Broadening the Semantic Coverage of Agent Communicative 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. ACLs based on communicative acts 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 agent communicative acts. It provides agents with the ability to express more of the semantics of human languages and yields a more powerful ACL. We first describe the main meaning categories and semantics for an ACL, which we derive from prior work on speech-act classifications. Next, we prove the resultant semantic coverage. Finally, we present some example applications, which demonstrate that the approach can combine the benefits of the FIPA ACL with Ballmer and Brennenstuhl's speech act classification, resulting in a more expressive and efficient ACL.

1 Introduction

As a critical element of multiagent systems and a key to the successful application of agents in commerce and industry, modern agent communication languages (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 venders or even different research projects cannot communicate with each other. In [3], Kinny shows that the FIPA ACL has a confusing amalgam of different formal and informal specification techniques whose net result is ambiguous, inconsistent, and 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.

Early work on communicative act based ACLs, such as KQML and FIPA, separated the communication problem into three layers—a message transport layer providing the mechanics of a communication, a domain-independent layer of communication semantics, and a domain-dependent content layer. The ACL speech acts were intended to describe the domain-independent middle layer. The problem is that the 22 communicative acts in the current FIPA ACL cover only a small fraction of the domain-independent concepts that an agent might want to express. For example, one agent can `inform` another of a domain concept using the FIPA ACL, but cannot `promise` another something. If an agent wants to make a promise, its only recourse is to express it in the content language, for which there typically is no standardized support.

Therefore, a larger set of communicative acts would be desirable in an ACL to improve understanding among agents. Recognizing that the ~4800 speech acts in [7] 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 [10], we focus on the standard messages used for communication instead of designing a conversation protocol.

Specifically, Section 2 describes prior work on a comprehensive classification of speech acts by Austin, Searle, and Ballmer and Brennenstuhl. The main meaning categories and their semantics are given in section 3, where 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. Finally, section 4 proves the semantic coverage by comparing it with the FIPA ACL, and several example applications are described in section 5.

2 Research Background

Current ACLs derive their language primitives from the linguistic theory of speech acts, originally developed by Austin [4]. The most important part of his work was to point out that human natural language can be viewed as *actions* and people can perform things by saying. Austin also classified illocutionary acts as verfictives, 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 [7, 12, 11].

Austin's work was extended by Searle [5, 13, 6, 12], 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 [7, 11] criticize six aspects of Searle's classification: clarity, 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. They propose an alternative classification, 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 has motivated 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 "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 a German classification, the names of the categories are not systematically chosen, and there is no formal semantic representation for the categories. However, most of these problems can be fixed by rebuilding the categories. 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 Semantic Description

This section describes the semantic categories for a relatively complete set of speech activity verbs, derived from the classification in [7]. The categories reflect an ontological and a conceptual structuring of linguistic behavior. The main categories and their relationships are represented in Figure 1. The top node, Speech Acts, represents the entire set of speech acts in human language and 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 most speaker-oriented and focuses on representing kinds of emotional states.

The *Enaction Model* is a function directed toward a hearer, by which a speaker tries to control the understanding of the hearer.

The *Interaction Model* is a function involving speaker and hearer in mutual verbal actions. This group includes three sub-categories to represent different degrees of the mutual competition: (1) in the Struggle Model, the speaker tries to get control over the hearer, or the speaker is more competitive in controlling mutual verbal actions; (2)in contrast, the hearer is more competitive in the Valuation Model; and (3) 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, i.e., the 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.

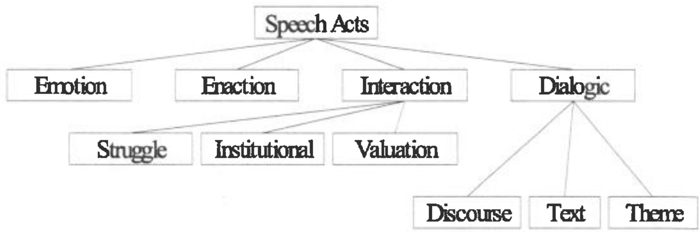

Fig. 1. Ontology of the Main Speech Act Categories

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 less constrained, and the Enaction Model and Dialogic Model are more institutionalized and controlled. Practically, these four basic models may be combined.

We next define several formal semantic model notations and then describe the detailed semantics for the meaning categories.

3.1 Formal Semantic Model Notations

The semantic model used in representing the categories in this paper follows the formal semantic language described for the FIPA ACL [8]. Components of the formalism are

- $p, p_1, \ldots$ are closed formulas denoting propositions;
- ϕ, ψ 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 (an agent's 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, but thinks that p is more likely than $\neg p$;

We use the abbreviations:

- $Bif_i \phi \equiv B_i \phi \vee B_i \neg \phi$, which means that agent i believes either ϕ or $\neg \phi$.
- $Uif_i \phi \equiv U_i \phi \vee U_i \neg \phi$, which means that either agent i is uncertain about ϕ (ϕ is more likely) or $\neg \phi$ ($\neg \phi$ is more likely).

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 is the agent who 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 termed *FP* (feasibility preconditions), and the latter *RE* (rational effect). We use the same model here, represented as

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

where i is the sender or speaker, j the recipient or hearer, *act* is the name of the speech act, C is the semantic content, and ϕ_1 and ϕ_2 are propositions.

3.2 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 $\{happy, love, ...\}$; e_0 is in the set of neutral emotions, which does not show any tendency, such as $\{hesitate, ...\}$; e_- is in the set of negative emotions, which intends or expresses a kind of negation, refusal, or denial, such as $\{angry, sad, afraid, ...\}$.

The Emotion Model is represented as follows:

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

where $em \in E$, and the semantic content ϕ can be empty. Here desire D is used instead of the stronger notion I, since emotions are easy to show for humans. This model represents that agent i sends a message to j that i has emotion *em* about ϕ or i is in the state of *em* when ϕ is empty. The FP shows that, when agent i does not believe agent j knows that i is currently in emotion *em* about ϕ, and i desires that j knows it, then this message can be sent. The RE shows that the desired result is that agent j believes that i is in emotion *em* about ϕ.

To simplify usage of this model, we can 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 ϕ. That is, for a positive effect, i hopes j knows that i has an intention on ϕ; for a negative one, i hopes j knows that i has a negative intention on ϕ; for a neutral one, i shows its attitude is uncertain about ϕ. Just like 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 [7]. Table 1 gives the foundational meaning units of emotions that combine the idea from [17, 18], and they are organized with consideration of positive, neutral, and negative values. In Table 1, each row represents a kind of

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

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 a positive 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 or some uncertainty, and *fear* shows a feeling to avoid something. In the last row, *shocked* shows a neutral feeling about surprise.

3.3 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:

$$EN = \{en_+, en_-\} \tag{4}$$

Unlike the Emotion Model, which describes emotions, the Enaction Model tries to make a 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 $\{intend, desire, askfor, encourage, ...\}$; en_- is an action in the set of negative enactions, such as $\{warning, cancel, ...\}$.

The Enaction Model can be defined as:

$$\begin{aligned}
&< i,\ en_{\pm}(j,\phi) > \\
FP:\ &\neg B_i\phi \wedge D_i\phi \wedge B_i(B_j\phi \wedge \neg D_j\phi) \qquad for\ en_{+} \\
&\neg B_i\neg\phi \wedge D_i\neg\phi \wedge B_i(B_j\neg\phi \wedge \neg D_j\neg\phi) \quad for\ en_{-} \\
RE:\ &Done(en_{\pm}(\phi))
\end{aligned} \tag{5}$$

where $en_{\pm} \in EN$. This model represents agent i sending a message to j to ask j to do $en_{\pm}$ on ϕ. The FP shows that this message could be sent for en_{+} when i does not believe that i can do ϕ and it desires ϕ, while i believes that j can do it, but j does not want to do it. FP is the same for en_{-}, except ϕ is replaced by $\neg\phi$. The expected result is that $en_{\pm}$ on ϕ is done. Practically, j could just add the action to its action queue for a positive enaction (in this case, $Done(en_{+}(\phi)) = Done(\phi)$), or delete it from its queue for a negative enaction.

3.4 Interaction Model

The Interaction Model is a function involving a speaker and a hearer mutually interacting. We assume an interaction set IN and the communicative act set $Acts$ so that $IN \subseteq Acts$, and for some $in \in IN$ and $act \in Acts$, $\exists rule : in \rightarrow act$, such that:

$$\begin{aligned}
&< i,\ in(j,(a, goal)) > \\
FP:\ &I_i goal \wedge \neg B_i Done(a) \wedge D_i Done(a) \wedge B_i(Agent(j,a) \wedge \neg D_j Done(a)) \\
RE:\ &Done(a) \wedge (< j, act(i,(a', goal - a) > \vee < j, succeed(i, goal) >) \\
&\vee < j, fail(i, goal) >
\end{aligned} \tag{6}$$

where a, a' are actions, and $goal$ can be a plan or a sequence of actions. This model represents agent i sending 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 cannot 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 $in \rightarrow act$. Generally, the message has the form $< j, act(i,(a', goal - a) >$, which means that after j has done a, it generates another action a' and reduces the $goal$. In some cases, for example after j has done a and the $goal$ is already achieved, then j sends back message $< j, succeed(i, goal) >$, which 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 $< j, fail(i, goal) >$, which means the $goal$ is unachievable.

There are three subcategories of the interaction model representing 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 $in \rightarrow act$ 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 to use in its reply. That is, the rule $in \rightarrow act$ 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.

3.5 Dialogic Model

The Dialogic Model covers a kind of reciprocal cooperation, and is a more regular and constrained verbal interaction. For this model, we at first assume a dialogic speech act set DS and the communicative act set $Acts$ so that $DS \subseteq Acts$, and for some $d \in DS$ and $act \in Acts$, $\exists rule : d \rightarrow act$, such that

$$\begin{aligned} &< i,\ d(j, \phi) > \\ &FP:\ B_i\phi \wedge D_iB_j\phi \\ &RE:\ B_j\phi \wedge < j, act(i, \phi') > \end{aligned} \tag{7}$$

For agent i to send a message to j about ϕ in this model, agent i believes ϕ, and i desires j to believe it. The expected result is that j believes ϕ and j replies to i with another message about a new ϕ, which is the reasoning result of agent j, and the communicative act used in the message follows the rule $d \rightarrow act$.

Corresponding to the three subcategories that focus on different types of content and organization, we can define three types for ϕ:

- The Discourse Model focuses on the organization and types of discourse, so ϕ points to some kind of type or organization that is predefined. 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 the specific knowledge involved, i.e., ϕ describes some knowledge about perceiving reality, producing texts, systematically searching for data, etc.
- The Theme Model focuses on the process of thematic structuring and its results, in other words, ϕ points to the structure or organization of some knowledge system.

4 Proof of Semantic Coverage

The FIPA ACL has four primitive communicative acts, and its other communicative acts are composed of the primitive acts or are composed from primitive messages by substitution or sequencing. [8] The four primitive acts are:

- The Assertive Inform:

$$\begin{aligned} &< i,\ inform(j,\phi) > \\ FP:&\ B_i\phi \wedge \neg B_i(Bif_j\phi \vee Uif_j\phi) \\ RE:&\ B_j\phi \end{aligned}$$

- The Directive Request:

$$\begin{aligned} &< i,\ request(j,a) > \\ FP:&\ FP(a)[i\backslash j] \wedge B_i Agent(j,a) \wedge B_i \neg PG_j Done(a) \\ RE:&\ Done(a) \end{aligned}$$

  where $FP(a)$ denotes the feasibility preconditions of a; $FP(a)[i\backslash j]$ denotes the part of the FPs of a that are mental attitudes of i; and $PG_i P$ means that i has P as a persistent goal.
- Confirming an Uncertain Proposition (Confirm):

$$\begin{aligned} &< i,\ confirm(j,\phi) > \\ FP:&\ B_i\phi \wedge B_i U_j\phi \\ RE:&\ B_j\phi \end{aligned}$$

- Contradiction Knowledge (Disconfirm):

$$\begin{aligned} &< i,\ disconfirm(j,\phi) > \\ FP:&\ B_i\neg\phi \wedge B_i(U_j\phi \vee B_j\phi) \\ RE:&\ B_j\neg\phi \end{aligned}$$

And among the 22 communicative acts of FIPA ACL, the composite ones corresponding to the above four primitive acts are as in Figure 2:

- Inform: accept-proposal, agree, failure, inform-if, inform-ref, not-understood, propagate, propose, proxy, reject-proposal, request-when, request-whenever, subscribe
- Request: cfp(call for proposal), query-if, query-ref
- Confirm: N/A
- Disconfirm: cancel, refuse

It can be seen that the composite communicative acts relate to the primitive ones unevenly. Most of the communicative acts are derived from *inform*, and even the primitive acts, *confirm* and *disconfirm*, are special cases of *inform*, which can be proved as follows.

Lemma 1. *In the primitive communicative acts of FIPA ACL, confirm* ($< i, confirm(j,\phi) >$) *is a special case of inform* ($< i, inform(j,\phi) >$).

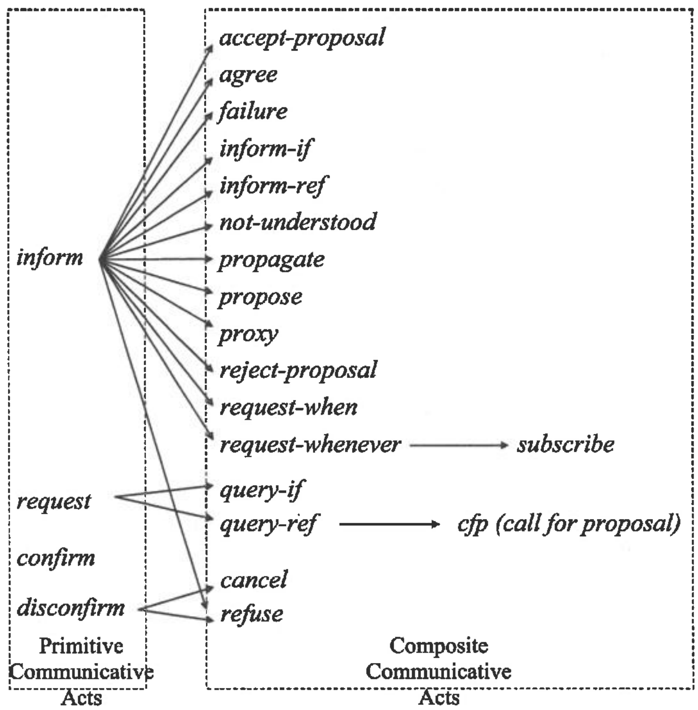

Fig. 2. Relationship of FIPA Primitive and Composite Communicative Acts

Proof. Comparing the definitions of *confirm* and *inform*, we see they have the same message body format and rational effect—*RE*. The only difference is the feasibility preconditions—*FP*. We can then try to prove that *FP* of *confirm* is a sufficient but not necessary condition of *inform*. That is, when the *FP* of *confirm* is satisfied, the *FP* of *inform* is also satisfied, or the satisfaction of *FP* of *confirm* can trigger an *inform* message; alternatively, the *FP* of *confirm* is not necessary for sending an *inform*.

FP of *inform* is:

$$
\begin{aligned}
& B_i\phi \wedge \neg B_i(Bif_j\phi \vee Uif_j\phi) \\
\equiv\ & B_i\phi \wedge \neg B_i((B_j\phi \vee B_j\neg\phi) \vee (U_j\phi \vee U_j\neg\phi)) && (8) \\
\equiv\ & B_i\phi \wedge (\neg B_iB_j\phi \vee \neg B_iB_j\neg\phi \vee \neg B_iU_j\phi \vee \neg B_iU_j\neg\phi) \\
\equiv\ & B_i\phi \wedge (\neg B_iB_j\phi \vee B_iB_j\phi \vee \neg B_iU_j\phi \vee B_iU_j\phi) && (9) \\
\equiv\ & (B_i\phi \wedge \neg B_iB_j\phi) \vee (B_i\phi \wedge B_iB_j\phi) \vee (B_i\phi \wedge \neg B_iU_j\phi) \vee (B_i\phi \wedge B_iU_j\phi) && (10)
\end{aligned}
$$

where equation (8) is derived from the definitions of Bif_ip and Uif_ip. We get equation (9) since agent i not believing j believes ϕ usually means the same as agent i believing j does not believes ϕ.

From equation (10), the last part $B_i\phi \wedge B_iU_j\phi$ is exactly the *FP* of *confirm*. When *FP* of *confirm* is satisfied, that is, when $B_i\phi\wedge B_iU_j\phi$ is true, then equation (10) will also be true. However, *FP* of *confirm* is not a necessary condition, since only if one of $B_i\phi\wedge\neg B_iB_j\phi$, $B_i\phi\wedge B_iB_j\phi$ and $B_i\phi\wedge\neg B_iU_j\phi$ is satisfied, equation (10) will also be satisfied.

Thus, *confirm* is a special case of *inform*.

Lemma 2. *In the primitive communicative acts of the FIPA ACL, disconfirm ($< i, disconfirm(j,\phi) >$) is a special case of inform ($< i, inform(j,\neg\phi) >$).*

Proof. Comparing the definitions of *disconfirm* and *inform*, we have

$$\begin{aligned} &< i,\ inform(j,\neg\phi) > \\ FP:\ & B_i\neg\phi \wedge \neg B_i(Bif_j\neg\phi \vee Uif_j\neg\phi) \\ RE:\ & B_j\neg\phi \end{aligned} \tag{11}$$

Thus, we get the same rational effect format—*RE*. Let's compare the feasibility preconditions—*FP*, and similarly equation (11) can be changed to:

$$\begin{aligned} & B_i\neg\phi \wedge \neg B_i(Bif_j\neg\phi \vee Uif_j\neg\phi) \\ \equiv\ & B_i\neg\phi \wedge \neg B_i((B_j\neg\phi \vee B_j\phi) \vee (U_j\neg\phi \vee U_j\phi)) \\ \equiv\ & B_i\neg\phi \wedge \neg B_i(B_j\neg\phi \vee B_j\phi \vee U_j\neg\phi \vee U_j\phi) \\ \equiv\ & B_i\neg\phi \wedge \neg B_i((U_j\phi \vee B_j\phi) \vee (B_j\neg\phi \vee U_j\neg\phi)) \\ \equiv\ & B_i\neg\phi \wedge (\neg B_i(U_j\phi \vee B_j\phi) \vee \neg B_i(B_j\neg\phi \vee U_j\neg\phi)) \\ \equiv\ & (B_i\neg\phi \wedge \neg B_i(U_j\phi \vee B_j\phi)) \vee (B_i\neg\phi \wedge \neg B_i(B_j\neg\phi \vee U_j\neg\phi)) \end{aligned} \tag{12}$$

From equation (12), the first part $B_i\neg\phi\wedge\neg B_i(U_j\phi\vee B_j\phi)$ is exactly the *FP* of *disconfirm*. When *FP* of *disconfirm* is satisfied, that is, when $B_i\neg\phi\wedge\neg B_i(U_j\phi\vee B_j\phi)$ is true, then equation (12) will also be true, which will trigger message $< i, inform(j,\neg\phi) >$. However, *FP* of *disconfirm* is not a necessary condition, since if $B_i\neg\phi\wedge\neg B_i(B_j\neg\phi\vee U_j\neg\phi)$ is satisfied, equation (12) will also be satisfied.

Thus, we proved that *FP* of *disconfirm* is a sufficient but not necessary condition to trigger message $< i, inform(j,\neg\phi) >$. In other words, *disconfirm* ($< i, disconfirm(j,\phi) >$) is a special case of *inform* ($< i, inform(j,\neg\phi) >$).

So far, we can conclude that there are actually two foundational communicative acts *inform* and *request*. If we can prove that our approach covers the semantic meaning of these two communicative acts, then our approach covers all the semantic meanings of the FIPA communicative acts, since the others can be derived from these two by adding constraints.

However, we think *inform* is too general. Considering $\neg B_i(Bif_j\phi \vee Uif_j\phi)$ in *FP* of *inform*, it actually lists all the possibility of j's knowledge about ϕ, such that: i does not believe j believes ϕ, or i does not believe j believes not ϕ, or i does not believe j is uncertain about ϕ, or i does not believe j is uncertain about not ϕ. Since at least one of them will be true, $\neg B_i(Bif_j\phi \vee Uif_j\phi)$ will always be true. So *FP* of *inform* can be simplified to $B_i\phi$, which is reasonable because only if agent i has the belief ϕ can it inform j about ϕ. While, we still think it should not ignore the desire to have j to believe ϕ, no matter what i believes or does not believe j's knowledge about ϕ, if i does not have any desire to have j believe ϕ, why does i want to send the message to j?

Based on the above analysis, we begin to prove that our approach covers the semantic meaning of the two foundational communicative acts *inform* and *request*.

Lemma 3. *The Dialogic model covers the semantic meaning of FIPA's inform.*

Proof. According to our above analysis of *inform*, $\neg B_i(Bif_j\phi \vee Uif_j\phi)$ did not supply any of i's opinion on j's knowledge about ϕ, and if i's desire for j to know about ϕ was also be ignored, we can then represent *inform* with more precise semantic meaning as:

$$\begin{aligned} &< i,\ \ inform(j,\phi) > \\ FP:\ &B_i\phi \wedge D_iB_j\phi \\ RE:\ &B_j\phi \end{aligned}$$

Then it has a format similar to the semantic representation of the dialogic model, and the difference is in RE. For the dialogic model, we assume a dialogic communicative act set DS and the communicative act set $Acts$ that $DS \subseteq Acts$, and for some $d \in DS$ and $act \in Acts$, $\exists rule : d \rightarrow act$, If $d \in DS$ is a terminal symbol, that is, there is no rule from d to something else, then in this case, $< j, act(i,\phi') >$ in RE can be ignored, so that we can get the same semantic meaning of *inform*. Thus, we can use the *dialogic model* to represent the semantic meaning of *inform*.

Lemma 4. *The interaction model covers the semantic meaning of FIPA's request.*

Proof. Let's first consider the definition of *request*, which is used to request a receiver to perform some action. Usually it presumes feedback from the receiver. FP of *request* involves three parts:

- $FP(a)[i\backslash j]$: denotes the part of the FPs of action a that are mental attitudes of agent i. We do not know exactly what the mental attitudes will be, however they should satisfy the following conditions for sending out a request: agent i should intend to have action a done—$I_iDone(a)$; and i can not do a by itself.
- $B_iAgent(j,a)$: i believes that j is the only agent that can perform a.
- $B_i\neg PG_jDone(a)$: this part (in page 36 of [8]) is also presented as $\neg B_iI_jDone$ (a) (in page 25 of [8]), which roughly points out a required condition: i does not believe j intents to do a.

The *goal* in the interaction model denotes a plan or a sequence of actions. To get comparable format of the interaction model, we can let the goal involve only one action, that is, let $goal = Done(a)$. Then the interaction model can be simplified to:

$$\begin{aligned} &< i,\ in(j,a) > \\ FP:\ & I_i Done(a) \wedge \neg B_i Done(a) \wedge D_i Done(a) \wedge \\ & B_i(Agent(j,a) \wedge \neg D_j Done(a)) \\ RE:\ & Done(a) \wedge < j, succeed(i, Done(a)) > \end{aligned} \tag{13}$$

In the BDI model, intention is generated from desire. If we separate desires into intentional desires (I) and non-intentional desires (NI), then we can represent $D_i p$ to be $I_i p \vee NI_i\, p$, such that equation (13) becomes

$$\begin{aligned} & I_i Done(a) \wedge \neg B_i Done(a) \wedge D_i Done(a) \wedge B_i(Agent(j,a) \wedge \neg D_j Done(a)) \\ \equiv\ & I_i Done(a) \wedge \neg B_i Done(a) \wedge (I_i Done(a) \vee NI_i\, Done(a)) \\ & \wedge B_i(Agent(j,a) \wedge \neg D_j Done(a)) \\ \equiv\ & I_i Done(a) \wedge \neg B_i Done(a) \wedge B_i(Agent(j,a) \wedge \neg D_j Done(a)) \\ \equiv\ & I_i Done(a) \wedge \neg B_i Done(a) \wedge B_i Agent(j,a) \wedge B_i \neg D_j Done(a) \\ \equiv\ & I_i Done(a) \wedge \neg B_i Done(a) \wedge B_i Agent(j,a) \wedge \neg B_i(I_i Done(a) \vee NI_i\, Done(a)) \\ \equiv\ & (I_i Done(a) \wedge \neg B_i Done(a) \wedge B_i Agent(j,a) \wedge \neg B_i I_i Done(a)) \vee \\ & (I_i Done(a) \wedge \neg B_i Done(a) \wedge B_i Agent(j,a) \wedge \neg B_i\, NI_i\, Done(a)) \end{aligned} \tag{14}$$

The first part of equation (14) has a format similar to the *FP* of *request*:

- $I_i Done(a) \wedge \neg B_i Done(a)$ corresponds to the first part of *FP* for *request*, which presents the detailed required conditions—agent i should intend to have action a done and i cannot do a by itself.
- $B_i Agent(j,a)$ is the same as the second part of *FP* for *request*.
- $\neg B_i I_i Done(a)$ is the same as the third part of *FP* for *request*. We did not use symbol *PG* in our approach, since *PG* is very similar to I, and here this part follows the format on page 25 of [8].

So far, we see that when *FP* of *request* is true, the equation (14) will also be true, and message of *interaction model* will be triggered. However the *FP* of *request* is not necessary for equation (14) to be satisfied.

Let's continue to consider the *RE* of *request*, which is the same as the first part of *RE* of *interaction model*. However, the second part is also reasonable for *request*, since in most cases *request* implies feedback from the receiver.

Thus, *request* is a special case of *interaction model*, and the *interaction model* covers the semantic meaning of *request* in the FIPA ACL.

In summary, our approach covers the semantic meaning of the two foundational communicative acts, so it also covers all the semantic meanings of the communicative acts in the FIPA ACL. Moreover, our approach also covers additional semantic meanings. For example, our emotion model supplies a way to communicate emotions, which the FIPA ACL does not. We believe it is important to cover emotions in agent communicative acts, and other researchers [14–16] have already discovered that emotions influence human decision-making, but this influence has traditionally been ignored.

5 Example Applications

This section provides several examples showing how these defined semantic categories can be used.

Example 1: Bob tells Sue that he loves her. Using the emotion model, the sender is Bob, the receiver is Sue, and $\phi =$ Sue to yield the message on the left below. The expected result will be that Sue has a belief that Bob is in love with her. Since the FIPA ACL does not have a communicative act with a similar meaning, the content must include the expression of emotion, as shown in the message on the right.

(love	(inform
:sender Bob	:sender Bob
:receiver Sue	:receiver Sue
:content (Sue))	:content (Bob loves Sue))

The left message separates domain independent from domain dependent information better and is less ambiguous.

Example 2: Jack commands Bill to turn off the TV. Using the enaction model, the message to be sent is

```
(command
:sender Jack
:receiver Bill
:content (turn off the TV))
```

The expected result will be that Bill turns off the TV. The communicative act *command* implies a master-slave relationship between the sender and receiver. The FIPA ACL does not have a similar communicative act, so all the information must be put in the content, as in ***Example 1***, although it is more ambiguous.

Example 3: Bob and Jack work together to open a case with ID 011. Bob gets the key but it is broken. Jack is an expert in fixing keys, so Bob asks Jack to fix the key.

According to the interaction model, the message sent to Jack will be

```
(interact
:sender Bob
:receiver Jack
:content (fix key) (open case 011))
```

The goal "open case 011" implies a sequence of actions, which are assumed known to both sender and receiver in advance. Thus Jack tries to fix the key. If Jack fixes the key successfully, he will send a reply to Bob that Bob can pick up the key to open the case now, as shown in the message below on the left. If Jack cannot fix the key, he will then tell Bob that the goal failed, as shown on the right.

(interact
:sender Jack
:receiver Bob
:content (pick-up key)
  (open case 011 - fix key))

(fail
:sender Jack
:receiver Bob
:content (open case 011)

This model is especially useful for multiple agents working together on a project.

Example 4: Bill wants to tell Bob about the structure of the subway system in Boston, which includes the red line, orange line, green line, and blue line. According to the dialogic model, the message sent to Bob would be

(structure
:sender Bill
:receiver Bob
:content (Subway in Boston: red line, orange line, green line, blue line)

The expected result will be that Bob records the structure information as one of his beliefs. We can also use FIPA's *inform* to represent the above message, but the relationship of the subway system and those lines would have to be part of the content.

6 Conclusion and Future Work

Comparing our approach to the FIPA ACL reveals that:

Better coverage: our approach covers more of human semantics.

Precise semantics: we adopt the same formalism as used by FIPA for our four basic categories and subcategories.

Easy usage: An ACL must be easy to use, and the FIPA ACL 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 or agent 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 poor, because the classification is obtained by translating German verbs and the names of the categories are not chosen systematically. We modified their classification by using typical English names, which should be more understandable.

Efficiency: Efficiency is desirable for an ACL. As can be seen in the above examples, our approach separates domain-independent from domain-dependent information better, which can shorten the message sent while improving the semantics.

In summary, our approach combines the benefits of the FIPA ACL and Ballmer and Brennenstuhl's speech act classification. It is more expressive in representing a broader range of domain-independent communication semantics, while remaining consistent with current approaches to ACLs. However, a better communicative act set with reasonable size still needs work. Instead of just considering the categories, some frequently used speech acts also need to be found to represent the categories.

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. 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.
8. FIPA: FIPA Communicative Act Library Specification. Foundation for Intelligent Physical Agent, Geneva, Switzerland, 2002.
9. Bellifemine, F., Caire, G., Poggi, A., and Rimassa, G.: JADE: A White Paper. TILAB "EXP in search of innovation" Journal. Vol.**3(3)**, 2003.
10. 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.
11. Esa A., Lyytinen K.: On the Success of Speech Acts and Negotiating Commitments. LAP'96, Tilburg, The Netherlands, 1996.
12. Searle J., Vanderveken D.: Foundations of Illocutionary Logic, Cambridge U. Press, London, 1985.
13. 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.
14. Antonio R. Damasio: Descartes' Error : Emotion, Reason, and the Human Brain, Avon Books, New York, 1994.
15. D. M. Wegner, T. Wheatley: Apparent mental causation: Sources of the experience of will. American Psychologist, Vol.**54(7)**, 480-492,1999.
16. Antoine Bechara: The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. Brain and Cognition, Vol.**55**, 30-40, 2004.
17. Li Yun (Ceremenial Usages: Their Origin, Development, and Intention). The Book of Rites, 200 B.C. http://www.sacred-texts.com/cfu/liki/index.htm
18. Paul Ekman and Richard J. Davidson: The Nature of Emotion: Fundamental Questions. Oxford University Press, New York, 1994.