

# An Approach to Broaden the Semantic Coverage of ACL Speech Acts

Hong Jiang and Michael N. Huhns

Department of Computer Science and Engineering,  
University of South Carolina, Columbia, SC 29208, USA  
{jiangh, huhns}@engr.sc.edu

**Abstract.** Current speech-act based ACLs specify domain-independent information about communication and relegate domain-dependent information to an unspecified content language. This is reasonable, but the ACLs cover only a small fraction of the domain-independent information possible. As a key element of modern ACLs, the set of communicative acts needs to be as complete as possible to enable agents to communicate the widest range of information with agreed-upon semantics. This paper describes a new approach to broaden the semantic coverage of ACL speech acts. It provides agents with the ability to express more of the possible meanings in human languages and yields a more powerful ACL. Specifically, we first compare Austin's and Searle's classifications, and Ballmer and Brennenstuhl's comprehensive classification of speech acts. The main meaning categories and their semantics are given next. Finally, a multifaceted evaluation of our approach is presented, which points out that the approach potentially can combine the benefits of the FIPA ACL with Ballmer and Brennenstuhl's speech act classification, resulting in a more expressive ACL.

## 1 Introduction

Agent communication languages (ACLs) are a critical element of multiagent systems and a key to the successful application of agents in commerce and industry. Modern ACLs, such as the FIPA ACL, provide a standardized set of performatives denoting types of communicative actions. Such ACLs have been designed as general purpose languages to simplify the design of multiagent systems. However, recent research shows that these ACLs do not support adequately all relevant types of interactions. Serrano and Ossowski [1] report a need for new *ad hoc* sets of performatives in certain contexts, which the FIPA ACL does not support. Singh [2] points out that agents from different vendors or even different research projects cannot communicate with each other. In [3], Kinny shows that FIPA reveals a confusing amalgam of different formal and informal specification techniques whose net result is ambiguous, inconsistent, and certainly underspecified communication. He proposes a set of requirements and desiderata against which an ACL specification can be judged, and briefly explores some of the shortcomings of the FIPA ACL and its original design basis.

Therefore, a complete set of speech acts as communicative acts in an ACL would be desirable in order to improve understanding among the agents in a multi-agent system. Recognizing that the ~4800 speech acts in [8] would be desirable but impractical to use individually, we describe a feasible approach to broaden the semantic coverage of ACLs by formalizing speech act categories that subsumes the ~4800, enabling the meanings of all the speech acts to be conveyed. Different from [11], we focus on the standard messages used for communication instead of designing a conversation protocol.

Specifically, Section 2 compares Austin's, Searle's, and Ballmer and Brennenstuhl's classifications of speech acts. Based on an abstract model that separates protocols, agent types, and decision mechanisms from communications, Section 3 describes the main meaning categories. In Section 4, we use FIPA's formal semantic language to represent the semantics of our speech act categories. This enables our approach to combine the benefits of the FIPA ACL with a broader set of speech acts. An evaluation of this approach is discussed in Section 5.

## 2 Comparison of Austin's, Searle's and Ballmer's Classification

Current ACLs derive their language primitives from the linguistic theory of speech acts. The original speech act theory was developed by Austin [4]. The most important part of Austin's work was to point out that human natural language can be viewed as *actions* and people can perform things by saying. According to his theory, a speech act has three aspects, as summarized in [7]: Locution, Illocution, and Perlocution. Austin also tried to classify speech acts. He classified illocutionary acts as verifictives, exercitives, commissives, behabitives, and expositives [4]. The classification has been criticized for overlapping categories, too much heterogeneity in categories, ambiguous definitions of classes, and misfits between the classification of verbs and the definition of categories [8, 13, 12].

Austin's work was extended by Searle [5, 14, 6, 13], who posited that an illocutionary speech act forms the minimum meaningful unit of language. He classified speech acts into five categories: Assertives, Directives, Commissives, Declaratives, and Expressives. Searle's speech act theory focuses on the speaker. The success of a speech act depends on the speaker's ability to perform a speech act that should be understandable and successful.

Ballmer and Brennenstuhl [8, 12] criticize the clarity of Searle's classification, definition of declaratives as a speech act type, principles used in the classification, selection of illocutionary verbs from all verbs, vague definition of the illocutionary point, and vagueness of the line between illocutionary force and propositional content. Based on their criticism, they propose an alternative classification of speech acts, which contains both simple linguistic functions such as expression and appeal, and more complex functions such as interaction and discourse. Models for alternative actions are formed and verbs are classified according to the phases of the model.

Ballmer and Brennenstuhl’s classification motivates us to rethink the speech acts used in ACLs. Since the classification is based on an almost complete domain (~4800 speech acts) and the authors claim they provide a “theoretically justified” classification that is “based explicitly and systematically on linguistic data,” we believe that to generate a speech act set for ACLs based on their classification will be a powerful way to represent meaning. However, this classification is not perfect: the classification for English is obtained by translating the verbs of the German one, the names of the categories are not systematically chosen, and there are no formal semantic representation for the categories. However, by rebuilding the categories, most of above problems can be fixed. Thus, we endeavor herein to derive a reasonable set of categories for agent communication from their theory, and to give a formal semantics using more typical English names.

### 3 Method Description

#### 3.1 Abstract Communication Model

Based on current popular communication models, we generalize an abstract communication model for agents consisting of agents and environments. An environment constrains the agents and affects the communications among them via a message control mechanism that provides protocols, routing, and message delivery. This gives flexibility to a multiagent system at an abstract level, in which the message control mechanism could be any one of many possibilities.

Following similar choices made for KQML and the FIPA ACL, the mechanism specifies from whom did a message originate, to whom should the message be sent, what information is being communicated, and how should the message be delivered. The “what” is separated into a communicative act and content, where the communicative act reflects domain *independent* information to be communicated, and content reflects domain *dependent* information. The message may also separately specify the language used in the content, encoding and decoding functions, ontology, and protocol.

We focus on the communicative acts. They should be in a formal form, as FIPA provides well. Further, they should cover all the possible domain-independent meanings in human communication, as we address next.

#### 3.2 Overview of Meaning Categories

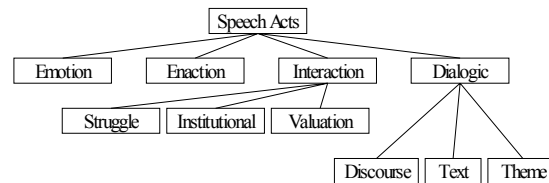
This section describes semantic categories for a relatively complete set of human speech activity verbs, derived from the classification in [8]. The categories reflect an ontological and a conceptual structuring of linguistic behavior. The main categories and their relationships are represented in Figure 1. The topmost node, Speech Acts, represents the entire set of speech acts in human language. The four major groups—Emotion Model, Enaction Model, Interaction Model, and Dialogic Model—represent four basic functions of linguistic behavior.

The *Emotion Model* is the least hearer-oriented and least extroverted function of the four, and focuses on representing the kinds of emotional states of a human or agent.

The *Enaction Model* is a function clearly directed toward a hearer. In other words, the speaker tries to get control over the hearer.

The *Interaction Model* is a function involving speaker and hearer in mutual verbal actions. Among this group are three sub-categories to represent different degrees of the mutual competition: Struggle Model, Institutional Model, and Valuation Model. In the Struggle Model, the speaker tries to get control over the hearer, or the speaker is more competitive in controlling mutual verbal actions. In contrast, the hearer is more competitive in the Valuation Model. In the Institutional Model, the hearer and speaker are equally competitive.

The *Dialogic Model* covers a kind of reciprocal cooperation where there is a better-behaved and more rigidly organized verbal interaction. Its three sub-categories focus on different types of the content and the organization: the Discourse Model focuses on the organization and types of discourse, the Text Model focuses on the textual assimilation and processing of reality, briefly, specific knowledge involved, and the Theme Model focuses on the process of thematic structuring and its results, in other words, the structure or organization of some knowledge system.



**Fig. 1.** Ontology of the Main Speech Act Categories

In the above ontology, the four basic models can be divided into unilateral and multilateral models. The Emotion Model and Enaction Model are unilateral, because they focus on a single speech action. The Interaction Model and Dialogic Model are multilateral because they consider the response from a hearer. The Emotion Model and Interaction Model are more original and racy, and the Enaction Model and Dialogic Model are more institutionalized and controlled. Practically, these four basic models may be combined.

## 4 Semantics of Meaning Categories

### 4.1 Formal Semantic Model Notations

The semantic model used in representing the categories in this paper follows the formal semantic language as described for the FIPA ACL [9]. Components of the formalism are

- $p, p_1, \dots$  are closed formulas denoting propositions;
- $\phi, \psi$  are formula schemes, which stand for any closed proposition;
- $i, j$  are schematic variables denoting agents.

The mental model of an agent is based on four primitive attitudes: belief (what the agent knows or can know); desire (what the agent desires); intention (which is defined as a persistent goal that could lead to some actions); and uncertainty. They are respectively formalized by operators  $B$ ,  $D$ ,  $I$ , and  $U$ :

- $B_i p$  agent  $i$  (implicitly) believes (that)  $p$ ;
- $D_i p$  agent  $i$  desires that  $p$  currently holds;
- $I_i p$  agent  $i$  intends a persistent goal  $p$ ;
- $U_i p$  agent  $i$  is uncertain about  $p$ ;

To enable reasoning about action, we also introduce operators *Feasible*, *Done*, and *Agent*:

- $Feasible(a, p)$  means that an action  $a$  can take place and, if it does, then  $p$  will be true.
- $Done(a, p)$  means that when  $p$  is true, then action  $a$  takes place.
- $Agent(i, a)$  means agent  $i$  performs action  $a$ .

Generally, the components of a speech act model involved in a planning process should contain both the conditions that have to be satisfied for the act to be planned and the reasons for which the act is selected. The former is named *FP* (feasibility preconditions), and the latter *RE* (rational effect) in FIPA ACL. We use the same model here, which is represented as follows:

$$\begin{aligned} &< i, act(j, C) > \\ FP &: \phi_1 \\ RE &: \phi_2 \end{aligned} \tag{1}$$

where  $i$  is the agent of the act, or speaker,  $j$  the recipient or hearer, *act* the name of the speech act,  $C$  stands for semantic content, and  $\phi_1$  and  $\phi_2$  are propositions.

## 4.2 Emotion Model

The Emotion Model focuses on representing the emotional states of a human or agent. We assume there is a finite set of emotions,  $E$ , represented as

$$E = \{e_+, e_0, e_-\} \tag{2}$$

where  $e_+$  is an emotion in the set of positive emotions, which is characterized by or displaying a kind of certainty, acceptance, or affirmation (about the content involved), such as  $\{happy, love, \dots\}$ ;  $e_0$  is in the set of neutral emotions, which does not show any tendency, such as  $\{hesitate, \dots\}$ ;  $e_-$  is in the set of negative emotions, which intend or want to express a kind of negation, refusal, or denial, such as  $\{angry, sad, afraid, \dots\}$ .

**Table 1.** Foundational Meaning Units of Emotional Speech Acts

+	0	-
happy	N/A	sad
love	N/A	hate
excited	nervous	angry
desire	hesitate	fear
N/A	shocked	N/A

The Emotion Model can be represented as follows:

$$\begin{aligned}
 & \langle i, em(j, \phi) \rangle \\
 & FP : \neg B_i(B_j Agent(i, em(\phi))) \wedge D_i(B_j Agent(i, em(\phi))) \\
 & RE : B_j Agent(i, em(\phi))
 \end{aligned} \tag{3}$$

where  $em \in E$ , and the semantic content  $\phi$  could be empty. This model represents that agent  $i$  sends a message to  $j$  that  $i$  has emotion  $em$  about  $\phi$  or  $i$  is in status of  $em$  when  $\phi$  is empty. The FP shows that, when agent  $i$  does not believe agent  $j$  knows about  $\phi$ ,  $i$  is currently in emotion  $em$  about  $\phi$ , and  $i$  desires that  $j$  knows it, then this message could be sent. The RE shows that the desired result is that agent  $j$  believes that  $i$  is in emotion  $em$  about  $\phi$ .

To simplify usage of this model, we could directly use  $e_+$ ,  $e_0$ , or  $e_-$  as communicative acts. In this case, we focus on the effect of the emotion speech act on the content  $\phi$ . That is, for a positive effect,  $i$  hopes  $j$  will increase its intention on  $\phi$ ; for a negative one,  $i$  hopes  $j$  will decrease its intention on  $\phi$ ; for a neutral one,  $i$  shows its attitude is uncertain about  $\phi$ . Just as for human interactions, we do not have to know the precise value of an attitude. Instead, we just need to know that something is viewed favorably, unfavorably, or neutrally.

However, detailed emotions are also desirable in some cases. To make this usable, we generate a set of foundational meaning units from 155 emotion speech acts listed in [8]. Table 1 gives the foundational meaning units of emotion with consideration of positive, neutral, and negative values.

In Table 1, each row represents one kind of meaning unit. In the first row, *sad* has the opposite meaning of *happy*. *Hate* has the opposite meaning of *love* in the second row. *Excited* represents an opposite attitude to something with strong feeling, *nervous* represents a strong uncertain feeling about something, and *angry* represents a strong negative feeling about something. In the fourth row, *desire* shows a feeling to get something, *hesitate* shows no intentions, and *fear* shows a feeling to avoid something. In the last row, *shocked* shows a neutral feeling about surprise.

### 4.3 Enaction Model

In the Enaction Model, the speaker more or less coercively attempts to get the hearer to do something by expressing an idea, wish, intention, proposal, goal,

etc. There are many speech acts in this group. To organize them and simplify the usage, we define the set of enactions as:

$$EN = \{en_+, en_-\} \quad (4)$$

Unlike the Emotion Model, which focuses on presenting a kind of description or knowledge, the Enaction Model tries to make the hearer do something. Thus, there are no neutral enactions: if agent  $i$  does not want  $j$  to do anything,  $i$  does not have to send any message to  $j$ .  $en_+$  is an action in the set of positive enactions, such as  $\{intend, desire, ask\ for, encourage, \dots\}$ ;  $en_-$  is an action in the set of negative enactions, such as  $\{warning, cancel, \dots\}$ .

The Enaction Model can be defined as:

$$\begin{aligned} & \langle i, en_{\pm} (j, \phi) \rangle \\ FP : & \neg B_i \phi \wedge D_i \phi \wedge B_i (B_j \phi \wedge \neg D_j \phi) \\ RE : & Agent (j, en_{\pm}(\phi)) \end{aligned} \quad (5)$$

where  $en_{\pm} \in EN$ . This model represents that agent  $i$  sends a message to  $j$  to ask  $j$  to do  $en_{\pm}$  on  $\phi$ . The FP shows that this message could be sent when  $i$  does not believe that  $i$  can do  $\phi$  and it desires  $\phi$ . On the other hand,  $i$  believes that  $j$  can do it, but  $j$  does not want to do it. The expected result is  $j$  does  $en_{\pm}$  on  $\phi$ . Practically,  $j$  could just add the action to its action queue for a positive enaction, or delete it from its action queue for a negative enaction.

#### 4.4 Interaction Model

The Interaction Model is a function involving speaker and hearer in mutual verbal actions. First of all, we assume an interaction set  $IN$ , and for some  $in_1, in_2 \in IN$ ,  $\exists rule : in_1 \rightarrow in_2$ , such that:

$$\begin{aligned} & \langle i, in_1 (j, (a, goal)) \rangle \\ FP : & I_i goal \wedge \neg B_i a \wedge D_i a \wedge B_i (B_j a \wedge \neg D_j a) \\ RE : & Agent (j, a) \wedge \\ & (\langle j, in_2(i, (a', goal - a)) \rangle \vee \langle j, in_2(i, +) \rangle \vee \langle j, in_2(i, -) \rangle) \end{aligned} \quad (6)$$

where  $a, a'$  are actions, and  $goal$  can be looked as a plan or a sequence of actions. This model represents that agent  $i$  sends a message to  $j$  to ask  $j$  to do action  $a$  for some  $goal$ . The FP shows that  $i$  intends to achieve the  $goal$ , so  $i$  desires to do  $a$  but can't do it itself, and  $i$  believes that  $j$  can do it. However,  $j$  does not desire to do it. The expected result is  $j$  does  $a$  first, and then generates another message back to  $i$ . This reply message follows the rule  $in_1 \rightarrow in_2$ . Generally, the message has the form  $\langle j, in_2(i, (a', goal - a)) \rangle$ , which mentions that after  $j$  has done  $a$ , it generates another action  $a'$  and reduces the  $goal$ . In some special case, for example after  $j$  has done  $a$ , the  $goal$  is already achieved, then  $j$  sends back message  $\langle j, in_2(i, +) \rangle$ , where  $+$  means the  $goal$  is achieved. Another extreme case is that  $j$  finds out that the  $goal$  is impossible to be achieved, then it sends back message  $\langle j, in_2(i, -) \rangle$ , where  $-$  means the  $goal$  is unachievable.

Among this model group, there are three subcategories to represent different degrees of the mutual competition: Struggle Model, Institutional Model, and Valuation Model. In the Struggle Model, the speaker tries to get control over the hearer, or the speaker is more competitive in controlling mutual verbal actions. In this case, the rule  $in_1 \rightarrow in_2$  is decided by the speaker or sender  $i$ .

In the Institutional Model, the hearer and speaker are equally competitive. For example, the establishment of a behavior in an institution equally affects the upholders of and the participants in the institution, especially when entering an institution and thereby adopting its norms, following its norms and rules, violating them, and being pursued by the upholders of the institution. Thus, the agents  $i$  and  $j$  should have some common rule system defined in advance.

In the Valuation Model, the hearer is more competitive, so it decides which communication act will be replied. That is, the rule  $in_1 \rightarrow in_2$  is decided by agent  $j$  after its evaluation of the previous message. Details of the Valuation Model cover both positive and negative valuations of actions, persons, things, and states of affairs.

#### 4.5 Dialogic Model

The Dialogic Model tends to a kind of reciprocal cooperation, and is a better-behaved and more rigidly organized verbal interaction. For this model, we at first assume a dialogic speech act set  $DS$ , and for some actions  $d_1, d_2 \in DS$ ,  $\exists rule : d_1 \rightarrow d_2$ , such that

$$\begin{aligned} & \langle i, d_1(j, \phi) \rangle \\ FP : & \neg B_i B_j \phi \wedge D_i B_j \phi \\ RE : & B_j \phi \wedge \langle j, d_2(i, \phi') \rangle \end{aligned} \quad (7)$$

This model represents that agent  $i$  sends a message to  $j$  about  $\phi$ . For this message to be sent, agent  $i$  does not believe  $j$  believes  $\phi$ , and  $i$  desires  $j$  to believe it. The expected result is that  $j$  believes  $\phi$  and  $j$  replies to  $i$  with another message about a new  $\phi$ , which is the reasoning result of agent  $j$ , and the communicative act used in the message follows the rule  $d_1 \rightarrow d_2$ .

According to the three sub-categories, which focus on different types of content and organization, we can define three types for  $\phi$ :

- The Discourse Model focuses on the organization and types of discourse. In other words,  $\phi$  points to some kind of type or organization that is pre-defined. For example, according to the status of a discourse, it could be { *beginning discourse, being in discourse, discourse inconvenience, reconciliation of discourse, ending discourse* }; according to the attitude for some content, it could be { *accept, refuse, cancel* }; according to the number of agents involved in the discourse, it could be { *discourse with several speakers, discourse with one speaker, ...* }; or a kind of irony, joke, etc.
- The Text Model focuses on the textual assimilation and processing of reality, briefly, the specific knowledge involved. Or,  $\phi$  focuses on some description



of specific knowledge. For example, it could be perceiving reality, producing texts, promulgating texts, systematically searching for data, etc.

- The Theme Model focuses on the process of thematic structuring and its results, in other words,  $\phi$  points to some structure or organization about some knowledge system.

## 5 Evaluation

For evaluation of an ACL based on our extended classification of speech acts, we focus on the following five aspects to compare with current ACLs:

*Better coverage:* By including speech acts with approximately 4800 verbs, while current ACLs include speech acts with 20 to 30 verbs, our approach provides better coverage.

*Precise semantics:* Precise semantics is an important property for an ACL, and one of the nice features of FIPA is that it provides one for its ACL. We adapt it for the four basic categories and subcategories of our approach.

*Easy usage:* For practical reasons the ACL must be easy to use. The FIPA ACL already has many successful uses. Instead of replacing it, we substitute our speech acts and keep its message structure. We organize the speech acts as an ontology with different abstract levels, so that a user can more easily navigate through them to choose the desired ones.

*Better understood:* Easy usage requires that the ACL be well understood. However, the original categories given by Ballmer and Brennenstuhl’s classification are very poor on this point, because the classification is obtained by translating German verbs and the names of the categories are not systematically chosen. In our research, we modified their classification by using typical English names, which should be more understandable.

*Efficiency:* Efficiency is desirable for usage of an ACL. Since we have not yet deployed our approach, we can not evaluate this aspect.

For evaluation, we will have a small group of users encode the conversations among the agents in some scenarios, for example:

“Agent Bob wants to ask agent Sue to a dance, but he doesn’t want to call directly, so he decides to find out Sue’s intention in advance. If Bob knows that Sue would say ‘yes’, then he would call; otherwise, he wouldn’t bother to call. To avoid embarrassment, he decides to ask Sue’s friend agent Jill to find out if Sue is available. Based on what Jill finds out, he will decide to call or not. Meanwhile, Sue wants to go to the dance. She prefers to go with Bob, but will go with agent Jack if he asks her before Bob does.”

The users will encode the conversations using both the FIPA ACL and our ACL. We will then survey the users to find out which is easier and preferred, and analyze the resulting agents to see which are better understood, etc.

## 6 Conclusion and Future Work

In this paper, we first compare Austin’s, Searle’s, and Ballmer’s classification of speech acts. Then we provide an abstract model, which separates protocols,

agents types, and decision mechanisms from the communications, so that we can focus on common messages for communication. Based on this model, we describe the semantic categories that are derived from Ballmer and Brennenstuhl's classification, which attempts to represent all possible meanings in human language. We also give a formal representation for each category and describe the subcategories. This formal representation follows the formal semantic language used for the FIPA ACL. Thus, our approach could combine the benefits from FIPA ACL and Ballmer and Brennenstuhl's speech act classification.

Above all, our approach is theoretically more expressive in representing a broader range of domain-independent communication semantics, while remaining consistent with current approaches to ACLs. However, a comprehensive evaluation is needed, and much work remains to be done to make this approach complete and practically applicable.

## References

1. Serrano, J.M., Ossowski, S.: An Organizational Metamodel for the Design of Catalogues of Communicative Actions. *PRIMA*. (2002) 92–108.
2. Singh, Munindar P.: Agent Communication Languages: Rethinking the Principles. *IEEE Computer*, vol.31(12). (1998) 40–47.
3. Kinny, David: Reliable Agent Communication – A Pragmatic Perspective *New Generation Computing*, Vol.19(2). (2001)139–156.
4. Austin, J. L.: *How to do Things with Words*. Clarendon, Oxford, UK, 1962.
5. Searle, J.: *Speech Acts: An Essay in the Philosophy of Language*. Cambridge U. Press, 1970.
6. Searle, J.: *Expression and Meaning*. Cambridge U. Press, Cambridge, 1979.
7. Wooldridge, M.: *Introduction to MultiAgent Systems*. JohnWiley and Sons, 2002.
8. Ballmer, Th., Brennenstuhl, W.: *Speech Act Classification – A Study in the lexical Analysis of English Speech Activity Verbs*. Springer-Verlag Berlin Heidelberg New York, 1981.
9. FIPA: *FIPA Communicative Act Library Specification*. Foundation for Intelligent Physical Agent, Geneva, Switzerland, 2002.
10. Bellifemine, F., Caire, G., Poggi, A., and Rimassa, G.: *JADE: A White Paper*. TILAB “EXP in search of innovation” Journal. Vol.**3(3)**, 2003.
11. Chang, M.K., Woo, C.C.: A Speech-Act-Based Negotiation Protocol: Design, Implementation, and Test Use. *ACM Transactions on Information Systems*, Vol.**12(4)**, 360–382, 1994.
12. Esa A., Lyytinen K.: *On the Success of Speech Acts and Negotiating Commitments*. LAP'96, Tilburg, The Netherlands, 1996.
13. Searle J., Vanderveken D.: *Foundations of Illocutionary Logic*, Cambridge U. Press, London, 1985.
14. Searle J.: *A Taxonomy of Illocutionary Acts*. 1971. *Language, Mind, and Knowledge: Minnesota Studies in the Philosophy of Science*. Keith Gunderson(Ed.). Minneapolis: U of Minnesota P, 1975.