Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Electronic Notes in Theoretical Computer Science 150 (2006) 21–36
Electronic Notes in Theoretical Computer Science
www.elsevier.com/locate/entcs

Agens Faber: Toward a Theory of Artefacts for MAS

Andrea Omicini[1] Alessandro Ricci[2] Mirko Viroli[3]

DEIS, Alma Mater Studiorum—Università di Bologna, via Venezia 52, 47023 Cesena, Italy

Abstract

Human intelligence has evolved along with the use of more and more sophisticated tools, allowing *Homo Faber* (from *Homo Habilis* to *Homo Sapiens Sapiens*) to cope with environment changes, as well as to adapt the environment to his needs. Analogously, in this seminal paper we introduce the notion of *Agens Faber*, conveying the idea that agent intelligence should not be considered as separated by the agent ability to perceive and affect the environment—and so, that agent intelligence is strictly related to the *artefacts* that enable, mediate and govern any agent (intelligent) activity.
Along this line, we first discuss the notion of *artefact for MAS* in general, then we try to devise out the admissible / required / desirable features of an artefact for MAS. We elaborate on the many sorts of possible relations between agents and artefacts, focusing in particular on the issue of the rational exploitation of artefacts, and also rough out a possible taxonomy of artefacts for MAS.

Keywords: Agent Intelligence, Multi-Agent Systems, Tools, Artefacts for MAS, Activity Theory.

1 Tools, Language, and Artefacts

1.1 The Language Hype

"In the past, the major impetus for work on this general question has presented the ... capacity for language as the explicandum, whereas the tool-making and tool-using faculty has been taken for granted as an expectable attribute ..., and therefore requiring less scientific examination. This may

[1] Email: andrea.omicini@unibo.it
[2] Email: a.ricci@unibo.it
[3] Email: mirko.viroli@unibo.it

doi:10.1016/j.entcs.2006.03.003

> reflect a deep logocentric philosophical bias on the part of western scholars, and one that has had a very long history. Now that this bias has been exposed for what it is, a whole new field of questions has been opened up ..." [6]

For a scholar in the agent field today, the sentence above could easily fit the current, overall status of research on MASs (Multi-Agent Systems). Actually, and quite interestingly, the citation comes instead from a well-respected anthropologist, late Gordon H. Hewes, studying and comparing the origins and evolution of the two most distinctive faculties of humans: the ability to speak, and the ability to use and make tools. Hewes observes how, after many decades of activities by researchers from many diverse but related areas (such as biological and social anthropology, archaeology, linguistics, psychology, neurology, and ethology), only in very recent times the issue of the relation between language, tools and the evolution of human cognitive abilities has been faced as a single, coherent problem (see for instance [5]).

More or less in the same way, a sort of hype toward language is today still quite evident in the MAS field. Apart from the overwhelming number of agent papers on communication, speech acts, and other language-related issues published in the last years by agent-related conferences and journals, a striking evidence comes for instance from the work by the only agent standardisation body, FIPA—dealing more or less exclusively with agent *communication* actions, in spite of its being the Foundation for Intelligent *Physical* Agents.

In order to avoid (further) pernicious effects of the same "logocentric philosophical bias" in the MAS field, some philosophical work is then required, aimed at providing a shared conceptual view where agent actions could be framed in their most general acceptation, actually accounting for both communicative and physical actions. Then, a notion of agent *tool* (or, an equivalent one) is required, which could allow a theory of agent physical action to be developed at the same level of refinement as the theory of agent communicative actions.

1.2 A General View over Tools and Language

The first characterisation of *Homo Habilis* is his ability to forge tools. Tools are not an exclusive feature of humans: beavers build dams, bees construct perfect hexagonal cells, many birds live in self-made nests. However, what is often taken as a distinguishing feature of human tools with respect to other animals' ones is the cognitive level at which the tools are conceived, designed, and used: apparently, tools are not part of human "embedded" behaviour, as in bees or birds, but they are rather the result of the rational elaboration about the relationship between the human being and his habitat—his living environment. Also, systematic and social design and use of tools is seemingly typical of the human species, and is often taken as a measure of human against animal intelligence.

More generally, our understanding of the strict relationship between tools and intelligence (human and not human) is such that we typically interpret tool-using and tool-making faculty as a fundamental revealing symptom of intelligence. For instance, ethologists commonly measure intelligence of animals by making them face problems that require the use of tools to be solved (see for instance [13]). Even more interestingly, a sort of tool-equivalent of the Turing test has been proposed by philosopher Ronald Endicott, which was aimed at evaluating intelligence in terms of the ability to exploit tools—the so-called "Tooling Test for Intelligence" [16].

A tool, according to Martelet [8], reveals the user awareness of self or/and of the world, whenever it is built with a goal, it is stored for further / repeated use, it is used for building new tools. Tools are at the same time the first and most distinctive expression of human intelligence, along with language; and also, the most powerful amplifiers of the (both individual and social) human ability to affect the environment—to survive environment change, first, and to change the environment for the human purposes, then.

Thus, when dealing with agent systems, and agent intelligence in particular, it turns to be awkwardly strange that most of the work till now has elaborated on linguistic concepts and acts—as in the agent-pervasive theory of speech acts—and has almost ignored, at least explicitly, the matter of agent tools.

Less surprising, instead, is the fact that social / organisational theories have well emphasised the role of tools in human activity, and their strict relation with language as well. For instance, central to Activity Theory (AT [9]) is the notion of *artefact* as a mediator for any sort of interaction in human activities: artefacts can be either physical (such as blackboards, walls, and traffic lights) or cognitive tools (such as operating procedures, heuristics, scripts, individual and collective experiences, and languages). Artefacts embody a set of social practise: their design reflects a history of particular use. As mediating instruments, they have both an *enabling* and a *constraining* function: on the one hand, artefacts expand out possibilities to manipulate and transform different objects, but on the other hand the object is perceived and manipulated not as such but within the limitations set by the tool.

So, on the one hand, artefacts are a quite general and powerful notion, encompassing both physical and psychological tools—thus allowing physical tools and language to be understood within the same conceptual framework.

On the other hand, AT obviously suggests that artefacts could be used as fundamental abstractions in the construction of agent systems—mostly as enabling and constraining tools for agent activities. In the following then, we generally talk about *artefacts for MAS*, or simply artefacts, so as to include the most general notion of tool available.

1.3 Agens Faber vs. Agens Symbolicus?

Even the development of some of the most apparently "abstract" results of human culture cannot be conceived or explained without the accompanying physical artefacts. For instance, the evolution of numbers and arithmetic cannot be explained separately by the tools used to record and compute them, and by the social processes in which they were used in the ancient human societies where they first developed. The apparent duality between Homo Faber or Homo Symbolicus [1]—who comes first?—is obviously to be solved without a winner: then, why should we choose between an Agens Faber and an Agens Symbolicus while we aim at intelligent agents?

Accordingly, adopting an evolutionary perspective over agents, and carrying along the analogy with the development and evolution of human intelligence, we claim here that a theory of agent intelligence should not be limited to the modelling of inner, rational processes (as in BDI theory), and should instead include not only the basics of practical reasoning, but also a suitable theory of the artefacts and the means for their rational use, selection, construction and manipulation. This is in fact the idea behind of the *Agens Faber* notion: agent intelligence should not be considered as separated by the agent ability to perceive and affect the environment—and so, that agent intelligence is strictly related to the artefacts that enable, mediate and govern any agent (intelligent) activity: in a sense, a specialised, agent-oriented version of Brooks' *situated intelligence* [2].

Along this line, in the remaining of this seminal paper we first collect some considerations of ours about the conceptual relation between agents and artefacts, and the rational exploitation of artefacts by agents (Section 2). Then, Section 3 sketches a possible model for artefacts for MAS (introduced in [15]), then lists and groups a number of features that artefacts could exhibit in order to enable and promote agent intelligent behaviour. Section 4 brings the seeds for a more comprehensive theory of artefacts for MAS, by roughing out a possible taxonomy of artefacts. Finally, Section 5 provides for final remarks along with some lines of future work.

2 On the Relation between Agents and Artefacts

2.1 Goals of Agents and Use of Artefacts

By considering the conceptual framework described in [3], agents can be generally conceived as *goal-governed* or *goal-oriented* system. Goal-governed systems refer to the strong notion of agency, i.e. agents with some forms of cognitive capabilities, which make it possible to explicitly represent their goals, driving the selection of agent actions. Goal-oriented systems refer to the weak notion of agency, i.e. agents whose behaviour is directly designed and programmed to achieve some goal, which is not explicitly represented. In both goal-governed and goal-oriented systems, goals are *internal*. *External goals* instead refer to goals which typically belong of the social context or environment where the agents are situated. External goals are sorts of regulatory states which condition agent behaviour: a goal-governed system follows external goals by adjusting internal ones.

This basic picture is then completed by systems which are not goal-oriented. This is the case of passive objects, which are characterised by the concept of *use*: they have not internal goals, but can be *used* by agents to achieve their goals. *Artefacts* are objects explicitly designed to provide a certain *function* [4], which guides their use. The concept of *destination* is related but not identical the concept of function: it is an external goal which can be attached to an object or an artefact by users, in the act of using it. Then an artefact can be used according to a destination which is different from its function.

An interesting distinction has been proposed, concerning agents / artefacts relationships, between *use* and *use value* [3]: there, use value corresponds to the evaluation of artefact characteristics and function, in order to *select* it for a (future) use. The distinction corresponds to two different kinds of external goals attached to an artefact: *(i)* the use-value goal, according to which the artefact should allow user agents to achieve their objective—such an external goal drives the agent selection of the artefact; *(ii)* the use goal, which directly corresponds to the agent internal goal, which guides the actual usage of the artefact. From the agent point of view, when an artefact is selected and used it has then a use-value goal that somehow matches its internal goal.

By extending the above considerations, the classical tool-using / tool-making distinction from anthropology can be articulated along three main distinct aspects, which characterise the relationship between agents and arte-

[4] The term "function" here refers to the functionality embodied by an artefact, and should not be confused with the same term as used e.g. in mathematics or in programming languages

facts:

- use
- selection
- construction and manipulation

While the first two aspects are clearly related to use and use value, respectively, the third is the rational consequence of a failure in the artefact selection process, or in the use of a selected artefact. Then, a new, different artefact should be constructed, or obtained by manipulation of an existing one.

2.2 Agents Reasoning about Artefacts

One of the key issues of in the Agens Faber approach is how artefacts can be effectively exploited to improve agent ability to achieve individual as well as social goals. The main questions to be answered are then: How should agents reason to use artefacts in the best way, making their life simpler and their action more effective? How can agents reason to select artefacts to use? How can agents reason to construct or adapt artefact behaviour in order to fit their goals?

On the one hand, the simplest case concerns agents directly programmed to use specific artefacts, with usage protocols directly defined by the programmer either as part of the procedural knowledge / plans of the agent for goal-governed systems, or as part of agent behaviour in goal-oriented system. In spite of its simplicity, this case can bring several advantages for MAS engineers, exploiting separation of concerns for programming simpler agents, by charging some burden upon specifically-designed artefacts. On the other hand, the intuition is that in the case of fully-open systems, the capability of the artefact to describe itself, its function, interface, structure and behaviour could be the key for building open MASs where intelligent agents dynamically look for and select artefacts to use, and then exploit them for their own goals.

At a first glance, it seems possible to frame the agent ability to use artefacts in a hierarchy, according to five different cognitive levels at which the agent can use an artefact:

unaware use — at this level, both agents and agent designers exploit artefacts without being aware of it: the artefact is used implicitly, since it is not denoted explicitly. In other words, the representation of agent actions never refer explicitly to the execution of operation on some kind of artefacts.

embedded / programmed use — at this level, agents use some artefacts according to what has been explicitly programmed by the designer: so, the artefact selection is explicitly made by the designer, and the knowledge about its use is implicitly encoded by the designer in the agent. In the case of cognitive agents, for instance, agent designers can specify usage protocols directly as part of the agent plan. From the agent point of view, there is no need to understand explicitly artefact operating instructions or function: the only requirement is that the agent model adopted could be expressive enough to model in some way the execution of external actions and the perception of external events.

cognitive use — at this level, the agent designer directly embeds in the agent knowledge about what artefacts to use, but how to exploit the artefacts is dynamically discovered by the agent, reading the operating instructions. Artefact selection is still a designer affair, while how to use it is delegated to the agent rational capabilities. So, generally speaking the agent must be able to discover the artefact function, and the way to use it and to make it fit the agent goals. An obvious way to enable agent discovery is to make artefact explicitly represent their function, interface, structure and behaviour.

cognitive selection and use — at this level, agents autonomously select artefacts to use, understand how to make them work, and then use them: as a result, both artefact selection and use are in the hands of the agents. It is worth noting that such a selection process could also concern sets of cooperative agents, for instance interested in using a coordination artefact for their social activities.

construction and manipulation — at this level, agents are lifted up to the role of designers of artefacts. Here, agents are supposed to understand how artefacts work, and how to adapt their behaviour (or to build new ones from scratch) in order to devise out a better course of actions toward the agent goals. For its complexity, this level more often concerns humans: however, not-so-complex agents can be adopted to change artefact behaviour according to some schema explicitly pre-defined by the agent designers.

3 Features of an Artefact for MAS

By extending to artefacts for MAS in general the formal model for *coordination artefacts* defined in [15] (and reported in brief in Subsection 3.1), we discuss an essential feature of the Agens Faber approach: that is, a model for agent reasoning where agent mental states and artefact behaviour are both accounted for in a coherent way (Subsection 3.2). The notion of Agens Faber, then, despite its apparent fuzziness, implicitly calls for a number of further desirable artefact features, which are shortly discussed in Subsection 3.3.

3.1 A Model of Artefacts for MAS

In order to allow for its rational exploitation by intelligent agents, an artefact for MAS possibly exposes *(i)* a *usage interface*, *(ii) operating instructions*, and *(iii)* a *service description*. On the one hand, this view of artefacts provides us with a powerful key for the interpretation of the properties and features of existing non-agent MAS abstractions, which can be then catalogued and compared based on some common criteria. On the other hand, it is also meant to foster the conceptual grounding for a principled methodology for the engineering of MAS environment, where artefacts play the role of the core abstractions.

Usage Interface — One of the core differences between artefacts and agents, as computational entities populating a MAS, lays in the concept of *operation*, which is the means by which an artefact provides for a service or function. An agent executes an action over an artefact by invoking an artefact operation. Execution possibly terminates with an *operation completion*, typically representing the outcome of the invocation, which the agent comes to be aware of in terms of perception. The set of operations provided by an artefact defines what is called its *usage interface*, which (intentionally) resembles interfaces of services, components or objects—in the object-oriented acceptation of the term.

In MASs, this interaction schema is peculiar to artefacts, and makes them intrinsically different from agents. While an agent has no interface, acts and senses the environment, encapsulates its control, and brings about its goals proactively and autonomously, an artefact has instead a usage interface, is used by agents (and never the opposite), is driven by their control, and automatises a specific service in a predictable way without the blessing of autonomy. Hence, owning an interface strongly clearly differentiates agents and artefacts, and is therefore to be used by the MAS engineer as basic discriminative property between them.

Operating Instructions — Coupled with a usage interface, an artefact could provide agents with *operating instructions*. Operating instructions are a description of the procedure an agent has to follow to meaningfully interact with an artefact over time. Most remarkably, one such description is history dependent, so that actions and perceptions occurred at a given time may influence the remainder of the interaction with the artefact. Therefore, operating instructions are basically seen as an exploitation protocol of actions / perceptions. This protocol is possibly furthermore annotated with information on the intended preconditions and effects on the agent mental state, which a rational agent should read and exploit to give a meaning to operating instructions. Artefacts being conceptually similar to devices used by humans, operation instructions play a role similar to a manual, which a human reads to know how to use the device on a step-by-step basis, and depending on the expected outcomes he/she needs to achieve. For instance, a digital camera provides buttons and panels (representing its usage interface), and therefore comes with a manual describing how to use them—e.g. which sequence of buttons are to be pushed to suitably configure the camera resolution.

Function Description — Finally, an artefact could be characterised by a *function description* (or service description). This is a description of the functionality provided by the artefact, which agents can use essentially for artefact selection. In fact, differently from operating instructions, which describes *how* to exploit an artefact, function description describes *what* to obtain from an artefact. Clearly, function description is an abstraction over the actual implementation of the artefact: it hides inessential details over the implementation of the service while highlighting key functional (input/output) aspects of it, to be used by agents for artefact selection. For instance, when modelling a sensor wrapper as an artefact, we may easily think of the operations for sensor activation and inspection as described via usage interface and operations instructions, while the information about the sensory function itself being conveyed through function description of the sensor wrapper.

3.2 Rational Exploitation of Artefacts

One of the key issues in the cognitive use of artefacts is how agents represent actions over artefacts and include them in their deliberate course of actions. The issue of rational exploitation of artefacts can be better formulated and understood by formally modelling the relationship between agents and artefacts: agent actions over artefacts, agent mental states, and artefact operating instructions. Following the model sketched above and introduced in [15], here we assume a model for the mental state of agents similar to the one exploited in the formal definition of the semantic language for defining FIPA ACL semantics [4]. The agent mental state is represented through a multi-modal logic with modalities for beliefs (Bel) and intentions (Int)—as far as exploiting an artefact is concerned, we neglect modalities for uncertain beliefs and desires. Hence, we consider the standard syntax for formulae

$$\phi ::= t \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid \phi \Rightarrow \phi \mid \mathsf{Bel}\,\phi \mid \mathsf{Int}\,\phi$$

where t is any term. To enable reasoning about actions and perceptions, we introduce the following operators (written in italics font), which are functions used to build terms: $\mathit{done}(a)$ stands for action a—i.e. an operation invocation—being executed, $\mathit{schedule}(a)$ stands for action a being scheduled, $\mathit{completed}(p)$ for perception of completion p—i.e. the completion of an operation invocation—being sensed, $\mathit{instr}(I)$ for I being the current state of operating instructions, $\mathit{trn}(I, l, I')$ for instructions I allowing interaction l and then moving to I', and $\mathit{feasible}(I, \phi)$ for mental state ϕ being a possible effect reachable from current operating instructions state I. In particular, we assume the case where the agent *(i)* is interacting with only one artefact, *(ii)* knows the current state of operating instructions ($\mathsf{Bel}\,\mathit{instr}(I)$ holds for precisely one I at a given time), *(iii)* understands the operational semantics of operating instructions ($\mathsf{Bel}\,\mathit{trn}(I, l, I')$ holds if $I \xrightarrow{l}_{\mathcal{I}} I'$), and *(iv)* computes the feasibility operator based on the operational semantics, that is through relation trn above. For instance, relation $\mathit{feasible}(I, \phi)$ could be defined as:

$$\begin{array}{lll} \mathit{feasible}(I, \phi) & \text{if} & I \xrightarrow{?t_{[\phi]}}_{\mathcal{I}} I' \\ \mathit{feasible}(I, \phi) & \text{if} & I \xrightarrow{l}_{\mathcal{I}} I' \quad \wedge \quad \mathit{feasible}(I', \phi) \end{array}$$

saying that ϕ is feasible from I if it exists a sequence of interactions leading to a completion with effect ϕ. This definition is to be taken as a mere reference: more specialised feasibility relations might be used which more effectively tackle computability and efficiency issues—whose details have however the scope of agent internal machinery, not that of interaction semantics.

Concerning the effect of interactions on the agent mental state, we simply assume that:

- as an agent schedules an action a—that is, $\mathsf{Int}\,\mathit{schedule}(a)$ holds in the agent mental state—then a is executed—the corresponding artefact operation is actually invoked. Correspondingly, the agent comes to believe $\mathsf{Bel}\,\mathit{done}(a)$.
- an an agent perceives a completion p, then it comes to believe $\mathsf{Bel}\,\mathit{completed}(p)$.

The formal model of agent behaviour is provided by reporting properties of the logic for mental states in the style of the properties of SL language of FIPA [4], handling planning, satisfiability of preconditions and application of effects.

Planning — Given an intention $\mathsf{Int}\,\phi$ and current operating instructions I, let $a_1, \ldots, a_n$ be the set of invocations possibly leading to effect ϕ, that is such that $\mathit{trn}(I, !a_{j_{[\phi_j]}}, I')$ and $\mathit{feasible}(I', \phi)$ (with $1 \leq j \leq n$)—i.e., a_j is allowed resulting in a state (I') where ϕ is feasible. Then it holds that

$$\models \mathsf{Int}\,\phi \Rightarrow \mathsf{Int}\,done(a_1|\ldots|a_n)$$

which says that the agent intention to achieve ϕ generates the intention to execute one of the allowed actions possibly leading to effect ϕ.

Satisfiability — Whenever an agent elects to perform an action (by the planning property above), this property imposes the intention to seek satisfiability of the preconditions of this action. This is formalised by the property

$$\models \mathsf{Int}\,done(a) \wedge \mathsf{Bel}\,instr(I) \wedge \mathsf{Bel}\,trn(I, !a_{[\phi]}, I') \Rightarrow \mathsf{Int}\,\phi$$

Scheduling — Whenever an agent intends to execute an action whose preconditions are satisfied, then it schedules that action, suitably updating operating instructions. This idea is formalised by the property:

$$\models \mathsf{Int}\,done(a) \wedge \mathsf{Bel}\,instr(I) \wedge \mathsf{Bel}\,trn(I, !a_{[\phi]}, I') \wedge \mathsf{Bel}\,\phi \Rightarrow$$
$$\mathsf{Bel}\,instr(I') \wedge \mathsf{Int}\,schedule(a)$$

Mental effect — Whenever an agent perceives the completion of an action, it should believe the effects of such a perception. This idea is formalised by the property

$$\models \mathsf{Bel}\,completed(p) \wedge \mathsf{Bel}\,instr(I) \wedge \mathsf{Bel}\,trn(I, ?p_{[\phi]}, I') \Rightarrow \mathsf{Bel}\,\phi \wedge \mathsf{Bel}\,instr(I')$$

These properties resemble in structure the five properties of SL language as they appear in [4]. The essential difference, however, is that here preconditions and effects are bound to actions/perceptions by the operating instructions: so, the interplay between agent mental states and interactions (agent actions over the artefacts) *makes into the logic*, by means of the scheduling property.

3.3 Other Properties of Artefacts for MAS

Furthermore, artefacts can exhibit a number of other important features, which possibly enhance agent ability to use them for their own purposes.

Inspectability — The state of an artefact, its content (whatever this means in a specific artefact), its usage interface, operating instructions and function description might be all or partially available to agents through *inspectability*. Whereas in closed MASs this information could be hard-coded in the agent—the artefact engineer develops the agents as well—, in open MASs third-party agents should be able to dynamically join a society and get aware at run-time of the necessary information about the available artefacts. Also, artefacts are often in charge of critical MAS behaviour [10]: being able to inspect a part or the whole of an artefact features and state is likely to be a fundamental capability in order to understand and govern the dynamics and behaviour of a MAS.

Controllability — Controllability is an obvious extension of the inspectability property. The operational behaviour of an artefact should then not be merely inspectable, but also controllable so as to allow engineers (or even intelligent agents) to monitor its proper functioning: it should be possible to stop and restart an artefact working cycle, to trace its inner activity, and to observe and control a step-by-step execution. In principle, this would largely improve the ability of monitoring, analysing and debugging at execution time the operational behaviour of an artifact, and of the associated MAS social activities as well.

Malleability — Also related to inspectability, malleability (also called *forgeability*) is a key-feature in dynamic MAS scenarios, when the behaviour of artefacts could require to be modified dynamically in order to adapt to the changing needs or mutable external conditions of a MAS. Malleability, as the ability to change the artefact behaviour at execution-time, is seemingly a crucial aspect in on-line engineering for MASs, and also a perspective key issue for self-organising MASs.

Predictability — Differently from agents—which as autonomous entities have the freedom of behaving erratically, e.g. neglecting messages—, usage interface, operating instructions and function description can be used as a contract with an artefact by an agent. In particular, function description can provide precise details of the outcomes of exploiting the artefact, while operating instructions make the behaviour of an artefact predictable for an agent.

Formalisability — The predictability feature can be easily related with *formalisability*. Due to the precise characterisation that can be given to an artefact behaviour, until reaching e.g. a full operational semantics model—for instance, as developed for coordination artefacts in [12]—it might be feasible to automatically verify the properties and behaviour of the services provided by artefacts, for this is intrinsically easier than services provided by autonomous agents.

Linkability — Artefacts can be used encapsulate and model reusable services in a MAS. To scale up with complexity of an environment, it might be interesting to compose artefacts, e.g. to build a service incrementally on top of another, by making a new artefact realising its service by interacting with an existing artefact. To this end, artefacts should be able to invoke the operation of another artefact: the reply to that invocation will be transmitted by the receiver through the invocation of another operation in the sender.

Distribution — Differently from an agent, which is typically seen as a point-like abstraction conceptually located to a single node of the newtwork, artefacts can also be distributed. In particular, a single artefact can in principle be used to model a distributed service, accessible from more nodes of the net. Using linkability, a distributed artefact can then be conceived and implemented as a composition of linked, possibly non-distributed artefacts—or viceversa, a number of linked artefacts, scattered through a number of different physical locations could be altogether seen as a single distributed artefact. Altogether, distribution and linkability promote the layering of artefact engineering—as sketched in Section 4.

As a final remark, it should be noted that all the artefact features presented above play a different role when seen from the different viewpoints of agents and of MAS engineers. For instance, operating instructions are mostly to be seen as a design tool for engineers, as well as a run-time support for rational agents. Instead, features like inspectability and malleability gain particular interest when the two viewpoints can be made one: when an intelligent agent is allowed and capable to play the role of the MAS engineer, it can in principle understand the state and dynamics of the MAS by observing the artefacts, then possibly working as an Agens Faber: that is, by re-working its tools (the artefacts) in order to suitably change the overall MAS behaviour.

4 Toward a Taxonomy of Artefacts for MAS

Many sorts of different artefacts can populate a MAS, providing agents with a number of different services, embodying a variety of diverse models, technologies and tools, and addressing a wide range of application issues. Correspondingly, a huge variety of approaches and solutions are in principle available for MAS engineers when working to shape the agent environment according to their application needs. So, the mere model of artefacts for MAS is no longer enough: a *taxonomy* of artefacts comes to be useful, which could help MAS engineers first defining the basic classes of artefacts, their differences and peculiarities, then classifying known artefacts, to understand and possibly compare them.

Among the many possible criteria for a taxonomy, we find it useful to focus on the mediation role of the artefact, and then discriminate artefacts based on the sort of the (non-artefact) MAS entities they are meant to tie

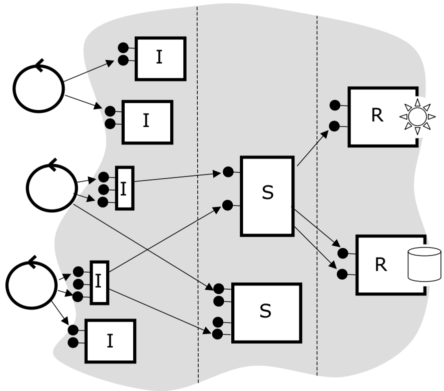

Fig. 1. Individual, social, and resource artefacts: a layered view of artefacts for MAS

together. According to the pictorial representation in Fig. 1, our first proposal here divides artefacts into *individual artefacts*, *social artefacts*, and *resource artefacts*.

Individual artefacts are artefacts exploited by one agent only—in other terms, an individual artefact mediates between an individual agent and the environment. Individual artefacts can serve several purposes, including externally enhancing agent capabilities—such as e.g. adding a private external memory—, enacting a filtering policy of the agent actions toward other artefacts (as in the case of agent coordination contexts [11]), providing individual agents with useful information on the organisation, and so on. In general, individual artefacts are not directly affected by the activity of other agents, but can, through linkability, interact with other artefacts in the MAS.

Social artefacts are instead artefact exploited by more than one agent—in other terms, a social artefact mediate between two or more agents in a MAS. In general, social artefacts typically provide a service which is in the first place meant to achieve a social goal of the MAS, rather than an individual agent goal. For instance, social artefacts might provide a coordination service [14], governing the activities of two or more agents—as e.g. in a multi-party protocols—, but can also realise global knowledge repositories, shared ontologies, or organisation abstractions containing information on roles and permissions.

Finally, resource artefacts are artefacts that conceptually wrap external resources—in other terms, a resource artefact mediates between a MAS and an external resource. External resources can be either legacy components and tools, applications written with non-agent technologies because of engineering convenience—such as Web Services—, or physical resources which the agents of a MAS might need to act upon and sense. In principle, resource artefacts can be conceived as a means to raise external MAS resources up to the agent cognitive level. In fact, they provide external resources with an usage interface, some operating instructions, and a service description, and realise their task by dynamically mapping high-level agent interactions upon lower-level interactions with the resources—using e.g. specific transports such as object-oriented local or remote method calls, HTTP requests, and the like.

Altogether, individual, social and resource artefacts can be used as the basis for building the glue keeping agents together in a MAS, and for structuring the environment where agents live and interact. In fact, our taxonomy, as apparent from Fig. 1, defines a structured, *layered* view over the MAS environment, and implicitly suggests a model for organising agent interaction within a MAS. As such, the artefact taxonomy could lead in principle to a well-principle foundation for a general agent-oriented methodology for the engineering of the agent environment as a first-class entity.

5 Conclusion

By drawing an analogy between intelligence in MASs and the development and evolution of human intelligence, in this seminal paper we elaborated on agent tools, and their relationship with agent intelligence. Our notion of Agens Faber comes to say that a theory of agent intelligence should not be limited to modelling the inner rational process of an agent, but should instead include not only the basics of practical reasoning, but also a theory of the agent *artefacts*, providing agents with the means for artefact rational use, selection, construction, and manipulation. Along this line, we first discuss a possible model of artefacts for MAS, and show how it affects the exploitation of artefacts by rational agents, then we list a set of relevant artefact features. Finally, a preliminary proposal for a taxonomy of artefacts for MAS is shortly discussed.

In the future, we plan to go deep along the many lines sketched in this paper: for instance, by accounting for the many mental processes that are known to be relevant in the social use of tools (such as teaching and imitation), or by discussing the possible impact of a precise and comprehensive artefact taxonomy on the engineering of complex MASs.

References

[1] K. Berggren. Homo Faber or Homo Symbolicus? The fascination with copper in the sixth millennium. *Transoxiana: Journal Libre de Estudios Orientales*, 8, June 2004.

[2] R. A. Brooks. Intelligence without representation. *Artificial Intelligence*, 47(1-3):139–159, 1991.

[3] R. Conte and C. Castelfranchi, editors. *Cognitive and Social Action*. University College London, 1995.

[4] FIPA. FIPA communicative act library specification. `http://www.fipa.org`, 2000.

[5] K. R. Gibson and T. Ingold, editors. *Tools, Language & Cognition in Human Evolution*. Cambridge University Press, 1993.

[6] G. W. Hewes. A history of speculation on the relation between tools and languages. In Gibson and Ingold [5], pages 20–31.

[7] N. R. Jennings, C. Sierra, L. Sonenberg, and M. Tambe, editors. *3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004)*. ACM, New York, NY, USA, 19–23 July 2004.

[8] G. Martelet. *Évolution et création, tome 1*. Editions du Cerf, Paris, 1998.

[9] B. A. Nardi. *Context and Consciousness: Activity Theory and Human-Computer Interaction*. MIT Press, 1996.

[10] A. Omicini, S. Ossowski, and A. Ricci. Coordination infrastructures in the engineering of multiagent systems. In F. Bergenti, M.-P. Gleizes, and F. Zambonelli, editors, *Methodologies and Software Engineering for Agent Systems: The Agent-Oriented Software Engineering Handbook*, chapter 14, pages 273–296. Kluwer Academic Publishers, June 2004.

[11] A. Omicini, A. Ricci, and M. Viroli. Formal specification and enactment of security policies through Agent Coordination Contexts. *Electronic Notes in Theoretical Computer Science*, 85(3), Aug. 2003. 1st International Workshop "Security Issues in Coordination Models, Languages and Systems" (SecCo 2003), Eindhoven, The Netherlands, 28–29 June 2003. Proceedings.

[12] A. Omicini, A. Ricci, M. Viroli, C. Castelfranchi, and L. Tummolini. Coordination artifacts: Environment-based coordination for intelligent agents. In Jennings et al. [7], pages 286–293.

[13] D. J. Povinelli. *Folk Physics for Apes: The Chimpanzee's Theory of How the World Works*. Oxford University Press, 2000.

[14] M. Viroli and A. Omicini. Coordination as a service: Ontological and formal foundation. *Electronic Notes in Theoretical Computer Science*, 68(3):457–482, Mar. 2003. 1st International Workshop "Foundations of Coordination Languages and Software Architecture" (FOCLASA 2002), Brno, Czech Republic, 24 Aug. 2002. Proceedings.

[15] M. Viroli and A. Ricci. Instructions-based semantics of agent mediated interaction. In Jennings et al. [7], pages 102–109.

[16] A. B. Wood, T. E. Horton, and R. S. Amant. Effective tool use in a habile agent. In E. J. Bass, editor, *2005 IEEE Systems and Information Engineering Design Symposium (SEADS 2005)*, pages 75–81, Charlottesville, VA, USA, 29 Apr. 2005. IEEE.