This is a pre print version of the following article:

Mixed environments for MAS: Bringing humans in the loop / Ricci, A., Rodriguez Aguilar, J.A., Pijoan, A., Zambonelli, F.. - STAMPA. - 9068:(2015), pp. 52-60. (4th International Workshop on Agent Environments for Multi-Agent Systems, E4MAS 2014 Paris; France May 6, 2014) [10.1007/978-3-319-23850-0_4].

Springer
Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

13/11/2026 05:55

(Article begins on next page)

Mixed Environments for MAS: Bringing Humans in the Loop

A. Ricci[1], J. A. Rodriguez-Aguilar[2], A. Pijoan[3] and F. Zambonelli[4]

[1] DISI, University of Bologna,
Cesena (FC), Italy
a.ricci@unibo.it
[2] Artificial Intelligence Research Institute, IIIA
Spanish Council for Scientific Research, CSIC
08193 Cerdanyola, Spain
jar@iiia.csic.es
[3] Deusto Institute of Technology
University of Deusto, Bilbao, Spain
ander.pijoan@deusto.es
[4] Univ. di Modena e Reggio Emilia
Modena, Italy
franco.zambonelli@unimore.it

Abstract. In many application domains for agents and MAS, the interaction between the systems and human users is a main element. In some cases, the interaction occurs behind a traditional computing device, such as a computer desktop or a smartphone. In other cases, the interaction occurs through the physical world. This is the case, for instance, of smart/intelligent environment applications, and more generally in the wide context of Internet-of-Things based apps. Can the concept of agent environment for MAS play a role in the design of such systems, where humans *are in the loop*? In this position paper we further develop this question, providing some reflections and suggestions for future works.

1 Introduction

Emerging technologies such as wireless sensor networks, Internet of Things (IoT), smart and wearable devices, or even volunteer geographic information, provide a whole new virtual layer where the physical world can be accessed or modified by any computational system. This information gathering capabilities from distributed sources is poised to revolutionize the way MAS environments interact with our real world.

These are main examples of application domains where humans are *in the loop*. By being in the loop we mean to be an essential part of the picture, either as: *(i)* a user of the system, engaging in continuous interactions with it through traditional computational devices (PC, tablet, smartphones,..) or with the physical environment itself, in the IoT perspective; or *(ii)* part of the system itself, like in the case of socio-technical systems.

Multi-agent systems appear as an effective approach for modelling and designing this kind of systems given their characteristics. In that, the notion of (application) environment in MAS as elaborated by the community [11] is called to play a relevant role. In fact, in these systems the interaction with a shared distributed physical and social world is a primary aspect, as well as the presence of rich forms of communication, either explicit or implicit, uncoupled, and persistent.

However, research contributions about environments for MAS so far have mainly focused on cases where the first-class citizens acting, perceiving, and interacting have typically been artificial agents —either software or robots. What about first-class agent environments where *both human and artificial agents* are first-class citizens? This is the type of environments we focus on in this paper. Therefore, we are interested in a new type of agent environments that henceforth we shall refer to as *mixed environments* since they are meant to realise mixed multi-agent systems populated by both human and artificial agents.

2 Background: Agent Environment Support Levels

To elaborate the question raised in the introduction above, we start by recalling the three levels describing the environment level of support reported in the literature [11]. These levels include:

- *basic level* — At the basic level, the agent environment enables agents to access the deployment context (see Fig. 1). That is, agents know low-level language, details and directly access hardware and software resources.
- *abstraction level* — The abstraction level bridges the conceptual gap between the agent abstraction and low-level details of the deployment context (see Fig. 2). Agents access an interface shielding specific details of the resource hidden behind.
- *interaction-mediation level* — The interaction-mediation level offers support to regulate the access to shared resources, ensure restrictions are met and mediate interaction between agents (see Fig. 3).

The three levels of support represent different degrees of functionality provided by the environment that agents can use to achieve their goals. The first question is: *where do we put humans in these layers?*.

3 Modelling Humans In the Loop

On the one side, one could consider humans as part of the deployment context— so the basic level. However, as soon as we consider modern application domains such as pervasive computing / smart environments, we realize that such a modeling is not fully satisfactory, from an abstraction point of view in particular.

An alternative modeling could be considering humans as part of the agents, so that the application environment becomes the glue also among human users

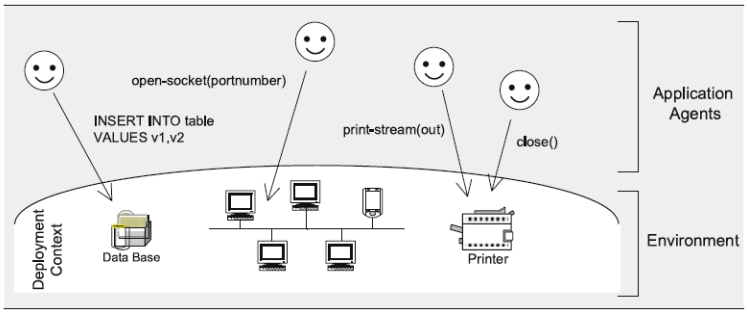


Fig. 1. Basic support level (from [11]).

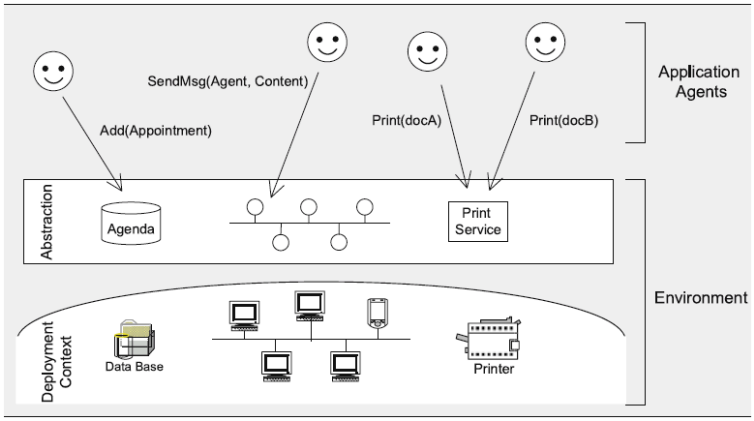


Fig. 2. Abstraction support level (from [11]).

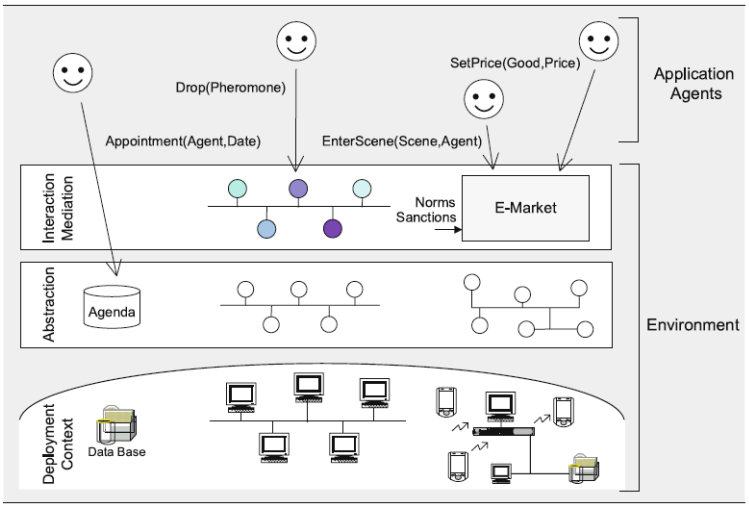


Fig. 3. Interaction-Mediation support level (from [11]).

and agents at different levels. But in this case we are considering part of the MAS something which is – actually – outside the system, being the *users* of the system.

A further solution in the middle could be to introduce "special agents", kind of *user agents*, that have the goal of modeling/representing the human users inside the system, with the goal of reflecting their actions and intentions, and collecting perceptions and information that should be brought back to the human users that they represent. This kind of agents is quite recurrent in agent-based applications. From an agent environment point view, they function as kind of *facilitators*, making it possible to bring inside the system actions that human user aim to perform and bring outside the system the perceptions/information targeted to them. From an abstraction point of view, they could be still considered part of the application environment. In fact, in a conceptual model like A&A [8], these facilitators are better modelled in terms of *artifacts* – not agents – designed in this case to make human actions observable (from an agent point of view) or agents' actions (from a human point of view).

Being represented either as agents or not, the problem is the model adopted to define the interaction between humans and the MAS. This issue is not new for the agent research community: the point here is to understand if the idea of environment as first-class abstraction inside a MAS can be helpful for that purpose.

4 Reifying Human Actions and Perceptions into the Agent Environment

Now we consider some relevant examples in the literature of agent-based systems with humans in the loop.

4.1 Pervasive Ecosystems and Human-aware Superorganisms

The first is given by pervasive ecosystems and human-aware superorganisms, as introduced and developed by Zambonelli and colleagues in [9, 12, 13].

Generally speaking, pervasive and ubiquitous computing is a well-known and obvious case where humans are in the loop. These technologies promise to notably change the future ICT landscape, letting us envision the emergence of an integrated and very dense socio-technical infrastructure for the provisioning of innovative general-purpose digital services. The infrastructure will be used to ubiquitously access services for better interacting with the surrounding physical world and with the social activities occurring in it.

Following this line, future urban environments can be depicted as *sociotechnical superorganism* [12], where networks of entities – ICT devices and citizens – will continuously and perhaps invisibly cooperate in highly decentralized and participatory sensing activities. The real-time sharing of the results of such activities at city scale will enable a shared understanding, via computing and

thinking, of urban issues and their dynamics. This in turn will make it possible to plan and direct responses or fix problems with specific collective actions.

In order to tackle the design of this kind of systems, the SAPERE project [13] envisioned an approach based on modeling and architecting pervasive service environments as a non-layered spatial *substrate*, laid above the actual pervasive network infrastructure, on top of which human users act as "prosumers" continuously producing and consuming data. The substrate embeds the basic laws of nature (or eco-laws) that rule the activities of the system. There, individuals of different species – agents/services, data, and devices – interact and combine with each other (in respect of the eco-laws and typically based on their spatial relationships), so as to serve their own individual needs as well as the sustainability of the overall ecology.

The approach is clearly data-centric, in that the environment supporting humans/agents interaction and coordination can be implemented as an open distributed set of data spaces, hosting streams of tuples – generated by sensors, actuators, so human actions and reactions – semantically combined, aggregated, manipulated, and diffused according to some nature-inspired laws.

4.2 Human-Agent Collectives

A second example stems from recent works on *human-agent collectives* (HACs) [6]. HACs refer to all those systems where humans and software agents continually and flexibly establish a range of collaborative relationships with one another, to meet their individual and collective goals, operating at a global scale. In HACs, depending on the task at hand, different constellations of people, resources, and information must come together, operate in a coordinated fashion, and then disband. The openness and presence of many distinct stakeholders, each with their own resources and objectives, means participation is motivated by a broad range of incentives – extrinsic (for example, money or tax-benefit), social, or image motivation (for example, public accreditation or leader-board position), or intrinsic (for example, personal interest in a social cause, altruism, or hobby) rather than diktat. Moreover, once presented with such incentives, the stakeholders need to be evaluated and rewarded in ways that ensure they sustain behaviours that are beneficial to the system they partially form.

4.3 Virtual Institutions

A third example comes from research on virtual institutions [2], a concept that combines electronic institutions [1] and 3D virtual worlds.

Virtual worlds technology has recently emerged in computing with enormous strength[7]. A virtual world is an online immersive environment where, using 3D visualisation, humans participate represented as graphically embodied characters (avatars) and interact with others and the environment by using simple and intuitive control facilities. Because humans are social, the concept of virtual worlds is very appealing to mediate their remote interactions. Nowadays, there are millions of people connecting to virtual worlds every day. Such immersive

and interactive environment provides many possibilities to represent the system state and the regulations defined by the coordination model. For instance, the other participants are represented also as avatars and their appearance can be used to display the role they are playing. We argue that 3D virtual worlds can be successfully used to incorporate humans into MAS. To illustrate this hypothesis, next we outline the work carried out to open electronic institutions to humans by using virtual worlds.

The aim of virtual institutions is to design regulated environments where both human and software agents can participate. In this context, electronic institutions are used to define the rules that structure participants interactions (for both humans and artificial agents), while 3D virtual worlds facilitate human participation in the system. Notice that virtual institutions provide different interfaces with the interaction-mediation level in Figure 3 to artificial agents and humans agents. On the one hand, artificial agents are allowed to directly interact with the electronic institutions' coordination infrastructure by means of an agent communication language. On the other hand, humans visualise and interact in a virtual world where they actions are collected and translated to be processed by the electronic institutions' coordination infrastructure. Thus, the environment's façade is different for human and artificial agents.

4.4 Mirror Worlds

A forth example is based on mirror worlds, as introduced in [4]. Originally inspired by Gelernter's book with the same name [5], mirror worlds can be conceived as an agent-based extension of augmented and mixed reality. Both human and artificial agents inhabit an environment which is both physical and augmented of a digital virtual layer (the mirror), coupled to the physical one.

Mirroring is given by the fact that physical things, which can be perceived and acted upon by humans in the physical world, have a digital counterpart (or augmentation, extension) in the mirror, so that they can be observed and acted upon by agents. Viceversa, an entity (artifact) in the MW that can be perceived and acted upon by software agents may have a physical appearance (or extension) in the physical world – e.g. augmenting it, in terms of AR – so that it can be observed and acted upon by humans (by means of e.g. smart-glasses). This implies a form of *coupling*, such that an action on an object in the physical world causes some kind of changes in entities in the mirror, perceivable then by software agents. Viceversa an action by agents on an artifact in the MW can have an effect on things in the physical world, perceivable by people.

In this case, the abstraction support level provided by the agent environment is twofold: from the artificial agents point of view, it provides a way to represent, perceive and interact with physical things, represented and abstracted by artifacts; from the human agents point of view, the virtual environment provides a way to *augment* the physical world with further functionalities, as well as to empower humans with further cognitive/sensing/acting capabilities. The interaction mediation support level in this case allows for designing environment-meditated coordination and cooperative strategies – possibly self-organising,

emerging – that exploit both the physical and digital layer, towards new forms of *Behavioural-Implicit Communication* and stigmergy [3, 10].

4.5 Commonalities: Stigmergy Among Humans and Agents

In spite of the differences, in all three approaches it is possible to recognise the introduction of different levels of mechanisms that can be framed as *stigmergic*. At a lower level, these mechanisms make it possible to track, reify, and make it persistent and observable information about human/agent actions and their effect, at the semantic level. At a higher level, this information can be continuously aggregated, elaborated and distributed according to high-level rules designed with some social goal in mind, so as to support high-level and emergent forms of coordination and cooperation between human and artificial agents.

5 Humans Inside

There are situations where it is necessary that real-world humans become an integral part of the system and live together within the same virtual environment with intelligent agents. These systems require facing complex decisions for an agent environment as to how model humans, how to design a human-aware communication infrastructure, how to provide decision and co-ordination support, and how to implement laws. More precisely,

- *Real-world human modelling*: It is necessary to capture the representative human attributes for the system and translate them into an environment entity that corresponds to reality. These attributes need to be aligned and harmonized with the layers where artificial agents perform their duties.
- *Communication infrastructure*: Different types of artificial agents in the environment must be able to recognize human entities and have mechanisms to communicate with them. Thus, agents can sense the humans and get information for their decision making processes.
- *Decision and coordination support*: Endowing an environment with capabilities to assist humans in their interactions is a must. Decision and coordination support appear as fundamental environment features to help humans reduce the scope of their reasoning with the aim of achieving their goals.
- *Constraints and laws implementation*: The agent environment should contain rules that all entities and agents must fulfill so as to define both their existence and behavior with others. The environment and its agents, need to add new restrictions to know how to interact with humans.

Putting real-world humans in the loop entails having elements that can not be controlled, modified or even expected to have a rational behavior. Although the environment enforces its laws on the artificial agents, it is not possible to change a human's state or attributes. Therefore, one of the main questions of this position paper is: *can humans break the environment rules?*

It is now, more than ever, when the need for an explicit environment entity that ensures the integrity and consistence of the entire ecosystem becomes unavoidable. It is neccesary to seek consensus on the role the environment must play when unstable situations are given and how to cope with them. The environment needs to be prepared to cope with impossible situations without coming into faulty performance. All events and behaviours that take place inside, need to be decomposed into atomic actions that go from one consistent state to another. Not only that, but considering that the real world is receiving actions from the system too, even with impossible situations the environment must preserve the consistence of the ecosystem and avoid creating damage or harming any human.

6 Towards Environments where Humans and Agents are Happy to Live in – Open Challenges

The objective of this position paper was to raise some basic questions about an issue that we believe could be important for future research about environments for MAS.

By considering humans in the loop, we are pushed towards identifying first-class environments that provide effective means for enabling and interaction, coordination, cooperation not only among agents, but among humans and agents too. Open challenges in that path include:

- devising effective environment mechanisms effectively modelling human-agent interaction and coordination, paying particular attention to providing decision and coordination support;
- the introduction of high-level environment abstractions and mechanism to effectively model and design *augmented* realities and worlds where (human/software) agents live in;
- methodologies for designing and developing scalable agent environment technology stacks to deal with open, large-scale human-agent environments, eventually integrating mainstream architectures and technologies related to e.g. Internet-of-Things and the cloud.

Existing explorations and visions about sociotechnical urban superorganisms, mirror worlds or human-agent collectives can be considered just a starting point. We believe that in order to tackle these challenges, a broader inter-disciplinary perspective is needed, bringing together studies and results from both the human and the agent side, with the specific objective of designing environments where both humans and agents are happy to live in.

References

1. Arcos, J.L., Esteva, M., Noriega, P., Rodríguez-Aguilar, J.A., Sierra, C.: Engineering open environments with electronic institutions. Engineering applications of artificial intelligence 18(2), 191–204 (2005)

2. Bogdanovych, A.: Virtual Institutions. Ph.D. thesis, University of Technology, Sydney, Australia (2007)
3. Castelfranchi, C., Pezzullo, G., Tummolini, L.: Behavioral implicit communication (BIC): Communicating with smart environments via our practical behavior and its traces. International Journal of Ambient Computing and Intelligence 2(1), 1–12 (Jan–Mar 2010)
4. Castelfranchi, C., Piunti, M., Ricci, A., Tummolini, L.: AmI systems as agent-based mirror worlds: Bridging humans and agents through stigmergy. In: Bosse, T. (ed.) Agents and Ambient Intelligence, Ambient Intelligence and Smart Environments, vol. 12, pp. 17–31. IOS Press (2012)
5. Gelernter, D.: Mirror worlds - or the day software puts the universe in a shoebox: how it will happen and what it will mean. Oxford University Press (1992)
6. Jennings, N.R., Moreau, L., Nicholson, D., Ramchurn, S., Roberts, S., Rodden, T., Rogers, A.: Human-agent collectives. Commun. ACM 57(12), 80–88 (Nov 2014), http://doi.acm.org/10.1145/2629559
7. Messinger, P.R., Stroulia, E., Lyons, K., Bone, M., Niu, R.H., Smirnov, K., Perelgut, S.: Virtual worlds - past, present, and future: New directions in social computing. Decision Support Systems 47(3), 204–228 (Jun 2009)
8. Omicini, A., Ricci, A., Viroli, M.: Artifacts in the A&A meta-model for multi-agent systems. Autonomous Agents and Multi-Agent Systems 17(3), 432–456 (2008), http://dx.doi.org/10.1007/s10458-008-9053-x
9. Sassi, A., Zambonelli, F.: Coordination infrastructures for future smart social mobility services. IEEE Intelligent Systems 29(5), 78–82 (2014), http://dx.doi.org/10.1109/MIS.2014.81
10. Tummolini, L., Castelfranchi, C.: Trace signals: The meaning of stigmergy. In: Weyns, D., Parunak, H.V.D., Michel, F. (eds.) Environments for MultiAgent Systems III, LNAI, vol. 4389, pp. 141–156. Springer (May 2007), http://www.springerlink.com/content/680g127886431453/
11. Weyns, D., Omicini, A., Odell, J.: Environment as a first class abstraction in multiagent systems. Autonomous Agents and Multi-Agent Systems 14(1), 5–30 (2007), http://dx.doi.org/10.1007/s10458-006-0012-0
12. Zambonelli, F.: Toward sociotechnical urban superorganisms. IEEE Computer 45(8), 76–78 (2012), http://doi.ieeecomputersociety.org/10.1109/MC.2012.280
13. Zambonelli, F., Castelli, G., Ferrari, L., Mamei, M., Rosi, A., Serugendo, G.D.M., Risoldi, M., Tchao, A., Dobson, S., Stevenson, G., Ye, J., Nardini, E., Omicini, A., Montagna, S., Viroli, M., Ferscha, A., Maschek, S., Wally, B.: Self-aware pervasive service ecosystems. Procedia CS 7, 197–199 (2011), http://dx.doi.org/10.1016/j.procs.2011.09.006