Multiagent Systems without Agents—*Mirror-Holons* for the Compilation and Enactment of Communication Structures*

Matthias Nickles and Gerhard Weiss

AI/Cognition Group, Department of Informatics, Technical University of Munich
D-85748 Garching bei München, Germany,
{nickles,weissg}@model.informatik.tu-muenchen.de

Abstract. It is widely accepted in Distributed Artificial Intelligence that a crucial property of artificial agents is their *autonomy*. Whereas agent autonomy enables features of agent-based applications like flexibility, robustness and emergence of novel solutions, autonomy might be also the reason for undesired or even chaotic agent behavior, and unmanageable system complexity. As a conceptual approach to the solution for this "autonomy dilemma" of agent-based software engineering, this work introduces the *HolOMAS* framework for open multiagent systems based on special meta-agents, so-called *Mirror-Holons*. Instead of restricting agent autonomy by means of normative constraints and defined organizational structures as usual, Mirror-Holons allow for the gradual *uncoupling* of agent interaction and *emergent* system functionality. Their main purpose is the derivation and adaption of *social structure knowledge* and evolving stochastical *social programs* from the observation and compilation of agent communication and additional design objectives. Social programs can either be executed by the Mirror-Holons themselves, or communicated to the agents and the system designer, similar to the functionality of mass media like television or newspapers in human societies.

Keywords: Multiagent Systems, Holons, Agent Communication, Cybernetics, Artificial Sociality, Autonomous Computing, Multiagent Coordination Media

1 Introduction

In [8, 9, 12], a novel approach to the design and control of open systems with truly autonomous agents has been introduced, which aimed at the establishment of mechanisms for autonomy-preserving self-control of the system by means of the *reflection* and propagation of social *expectation structures*. The main component of this architecture is the so-called *Social System Mirror*, a MAS-middleware component which continuously observes agent communications, derives generalized social structures from these observations (plus additional normative design objectives if required), and communicates ("reflects") these structures back to the agents. Leaning on *Social Systems Theory* [2, 8], interaction structures like organizational structures [4], norms and agent

* This work is supported by Deutsche Forschungsgemeinschaft (DFG) under contracts no. Br609/11-2 and MA759/4-2.

roles are always the regularities (*structures*) of communication processes which have to be represented as adaptive and normative *expectations* [8, 12] regarding the continuations of these processes. The goals of a Social System Mirror (or "mirror" for short) are the indirect, autonomy-preserving influencing of agent behavior by means of the system-wide propagation of social structures towards quicker structure evolution and higher coherence of communication structures without restricting agent autonomy, and the provision of an evolutionary model of social structures for the MAS designer. While a Social System Mirror models a single communication system and remains (apart from the propagation of expectations) passively, the successor architecture *HolOMAS*, which we introduce in this work, is able to model multiple communication systems at the same time through multiple *Mirror-Holons* in order to model large, heterogeneous systems. In addition, Mirror-Holons can take action themselves by means of the execution of emergent social programs which are generated from expectation structures.

A Mirror-Holon can be characterized informally as a

higher-order agent which "impersonates" an entire distributed, social program via the synchronous or asynchronous execution of extrapolated multi-agent interaction trajectories learned and revised i.a. by observation of multi-agent interactions.

Other agents, including other Mirror-Holons, can (besides their contribution of learning examples via their communications) optionally be involved in this execution process as effectors, which execute commands in social programs in their respective environmental domain, if they do not prefer to deny the respective command. In any case they can influence the social programs accommodated by the Mirror-Holon through their communication with peer agents. In addition, a Mirror-Holon can use given structures in addition to learned structures also (e.g., norms and protocols).

Since Mirror-Holons are agents that in some sense comprise "lower-level" agents and can be comprised recursively by higher-level, similarly constructed Mirror-Holons themselves recursively (forming a so-called *holarchie*), HolOMAS is strongly related to the *Holon* concept [27]. Nevertheless, there is an important difference between Mirror-Holons and traditional agent holons (e.g. [24]): A Mirror-Holon does not contain sets of *agents*, but instead actively represents a certain *social functionality* which is identified in form of regularities in the observed communications, without disregarding the autonomy of his adjoint lower-order actors. This allows for a flexible, more or less loose coupling of desired system functionality and lower-order agent behavior (although a governing of lower-order agents by means of social norms and sanctions is also optionally possible).
Since the holon concept of HolOMAS is based on the observation of agent interactions which can be used to coordinate agents behavior in turn, HolOMAS is also related to the theory of *coordination spaces* [25].

We expect that the concept of Mirror-Holons opens up prospects of autonomous software systems where on the one hand agent autonomy should not (or can not) be restricted, and on the other hand a fast, reliable system behavior is required (so to speak "real-time multiagent systems"). They are also expected to provide consistent, reliable and homogenous computational representations of open systems (e.g., virtual organizations) which otherwise can not guarantee such properties due to, e.g., internal conflicts and incoherencies.

The further sections of this paper are organized as follows: The next section introduces the basic concepts of Mirror-Holons. Section 3 outlines the central aspect of our framework, the empirical derivation of expectation structures (represented as so-called *expectation networks*), Section 4 describes how holon programs can be induced from empirically obtained social structures, and Section 5 outlines how multiple Mirror-Holons emerge and communicate. Finally, Section 6 points out open research problems and motivates future work.

2 Mirror-Holons

Since symbolic, deniable communications with a more or less indefinite result in terms of subsequent actions is the only way for truly autonomous agents to overcome their opaqueness, agent sociality can be modeled in terms of emergent, evolving expectation structures of communication processes *only* [8]. Because ultimatively the meaning of communications lies in their expected consequences, in [8, 9, 19, 12] we have therefore introduced expectation structures regarding communicative actions (therefore, sometimes called "communication structures") as a universal means for the modeling of social structures. Such (social) expectation structures integrate both normative expectations (expectations which describe how someone *should* behave) and adaptive expectations derived empirically from the actual behavior, which might be in open systems with a heterogeneous, fluctuating set of black-box-agents the only way to determine communication semantics. According to the concept of *autopoiesis* [2], expectation structures of social systems are more or less stable and reproduce themselves in order to provide a context for further communication despite the mental opaqueness of actors.
A formal framework for the representation of expectation structures can be found in [19] and (in a revised, abbreviated version, together with a learning algorithm) in [11].

A Mirror-Holon is a higher-order agent that comprises the behavioral spectrum of multiple lower-order yet intelligent and autonomous agents. To distinguish "ordinary" agency from such higher-order-agency, we introduce the taxonomy below of social structures in terms of the sort of agent communications that contribute to these structures. Social expectation structures model sets of communication processes, which are, in our usage of this term, sequences of elementary communications coupled by a relation called *communicative adjacency*. Communicative adjacency indicates that communication subsequent to another communication expresses implicitly or explicitly the understanding and referencing of the preceding communications. Communication processes can, from an observers perspective, be identified as trajectories of communication acts (especially utterances using some formal communication language). Since communication processes are the most elementary kind of observable sociality, we use them as the empirical evidence for the modeling of social systems [2], and in particular of so-called *spheres of communication* (cf. section 3).

Social expectation structures and the communications they are modeling can be given a (informal) hierarchy as follows:

First-order expectation structures are social expectation structures which describe the inner coherence and correlations of communication processes, but not their boundaries of validity. Unaware of latent information about the participating agents respectively their hidden intentions and goals, and if agents do not contradict themselves, first-order expectation structures can only describe expectations which are communicated explicitly, because agents aim for a consistent, justified and reliable communicational behavior towards other agents temporarily. First-order expectation structures are thus not able to model phenomena like insincerity or fraudulence as long as they are not communicated explicitly. Typical examples for first-order expectation structures are *spheres of commitment* [7] and closed-system structures like simple virtual organizations. E.g., an auction protocol which is not aware of insincere agents that might break contracts (and consequently does not provide counter-measures like sanctions) can be considered as a first-order expectation structure.

Higher-order expectation structures are social structures that model first-order (second-order...) social structures we call second-order (third-order...) social structures. If an observer models processes of higher-order communications, and he trusts the communicated propositions (about other communications), then she can easily obtain higher-order expectation structures from the message content.

An example for the use of second-order expectation structures is the following scenario: In a discussion, employees of some organization hold opinion A. Suddenly their boss steps into the office. In the continuation of the discussion, the employees hold opinion $\neg A$. Whereas the discussion before and after the appearance of the boss would be modeled using two mutually inconsistent first-order expectation structures, second-order expectation structures would relate both first-order structures and explain the transition from one to the other with the entry of the boss into the office.

Seemingly, from the *passive* modeling of communication processes in order to forecast the interactional behavior of a set of agents to the *active* participation it is only a small step. We could describe an active higher-order agent as an entity that derives so-called *actual* expectation structures from empirical observations, maintains an other set of expectation structures (so-called *goal structures*, e.g. predefined by the system designer or empirically obtained also), and aims at a minimization of the differences of these two sets of expectation structures in a rational way by taking action himself (especially by means of communication) in one or both of these two agents domains (actual and goal)[1]. Doing so, Mirror-Holons are not only emergent from lower-order

[1] Acting in order to manipulate the agents domain "physically" and directly (not using symbolic interactions) can be modeled as a certain kind of *indirect* communication, too. So "ordinary",

agents' behavior, but they can be goal-directed with goals emergent from observed behavior also. A Mirror-Holon is thus also an agent with the peculiar feature that its social belief and its goals are (at least partially, depending from the type of Mirror-Holon, as explained below) obtained at run-time in form of emergent expectation structures. Therefore, Mirror-Holons can be used to influence a multiagent system *from out of itself* instead of an external instance like in usual approaches to multiagent control (for example, in [9], we have shown how a Social System *Mirror* (a simple kind of Mirror-Holon, cf. below) can be used as a CASE-tool for agent-oriented software development for the purpose of the derivation and propagation of expectation structures in an evolutionary MAS design process). In some sense, this concept is a reversal of the concept of "cognitive" agents that are enabled to interact socially (e.g., [16]), since the cognition and acting of Mirror-Holons are seen as an outcome of observed communications and not the other way round as usual.

This concept of *goal structures* emerging from *actual structures* (the actual, empirical social communication structures of the MAS at a certain time) would be of no use if the goal structures and the actual structures were identical. There are several possibilities to obtain goal structures different from actual structures, e.g. by

– *Synthesization* of goal structures from the informational contributions of multiple, heterogenous communication sources (either agents or peers in open P2P networks). Using techniques like *social reification* [13–15], this synthesization can provide reasonable results even in case of inconsistent informational input (e.g. conflictive behavior). Possible applications are, e.g., the propagation of the resulting structures in order to improve the social reasoning of agents and/or the system designer (similar to the effect of public *mass media* like television or books on human societies.). This improvement is especially useful for the unveiling of *social conflicts* [26]. Examples for the application of synthesized social structures are *Open Ontologies* and *Open Knowledge Bases*, which—in contrast to traditional information media—maintain, weigh and socially reify semantically conflictive (conceptual or instantiated) knowledge computationally [13–15].

– *Biasing* of actual, empirical structures by means of normative structures which were predefined by the system designer in order to filter out undesired behavior and strengthen desired behavior using sanctions or argumentation. Such structures can result from schemes like *RNS* [18].

– *Simplification and acceleration* of communication structures. Goal structures can be obtained from the compilation of actual expectation structures using modifications in order to make them more simple, fast and reliable. If the respective Mirror-Holon is sufficiently powerful, it could even enact observed communication processes (for a certain period of time) in its domain *without* the further participation of the agents that contributed to these processes, whereas a "weaker" Mirror-Holon could act as a "communication catalyzer" that makes use of the agents as

non-interactive acting is included within this rational social behavior, too, as long as the "physical actions" have a significant impact on behavior of the other agents.

effectors to put physical actions occurring in these processes into action. In both cases, the Mirror-Holon would act as a more or less complete replacement of the observed multiagent system (we speak about a *Functional Mirror-Holon* (cf. Section 2.3) in this case).

– *Merging* of the structures of multiple social systems.

This list is not exhaustive, and combinations of these approaches are also imaginable.

Each Mirror-Holon possesses two communication *ports*: The unidirectional *source port* is used to observe communications that occur in the *source domain* (goals), whereas the bidirectional *target port* is used to both, observe communication and participate actively in communication in the *targeted domain*. We use the term "domain" in a quite broad sense, denoting observable events generated by agents during interaction and physical phenomena. For a Mirror-Holon, a certain "physical" domain is of course only indirectly accessible by means of the observation of communications. Source and targeted communications are typically generated by different sets of agents (including other Mirror-Holons), but as with the original Social System Mirror, it might be reasonable to have a non-empty intersection (i.e., an interaction domain influences itself). From the input obtained from these ports the Mirror-Holon derives two *expectation networks* (ENs) [8, 9, 19, 11, 12]. An EN is a concrete, graph-based formal representation form for expectation structures (cf. below for details). Participating using the target port means communication with the aim to reduce the difference of the expectation structures obtained from source and target port by means of taking action using the target port (more precisely: minimizing the probability that the expected continuation of agent communication observed via the target port deviates from the expected source port communications). The concrete goals of a Mirror-Holon is thus determined at run-time from the source port—a Mirror-Holon acts towards his targeted audience like a representant of the source structures. In addition, each Mirror-Holon is optionally equipped with a number of normative expectation structures which serve as an a-priori presetting for the source expectation structures build in by the system designer.

2.1 Expectation networks

Expectation networks, introduced in [8], are graphs that represent expectation structures formally—specifically, they represent the empirically obtained probability distribution of all significant future event sequences resulting from the observance of the agents and their environment. They can also be used to pre-define expectation structures designed manually by the system designer, modeling e.g. normative expectations directed to the agents in order to restrict their behavior. Expectation networks may need to be adapted if unexpected newly observed events occur, and might subsume communicative actions as well as non-symbolic "physical" events. Expectation networks can be modeled in multiple levels of generalization to enable the description of under-specified utterances, communication patterns, collaboration-emergent meaning (e.g., shared opinions of agent groups) and agent roles (which are basically generalized agent behavior patterns). Expectation networks also provide a common ground which contextualizes an

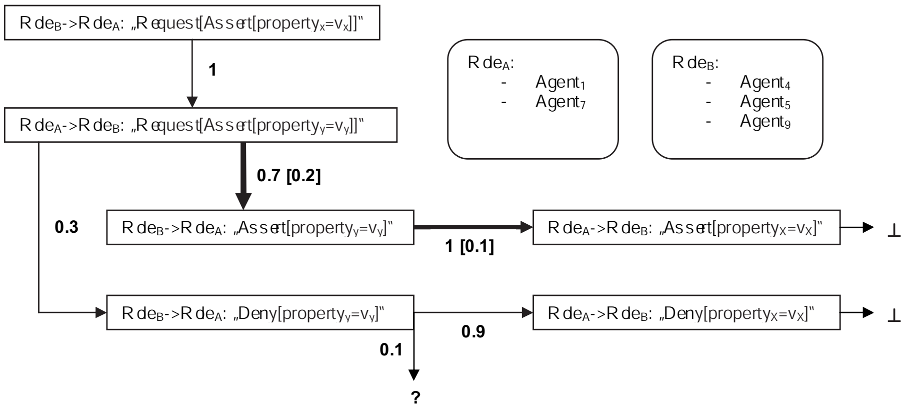

Fig. 1. An expectation network

communication act within a discourse. They are furthermore able to represent not only dynamic (i.e. adaptable) expectation, but also *normative* expectations (e.g. laws), which remain stable even if they contradict actual behavior. Please refer to [12, 19] for a detailed, formal description of expectation networks.

Figure 1 shows a very simple expectation network that represents the structure of a discourse of two agents (or two agent roles, respectively). For simplicity, we use a graphical notation which is slightly different compared to the full notation. Nodes (depicted as squares) are labeled with message templates (in a formal agent communication language) or the special symbols "⊥" (denoting the end of a conversation) and "?" (denoting an unknown or uninteresting continuation). Nodes are connected by edges (shown as arrows) labeled with numerical *expectabilities*, which denote the probability that the respective message(s) occur subsequently. These probabilities are derived from observed frequencies of the respective message sequences in the past. The thickness of edges represents the *normativeness* of the respective expectability and the numerical value in square brackets denotes its *deviancy*. An edge with high normativeness (thick arrow) represents an expectation which has proved itself as empirically stable in the long term, which is a typical property of expectations obtained from laws and other social norms. The deviancy is the difference of long-term and short-term expectability, corresponding to the expectability of agent behavior which deviates from a social norm. Substitution lists appear in rounded boxes. A substitution list denotes a social *role* the listed agents can impersonate. For this purpose, the message templates contain role variables ($Role_A$ and $Role_B$) that can be bounded to each of the list entries, provided this bounding is done in a consistent way along the respective network path. An expectation network can be *generalized* in two ways: First, a single expectation network might describe the expectations regarding multiple message sequences due to different instantiations of role variables (a Mirror-Holon might be able to obtain these roles automatically from the unification of syntactically matching message sequences observed for different agents as described in [8, 12]). Second, each message sequence is expected to be repeatable without precondition if the root of the corresponding path does not have any incoming edges. The numerical expectabilities correspond to the *frequency* of observed message trajectories in the past that unify syntactically with the respective paths. Theoretically, an expectation network must contain a path for every possible sequence of messages, but in practice, edges with a very low or unknown probability are omitted.

2.2 Structure enactment and execution

Likewise there are multiple ways to obtain holon goal structures from actual empirical structures (social programs, mainly), there are also multiple possibilities for the enactment of goal (i.e. source) structures by means of Mirror-Holon communication using the target port:

– *Influencing through information* aims for a change of the behavior of agents (in the targeted domain) by means of informing them about otherwise tacit social structures (in a way similar to the influence *mass media* have in human societies). These information depend from the source structures obtained from the source port, but do not necessarily be exhaustive or true (e.g., a Mirror-Holon might be able to lie in order to influence the target domain).

– *Argumentation, negotiation and sanctioning* are discourse practices of social agents which can be likewise performed by Mirror-Holons. Since Mirror-Holons are primarily thought as control instruments used by the system designer, a Mirror-Holon usually has more power in terms of the enactment of positive or negative sanctions it can impose on "ordinary" agents.

– *Direct enactment* of goal structures requires that the Mirror-Holon has direct access to the target domain. In this case, the Mirror-Holon puts communication acts and other events into action *instead of* or in collaboration with the agents within the target domain.

In general, Mirror-Holons have the following architecture (cf. fig. 2). As we'll see later, this general architecture, which is influenced from *Social Systems Theory* [2] and *Second-Order Cybernetics* [28], allows for a lot of variety.

Definition. A (general) *Mirror-Holon* is defined as a structure

$$(sourceKB, targetKB, defaultKB, sourceUpdate, targetUpdate) \text{ where}$$

– $sourceKB : EN$ is the current model of the source domain ("KB" stands for knowledge base, which means here a set of social structures, i.e. *social knowledge*). This model is given as an expectation network (EN) or some stochastically equivalent representation formalism which represents a probability distribution of events (especially communicative actions) that is incrementally updated from observed messages.

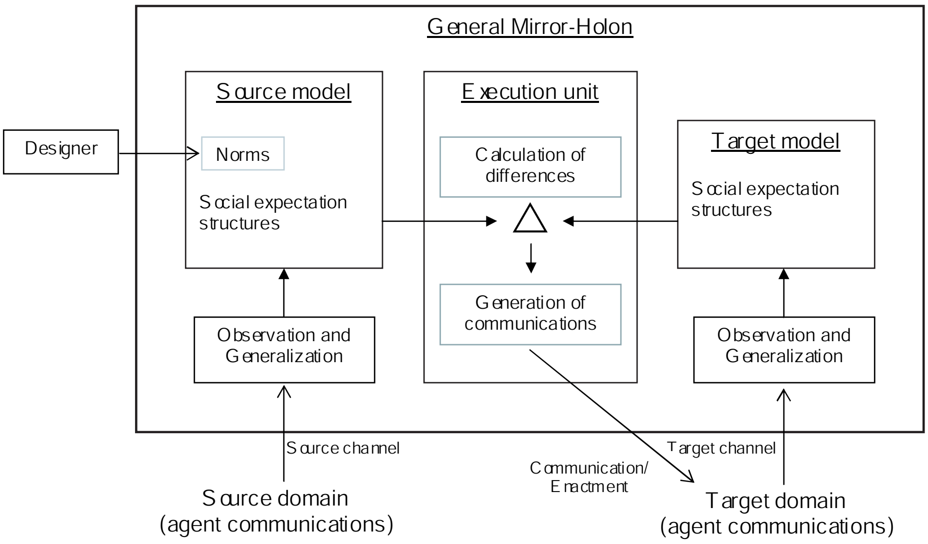

Fig. 2. General Mirror-Holon

– $targetKB : EN$ is the current model of the targeted domain. It is also represented as an expectation network, and adapted dynamically. Some types of Mirror-Holons do not use this model (cf. below).

– $defaultKB : EN$ is the initial content of $sourceKB$. It serves as a normative, a priory bias for the learning of $sourceKB$ from observations, and helps to avoid the bootstrapping problem which might occur otherwise in case no reasonable, structurally relevant input can be accumulated through the source port initially.

– $sourceUpdate : EN \times ACL \rightarrow EN$ is a function which updates the source model after the observance of a message $m \in ACL$, whereby ACL is some agent communication language. This process of source model adaptation is called *generalization* (of observed communication processes). It has the following two aspects: *Timely generalization*, i.e., the extrapolation of communication trajectories into the future, and *role/social program generalization*, i.e., the abstraction of behavioral patterns from concrete agents.

– $targetUpdate : EN \times ACL \rightarrow EN$ likewise updates the target model (including events the respective Mirror-Holon has generated by itself).

We provide two different functions for the update of source and target model in order to allow different algorithms. E.g., it might be useful to do strong filtering and generalization of the source model to obtain consistent, stable goal structures, whereas the targeted domain is modeled as accurate as possible to enable an effective influencing of the targeted audience.

If at least one Mirror-Holon appears in a multiagent system, we talk about a *HolOMAS* (“Holonic Open Multiagent System”).

A single HolOMAS can accommodate more than one Mirror-Holon to enable diversification of expectation structures into multiple *spheres of communication* (cf. Section 3 and 5). To allow the observation of agent communication, the Mirror-Holons are supposed to have access to some sort of shared memory (*whiteboard*) the agents use to make some or all of their messages and other actions observable for the Mirror-Holon (source and target port), and in order to allow the Mirror-Holon to emit events directed to the agents (using its target port). Since it would be rather off-topic, we do not consider technical details and privacy issues regarding such communication media here.

In general, the process a Mirror-Holon performs can be described as follows (the so-called *Mirror-Holon Cycle*):

1. $targetKB := \varnothing$
2. $sourceKB := defaultKB$
3. $message_{source} := pull(sourcedomain)$
4. $sourceKB := sourceUpdate(sourceKB, message_{source})$
5. $message_{target} := pull(targeteddomain)$
6. $targetKB := targetUpdate(targetKB, message_{target})$
7. $subst := unifier(sourceKB, targetKB)$
8. $\Delta := subst(sourceKB) - targetKB$
9. $put(targetdomain, \Delta)$
10. go to step 3

whereby the following additional procedures are used: $pull$ denotes a function which waits for and reads the latest event (e.g. message) that occurred in the respective domain. $unifier$ computes a list of role variable and agent name substitutions such that the application of this substitutions list as a function $subst$ on $sourceKB$ makes this expectation network as similar as possible to $targetKB$ by means of an appropriate renaming of variables and agent names, i.e., $unifier$ finds the most general unifier for the matching parts of the two expectation networks. The infix function '$-$' (step 8) calculates the difference of two expectation networks and results in a list of subsequent Mirror-Holon communications which minimizes this difference of $sourceKB$ and $targetKB$ with the highest probability using the lowest number of single communications. put emits this communication sequence to the targeted audience. "Difference" means the graphical tree-distance of the two ENs here.

As a simple example, imagine a source model which contains the following expectation network paths (we use a textual representation here instead of the graphical notation that should be self-explanatory):

$Role_A \rightarrow Role_B : DeliverGoods(...)$
$\dashrightarrow (0.9)\ Role_B \rightarrow Role_A : FulfillPayment(...)$
$\dashrightarrow (0.1)\ Role_B \rightarrow Role_A : DenyPayment(...)$

For some trading scenario, this course of events can be considered as ideal (just 1% expectation of denial of payment). These expectations might have been obtained in a closed source domain with sincere and reliable trading agents, or could have been predefined via $defaultKB$ by the MAS designer. In contrast, the target model shall contain the following structures:

$Role_X \rightarrow Role_Y : DeliverGoods(...)$
$\dashrightarrow (0.5)\ Role_Y \rightarrow Role_X : FulfillPayment(...)$
$\dashrightarrow (0.5)\ Role_Y \rightarrow Role_X : DenyPayment(...)$

A Mirror-Holon with the task to correct undesired behavior occurring in the target domain would find a high deviancy of the target expectation network in comparison with the source structures, and should perform appropriate sanctions as follows:

$Agent_Y \rightarrow Agent_X : DenyPayment(...)$
$\dashrightarrow (1)\ MirrorHolon \rightarrow Agent_Y : Sanctioning(...)$ where the $Agent_{...}$ are instances of the respective agent roles.

In the case target and source domain are equal, and $Role_Y$ (respectively the instancing agents) deviates from its expected behavior because it is unaware of some fact (e.g., legal powers), the Mirror-Holon could alternatively (or in addition) act as a Social System Mirror and just inform $Role_Y$ about the possible consequences of its behavior.

However, after some cycles, the influence of the Mirror-Holon should lead to reasonably adapted target structures, e.g.:

$Role_X \rightarrow Role_Y : DeliverGoods(...)$
$\dashrightarrow (0.8)\ Role_Y \rightarrow Role_X : FulfillPayment(...)$
$\dashrightarrow (0.2)\ Role_Y \rightarrow Role_X : DenyPayment(...)$

2.3 Special Mirror-Holons

Obviously, the computation of function '$-$' within the Mirror-Holon Cycle is problematic, and we doubt that it can be computed efficiently for the general case[2]. It is also not completely clear yet what "minimizing the difference" of two expectation networks means. For now, we can define this only for special cases: If $-$ results in a sequence of holonic communications that makes the target communication equal to the source communication, the Mirror-Holon surly succeeded in obtaining its goal at least temporary. For these reasons, we introduce in the following more "manageable" Mirror-Holon subtypes derived from the general case described above.

Social System Mirror (cf. Figure 3)

A Social System Mirror [8, 9] is a Mirror-Holon with the following specific properties:

– *The targeted domain is a part of its source domain*, i.e. a domain influences itself at least to some degree. Such influencing mimics *mass media* like television, books and newspapers in human societies, where information (possibly strongly biased by norms and a-priory knowledge) appear to be related to the needs and the behavior of persons who do not necessarily have been involved directly in the creation of these information (e.g. readers of a newspaper). In the special case the targeted domain is the same as the source domain, and $defaultKB = \varnothing$, we would obtain a truly self-influencing of the domain.

– *The Mirror-Holon emits meta-communications (communications about communications) only* (i.e., technically, the content of the generated utterances consist of information about expectation structures only. In particular, a Social System Mirror does not "impersonate" the acting within the source domain.)

– *A Social System Mirror does not impose sanctions on its targeted domain*. In the most basic case, the target port even works in one direction (emit messages) only, and the holon simply generates mass communication ($1 : n$; where n is the number of agents in the targeted domain). The idea behind this is that the targeted audience should select relevant information from the mirror communication itself, electively querying a "socially-aware, open knowledge base" [13–15].
Of course, this property could be omitted or relaxed, if necessary, and in any case, a Social System Mirror is able to *inform* about sanction and social norms.

The two main purposes of Social System Mirrors are 1) the informational influencing of evolving open agent systems during the design phase [9], comparable to a

[2] Of course, we could incrementally compute random Mirror-Holon behavior e.g. using a genetic algorithm, and hope that this eventually leads to the desired structures.

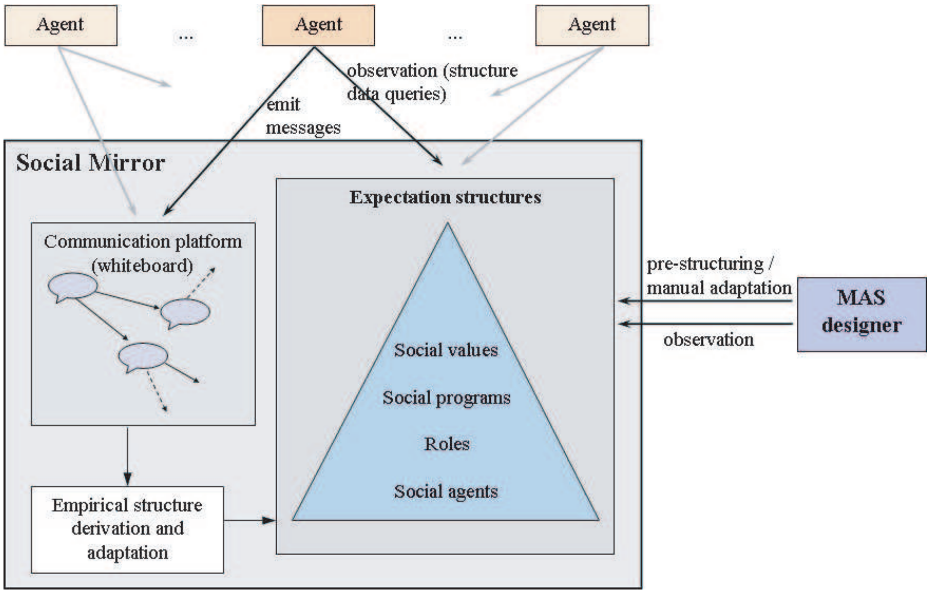

Fig. 3. Social System Mirror

CASE tool, and 2) the enhancement of the social capabilities of the targeted agents by means of enhancing their knowledge about the social system structures.

Normative Mirror-Holon A *Normative Mirror-Holon* (a subtype of the Social System Mirror) does not make use of its source port ($sourceUpdate$ does nothing), and $defaultKB$ is not empty. Thus it is a means for the propagation of social norms and static knowledge. It can be used, e.g., to communicate normative action constraints to autonomous agents (if the scheme can be translated into expectation networks as it is possible for role-based obligation schemes like RNS [18]).
Optionally, it can be equipped with the ability to argue and/or impose sanctions for norm-deviant agent behavior.

Functional Mirror-Holons are Mirror-Holons which represent suitable social structures (*social programs* [2, 8, 9]) as executable programs, which are inductively derived, evaluated and adopted during run-time of the MAS. We call both the original social structures and the derived computational programs "social programs" [3].

Synchronous-Functional Mirror-Holon (SFMH) (cf. Figure 4)
A SFMH is a Functional Mirror-Holon that works as a functional representation of social expectation structures. Its functionality is based on the interpretation of expectation structures as social programs. These can e.g. be ENs (ex-

[3] The term "functional" is used to emphasize the priority of social functionality over non-functional, redundant structural determination according to [2].

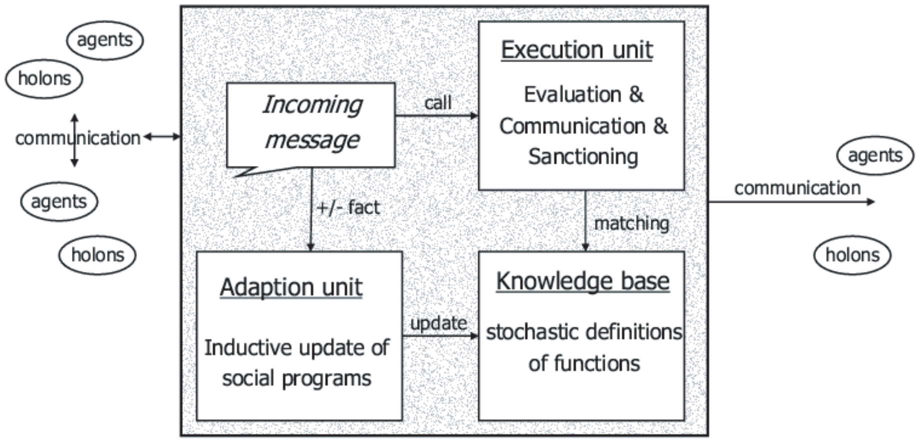

Fig. 4. Synchronous-Functional Mirror-Holon

ecuted by means of stochastic simulation, cf. Section 4), or declarative, functional programs in the computer-scientific sense. In the latter case (which can only be sketched here for lack of space), each SFMH represents its $sourceKB$ as an adaptive set of function definitions, which are continuously and inductively learned from observed communication acts (in the context of other communication acts and other observable events). Each function represents a certain generalized sequence of correlated agent action events, similar to the paths in the expectation network.
The SFMH communicates the content of $sourceKB$ to the agents in the same way a Social System Mirror communicates expectation structures to enhance or update their own social belief. But in addition, the agents can also *call* the inductively learned "social functions" like communications macros, and the evaluation of each function can in turn create "calls" of further agent behavior. Therefore, a SFMH works in interaction with other agents with the aim to make their social behavior more efficient.

Asynchronous-Functional Mirror-Holon (AFMH) (cf. Figure 5)
An AFMH is a variant of the SFMH, with the important difference that it separates the process of continuous observation and learning of expectation structures on the one hand, and the execution of the derived social programs on the other (in Figure 5 called "evaluation") timely and organizationally. Following Social Systems Theory, the loose coupling between these two processes is called *irritation* (of the structures represented by the social programs).

An AFMH does not inform the targeted agents to put the inductively derived action functions into practice, but instead performs the recorded and extrapolated sequences of agent actions by itself, or, alternatively, *forces* the targeted

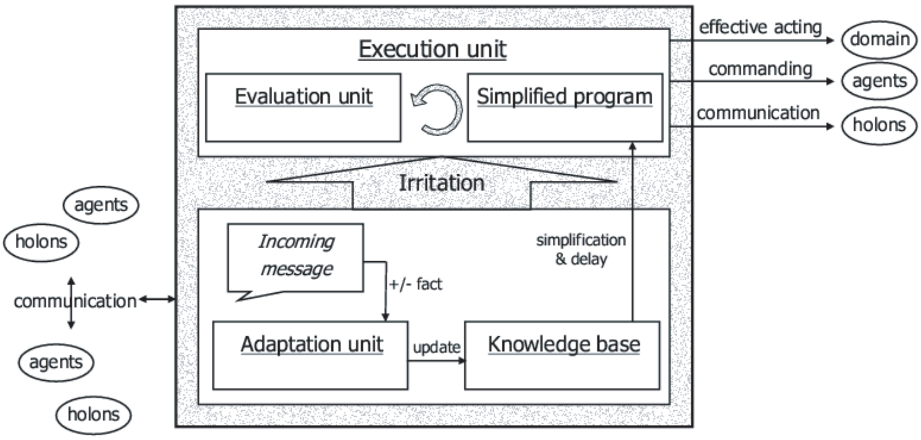

Fig. 5. Asynchronous-Functional Mirror-Holon

agents to execute them by means of normative power.

AFMHs have the big advantage that they do not require a model of the target domain and thus do not need to calculate the difference of source and target model to obtain optimal target communications—instead, they more or less copy and replace the observed MAS. More specifically, the calculation of the difference of source and target model is trivial, because it always results in the source model. However, usually an AFMH simplifies and speeds up the source structures before execution. AFMHs are most useful if the Mirror-Holon just needs to execute communicative acts which have a physical impact and not just a symbolic meaning.

Because of their simplicity and generality, we have chosen AFMHs for this work as the concrete example type for Mirror-Holons in the next sections, whereas we do not go into detail about the other types.

3 Empirical Semantics

As we have seen, a Mirror-Holon can be thought as an acting impersonator of expectation structures. Therefore, the by far most important task for a Mirror-Holon is the derivation of communication meaning. In this section, we provide only an informal overview of the central aspects of our communication model. Please consult [11, 17] for details and a formal framework.

Although many approaches to the semantics of agent communication languages (ACL) have already been proposed, it is widely realized in distributed artificial intelligence that a comprehensive understanding of agent communication is still outstanding.

While it is relatively easy to define a proper formal semantics for the so-called “content level” of agent languages (in contrast to the speech act illocution encoded by means of performatives), like it has been done for, e.g., KIF [20], there is still no general model of the actual *effects* of utterances in social encounters, a field which is traditionally studied in linguistical pragmatics and sociological theories. Currently, two major speech-act-based [21–23] approaches to this aspect of agent communication (i.e. “semantics” in a broader sense, covering both traditional linguistic sentence semantics and pragmatics) exist, if we do not count plain interaction protocols (in some sense primitive social semantics) and other low-level formalisms like message passing. The older *mentalistic* approach (e.g. [6, 3]) specifies the meaning of utterances by means of a description of the mental states of the respective agents (i.e., their beliefs and intentions), while the more recent approaches to ACL semantics (e.g. [5, 1]) try to determine communication from an *objectivistic* point of view, focussing on public language rules. The former approach has two well-known shortcomings, which eventually led to the development of the latter: At least in open multiagent systems, agents appear more or less as black boxes, which makes it in general impossible to impose and verify a semantic described in terms of cognition. Furthermore, they make simplifying but unrealistic assumptions to ensure mental homogeneity among the agents, for example that the interacting agents were benevolent and sincere. Objectivist semantics in contrast is fully verifiable, it achieves a big deal of complexity reduction through limiting itself to a small set of normative rules, and has therefore been a significant step ahead. But it oversimplifies social processes, and it does not have a concept of semantics dynamics and evolution. In general, we doubt that the predominately normative, static and definite concepts of current approaches to ACL semantics, borrowed from the study of programming languages and interaction protocols, are adequate to cope with concepts crucial for the successful deployment of agents to heterogeneous, open environments with changing populations like the internet. Of course, this issue is less problematic for particular environments, where agent benevolence and sincerity can be presumed and agent behavior is relatively restricted, but for upcoming information-rich environments like the Semantic Web, three particular communication-related properties, which are traditionally associated with human sociality, deserve increased attention: 1) meaning is usually the *result* of multiple heterogeneous, possibly indefinite and conflicting communications, 2) benevolence and sincerity can not be assumed, and 3) homogenous mental architectures and thus the uniform processing of communicated information cannot be assumed also.

The meaning of utterances has two dimensions that need to be covered by a comprehensive approach to the semantics of agent communication (the term "semantics" here always in the broader computer scientific meaning including pragmatics, not just the more abstract linguistic sentence meaning): First, the sentence level, which is the aspect of meaning that is traditionally subject of linguistical semantics. This aspect of meaning is contextualized with an environmental description in the form of a (assumably) consented ontology. In addition, a calculus to describe objects and events within the environment the respective utterance refers to has to be provided, for example predicate logic and temporal modalities. The second dimension of meaning, its pragmatics (i.e., the actual use and effect of utterances in social encounters), contributes by far the most difficulties to current distributed artificial intelligence. This is mainly due to agent autonomy, which makes it extremely difficult to obtain deterministic descriptions of agent behavior. Thus, current objectivist approaches either deliberatively avoid pragmatics at all, or try to impose pragmatical rules in a normative manner (leaving beside mentalistic approaches, which are not suitable for black- or gray-box agents in open system for obvious reasons).

The communication model we propose is grounded in *Social Systems Theory* [2], as it has been adopted for the modeling of expectation structures of artificial agents [8][4].

In our communication models [8, 19, 11, 17], called *Empirical Semantics* and *Empirical-Rational Semantics*, which we can only outline here, a single communication attempt can be seen as a request to act in conformance with the information expressed by the utterance, or respectively to establish a requested future state (this includes both assertions of propositional information and requests to perform actions). In contrast to non-communicative events, an utterance has no (significant) direct impact on the physical environment. Instead, its consequences are achieved socially, and, most important, the addressee is free to deny the communicated proposition. Since an utterance is always explicitly produced by a self-interested agent to influence the addressee, communicated content can not be "believed" directly (except in the case the addressee could have derived its truth/usefulness herself and a communication would thus be rather unnecessary), but needs to be accompanied with social reasons given to the addressee to increase the probability of an acceptance of the communicated content. This can be done either explicitly by previous or subsequent communications (e.g., "If you comply, I'll comply too"), or implicitly by means of generalizations from past events (e.g., trust). The expected communication events which are triggered by a certain communication (in the context of the preceding communication process) in order to support the aims of this communication we call (informally for now) the *rational hull* of the triggering communication. The rational hulls of communications specify inter alia indirectly the social relationships which steer the acceptance or denial of communicated content according the public communication attitudes the agents exhibit (their public intentional stances, so to say), e.g., by argumentation and sanctioning. Typically, a rational hull is initially quite indefinite and becomes increasingly definite in the course of interaction, provided that the agents work towards mutual understanding of their ostensible goals (this of course does not entail consensus).
Our model is centered around the following terms, which we propose primarily as empirical replacements and supplements for terms used in traditional ACL semantics, like message content and commitment.

Social expectation structures (communication structures) As already described, social expectation structures are the part of the expectation structures consisting of social expectations that result from communication processes and constrain future communications. The visible effect a certain utterance brings about in the social expectation structures is the semantics of this utterance if no a-priori or mental

[4] The term "agent" always shall denote both ordinary agents and Mirror-Holons.

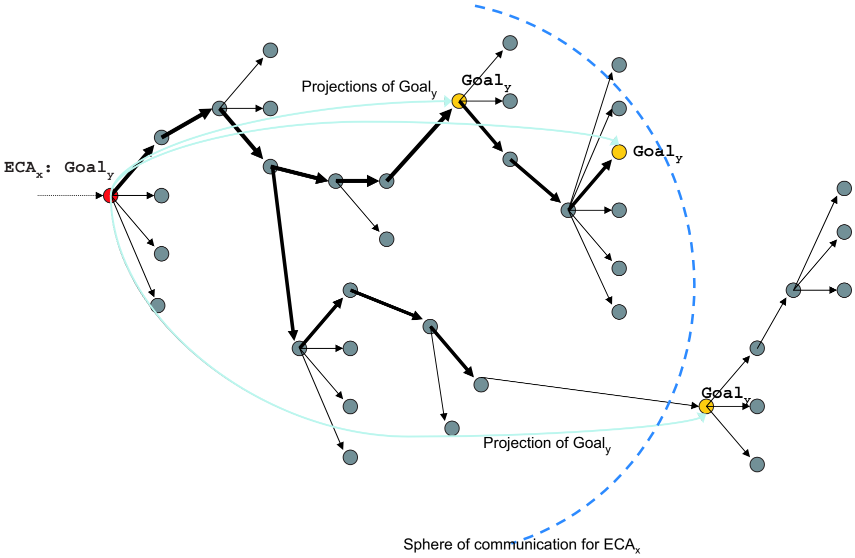

Fig. 6. An EN with projections and a sphere of communication

knowledge about the respective actor is available.

Utterances An agent action event with the following properties: 1) it is occurring under mutual observation of the communicating agents, 2) without considering social expectation structures, the event would have a very low probability (e.g., it is unlikely that sentences are uttered without the intention to communicate), 3) its expected consequences in terms of physical expectation structures only are of low relevance (think of the generation of sound waves through human voice), and 4) considering social expectation structures, the event needs to be informative, i.e., its probability must be lower 1 and must result in a change of expectations. For utterances using a formal language and reliable message passing, criteria 1) to 3) are clearly met. In our model, each utterance encodes one or more *projections*.

Projections Our replacement to the term “content” used in the context of formal languages like KIF (used to represent the propositional so-called “inner level” of messages described in speech act based ACLs like KQML). A projection is the part of the expectation structures which is selected through an utterance to inform the addressee. Of course, each utterance can also encode multiple projections at the same time. A projection can be considered to be a set of goal states within the EN the uttering agent strives for rationality, at least allegedly for the time of the respective

sphere of communication (see below). Therefore, these goals need not be the true goals of the uttering agent, but at least for some time the agent acts "as if" they were.
The most basic kind of projection is obtained through *demonstrative acting*, where the uttering agent encodes its goals by means of "playing-act". Another possibility is to encode a projections within the traditional speech act form, or as a so-called elementary communication act (ECA) as described in [11]. The latter is more or less a set of pointers at agent-individual goal states of an EN. We consider all kinds of speech acts to be encodable as sets of ECAs together with constellations of given expectation structures (e.g. a command could be encoded as the ECA describing the goal of the command together with additional ECAs to sanction the addressee, and/or social structures which give the commanding agents the required social power to make the command effective). More generally, utterances can encode a projection using a "wishful" expectation network (describing the state the agent appears to desire). These expectation networks need to be matched with the actual EN to identify the states within the actual EN the agent strives for.
If a projection refers to communicative behavior itself (e.g., a question demands a communication), we talk about meta- or higher-order communication (cf. below).

Rational hulls The *rational behavior* an agent is *expected* to perform in order to make a certain uttered projection become reality (using e.g. assertions, sanctions, normative behavior, negotiations, actions to increase trust and his reputation...). Here, rationality only means ostensibly rational behavior towards ostensible agent goals, which are not necessarily the true goals of the agent. The rational hull is defined as the set of social expectations arising from the assumption that the uttering agent tries (at least for some time) to maximize the probability that subsequent events are consistent with its uttered projection (speech act perlocution is a special case of this principle). Practically, the rational hull of an ECA is computed via a combination of empirically learned, revisable experiences from past agent behavior in similar contexts and the application of the rule of rational choice (cf. [11]). Rational hulls are recursive in the sense that each element of a rational hull has its own rational hull and so on, in their sum amounting to the empirical meaning of communication.

Figure 6 shows an EN modeling the future of some communication process. ECA_X is the utterance which encodes $Goal_Y$. This goal itself stands for several (seemingly) desired states of the EN (yellow nodes). Since within the so-called *sphere of communication* of ECA_X (see below) it is expected that the uttering agent rationally strives for these states, certain EN paths leading to these states become more likely (bold edges). Such behavior paths need to be (more or less) rational in terms of their expected utility (e.g. in comparison with competing goal states), and they need to reflect experiences from analogous agent behavior in the past.

Communication processes A set of probabilistically correlated utterances with the following properties: 1) each agent acts in consistence with the rational hulls induced from his utterances (which especially means that he does not contradict itself), and 2) each projection is consistent with 1), i.e. it does not deny that property

1) is met. 2) is somehow an empirical version of mental *understanding* and *trust*: With each communication, an agent acknowledges with his own communicating behavior implicitly that the other agent tries to get accepted his own projections.

Spheres of communication Each utterance (more precisely: each ECA) can have its own spheres of communication which describes the boundary of foreseeability of its consequences (=semantics) in terms of the expected subsequent communication process, i.e. the timely extend of a set of expected communication acts that are consistent and correlated. Every communication process together with the foreseeable expectation structures arising from this process creates thus a set of spheres of communication. Together these spheres form a so-called *social interaction system*. Whereas a communication sphere is similar to the special sort of social system called *interaction system* known from social systems theory [2], and resembles some of the properties of *spheres of commitment* [7], in our model spheres of communications have dynamic, empirically discovered boundaries in the sense that communications which do not fulfill the consistency criteria for communication processes at run-time mark their boundaries and are thus not part of it (e.g., misunderstandings and lies, if they become obvious). The most simple examples for communication spheres are those that rely on normative structures (e.g. protocols), like auctions (cf. [9] for a case study on empirical expectation-oriented modeling of a trading platform) and (agent-supported) forums on the internet, as long as the agents do not break the communication laws of these systems.
A general Mirror-Holon accommodates exactly one communication sphere for its source domain, and one for the targeted domain. Therefore, a single Mirror-Holon is not able to model (communicatively revealed) misunderstandings and two or more sets of communications that are not empirically correlated. In such cases, multiple Mirror-Holons are required to bridge such inconsistencies and incoherencies.

Higher-order expectation structures Social expectations which model multiple, probably inconsistent communication spheres at the same time.
At the moment, our formal model of expectation structures [19] does not allow for an explicit modeling of higher-order expectation structures. But a communication sphere can of course describe processes of higher-order communications (e.g. communications about communications, generated by a Social System Mirror), which can be modeled using an expectation network. Such an expectation network thus models expectations that model expectation structures themselves, therefore some sort of higher-order expectation structures.

Physical expectation structures Optionally, expectation networks can additionally contain domain-dependent expectation structures (i.e., ontological information about the non-symbolic environment). These structures include expectations regarding "physical" agent actions and other events in the agents' environment. Their main characteristic in comparison to communicative events (utterances) is that such events do not consist of projections.

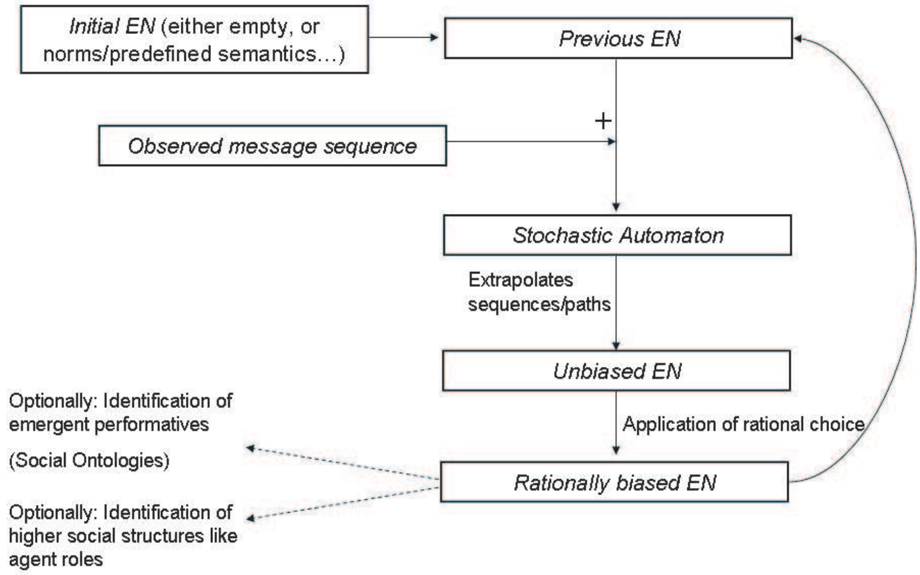

Fig. 7. Learning communication structures rational-empirically

4 Learning and Asynchronous Enactment of Social Structures

4.1 Derivation loop

Figure 7 depicts the *EMPRAT* algorithm for the derivation of expectation networks (i.e. the semantics of communication sequences) from observed agent messages (please refer to [11] for details). The figure shows its most basic kind—it needs to be adapted depending from the concrete type of Mirror-Holon. If, for example, the system designer wants to propagate normative social structures to the agents, it would be required to "inject" static expectations (norms) into the derived ENs.

The algorithm starts with the *initial EN* that contains given knowledge about the communication system and predefined communication structures (like communication protocols). Observed agent messages are used as learning examples for the inductive building of a stochastic automaton, which is converted into an expectation network. This "unbiased EN" therefore considers empirical experiences only (besides the initial EN). To speed up the learning process, from this unbiased EN a *rationally-biased* EN is generated by the application of rational choice rules which reflect the agents' decision processes and rational attitudes (limited by the borders of the respective spheres of communications). The process repeats for newly observed messages, using the latest rationally-biased EN as a new initial EN.

The current version of this algorithm does not yet consider higher-order expectation structures (required e.g. for the modeling of *questions*) and misunderstandings. Traditional speech act performatives like assertions as in Figure 1 are also not supported yet, but can often easily be "emulated" using ECAs.

4.2 Execution

The straightforward, yet quite "naive" method to enact an EN is to use *stochastic simulation* [29]. The algorithm can be sketched in pseudo-code as follows:

for $i = 1..n$ {

$execute(startnode)$

}

with

function $execute(node)$ {

for each $childnode_j$ in $children(node)$ {

if $(1/expect(childnode_j))|i$ { (*)

$emit(childnode_j)$

$execute(childnode_j)$

}

}

}

Here, the whole expectation network is traversed top-down n times, beginning at $startnode$ (please refer to the formal framework of ENs as described e.g. in [19]). The higher n, the higher is the accuracy of the simulation process, i.e., the more closely the probability distribution of the emitted actions resembles the probability distribution represented by the expectation network.

Each call of the function $execute$ traverses the child nodes of a certain node, and recursively calls $execute$ with probability $expect(childnode_j)$ for each child node $childnode_j$ (i.e. calls $execute$ iff i divides the inverse of $expect(childnode_j)$ without a remainder). $expect(childnode_j)$ is the expectability of $childnode_j$ within the EN (i.e. the probability the Mirror-Holons assigns), from the interval $[0, 1]$.

Function $emit$ executes the action associated with the respective node. Depending from the concrete type of Mirror-Holon, the holon could perform this action by itself, or delegate the execution to the agent which originally contributed this action (normatively, or via information only).

Of course, this way of executing an EN has several shortcomings. First of all, repeating the whole execution n times does not necessarily reflect the behavior of the original social system. Seemingly, improvements in this respect would be to call $execute(startnode)$ significantly less than n (or only once), and to replace condition (*) with the non-deterministic result of a random number generator which generates $true$ with probability $expect(childnode_j)$. Furthermore, the difference of emitting communicative actions and "physical" actions is not considered, and no simplification/clearance of the EN (in order to speed up the execution or to make it more reliable) is performed here.

5 Differentiation and Communication of Mirror-Holons

So far, we did not say much about the *boundaries* of Mirror-Holons, i.e., the selection of communication it observes and its sphere of activity within the respective MAS, and the interaction of multiple Mirror-Holons. In general, a single Mirror-Holon could model a complete multiagent system, provided that the trajectories of observed source communications are communication processes in the sense of Section 3, i.e., the Mirror-Holon represents a single sphere of communication . Thus a reason for having more than one Mirror-Holon in a MAS would be the presence of multiple communication spheres. This can occur if 1) some the communications show up inconsistencies regarding understandability (an agent contradicts himself, which can not be modeled within a single sphere of communication) and 2) some of the communications are not correlated statistically. In case 1), to model this inconsistency, we need to introduce a *meta* Mirror-Holon that accommodates higher-order social structures to provide a model which explains these inconsistencies, whereas issue 2) could simply be handled by multiple Mirror-Holons for each identified sphere.

In case a MAS is equipped with at least two Mirror-Holons, these Mirror-Holons typically communication with each other for the following two reasons. First, provided that the interacting Mirror-Holons trust each other, one could supply the other with information about social structures (i.e., meta-communicate about communication processes). This is useful if one Mirror-Holon needs to model a certain source domain, and the other Mirror-Holons emits (e.g., as a Social System Mirror) information about this domain. Then the first mirror can simply query the required social structures from the other Mirror-Holon instead of having to obtain them itself from agent observation. Second, the communicative actions a Mirror-Holon performs in its targeted domain might be observed by another Mirror-Holon. This indirect way of communication is closely related to the concept of meta Mirror-Holons, because observing and modeling the behavior of an actor which represents social structures can be seen as deriving some sort of higher-order expectation structures.

Figure 6 shows a MAS which is equipped with five interacting Mirror-Holons.
Mirror-Holon 1 needs to model its targeted domain (Communication set B), which happens to be the source domain of Mirror-Holon 2. As a Social System Mirror, Mirror-Holon 2 can communicate structure information about Communication set B to Mirror-Holon 1, and thus Mirror-Holon 1 does not need to obtain expectation structures for its

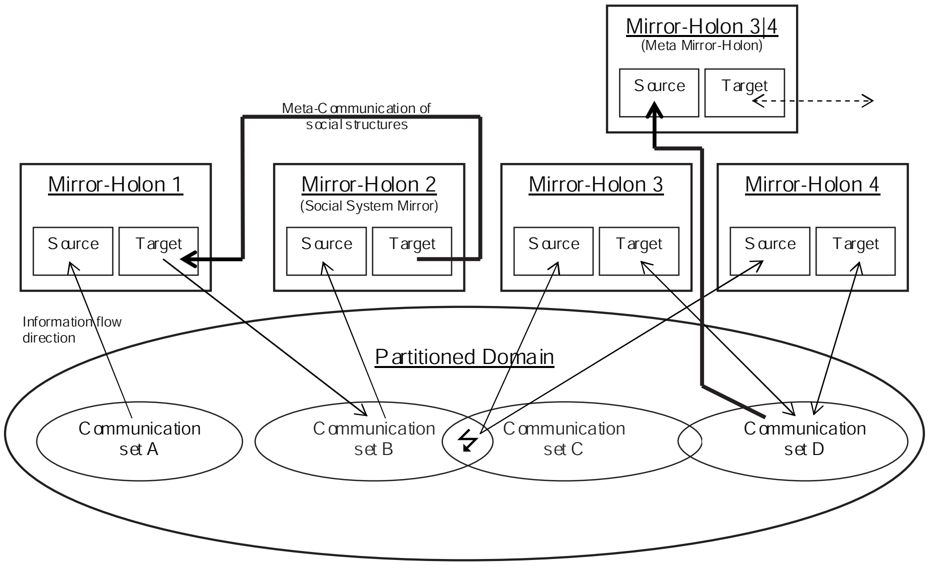

Fig. 8. Interacting Mirror-Holons

targeted domain itself (strictly speaking, Mirror-Holon 2 is part of the targeted domain of Mirror-Holon 1, and Mirror-Holon 1 ignores all but the higher-order communication generated by Mirror-Holon 2 within this domain).

The intersection of communication sets B and C shall contain inconsistencies. Therefore, this intersection can not be modeled by a single Mirror-Holon. The source models of Mirror-Holon 3 and 4 shall each represent a consistent subset of this intersection, and put this subset into action in Communication set D. Mirror-Holon 3|4 is a meta Mirror-Holon in the sense that it models the behavior of Mirror-Holon 3 and 4 as it appears in communication set D.

6 Conclusion

In this paper, we have introduced Mirror-Holons as means for the autonomy-preserving influencing of multiagent systems at run-time. Because Mirror-Holons are on the one hand based on the only two coordination principles that fully preserves agent autonomy, namely *observation* and non-obstructive *information*, and on the other hand Mirror-Holons can be fully equipped with the ability to act and communicate, they bring together both poles of agent-oriented software development—complete passiveness and the imposition of activity—and allow for a leveled, dynamic weighting of both extremes. General Mirror-Holons are a broad and abstract approach. But, while a lot of work lies ahead, we nevertheless strongly believe that Mirror-Holons have the potential to become useful coordination mechanisms especially *because* of their abstractness. As we have seen, there is a large spectrum of Mirror-Holon subtypes and only a few fixed design constraints which makes it likely that adequately tailored Mirror-Holons are applicable even for ill-defined, underspecified scenarios, where almost no assumptions can be made about agent behavior.

Being an introductional work, this paper leaves much room for further specifications, applications and enhancements. As most important we consider to be the continuation of the implementation and evaluation of basic functional Mirror-Holon types like the AFMH and the SFMH, and the theoretical implications of the general Mirror-Holons, especially regarding the differentiation of multiagent system communications into multiple spheres of communication (respectively their representing Mirror-Holons), as we suspect that this kind of differentiation resembles to some degree the process of *Functional Differentiation* described by Social Systems Theory.

References

1. M. P. Singh. A social semantics for agent communication languages. In Proceedings of the IJCAI Workshop on Agent Communication Languages, 2000.
2. N. Luhmann. Social Systems. Stanford University Press, Palo Alto, CA, 1995.
3. Y. Labrou, T. Finin. Semantics and conversations for an agent communication language. In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI-97), 1997.
4. M. Schillo and D. Spresny. Organization: The Central Concept of Qualitative and Quantitative Scalability. In this volume.
5. F. Guerin and J. Pitt. Denotational Semantics for Agent Communication Languages. In Proceedings of the 5th International Conference on Autonomous Agents (Agents'01), ACM Press, 2001.
6. P. R. Cohen, H. J. Levesque. Communicative actions for artificial agents. In Proceedings of the First International Conference on Multiagent Systems (ICMAS-95), 1995.
7. M. P. Singh. Multiagent Systems as Spheres of Commitment. In Proceedings of the ICMAS Workshop on Norms, Obligations, and Conventions, 1996
8. K. F. Lorentzen, M. Nickles. Ordnung aus Chaos – Prolegomena zu einer Luhmann'schen Modellierung deentropisierender Strukturbildung in Multiagentensystemen. In T. Kron (Ed.), Luhmann modelliert. Ansätze zur Simulation von Kommunikationssystemen, Leske & Budrich, 2002.
9. W. Brauer, M. Nickles, M. Rovatsos, G. Weiß, K. F. Lorentzen. Expectation-Oriented Analysis and Design. In Proceedings of the Second International Workshop on Agent-Oriented Software Engineering (AOSE-2001), LNCS 2222, Springer, 2001.
10. M. Nickles and G. Weiss. Empirical Semantics of Agent Communication in Open Systems. In Proceedings of the Second International Workshop on Challenges in Open Agent Environments, 2003.
11. M. Nickles, M. Rovatsos, and G. Weiss. Empirical-Rational Semantics of Agent Communication. In Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-04), 2004.
12. M. Nickles, M. Rovatsos, G. Weiß. Expectation-Oriented Modeling. In International Journal "Engineering Applications of Artificial Intelligence" (EAAI), Elsevier, 2005.
13. M. Nickles and T. Froehner. Social Reification for the Semantic Web. Research Report FKI-24x-04, AI/Cognition Group, Department of Informatics, Technical University Munich, 2004.
14. T. Froehner, M. Nickles, and G. Weiß. Towards Modeling the Social Layer of Emergent Knowledge Using Open Ontologies. In Proceedings of The ECAI 2004 Workshop on Agent-Mediated Knowledge Management (AMKM-04), 2004.

15. T. Froehner, M. Nickles, and G. Weiß. Open Ontologies—The Need for Modeling Heterogeneous Knowledge. In Proceedings of The 2004 International Conference on Information and Knowledge Engineering (IKE 2004), 2004.
16. M. Rovatsos, G. Weiß, and M. Wolf. An Approach to the Analysis and Design of Multiagent Systems based on Interaction Frames. In Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02), 2002.
17. M. Nickles, M. Rovatsos, M. Schmitt, W. Brauer, F. Fischer, T. Malsch, K. Paetow, G. Weiss. The Empirical Semantics Approach to Communication Structure Learning and Usage: Individuals- vs. System-Views. In Nigel Gilbert (Ed.), Journal of Artificial Societies and Social Simulation (JASSS), Issue on Socionics, 2005.
18. M. Nickles, M. Rovatsos, G. Weiß. A Schema for Specifying Computational Autonomy. In Proceedings of the Third International Workshop on Engineering Societies in the Agents World (ESAW). Lecture Notes in Computer Science. Springer, 2003.
19. M. Nickles, M. Rovatsos, W. Brauer, G. Weiss. Communication Systems: A Unified Model of Socially Intelligent Systems. In this issue.
20. M. R. Genesereth, R. E. Fikes. Knowledge Interchange Format, Version 3.0 Reference Manual. Technical Report Logic-92-1, Stanford University, 1992.
21. J. L. Austin. How to do things with words. Clarendon Press, Oxford, 1962.
22. J. R. Searle. A taxonomy of illocutionary acts. In K. Gunderson (Ed.), Language, Mind, and Knowledge (Minnesota studies in the philosophy of science VII), pages 344–369. University of Minnesota Press, 1975.
23. M. Colombetti, M. Verdicchio. An analysis of agent speech acts as institutional actions. In Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS2002), 2002.
24. C. Gerber, J. Siekmann, G. Vierke. Flexible Autonomy in Holonic Multiagent Systems. In AAAI Spring Symposium on Agents with Adjustable Autonomy, 1999.
25. A. Omicini, E. Denti. Formal ReSpecT. Electronic Notes in Theoretical Computer Science 48, Declarative Programming Ů Selected Papers from AGP'00. Elsevier Science B.V., 2001.
26. R. Dieng, H. J. Mueller (Eds.), Conflicts in Artificial Intelligence. Springer, 2000.
27. A. Koestler. The Ghost in the Machine. Arkana, 1967.
28. F. Heylighen, C. Joslyn. Second Order Cybernetics. In F. Heylighen, C. Joslyn, V. Turchin (Eds.), Principia Cybernetica, 2001.
29. B. D. Ripley. Stochastic Simulation. Wiley, 1987.