communication is only part of the story: not all kinds of communication exploit codified (and hence rigid) actions. Humans and animals are for instance able to communicate also without a *predefined* conventional language, by observing their normal behaviour and practical actions. More generally, also forms of *implicit* communication play a key role as kind of interaction. Looking to societies of individuals provided with cognitive capabilities (humans, agents, ...), *observation* and *awareness* can be counted among the main basic mechanisms that enable forms of implicit communication, which allows for coordination and autonomous organisation activities. An agent's behaviour could be observed by another agent, and interpreted / used as information by the observing agent; but also, being aware to be observed, an agent could use its behaviour as a means to communicate.

So, our claim here is that implicit communication – based on observation and awareness – can be very effective as basic brick to build flexible coordination and self-organisation in the context of artificial societies, composed by cognitive agents. While we agree with [15] that coordination is a causal process of correlation between agents' actions typically involving an information flow between an agent and its environment, we do not consider always this flow as a process of communication. Consider a case where an hostile agent, whose actions are "observable", is entering a MAS. If another agent becomes aware of his presence, can observe him, should we say that the hostile agent is communicating his position? Or, differently, is the escaping prey communicating to the predator her movements? Also, even if an agent's perception of the action of another

[1] An agent's behaviour can be goal oriented for different reasons. An intentional agent (i.e. a BDI agent) is a goal governed agent (the goal is internally represented) which instantiates a communicative plan to reach the goal that another agent is informed about something. However, also simple reactive agents (i.e. insect-like) can act purposively (hence can communicate) if their behaviour has been shaped by natural or artificial selection, by reinforcement learning or by design (in the interest of the agent itself). In these latter cases the behaviour has the *function* of communicating in the sense that it has been selected *because of* a certain communicative effect.

agent is necessary implemented as information transition from a sender to a receiver, this implementation of interaction should not be necessarily considered as "communication" and the passed information should not be always labelled as a "message". From the external viewpoint of the designer a message passing of this sort is designed in order to inform the agent who is observing. However from the viewpoint of the agent a simple perception is not necessarily communication.

With respect to existing approaches on self-organisation using intelligent agents (such the AMAS approach [7]), we do not adopt direct communication as the main form of interaction, instead we aim at exploring implicit communication as a form of indirect interaction, based on observation and awareness as its basic bricks. With respect to existing approaches based on indirect interaction – such as stigmergy or computational fields [16] – we aim at considering societies composed by individuals with high level cognitive capabilities able to observe and reason about observations and actions.

B. Behavioural Implicit Communication

In cognitive MAS, communication is normally conceived as implemented through specialised actions such as speech acts defined in the FIPA ACL protocol [17]. Such protocols are inspired by natural language or expressive signals where meaning is associated to a specific action by convention.

Here we are interested in the case where the agent is aware of being observed (other agents believe that he is performing a given practical action) and he "intends that" [18] the other are interpreting his action. This sort of communication without a codified action but with a communicative intention is what we intend for behavioural Implicit Communication [11]. What is relevant here is that the agent's execution plan is aimed to achieve a pragmatic goal as usual: i.e. an agent A is collecting trash to put it in a bin (as in [19]).

A general definition for BIC is: the agent (source) is performing a usual practical action α, but he also knows and lets or makes the other agent (addressee) to observe and understand such a behaviour, i.e. to capture some meaning μ from that "message", because this is part of his (motivating or non motivating) goals in performing α. To implicitly communicate, the agent should be able to contextually "use" (or learn to use or evolve to use) the *observed* executed plan also as a sign, the plan is used as a message but it is not shaped, selected, designed to be a message.

An agent B has the same goal but observing the other's action he decides to clean another side of the road. Since the agent A knows that an agent B is observing him, the practical action he is executing can be used *also* as a message to B such as "I am cleaning here". Such a possibility can lead agents to avoid a specific negotiation process for task allocation and can finally evolve in an implicit agreement in what to do.

Three different conditions are necessary to support such a form of communication.

- The first is relative to environmental properties. The "observability" of the practical actions and of their traces is a property of the environment where agents live, one environment can "enable" the visibility of the others while another can "constrain" it, like sunny or foggy days affect our perception. An environment could also enable an agent to make himself observable or on the contrary to hide his presence on purpose.
- The second is related to the capacity of agents to understand and interpret (or to learn an appropriate reaction to) a practical action. A usual practical action can be a message when an agent knows the way others will understand his behaviour. The most basic message will be that the agent is doing the action α. More sophisticated form would imply the ability to derive pragmatic inference from it (what is the goal of doing? What can be implied?).
- The third condition is that the agent should be able to understand (and observe) the effect that his actions has on the others so that he can begin acting in the usual way *also* because the other understand it and react appropriately.

behavioural Implicit Communication is in this sense a parasitical form of communication that exploits a given level of visibility and the capacity of the others to categorise or react to his behaviour.

So, BIC can be considered a generalisation of stigmergy. The need for an environment for a MAS is often associated with the goal of implementing stigmergy as decentralised coordination mechanism. Besides, being *the production of a certain behaviour as a consequence of the effects produced in the local environment by previous behaviour* or *indirect communication through the environment* [4], stigmergy seems very similar to the form of communication we are arguing for.

However these general accepted definitions make the phenomenon too broad. It is too broad because it is unable to distinguish between the communication and the signification processes. As we have seen in 2.1 we do not want to consider the hostile agent's actions or the escaping prey as communicative actions notwithstanding that the effects of their actions elicit and influence the actions of other agents. Besides, every form of communication is mediated by the environment exploiting some environmental channel (i.e. air).

As in BIC, real stigmergic communication does not exploit any *specialised communicative* action but just usual practical actions (i.e. the nest building actions). In fact we consider stigmergy as a subcategory of BIC, being communication via long term *traces*, physical *practical* outcomes, *useful* environment modifications which preserve their practical end but acquire a communicative function. We restrict stigmergy to a special form of BIC where the addressee does not perceive the *behaviour* (during its performance) but perceives other *post-hoc traces* and outcomes of it.

Usually stigmergy is advocated as a coordination mechanisms that can achieve very sophisticated forms of organisation without the need for intelligent behaviour. However there also exist interesting forms of stigmergic communication at the intentional level. Consider a sergeant that – while crossing a mined ground – says to his soldiers: "walk on my prints!".

From that very moment any print is a mere consequence of a step, plus a stigmergic (descriptive "here I put my foot" and prescriptive "put your foot here!") message to the followers.

C. Forms of Observation-based Coordination

Coordination is that additional part or aspect of the activity of an agent specifically devoted to deal and cope with the dynamic environmental interferences, either positive or negative, i.e. with opportunities and dangers/obstacles [20]. Coordination can either be non social as when an agent coordinate with a moving object. The conceptual framework introduced so far makes it possible to frame some basic forms of coordination in terms of observation and awareness, which will be the key for enabling self-organisation of systems:

- *Unilateral* — X intends to coordinate with Y by observing Y's actions.
- *Bilateral* — In this case we have the unilateral form of coordination for both agents, so: X intends to coordinate with Y by observing$Y$'s actions, and viceversa: Y intends to coordinate with X by observing X's actions.
- *Unilateral-AW* — In this case we have a unilateral form of coordination, but with a first form of awareness: X intends to coordinate with Y by observing Y's actions, and Y is aware of it (i.e. knows to be observed).
- *Reciprocal* — In this case the we have both a bilateral form of observation based coordination and awareness by both the agents: X intends to coordinate with Y by observing Y's actions, Y is aware of it, Y intends to coordinate with X by observing X's actions and X is aware of it.
- *Mutual* — This case extends the reciprocal form by introducing the explicit awareness of each other intention to coordinate: X intends to coordinate with Y by observing Y's actions, Y is aware of it, Y intends to coordinate with X by observing X's actions, X is aware of it, and X is aware of Y intention to coordinate and Y is aware of X intention to coordinate.

behavioural implicit communication is necessary for mutual coordination while it is possible and useful in the other kinds of observation-based self-organisation.

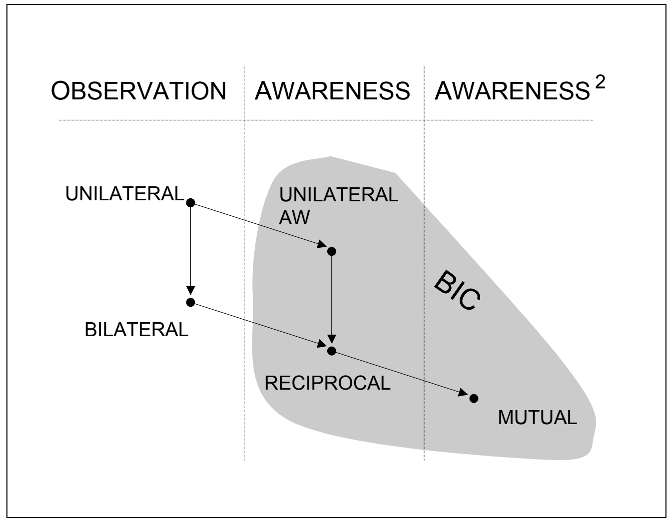

Fig. 1. Forms of coordination in relation to observation capability and awareness. Squared awareness means awareness of awareness. BIC appears with awareness, but is fully exploited when considering mutual coordination.

D. The Role of behavioural Implicit Communication in Dynamic Social Order

Global social order cannot be mainly created and maintained by explicit and formal norms, supported only by a centralised control, formal monitoring, reporting and surveillance protocols. Social order needs to be self-organising, spontaneous and informal, with spontaneous and decentralised forms of control and of sanction [21]. In this respect, BIC plays a crucial role. Sanctions like the act of excluding or avoiding cheaters are messages; the same for the act of exiting (quitting commitments). The act of monitoring the others' behaviour is a message for social order; the act of fulfilling commitments, obeying to norms, are all implicitly communication acts. Behavioural Implicit Communication has a privileged role also for establishing commitments, locally negotiating rules, monitoring correct behaviours, enforcing laws, letting spontaneously emerge conventions and rules of behaviours.

Accordingly, a self-organising society of artificial agents should be able to let emerge a sort of 'social contract' analogous to the one we find in human societies. Such a social contract will first be established mainly by implicit communication, then tacitly signed and renewed.

In what follows, we give some examples of this crucial role.

- *Imitation for rule propagation* — One of the main functions of imitation (i.e., repeating the observed behaviour of Y – the model) is for achieving a basic form of implicit communication. The condition is that Y (the model) can observe (be informed about) the imitative behaviour of X. By simply imitating the peer, the agent can propagate a tacit message like "I use the same behaviour as you, I accept (and spread) it as convention; I conform to it". This BIC use of imitation is probably the first form of mimetic propagation through communication and plays a key role in convention establishment. X interprets the fact that Y repeats its innovation as a confirmation of its validity (good solution) and as an agreement about doing so. Then, X will expect that Y will understand again its behaviour next time, and that Y will use again and again it, at least in the same context and interaction.
- *The fulfilment of social commitments* — Differently from the acts of conforming to already existing norms, agents (when observable) can implicitly communicate the fulfilment of their social commitments. A conforming behaviour is a form of demonstrative act primarily intended to show that one have done the expected action. Thus, the performance of the act is also aimed at informing that it has been performed.

  This is especially important when the expectation of X's act is based on obligations impinging on X, and Y is monitoring X's non-violation of his duty. Either X is respecting a prohibition, or executing an order, or keeping a promise. A social-commitment of X to Y of doing the act, in order to be really (socially) fulfilled, requires not only that agent X performs the promised action, but also that the agent Y knows this. Thus, when X is performing the act in order to keep his promise and fulfil his commitment to Y, he also intends that Y knows this. Even in absence of explicit and specific messages, any act of social commitment fulfilment can be an implicit communication act about that fulfilment.

A second order meaning of the conforming act can also be: "I'm a respectful guy; I'm obedient; I'm trustworthy", but this inferential meaning is reached trough the first meaning "I'm respecting, obeying, keeping promises". This second order meanings can circulate and boost the reputation process that is a key informal sanction system for dynamic social order [22].

- *Local reissuing of norms* — Moreover, one of the functions of norm obedience is the confirmation of the norm itself, of the normative authority of the group, and of conformity in general. Consequently, one of the functions of norm obeying behaviours is that of informing the others about norm obedience. At least at the functional level, X's behaviour is implicit behavioural communication.

  Frequently, X either is aware of this function and collaborates on this, thus he intends to inform the others about his respect of norms, or he is worrying about social monitoring and sanctions or seeking for social approval, and he wants the others see and realise that he is obeying the norms. In both cases, his conforming behaviour is also an intentional behavioural/implicit communication to the others.

  At the collective level, when an agent respects a norm, he pays some costs for the commons and immediately moves from the mental attitude of norm addressee (which recognised and acknowledge the norm and its authority, and decided to conform to it) to the mental set of the norm issuer and controller [23]: he wants the others to respect the norm, pay their own costs and contribution to the commons.

III. A BIC-oriented Shared Environment for Self-Organisation

So, to promote advanced forms of self-organisation in MAS featuring cognitive agents, MAS environment should be shaped so as to allow for observability and awareness of agents behaviour.

Generally speaking, agents that live in a *common environment* (*c-env*) are agents whose actions and goals interfere (positively or negatively). In a pure *c-env*, agent actions and their traces are state transitions which can ease or hamper the individual agents' goals. An example is a ground that is common for different insect species but where no interspecies communication is possible. Agents can observe just the state of the environment, and then act on that basis, achieving a given self-organisation, still with no access to the actions of their peers. Even a trace is seen as part of the environment and not as a product of other agents. So, a generic property of a *c-env* is that it provides agents with the means to keep track of its state and possibly affect it.

As far as observation-based self-organisation is concerned, we here propose a stronger notion of environment, called *shared environment* (*s-env*). This is a particular case of a *c-env* that enables (i) different forms of observability of each other action executions, as well as (ii) awareness of such observability, thus supporting unilateral, bilateral, reciprocal, and mutual coordination.

A. Observability in Shared Environments

Each *s-env* is defined by the level of observability that it can afford. The level of observability is the possibility for each agent to observe another agent behaviour, namely, to be informed when another agent executes a given action. For instance, the most general kind of *s-env* can be defined by the fact that each agent *can* observe the execution of all the actions of all others agents. A prototypical model of this sort of environment is the central 'square' of a town. Other levels of observability may limit the ability of agents to observe given actions of other agents – e.g. considering sort of invisible actions – or to observe only given agents and not others – e.g. considering obstacles preventing observation.

The level of observability of an *s-env* is easily understood by a *power* relation $Pow : A \times A \times Act$, where A is the set of agents – ranged over by meta-variables x, y, and z – and Act is the set of usual practical actions which may be subject of observation through the *s-env* – ranged over by meta-variables α and β. When $\langle x, y, \alpha \rangle \in Pow$, also written $Pow(x, y, \alpha)$, it means that action $\alpha \in Act$ executed by agent y is observable by agent x through the *s-env*.[2] This means that in that *s-env*, it is structurally possible for x to observe the executions of action α by y. We naturally say that x has the role of observer agent, y that of observed agent, α that of observed action. We extend the notation for power relation using sets of agents or actions, e.g. writing $Pow(x, B, \alpha)$ with $B \subseteq A$ for $Pow(x, y, \alpha)$ holding for all $y \in B$, or $Pow(x, y, Act)$ in place of $Pow(x, y, \alpha)$ for all $\alpha \in Act$.

Pow relation can be then conceived as specifying the rules defining the set of 'opportunities and constraints' that afford and shape agents' observability within the environment. A specific rule is an opportunity or a constraint *for a specific agent* and in particular it is only relative to the agent's active goals while interacting with that environment.

Whereas relation Pow is introduced to statically describe the set of opportunities and constraints related to agents' observability, an *observation* relation Obs (a subset of Pow) has to be introduced to characterise the state of the s-env at a given time. When $Obs(x, y, \alpha)$ holds, it means that agent x

[2] Observability of an action should be intended here in its most general acceptation, that is, accounting for all the properties that need to be observed – so, not only the executing agent, but also time of execution, information possibly carried along, and so on.

is actually observing executions of action α by agent y. That is, $Obs(x, y, \alpha)$ means that an execution of action α by agent y will be perceived by x. Hence, notice that we differentiate between the potential ability to observe, which is a typical property of the environment where the agents live in, and the actual observability, which might be driven by the explicit motivation of agents. Indeed, since $Obs \subseteq Pow$, observation is constrained by the level of observability featured by the *s-env*.

The meaning of the observation relation can be understood by taking into account the agent's viewpoint over observation. We first introduce the concept of agent *epistemic state (ES)*, representing the beliefs the agent has because of his observation role. The ES of an agent x includes its *environmental knowledge* about observation, which is then given by information *(i)* on the agents he is observing, *(ii)* on the agents that are observing him, and *(iii)* on the action executions that he is observing.

The first two kinds of knowledge can be addressed by supposing the agent may, at a given time, have some knowledge about the current state of relation Obs. In particular, write $B_z obs(x, y, \alpha)$ for agent z believing that x is observing, from that time on, executions of action α performed by z. On the other hand, to represent the third kind of knowledge, we write $B_z(done(y, \alpha))$, meaning that agent z believes that y has executed action α.[3]

B. Epistemic Actions

The epistemic state of an agent evolves through *epistemic actions*, which are actions aimed at acquiring knowledge from the environment [25]. Such an aim is expressed as an agent intention: accordingly, we also define the concept of *motivational state* (MS) of an agent, which includes all the intentions an agent has at a given time. Then, an epistemic action is fired by an agent intention, by which the *s-env* reacts updating its state as well as the epistemic state of the agent. So, we have different kinds of epistemic actions, each fired by a different motivation: they are used e.g. to know who is observing who, to have an agent observing another, to avoid an agent observing another, and so on.

A first case of epistemic action is used by the agent which is willing to know whether he is observing another agent, whether another agent is observing him, or generally, whether an agent x is observing actions α of an agent y. So, suppose the MS of z includes intention $I_z check(x, y, \alpha)$, which means that agent z intends to know whether x observes executions of α by y. Then, eventually an epistemic action is executed by which the ES of agent z will include the belief about whether $Obs(x, y, \alpha)$ holds or not.

Similarly, an agent may have the intention $I_x obs(x, y, \alpha)$ in exploiting the observability power of the environment to observe y's actions α. When such an intention appears in the MS of agent x, the s-env conceptually intercepts it and enacts the corresponding observations. This means that (i) the s-env adds $B_x obs(x, y, \alpha)$ to the agent's epistemic state (agent x knows that he is observing actions by agent y), and (ii) relation Obs is added the rule $Obs(x, y, \alpha)$ (the s-env makes agent x observing actions α by agent y). In other words, we can think that the appearance of an intention in the motivation state of the agent causes the execution of an epistemic action toward the environment, enabling agent observations.

Similarly, an agent may want to stop observing actions. When the intention $I_x drop(x, y, \alpha)$ appears in the agent motivational state, the effects of $obs(x, y, \alpha)$ are reversed.

Now we are ready to link the MS state of the agent, Obs rules and the ES state of the agent. According to the semantics of the actions, the execution of an action α by agent y (written $done(y, \alpha)$) causes the creation of a new belief $B_x done(y, \alpha)$ in the epistemic state of all the agents x of the environment such that $Obs(x, y, \alpha)$ holds.

C. Formal Model

To make our arguments more precise we introduce a formal framework to describe the notions of ES, MS, epistemic actions, and observation in a precise way, which is meant to serve as an actual design for an infrastructure providing a *s-env*. In particular, we provide a syntax and an operational semantics for modelling MAS according to the conceptual framework defined in previous sections.

Throughout this model, composition operator $||$ is assumed to be commutative, associative, to absorb the empty configuration 0, and to consume multiple copies of the same element – that is, $x \,||\, x \equiv x$. Accordingly, any grammar definition of the kind

$$X ::= 0 \mid x_1 \mid \ \dots \ \mid x_n \mid X \,||\, X$$

defines elements of the syntactic category X as compositions (without repetitions) of terms $x_1, \dots, x_n$. Given one such composition X, we write $x_j \in X$ and $x_j \notin X$ with the obvious meaning. The syntax of MAS configurations is reported in Figure 2.

Metavariable S ranges over configurations of the MAS, which at our abstraction level are simple compositions of agent configurations (ES and MS) and environment configurations (Pow and Obs). Environment configurations are composition of terms, each denoting either the power of agent x to observe action α executed by agent y ($Pow(x, y, \alpha)$), or the fact that the environment is making x observe actions α executed by agent y ($Obs(x, y, \alpha)$). Agent configurations are compositions of mental properties, namely beliefs (B) and intentions (I) qualified by the agent x, and about a formula ϕ. As described above, these properties are used to represent the ES and MS of agent x, namely its knowledge and motivations. Notice that we model a MAS configuration as a composition of both agent and environment properties without a neat separation: in fact, at our level of abstraction such a distinction is not necessary, for epistemic actions involving both kinds of properties in a uniform way.

[3]The syntax we introduced clearly reminds standard modal logics for beliefs as in [24], however, it is not our goal here to introduce any logics for agent reasoning. This is why we still refer to the weaker notion of epistemic state instead of beliefs state – and motivational state instead of intentional state as described below.

$$
\begin{array}{lcll}
S & ::= & 0 \mid A \mid E \mid S \,||\, S & \text{MAS configuration} \\[6pt]
E & ::= & 0 & \text{environment configuration} \\
 & \mid & Pow(x,y,\alpha) & x \text{ has the power to observe } y\text{'s } \alpha \\
 & \mid & Obs(x,y,\alpha) & x \text{ is observing } y\text{'s } \alpha \\
 & \mid & E \,||\, E & \text{composition} \\[6pt]
A & ::= & 0 & \text{agent configuration} \\
 & \mid & B_x\phi & \text{belief of } x \\
 & \mid & I_x\phi & \text{intention of } x \\
 & \mid & A \,||\, A & \text{composition} \\[6pt]
\phi & ::= & & \text{formulas} \\
 & & obs(x,y,\alpha) & x \text{ is observing } y\text{'s } \alpha \\
 & \mid & coord(x,y,\alpha) & x \text{ coordinates with } y \text{ through } \alpha \\
 & \mid & check(x,y,\alpha) & \text{check whether } x \text{ is observing } y\text{'s } \alpha \\
 & \mid & drop(x,y,\alpha) & \text{prevent } x \text{ from observing } y\text{'s } \alpha \\
 & \mid & done(x,\alpha) & x \text{ executes actions } \alpha \\
 & \mid & \neg\phi \mid I_x\phi \mid B_x\phi & \text{structured formulas}
\end{array}
$$

Fig. 2. Syntax of MAS configurations.

Elements ϕ are formulas which can be believed and/or intended by an agent. Atomic formulas are: *(i)* $obs(x, y, \alpha)$, used to express that x is observing executions of α by y, *(ii)* $coord(x, y, \alpha)$, used to express that x coordinates its behaviour with y by observing executions of α, *(iii)* $check(x, y, \alpha)$, used to check if x is observing executions of α by y, *(iv)* $drop(x, y, \alpha)$, used to prevent x from observing executions of α by y, and *(v)* $done(x, \alpha)$, used to express that x executes/has executed α. Moreover, formulas can be structured ones: $\neg\phi$ expresses negation of ϕ, $I_x\phi$ and $B_x\phi$ that agent x intends/believe ϕ. A number of assumptions on such formulas are clearly to be made as usual, e.g. that $\neg\neg\phi \equiv \phi$ or $B_x\phi \equiv B_xB_x\phi$. This amounts to define a logics for beliefs and intentions: however, this aspect can be treated in a fairly standard way, therefore its details are not reported for they play no significant role in this paper – they are more about agent internal architecture rather than agent interaction through the environment.

On top of this syntax for MAS configurations, we introduce an operational semantics, describing what are the allowed evolutions of such configurations. This describes the dynamic aspects of our model, providing details on preconditions and effects to epistemic actions and observation in general. As usual [26], operational semantics is defined by a set of rewrite rules, reported in Figure 3. Each rule defines a MAS configuration to be rewritten as interaction of the agent with the *s-env* occurs: the left-hand side reports preconditions, the right-hand effects, and the above part (when present) further preconditions for the applicability of the rule.

Rule [CHECK] says that if agent z intends to check/know if x is observing y's action α, and this is the case, then such an intention will be turned into a positive belief. Dually, rule [N-CHECK] deals with the case where this is not the case ($Obs(x, y, \alpha)$ does not occur in the system configuration), so that z will believe that $obs(x, y, \alpha)$ does not hold.

Rule [DROP-Y] says that if agent z knows that x is observing y's action α (which is the case) and wants to stop him, term $Obs(x, y, \alpha)$ is dropped from the environment and z's belief is updated correspondingly. By rule [DROP-N] we deal with the similar case, but supposing the agent had a wrong belief (x was not actually observing y's actions α), which is dealt with trivially.

Rule [ASK] is about agent z willing that x observes y's actions α: if this is allowed ($Pow(x, y, \alpha)$), x's beliefs will be updated along with the environment state.

Rule [OBS-R] and [OBS-F] recursively define how the environment broadcasts information about an action to all the observers. When agent x wants to execute α, each observer y (rule [OBS-R]) will be recursively added the belief $B_y done(x, \alpha)$: when none needs to be managed, x intention can simply become a fact, that is, he will believe the action to be executed ([OBS-F]).

The final, trivial rule [AGENT] is used to represent the fact that at any given time some agent configuration can change autonomously, thus modelling any belief revision or intention scheduling.

Notice that formulas $B_z coord(x, y, \alpha)$ or $I_z coord(x, y, \alpha)$ never appear in this semantics. This is because the fact that an agent coordinates its behaviour with another is not an aspect influencing/influenced by the environment: it is rather a mental property characterising the forms of observation-based coordination an agent participates to thanks to the *s-env* support.

D. Formalising Observation-based Coordination

We put to test our formal framework showing how the forms of coordination devised in Subsection II-C can be represented through our syntax.

$$\frac{Obs(x,y,\alpha) \in S}{I_z check(x,y,\alpha) \,||\, S \rightarrow B_z obs(x,y,\alpha) \,||\, S} \quad \text{[CHECK]}$$

$$\frac{Obs(x,y,\alpha) \notin S}{I_z check(x,y,\alpha) \,||\, S \rightarrow B_z \neg obs(x,y,\alpha) \,||\, S} \quad \text{[N-CHECK]}$$

$$\frac{-}{I_z drop(x,y,\alpha) \,||\, B_z obs(x,y,\alpha) \,||\, Obs(x,y,\alpha) \,||\, S \rightarrow B_z \neg obs(x,y,\alpha) \,||\, S} \quad \text{[DROP-Y]}$$

$$\frac{Obs(x,y,\alpha) \notin S}{I_z drop(x,y,\alpha) \,||\, B_z obs(x,y,\alpha) \,||\, S \rightarrow B_z \neg obs(x,y,\alpha) \,||\, S} \quad \text{[DROP-N]}$$

$$\frac{-}{I_z obs(x,y,\alpha) \,||\, Pow(x,y,\alpha) \,||\, S \rightarrow B_z obs(x,y,\alpha) \,||\, Pow(x,y,\alpha) \,||\, Obs(x,y,\alpha) \,||\, S} \quad \text{[ASK]}$$

$$\frac{I_x done(x,\alpha) \,||\, S \rightarrow I_x done(x,\alpha) \,||\, S'}{I_x done(x,\alpha) \,||\, Obs(y,x,\alpha) \,||\, S \rightarrow I_x done(x,\alpha) \,||\, Obs(y,x,\alpha) \,||\, B_y done(x,\alpha) \,||\, S'} \quad \text{[OBS-R]}$$

$$\frac{Obs(y,x,\alpha) \notin S}{I_x done(x,\alpha) \,||\, S \rightarrow B_x done(x,\alpha) \,||\, S} \quad \text{[OBS-F]}$$

$$\frac{-}{A \,||\, S \rightarrow A' \,||\, S} \quad \text{[AGENT]}$$

Fig. 3. Operational Semantics of Agent Configurations.

Given two agents x and y, an action α, and the system configuration S we introduce the following predicates:

- Unilateral

  $$Uni(x,y,\alpha,S) \triangleq Obs(x,y,\alpha) \in S \;\wedge\; I_x coord(x,y,\alpha)$$

  Agent x is in unilateral coordination with y (in system S, through action α), if he is observing y's actions α and he intends to coordinate with y through such actions.
- Unilateral with Awareness

  $$UniAW(x,y,\alpha,S) \triangleq Uni(x,y,\alpha,S) \;\wedge\; B_y obs(x,y,\alpha) \in S$$

  The form of coordination is unilateral with awareness if x is in unilateral coordination with y and if y knows to be observed by x.
- Bilateral

  $$Bi(x,y,\alpha,S) \triangleq Uni(x,y,\alpha,S) \;\wedge\; Uni(y,x,\alpha,S)$$

  x and y are in bilateral coordination if they are both in unilateral coordination with each other.
- Reciprocal

  $$Rec(x,y,\alpha,S) \triangleq UniAW(x,y,\alpha,S) \;\wedge\; UniAW(y,x,\alpha,S)$$

  x and y are in reciprocal coordination if they are both in unilateral coordination with awareness.
- Mutual

  $$Mut(x,y,\alpha,S) \triangleq Rec(x,y,\alpha,S) \;\wedge\; B_x I_y coord(y,x,\alpha) \;\wedge\; B_y I_x coord(x,y,\alpha)$$

  Finally, x and y are in mutual coordination if they are in reciprocal coordination and, moreover, they both know that the other agent intends to coordinate through the observed action α.

IV. Conclusions

In this paper we focused on some properties of MAS infrastructures for cognitive agents supporting forms of self-organisation, based on the BIC theory. Even though not dealing with internal aspects of agents, we consider agents provided with some cognitive capabilities, differently from current environment-based approach in self-organisation, typically based on reactive agents (e.g. ants).

MASs built on top of a BIC-oriented infrastructure exhibit the basic enabling principles which typically characterise self-organisation:

- *Local interaction* — In the framework there is an explicit notion of locality of interaction: agent observability and awareness are related to a notion of locality that is dynamic, depending on the adopted topology, which is defined by the infrastructure and can be changed over

time. The enacting of $Pow(x, y, \alpha)$ rules by MAS infrastructure implicitly defines such a topology in terms of what actions can be observed by whom at any time.
- *Decentralised control* — Control is decentralised and encapsulated in cognitive agents, which exhibits an autonomous behaviour with respect to the environment.
- *Emergent patterns* — Patterns of MAS self-organisation emerge from agent interacting through a suitably shaped environment, by exploiting observation capabilities provided by the infrastructure.

Besides these basic principles, other interesting aspects that are often considered when dealing with self-organising systems can be re-casted in our framework:

- *Individual-based models* — Individual-based models are currently considered the right approach for the quantitative and qualitative study of SOS [26], tracking each individual state and behaviour. The model presented in the paper is indeed individual-based, since a MAS is composed by individual agents with their own cognitive state and behaviour, eventually playing different kinds of roles inside the system.
- *Openness* (in the thermodynamic acceptation) — In order to keep thermodynamic systems self-organised there must be a continuous flow of energy from the environment: our MASs are characterised by an analogous form of openness, since agents are meant to exchange information within the environment – which is outside the system – by means of perceptions and actions.
- *Non-linearity and feedbacks* — Non-linearity and (positive) feedback that typically characterise SOS can be obtained with forms of mutual coordination, realising kind of non-linear chains of observation and awareness.
- *Dissipative structures* — In our framework, infrastructure structure / services exploited by agents for enhancing their observation / awareness capability can play the role of dissipative structures, typically considered in SOS [27] as a key to export entropy out of the system.

Most of complex system scenarios calls for systems with self-organising capabilities but immersed in an environment that can have (social) norms and constraints, typically specified at design time and that enforced at runtime. We think that in order to cope with such (apparently conflicting) aspects, MAS infrastructure can play a key role [28]. On the one side, it can provide mechanisms and abstractions enabling forms of interaction enabling MAS self-organisation – thus promoting system's unpredictability. On the other side, such mechanisms and abstractions can play a regulatory role, by enforcing laws and norms constraining and ruling agent interaction space – thus promoting system's predictability. We believe that our approach will support MAS engineers in finding the most suitable balance between such a dilemma of "global vs. local control" in MASs.

References

[1] F. Nedelec, T. Surrey, and É. Karsenti, "Self-organisation and forces in the microtubule cytoskeleton," *Current Opinion in Cell Biology*, vol. 15, no. 2, pp. 118–124, Feb. 2003.

[2] P.-P. Grassé, "La reconstruction du nid et les coordinations inter-individuelles chez bellicositermes natalensis et cubitermes sp. la theorie de la stigmergie: essai d'interpretation des termites constructeurs," *Insectes Sociaux*, vol. 6, pp. 41–83, 1959.

[3] H. Haken, *Synergetics: An Introduction. Nonequilibrium Phase Transition and Self-Organization in Physics, Chemistry, and Biology*. Springer-Verlag, 1977.

[4] O. Holland and C. Melhuis, "Stigmergy, self-organization, and sorting in collective robotics," *Artificial Life*, vol. 5, no. 2, pp. 173–202, 1999.

[5] G. Di Marzo Serugendo, A. Karageorgos, O. F. Rana, and F. Zambonelli, Eds., *Engineering Self-Organising Systems: Nature-Inspired Approaches to Software Engineering*, ser. LNAI, vol. 2977. Springer-Verlag, May 2004. [Online]. Available: http://www.springer.de/cgi/svcat/search_book.pl?isbn=3-540-21201-9

[6] G. Di Marzo Serugendo, N. Foukia, S. Hassas, A. Karageorgos, S. Kouadri Mostéfaoui, O. F. Rana, M. Ulieru, P. Valckenaers, and C. Van Aart, "Self-organisation: Paradigms and applications," in *Engineering Self-Organising Systems*, ser. LNAI, G. Di Marzo Serugendo, A. Karageorgos, O. F. Rana, and F. Zambonelli, Eds. Springer-Verlag, May 2004, vol. 2977, pp. 1–19. [Online]. Available: http://www.springer.de/cgi/svcat/search_book.pl?isbn=3-540-21201-9

[7] D. Capera, M.-P. Gleizes, and P. Glize, "Self-organizing agents for mechanical design," in *Engineering Self-Organising Systems*, ser. LNAI, G. Di Marzo Serugendo, A. Karageorgos, O. F. Rana, and F. Zambonelli, Eds. Springer-Verlag, May 2004, vol. 2977, pp. 169–185. [Online]. Available: http://www.springer.de/cgi/svcat/search_book.pl?isbn=3-540-21201-9

[8] T. D. Seeley, "When is self-organization used in biological systems?" *Biological Bulletin*, vol. 202, pp. 314–318, 2002.

[9] Hadeli, P. Valckenaers, C. Zamfirescu, H. Van Brussel, B. Saint Germain, T. Hoelvoet, and E. Steegmans, "Self-organising in multi-agent coordination and control using stigmergy," in *Engineering Self-Organising Systems*, ser. LNAI, G. Di Marzo Serugendo, A. Karageorgos, O. F. Rana, and F. Zambonelli, Eds. Springer-Verlag, May 2004, vol. 2977, pp. 105–123. [Online]. Available: http://www.springer.de/cgi/svcat/search_book.pl?isbn=3-540-21201-9

[10] R. Tolksdorf and R. Menezes, "Using swarm intelligence in Linda," in *Engineering Societies in the Agents World IV*, ser. LNAI, A. Omicini, P. Petta, and J. Pitt, Eds. Springer-Verlag, June 2004, vol. 3071, pp. 49–65, 4th International Workshop (ESAW 2003), London, UK, 29–31 Oct. 2003. Revised Selected and Invited Papers. [Online]. Available: http://www.springer.de/cgi/svcat/search_book.pl?isbn=3-540-22231-6

[11] C. Castelfranchi, "When doing is saying – the theory of behavioral implicit communication," 2003, draft. [Online]. Available: http://www.istc.cnr.it/doc/62a_716p_WhenDoingIsSaying.rtf

[12] B. Nardi, Ed., *Context and Consciousness: Activity Theory and Human Computer Interaction*. Cambridge, MA: MIT Press, 1996.

[13] E. Hutchins, *Cognition in the Wild*. Cambridge, MA: MIT Press, 1995.

[14] P. Watzlavich, J. Beavin Bavelas, and D. D. Jackson, *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies, and Paradoxes*. New York: W.W. Norton & Co., 1967.

[15] H. V. D. Parunak, S. Brueckner, M. Fleischer, and J. Odell, "A design taxonomy of multi-agent interactions," in *Agent-Oriented Software Engineering IV*, ser. LNCS, P. Giorgini, J. Müller, and J. Odell, Eds. Springer-Verlag, 2004, pp. 123–137, 4th International Workshop (AOSE 2003), Melbourne, Australia, 15 July 2003, Revised Papers.

[16] M. Mamei and F. Zambonelli, "Self-organization in multi-agents systems: A middelware approach," in *Engineering Self-Organising Systems*, ser. LNAI, G. Di Marzo Serugendo, A. Karageorgos, O. F. Rana, and F. Zambonelli, Eds. Springer-Verlag, May 2004, vol. 2977, pp. 233–248. [Online]. Available: http://www.springer.de/cgi/svcat/search_book.pl?isbn=3-540-21201-9

[17] FIPA, *FIPA Communicative Act Library Specification*, 2000, `http://www.fipa.org`.

[18] B. J. Grosz and S. Kraus, "Collaborative plans for complex group action," *Artificial Intelligence*, vol. 86, pp. 269–357, 1996.

[19] A. S. Rao, "A unified view of plans as recipes," in *Contemporary Action Theory*, G. Hölmstrom-Hintikka and R. Tuomela, Eds. Kluwer Academic Publishers, 1997, vol. 2: Social Action.

[20] C. Castelfranchi, "Modelling social action for AI agents," *Artificial Intelligence*, vol. 103, pp. 157–182, 1998.

[21] ——, "Engineering social order," in *Engineering Societies in the Agents World*, ser. LNAI, vol. 1972. Springer-Verlag, Dec. 2000, pp. 1–18, 1st

International Workshop (ESAW'00), Berlin (Germany), 21 Aug. 2000, Revised Papers.

[22] R. Conte and M. Paolucci, *Reputation in Artificial Societies. Social Beliefs for Social Order*. Boston: Kluwer Academic Publisher, 2002.

[23] R. Conte and C. Castelfranchi, *Cognitive and Social Action*. London: University College of London Press, 1995.

[24] M. D. Sadek, "A study in the logic of intention," in *3rd Conference on Principles of Knowledge Representation and Reasoning*, Cambridge, MA, USA, 1992, pp. 462–473.

[25] C. Castelfranchi and E. Lorini, "Cognitive anatomy and functions of expectations," in *Cognitive Modeling of Agents and Multi-Agent Interactions*, 2003, workshop at IJCAI 2003. Proceedings.

[26] G. Plotkin, "A structural approach to operational semantics," Department of Computer Science, AArhus University, Denmark, Tech. Rep. DAIMI FN-19, 1991.

[27] G. Nicolis and I. Prigogine, *Exploring Complexity: An Introduction*. W.H. Freeman & Co., 1989.

[28] A. Omicini, A. Ricci, M. Viroli, C. Castelfranchi, and L. Tummolini, "Coordination artifacts: Environment-based coordination for intelligent agents," in *3rd international Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004)*, N. R. Jennings, C. Sierra, L. Sonenberg, and M. Tambe, Eds., vol. 1. New York, USA: ACM, 19–23 July 2004, pp. 286–293. [Online]. Available: http://portal.acm.org/citation.cfm?id=1018409.1018752