# Constitutive Rules for Agent Communication Languages

Jeremy Pitt
Intelligent & Interactive Systems Group, Department of Electrical & Electronic Engineering
South Kensington Campus, Exhibition Road, Imperial College London, SW7 2BT, England
Email: j.pitt@imperial.ac.uk URL: http://www.iis.ee.imperial.ao.uk/~j.pitt

## Abstract

We follow Searle's contention that speaking a (natural) language is to engage in a rule-governed form of behaviour, and that those rules are conventional (institutional) rather than natural or physical. We show how this analysis can also be used to specify rules of interaction for systems of electronic agents communicating with an artificial language. We conclude that using constitutive rules to define the semantics of an agent communication language not only distinguishes agent communication from method invocation, but also offers significant computational advantages over using intentional states.

## 1 Introduction

We maintain that there is a distinct conceptual and functional difference between communication using an ACL (Agent Communication Language) and communication using an API (Application Programmer Interface). Method invocation (via an API) is essentially *perlocutionary,* that is, the [4]speech act' (the method call) is completely definable in terms of its perlocutionary effect, i.e. the further consequences or effects on the receiving (remote) object. This property makes method invocation practical for client-server distributed systems, because the execution of the remote method is transparent to the (local) caller - the data and methods of the called object may just as easily have been on the same machine as on another connected to the network. It also explains why the call semantics is so important: the fundamental difference between idempotent and at-most-once call semantics is whether or not the server is maintaining state. The communicating language itself is just a conventional device for securing a natural response.

However, according to Searle's institutional theory of communication [Searle, 1969], speaking a language is to engage in a rule-governed form of behaviour. The semantic structure of the language is given by a conventional characterisation of sets of *constitutive* rules, and 'speech acts' are acts typically performed by uttering an expression in accordance with those rules. Thus the basic unit of communication is the *illocutionary* act, which comprises the type of the act (its illocutionary force: stating, ordering, promising, etc.) and its propositional content. The meaning of each unit is given by the constitutive rules: these are rules that both define the forms of behaviour and determine what that behaviour *counts as.* Searle contended that it was not possible to reduce the analysis of illocutionary acts to perlocutionary effects. Therefore there *is* a significant difference between agent communication and object invocation: the aim of this paper is to make that difference clear and suggest how to leverage it for computational advantage.

In [Searle, 1969], the characterisation of the constitutive rules (in particular the preparatory and sincerity conditions) was expressed in terms of intentional states. While justifiable for natural language communication between humans, we believe that this is inappropriate for open systems of electronic agents and has misdirected standardisation efforts, e.g. FIPA (the Foundation for Intelligent Physical Agents). However, we believe that the conventional and institutional nature of communication can still be used to specify constitutive rules of interaction for such artificial systems.

The argument advanced in this paper to support that belief is as follows. Section 2 gives some preliminaries, introducing our methodology and notation. Section 3 gives some illustrative examples to motivate the specification, in Section 4, of some basic 'standard' communicative acts, and in Section 5 of a richer form of communication for conversations. Section G concludes with a summary, a brief review of further and related work, and draws some conclusions. In particular, the emphasis on the institutional perspective of Speech Act theory, rather than on agents' beliefs, desires, or intentions, indicates that a new paradigm in designing agent communication languages offers significant computational advantages.

## 2 Preliminaries

In this section, we briefly describe our methodological motivation for applying the ideas of Speech Act Theory to agent communication. We then describe the notation used in the rest of the paper, in particular, the abstract-representation of agents and normative relations (in particular institutional power, permission and obligation).

## 2.1 Methodology

We operate on the assumption that there are aspects of human intelligence and society than can inform algorithms, architectures and data structures for the design and implementation of 'better' solutions to software engineering problems. The process begins with the analysis of human intelligence and the expression of a theory (the process of theory formation in philosophy, psychology, linguistics, etc. We don't presume to undertake this step ourselves). The next step is a process we might call reification: the specification of that theory in a formal language. There are two problems with this process: firstly, information loss, as parts of the theory which are intractable are eliminated; and secondly, complexity gain, as there may be theoretical assumptions which are easy to state but much harder to formalise. As a consequence, the reified theory, or specification, is *not* a direct representation or literal specification of the original: instead it has been inspired by that t heory.

The process is marked in this paper. We have been analysing both Speech Act Theory [Scarle, 1969], Institutional Power [Jones and Scrgot, 1996], and Conventional Signalling systems [Jones, 2003]. The resulting specification of - or rather, proposal for a way of specifing - *an* agent communication language owes something to both accounts of communication (in particular, we will refer to signals and agent's signalling actions, rather than speech acts) but also looks to the extant computer implementations described in [Artikis *et* a/., 2002]. (Such implementation of the specification is step three of our methodology, while step four is to embed that implementation back in the human context in which it originated.)

## 2.2 Notation: Agent Description

We assume a domain with distinguished agents a, 6, etc., facts (propositions) $\phi$, and actions (procedures) $\pi$. An agent is a triple comprising *(K,* G, P), respectively the agent's knowledge base, goal base, and planned actions. For an individual agent a, we write:

| | |
|---|---|
| $\Delta_a$ | represents the program state of agent a, which encapsulates the agent's deliberative state; |
| $\Delta_a \vdash \phi$ | a's current program state (knowledge base) proves $\phi$ ($\phi \in K$); |
| $\Delta_a \rightsquigarrow \phi$ | a's current program state plans that in some future program state $\Delta_a \vdash \phi$ (currently $\phi \in G$); |
| $\Delta_a \mapsto \pi$ | a's program state plans action *IT* (not necessarily that *a* plans to do 7r itself) ($\pi \in P$); |
| $\tau(\Delta_a, \pi) = \Delta'_a$ | denotes a transition of a's program state as a consequence of action $\pi$ being performed. |

Intuitively, $\Delta_a \vdash \phi$ can be thought of, if it is preferred, as agent *a* believes $\phi$ is true). We write $\Delta_a \vdash (\Delta_b \vdash \phi)$ to denote an actual 'belief of *a* itself concerning the program state of 6, which can be true (in a) independently of whether $\Delta_b \vdash \phi$ or $\Delta_b \vdash \neg\phi$. $\Delta_a \rightsquigarrow \phi$ is an abstraction covering a future state of a, which may be thought of as a goal to achieve $\phi$ (say), and $\Delta_a \mapsto \pi$ is an abstraction for a plan to execute action $\pi$, either to be done itself or by another agent.

Note that, in a transition $\tau(\Delta_a, \pi)$ to $\Delta'_a$, if *a* performed $\pi$, we require that $\Delta_a \mapsto \pi$ and would normally expect that $\Delta'_a \not\mapsto \pi$ (i.e. we are not concerned with 'accidents': we are concerned with planned, purposeful action). If $\Delta_a \vdash \neg\phi$ and after (doing) $\pi$, $\phi$ is the case, we assume there is some process of belief revision which ensures that after the transition $\tau(\Delta_a, \pi)$, $\Delta'_a \vdash \phi$ and not $\Delta'_a \vdash \neg\phi$, i.e. that $\Delta'_a$ is *internally* consistent.

## 2.3 Notation: 'Logical' Description

The characterisation of the constitutive rules for an agent communication language requires the representation of action, 'counts as' and institutional power [Jones and Sergot, 1996], and dcontic relations. The notation used is as follows.

For an agent's actions, both signalling acts and (physical) acts on the environment, a relativised (to agents) £ operator is used (note that $\pi$ here is the proposition that 7r happened):

| | |
|---|---|
| $\mathcal{E}_a\phi$ | a sees to it that $\phi$ is true |
| $\mathcal{E}_a\pi$ | *a* sees to it that $\pi$ is performed |

For reasons illustrated in the first example of the next section, we also introduce a second relativised action operator C, for the idea of 'capability to verify', with intuitive reading (we consider the formal reading later):

| | |
|---|---|
| $\mathcal{C}_a\phi$ | means *a* is able to demonstrate (give evidence) that $\phi$ (is true) |
| $\mathcal{C}_a\pi$ | means *a* is able to demonstrate physical capability to perform $\pi$ |

Central to the [Jones, 2003] analysis of conventional signalling systems and the [Jones and Sergot, 1996] analysis of institutional power is the idea of *counts as.* The idea here is that one agent seeing to it that $\phi$, say, can count as, in certain contexts, as another agent - or even the institution itself - seeing to it that $\psi$. This notion was formalised in [Jones and Sergot, 1996] with a relativised (to institutions) conditional connective $\Rightarrow$:

$$\mathcal{E}_a\phi \Rightarrow_I \mathcal{E}_I\psi$$

Such a formula is called an *institutional power,* whereby *a* seeing to it that $\phi$ counts as (just as if) *1* sees to it that $\psi$. This is used to formalise Searle's constitutive rules of the form *"X* counts as *Y* in context C", where Searle's context *C* will here be denoted by some institution /. We write:

$$_I\mathbf{Pow}_a p \stackrel{def}{=} \mathcal{E}_a\phi \Rightarrow_I \mathcal{E}_I\psi$$

to identify a 'named' power *p* denoting a specific counts as formula.

An agent a's permissions and obligations with respect to an institution / are represented by:

| | |
|---|---|
| $_I\mathbf{Per}_a\mathcal{E}_a\pi$ | $a$ is permitted to perform $\pi$ in (by) J |
| $_I\mathbf{Per}_a\mathcal{E}_a\phi$ | $a$ is permitted to make $\phi$ true in (by) / |
| $_I\mathbf{Obl}_a\mathcal{E}_a\pi$ | $a$ is obliged to (by) / to perform $\pi$ |
| $_I\mathbf{Obl}_a\mathcal{E}_a\phi$ | $a$ is obliged to (by) / that $\phi$ be true |

## 3 Two Illustrative Examples

### 3.1 Example 1: TCP Slowstart

The standard TCP specification requires implementation of an algorithm called "slow start", documented in IETF RFC2001 (cf. [Stevens, 1996]). The specification of the algorithm mandates that, on the sender's side, the first transmission consists of just one packet. If this is acknowledged successfully, then the sender can transmit two packets. The sender then continues to increase the number of packets transmitted until the network capacity is reached (packets are not acknowledged successfully), and then throttles back. This ensures that no TCP sender overwhelms a network.

Let $w$ be a Windows NT server, and let $\phi$ be the statement *"I (w) implement TCP slow start algorithm as specified in IETF RFC2001".* Let $\sigma$ be the signalling system (i.e. agent communication language) in which the signal *q* denotes $\phi$. Assume *a* is used for communication by institution IETF, represented by /, so that $\mathcal{E}_a\varsigma$ is the action of transmitting the signal denoting $\phi$.

Now, suppose we want to allow agents, in particular servers (TCP senders), to assert or inform other agents that they implement the IETF (institution /) standard specifications. We also want / to be able to inspect and verify that servers do indeed implement the standard specifications correctly.

So, we can write $\Delta_w \vdash \phi$, meaning that, subjectively, from *w's* perspective, $\phi$ is true *(w* believes $\phi$). Suppose then that $w$ makes the assertion, in *a* that $\phi$. As per Searle ("Counts as an undertaking that $[\phi]$ represents an actual state of affairs" ([Searle, 1969]: p.66)) we want this to count as a commitment, in the context of the institution /, that it implements the slow start algorithm correctly. We interpret commitment here as an obligation to be able to verify that $\phi$.

Now, to capture the idea of an assertion as a commitment, the institution / has the following as one of its constitutive rules, which also apply to *any* other server:

$$\mathcal{E}_w\varsigma \Rightarrow_I \mathcal{E}_I({}_I\mathbf{Obl}_w\mathcal{C}_w\phi)$$

Thus the assertion, in 7, that the server *w* implements the slow start algorithm, counts as an obligation, from the server to the institution, that this is a true statement.

However, and more importantly, we also want to be able to make judgements of the form cither:

$$\Delta_I \vdash \phi \quad \text{or} \quad \Delta_I \not\vdash \phi$$

to say that according to the institution, $\phi$ is objectively (at least according to the institution) true (or not).

Certainly this demands that $\Delta_w \vdash \phi$, as a required state, but also that *w* be capable of demonstrating that $\Delta_w \vdash \phi$. As it turns , because of a widespread bug in the slow start implementation in Windows NT servers, $\phi$ is not a true (objectively true) statement [Stevens, 1996]. Being unable to verify (if demanded to do so) that $\phi$, *w* would be in violation of its obligation if it made the assertion. (Note we have set aside the practical problem of a program verifying properties (correctness) of itself.)

### 3.2 Example 2: Football (Soccer)

This kind of example is common in the literature but worth examining again. Consider the institution *F* of football (soccer). We define a "goal" is scored when the ball has crossed a painted white line between two uprights and under a crossbar. Then, we can examine the situation at any given instant and determine whether or not a *goal* (according to *F)* has been scored. (In Searlc's terms, the fact that a ball has crossed a line is a brute fact, the fact that by crossing the line between some sticks a *goal* has been scored, is an institutional fact.)

In certain games, a referee is given the power to examine the situation and decide whether or not it fits the definition of a "goal". If the referee signals that this is the case, then it is a *goal* (in the sense that the score of one team is incremented, etc.). Although a fairly coarse over-simplification, we could say that the referee is empowered by *F* to sec to it that *goals* are scored. However, the referee is not permitted to signal "goal" (cannot exercise the institutional power) unless the world-to-word fit is such that the situation in the world matches the definition of the word. So if one of the players sees to it that the conditions for a *goal* are satisfied (implicitly, according to the rules of football), then this counts as, in F, as initiating the referee's permission to signal "goal".

Let us assume that the players *are* permitted to signal "goal" arbitrarily (of course this is not strictly the case), but because they are not empowered to create *goals,* if they do signal "goal", this does not count as a *goal* (in effect, all they can do is 'claim').

So let us have agent 7' who has the power of referee (i.e. agent 7* occupies the role of referee, which is granted certain powers by institution F). Let $\varsigma_g$ be the action of transmitting a signal in the signalling system (ACL) used by F, which denotes "goal" (e.g., whistling, saying "goal", etc.). Then, for the institution F, referee agent r, and player agents p, we have :

$$\begin{array}{l} {}_F\mathbf{Pow}_r\, referee \stackrel{def}{=} \{\mathcal{E}_r\varsigma_g \Rightarrow_F \mathcal{E}_{\mathcal{F}}\, goal\} \\ \mathcal{E}_p\text{"goal"} \Rightarrow_F \mathcal{E}_F({}_F\mathbf{Per}_r\mathcal{E}_r\varsigma_g) \\ \forall p.{}_F\mathbf{Per}_p\mathcal{E}_p\varsigma_g \end{array}$$

A situation arose in the 1966 World Cup Final between England and Germany, when the England players were claiming a goal had been scored. The England players were permitted to signal "goal", but not being empowered, these signals did not count as *goals.* The referee was so empowered, so when he signalled "goal", it did count as a *goal* - even though a "goal" had *not* been scored. The problem, according to the characterisation above, was that the referee was exercising a power that in fact he was not permitted to. At this point, the referee should have been sanctioned according to the rules of F: but this too is a subject for further work. However, as we see in Section 5, this is the kind of situation we expect to encounter in future multi-agents systems, where agents have powers to see to it that institutional facts are true, but have constraints on their permissions or occasions when to exercise that power.

## 4 Single Signalling Acts

The two illustrative examples of the previous section indicate typical examples of the kind of meaning we seek for agent's illocutionary acts. In the first example, it was that the assertion counted as an undertaking that the propositional content was true (and that the asserter could verify that it was true); in the second example, that the declaration counted as the insitution seeing to that the corresponding institutional fact was true.

Motivated by this, in this section, we specify a set of constitutive rules for single, 'one off' signals, i.e. signalling actions not anticipating a reply. The specific context for this argument is as follows. For "an agent communicates by message passing", i.e. a communicative action performed by an agent, we are assuming a computational model involving the point-to-point transmission (from agent a to agent *b)* of 'information structures' representing these messages. 'Saying' and 'hearing' arc the results of, respectively, writing to and reading from some channel (e.g. a TCP/IP socket connection), or the consequences of remote method invocation (RMI) using constitutive rules rather than a conventional device.

For the operational context, we mean an open system as indicated in [Artikis *et* a/., 2002], whereby: no single agent has universal knowledge of the entire system, agents may have conflicting goals, internal architectures are not known, local states are individually consistent but may be globally inconsistent, and there is no centralized authority or control (of knowledge and behaviour). On the issue of decentralization, we take the position of local autonomy over goals, decisions and state, but that all communication is institutional [Searle, 1999].

Assume for the application that the agents are engaged in information trading or web services. We assume that the application will include finding and requesting services, and finding and supplying information. Thus there will be some individual message exchanges (assertives, directives, etc.) and some structured exchanges (e.g. contract nets and auctions). For each individual signalling act, we give the required form (of propositional content arid signalling content), the conventional interpretation of the signal, and the institutional powers (if any) associated with the signal itself. Note that this is an external specification: so the conventional interpretation indicates that, in the context of this particular institution, if an agent wants to convey a certain meaning, then (by convention) it uses a particular signal; and if an agent witnesses a particular signal, then (by convention) a certain meaning is conveyed. We consider four of the five types of illocutionary point identified in [Searle, 1999]: assertive (assert and inform), directive (command and request), commissive and declarative.

| 1. Assertive (assert) | |
|---|---|
| *propositional content* | proposition $\phi$ |
| *signal* $\varsigma$ | assert($\phi$) |
| *conventional interpretation* | $\Delta_a \leadsto (\Delta_b \vdash \phi)$ |
| *institutional condition* | $\mathcal{E}_a\varsigma \Rightarrow_I \mathcal{E}_I({}_I\mathbf{Obl}_a\mathcal{C}_a\phi)$ |

| 2. Assertive (inform) | |
|---|---|
| *propositional content* | proposition $\phi$ |
| *signal* $\varsigma$ | inform($\phi$) |
| *conventional interrelation* | $\Delta_a \leadsto (\Delta_b \vdash \phi)$ |
| *institutioTial condition* | none |

| 3. Directive (request) | |
|---|---|
| *propositional content* | action $\pi$ |
| *signal* $\varsigma$ | request($\pi$) |
| *conventional interpretation* | $\Delta_a \leadsto (\Delta_b \mapsto \pi)$ |
| *institutional condition* | none |

| 4. Directive (command) | |
|---|---|
| *propositional content* | action $\pi$ |
| *signal* $\varsigma$ | command($\pi$) |
| *conventional interpretation* | $\Delta_a \leadsto (\Delta_b \mapsto \pi)$ |
| *institutional condition* | $\mathcal{E}_a\varsigma \Rightarrow_I \mathcal{E}_I({}_I\mathbf{Obl}_b\mathcal{E}_b\pi)$ |

| 5. Commissive (promise) | |
|---|---|
| *propositional content* | action $\pi$ |
| *signal* «, | promise($\pi$) |
| *conventional interpretation* | $\Delta_a \leadsto (\Delta_b \vdash (\Delta_a \mapsto \pi))$ |
| *institutional condition* | $\mathcal{E}_a\varsigma \Rightarrow_I \mathcal{E}_I({}_I\mathbf{Obl}_a\mathcal{E}_a\pi)$ |

| 6. Declarative (declare) | |
|---|---|
| *propositional content* | proposition $\phi$ |
| *signal* $\varsigma$ | declare($\phi$) |
| *conventional interpretation* | $\Delta_a \leadsto (\Delta_b \vdash \phi)$ |
| *institutional condition* | $\mathcal{E}_a\varsigma \Rightarrow_I \mathcal{E}_I\phi$ |

Note that there might be some minor objections to the characterisation of the constitutive rules for the given illocutionary act types. However, recall that this is an *artificial* system of communication, that is being *designed.* It is not intended to be a formal characterisation of a preexisting system of communication, i.e. natural language, although it draws its inspiration from that source. If in turn the inspiration is not respecting the intuition, then it is cas}' enough to substitute *signaLacLI* for assert, *siyndLact-2* for inform, and so on, as the software will process the messages in just the same way and according to the same constitutive rules.

Searle's [Searle, 1969] formulation of rules in speech act theory included what he called preparatory conditions and sincerity conditions. Preparatory conditions were extant circumstances required for a speech act to be a valid act of the associated illocutionary type; the sincerity condition was a 'psychological' expression associated with the speech act.

Our agents do not have psychological states, but given the notation of Section 2 it is not uninteresting to consider putative sincerity conditions expressed in the signals defined above. For example, if we added to inform the preparatory condition -i($A_a$ h (At, h </>)), and the sincerity condition $A_a$ h <t>, then these are effectively FIPA inform Feasibility Preconditions FP1 and FP2 [FIPA, 1997], while the conventional interpretation is essentially the intended RE (Rational Effect). Thus it is straightforward, within the framework of constitutive rules, to get the current FIPA ACL semantics to 'drop out' as a special case.

## 5 Conversations: An Auction Protocol

In this section, we apply the idea of constitutive rules to a conversation, based on a variant of the English Auction protocol [Venkatraman and Singh, 1999]. One observation from the analysis here is that specification of constitutive rules for an ACL must be sensitive to time and predication, cf. [Artikis *et* a/., 2002; Sergot, 2003].

An informal description of the auction is: the auctioneer opens the auction for goods *g* at price *p* bidders are empowered to bid for *g* at price *p;* after receiving a bid, the auctioneer is empowered to announce a new price $p', p' > p$; and if there is no bid at the new price, the auctioneer has the power to accept a bid at price *p.*

There are also other rules that govern the conduct of an auction, namely that,: offering specific goods *g* for auction implies that the auctioneer either owns or is licensed to sell *g* making a bid of *p* for *p* implies that the bidder owns *p;* bidding *p* more than once and bidding a price lower than the announced price *p* arc meaningless actions; and the auctioneer is not empowered to accept two bids for the same *g.*

One way to formalise this protocol is to specify: the institutional powers of the auctioneer and the bidders; the assertions implicit in making announcements and bids; and the changing permissions of each participant as the protocol progresses which determine whether or not they are allowed to exercise their powers. There are three illocutionary acts: announce, bid and accept, with propositional content goods *g* (lot number) and price *p.* We assume, that all communications are broadcast, with the intended recipient *b* of accept messages explicitly identified in some way.

In general, the assertive force implicit in the illocutionary acts is given by:

$$\mathcal{E}_a \text{announce}(G,P) \Rightarrow_I \mathcal{E}_I({}_I\mathbf{Obl}_a \mathcal{C}_a own(G))$$

$$\mathcal{E}_b \text{bid}(G,P) \Rightarrow_I \mathcal{E}_I({}_I\mathbf{Obl}_b \mathcal{C}_b own(P))$$

The institutional conditions element of the constitutive rules are given by the following institutional powers. Note that these are unlike some of the institutional powers previously discussed in this paper, for two reasons. Firstly, these are 'power schema[1], in the sense that the actions of other agents will instantiate specific instances of each schema; and secondly, because they are transient rather than permanent, and each specific instance is either initiated or terminated under certain circumstances.

The power of a bidder is given by:

$${}_I\mathbf{Pow}_b bid(G,P) \stackrel{def}{=} \mathcal{E}_b \text{bid}(G,P) \Rightarrow_I \mathcal{E}_I bid(b,G,P)$$

while the auctioneer has two powers: firstly the power to make announcements, by announcing an auction, and secondly the power to make contracts, by accepting a bid:

$${}_I\mathbf{Pow}_a announce(G,P) \stackrel{def}{=}$$
$$\mathcal{E}_a \text{announce}(G,P) \Rightarrow_I \mathcal{E}_I announcement(a,G,P)$$

$${}_I\mathbf{Pow}_a accept(B,G,P) \stackrel{def}{=}$$
$$\mathcal{E}_a \text{accept}(B,G,P) \Rightarrow_I \mathcal{E}_I contract(a,B,G,P)$$

Now the way it works is that if an agent has the power and performs the action, then it will *count as* seeing to it that the corresponding instituional fact is true. The institutional fact in turn initiates or terminate other powers (so each agent in effect has the power to empower other agents). If, initially, /Pow$_a$*announce*(<7,p), then:

| signal | powers (initiated/terminated) | |
|---|---|---|
| *a* signals announce$(g,p)$ results in ... | | |
| ... initiating | *jT*ow$_h$bid(g, p)* | for all *b* |
| ... terminating | /Pow$_a$*announce*$(g,p)$ | |
| *b* signals bid(#,p) results in ... | | |
| ...initiating | /Pow$_u$ *announce*$(g,p')$ | $p' > p$ |
| ... terminating | *iPow$_b$>bid(g,p)* | for all b' |
| ... terminating | /Pow$_a$ *accept*$(b'',g,p)$ | some *b"* |
| ...initiating | ;Pow$_a$ *accept*$(b,g,p)$ | |
| *a* signals accept(/;, *g,p)* results in ... | | |
| ... terminating | /Pow$_a$ *accept* (*b*, *g*, *p)* | |
| ... terminating | /Powj/ *bid(g, p')* | |

Note [Artikis *et* a/., 2002] formalised a specification of the contract-net protocol in terms of which powers, etc., were initated and which were terminated by specific actions (which allowed for an event calculus specification and direct implementation of the protocol). [Sergot, 2003] also contains an analysis of an auction protocol.

What's missing from the formalisation of conversational illocutionary acts, as opposed to the signalling acts of the previous section, is the conventional interpretation (that these acts count as attempts to affect computational states in some way). It is not entirely clear what these should be (e.g. for announce, to inform *b* that *g* is for sale, to induce *b* to make a bid, or to indicate its intention to sell *g* once a buyer is found); or indeed that such rules are adding anything useful in this context. We leave this for further investigation.

In addition, we could specify that agents are only permitted to exercise their power, as an auctioner or a bidder, if they either own the goods *g* or the bidding price *p* respectively. In this way, we could circumvent the implicit assertive force in the announce and bid signals, by requiring that agents seek permission to sell goods or offer bids, and are subject to sanction if they create contracts without permission, which they cannot honour.

## 6 Summary and Conclusions

Searle [Searle, 1969] defined the semantics of illocutionary acts in terms of the conventional realization of series of sets of constitutive rules, and that a meaningful speech act was performed by uttering an expression in accordance with those rules. The constitutive rule sets proposed ([Searle, 1969]:p66-67) contained several conditions, including the propositional content, preparatory, sincerity and essential conditions. However, a number of these rules were concerned with 'psychological[1] states, and the attempted formalization of such rules in multimodal BDI logics are far from computationally tractable.

In this paper, we have proposed to use Searle's theory of constitutive rules to define a semantics of illocutionary acts (signals) for an agent communication language.

The set of conditions, for a wide series of act types, includes the propositional content and form of a signalling act, the conventional interpretation (of intended perlocutionary effect), the institutional conditions (expressed as a counts as formula) requiring institutional powers, and (in conversations) what deontic states are initiated or terminated by the performance of the signalling act.

There is, however, more work that is required. We have characterised the formalism used in Section 2 as a 'notation'. Ideally, the agent abstraction would define an operational model, which would be directly executable, and the 'logical' representation would be complete. For an executable operational model, we need an operational semantics, and for this purpose we are investigating labelled transition systems. A compiler for such rules, with an explicit link to (an implementation of) the agent abstraction, could then be considered.

For the logical model, we require (at least) a semantics of the *C* operator. We recognize this is problematic, but for the kind of applications in which we axe interested (where assertions of the forms "I have this right", "I represent this agency", "I can perform this service" are anticipated) we feel something of this kind is required. We have interpreted Searle's meaning of an assertion ("commitment to the truth of $\phi$") as an obligation to be able to verify $\phi$, and this is not a matter of merely seeing to it that $\phi$ ($\mathcal{E}_a\phi$). However, the work-around suggested in the previous section could eliminate the *C* operator.

Moreover, there axe many open questions, some of which have already been indicated. The nature of the relationship between electronic institutions and deontic states of agents needs to be further elaborated, as this is central to definition of the communication model. Many problems in Section 5 stemmed from trying to analyse dynamic phenomena in an essentially static formalism, so an explicit treatment and representation of time, linking to [Artikis *et a*l., 2002; Sergot, 2003], is required. The representation of contracts, from where this work originated, also needs to be refined.

Defining ACL semantics based on Speech Act theory was first proposed for both KQML [Labrou and Finin, 1998] and FIPA [FIPA, 1997], and given in terms of an axiomatic characterisation of intentional BDI agents. We would argue that the emphasis on the intentional states rather than the institutional nature of the communication has omitted the most important component of the theory. [Greaves *et al,* 2001] proposed an alternative model, using the idea of conversation policies, which define fine-grained constraints on ACL usage. Another alternative involves proposals primarily based on the idea of commitments, e.g. [Venkatraman and Singh, 1999; Colombetti, 2001], while Jones [Jones, 2003] proposes a conventional account of natural language speech acts based on the intentional exploitation of signalling systems. Motivated in part by the latter work, we have returned to Speech Act theory and attempted to characterise constitutive rules for a system of communication between 'autonomous' agents governed by an electronic institution. Our intuition is that the interaction between this work and that of conversation policies offers some intriguing potential for further development.

In conclusion, we believe that we have a basis for designing ACLs, and giving a formal characterization of the semantics of communicative acts which is powerful, expressive and flexible. Most importantly, it demonstrates the fundamental conceptual and functional difference between language-based communication and method invocation, and justifies why agent-based systems offer a unique advantage over object-oriented ones.


## Acknowledgements

We are grateful to the partners on the ALFEBIITE project (EU IST-1999-10298), especially Marek Sergot, Andrew Jones, Alexander Artikis and Lloyd Kamara; the members of FIPA TC Semantics; and several other commentators for their valuable input and insight.


## References


[Artikis *et al.,* 2002] A. Artikis, J. Pitt, and M. Sergot. Animated specifications of computational societies. In C. Castelfranchi and L. Johnson, eds., *Proceedings AAMAS'02,* ppl053-1062. ACM Press, 2002.

[Colombetti, 2001] M. Colombetti. Semantic, normative and practical aspects of agent communication. In F. Dignum and M. Greaves, eds., *Issues in Agent Communication,* LNAI1916, ppl7-30. Springer, 2001.

[FIPA, 1997] FIPA. FIPA'97 specification part 2: Agent communication language. FIPA (Foundation for Intelligent Physical Agents), http://www.fipa.org, 1997.

[Greaves *et a*l., 2001] M. Greaves, H. Holrnback, and J. Bradshaw. What is a conversation policy. In F. Dignum and M. Greaves, eds., *Issues in Agent Communication,* LNAI1916, ppl18-131. Springer, 2001.

[Jones and Sergot, 1996] A. Jones and M. Sergot. A formal characterisation of institutionalized power. *Journal of the IGPL,* 4(3):429-455, 1996.

[Jones, 2003] A. Jones. A logical framework. In J. Pitt (ed.): The Open Agent Society. Wiley, 2003 (to appear) [m/s at http://alfebiite.ee.imperial.ac.uk/].

[Labrou and Finin, 1998] Y. Labrou and T. Finin. Semantics and conversations for an agent communication language. In M. Huhns and M. Singh, eds., *Readings in Agents,* pp235-242. Morgan Kaufmann, 1998.

[Searle, 1969] J. Scarle. *Speech Acts: An Essay in the Philosophy of Language.* CUP, 1969.

[Searle, 1999] J. Searle. *Mind, Language and Society: Philosophy in the Real World.* Basic Books, 1999.

[Sergot, 2003] M. Sergot. The language $(C+)++$. In J. Pitt (ed.): The Open Agent Society. Wiley, 2003 (to appear) [m/s at http://alfebiite.ee.imperial.ac.uk/].

[Stevens, 1996] W. Richard Stevens. *TCP/IP Illustrated, Volume 3.* Addison-Wesley, 1996.

[Venkatraman and Singh, 1999] M. Venkatraman and M. Singh. Verifying compliance with commitment protocols: Enabling open web-based multi-agent systems. *Autonomous Agents and Multi-Agent Systems,* 2(3):217-236, 1999.