The role of the environment in agreement technologies

**Estefanía Argente · Olivier Boissier · Carlos Carrascosa ·
Nicoletta Fornara · Peter McBurney · Pablo Noriega · Alessandro Ricci ·
Jordi Sabater-Mir · Michael Ignaz Schumacher · Charalampos Tampitsikas ·
Kuldar Taveter · Giuseppe Vizzari · George Vouros**

Abstract The notion of Multi-Agent System environment is currently considered as a mediating entity, functioning as enabler but possibly also as a manager and constrainer of agent actions, perceptions, and interactions. In this paper, we analyze how the environment could be a first class abstraction to support the building, the development and the management of Agreements in decentralized and open systems between autonomous agents. To this aim we

E. Argente
Departamento Sistemas Informáticos y Computación, Universitat Politècnica de València, Valencia, Spain
e-mail: eargente@dsic.upv.es

O. Boissier (✉)
LSTI/Institut Henri Fayol, ENS Mines Saint-Etienne, Saint-Etienne, France
e-mail: olivier.boissier@emse.fr

C. Carrascosa
Departamento Sistemas Informáticos y Computación, Universitat Politècnica de València, Valencia, Spain
e-mail: carrasco@dsic.upv.es

N. Fornara
Faculty of Communication Sciences, Università della Svizzera italiana, Lugano, Switzerland
e-mail: nicoletta.fornara@usi.ch

P. McBurney
Department of Informatics, King's College London, London, UK
e-mail: peter.mcburney@kcl.ac.uk

P. Noriega · J. Sabater-Mir
Artificial Intelligence Research Institute (IIIA-CSIC), Spanish National Research Council,
Barcelona, Spain
e-mail: pablo@iiia.csic.es

J. Sabater-Mir
e-mail: jsabater@iiia.csic.es

A. Ricci
DISI, University of Bologna, Cesena, Italy

analyze the synergies between the environment and the foundational dimensions of agreement technologies such as semantics, norms, organizations, argumentation & negotiation, trust.

Keywords Environment · Semantic · Norm · Organization · Argumentation · Negotiation · Trust · Agreement technologies

1 Introduction

In Agent-Oriented Software Engineering, the notion of Multi-Agent System (MAS) *environment* has gained a key role, becoming a mediating entity, functioning as enabler but possibly also as a manager and constrainer of the agent actions, perceptions, and interactions[1] while addressing the requirements of openness and scalability. According to such a perspective, the environment is not a merely passive source of the agent perceptions and target of the agent actions which is, actually, the dominant perspective in agency, but a *first-class abstraction* that can be suitably designed to encapsulate some fundamental functionalities and services, such as coordination and organization, besides agent mobility, communications, security, etc. (Ricci et al. 2011).

The analysis done in this paper, shows that the environment dimension intersects with all the dimensions that can be addressed to define an *agreement* between autonomous agents, that is, all the different *Agreement Technologies*[2] (AT) giving support to the building, development and management of agreements in decentralized and open systems between autonomous agents. Those dimensions are the ones related to the development of technologies dealing with: Semantics, Norms, Organizations, Argumentation & Negotiation, and Trust. The objective of this contribution is then to provide a comprehensive analysis and discussion about the role that the environment could play for the engineering of Agreement Technologies, along these dimensions that define the agreement technology conceptual space. Besides identifying the points and features that could concern the individual dimensions alone, the aim is to finally provide a conceptual framework which links together such points across the

---

[1] see (Weyns et al. 2007) for comprehensive surveys.
[2] Agreement Technologies refer to computer systems in which autonomous software agents negotiate with each other, typically on behalf of humans, in order to come to mutually acceptable agreements.

M. I. Schumacher · C. Tampitsikas
Institute of Business Information Systems, University of Applied Sciences Western Switzerland, Sierre, Switzerland
e-mail: michael.schumacher@hevs.ch

C. Tampitsikas
e-mail: charalampos.tampitsikas@hevs.ch

K. Taveter
Department of Informatics, Tallinn University of Technology, Tallinn, Estonia
e-mail: kuldar.taveter@ttu.ee

G. Vizzari
Complex Systems and Artificial Intelligence research center, University of Milano-Bicocca, Milano, Italy
e-mail: giuseppe.vizzari@disco.unimib.it

G. Vouros
Department of Digital Systems, University of Piraeus, Piraeus, Greece
e-mail: georgev@aegean.gr

dimensions and gives an answer to two basic background questions: *(i)* how the design and development of agreement agent-based technologies could be improved by introducing the environment as a first-class abstraction besides agents and direct communication models; *(ii)* how the environment can facilitate reaching agreements with respect to the AT dimensions (semantics, trust, etc).

The paper is organized as follows. In Sect. 2 we provide a global definition of environment as it is considered in the multi-agent literature in order to define the framework in which the following sections place their contribution. We then draw some motivations from the analysis of different applications in the context of Agreement Technologies (Sect. 3). The following sections then analyze and describe the relations between the environment and semantics (Sect. 4), and norms (Sect. 5), and organizations (Sect. 6), and argumentation & negotiation (Sect. 7), and trust (Sect. 8).

We finish the paper by discussing the different features and challenges facing the use of the environment for agreement technologies.

2 Environment definition

The notion of *environment* is a primary concept in agent and multi-agent systems. It is the computational or physical place where agents are situated. It provides the basic ground for defining the notions of agent perception, action and then interaction. Fundamental features of the agent abstraction are directly or indirectly related to the environment: reactivity is an obvious example, but also pro-activeness, referring to the notion of *goal*, typically defined in terms of states of the world that one or more agents aim to bring about.

Besides being considered a mere source of the agents percepts and target of the agent actions, the environment can be framed as a *first-class abstraction* in Agent Oriented Software Engineering (Weyns et al. 2007; Odell et al. 2003). That is to say that it is a part of the MAS, that can be designed and programmed to uniformly and effectively encapsulate and modularize functionalities and services out of the agents (Weyns and Parunak 2007; Weyns et al. 2005; Ricci et al. 2011).

In that view, the responsibilities and functionalities of environments can be summarized by the following three different levels of support, identified in Weyns et al. (2007) (see Fig. 1):

- a *basic level*, where the environment is exploited to simply enable the agents to access the *deployment context*, i.e. the external hardware/software resources which the MAS interacts with (sensors and actuators, a printer, a network, a database, a Web service, etc.);
- an *abstraction level*, exploiting an environment abstraction layer to bridge the conceptual gap between the agent abstraction and the low level details of the deployment context, hiding such low level aspects to the agent programmer;
- an *interaction-mediation level*, where the environment is exploited to both regulate the access to shared resources, and mediate the interaction between the agents. It becomes then an enabler and facilitator of agent coordination, organization, cooperation.

Each of these levels represents a different degree of functionality that the agents can use to achieve their goals.

In this perspective, the environment becomes both a *first-class abstraction* for the modelers / designers / developers and a *first-class entity* for the agents. In the former view, it can be exploited at design/development time. It is then used to modularize, encapsulate, structure and compose functionalities and services, and possibly, to define the software architecture
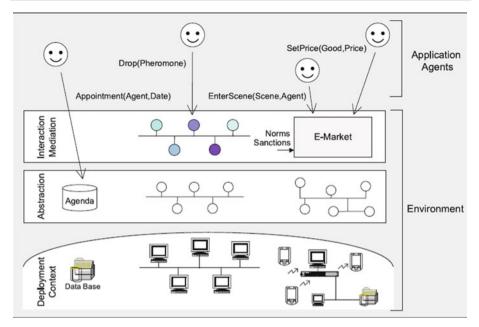
Fig. 1 Environment support levels, as depicted in Weyns et al. (2007)

of the MAS. To this purpose, in the literature, concrete environment architectures have been proposed (Weyns and Holvoet 2004), as well as environment programming models (d'Inverno et al. 2012) and APIs (Ricci et al. 2011). In the latter view, it is perceived and exploited by the agents as part of their world, possibly reasoning about and adapting it at runtime. For instance, in artifact-based environments as introduced by the A&A conceptual model (Omicini et al. 2008) and the CArtAgO platform (Ricci et al. 2009) as well as in GOLEM (Bromuri and Stathis 2008), the environment is structured and perceived by the agents in terms of the dynamic set of passive entities (artifacts or objects) with affordances that allow for observing and using them.

3 A motivating application

The notion of environment, as defined in the previous section, has found applications in several fields. The first of them concerns *multi-agent simulations*. For instance, pedestrians simulations (see, e.g., Bandini et al. 2009; Arcos et al. 2007) are generally employed by decision makers to analyze the plausible outcomes of their choices. Applications with *spatial aspects* profit significantly from environment-enabled mechanisms to realize *control systems*. For instance, in Weyns et al. (2005), the environment is actually employed by agents controlling automatic guided vehicles managing the transportation of loads in an automated warehouse. The agents share the physical environment and coordinate in order to avoid collisions: they use the environment to project their movement intention to *detect and resolve potential conflicts* before they actually happen. *Ubiquitous and pervasive computing* applications also consider the spatial aspects of the environment in which devices are situated. The environment is used to *manage context awareness* information, providing

indications of the local state of the system or even suggesting potentially exploitable services (Locatelli and Vizzari 2007). Other examples are applications sharing information. For instance in Krummenacher et al. (2009) medical information about patients, their histories and known conditions are shared. Patients are described according to the local rules of a given local healthcare system, but the patient might require some form of treatment when he/she is abroad. In this case, environments representing the distinct but integrated healthcare systems, react to requests coming from end-users by providing basic mechanisms for the translation of relevant pieces of information from the original ontology to the one adopted by the local system.

From this general motivation, let's turn now to an application example, which is tightly connected to Agreement Technologies.This application considers a *personalized and social e-health system* (Sarrasin et al. 2010; Kullerkupp 2012). In order to diminish the possibility of buying a counterfeit product and to assist a patient by providing alternative medicines to choose from, the system is able to read the barcode on a package in order to identify a medicine and to give access to a set of services. Moreover, with access to the patient e-health profile, the system will filter out unsuitable medicines that can interact with the other ongoing treatments or can cause allergies. This will help a buyer to find the right product with a minimal intervention by the seller/pharmacist. In addition, the system offers the possibility to rate and review medicines anonymously. The ratings and comments offer actual feedback by the consumers.

The proposed system can be designed and implemented as a complex system with an architecture involving a combination of both software agents and services invoked by them, which are embedded in the agents environment. To come up with a well-justified decision for the architecture, the system can be analyzed and designed using the principles of agent-oriented modeling (AOM) for designing socio-technical systems (Sterling and Taveter 2009). In addition to modeling the knowledge and behaviours of agents as well as interactions between them by AOM, we should also represent by appropriate models how different aspects of agreement technologies manifest themselves in the environment that should be designed for the agents:

– As the catalogues of medicaments are constantly growing, ontologies representing those catalogues are integrated directly within the environment, in order to offer to all the agents an up-to-date domain model (*Semantics dimension*).
– The environment stores different types of norms for different medicament market segments, some being stricter than the others (*Norms dimension*).
– The role model differentiates between different sorts of patients, nurses, pharmacists and medical doctors. Specific patient groups (such as diabetic type 1 patients) need to interact within specific environment spaces with specific roles and norms (*Organizations dimension*).
– When searching for adequate medicines on a market, agents use the environment to objectivize their negotiation processes (e.g. making it public to some control agents), in order to ensure the truthfulness of the negotiation and transaction (*Argumentation & negotiation dimension*).
– The environment offers a mechanism for defining ratings on other users, and services to evaluate the impact of those ratings (*Trust dimension*).

This example illustrates and motivates the potential of the environment point of view for defining agreement technologies. The next sections will describe and analyze the relations between the environment and the different aspects of agreement technologies in more detail.

4 Environment for semantics

Semantics is about the meaning of symbols. It associates symbols and expressions in a language to entities and sets of entities in the real-world (e.g. medicines, allergies, people, services). In the context of the environment, these entities may be distributed following Fig. 1 into the *basic level* of the environment, where the agents are deployed (e.g. medicines, people and services), into the *abstraction level* (e.g. service profiles, categories of objects, roles played by people), or into the *interaction/mediation level* (e.g "translation" services, argumentation services). Moreover, in the context of Agreement Technologies, semantics plays a twofold role: (*i*) agreement technologies (e.g. argumentation dos Santos et al. 2009) bring agents into agreements on their representations, and (*ii*) semantics drives agents to reach "meaningful" agreements (e.g. on trust or norms Sensoy et al. 2012 and Koster et al. 2012). From these general considerations, two kinds of requirements are raised:

1. As far as the environment dimension is concerned in the former case, agents are required to be capable to compute "meaningful" correspondences between the symbols they use, so as to interact and coordinate effectively while acting in common environment regions.
2. As far as the environment acts as a manager or interaction mediator in the second case, the environment is required to provide the appropriate facilities (e.g. services, repositories, information sources, argumentation spaces) for the agents to reach semantic agreements.

Let's analyze in what follows these two sets of requirements at the junction of environment, agreement technologies and semantics.

Agents computing semantics: While computing "meaningful" correspondences between their representations about the entities belonging to the environment, agents may encounter different "types" of disagreement:

1. agents share the same formal language (e.g. for specifying their ontologies), although they may not use the same vocabularies or conceptualizations,
2. agents use different languages and maybe different vocabularies and conceptualizations,
3. agents exploit implicit/informal semantics that are hand-wired into their decision/action cycle.

It is clear that moving from the first to the third type increases the difficulty of achieving semantic agreements. Nevertheless, as the difficulty increases (something expected in open settings), the role of the environment towards establishing semantic correspondences should become more important, as it provides the common basic level, i.e. the actual context at which all the agents are deployed.

Agents that use commonly agreed formal languages may establish agreements on their representations by computing proper correspondences between the represented entities, also driven by the semantics of their common representation framework. For instance, agents need to consider the same identifiable artifact when they both refer to an entity, e.g. to the "X service" or to the "X person". However, these semantic correspondences may be *subjective* (the correspondences between agents A and B computed by A, may not agree to the correspondences computed by B), or *objective* (a distinguished trusted entity may be delegated to compute all correspondences). It must be pointed out that the computation of objective correspondences may be delegated to (an entity / service in) the environment. An important aspect in any case, is that agents in a community or in an ad-hoc team need to reach a coherent set of correspondences, so as to function coherently and interact consistently. In large scale and open settings, with subjective correspondences, this is a challenge.

Nevertheless, the agents may need to uncover the semantics of the terms and expressions they use. To that aim, they will exploit syntactic correspondences between their distinct representations. Semantics thus emerges and evolves out of syntactic correspondences. In such a case, the environment may provide the appropriate services to further drive the emergence of meaningful correspondences.

Finally, in the case where implicit semantics are hard-wired into the deliberation - reaction cycle of an agent, agents need to exploit the environment itself to establish connections between them and further constrain the actions of others. One further (challenging) possibility to be considered consists in agents building formal representations for their semantics and then, using these representations to establish connections and agreements to the others.

Environment as interaction manager: Considering the environment as a manager or as an interaction mediator, it must provide the appropriate facilities (e.g. services, repositories, information sources, argumentation spaces) for the agents to reach semantic agreements. These interaction-mediation facilities constrain, or provide evidence to agents towards reaching agreements. In few words, the environment itself provides the context(s) where symbols can be interpreted. Thus, the agents may interpret and constrain the interpretations of their representations while interacting with the environment (also in conjunction to other agents). For instance, to learn the semantics of service input/output parameters, agents may need to gather training examples, e.g. by exploiting other agents' interactions. As another example, agents may share interactions so as to establish a common understanding of their trust assessments. Agents may further exploit contextual features to understand the meaning of tokens attributed to some resources. Services, parameters, interactions, and tokens are entities that agents share, and establish a context of interaction.

This leads to situated cognition where knowledge exists inseparably from context. Therefore, establishing agreements is seen in terms of agents' increasingly effective performance across situations rather than in terms of an "objective" combination of knowledge, since what is known is co-determined by the agents and their context.

5 Environment for norms

Norms are fundamental elements for regulating open interaction between autonomous heterogeneous agents where no assumption can be made on their behaviour. Their regulation may consist in regimentation (it is not possible for the agents to violate them) or in enforcement after violation by agents. Norms are thus a fundamental component of Agreement Technologies[3] to regulate the spaces of interaction where agreements may be defined and negotiated by these agents, that typically act on behalf of their human owners. In that context, based on deontic concepts, norms may be used to different aims: (*i*) to guide the activities of the interacting agents, (*ii*) to affect the agents' expectations on the behaviour of the other agents, (*iii*) to impact the state of the entire system by constraining the interactions for the achievement of some global desirable properties (Fornara and Colombetti 2009; da Silva Figueiredo et al. 2011; Garca-Camino et al. 2005).

While considering the relations between norms and environment, different considerations arise. Norms can represent physical or institutional constrains for the behaviour of agents. In such a case, they are situated (Okuyama et al. 2008), that is, they are meant to be fol-

---

[3] Memorandum of Understanding of the Agreement Technologies COST Action IC0801, http://w3.cost.eu/fileadmin/domain_files/ICT/Action_IC0801/mou/IC0801-e.pdf last accessed 27th September 2012.

lowed within certain physical spaces and for a certain period of time. In a similar way institutional norms that regulate the social interactions of agents should be situated in limited spaces of interaction (Tampitsikas et al. 2012) and active for a given interval of time. For example, a space may be used for representing a group of agents that collaborate on a given task for a month. This feature is important in the case of open systems where many norms may be specified and where only some of them may be relevant in a given context of interaction. Taking the agent point of view, it is also of first importance for an agent to be able to clearly identify the set of norms that are regulating its behaviour in a given situation. Referring to the different levels defined in Sect. 2, norms definitively address the *interaction/mediation level*. The environment is thus a crucial component for the management of norms in the context of agreement technologies. To that aim, some mechanisms are required:

– Mechanisms for representing norms and making them *perceivable* by the agents. In existing agent environment meta-models, norms may be represented as first-class abstractions by using for example objects (Tampitsikas et al. 2012), normative objects (Okuyama et al. 2008) or artifacts (Ricci et al. 2011). The perception of norms depends on their contextual level (e.g. role of the agent, position of the agent inside the environmental spaces, state of the environment, state of the norm). The perception of norms when situated inside an agent environment contributes to the optimization of agent reasoning and goal planning. Since norms can be discovered and perceived, the agents can adapt their strategies according to the norms, measuring in a clear and concrete way the consequences of their actions into their institutional context.
– Mechanisms for *contextualizing* at run-time norms: norms may be expressed at design-time in an abstract form (referring to roles, containing some unspecified parameters Fornara and Tampitsikas 2012). To be used, they require to be instanciated in the current context. Think for example to the mechanisms for transforming a generic obligation of the auctioneer of an auction to declare the winner of a run of the auction into a specific obligation related to specific agents that should be fulfilled within a specific deadline.
– Mechanisms for the *enforcement* of norms: this requirement is directly related to the event-based model of the environment and in particular a model of norms that apply as a response to the production of events. This level of norm support requires that the environment should be able to support: (i) the monitoring of the state of norms in accordance with the event types that trigger the transition from one norm state to another; (ii) the detection of norms violation, by checking the consistency of the produced events against the current state of norms and deciding whether it will allow the effects of the events to happen or not; (iii) the application of sanctions or rewards when norms are violated or fulfilled.

In addition to these different mechanisms, norms can be manipulated to adapt and change their content at run-time (Tinnemeier et al. 2010). These changes may be realized by the agents themselves through different coordination patterns. They may also be realized by the environment itself, while monitoring the effects of the norms on the agents' behaviour and on the state of the system, in order to preserve the stability of the system. From the software engineer point of view, it could be of first importance to reuse the existing environment frameworks (Bromuri and Stathis 2009; Ricci et al. 2011) for representing, managing, and enforcing situated norms. The challenges are thus to easily extend the functionalities/services provided by environment frameworks with the mechanisms described above.

6 Environment for organizations

Organizations is a complementary dimension to the norms, just described in the previous section. Agent organizations represent complex entities where a multitude of agents interact, within a structured environment aiming at some global purpose (Dignum 2009). Concepts such as roles (or functions, positions), groups, tasks (or activities), interaction protocols (or dialogue structures), as well as norms or rules, are key elements for modeling, designing and implementing agent organizations. In the context of Agreement Technologies, their use is twofold:

– Agent organization as *agreement facilitator*. Organizational structures may be used to restrict the way agreements are reached by fixing the capabilities of the roles that agents can play inside the organization and the relationships among these roles (e.g. power, authority, compatibility, etc.). In this sense, the organizational structure imposes limits to which agents can negotiate with others, and which weight or importance will have their decisions inside the argumentation processes.

– Agent organization as *agreement representation*. In the case of reaching an agreement, the organizational structure may be used to express the agreement in terms of roles that can be played by agents, of relationships between these roles (i.e. social/power/authority relationships), of norms or restrictions that affect them, of capabilities of the roles, etc. For instance, when two companies merge, their agreement will result in a new organizational structure, with new roles and relationships that might integrate the entities existing in the original companies.

As in human organizations, the environment is a fundamental issue to consider in order to understand how agent organizations behave either as agreement facilitators or as agreement representation. The environment encompasses all the actors that are outside the organization (suppliers, customers, competitors and institutions that regulate the behaviour of the members of the organization) and also the economic, geographic and political conditions (Wagner and Hollenbeck 2001). Thus, the environment provides the materials, technology and members that the organization needs to develop its products and services as well as to sell and get benefits. It is the source of the resources that are required by the organization to survive as well as the source of every opportunity to succeed and every threat to the existence of the organization (Hodge et al. 2002).

In that global picture, as for norms, different considerations arise as soon as we envision inter-relation between agent organisations and environment. They clearly address the *interaction/mediation level* of Fig. 1. We can distinguish first the inclusion of the environment into the representation of organizations and, second, the definition of mechanisms in the environment for the management of the organization.

When modelling agent organization, the environment model normally provides constructs to represent the collection of resources in the space of the agent organization formed by non-autonomous entities that can be perceived and acted upon (manipulated, consumed, produced, etc.) by the component agents. AGRE (Ferber et al. 2005), MAS-ML (da Silva et al. 2008), and AOM (Sterling and Taveter 2009) are examples of organizational models (modelling techniques) that provide constructs to represent organization environment elements.

Recently, some approaches have produced some mechanisms in the environment for helping the organization to make better profit of it. These approaches are mainly based on the Agents & Artifacts (A&A) conceptual framework (Omicini et al. 2008), that models the environment around (*i*) *agents*, the proactive entities of the system, (*ii*) *artifacts*, passive entities that help agents to reach their objectives by means of their provided functionality, and (*iii*)

workspaces, used to model the topology of the environment. For example, in Esparcia et al. (2011b, 2011a), *artifacts for organizational mechanisms* provide several mechanisms for: helping agents to obtain information about organizations of their environment; modifying the global behaviour of the system by changing the incentive system of the organization; and updating the action space of agents. Another example is the set of *organizational artifacts* implemented in ORA4MAS (Hübner et al. 2010), aimed at managing the current state of the organization in terms of groups, social schemes, and normative state. They provide actions for agents to proactively take part in an organization (adopting/leaving roles, committing to goals) and actions for organizational agents to manage the organization. Those examples show that the environment can provide several mechanisms to allow the agents to know what are the organizations in the system and how to participate to them. In fact, situated organizations make visible their state and properties in order to let the agents perceive and decide to join them. Situated organizations can also simplify the detection and solving of interdependencies between organizations (Tampitsikas et al. 2012). Furthermore, physical objects situated in the environment may be used as mediators for the access to some organizations: using these physical objects may count, for the user agents and for the organization, as endorsing some role in the organizations (Piunti et al. 2009). This can help bridging the conceptual gap between the elements available at the organizational level, normally defined in terms of norms, roles, resources, global goals, stakeholders, etc.; and the notions and constructs adopted at the agent level, for example with mentalistic notions such as beliefs, desires, intentions of a BDI-like agent model.

As a result, social structures can be explicitly expressed and shaped through organizational concepts, which can be perceived and managed by means of elements of the environment, such as artifacts (e.g. the above mentioned "organizational artifacts" or the "artifacts for organizational mechanisms"). These social structures can then be exploited by agents to cope with the difficulties of solving complex tasks in a coherent and efficient manner.

7 Environment for argumentation & negotiation

Instantaneous agreement between multiple humans or machines is quite rare. Different people have different beliefs about the world, different preferences and goals, and different decision constraints, such as deadlines or available processing resources. Accordingly, before agreement between multiple people or machines can be reached on some matter, there is often a requirement for the identification and possible resolution of these differences. Agreement may be possible even without resolution of differences, as when two parties to an economic transaction both believe they have done better out of the deal than did the other party. Indeed, some technologies and methods for reaching agreement disregard differences between participants. For example, simple voting mechanisms pool the varied preferences of the participants without trying to resolve or even identify the various preferences. Similarly, auction protocols seek to find an agreement between potential buyers and sellers of some item without trying to explain or resolve the different beliefs of the participants over the item's value.

Many human interactions, however, do use methods for identifying and possibly resolving differences in the beliefs, preferences, or objectives of the participants. Chief among these methods are those based on argumentation—the requesting and giving of reasons for statements—and negotiation—the attempt to decide between mutually-incompatible goals. Human beings have been engaged in argumentation and negotiation since our earliest days

as a species. It is not surprising therefore that we should seek to vest software agents with the same capabilities to argue, to negotiate, and to jointly deliberate about what to believe and what to do (McBurney and Parsons 2009). Given a system comprising multiple autonomous software agents, jointly interacting via argumentation to achieve their individual or shared objectives, there are several resources that the system environment may provide to the participating agents. These resources include:

– Shared ontologies and semantics for communications between different agents.
– Libraries of protocols for agent interaction, possibly including associated meta-data and annotations, such as statements of possible protocol outcomes and statements and proofs of protocol properties.
– Mechanisms for monitoring adherence to legitimate agreements reached between agents engaged in argument, along with appropriate enforcement, punishment and compensation procedures. These mechanisms have clear links to the norms discussed in Sect. 5.
– Third-party escrow accounts, which may hold money or other items in transit between negotiating agents. For example, an agent A who contracts to purchase a service from agent B may not pay directly, but, with B's agreement, may put the money in an escrow account to be held until the service is performed, whereupon the money is transferred to B.
– Mechanisms for third-party certification, as happens with professional groups (such as doctors) and artisans in many countries.
– Record-keeping, for example, acting as an independent register of all legitimate agreements made. For some types of interactions, a copy of the discussions leading to agreements may also have to be stored; this is the case, for example, in safety-critical domains such as medicine or the design of aircraft.
– Mechanisms for mediation and intervention in agent interactions, e.g. to suggest potential compromise agreements to agents engaged in negotiation or conflict.

There has been research on some of these topics. For example, Oren and colleagues (Oren et al. 2004) have considered policing and enforcement mechanisms for multi-agent systems. Miller and McBurney (2008, 2010) have considered efficient means of storage and retrieval of protocols stored in libraries. And, considerable work in argumentation has looked at intelligent mediation between conflicting agents, for example (Gordon and Karacapilidis 1997; Kraus et al. 1998; Oliva et al. 2007).

In the reverse direction, argumentation and negotiation activities can usefully be applied to the design and selection of various features of the Environment. For example, participants may negotiate a shared ontology, as in Beun and Eijk (2004), Reed et al. (2002), for use in subsequent interactions. Similarly, participants may argue over the choice of a protocol, prior to commencing an interaction using the agreed protocol (Hulstijn et al. 2000). Clearly, there are close links between these issues and the topics of semantics, norms, and trust.

Implicit in the study of argumentation and negotiation by agreement technologists is a belief that these approaches should enhance the likelihood of agreements being reached, or of deals being struck, between the different participants. There has only been limited research on this question, but what has been done supports this implicit assumption — for instance, the work of Karunatillake Karunatillake (2006) demonstrates the beneficial impacts of argumentation in reaching agreements over conflicting uses of resources. Much research remains to be done to understand better and to quantify these benefits.

8 Environment for trust

On the contrary of other Agreement Technologies like negotiation or argumentation that are directly linked to the idea of agreement, trust and reputation are not agreement generator mechanisms. However, without trust and reputation a great amount of agreements would not be possible or rather difficult. Any agreement process has associated a certain level of uncertainty. Trust and reputation mechanisms help the individual to deal with (at least part of) that uncertainty. For example, in an argumentation process (as described above) some of the arguments can be based on information that cannot be confirmed. The acceptance of these arguments depends only on the reliability of the sender. The reputation of that sender can make up for that lack of direct knowledge about the arguments, becoming the key to finally reach the agreement. The relation between trust and reputation as Agreement Technologies and the environment can be analysed from two perspectives:

1. First looking at the environment as a *support* for trust and reputation mechanisms (and in general as a support for any Agreement Technology)
2. and second looking at the trust and reputation systems as *modifiers* of the environment (again, this role of environment modifier is not exclusive of trust and reputation but can be applied also to other Agreement Technology).

In general for the Agreement Technologies, and in particular for trust and reputation, the environment is the substratum that supports them. One of the main sources to build trust on others is the result of direct interactions. These interactions are the consequence of a necessity from one of the agents (the trustor) that "trusts" the other (the trustee) to perform an action. After agreeing on the commitment, the environment will become the place where the action will be executed and more importantly, where the result of the action will be made evident. The observation of the performance and its subsequent evaluation is one of the most important inputs used by the trustor to modify the trust on the trustee. Therefore the environment, on top of being the place where the trustee performs the action, is also *channel* that allows the trustor to perceive the performance of the trustee. In the case of reputation something similar happens. We can define reputation as "what a social entity says about a target regarding his/her behaviour". The definition states the fact that reputation cannot exist without communication. Reputation is based on information circulating among individuals. Therefore, we can see how the environment is acting again as a support by allowing this communication. In both cases it is clear that the environment can be an obstacle or can facilitate the functioning of trust and reputation mechanisms (and Agreement Technologies in general). An environment that allows an objective observation of the individuals' behaviour and facilitates communication is favouring and making more effective the use of trust and reputation mechanisms (and the other way around).

The environment is the support element of trust and reputation but at the same time, trust and reputation mechanisms modify the environment. Trust and reputation have a dual nature: as local mechanisms used by individuals to select their partners and, from a societal perspective, as social instruments for social order. As mechanisms for social order, their task is to influence the behaviour so to enforce some kind of conduct in detriment of another. Social relations among individuals are part of the environment. Reputation as one of the sources for trust, and trust as a mechanism that defines in a big extend the relations among individuals, have an important influence in the intensity, creation and destruction of these social relations. Apart from this direct influence of trust and reputation in the environment, more than often societies also modify the environment to improve trust among their members and facilitate reputation mechanisms to detect wrongdoers (for instance by adding

communication mechanisms to allow reputation spread, by creating physical and social spaces to facilitate trust relations among companies, etc.).

9 Discussion and concluding remarks

From the characterization of agent environment done in Sect. 2 and the analysis of its role with respect to the different dimensions of Agreement Technologies in the other sections, we can broadly frame the benefits of the agent environment for agent-based agreement technologies at two different (but related) levels:

– *conceptual level* where the use of environment improves the modeling and design of the strategies and mechanisms that allow for reaching agreements, conceiving solutions that don't rely necessarily only on agents and message passing as unique abstractions,
– *practical level* where the use of agent environment improves the separation of concerns, the modularity, reusability, openness and extensibility in designing and implementing agent-based agreement technologies.

Let's recall shortly, for instance, what was presented on the semantics dimension (see Sect. 4) to illustrate those points. When we consider heterogeneous agent-based agreement technologies, benefits of environment at the conceptual level and its implication at the practical level are clearly visible: the environment is an unavoidable mediating entity between the semantics which are hard-wired into the internal reasoning of the different agents. Other impacts in relation with norms, organizations, argumentation and negotiation, trust and reputation can be found in the previous sections of the paper.

Going a step further in the characterization of agent-environment for agent-based agreement technologies, it appears that the *interaction mediation-support layer* of the agent environment as shown in the Fig. 1 plays a crucial role. In some sense, this layer appears to be an *agreement mediation layer*. This layer implements the high level constructs of a *coordination metamodel* which is involved in the building of agreements. Moreover, this layer ensures also that the actual specification of a particular instance of the coordination conventions pertaining to that coordination metamodel, is properly implemented and the corresponding conventions dully enforced (d'Inverno et al. 2012). It may thus be seen and understood as the "institutional layer" as defined in d'Inverno et al. (2012)) in the sense that:

1. a virtual agreement space—or "institutional", in the sense of Searle (1995)— is created (at run-time) for the admission and processing of certain messages, and
2. that the admission of these messages and their subsequent processing in the environment comply with those conventions that have been specified.

To achieve this double purpose of translating the specification into a run-time agreement space and enforcing the coordination conventions within that agreement space, the agreement mediation layer of the environment needs to include appropriate data structures, operations, and governance functionalities. The data structures mirror the conceptual coordination devices and the operations apply to those data structures. Governance functionalities ensure that the actual interaction flow complies with the coordination conventions that have been specified. Thus, in particular, this layer would need to mirror properly the coordination metamodel by providing support to basic services such as:

– Time keeping, and consequently control of time-outs, synchronization and interaction clock (if part of the metamodel)

- Interface with the external environment (depending on the metamodel: services, activation platform, other organizations)
- Primitive means of decomposition of complex activities (e.g., scenes and transitions in the case of electronic institutions)

It has also to provide support to elaborated services in relation to agreement technologies such as:

- Operations that implement interactions between agents and environment: enter or leaving the agreement space, moving form one subactivity to another, performing an atomic action, ...
- Operations that the environment needs to perform in order to keep activity flowing in the agreement space: create and terminate the agreement space and its subactivities, controlling the changes of roles attempted by agents, managing the movements of agents between activities and most importantly, guaranteeing that atomic actions comply with stated norms, ...
- Data structures needed to keep track of agent activity, namely, the state of the local and global contexts as interactions proceed
- Representation and handling of atomic interactions among agents, handling of interaction flow

Besides these basic and common features, the paper has shown that elaborated services in relation to agreement technologies need to be considered and added in that layer. This raises different challenges and multiple and interesting research directions for supporting and developing Agreement Technologies. In the following, we enumerate some of these challenges in relation with of the dimensions.

Challenges w.r.t. semantics and environment for Agreement Technologies: Considering semantics, it is required elaborated services to:

- represent environmental aspects at different levels of support: either using a specific representation framework, or using any combination of formal and informal representation means, agents need to explicitly represent knowledge about all the environment aspects,
- combine knowledge in large-scale open settings & reconcile subjective views: if we need to deal with open and large scale settings, then establishing common and coherent views of the shared environment is unavoidable for agents to coordinate and collaborate effectively. However, establishing semantic agreements between any pair of distant agents in large and open communities of agents is a challenge. The environment dimension has not been thoroughly investigated towards reaching this end,
- help in learning the semantics of everything, out of cases: being able to represent everything in the environment and understanding the representations used, is one thing. Inspecting and exploiting the interactions of others with the environment (within specific contexts of interaction) agents must be able to learn the meaning of representations used.
- support computing emergent, commonly agreed languages (symbols, syntax and semantics) for interaction. In open and large-scale settings, agents may use any formal or informal representation language. Computing commonly agreed representation means, is a great challenge.

Challenges w.r.t. norms and environment for Agreement Technologies: Considering the norms, it is of first importance of defining and representing in a standard way the events and actions that happen in an environment, as well as the context of the interactions in terms

of properties of resources and their value in order to be used in the definition of norms that should regulate those events and actions. There is also the need of elaborated services to add in this agreement mediation layer such as:

– general mechanisms for contextualizing abstract norms defined at design-time into norms situated in specific spaces,
– extendable functionalities/services provided by the environment for adding those required for norms management.

Challenges w.r.t. organizations and environment for Agreement Technologies: Moving to the organization, elaborated services, situated in the agent environment and accessible to the agents, are needed to:

– provide facilities to enter or exit a given organization to allow run-time recruitment of new members as well as voluntary desertion and/or expulsion of members,
– support on-demand creation, deletion and modification of organizations,
– give support to the institutional components of an organization, i.e. norms, powers, agreements.

Challenges w.r.t. argumentation, negotiation and environment for Agreement Technologies: Considering the argumentation dimension, besides scaling up the existing work, which typically considers single interactions between two or a small number of agents, elaborated services have to be considered to complement the basic services dedicated to interaction management. These elaborated services will provide the possibility to manage libraries and database of ontologies, protocols and agreements. Finally, once abstractions such as ontologies and interaction protocols are stored and accessible in the environment, participating agents in a system then need the ability to reason about these abstractions and to invoke them as required. Some work in this direction has been undertaken, e.g. Miller and McBurney (2011), but much remains to be done.

Challenges w.r.t. Trust, reputation and environment for Agreement Technologies: Finally considering Trust and Reputation, challenges and elaborated services need to be considered in order to exploit the existing strong link between trust and reputation and the environment in electronic environments. Computational trust and reputation models (Pinyol and Sabater-Mir 2012) till recently have been designed as isolated components, barely linked to the context and the rest of the agent elements. Therefore they have little capacity to change the environment to facilitate its functioning as described before. One improvement in this sense would be giving the trust and reputation system the capacity to influence the actions of the agent to modify the environment so it becomes more trust and reputation "friendly". For example the trust and reputation mechanism could suggest actions to the agent in order to improve the communication capabilities the environment has or suggest the formation of new social relations to enforce already present trust relationships. So instead of relying on the possibilities the environment offers "for free" to facilitate trust and reputation, the agent (by means of its trust and reputation mechanism) should be proactive and modify the environment to make more robust and reliable the two social artifacts.

References

Arcos JL, Noriega P, Rodriguez-Aguilar JA, Sierra C (2007) E4mas through electronic institutions. In: Weyns D, Parunak H, Michel F (eds), Environments for multi-agent systems III., number 4389 in Lecture Notes in Computer Science, pp. 184–202. Springer, Berlin/Heidelberg, 08/05/2006

Bandini S, Manzoni S, Vizzari G (2009) Agent based modeling and simulation: an informatics perspective. J Artif Soc Soc Simul 12(4):4

Beun R-J, Eijk R (2004) A co-operative dialogue game for resolving ontological discrepancies. In: Dignum F (ed) Advances in agent communication, LNAI 2922. Springer, Berlin pp 349–363

Bromuri S, Stathis K (2008) Situating cognitive agents in GOLEM. In: Weyns D, Brueckner S, Demazeau Y (eds) Engineering environment-mediated multi-agent systems, vol. 5049., of LNCSSpringer Berlin, Heidelberg, pp 115–134

Bromuri S, Stathis K (2009) Distributed agent environments in the ambient event calculus. In: Proceedings of the third ACM international conference on distributed event-based systems, DEBS '09, pp 12:1–12:12, New York, NY, USA, ACM

da Silva VT, Choren R, Lucena CD (2008) MAS-ML: a multi-agent system modeling language. Int J Agent-Oriented Softw Eng 2(4):381–421

da Silva Figueiredo K, Torres da Silva V, de Oliveira Braga C (2011) Modeling norms in multi-agent systems with normml. In: De Vos M, Fornara N, Pitt J, Vouros G (eds) Coordination, organizations, institutions, and norms in agent systems VI, vol. 6541., of Lecture Notes in Computer ScienceSpringer, Berlin Heidelberg, pp 39–57

Dignum V (2009) Handbook of research on multi-agent systems : semantics and dynamics of organizational models, chapter the role of organization in agent systems, pp 1–16. Information Science Reference Publisher, ISBN 978-1-60566-256-5, doi:10.4018/978-1-60566-256-5.ch001

d'Inverno M, Luck M, Noriega P, Rodriguez-Aguilar JA, Sierra C (2012) Communicating open systems. Artif Intell 186:38–94

dos Santos CT, Quaresma P, Vieira R, Isaac A (2009) Comparing argumentation frameworks for composite ontology matching. In: McBurney P, Rahwan I, Parsons S, Maudet N (eds), ArgMAS, vol. 6057 of Lecture Notes in Computer Science. Springer, pp 305–320

Esparcia S, Argente E, Centeno R, Hermoso R (2011) Enhancing MAS environments with organizational mechanisms. Int J Artif Intell Tools (IJAIT) 20(4):663–690

Esparcia S, Centeno R, Hermoso R, Argente E (2011) Artifacting and regulating the environment of a virtual organization. In: 23rd IEEE international conference on tools with artificial intelligence (ICTAI 2011). IEEE

Ferber J, Michel F, Baez J (2005) AGRE: integrating environments with organizations. In: Environments for multi-agent systems, vol. 3374 of LNCS, pp 48–56. Springer

Fornara N, Colombetti M (2009) Specifying and enforcing norms in artificial institutions. In: Baldoni M, Son T, van Riemsdijk M, Winikoff M (eds) Declarative agent languages and technologies VI, vol. 5397. Springer, Berlin, pp 1–17

Fornara N, Tampitsikas C (2012) Using OWL artificial institutions for dynamically creating open spaces of interaction. In: Proceedings of the AT 2012 first international conference on agreement technologies, October 15–16, 2012 in Dubrovnik, Croatia., page to appear

Garca-Camino A, Noriega P, Rodrguez-Aguilar JA (2005) Implementing norms in electronic institutions. In: Proceedings of the 4th international joint conference on autonomous agents and multiagent systems (AAMAS '05, (ed) Michal Pechoucek ST, Steinernald. Utrecht, NL, ACM Press, pp 667–673

Gordon TF, Karacapilidis N (1997) The zeno argumentation framework. In: Proceedings sixth international conference on AI and law, pp 10–18, New York, ACM Press

Hodge BJ, Anthony W, Gales L (2002) Organization theory: a strategic approach. Prentice Hall, Englewood Cliffs

Hübner J, Boissier O, Kitio R, Ricci A (2010) Instrumenting multi-agent organisations with organisational artifacts and agents. Auton Agents Multi-Agent Syst 20:369–400

Hulstijn J, Dastani M, van der Torre L (2000) Negotiation protocols and dialogue games. In: Proceedings Belgian-Dutch AI conference (BNAIC-2000), Kaatsheuvel

Karunatillake NC (2006) Argumentation-based negotiation in a social context. Phd, School of electronics and Computer Science, University of Southampton, Southampton, UK

Koster A, Schorlemmer WM, Sabater-Mir J (2012) Engineering trust alignment: theory, method and experimentation. Int J Hum-Comput Stud 70(6):450–473

Kraus S, Sycara K, Evenchik A (1998) Reaching agreement through argumentation: a logical model and implementation. Artif Intell 104:1–69

Krummenacher R, Simperl EPB, Cerizza D, Valle ED, Nixon LJB, Foxvog D (2009) Enabling the European patient summary through triplespaces. Comput Methods Programs Biomed, 95(2–S1):33–43

Kullerkupp R (2012) Analysis and preliminary design of a social networking system for supporting information exchange about medicines. Master's thesis. Department of Informatics, Tallinn University of Technology, Estonia

Locatelli MP, Vizzari G (2007) Awareness in collaborative ubiquitous environments: the multilayered multi-agent situated system approach. TAAS 2(4)

McBurney P, Parsons S (2009) Dialogue games for agent argumentation. In: Rahwan I, Simari G (eds) Argumentation in artificial intelligence, chapter 13. Springer, Berlin, pp 261–280

Miller T, McBurney P (2011) Propositional dynamic logic for reasoning about first-class agent interaction protocols. Comput Intell 27(3):422–457

Miller T, McBurney P (2008) Annotation and matching of first-class agent interaction protocols. In: Proceedings seventh international joint conference on autonomous agents and multi-agent systems (AAMAS 2008, (ed) Padgham L, Parkes D, Mueller JP, Parsons S. NY, SA, ACM Press, New York

Miller T, McBurney P (2010) Characterising and matching iterative and recursive agent interaction protocols. In: Proceedings ninth international conference on autonomous agents and multi-agent systems (AAMAS 2010, (ed) van der Hoek W, Kaminka G, Luck M, Sen S. Toronto, Canada, IFAAMAS, ACM Press, pp 1207–1214

Odell J, Parunak HVD, Fleischer M, Brueckner S (2003) Modeling agents and their environment. In: Agent-oriented software engineering III, vol. 2585 of LNCS, pp 16–31. Springer, Berlin

Okuyama FY, Bordini RH, da Rocha Costa AC (2008) A distributed normative infrastructure for situated multi-agent organisations. In: Proceedings of the 7th international joint conference on autonomous agents and multiagent systems - vol. 3, AAMAS '08. Richland, SC, International Foundation for Autonomous Agents and Multiagent Systems, pp 1501–1504

Oliva E, McBurney P, Omicini A (2007) Co-argumentation artifact for agent societies. In: Rahwan I, Reed C, Parsons S, (eds), Proceedings fourth international workshop on argumentation in multi-agent systems (ArgMAS 2007), pp. 115–130, AAMAS (2007) Honolulu. Hawai'i, USA

Omicini A, Ricci A, Viroli M (2008) Artifacts in the A&A meta-model for multi-agent systems. Auton Agents Multi-Agent Syst 17(3):432–456

Oren N, Norman T, Preece A, Chalmers S (2004) Policing virtual organizations. In: Proceedings second european conference on multi-agent systems (EUMAS 2004, (ed) Ghidini C, Giorgini P, Hoek W. Barcelona, Spain, EUMAS, pp 499–508

Pinyol I, Sabater-Mir J (2012) Computational trust and reputation models for open multi-agent systems: a review. Artif Intell Rev, pp 1–25

Piunti M, Ricci A, Boissier O, Hübner JF (2009) Embodied organisations in MAS environments. In: Proceedings of the 7th German conference on multiagent system technologies MATES'09, vol. 5774 of LNAI, pp 115–127. Springer

Reed C, Norman TJ, Jennings NR (2002) Negotiating the semantics of agent communications languages. Comput Intell 18(2):229–252

Ricci A, Piunti M, Viroli M, Omicini A (2009) Environment programming in cartago. In: Bordini RH, Dastani M, Dix J, El Fallah-Seghrouchni A (eds) Multi-agent programming: languages, platforms and applications, vol 2. Springer, Berlin, pp 259–288

Ricci A, Piunti M, Viroli M (2011) Environment programming in multi-agent systems: an artifact-based perspective. Auton Agents Multi-Agent Syst 23(2):158–192

Sarrasin JJ, Schumacher M, Hay C, Richard P (June 2010) Health-identity: mobile services for consumers of medicines. In: Special topic conference "Seamless care - safe care: the challenges of interoperability"

Searle JR (1995) The construction of social reality. Free Press, NY

Sensoy M, Norman TJ, Vasconcelos WW, Sycara KP (2012) Owl-polar: a framework for semantic policy representation and reasoning. J Web Sem 12:148–160

Sterling L, Taveter K (2009) The art of agent-oriented modeling. The MIT Press, Massachusetts

Tampitsikas C, Bromuri S, Fornara N, Schumacher MI (2012) Interdependent artificial institutions in agent environments. Appl Artif Intell 26(4):398–427

Tinnemeier N, Dastani M, Meyer J-J (2010) Programming norm change. In: Proceedings of the 9th international conference on autonomous agents and multiagent systems: vol. 1, AAMAS '10. Richland, SC, ACM, pp 957–964

Wagner J, Hollenbeck J (2001) Organizational behavior. South-Western Pub

Weyns D, Holvoet T (2004) Formal model for situated multiagent systems. Fundam Inform 63(2–3):125–158

Weyns D, Omicini A, Odell JJ (2007) Environment as a first-class abstraction in multi-agent systems. Auton Agents Multi-Agent Syst 14(1):5–30

Weyns D, Parunak HVD (eds) (2007) Special issue on environments for multi-agent systems, vol, 14 (1) of autonomous agents and multi-agent systems. Springer Netherlands

Weyns D, Parunak HVD, Michel F, Holvoet T, Ferber J (2005) Environments for multiagent systems: state-of-the-art and research challenges. In: Weyns D, Parunak HVD, Michel F, Holvoet T, Ferber J (eds) Environment for multi-agent systems, vol. 3374. Springer, Berlin pp 1–47

Weyns D, Schelfthout K, Holvoet T, Lefever T (2005) Decentralized control of E'GV transportation systems. In: Pechoucek M, Steiner D, Thompson SG (eds) AAMAS industrial applications, pp 67–74. ACM