Communication Attitudes: A Formal Approach to Ostensible Intentions, and Individual and Group Opinions

Matthias Nickles [1] Felix Fischer [2] Gerhard Weiss [3]

AI/Cognition group, Department of Informatics
Technical University of Munich
85748 Garching bei München, Germany

Abstract

Conventional approaches to the modeling of autonomous agents and agent communication rely heavily on the ascription of mental properties like beliefs and intentions to the individual agents. These "mentalistic" approaches are, when applicable, very powerful, but become problematic in open environments like the Semantic Web, Peer2Peer systems and open multiagent systems populated by truly autonomous, self-interested grey- or black-box agents with limited trustability.

In this work, we propose *communication attitudes* in form of dynamic, revisable *ostensible beliefs* (or *opinions*) and *ostensible intentions* as foundational means for the logical, external description of agents obtained from the observation of communication processes, in order to retain the advantages of mentalistic agent models as far as possible, but with verifiable results without the need to speculate about covert agent internals. As potential applications, communication attitudes allow for a simultaneous reasoning about the (possibly inconsistent) "public image(-s)" of a certain agent and her mental properties without blurring interferences, new approaches to communication language semantics, and a fine-grained, statement-level concept of trustability.

As a further application of communicative attitudes, we introduce multi-source opinions and opinion bases. These allow for the computational representation of semantically heterogeneous knowledge, including the representation of inconsistent knowledge in a socially reified form, and a probabilistic weighting of possibly indefinite and inconsistent assertions explicitly attributed to different provenances and social contexts.

Key words: Agent Communication, Open Environments, Multiagent Systems, Knowledge Fusion, BDI, FIPA

[1] Email: nickles@cs.tum.edu
[2] Email: fischerf@cs.tum.edu
[3] Email: weissg@cs.tum.edu

1 Introduction

Although it is nowadays widely accepted in artificial intelligence that open systems with heterogeneous, more or less inscrutable actors (like the Semantic Web or Peer2Peer systems) are an increasingly important field of application for multi-agency, agents (including human users) are still mainly characterized in terms of their mental attitudes. For the most part, however, these are concealed from the point of view of an external observer. And even if the mental states of all participants were accessible, it would be an extremely difficult task to describe the dynamic behavior of a system as a function of all of its individuals, particularly in applications with a large number of actors (like it is possibly the case in internet-based applications). Hence, although the description of an agent in terms of her beliefs, intentions and desires can be unsurpassably powerful, it comes to its limits in such scenarios.

One way to cope with this situation in order to ensure the controllability of the agents and to secure the attainment of supra-individual system goals is to impose normative restrictions on agent behavior, and/or to model the system in terms of static organizational structures. While such approaches certainly have their merits, and systems without any predefined policies and protocols are hardly imaginable, they can also lead to serious limitations of the desired properties of autonomous systems like their robustness and flexibility by imposing constraints on agent adaptivity in a top-down fashion. [4] And after all, agent technology is basically a bottom-up approach, allowing for the emergence of behavior and solutions, and should not be turned upside down. This is not to dismiss the importance of supra-individual structures and norms, of course. Especially public communication models like an agent communication language semantics cannot avoid normativity completely. But we think that normativeness should be limited to a minimum, and the focal point should be the emergence and adaptivity of the models, and the measurement and autonomy-respecting influencing of the degree of adherence of the agents to given models at run-time.

Another possible approach, which could be seen as complementary to the ones just described, would be to constrain the modeling of black-box agents from an external point of view by limiting the validity of norms or cognitive models (thus inducing some sort of bounded observer rationality), e.g. by using information about the respective agent's trustability and reputation. The approach introduced in this paper is basically compatible and in line with such means, but takes a different perspective: Instead of adding restrictions to statements about internal agent properties, or to *social commitments* [20] or other norms, we "lift" mental attitudes to the social stage by introducing logical and verifiable communication-level pendants to belief and intention, namely the communication attitudes *opinion* (synonymous

[4] Even the enactment of an obligatory agent communication language semantics can seriously limit agent autonomy, for example by prescribing to some degree what the communication parties have to "think" during communication (like [5]).

to *ostensible belief* in this work)[5] and *ostensible intention*.

This "second set" of attitudes will be given a meaning which resembles the modalities of agent belief and intention in many ways, and can be used as a direct replacement for belief and intention in most imaginable scenarios (as we will later show for FIPA ACL semantics). It is nevertheless obtained and related entirely from/to observations, and cleanly separated from actual mental attitudes. The difference of actual and alleged mental attitudes has sometimes been recognized informally in fields like belief ascription, but not been handled explicitly or in a formal framework so far, to our best knowledge. Communication attitudes also provide a logical model which is descriptive in the sense that it describes the actual communicated "public image" of an agent[6], in contrast to (and as a possible complement of) prescriptive approaches based on social norms and commitments.

With our approach, we aim thus at a convenient and practicable means to reason about interaction and communication processes. Thereby we want to prevent the difficulties and confusion that come along with traditional mental reasoning about truly autonomous agents (but without loosing the great expressiveness of agent modeling in terms of mental properties as far as possible), and we also aim at avoiding the rather normative or indifferent (in regard to what "being committed" actually means) perspective of most commitment-based approaches (but acknowledge the necessity to provide a minimally committing communication act meaning as a matter of principle).

Our approach to abstract from traditional "in depth" mental modeling and reasoning promises a great deal of complexity reduction, especially when large sets of mentally opaque agents have to be handled. It also allows for the simultaneous reasoning about real beliefs and intentions of an agent on the one hand and his communicated viewpoints (ostensible beliefs and intentions) on the other without confusion, and for the modeling of social behavior which cannot be modeled as a revision of beliefs and intentions, like the bargaining of opinions.

Optionally, communicative attitudes can be annotated with spheres of reliability on statement basis, in order to restrict their validity. Obviously, communication attitudes are highly context dependant, and when the social context changes (like when the agent faces a different communication partner), its communication attitudes might also change, possibly leading to inconsistent ostensible beliefs and intentions of the same agent. In this regard, we believe that existing approaches to trust and reputation could be integrated with our approach in a straightforward manner. Assuming significant spheres of reliability as a heuristics makes use of a remarkable property of communication: even if an agent cannot be trusted, she will quite likely stick to her behavior and opinions exhibited before as long as she is

[5] We could likewise talk about "assertions". We prefer the term opinion, however, since it allows us to speak about *passive* opinions as well (details on this will follow below), and avoids the fact of "something being asserted" to be mixed up with the speech act type of the same name.

[6] Whereby the term "public" should not disavow that a certain "image" of an agent could be known and intended only to/for a restricted group of agents, or be restricted otherwise to some social context.

under continuous observation.

As an important field of application besides agent communication semantics we consider ontology and knowledge acquisition and integration in open environments like the Semantic Web. Similar to the attempts to determine agents' beliefs, traditional knowledge acquisition and *multi-agent belief revision and fusion* (see, e.g., [4]) aim for determining the "truth" from the contributions of multiple sources. But settling on the opinion level, the main focus of such approaches could shift to the determination, weighting and integration of multiple individual viewpoints and "subjective truths", traditional attempts to judge the reliability of these individual contributions taking a subordinate position. To this end, we propose *opinion bases* as some sort of "multiagent belief bases" lifted to the level of communication attitudes. These are based on previous approaches to extend repositories of (RDF) statements with (i) the provenance of the respective knowledge provider, and (ii) a probabilistic weight denoting the degree of social acceptance of the respective statement (see, e.g., [17]), and on the concepts of *Open Knowledge Bases* and *Open Ontologies* [6]. To our best knowledge, this resembles the first agent-based approach to (formal) multi-source knowledge [7] which allows for the extension of knowledge contributions with probabilistic degrees of behavioral agreement (or disagreement, respectively) given to them by explicitly listed individuals or social groups with consideration and integration of communicative (i.e., social) contexts.

The rest of this paper is organized as follows: The next section introduces opinions and ostensible intentions as communication attitudes, contrasting intra-agent beliefs and intentions. Section 3 presents a logic for the representation of opinions and ostensible intentions as logical modalities, and suggests how the FIPA-ACL semantics could be "lifted" to the social level. As a further example application, section 4 shows how semantically heterogeneous sets of agent assertions can be fused, stored and rated within so-called *opinion bases*. Section 5 concludes with a summary and an outlook to open research issues.

2 Communication Attitudes

The main difference between an agent's belief and an *opinion* (ostensible belief) of her is that the former is a mental attitude, whereas the latter is a *communication attitude* which is triggered and revised by communicative acts and existing social structures (like norms), i.e. by *social conditions*. The opinion of an autonomous grey- or black-box agent neither has to reflect this agent's belief, nor does uttering an opinion necessarily mean that the agent truthfully intends to make someone belief it.

Opinions emerge from more or less hidden agent intentions and social processes, and they are tailored to the intended communicative effect and the opinions and ostensible intentions of the audience. Thus, opinions are not necessarily

[7] "Knowledge" in a weak sense, conceding that in open systems competing subjective "truths" might exist, which should nevertheless be distinguished from agent-internal *beliefs*.

governed by the principles of, e.g., perception, epistemic logic, knowledge acquisition, and belief revision, which are instead devoted to the acquisition and maintenance of truthful and consistent information. They are, however, to a certain extent determined by social conditions, including legal, organizational and economic laws. Hence, although opinions might as well reflect the true beliefs of benevolent, trustworthy agents (as assumed in most traditional approaches to agent communication semantics), this should rather be considered a special case in open systems.

From an observer's point of view, a *pro-active* opinion can be characterized as a communication attitude of an agent within a particular course of communication, denoting that the agent *allegedly intends* to make the addressed agents *act in conformance* with the expressed information. We can thus reduce pro-active opinions to *ostensible intentions*.

Informally, this ostensibly intended "acting in conformance with the opinion" means to act *to a certain degree as if* the propositional content of the uttered opinion was true. This includes (i) the adoption of the opinion as an opinion (thus, uttering an opinion means to ostensibly bring about an opinion recursively), and (ii) probably an appropriate "physical" (i.e., non-communicative) behavior. If, for example, someone asserts that the addressee will, should or has to do something, a way for the addressee to act in conformance with this opinion is to share it, and thus to actually perform the action she is said to be obliged or even willing to do. Following the utterance "You are obliged to close the window", it is not sufficient to just reply with "Yes I am". The only way for the addressee to act in conformance with the former opinion is to actually close the window (presumed that there exists an implicit time limit for obeying the order). Doing so implicitly includes expression of the (conforming) opinion "I obey". In general, it is often possible to translate opinions into requests and backwards using a scheme like

$$\text{“}\mathit{Request}(\mathit{agent}_1, \mathit{agent}_2, \mathit{action})\text{”} \rightleftharpoons$$
$$\text{“}\mathit{Assert}(\mathit{agent}_1, \mathit{agent}_2, \text{'}\ \mathit{agent}_2 \text{ has reasons to perform } \mathit{action}')\text{”}.$$

Pursuing an option can have the goal of actually convincing the addressee to believe the expressed information, but since in our scenario the ostensible are autonomous agents, the success of an opinion can only be observed indirectly and temporarily as a change of opinion of another agent. Thus, the meaning of opinions is inherently recursive: *An opinion is maintained in order to change an opinion, or to express approval with another opinion*. This reflects the insight that the structures of communication among black-box actors can not definitely be grounded in determinable causes (like agent beliefs), but only in form of more or less (un-)reliable *expectations* [16] about the future interaction behavior of the communicating agents, until, eventually, the opinion results in non-symbolic acting (like "closing the window" in the above example).

In analogy to ostensible beliefs an *ostensible intention* is a communication attitude denoting that the intention holder allegedly commits herself to striving for some desired action or goal state. This may include getting another agent to per-

form a desired action (i.e., in our framework the ostensible intention of $agent_1$ to have action *a* performed does not necessarily imply that $agent_1$ can perform *a* by herself in our framework). Just like opinions, ostensible intentions are part of the "public (or discourse) identities" of otherwise opaque agents, and they are thus in general not identical with any "true" intentions in the sense of earnest self-commitments [19].

In practice, the main difference between an intention and an ostensible intention is that the latter implies some extrovert performance towards the ostensibly desired action or state (the agent *plays act*, so to say). Such a performance may be limited to a very short period of time (e.g., think of provocations and joke questions), and the effort an agent puts into the realization of an ostensible intention might also be deliberatively limited, namely if the agents does not sincerely mean to do so. But it might likewise be possible that an ostensible opinion or belief is much more stable than any mental attitude of the same agent. For example, the communicative act *request* denotes the ostensible intention of the uttering agent to make the addressee fulfil the request, and an observer can rather safely conclude from the observation of the request that the agent will in some way strive for the fulfilment of the request as long she is in social contact with the addressee (e.g., in terms of mutual observation).

It should be noted in this regard that the addressed agents might believe that a particular ostensible intention is not an "honest" intention, and the uttering agent might again know about this, without necessarily making communication dysfunctional. In fact, ostensibility does not necessarily imply deception, but only denotes that a communication attitude is bounded in terms of validity and effort-making from the more or less uninformed point of view of an observer. To some extent, this reflects the conception of *intentional stances* [3]. An ostensible intention or belief can be known for sure to be feigned but still be useful, be it just for the reason that even such an attitude is likely to determine the acting of the agent who claims it *for social reasons*, as long this agent is under observation.

Although every ostensible attitude maps to mental attitudes (which, regarding their desired effects, might be contrary to the ostensible attitudes), human communication to a certain extent works without the need to reason about these "background intentions", if only to reduce the complexity of cognitive reasoning while communicating [14]. This makes communication processes the constituents of systems which are to some degree independent from their participants' mental states (the so-called *autopoiesis* of communication), and it makes communication attitudes largely (but not fully, of course) determined by social (i.e., communicative) events.

A makeshift, which is quite often used at least informally, would be to "misuse" mental attitudes in combination with additional assumptions like the trustability of "whole agents" to model agent interaction instead. A related, but more appropriate approach would be to ascribe "weak" intentions to an observed agent that are not necessarily pursued until the intended effect is eventually achieved or impossible [2]. This approach would assumably be workable, but it would again yield the prob-

lem of mixing the levels of cognition and sociality, and force the socially reasoning communication observer to find an immediate, possibly complicated mapping of alleged attitudes to intentions. If someone asserts "$x = y$", it is certainly a good idea to reason about this assertion "as is" as long as possible instead of trying to boil it down to some set of weak intentions immediately. Nevertheless, in order to give a proper semantics for communication attitudes, we would need to provide a mapping from assertions to behavioral expectations, as outlined later in this paper. We think that the amalgamation of pseudo mental and pseudo communication attitudes, which is still quite common, is dangerous and inappropriate. It is likely to impede the simultaneous reasoning about mental and social issues, and it is misleading with respect to the very specific nature of communication processes, which form a system of their own that is in a certain sense decoupled from mental cognition [14]. To this end, ostensible beliefs and intentions shall provide a second set of attitudes cleanly separated from mental attitudes, but still resembling properties and features of "real" mental attitudes. This does not mean to ground ACL semantics in *conventions* [12] only. We do not require communication acts to be sincere (or their content being reliable), but we assume that *rational* communicating agents act in conformance with their ostensible intentions and beliefs as they show up to an observer in their assertions (requests, commands...) at least to a limited amount, shifting the focus of communication semantics from *a priori* to *a posteriori* modeling. Opinions and ostensible intentions are thus *ostensible* in the sense of being not necessarily undesigning, opinions are at the same time *subjective statements* in the sense that they (their content, respectively) are not necessarily believed to be true by the observer. In addition, both attitudes are assumed to be *rational* in the sense that their claim or alleged intention is pursued by the utterer in some foreseeable, but probably limited manner, given their content and social context, i.e., each opinion and ostensible intention can be annotated with its own degree (sphere) of reliability. This view is heavily influenced by the more radical approach suggested in [15], but focusses on the logical modeling of communication states instead of the determination of these states.

Social and institutional commitments [20] are another, more recent and - in our opinion - more adequate approach to the modeling of agent communication (as compared to traditional mentalist approaches). But we also feel that in some scenarios this so-called *social semantics* does not provide a sufficiently rich and powerful model of real-world communication, and it is still not completely clear what "being committed" means or should mean exactly. A usual way to put commitment-based semantics into practice is thus to prescribe a semantics of communication which deliberatively limits itself to the expression of the meaning of communication in terms of (themselves rather under-specified) commitments.

On the contrary, the expressiveness of mentalistic semantics is very high (when applicable), modeling *entire agents* situated in interaction. Our work aims at lifting most of this expressiveness to the social level by modeling the rational public identities of the agents (called *social agents* in [16]) in terms of their ostensible beliefs and intentions. In addition to that, mentalistic approaches have the appeal of being

widely used and rather intuitive. In section 3.2, we will see how a subset of FIPA's communication acts semantics (which is mentalistic) can be "lifted" to the level of ostensible beliefs and opinions. This is not to provide a full-fledged alternative ACL semantics, but to hint how a semantics grounded in communication attitudes could look like.

To establish communication attitudes, we take the viewpoint of an observer which selectively overhears communication processes and models and revises from these observations a current "communication attitudes store" of the communication system. This observer can either be passive, or be one of the agents participating in the observed interaction. We associate opinions with *social contexts* (or "contexts" for short) which partially or fully describe the communicative state in which the respective attitude is maintained and provide expected (i.e., possibly uncertain) conditions for the initiation and termination of the respective attitude (its logical validity, respectively). In this paper, we model the social conditions of attitudes both logically and procedurally (using Dynamic Logic) as sequences of observable actions and other events, plus the uttering and addressed agents (or agent groups). Alternative approaches to context modeling are imaginable and have been proposed in speech act theory, but are not investigated in this paper for lack of space.

The validity of communication attitudes of course not only depends on their contexts, but also on the observed agent's mental states. However, an observer in an open system has to rely on publicly visible events to determine relevant conditions and consequences. According to sociological approaches to communication like *social systems theory* [14], the knowledge of communication processes alone is, in general, sufficient to derive reliable (but not necessarily fully certain) communication structures from an observers' point of view, since communication (such as that of opinions) relies heavily on previous and expected future communication to be understood and effective [14,15].

More precisely, social conditions obtained in terms of observed and anticipated communication acts, other kind of observable events, and logical conditions trigger and end communication attitudes. These can in turn contribute to the social conditions for other communication attitudes (cf. section 3.2 for a basic approach to such trigger events). They also determine the dynamic, possibly uncertain, and revisable *expectation* about how long a newly triggered attitude is likely to be maintained, and what concrete actions will *expectedly* accompany the sustainment of this attitude (the run-time effects of holding the respective attitude, and thus its meaning in a consequentialist, pragmatic sense). Behavioral *expectations* [16] provide a general semantics of communication attitudes and, indirectly, one of communication acts which can be given a semantics in terms of communication attitudes. Strictly speaking, logical conditions that constrain communication attitudes (like "Agent holds opinion $a \rightarrow \neg$ Agent holds opinion $\neg a$") can also be provided as fallible expectations only, since agents are fully autonomous to violate every communication model. We will nevertheless provide a number of such conditions as axioms, justified by the fact that communication attitudes need to be reliable, inherently logically and intentionally consistent (at least to some minimal degree) to be usable for

reasonable interaction.

An opinion held by an agent in a certain state of interaction can thus informally be written as a tuple $Op(Provenance, Target, Proposition)$ holding given *Precondition*, where *Provenance* is an unambiguous information source identifier (ostensible intentions can be treated analogously). Technically, the provenance could be represented as a URI, e.g. denoting an artificial or human agent, a group of agents, a web site, or a peer node within a Peer-to-Peer system. If anonymous contributions are desired, nicknames could be used, although this might reduce the reliability of knowledge rating techniques (cf. section 4), as known from the spamming of Google. *Target* denotes the addressee(s) of the opinion. For example, an agent might maintain two inconsistent ostensible opinions at the same time, facing two different other agents in two parallel discourses, whereby our observer overhears both. *Proposition* is a logical statement, the actual content of the opinion, *Precondition* denotes, as a social context, a course of action which is a precondition for holding this opinion (from the viewpoint of the observer, which may be wrong) [8].

To sum up, ostensible beliefs and intentions

- are triggered, constrained and ended to a significant amount by *social conditions* like communication acts and their contexts, norms, organizational structures, etc.;
- provide a *second set* of attitudes as a *verifiable, social-level replacement* for mental attitudes;
- are restricted by logical conditions like opinion consistency given certain social contexts;
- provide a means for modeling the *public (or discourse) identities* of mentally opaque agents. With context-dependant communication attitudes, a single agent can exhibit multiple, possibly inconsistent external identities subsequently or even at the same time (e.g., facing different conversational partners);
- can (but not must) be semantically grounded in *expectations* regarding the respective agent's future behavior; and
- can be optionally refined using additional information like the reputation of an opinion holder or the reliability of the respective opinion.

3 A Dynamic Logic with Modalities for Opinions and Ostensible Intentions

In the following, we suggest *Propositional Dynamic Communication Attitude Logic* (pDCAL) as an extension of *Propositional Dynamic Logic* (PDL) [9] with additional modalities for opinions and ostensible intentions. The intuitive meaning of

[8] This could also be a possible (but not necessary) course of events in the *future*, but this aspect is not covered in this paper.

these modalities has already been described. Our main goal here is to "lift" the usual modal logic and protocol languages for modeling agent beliefs and intentions and their revision (see, e.g., [10]) to the communication level, in order to deal with opaque agents in open systems. Within the scope of this work, we have chosen Dynamic Logic as a basis for this, but other formalisms (like the event calculus [13]) could certainly be used instead. As we have seen in the previous section, their dependency on social conditions is a crucial property of communication attitudes, demarcating them, inter alia, from mental attitudes (which also have contexts, of course, but of a different kind). Since an exhaustive description of all the possible ways by which communication attitudes are steered by social conditions of various kinds is far beyond the scope of this paper, we confine ourselves to the presentation of simple trigger events (communication acts modeled using FIPA communication acts), which is after all in line with the FIPA semantics.

3.1 Syntax

Definition 3.1 The syntax of well-formed pDCAL formulas $F, F_1, F_2, \ldots$ and processes $\alpha, \beta, \ldots$ is given by

$$F, F_1, F_2, \ldots ::= p \mid \top \mid \bot \mid \neg F_1 \mid F_1 \wedge F_2 \mid F_1 \vee F_2 \mid F_1 \rightarrow F_2 \mid F_1 \leftrightarrow F_2 \mid \langle \alpha \rangle p \mid done(\alpha) \mid Op(a_1, a_2, F) \mid OInt(a_1, a_2, F) \mid Bel(a, F) \mid Int(a, F)$$

$$\alpha, \beta, \ldots ::= action \mid a_i.action \mid any \mid \alpha;\beta \mid \alpha \cup \beta \mid \alpha* \mid F?$$

Here,

- $a, a_i \in Sources$ are *opinion sources*, i.e., an agent, an agent group (denoted as a set of agents), or a (group of) non-agent opinion resources like peers in a Peer2Peer network or web documents;
- p is a proposition;
- $action, a_i.action \in A$, with A denoting the set of elementary actions; every elementary action can be indexed with an agent (e.g., $agent_3.assert$ represents the communication act "assert" uttered by $agent_3$);
- *any* is an arbitrary action;
- $\alpha;\beta$ denotes sequential process combination (i.e., (sub-)process α is followed by β);
- $\alpha \cup \beta$ denotes non-deterministic choice between α and β;
- $\alpha*$ denotes zero or more iterations of α;
- $F?$ is a test operation (i.e., the process proceeds if F holds true);
- $\langle \alpha \rangle p$ denotes that α can be processed, and p holds afterwards;
- $Op(a_1, a_2, F)$ denotes that agent a_1 holds the opinion (ostensible belief) F facing agent a_2 [9];

[9] This opinion can of course be dropped in other contexts. Also see section 3.2 for a differentiation

- $OInt(a_1, a_2, F)$ hold iff agent a_1 facing agent a_2 exhibits the intention to make F come true (i.e., a_1 holds the ostensible intention that he will perform/make true F by himself);
- $Bel(a, F)$ denotes that agent a (sincerely) believes that F; and
- $Int(a, F)$ denotes that agent a (sincerely) intends that F.

$\neg, \wedge, \top, \bot, \rightarrow$ and $\leftrightarrow$ shall have the usual meaning. See below for the meaning of *done* (adopted from FIPA-SL semantics [5]). Further, let $Sources = \{a_i\}$ the set of all provenance identifiers, $\Theta = \{\alpha, \beta, \gamma...\}$ the set of all syntactically valid processes, and $\Phi = \{F, F_1, F_2, ...\}$ the set of all syntactically valid pDCAL formulas.

For convenience, we also define $holder(Op(a_1, a_2, F)) = a_1$, $addressee(Op(a_1, a_2, F)) = a_2$ and $content(Op(a_1, a_2, F)) = F$.

3.2 Axioms and Properties

We do not aim at a complete axiomatization of our logic here. Instead, we define basic axioms, based on KD45 and the usual other axioms for Belief-Intention modalities. We also propose some further properties, not all being deducible from the axioms or worth making them axioms themselves (among other reasons, because the FIPA-like communication acts we use lack a suitable, consented semantics), but nevertheless reasonable in our opinion. Axioms for mental beliefs and mental intentions could be trivially added to our language and have been omitted here for lack of space.

Axioms

(1) (*Kop*) $Op(a_1, a_2, F_1 \rightarrow F_2) \rightarrow (Op(a_1, a_2, F_1) \rightarrow Op(a_1, a_2, F_2))$

(2) (*DOp*) $Op(a_1, a_2, F) \rightarrow \neg Op(a_1, a_2, \neg F)$

(3) (*4Op*) $Op(a_1, a_2, F) \rightarrow Op(a_1, a_2, Op(a_1, a_2, F)) \wedge Op(a_2, a_1, Op(a_1, a_2, F)))$

(4) (*5Op*) $\neg Op(a_1, a_2, F) \rightarrow Op(a_1, a_2, \neg Op(a_1, a_2, F))$

(5) $OInt(a_1, a_2, F_1 \rightarrow F_2) \rightarrow (OInt(a_1, a_2, F_1) \rightarrow OInt(a_1, a_2, F_2))$

(6) $OInt(a_1, a_2, F) \rightarrow \neg OInt(a_1, a_2, \neg F)$

(7) $Op(a_1, a_2, OInt(a_1, a_2, F)) \leftrightarrow OInt(a_1, a_2, F)$

(8) $Op(a_1, a_2, \neg OInt(a_1, a_2, F)) \leftrightarrow \neg OInt(a_1, a_2, F)$

(9) $OInt(a_1, a_2, F) \rightarrow \neg Op(a_1, a_2, F)$ (or $Op(a_1, a_2, F) \rightarrow \neg OInt(a_1, a_2, F)$)

(10) $OInt(a_1, a_2, Op(a_2, a_1, F)) \rightarrow Op(a_1, a_2, F)$

(11) $Op(a_1, a_2, F) \rightarrow \langle a_2.Query(a_1, F); a_1.Reply(a_2, F) \rangle \top$

Axioms 1 to 4 correspond to the usual axioms K, D, 4 and 5 of a modal logic of belief. Axiom 4Op also reflects that opinions are directed from a sender to an addressee, and are visible for this addressee.

---

between *passive* and *active* opinions.

The remaining axioms concern additional properties of opinions, ostensible intentions and the relationship between them. Axiom 11 states that agents (have to) admit their opinions at least upon request. This denotes passive opinions, whereas pro-active opinions (denoted as Op^{active}) correspond to the ostensible intention to spread the opinions pro-actively among their respective addressees. For pro-active opinions, $OInt(a_1,a_2,Op(a_2,a_1,F)) \leftrightarrow Op^{active}(a_1,a_2,F)$ holds (cp. axiom 10), i.e., pro-active opinions can be modeled in terms of ostensible intentions.

We further propose the following set of axioms to bridge the gap between ostensible beliefs and ostensible intentions on the one hand and mental attitudes on the other. We think, however, that in most cases these will not be needed to reason about communication processes.

(12) $OInt(a_1,a_2,F) \leftrightarrow Bel(a_2,OInt(a_1,a_2,F))$

(13) $OInt(a_1,a_2,F) \leftrightarrow Bel(a_1,OInt(a_1,a_2,F))$

(14) $Op(a_1,a_2,F) \leftarrow Bel(a_2,Op(a_1,a_2,F))$ (i.e. opinions are not necessarily visible, but if an agent comes to believe that someone holds an opinion this belief is justified)

(15) $Op(a_1,a_2,F) \leftrightarrow Bel(a_1,Op(a_1,a_2,F))$

(16) $Op(a_1,a_2,Bel(a_1,x) \leftrightarrow y) \leftrightarrow Op(a_1,a_2,y)$

Properties

The following property expresses that holding an ostensible intention regarding a sequence of actions means that the agent truly intents a prefix of this sequence. As a heuristic, at least those actions are truly intended that need to be performed under mutual observation of the communicating agents (e.g., $h \geq 1$ in the following instead of $h \geq 0$ if the next announced action is under observation).

$$OInt(a_1,a_2,done(\langle action_1;action_2;...;action_n\rangle)) \rightarrow \exists h \geq 0 : Int(a_1,done(\langle action_1;...;action_h\rangle))$$

Here, $done(action)$ denotes that *action* has been executed successfully.

If the mental state of a_1 is not accessible, the value of the *sphere h* could be determined context-depended via empirical observation of past ostensible intentions and their consequences in terms of actions executed by a_1. A similar sphere can analogously be defined for pro-active opinions, i.e., $OInt(a_1,a_2,Op(a_2,a_1,F_1))$, and would require us to determine the course of actions performed by a_1 allegedly towards getting this opinion accepted. Since the same agent might be trustable in one context and cheat in another, these spheres allow for a much more accurate analysis than the assignment of reliability and trustability to "whole agents". But even the the relationship of intentions and ostensible intentions suggested above might be still not fine-grained enough, since it does not determine *how* (in a qualitative sense) ostensible intentions are put into action in a concrete situation.
The above mapping of ostensible to mental intentions would already provide a very simple semantics for ostensible intentions and beliefs, but it would come

at the price that it would anchor communication attitudes in mental attitudes. Another problem would be that no concrete method for the determination of the sphere *h* is given. For these reasons, we prefer a semantics of *Oint* and *Op* based on *expectations* [16], i.e., to provide the meaning of communication attitudes in terms of the expected behavioral consequences of holding the respective attitude, represented as a probability distribution of observable events. One way to achieve this practically at run-time is basically to obtain the sequences of the communications which had brought about the communication attitudes (cf. below in regard to FIPA-style communication acts), and then to use the empirical semantics algorithm [15] to calculate extrapolations (i.e., predicted continuations) of these sequence. In addition to such empirical expectations, *normative* and adaptive-normative expectations [16] can be used to describe how an agent which demonstrates a certain communication attitude *should* behave. Alternatively, an adoption of the usual Kripke-style semantics for beliefs and intentions [8,10] is also imaginable.

The following properties denote relationships between communication attitudes and selected communicative acts in terms of pre- and post-conditions (FP, RE). While these acts shall not denote FIPA communication acts, they deliberatively resemble FIPA-ACL syntax and are give in FIPA style using (non-exhaustive) pre- and post-conditions (denoted FP and RE, respectively) [5]. This is not to provide an alternative semantics, but to give some intuitive basics on how such a semantics in terms of communicative attitudes could look like. Trust and sincerity axioms are not required.

- $a_1.inform(a_2,F)$ (agent a_1 informs agent a_2 that F)

  FP: $\neg Op(a_2,a_1,F) \wedge \neg Op^{active}(a_1,a_2,F)$

  RE: $Op(a_1,a_2,F) \wedge OInt(a_1,a_2,Op(a_2,a_1,F))$

  In RE, $Op(a_1,a_2,F)$ is redundant (cf. axiom 10), but stated for clarity.
  In absence of trustability or expectations, the use of communication instead of mental attitudes would thus make the FIPA-inform essentially a request do adopt an opinion (a "requesting assertion", respectively) with a largely uncertain effect on the addressee (in case no previous knowledge about the agents' truthfulness etc. is given).

- $a_1.request(a_2,\alpha)$ (agent a_1 requests that agent a_2 performs α)

  FP: $\neg OInt(a_1,a_2,done(a_2.\alpha))$

  RE: $OInt(a_1,a_2,done(a_2.\alpha))$

  Again, the postcondition is a statement about how a_1 is situated after the request, while the corresponding post-conditions of the FIPA semantics describe the ef-

fect on the addressee (which is simply $Bel(a_2, OInt(...)) \wedge Op(a_2, a_1, OInt(...))$, according to the axioms).

In order to provide a means for the agreement or disagreement with asserted knowledge, the following additional communicative acts will be useful:

- $a_1.approve(a_2, F)$ (agent a_1 informs agent a_2 that she agrees F, F asserted previously by agent a_2)

  FP: $OInt(a_2, a_1, Op(a_1, a_2, F))$ (which implies $Op(a_2, a_1, F)$ according to axiom 10)

  RE: $Op(a_1, a_2, F)$

- $a_1.disapprove(a_2, F)$

  FP: $OInt(a_2, a_1, Op(a_1, a_2, F))$

  RE: $Op(a_1, a_2, \neg F)$

As already mentioned before, we have two different kinds of opinions now: *pro-active opinions*, which have been pro-actively expressed in terms of an assertion, and *passive opinions*, which might have only been uttered as a reply or (dis-)agreement (the former occurs if *approve* is used to answer a question). We can write $OInt(a_1, a_2, Op(a_2, a_1, F))$ to explicitly express pro-active opinions (Op^{active}). Of course, and in accordance with axiom 10, these are (passive) opinions as well.

If a passive opinion is expected without having asked the opinion holder, we talk about a *latent opinion*. In some cases, e.g., when a communication process is not fully observable, or is expected to occur in the future, reasoning about the *probabilities* associated with passive opinions might become necessary, but otherwise, it can be safely assumed that latent opinions are uttered (verifiable) at least on request (like in an opinion poll).

As it can be seen, the semantics of communicative acts in terms of opinions and ostensible intentions is rather lightweight. Thus, to provide a rich semantics for communicative acts, the observable process the acts are embedded in has to be evaluated, in order to compensate for the lack of information about agents' mental states that is inherent in open systems. In particular, such an approach would have to provide a means for i) tracking opinions, ii) recognizing and predicting opinion changes caused by a change of context, iii) identifying the particular actions an agents carries out to get her opinions accepted, and iv) reacting appropriately to other agents' opinions.

4 Acquiring and fusing heterogeneous opinions

As an example application of pDCAL and the communication act semantics proposed in 3.2, we now present an approach to knowledge bases and ontologies for knowledge acquired from multiple, heterogeneous sources (a group of agents, for example), with the special features that i) (first order) statements are weighted probabilistically regarding their degrees of ostensible communicative (i.e., social) acceptance, that ii) a knowledge base or ontology might contain mutually inconsistent statements (whereas the knowledge base itself is not inconsistent, since the statements are reified by opinion modalities), iii) every statement is extended with its provenance (i.e., the asserting agent, or another source of opinion, like communities, organizations or data bases), and iv) the bases can be socially contextualized using (fragments of) discourses.

Definition 4.1 An *opinion base* is a tuple $OpBase(\rho, \phi, filter, weight, rows)$, where

- $\rho \in \Theta$ is a (possibly non-deterministic, i.e., concurrent) communication process;
- $\phi \in 2^{\Phi}$ is a set of pDCAL formulas such that for all elements F of ϕ, F has the form $\langle filter(\rho)\rangle F'$;
- $filter : \Theta \rightarrow \Theta$ is a function which maps a process to another process such that a set of zero or more elementary actions in the origin process are replaced by $\top$? (e.g., all actions of a particular agent whose opinions should be omitted from the knowledge base), in order to tailor the opinion base to a set of interesting or relevant contributions, topics or provenances;
- $weight : \Phi \rightarrow [0;1]$ is an assessment function for the rating of knowledge contributions; and
- $rows : \{((supporters_s, opponents_s), s, w_s) | supporters_s, opponents_s \in 2^{Sources}, s \in \Phi, w_s \in [0;1]\}$ is the "output" of the opinion base with

$$\begin{aligned} supporters_s &= \{a | \exists o \in O : (content(o) \models s) \wedge holder(o) = a\}, \\ opponents_s &= \{a | \exists o \in O : (content(o) \models \neg s) \wedge holder(o) = a\}, \\ w_s &= weight(s), \text{ and} \\ O &= \{Op(a_1, a_2, F) | \exists F \in \phi : \phi \models Op(a_1, a_2, F)\}. \end{aligned}$$

As one among several particular *weight* functions that appear reasonable, we propose the following valuation which assigns assertions the *degree of ostensible consensus* among those sources who have an opinion about it (positive or negative):

$$weight_A(s) = \frac{|supporters_s|}{|supporters_s| + |opponents_s|}$$

with $supporters_s$ and $opponents_s$ defined as above. A different definition of *weight* can be found below.

If the opinion base viewer is mainly interested in the reliability of particular opinions instead of the unbiased distribution of social support, she could multiply every $weight(s)$ with a factor corresponding to the reliability of either one of

the opinion holders of s (e.g., $weight(s) = weight_A(s) \cdot max(\{reputation(a_i) | a_i \in supporters_s\})$, or with a reliability factor for every statement, obtained, e.g., from background knowledge, via user feedback or voting.

We expect that different users of opinion bases would create different definitions for *filter* and *weight* in order to tailor the opinion base to their personal viewpoints and needs.

Note that unlike in "pure" Bayesian belief aggregation systems (e.g. [18]), the weights here express the probability that one member chosen by chance out of the group of supporters and opponents of the respective statement supports the statement *communicatively in a concrete social situation*. Opinion bases are therefore closer related to behavioral reasoning.

Example 4.2 Consider the following set of sales talks, advertisements and consumer chats, observed until a certain time point, and denoted as a linear program:

$$
\begin{aligned}
&a_1.inform(\{a_2,...,a_7\}, fast(digicam));\\
&a_2.inform(\{a_3,...,a_7\}, \neg fast(digicam));\\
&a_3.inform(\{a_1,a_2,a_4,...,a_7\}, \neg fast(digicam));\\
&a_4.inform(\{a_1,a_2\}, \neg fast(digicam) \wedge heavy(digicam));\\
&a_5.inform(\{a_1,...,a_7\}, fast(digicam));\\
&a_1.inform(\{a_1,...,a_7\}, costs(digicam, 120£));\\
&\qquad a_2.inform(\{a_3,...,a_7\}, costs(digicam, 138£));\\
&\qquad\quad a_1.approve(a_2, costs(digicam, 138£));\\
&a_6.request(a_1, handy-camera);\\
&\qquad a_1.inform(a_6, \neg heavy(digicam));\\
&a_7.request(a_1, solid-camera);\\
&\quad a_1.inform(a_7, heavy(digicam))
\end{aligned}
$$

Table 1 shows the resulting opinion base contextualized by the communication subprocesses above, given $weight := weight_A$ and $filter := id$. For propositions, we use a FOL syntax. It can be seen that a_1 takes two inconsistent opinions, depending on if he faces a_6 or a_7. In contrast, he corrects ("overwrites") himself when confronted with the information by a_2 about the price of the camera.

Example 4.3 Alternatively, we can re-define the assessment function as follows:

$$weight_B(s) = \frac{|supporters_s|}{|all|}$$

with *all* being the set of all opinion holders (not only of opinion s), which is $\{a_1,a_2,...,a_5\}$ in our example. As shown in table 2, this results in a Google-like opinion ranking when ordered by weight.

Supporters vs. Opponents	Assertion	Social approval
$\{a_1, a_5\} vs. \{a_2, a_3, a_4\}$	$fast(digicam)$	0.4
$\{a_2, a_3, a_4\} vs. \{a_1, a_5\}$	$\neg fast(digicam)$	0.6
$\{a_4\} vs. \{a_1, a_5\}$	$\neg fast(digicam) \wedge heavy(digicam)$	0.33
$\{a_1\} vs. \{a_4, a_1\}$	$\neg heavy(digicam)$	0.33
$\{a_1, a_4\} vs. \{a_1\}$	$heavy(digicam)$	0.66
$\{\} vs. \{a_1, a_2\}$	$costs(digicam, 120£)$	0
$\{a_1, a_2\} vs. \{\}$	$costs(digicam, 138£)$	1

Table 1
Opinion base contextualized by the communication sub-processes of example 4.2

Supporters	Assertion	Ranking
$\{a_2, a_3, a_4\}$	$\neg fast(digicam)$	0.6
$\{a_1, a_5\}$	$fast(digicam)$	0.4
$\{a_1, a_4\}$	$heavy(digicam)$	0.4
$\{a_1, a_2\}$	$costs(digicam, 138£)$	0.4
$\{a_4\}$	$\neg fast(digicam) \wedge heavy(digicam)$	0.2
$\{a_1\}$	$\neg heavy(digicam)$	0.2
$\{\}$	$costs(digicam, 120£)$	0

Table 2
Opinion base contextualized by the communication sub-processes of example 4.2, but with the alternative *weight* function of example 4.3 and *Opponents* omitted.

5 Conclusion

Finding appropriate computational approaches to the challenges posed by open environments (like truly autonomous artificial and human agents, and the emergence of meaning and knowledge from interaction) is definitely a great challenge and requires long-term research efforts. We hope that the approach outlined in this work will be a fruitful contribution. On the one hand, agent behavior and knowledge from autonomous sources need to provide a stable grounding for interaction and applications. On the other hand, agents in open environments are self-interested grey or black boxes which cannot and should not be fully controlled as a matter of fact. Opinions in open scenarios can be unreliable, inconsistent information emerging from a potentially large number of competing goals and viewpoints, and a priori there might be no such thing as a commonly agreed “truth” or common interests.

To move towards a solution for these problems must be a core concern of a strictly communication-oriented paradigm of agent systems engineering and of application fields like agent-mediated knowledge management, and has been the basic motivation underlying the work described here. Future works in this regard could include the further exploitation of communication attitude semantics, the practical evaluation in the context of real-world applications, and the potential integration with the related grounding-based approach to discourse modeling [11].
With this work, we explicitly do not want to provide mechanisms to judge which opinions are worth keeping, or which of them are "reliable", "valuable" or even "true", as it is sometimes done in the context of collaborative belief integration and revision frameworks (see, e.g., [18,4]), and approaches to trustability and reputation (see, e.g., [7,1]). While we think that these approaches are extremely useful, the primary goal of our current and future work in the area of agent opinions is to map divergent opinion systems to their social semantics, like their provenances, the social communities and conditioning processes of their utterance, the amount and kind of support they receive from others, their emergence from and dissemination via communication, and the consensus or dissent they give rise to.

References

[1] Cobos, R. and X. Alaman, *Knowcat: a knowledge crystallisation tool*, in: *Proceedings of the International Conference e-Business and e-Work*, 2000.

[2] Cohen, P. R. and H. J. Levesque, *Intention is choice with commitment*, Artificial Intelligence **42** (1990), pp. 213–261.

[3] Dennett, D., "The Intentional Stance," The MIT Press, Cambridge, MA, 1987.

[4] Dragoni, A. F. and P. Giorgini, *Belief revision through the belief-function formalism in a multi-agent environment*, in: J. P. Müller, J. Wooldridge and N. R. Jennings, editors, *Intelligent Agents III* (1997), pp. 103 – 116.

[5] FIPA (Foundation for Intelligent Physical Agents), *Communicative act library specification SC00037J*, http://www.fipa.org/specs/fipa00037/.

[6] Froehner, T., 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)*, 2004.

[7] Golbeck, J., B. Parsia and J. Hendler, *Trust networks on the semantic web*, in: *Proceedings of the 7th International Workshop on Cooperative Information Agents (CIA)*, 2003.

[8] Goldblatt, R., *Logics for time and vomputation*, CSLI, 1987.

[9] Harel, D., D. Kozen and J. Tiuryn, "Dynamic Logic," MIT-Press, 2000.

[10] Herzig, A. and D. Longin, *A logic of intention with cooperation principles and with assertive speech acts as communication primitives*, in: *Proceedings of the First*

International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), 2002.

[11] Gaudou, B., A. Herzig, and D. Longin. *A logical framework for grounding-based dialogue analysis*. In this volume.

[12] Jones, A.J.I. and X. Parent, *Conventional signalling acts and conversation*, in: *Proceedings of the AAMAS-03 Workshop on Agent Communication Languages and Conversation Policies*, 2003.

[13] Kowalski, R. A., and M.J. Sergot, *A logic-based calculus of events*, New Generation Computing **4-1**, pp. 67–96, 1986.

[14] Luhmann, N., "Social Systems," Stanford University Press, Palo Alto, CA, 1995.

[15] Nickles, M., M. Rovatsos and G. Weiss, *Empirical-rational semantics of agent communication*, in: *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2004.

[16] Nickles, M., M. Rovatsos and G. Weiss, *Expectation-oriented modeling*, Engineering Applications of Artificial Intelligence (EAAI) (2005), to appear.

[17] Nickles, M. and G. Weiss, *A framework for the social description of resources in open environments*, in: *Proceedings of the 7th International Workshop on Cooperative Information Agents (CIA)*, 2003.

[18] Richardson, M. and P. Domingos, *Building large knowledge bases by mass collaboration*, Technical Report UW-TR-03-02-04, Dept. of CSE, University of Washington (2003).

[19] Shoham, Y., *Agent-oriented programming*, Artificial Intelligence **60** (1993), pp. 51–92.

[20] Singh, M. P., *A social semantics for agent communication languages*, in: F. Dignum and M. Greaves, editors, *Issues in Agent Communication*, Springer-Verlag, 2000 pp. 31–45.