Empirical Semantics of Agent Communication in Open Systems

Matthias Nickles
AI/Cognition Group
Department of Informatics
Technical University of Munich
nickles@cs.tum.edu

Gerhard Weiss
AI/Cognition Group
Department of Informatics
Technical University of Munich
weissg@cs.tum.edu

ABSTRACT

The paper proposes a novel approach to the semantics of communication of self-interested and autonomous agents in open systems. It defines the semantics of communicative acts primarily as the observable effect of their actual use in social encounters, and differs thus fundamentally from mentalistic and current objectivist approaches to the semantics of agent communication languages. Empirical communication semantics enables the designer of agent-oriented software applications as well as agents to reason about social structures on the level of dynamically formed expectations, which we consider to be a crucial capability especially for social reasoning within and about open systems with truly autonomous black- or gray-box agents.

Keywords

Agent Communication Languages, Open Systems

1. INTRODUCTION

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. Due to agent *autonomy*, which is certainly the crucial property of artificial agents, agents need to communicate to achieve their goals interactively, where communication has to be seen in contrast to the exchange of information among ordinary objects. While at first glance it seems to be 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), as it has been done using, e.g., first order predicate calculus for KIF [8], there is still no general model of the actual usage and effects of utterances in social encounters, a field which is traditionally studied in linguistical pragmatics and sociology. Currently, two major approaches to the meaning of agent communication in a broader sense, covering both traditional sentence semantics and pragmatics, exist, if we do not count plain interaction protocols (in some sense very simple social semantics. Interaction protocols can for example be used to provide a partial semantics for FIPA-ACL [5, 13]) and 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, and thus indirectly their behaviour), while the more recent approaches (e.g. [1, 4]) try to determine communication from an *objectivistic* point of view, focussing on public 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 in favor of traditional sentence-level semantics, and it doesn't have a concept of meaning dynamics and generalization (cf. below). 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 heterogenous, 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 reminder of this paper is organized as follows: The next section provides a more detailed descriptions of the difficulties of normative and non-normative meaning assignment. Section 3 outlines our semantical approach in response to these issues. Finally, section 4 draws some conclusions regarding future work.

2. ISSUES IN ACL SEMANTICS FOR OPEN ENVIRONMENTS

The meaning of utterances has two dimensions that need to be covered by a comprehensive approach to the semantics of agent communication: 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 application domain 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 in ACL research. This is mainly due to agent autonomy, which makes it 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). In open systems, this is very problematic for reasons pointed out in the following, although of course, not all of the described issues do deserve attention in all application domains.

2.1 Dynamics and Indefiniteness of Meaning

The meaning of an utterance can most generally be characterized as its consequences, i.e., the course of events caused by it [16, 15, 17]. The currently most advanced approaches to ACL semantics determine this meaning as the result of a functional application of the respective speech act to the current discourse context, which has been derived as the outcome of the previous speech acts (e.g., [4]). The discourse context is thereby seen as a set of facts (e.g., about anaphora, roles and other social structures), which serve as parameters for the application of the speech acts, while the speech acts themselves function as a fixed mapping rule being statically assigned to the ACL words. Such approaches work if the set of speech act types is known a-priori, each allowed utterance denotes a fixed and known illocutionary act, and the constraining a certain discourse context imposes on the course of future speech acts is sufficient regarding the design goals, normative and known from the beginning on. They do not work, if these conditions are not known a-priori to either the system designer or the agents, or if they vary in a manner during run-time of the MAS which is not foreseeable at design-time, as it can occur in heterogenous, open systems.

2.2 Speech Act Taxonomy

The usual set of performative types current ACLs provide represents an important, but nevertheless incomplete part of the set proposed by speech act theory [9, 10]. A performative has primarily the purpose to make the illocutionary act an agent performs explicit, and thus the set of performatives available to black-box agents would have to cover every possible communicational intention, which might be hardly foreseeable. As a work-around for this problem, contemporary agent languages like FIPA-ACL allow for an extension of the basic set of performatives, but not in computational adoption to run-time meaning dynamics. Another observation which has been made in this respect is that the quite strict distinction traditional ACLs draw between performative and content level (as with the duo KQML and KIF), is quite arbitrary [12], as well as the boundary of assertive and non-assertive performatives in speech act theory [11].

2.3 Rationality in Interaction

Communication has an unique property: It constructs a social situation, which is inherently consistent and reasonable, even if it opposes the "real world" outside communication and the cognitive beliefs of the agents: 1) Communicated information is supposed to be consistent with information previously communicated by the same agent, or this agent at least justifies his change of mind, 2) the agent defends and asserts his utterances by means of argumentation or other rational means like rewards and sanctions, and 3) information not expressed explicitly can be deduced from information communicated before and background knowledge. If, for example, in an open auction on the internet some agent a asserts "I will deliver the goods if you win the auction.", an observer does not need to believe him. But the observer believes that the further communication of a complies with this assertion. To make communication work, this belief is to some extent independent from reasoning about the true motives "within the agents mind". Agent a is supposed to act at least for some time in a rational manner in accordance with the social image that he *projects* for himself by means of communication (e.g., a sanctions the denial of his proposal, rewards its acceptance etc). The information about such (bounded-)rational attitude is implicitly associated with each communication of a self-interested agent, and is thus part of communication semantics. Some approaches to ACL semantics somehow target this kind of social rationality via the introduction of social commitments, but this term is quite underspecified [12], and often comes along with mental concepts like the "whole-hearted satisfaction" [2] to equip commitments with reliable intentions.

2.4 Generalization of Meaning

Current approaches to ACL semantics are intended primarily for dyadic situations. Some of them allow for message broadcasts, but they lack a concept for unification and weighting of multiple messages or, respectively, responses, to reflect a (possibly inconsistent) common point of view of multiple agents, or to enable collaboration in joint communicative action. It is hardly imaginable, how thousands or even millions of agents shall contribute to web semantics, if information agents are unable to generalize upon their communications by means of statistical evaluation. Of course, such abilities could be provided through special frameworks like recommender systems, but in fact, generalization is an inevitable part of communication meaning. In human communication, generalization of knowledge especially plays an important role for the assignment of meaning to underspecified messages, and the temporal generalization of communication trajectories (i.e., the prediction of their effects from experience) in our opinion even provides the most appropriate semantics for agent communication in open systems, as we will specify next. In response to the described issues, we propose the following aspects a semantical framework should consider for being suitable in open environments.

- *Expectation-orientation and evolution of semantics.* The meaning of utterances lies primarily within their expectable *consequences* and might *evolve* during system operation. If no a-priori assumptions can be made, empirical evaluation of observations of message exchange among agents is the only feasible way to capture this meaning.
- *Support for generalization.* In complex or underspecified social situations, it is required to generalize expec-

tations regarding observed utterances to assign them a reasonable (however revisable) default meaning.

- *Rational attitude.* Mentalistic approaches to agent semantics are not feasible for black-box agents , but nevertheless, even for such agents rational planning and acting can be assumed. Therefore, semantics assignment has to consider "social rationality", i.e., (bounded-)rational attitude in behavior.
- *Dynamic, context-driven performative ontology.* To overcome restrictions due to a pre-defined set of performatives, we plead for dynamic performative types which gain their illocutionary force from the social context of their actual use.
- *Run-time derivation and propagation of semantics.* Since empirical communication semantics are obtained dynamically at run-time, a technical facility is required to derive and, if this facility is not located within the agents themselves, to propagate them similar to dictionaries and grammars for human languages. Focussing on the design process of open systems, we've developed such an instance in form of a special middle-agent (the so-called *Social System Mirror* [15, 16, 17]).

3. EMPIRICAL MODELLING OF COMMUNICATION

The communication model we propose is grounded in *Social Systems Theory* [14], as it has been introduced for the empirical derivation of communication structures of artificial agents [15, 16]. This model is based on the assumption that sociality in truly open systems is basically the observable result of communication, and therefore social structures (e.g., roles, public ontologies, organizational patterns) can be represented by means of probabilistic *expectation structures* regarding future communication processes, which are dynamically extrapolated from observed message trajectories. This observer is usually an agent which participates in the communication himself, but it could also be the system designer or a human application user.

In our model [17], which we can only sketch very briefly and informally in this work for lack of space, a single communication attempt can be seen as a declarative request to act in conformance with the information asserted by an utterance (our approach does not follow speech act theory here because in our model an utterance always performs a declaration). 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 (otherwise the addressee could have derived its truth/usefulness herself and a communication would thus be 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 whole of the expectations which are triggered by a communication in the context of the preceding communication process we call its *rational hull*. The rational hull specifies the traditional speech act context (anaphora, common ground information like public ontologies etc.), and, more important, the rational social relationships which steer the acceptance or denial of communicated content according the rational attitudes the agents exhibit. Typically, rational hulls are initially very indefinite and become increasingly definite in the course of interaction, provided that the agents work towards some (temporary) mutual understanding.
Our model is centered around the following terms, which we propose primarily as "empirical replacements and supplements" for terms associated with traditional ACL semantics, like "message content" and "commitment".

Expectation structures The empirically obtained probability distribution of all future event sequences resulting from the observance of the agents and their environment. It needs to be adopted for each newly observed event and subsumes both communicative actions and ordinary "physical" events. Expectation structures can be modelled in multiple levels of generalization to enable the description of underspecified utterances, communication patterns, collaborative meaning (e.g., opinions of agent groups) and agent roles (which are basically generalizations of single-agent behavior). Expectation structures also provide for the common ground which contextualizes communication acts.

Social expectation structures The part of the expectation structures consisting of social expectations that result from communication processes and constrain future communications. The effect a certain utterance brings about in terms of social expectation structures is the semantics of this utterance. In contrast, *physical expectation structures* are domain-dependent expectation structures (i.e., ontological information about the agents environment). Physical expectation structures include purely normative interaction structures but no social expectations.

Utterances An agent action event with the following properties: 1) it is occurring under mutual observation, 2) without considering social expectation structures, the event would have a very low probability, 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 technical message transmission, criteria 1) to 3) are clearly met.

Projections Our equivalent to the terms "performative" *and* "content". A projection is a part of the expectation structures the uttering agent selects as an assertion regarding the future of the communication system to make the addressee act in conformance with it (e.g., fulfill a request). The observer needs to compare the utterance with both social and physical expectation structures to find the most probable match. The most basic kind of projection is obtained through *demonstrative acting*, where the uttering agent encodes its message by means of "playing-act".

Rational hulls The rational behavior an agent is expected to perform to propagate, support and defend a certain

uttered projection (e.g., sanctions or argumentation). 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 the respective projection. A commitment can be seen as one possible means to such maximization: An agent commits herself to perform certain actions to bring about a certain behavior from another agent.

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 own utterances (e.g., he does not contradict himself), and 2) each projection is consistent with 1), i.e. it does not deny that property 1) is met. Criterium 2) is somehow an empirical kind of mental *understanding*: With each communication, an agent acknowledges implicitly that the other agent tries to get accepted his projections, even if he does not agree with these projections.

Communication spheres A communication process together with the expectation structures arising from this process. Whereas a communication sphere is basically the same as the *interaction system* known from social systems theory [14], and resembles some of the properties of *spheres of commitment* [7], in our model it has dynamic, empirically discovered boundaries in the sense that communications which do not fulfill the consistency criteria for communication processes are not part of the respective sphere (but they can be modelled using higher-order communication structures). Examples for communication spheres are open markets and auctions (cf. [16] for a case study on empirical expectation-oriented modelling of a trading platform) and (agent-supported) forums on the internet.

Higher-order communication structures Social expectations which govern multiple communication spheres at the same time.

4. CONCLUSION

We've proposed a novel approach to the semantics of agent communication, based on the empirical evaluation of observed messages. We believe that this approach adopts better in comparison to traditional mentalistic or objectivist semantics to the peculiarities of large, heterogenous open system, as it enables both the software designer and the agents themselves to analyze the meaning of messages on the level of empirical expectations without the need to know about mental agent properties or architectural details. The biggest challenge for future research efforts in this respect is the development of an expectation-oriented ACL, respectively the assessment of an empirical (or semi-empirical, semi-normative) semantics to traditional ACLs. For this purpose, we are currently investigating the use of flexible behavioral patterns assigned to performatives of FIPA-ACL [5] to serve as a default semantics for empirical meaning assignment.

Acknowledgements: This work has been supported by the German National Science Foundation (DFG) under grant BR609/11-2. Thanks to Michael Rovatsos for useful comments on the topic of this work.

5. REFERENCES

[1] M. P. Singh. A social semantics for agent communication languages. In Proceedings of the IJCAI Workshop on Agent Communication Languages, 2000.

[2] M. P. Singh. Multiagent Systems - A Theoretical Framework for Intentions, Know-How, and Communications. LNCS 799, Springer, New York, 1994.

[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] 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.

[5] FIPA, Foundation for Intelligent Agents, http://www.fipa.org.

[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. Singh. Multiagent Systems as Spheres of Commitment. In Proceedings of the ICMAS Workshop on Norms, Obligations, and Conventions, 1996

[8] M. R. Genesereth, R. E. Fikes. Knowledge Interchange Format, Version 3.0 Reference Manual. Technical Report Logic-92-1, Stanford University, Stanford, 1992.

[9] J. L. Austin. How to do things with words. Clarendon Press, Oxford, 1962.

[10] J. R. Searle. A taxonomy of illocutionary acts. In K. Gunderson (Ed.), Language, mind, and knowledge, pages 344–369. University of Minnesota Press, 1975.

[11] K. Bach, R. Harnish. Linguistic Communication and Speech Acts. MIT Press, Cambridge, Massachusetts, 1979.

[12] 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.

[13] J. Pitt, F. Bellifemine. A Protocol-Based Semantics for FIPA97 ACL and its Implementation in JADE. Technical report, CSELT S.p.A.,1999.

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

[15] 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.

[16] 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, Berlin, 2001.

[17] M. Nickles. An Observation-based Approach to the Semantics of Agent Communication. Research Report FKI-24x-03, AI/Cognition Group, Technical University of Munich, 2003. To appear.