NEGOTIATING THE SEMANTICS OF AGENT COMMUNICATION LANGUAGES

CHRIS REED

Department of Applied Computing, University of Dundee

TIMOTHY J. NORMAN

Department of Computing Science, University of Aberdeen

NICHOLAS R. JENNINGS

Department of Electronics and Computer Science, University of Southampton

This article presents a formal framework and outlines a method that autonomous agents can use to negotiate the semantics of their communication language at run-time. Such an ability is needed in open multi-agent systems so that agents can ensure they understand the implications of the utterances that are being made and so that they can tailor the meaning of the primitives to best fit their prevailing circumstances. To this end, the *semantic space* framework provides a systematic means of classifying the primitives along multiple relevant dimensions. This classification can then be used by the agents to structure their negotiation (or *semantic fixing*) process so that they converge to the mutually agreeable semantics that are necessary for coherent social interactions.

Key words: agent communication languages, semantic space, semantic fixing.

1. INTRODUCTION

Inter-agent communication is the most common means by which autonomous agents cooperate, coordinate, and negotiate. To this end, such agents are typically equipped with a particular agent communication language (ACL). This language defines the set of communication primitives (both syntax and semantics) that the agents can either generate or receive during their social exchanges. In systems in which there is a single authority that has complete control over the design and implementation of all the agents, achieving meaningful interoperation is not overly problematic. However, when no such authority exists, which is the case for many of the applications for which agents are best suited (Jennings 2000), effective communication is more difficult to achieve. To date, there have been two broad approaches to addressing this problem. The first is to define a standard for agent communication and then dictate that all the agents must adhere to it (examples of such standards include FIPA-ACL (FIPA 2000) and KQML (Labrou and Finin 1997)). The second is to assume there is no existing communication language and to allow the agents to generate one for themselves (e.g., Steels and Kaplan 1999). Both of these approaches have their respective drawbacks, however. In the former case, the need to be encompassing means that many of the primitives' definitions are not well suited to expressing the agent's communication needs in the particular situation in which it finds itself. Also an agent has no means of checking at run-time the precise semantics being used by the agent with which it is communicating (it simply has to assume they are conforming to the standard). In the latter case, convergence to a set of usable primitives can be extremely slow and it is a process that must be repeated for each new agent that is encountered. To overcome both sets of shortcomings, we advocate a hybrid approach in which the agents dynamically negotiate the precise semantics of their communication primitives according to their prevailing needs using a systematic multidimensional framework. We term the framework the *semantic space* of the communication primitives (since it represents the range of possible semantics that the primitives may take) and the process of setting the semantics of the primitives as *semantic fixing* (since it involves negotiation about the meaning of the primitives and, eventually, an agreement about their meaning).

The motivating hypothesis behind this work is that there are no definitive (absolute) definitions of agent communication primitives that are universally applicable. Thus, it is desirable for agents to have the ability to define and agree upon the semantics at run-time according to their prevailing interaction needs. Since the agents are autonomous, and typically represent different stakeholders, the process by which they come to an agreement will necessarily be some form of negotiation since the agreement must be one that is mutually acceptable (Jennings et al. 2001; Walton and Krabbe 1995). Without a systematic underpinning framework to guide this negotiation process, however, the cost of reaching such agreements would be prohibitively expensive.

Against this background, the contribution of this work can be viewed as twofold. First, we introduce the notion of semantic space and provide a systematic framework for defining primitives within it. A semantic space provides a multidimensional structure within which all potential communication primitives can be located. Thus, rather than viewing communication primitives as being arranged in some rigid taxonomy whose semantics are fixed at design time, we explicitly acknowledge that (seemingly subtle) variations in the meaning of a primitive can make a significant difference to its efficacy in a given encounter. For this reason we believe these meanings should be decided at run-time so that they can be tailored to reflect the agent's current context. In more detail, all ACL primitives can be thought of as identifying a single point or a region in semantic space, regardless of the tradition, scheme or library of which they form a part. The distribution of primitives in semantic space is not even but clumpy, with primitives occurring in clusters. Thus, for example, the commissive acts defined within the FIPA, KQML, and KAoS specifications would lie close together in a region of semantic space (though would not be at exactly the same point—see Section 3 for more details) and this cluster would be relatively distant from a cluster representing a selection of assertive acts such as FIPA's inform and KQML's tell—see Section 2 for a more detailed discussion of these. The concept of semantic space extends the parsimonious approach of Hindriks et al. (2000), in which a small number of basic primitives can be composed to produce (epiphenomenally) a range of speech acts, such as are provided by KQML or FIPA. The approach advocated by Hindriks et al. will allow access, by composition, to a large portion of semantic space, but is inevitably unable to characterize the entire space. The approach would be unable, for example, to capture both KQML and FIPA variants of an inform.

The second contribution of this work is that we outline a mechanism by which agents can use the semantic space framework to negotiate the meaning of their communication primitives according to the type of dialog in which they are engaged. The method that we advocate for this semantic fixing is a simple voting mechanism. In particular, we motivate this choice from among the range of potential negotiation mechanisms, we describe the assumptions that are necessary for this approach to operate and we describe a specific operational protocol that is appropriate for this task.

The remainder of this article is structured as follows: Section 2 uses probably the most familiar type of communicative act—the assertive—to illustrate the fundamental differences that exist between the specifications provided by existing agent communication languages. The examples used are based on three prominent, yet fundamentally different, ACLs. Our purpose here is to highlight the fact that each definition, when viewed in isolation, appears reasonable; however it is impossible to say that any one of them is "correct" and that the others are "wrong." Rather, each definition is suitable for agents interacting in different contexts. With this motivation, Section 3 presents an analysis of some of the key dimensions of semantic space with a focus on the types of move made in the establishment of contracts between agents. The reason for this focus is that the negotiation of contracts is of particular importance to both the current efforts in specifying agent communication languages and in the application of agent technology in general. In Section 4, the issue of semantic fixing is

discussed and the simple voting mechanism is detailed; an example of which is examined in Section 5. Section 6 discusses the many issues raised by our approach and relates our work to that of others in the field. Finally, Section 7 concludes the article and outlines the avenues of further work.

2. EXAMPLE: ASSERTIVES IN SEMANTIC SPACE

To provide an example of disparate primitives in a uniform semantic space, three closely related primitives are here characterized in terms of the FIPA SL—the Semantic Language used to specify the communicative acts of FIPA-ACL. The use of the FIPA SL is not crucial to the exposition, which could equally employ some other, possibly novel, language. The choice is expedient simply because the formalism is already available and familiar, and intuitive explanations of the modalities and abbreviations used in FIPA SL are given where necessary. The example adopted is that of the assertives: inform (FIPA 2000), tell (Labrou and Finin 1997) and INF (Smith et al. 1998). For the sake of clarity, throughout the remainder of the article the definitions of primitives in these three ACLs are specified and discussed using a common set of terms to denote states (p, q, etc.), actions (a, a', etc.) and agents (i, j, etc.).

The full FIPA definition of inform is given in Definition 1. Note that this includes components which are omitted from the FIPA summary, including the intentional effect—property 4 (FIPA 2000, p. 49). In the case of inform, the intentional effect is the second "rational effect" (RE) given in Definition 1: the recipient, j, believes that the sender, i, intends that j believe p. This definition (and Definitions 2b and 3b) also specify a set of "feasibility preconditions" (FPs) that state the circumstances in which, by performing the act, the rational effects may be established.[1]

$$< i, \text{inform}(j, p) > \tag{1}$$

$$\text{FP}: \quad \mathsf{B}_i p \wedge \neg \mathsf{B}_i(\text{Bif}_j p \vee \text{Uif}_j p) \wedge \neg \mathsf{B}_i \mathsf{B}_j p$$

$$\left(\bigwedge_{n>1} \neg \mathsf{AB}_{n,i,j} \mathsf{B}_i p \right) \left(\bigwedge_{n>2} \neg \mathsf{AB}_{n,i,j} \mathsf{B}_j p \right)$$

$$\text{RE}: \quad \mathsf{B}_j p \wedge \mathsf{B}_j \mathsf{I}_i \mathsf{B}_j p$$

$$\text{tell}(i, j, p) \tag{2a}$$

$$\text{Pre}(i): \quad \text{bel}(i, p) \wedge \text{know}(i, \text{want}(j, \text{know}(j, q)))$$

$$\text{Pre}(j): \quad \text{intend}(j, \text{know}(j, q))$$

$$\text{Post}(i): \quad \text{know}(i, \text{know}(j, \text{bel}(i, p)))$$

$$\text{Post}(j): \quad \text{know}(j, \text{bel}(i, p))$$

$$< i, \text{tell}(j, p) > \tag{2b}$$

$$\text{FP}: \quad \mathsf{B}_i p \wedge \mathsf{B}_i \mathsf{C}_j \text{Bif}_j p \wedge \mathsf{I}_j \text{Bif}_j p$$

$$\text{RE}: \quad \mathsf{B}_i \mathsf{B}_j \mathsf{B}_i p \wedge \mathsf{B}_j \mathsf{B}_i p$$

---

[1] Modality B indicates belief and is a KD45 modal operator (Chellas 1980). $\text{Bif}_i p$ abbreviates $\mathsf{B}_i p \vee \mathsf{B}_i \neg p$. $\text{Uif}_i p$ abbreviates $\mathsf{U}_i p \vee \mathsf{U}_i \neg p$, where $\mathsf{U}_i p$ represents agent i being uncertain about the truth condition of p but that p is more likely than $\neg p$. $\mathsf{AB}_{n,i,j}$ represents the belief of agent i about the belief of agent j, etc., to a level of nesting n. Finally, I denotes intention, and $\mathsf{I}_i p$ represents i's commitment to satisfying p.

The closest equivalent—the primitive employed for assertive communication—in KQML is tell. The definition, from (Labrou and Finin 1997), is given in Definition 2a, and a translation into the FIPA SL in Definition 2b.[2] The preconditions of sender, i, and recipient, j, in Definition 2a map onto the feasibility preconditions in Definition 2b. Similarly, the postconditions of sender, i, and recipient, j, in Definition 2a map onto the rational effects in Definition 2b—communicative act postconditions (in KQML) and the rational effect of a communicative act (in FIPA ACL) are equivalent notions. It is, however, important to point out that, to paraphrase the FIPA specification document (FIPA 2000, p. 31), "[an] agent will select acts based on the relevance of the act's expected outcome or *rational effect* to its goals[, but] it cannot assume that the rational effect will necessarily result from sending the message." This is, of course, true for any action that an agent may perform in a nondeterministic environment, even if it is, arguably, a more pervasive issue in agent communication.

There are a number of comments and caveats regarding this translation of the KQML tell. It should be noted that the languages offered in the two works are significantly different; there is not always a simple one-to-one translation between them. In particular, justification is required for this translation of bel and know. The bel operator is used when reference is being made to a sentence in the knowledge representation language of a particular agent, whereas know is used when reference is made to an abstract cognitive state. This is useful if it is important to make explicit the fact that agents may not share the same knowledge representation, but in this article they are treated as equivalent. The know operator is also used to refer to some q in Pre(i) and Pre(j). This q can be either bel(j, p) or ¬bel(j, p) and, because we make no distinction between know and bel, know(j, q) can simply be translated as $\mathsf{Bif}_j\, p$.

The aim, however, is not to provide a rigorous intertranslation, but rather, to capture the spirit of the primitives and to investigate their differences at that abstract level. Given this aim, it is also instructive to examine the work of Cohen and Levesque (C&L) (1990) and Smith et al. (1998) on agent communication in the same way, despite the fact that their definitions are in terms of denotational, rather than operational, semantics. Definition 3a summarizes the definition of inform used in the KAoS system (Smith et al. 1998),[3] a FIPA SL translation is offered in Definition 3b.

$$
\mathsf{INF}(i, j, e, p) \;\overset{\text{def}}{=}\; \begin{aligned}[t] &\mathsf{BEL}(i, \neg\mathsf{BMB}(j, i, p))\,\wedge \\ &\mathsf{GOAL}(i, \mathsf{HAPPENS}(i, e; \Diamond\mathsf{BMB}(j, i, p)?))\,\wedge \\ &\mathsf{INT}(i, e; \mathsf{BMB}(j, i, \mathsf{BEL}(i, p))) \end{aligned} \tag{3a}
$$

$$
< i, \mathrm{inf}(j, p) > \tag{3b}
$$

$$
\mathsf{FP}: \quad \mathsf{B}_i\left(\bigwedge_{n>1}\neg\mathsf{AB}_{n,j,i}\,p\right) \wedge \mathsf{C}_i\,\mathsf{Possible}\left(\bigwedge_{n>1}\mathsf{AB}_{n,j,i}\,p\right)
$$

$$
\wedge\, \mathsf{I}_i\left(\bigwedge_{n>1}\mathsf{AB}_{n,j,i}\,\mathsf{B}_i\,p\right)
$$

$$
\mathsf{RE}: \quad \bigwedge_{n>1}\mathsf{AB}_{n,j,i}\,p
$$

---

[2] $\mathsf{C}_i\, p$ represents i's desire that p holds.

[3] A different notation is used for the sake of consistency within this article; (Smith et al. 1998) use a Lisp-like notation, e.g., (INF $i\ j\ e\ p$).

Again, there are several comments on the translation. Some aspects are reasonably straightforward: the use of an infinite conjunction for mutual belief is in direct accordance with Cohen and Levesque's (1990, p. 232) definition, and the HAPPENS operator can be captured reasonably well using SL's Feasible operator, upon which Possible is defined as $\text{Possible}(p) \equiv (\exists a)\text{Feasible}(a, p)$. The most important liberty taken in this translation is to conflate Cohen and Levesque's notion of the 'minimum acceptable result' into the rational effect. Thus, in the example above, the minimum acceptable result is that the speaker and hearer mutually believe that the speaker believes the content of the utterance—i.e., $\text{BMB}(j, i, \text{BEL}(i, p))$. Because FIPA does not explicitly formalize any equivalent of an 'attempt' within its semantics, and cannot therefore distinguish between an ultimate goal and a minimum acceptable result, the rational effect is taken to be the satisfaction of the ultimate goal of the utterance. The intention to achieve the minimum acceptable result and the ultimate goal are included in the feasibility preconditions. Although $\text{BMB}(j, i, \text{BEL}(i, p))$ is not *a priori* derivable from $\text{BMB}(j, i, p)$, the difference lies in $\text{ABEL}(n, j, i, \text{BEL}(i, p))$ for even values of n—i.e., $\text{B}_j\text{B}_i\text{B}_i p$, $\text{B}_j\text{B}_i\text{B}_j\text{B}_i\text{B}_i p$, and so on, in our SL translation. Although these do not follow directly from $\bigwedge_{n>1} \text{AB}_{n,j,i} p$, FIPA SL is a KD45 logic, so they can be concluded by introspection. Thus the rational effect need include only the $\text{BMB}(j, i, p)$ clause. In effect, for the purposes of investigating the character of the primitives, we are assuming that communication is reliable (fallibility handling would be captured on yet further axes of semantic space). Finally, Cohen and Levesque's GOAL operator needs to be translated as the FIPA SL C_i modality, because the selection of the primitive (by FIPA property 1) rests upon an extant intention of achievement of the rational effect. Thus a communicative act a with rational effect p, is only selected if the agent holds $\text{I}_i p$; therefore, including $\text{I}_i p$ as a feasibility precondition would be vacuous.

Definitions 1, 2b, and 3b each purport to capture an assertive communicative act. Yet clearly, they differ from one another along many axes. The first, and perhaps most obvious, is that FIPA and KQML explicitly demand sincerity: $\text{B}_i p$ is a precondition in both definitions. For Cohen and Levesque, no such requirement is imposed. Instead, they explain that sincerity may be derivable from agents' mental states in particular circumstances, but is not ubiquitous. Pitt and Mamdani (2000) discuss a number of examples that mitigate against the blanket adoption of sincerity, but whichever point is selected on this particular binary dimension in semantic space, it is clearly crucial that agents involved have a common understanding of it. A second dimension of difference between these three examples is what might be termed 'level of success' in the outcome. FIPA is perhaps the most optimistic, since it includes $\text{B}_j p$—that the hearer does, in fact, believe the content of the inform. Next is the C&L act, which (if successful) leads to the speaker's belief that speaker and hearer mutually believe. Finally, the most pessimistic is KQML, which captures the hearer's belief that the speaker believes the content of the utterance (the hearer believes the speaker is sincere), and the speaker's belief that that is what the hearer believes (the speaker believes the hearer will trust it). This dimension has many more points—from full, bilateral mutual belief, down to nothing more than very weak nested constructs, such as FIPA's intentional effect (the hearer recognizes the speaker's intention to effect belief change). Crucially, each point along this scale seems reasonable in a given context—in tightly knit, reliable, trustworthy relationships, shifting to bilateral mutual belief with a single utterance might be highly cost effective. In contrast, a more loosely coupled, unreliable, possibly new relationship might demand greater caution in the conclusions to be drawn from an inform. Again, though, agents need to be aware of the precise nature of the inform that they are employing.

This example, contrasting the different conceptions of assertives, illustrates how individual extant primitives lie at particular, individuatable points within the space of speech-act

semantics. This shows that, although one assertive may seem to be, intuitively, no different from any other, a deeper investigation may reveal subtle distinctions. Moreover, it is simply not possible to say which of these is "correct" and which is "wrong." Given this, the following section presents an analysis of three of the main dimensions of the semantic space of communicative acts.

3. MAPPING SEMANTIC SPACE

The key argument of this article is that the mapping of semantic space provides a means to systematically analyze inter-agent communication. To cover the range of possible communication primitives, this space needs to be highly multidimensional—the above analysis of assertives illustrates this—although not all of the dimensions will be relevant to all types of primitives. Our purpose, then, in this section is to focus on three of the dimensions and to focus on primitives that are concerned with making contracts. Our aim is not to exhaustively detail all the dimensions, nor to show how other types of communication primitive can map into the semantic space structure we propose (both these activities are worthy longer-term goals, but they fall beyond the scope of this article).

It is worth noting at this point that we assume that there are no dependencies between the dimensions of semantic space (hence the use of the term dimension). Having said this, a number of theories of agency do rely on constraints between modalities such as belief and intention (Grosz and Kraus 1996; Panzarasa et al. 2001), but this is to do with the model of a particular agent rather than the model of communicative acts that agents use. It could be the case that the agent's internal model precludes the use of a particular communicative act specification that exists within semantic space, but this doesn't mean that semantic space should be restricted in this way.

The dimensions we focus on refer to a single agent and a single proposition. Conversely, if an agent is not permitted to bring about p and some authority (possibly some other agent) permits it to do so, its freedom to act toward that proposition is increased. Such moves are especially critical during the negotiation of contracts between agents (Norman et al. 1998).

These commissive and permissive moves serve to decrease and increase the freedom of an agent toward some proposition respectively. For these reasons, two of the dimensions considered employ Lindahl's jurisprudential analysis of normative positions (1977) since it is the normative position of an agent that characterizes its freedom toward a proposition.[4] The first dimension characterizes the preconditions in the semantics of a primitive, and the second is used to define its postconditions. The third dimension relates to belief and, as the discussion is restricted to changes in normative position, will appear only in the preconditions of a primitive.

3.1. Normative Positions

In this section we concentrate on the first two dimensions. These characterize how the normative position of a single agent changes as the result of a communicative act: the normative position specified in the preconditions is replaced by that specified in the postconditions of the act. Lindahl's model specifies seven distinct points on each axis. Using the modality P to refer to permission and $E_i p$ to refer to agent i bringing about the state of affairs p, a

---

[4]For a more detailed analysis of normative positions and their uses in the specification of computing systems, see Sergot (1999).
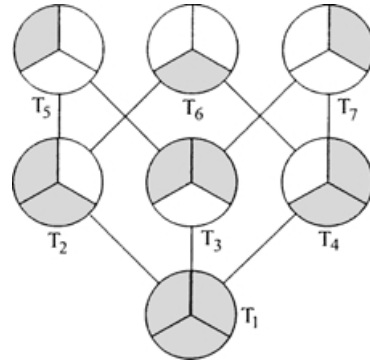
FIGURE 1. Hasse diagram representing Lindahl's seven one-agent types of normative position for a single agent, i and a single proposition, p. A shaded top-left segment indicates that the agent, i is permitted to bring about the state of affairs, p: $\mathsf{P}\,\mathsf{E}_i\,p$. A shaded top-right segment indicates that i is permitted to bring about $\neg p$: $\mathsf{P}\,\mathsf{E}_i\neg p$. A shaded bottom segment indicates that i is permitted to remain passive towards the state of affairs p: $\mathsf{P}\,(\neg\mathsf{E}_i\,p \vee \neg\mathsf{E}_i\neg p)$. Unshaded segments indicate that i is *not* so permitted: i is forbidden.

single normative position may be composed from a permutation of the following and their negation:

1. permission to bring about p: $\mathsf{P}\,\mathsf{E}_i\,p$;
2. permission to bring about $\neg p$: $\mathsf{P}\,\mathsf{E}_i\neg p$; or
3. permission to remain passive toward p: $\mathsf{P}\,(\neg\mathsf{E}_i\,p \vee \neg\mathsf{E}_i\neg p)$.

The logic of the E operator is one of successful action: it includes the axiom schema $\mathsf{E}_i\,p \rightarrow p$ and the rule of inference if $\vdash p \leftrightarrow q$ then $\vdash \mathsf{E}_i\,p \leftrightarrow \mathsf{E}_i\,q$. Lindahl (1977), in fact, uses the notation $\mathsf{Do}(i, p)$ for $\mathsf{E}_i\,p$—the latter notation adopted here follows that used by Jones and Sergot (1996). P can be read "it is permitted that" and it is the dual of O, or "it is obligatory that," i.e., $\mathsf{O}p \equiv \neg\mathsf{P}\neg p$. The logic of O contains the rule of inference if $\vdash p$ then $\vdash \mathsf{O}p$, and the axiom schemas $\mathsf{O}(p \rightarrow q) \rightarrow (\mathsf{O}p \rightarrow \mathsf{O}q)$ and $\mathsf{O}p \rightarrow \neg\mathsf{O}\neg p$. Again, Lindahl uses a different notation: Shall and May rather than O and P respectively.

One possible legal position, T_1, represents complete freedom (an agent is permitted to bring about p, to bring about $\neg p$, or to do nothing to p at all), and another position, T_5, represents a specific obligation (permitted to bring about p, but not permitted to bring about $\neg p$ or to remain passive). The seven distinct types can be arranged as a partial ordering, shown in Figure 1 (reproduced from Lindahl (1977)), from complete freedom (at the bottom), to maximal restriction (at the top).

The arcs in Figure 1 represent potential communicative actions. Moving from T_1 to T_2, for example, represents a communicative exchange which decreases the freedom of i (from complete freedom toward p, to a restriction whereby i may not bring about $\neg p$). There are thus nine legitimate *primitive* communicative actions that reduce an agent's freedom (traversing arcs upward), and nine more that increase its freedom (traversing downward). This represents a subset of all those that are possible: a designer of an agent system may wish agents to be able to move from any one of the seven positions to any other in a single communication exchange. Thus the initial and resultant Lindahl states can be seen as distinct axes in semantic space, offering, in this plane, 42 possible primitives that effect a change in the normative position of the agent. This space, and the eighteen primitive moves mentioned, are shown in Figure 2.
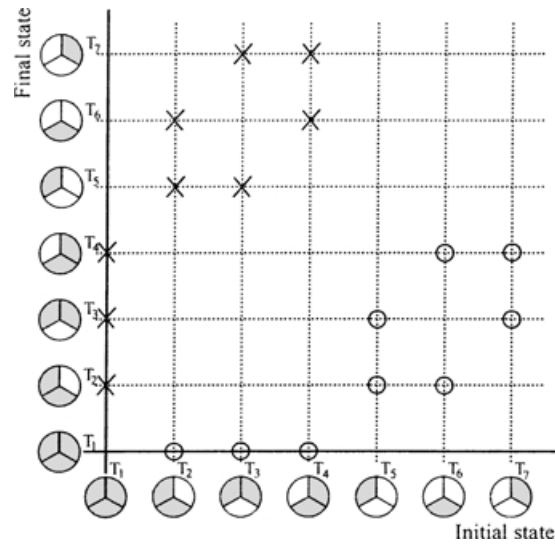
FIGURE 2. Primitive normative position movements (for commissives, X, and permissives, O) for a single agent, i, and a single proposition, p.

A point in this space can be identified with a particular communication action. Thus, for example, (1,4) represents agent i's committing to not bringing about p, from an initial position of complete freedom toward p (a movement from T_1 to T_4).[5] Furthermore, a single utterance may constitute (by underspecification) a move from any number of start states to any number of final states. Thus, the set of points, $\{(2,5), (3,5)\}$ represents a communicative action which commits agent i to bringing about p, specifying only the restriction that the agent is not so committed at the start state, but which still conforms to the primitive moves specified in Figure 1. Of course, as recognized by Lindahl, there is no requirement that a communicative action should constitute a primitive move, so a communicative action which simply introduces a commitment on i to bring about p would be represented by the set of points $\{(1,5), (2,5), (3,5), (4,5), (6,5), (7,5)\}$.[6]

It is important to recognize that a single communicative action may correspond to a region of semantic space, rather than just a point. In fact, this idea is implicit in the construction of semantic space itself, because it is unlikely that every primitive will be precisely specified on every axis. Axes concerning permission, for example, are very infrequently likely to be useful in requesting, so many definitions of requests may range over all possible points on those axes. But as demonstrated above, communicative acts may also range noncontiguously over subsets of points on particular axes. This accords well with intuition about communication action, because some actions are more general than others. In FIPA, for example, the query-if communicative act is defined in terms of the more general request and inform primitives.

3.2. Committing to Bring about p

To demonstrate how semantic space, including those dimensions introduced in the previous section, can be populated by distinct, useful communicative acts, the act of committing

---

[5]The origin is taken to be (1,1) indicating a reflexive move from normative position T_1 to T_1.

[6]The reflexive move (5,5) is not included in this set, but it could be justified if an agent may commit to bringing about p when it is already so committed. Thus the intensional state and extensional one-agent state remains unchanged, but there may, of course, be implications on the agent's relationships with others.

to bring about a state of affairs p is considered. The assumption made is that the rational effect of this act is that the speaker moves to the Lindahl state T_5. It is worth noting here that the speaker is, in some way, committed to the hearer to bring about p after this communicative act. The key issue for such an act, however, is that the speaker is committed to act toward the achievement of the goal p. The hearer can, therefore, assume that after this communicative act is performed, the speaker is so committed.[7] Other types of commissive act are, of course, also possible. For example, an agent can commit to remain passive toward the state of affairs p regardless of whether p holds or not—a normative state of affairs that may be useful in many applications. The reason for focusing on acts that result in the speaker being in state T_5 is that this is the most common commissive act considered by designers of ACLs.

Before specific examples from FIPA, KQML, and KAoS are discussed, there are a number of issues that are important in selecting the region of semantic space that defines the semantics of a particular commissive act. These include the third dimension of semantic space that is considered in this article: whether or not the agent concerned believes that the proposition holds (it is assumed that, unlike its normative state, the agent's belief in the proposition is unchanged). Each of these dimensions refer to a single agent and a single proposition.[8] The issues that are considered in this section that are relevant to commissive acts that result in the speaker entering normative position T_5 are:

1. Whether or not the speaker is committed to *p prior to* the performance of the commissive act.
2. Whether or not the speaker is *a priori* forbidden from bringing about the state of affairs p.
3. Whether or not the speaker believes that p holds.

FIPA's accept-proposal is "the action of accepting a previously submitted proposal to perform an action," and so must be viewed in the context of the semantics of a proposal. The key FP of the **propose** primitive is that *if* the speaker, i, believes that the hearer, j, intends that i perform the act *then i* will intend to perform the act. An acceptance of a proposal (the **accept-proposal** primitive) is then simply the communication by j of its intention that i performs the act. This is as close as FIPA gets to a commissive act: a sequence of two specialized assertives. In combination, however, this short dialog results in the original proposer having the intention to perform the act.[9] The closest we can get to a mapping from this short dialog to the region of semantic space discussed in this article is that in the initial state (prior to the proposal being issued) the proposer may be in any one of the seven normative positions and in the final state (after the proposal is accepted) the proposer is in the normative position T_5. The FIPA specification allows the agent making the commitment to be so committed prior to the acceptance of the proposal, it may be *a priori* forbidden from achieving the goal, and the goal may already be achieved. There is, therefore, no restriction, according to the criteria discussed above, on the circumstances in which a FIPA-compliant agent can commit to activity.

---

[7]See Wooldridge (2000) for a discussion of the verification of inter-agent communication.

[8]An obvious extension to this three dimensional semantic space is to consider the normative positions (before and after the act) and the epistemic state of the hearer—resulting in six dimensions of semantic space being considered. This is discussed in Section 6.

[9]This can be derived from the semantics of propose and accept-proposal and the property of SL that states $\models \psi \Leftrightarrow B_i \psi$ when ψ refers to a mental attitude of i (FIPA 2000, p. 47).

Using the notation that $T_n(i, p)$ refers to an agent, i, in a normative position, T_n, with respect to p, a commissive act that captures the spirit of FIPA's propose-accept-proposal may be defined as follows:

$< i$, FIPA-commit$(j, p) >$

$$FP: \ T_1(i, p) \vee T_2(i, p) \vee T_3(i, p) \vee T_4(i, p) \vee T_5(i, p) \vee T_6(i, p)$$
$$\vee T_7(i, p)$$
$$RE: \ T_5(i, p)$$

In common with FIPA, KQML does not include commissive acts within its specification. The canonical example, however, for establishing some kind of contractual agreement between agents is an achieve-tell sequence. The **achieve** represents a request from an agent i to an agent j to bring about some state of affairs, p, in the world. The **tell** offered in response by j can then either be acknowledgment that p has been brought about (or has failed to be brought about), or can indicate an intention to bring about p. It is the latter case that is of interest since the former barely represents a contract at all. To use Labrou and Finin's example, i issues the performative **achieve**$(i, j, \text{torque}(\text{motor1}, 5))$. To express j's intention to conform to this request, j may reply with **tell**$(j, i, \text{intend}(j, \text{torque}(\text{motor1}, 5)))$. It is, therefore, this **tell** which commits j to performing an action appropriate to bringing about the state of affairs **torque**$(\text{motor1}, 5)$, since **tell** includes sincerity as a precondition.[10] The preconditions of the **achieve** include that j does not already intend to bring about the state of affairs; and the **tell** requires j to intend to bring it about. This suggests that the area of semantic space being identified is the set of points (in the Lindahl plane) of the preconditions being any normative position except T_5, and postcondition being T_5—in Figure 2, the set of points $\{(1,5), (2,5), (3,5), (4,5), (6,5), (7,5)\}$. According to the KQML specification, therefore, an agent may commit to achieving a goal that it is forbidden from achieving, or that has already been achieved. It may not, however, commit to a goal it is already committed to. A commissive act that captures the spirit of KQML's achieve-tell may be defined as follows:

$< i$, KQML-commit$(j, p) >$

$$FP: \ T_1(i, p) \vee T_2(i, p) \vee T_3(i, p) \vee T_4(i, p) \vee T_6(i, p) \vee T_7(i, p)$$
$$RE: \ T_5(i, p)$$

There are two problems with the analyses of these FIPA and KQML exchanges. In the first place, both definitions rest upon a complex of primitives, rather than upon the semantics of a singular performative. In the case of FIPA, a **propose-accept-proposal** sequence, and, in the case of KQML, an **achieve**-tell sequence. The second problem is in what is lost in the analysis, and in particular the binding between the two agents (in both cases). The result of the **accept-proposal** (in FIPA) is that both agents intend that the proposer perform the action concerned. The result of the **tell** (in KQML) is that both agents know that one intends to bring about the specified state of affairs—although there is no restriction on the action of the agent that has not committed to act, cf. Norman et al. (1998). To capture this result, the model of semantic space needs to be refined to detail Lindahl's two-agent types, in which such reciprocity is elegantly handled. This refinement of the model is discussed in Sections 4 and 6.

The confirm act in the KAoS system (Smith et al. 1998) is an inform of a **PWAG**; i.e., an agent expressing either a **PGOAL** (persistent goal—Cohen and Levesque's (1990) notion of intention) to bring about some state of affairs p, or, if that cannot happen (p is already true,

---

[10]The other primary means for this type of commitment to be incurred is through an advertize, though the dynamics of such a conversation are different, and the performatives necessary occupy different positions in semantic space.

can never be true, or the relativizing condition, q, fails) a PGOAL to bring about mutual belief that p cannot happen. In the most straightforward case, then, a confirm act entails a PGOAL toward a state of affairs.

To define a formal mapping between the semantics used to define communicative acts in the KAoS language and a semantics for Lindahl deontic and action modalities is beyond the scope of this article; again, our aim is to investigate the spirit of the communicative acts, and to determine their location on at least some of the axes of semantic space. With this caveat in mind, the PGOAL represents a specialization of $O\,E_i\,p$ (normative position T_5). That is, since we are not here concerned with issues such as sanctions and contrary-to-duty obligations, a PGOAL can be seen as one specific operationalization of an agent's commitment to bring about a state of affairs. Thus the KAoS act confirm occupies only point T_5 along the Lindahl-postcondition axis (all other points are prohibited by the strong commitment captured in a PGOAL by which passivity with respect to p and permission to bring about $\neg p$ are excluded).

Although this communicative act's position on the postcondition axis is clear, its position on the precondition axis is less so. If an assumption of consistency is made about an agent's PGOALs (Cohen and Levesque 1990), then confirm, at least, cannot be applied from a situation in which the agent is not permitted to bring about p (T_4, T_6 or T_7).

Also, according to the PWAG, such a confirm also entails that the agent does not believe that the state of affairs already holds, $\neg BEL(i, p)$. However, the PWAG is defined upon a PGOAL which requires that $BEL(i, \neg p)$. Since it can be reasonably assumed that if an agent believes $\neg p$, it is not the case that it believes p, we will focus on the latter. This definition of the confirm act motivates the introduction of the third dimension to semantic space that will be considered here: the agent may believe p or it may believe $\neg p$. As suggested by the KAoS specification, it may be useful to require that prior to an agent performing a commissive act such that the postcondition of that act is that it moves to the normative position T_5, it must believe $\neg p$. A specification that captures the spirit of the KAoS commissive act may be defined as follows:

$$< i, \text{KAoS-commit}(\, j, p) >$$

$$\text{FP:}\ (T_1(i, p) \vee T_2(i, p) \vee T_3(i, p)) \wedge \text{bel}(i, \neg p)$$

$$\text{RE:}\ T_5(i, p)$$

This is the most restricted operator (in terms of its feasibility preconditons) specified within this section, but the question remains: how are agents to agree on the semantics of commissive acts (or any other communicative act) used in a particular dialog? This is the question addressed in the following section.

4. SEMANTIC FIXING

The concept of semantic space offers a rich framework for characterizing communicative action. This framework can be used in two distinct ways. Firstly, it could be used by the designers of agent systems and agent communication languages to ensure that their definitions are comprehensive and that they are aware of all the choices they are making (and importantly not making) when they specify the individual primitives. This is an important contribution and role in its own right. Secondly, the semantic space model can be used by the agents themselves to define or determine the semantics of their communication language at run-time. It is on this second role that we focus in the remainder of this section.

Different situations (auctions, brokering, etc.) and classes of situations (new encoun-ters, unbalanced exchanges, etc.) can be enhanced by the adoption of exactly the right

communication primitives: this, after all, is the premise from which designers of agent communication languages are working. Rather than attempting to design a generic set of primitives with which agents can make do in any situation, we advocate providing agents with a means to determine precise semantics for a precise set of primitives appropriate to the given context. By fixing their own set of communication primitives, agents not only eliminate the possibility of misunderstandings, but also open up the potential for computational savings (using one tailor-made primitive where otherwise, several generic acts would be required, for example).

There is a range of processes that might be employed to carry out the task of fixing the semantics for a set of primitives. At one end of the range, there are techniques which are fast, simple to implement, and extremely inflexible. It is here that languages predefined by designers belong: once the agent system is started, all semantic shift and variation is precluded. The range and meaning of all primitives is (hopefully) well defined beforehand, and agents are obliged to use only those primitives provided by the designer.

At the next level of complexity lie mechanisms that rely on some limited bootstrapping foundation: in order to communicate to determine the semantics of primitives, it is necessary to have at least some primitives with which to carry out that communication. Clearly, the semantics of these latter primitives must be determined *a priori*, or else the semantic fixing would be unreliable. And yet, equally, it is important to avoid the infinite regress that is looming. A pragmatic approach suggests itself. So long as the set of basic bootstrapping primitives is both small and easily defined, the overhead required seems entirely justified when balanced against the gains available from dynamic semantic fixing. Within this constraint, however, there lies an enormous variety of potential mechanisms. One such is explored in detail below.

As the complexity of the bootstrapping primitive set increases, so too does the potential complexity of the mechanisms which can avail themselves of such rich communication. One example is the use of argumentation structures such as those suggested by Kraus et al. (1998), Parsons and Jennings (1996), and Reed (1998). Agents use forms of 'accept' and 'propose' to exchange, critique, and incrementally develop the semantics of individual primitives. There is a trade-off here: with rich bootstrapping primitives such as these, the mechanism of semantic fixing can be efficient, flexible, and fast—but it requires substantial overhead in the initial specification of agent communicative capabilities. If the process is to be employed as a means of bridging existing agent societies which rely on different protocols, or as a way of interacting with newly recruited agents with unknown communication competence, then it is a lot to ask that any arbitrary agent should be equipped with a sophisticated set of bootstrapping primitives. It seems more reasonable to demand that they have just one or two simple primitives which can be reliably employed in the fixing of others. With the current diversity in agent communication languages, and the difficulties faced in pinning down precisely the semantics of complex primitives, the middle point in the range appears much more appealing.

A strong candidate for a mechanism of semantic fixing which occupies a point in the middle of the range is one based upon voting. Voting has a long tradition of being used as a social choice mechanism for group decision making (Sandholm 1999) and the one suggested here has the merit of being very simple. Like other mechanisms that attempt greater flexibility than that offered by a predetermined communication language, the simple voting mechanism (SVM) makes two key assumptions. The first is that agents share an ontology of semantics. That is, agents can talk about the meanings of primitives in a common language—the FIPA SL, or a KD45 logic, or whatever. A common ontology is as vital for talking about the semantics of primitives as it is for talking about more concrete domains such as manufacturing, e-commerce, or health care. The discussion in Sections 2 and 3, above, shows that the enterprise of manually converting definitions of primitives into a

common specification language is entirely feasible, and so it is argued that this is a reasonable assumption, even if full intertranslation has not yet been achieved. Of course, in time, it might be hoped that a common ontology of communication might be used as the native means of describing communication primitives and hand-crafted languages. We return to potential problems posed by the assumption of a common ontology at the end of this section.

The second assumption is that agents are already arranged in some sort of group by which they are aware of one another. The coupling of the group may be extremely loose—it may simply be two agents at a communicative encounter, or it may be an entire society of agents. Alternatively, the coupling may be very tight with agents in the group having agreed on joint goals and the apportionment of responsibility. In either case, the need arises for the language of the group to change. (The process by which such a need is identified is not our concern here. One might imagine a specified protocol that indicates semantic fixing is appropriate at a particular stage in some proceedings—for example, after a group of collaborating researchers decide to write a paper together, they may then always proceed to lay out rules for how drafts are to be exchanged.)

On recognizing that a process of semantic fixing is appropriate under SVM, an agent must broadcast the fact to the other group members (to broadcast, the agent must be aware of all the members of the group; hence assumption number two, above). This, then, is the first bootstrapping primitive: the 'initiate' broadcast. SVM dictates that agents then respond by offering suggestions of semantic primitives. Again, we are not directly concerned with the reasoning processes by which individual agents come up with suggestions: the SVM will work as well with random suggestions as with those that are carefully constructed—though of course the resulting language will have much greater fit for purpose in the latter case. However, it is important to mention one particularly significant factor in the generation of appropriate suggestions, namely the type of dialog that is to ensue. Though there are potentially very many ways of describing the various types of dialog, one influential theory that has been shown to be useful in characterizing agent dialog encounters is Walton and Krabbe's dialog typology (1995). In this work, six types of dialog are distinguished: persuasion, negotiation, deliberation, inquiry, information-seeking, and eristic. Each dialog type has its own characteristics, assumptions, and context, all of which can be exploited in computational work (McBurney and Parsons 2001; Reed 1998). By recognizing which type of dialog is appropriate in a given situation, it is possible to select appropriate clusters of primitives. So for example, in the motivating example offered by Gmytrasiewicz and Huhns (2000) of Polish- and English (only)-speaking humans successfully haggling and purchasing goods from a Catalan (only)-speaking shopkeeper, all parties recognize that they are entering into some sort of negotiation dialog. This immediately restricts the space of appropriate primitives from which to choose to those in the propose, counterpropose and accept clusters. The determination of precise semantics of individual primitives within those clusters is then heterogeneous across agents. Thus agents who are more likely to be selling will favor semantic characterizations which reduce their own liability, at the expense of buyers.

An agent submits a suggestion by broadcasting it: the 'suggestion' broadcast is thus the second bootstrapping primitive. There are no restrictions on how many suggestions for primitives an agent may ultimately submit: the only restriction that SVM imposes is that submissions should occur in rounds, and that at each round only a single suggestion be adopted—the adoption is the primitive that was suggested most often (with some straightforward deterministic procedure for resolving ties). One possible mechanism for resolving such a tie is to use a casting vote 'token.' Along with the initiate primitive, a sequence of agent identifiers are sent that indicates the sequence in which the agents have the casting vote. This virtual casting vote 'token' is passed around the group of agents involved in the SVM. The agent with the casting vote in each round will, in the case of a tie, send a second suggestion

broadcast indicating its choice between the most popular suggestions made in that round. This second suggestion will resolve the tie.

It may seem that with such an enormous semantic space (see below), duplicate suggestions are likely to be extremely rare. However, a more careful analysis suggests that though not homogeneous, the reasoning capabilities of agents in a group are likely to be sufficiently similar to lead typically to substantial pruning of semantic space, and a consequent increase in the likelihood of duplicate submissions. Given that agents may often select the same dimensions upon which to negotiate (the Lindahlian dimensions for contract-related primitives, for example) it is reasonable to expect substantial pruning in real situations. Ultimately though, even in situations where this serendipitous convergence is absent, and the number of suggested primitives becomes large, SVM guarantees a language that all members of the group can share, even if it is suboptimal. Similarly, SVM and the proposed approach to negotiating semantics in general do not presuppose any particular reasoning capabilities on the part of the agents involved, but the more sophisticated those capabilities are, the quicker a language might be constructed, and the better its 'fit for purpose' when that language is subsequently employed to solve the problems at hand.

A single round of suggestion broadcasts thus leads to the agent in the group adding a single primitive to their language. This modification to an agent's language can then affect its reasoning for the next suggestion. So, for example, if an agent had suggested a particular variant of a 'propose' primitive but had been outvoted by several other members of the group all suggesting some common alternative 'propose,' it is up to the agent to decide whether or not the other 'propose' is adequate. If it is, then there is no need to resubmit the initial suggestion. Alternatively it could be that the current 'propose' places the agent at a disadvantage (for an agent that typically proposes, it would disadvantageous if that act imposed severe commitments on the proposer while imposing none on the recipient, for example). In this case, the agent may decide to resubmit the original suggestion. In any round, an agent may submit a null, indicating contentment of the current state of the putative language. SVM terminates at a round in which all agents submit null.

To summarize SVM, consider a finite set of agents, `A`, a finite set of communicative acts, `C`, and a subset of `C` that is the resulting language, `L`. Initially, `L = {}`, and some agent in `A` will 'initiate' the procedure. Each agent in `A` will have a function that is used to choose its next suggestion to be broadcast to the group: for agent `a` the function is `choose_a()`. This function will evaluate a set of communicative act specifications to either one of those specifications or the special word `null`. The only restriction on the definition of this function is that the empty set evaluates to `null`:

```
choose_a({}) = null
```

On receipt of a suggestion from all members of the group, each agent will check the set of suggestions, terminate the SVM if all are `null` or update the language `L` with the most popular primitive otherwise. The simple process of determining the most popular primitive is performed by the function `select-primitive()` (which, as stated above, has deterministic behavior in the case of a tie.) The `valid-p()` predicate function evaluates to `false` if a suggestion is `null`, already contained in `L` or not contained in `C`. The remaining set, VALID, then contains all suggestions that are candidates for including in `L`. This filter, therefore, removes from consideration any invalid communicative act specifications. The simple voting mechanism, therefore, operates as shown in Figure 3.

It should be clear that SVM is thus theoretically open to abuse. An agent could decide to be entirely dogged, and repeatedly submit its own suggestions until they were all accepted (eventually this would necessarily be the case). And yet such dogged behavior would fail to result in any advantage accruing to the agent. For although all agents in the group would, at

```
SVM(A,C,L)

   SUGGESTIONS = {(a, choose_a(C MINUS L)) : a<-A}

   VALID = {s : s<-SUGGESTIONS s.t. valid-p(s, C, L)}

   IF emptyset-p(VALID) THEN

     RETURN L

   ELSE

     L' = L UNION {select-primitive(VALID)}

     SVM(A,C,L')

   ENDIF

ENDSVM
```

FIGURE 3. The simple voting mechanism.

the termination of SVM, be able to understand (have a precisely defined semantics for) all the variously proposed primitives, there would be no incentive to employ all those primitives. Assume, for example, that our dogged agent is a buyer, and has submitted a range of 'offer' primitives which are highly advantageous to buyers (a good example would be an 'offer' that on acceptance commits a seller absolutely, while the buyer is free to continue to search out a better deal). Sellers may simply refuse to enter communicative exchanges in which such a biased 'offer' is employed. In that case, if the buyer wanted to buy, it would be obliged to use a less skewed communication primitive.

It is clear, again, that such freedom could lead to an impasse with neither buyers nor sellers agreeing to use a common set of primitives. This though is directly analogous to determining a price: giving agents the flexibility to negotiate over the price of a good admits a scenario in which buyer and seller fail to reach a price upon which they both agree. If the agents have the freedom, and are not by their own goals compelled to make the sale (or purchase), then they may simply fail to reach an agreement in that scenario. By way of a contrived example, if one were at liberty to negotiate whether the purchase of a particular house in Berwick (a town on the English-Scottish border) was to be governed by the law of Scotland (in which an accepted offer is binding, having been made through a solicitor), or by that of England (in which a seller may "gazump" on an accepted offer if a better offer is received), then the buyer might prefer the former, and the seller the latter. If neither were desperate to close the deal, then it could fall through before any negotiation had taken place at all.

The semantic fixing mechanism SVM is thus a profligate language construction process, admitting any suggestion at all, though prioritizing the order in which suggestions are accepted in the hope that it might lead agents to decide not to submit further suggestions. Though relying for efficiency on the internal design of agents, SVM encourages agents to restrict their suggestions by offering no gain for doggedness, and, in extreme cases, offering nothing but greatly increased computational cost for individual agents who pursue strategies of doggedness (since it is they that would be attempting to use the full gamut of primitives which they had suggested). Notice that the much greater overheads involved in the argumentation-based semantic fixing mentioned above could guarantee a much more parsimonious language construction process, by diligently assessing the support for adoption of an individual suggestion. The simplicity, speed and ease of implementation of SVM, however, coupled with the lack of incentive for its abuse, still outweighs the problems associated with more complex mechanisms.

Although SVM is a profligate language construction process, some remarks may be made about its computational limits. It can be shown that SVM will terminate under reasonable conditions. Suppose, in general, that semantic space has d dimensions where the *average* number of discrete positions in each dimension is n.[11] The initial and final states in Figure 2 are treated as two seperate dimensions in this discussion, and these relate to the feasibility preconditions and rational effects of a communicative act respectively. Therefore, the fact that for each communicative act (e.g., moving from T_1 to T_2) there exists a symmetric act (e.g., moving from T_2 to T_1) is accounted for in d. The maximum number of single-point communicative acts that may be defined in such a semantic space is, therefore, n^d. For example, if the semantic space consists of the two dimensions illustrated in Figure 2 (i.e., the normative position of a single agent with respect to a single proposition before and after the utterance), the number of primitive communicative acts is 49.

In Section 3.1, however, it was noted that there is a possibility that a communicative act may be underspecified. This means that a single communicative act may specify a set of possible transitions in semantic space. For example, a communicative act that simply permits an agent to bringing about p may be characterized as the set of transitions $\{(4, 1), (6, 2), (7, 3)\}$. Thus, the set of communicative acts where underspecification is permitted will be the power set of the set of primitive communicative acts. The number of such acts is 2^{n^d}. Again, if the semantic space is that illustrated in Figure 2, the number of possible communicative acts is 2^{49}.

A proportion of these acts are redundant; i.e., those that effect no change in the normative position of the agent concerned. The communicative action characterized by the set of transitions $\{(1, 1), (3, 3)\}$ is an example. Furthermore, the action characterized by the set of transitions $\{(1, 2), (3, 3)\}$ is equivalent to that characterized by $\{(1, 2)\}$. These redundant communicative actions occur simply because the two dimensions discussed represent the pre- and postconditions of an action.

Assume that, for each dimension in semantic space representing a precondition of an action there is a matching dimension representing the postcondition of that action. By way of a further example, an agent may believe (or not) p before and after the communicative act. In general, therefore, the number of communicative acts for a semantic space with d dimensions, where the average number of positions on each dimension is n, is $2^{n^d - n^{d/2}}$. For example, if the semantic space consists of the two dimensions illustrated in Figure 2, the number of possible communicative acts is 2^{42}.

The examples used so far concern the normative position of a single agent with respect to a single proposition before and after an utterance. A more realistic situation to consider is when the normative positions of both speaker and hearer in a communicative exchange are considered in semantic space. Consider, for example, the trading of options. Permission to bring about p represents the right of the agent to buy an underlying asset at a certain time in the future at a certain price ('call' the option). Holding an option also provides the agent with the position to confer this right to others (sell or 'put' the option). In 'putting' an option, however, the freedom of the agent toward the option is removed. The action of 'putting' an option thus changes the normative state of both the speaker and the hearer. Prior to the act, the speaker, say i, is permitted to bring about p ('call' the option) or to remain passive toward p (no agent, in this example, is permitted to bring about $\neg p$). Prior to the act, the hearer is only permitted to remain passive. After i sells the option to j, their positions are reversed. Thus, the act of putting an option could be characterized in the following way: the speaker, i,

---

[11] Assuming that we have a finite number of dimensions, SVM is guaranteed to terminate if n is finite. In other words, SVM cannot be guaranteed to terminate if there is one dimension in which the number of positions that an agent can take towards the propositional content of a message is infinite. (We return to this point in Section 6.)

moves from state 2 to 6, and the hearer, j, moves from state 6 to 2. In this case, the number of dimensions of semantic space is four: the normative positions of the speaker and hearer before and after the utterance. Therefore, if $d = 4$, the number of possible communicative acts is 2^{2352}!

One of the criticisms that has been leveled against the process of semantic fixing focuses on the semantic space itself rather than on the process by which that space is navigated. In order that agents are able to negotiate meanings, they must share a common understanding of semantic space: that is, the dimensions of semantic space must be available as a common ontology. Surely, the criticism runs, such provision is far more complex than the tasks either of designing a common language or of modeling the process of negotiating semantics. In answering the criticism, it is first important to note that agents need only share a common understanding of a small part (that is to say, of a small number of the dimensions) of semantic space for the approach to work in a given situation. Underspecifying communication primitives with respect to the dimensions of semantic space that are not shared is equivalent to underspecifying with respect to dimensions that are shared. Thus the problem of designing some monolithic shared ontology is reduced to the problem of ensuring that enough of the ontology of semantic space is shared for producing a language with a good 'fit for purpose' in the situation at hand. But this alone fails to defuse the criticism entirely: indeed, the design and implementation of even just fragments of such an ontology is undeniably a major practical and theoretical challenge. The argument we are presenting, however, demonstrates that the emphasis should be upon creating this common ontology of communication, rather than upon attempting to enforce a single common language upon all agent communicative encounters. Standardization is inevitable in a domain in which, as yet, an assumption that agents can observe a common scene and ground their utterances in it, is simply unreasonable. The focus of that standardization needs to be squarely upon how communication can be described, rather than upon a library of communication primitives.

Clearly, semantic space is very rich, and supports the construction of an enormous range of subtly different communicative acts. The size of the construction task is fully to be expected if the characterization supports such a fine-grained building material. The aim is to provide agents with a system by which they can tune a language with great precision to the needs at hand, and the ability to do this outweighs the potential pitfalls of any particular language construction process, such as SVM. SVM is thus well suited to domains in which agents might reasonably be expected not to suggest a huge number of different primitives. In a domain in which this assumption was not reasonable, some other mechanism might be more appropriate—say, one in which agents have to rank their own proposals and submit a fixed number. Of course, it would also be possible to design the construction process in such a way that, for example, only point-based communicative primitives were to be submitted for consideration. This would bring down the complexity dramatically (it would no longer be necessary to work on the power set of the points in semantic space), but at the cost of requiring longer sequences of primitives in the final language. This decoupling between semantic space and the mechanism by which language is constructed from it is one of the advantages of the approach.

5. EXAMPLE OF SVM

To illustrate SVM, consider three agents, each of which wish to introduce communicative acts such as commands, permissive, and commissive acts into a shared communication language. Each act specifies (or partially specifies) transitions of the speaker and hearer according to the Lindahl state transitions illustrated in Figure 2. These acts are represented

here with a set of transitions for the speaker and an equivalent set for the hearer. For instance, an act may specify that, before the act, the speaker, i, is permitted to remain passive toward the propositional content of the act and after the act, i is committed to remain passive. In other words, i is, before the act, in any of the Lindahl states 1, 2, or 4 and after the act in the state 6. Thus, the set of transitions for the speaker is: $\{(1, 6), (2, 6), (4, 6)\}$. For the hearer, j, before the act, j is permitted to bring about p and after the act, j is committed to bring about p. In other words, j is, before the act in any of the states 1, 2, or 3 and after the act in the state 5. Thus, the set of transitions for the hearer is: $\{(1, 5), (2, 5), (3, 5)\}$. This particular communicative act results in the hearer being obliged to bringing about p and the speaker being obliged to remain passive toward p: the hearer must bring about p and the speaker cannot interfere. This is communicative act 3.b in the example below.

The initial state of the semantic fixing between these three agents is that agents 1, 2, and 3 are interested in the following sets of communicative acts being included in the language:[12]

Agent 1. This agent wishes to introduce two actions into the language.

    `1.a` A command that commits the hearer to bring about p such that the hearer is not *a priori* forbidden from doing so.
        Speaker: {}
        Hearer: $\{(1, 5), (2, 5), (3, 5)\}$

    `1.b` An act that commits the speaker to bring about p such that the agent is *a priori* forbidden from doing so.
        Speaker: $\{(1, 5), (2, 5), (3, 5)\}$
        Hearer: {}

Agent 2. This agent wishes to introduce two actions into the language.

    `2.a` An act the permits the hearer to bring about p such that the agent is *a priori* committed to remain passive.
        Speaker: {}
        Hearer: $\{(6, 2)\}$

    `2.b` An act that commits the hearer to remain passive toward p such that the agent is *a priori* permitted to doing so or remaining passive.
        Speaker: {}
        Hearer: $\{(2, 6)\}$

Agent 3. This agent wishes to introduce two actions into the language:

    `3.a` A command that commits the hearer to bring about p and the speaker cannot interfere.
        Speaker: $\{(1, 6), (2, 6), (4, 6)\}$
        Hearer: $\{(1, 5), (2, 5), (3, 5)\}$

    `3.b` A put-option act (see Section 4 above).
        Speaker: $\{(2, 6)\}$
        Hearer: $\{(6, 2)\}$

SVM then proceeds in the following way:

Round 0. Agent 1 broadcasts `initiate(1,2,3)` (1-2-3 being the casting vote sequence). The language, L is initialized.

---

[12]Each communicative act specification refers to the changes in normative position of the agents that will take on the roles of speaker and hearer when the act is used during communication.

Round 1. Agent 1 has the casting vote. Agent 1 broadcasts `suggestion(1.a)`; agent 2 broadcasts `suggestion(2.a)`; and agent 3 broadcasts `suggestion(3.a)`. There is a tie. However, rather than using its casting vote to force the inclusion of `1.a`, agent 1 decides to endorse agent 3's suggestion. Agent 1 broadcasts `suggestion(3.a)`, and so this act is included in L.

Round 2. Agent 2 has the casting vote. Agent 1 broadcasts `suggestion(1.b)`; agent 2 broadcasts `suggestion(2.a)`; and agent 3 broadcasts `suggestion(3.b)`. There is a tie, and so the agent with the casting vote, agent 2, broadcasts `suggestion(2.a)`. `2.a` is included in L.

Round 3. Agent 3 has the casting vote. Agent 1 broadcasts `suggestion(1.b)`; agent 2 broadcasts `suggestion(2.b)`; and agent 3 broadcasts `suggestion(3.b)`. There is a tie. Although `2.a` and `2.b` use the same transitions as `3.b`, `3.b` is being introduced for a different purpose—for the trading of options—and so agent 3 uses the casting vote to broadcast `suggestion(3.b)`. `3.b` is included in L.

Round 4. Agent 1 has the casting vote. Agent 1 broadcasts `suggestion(1.b)`; agent 2 broadcasts `suggestion(2.b)`; and agent 3 broadcasts `suggestion(null)`. There is a tie, and so agent 1 uses the casting vote and broadcasts `suggestion(1.b)`. `1.b` is included in L.

Round 5. Agent 2 has the casting vote. Agent 1 broadcasts `suggestion(null)`; agent 2 broadcasts `suggestion(2.b)`; and agent 3 broadcasts `suggestion(null)`. `2.b` has the only vote, and so this is included in L.

Round 6. Agent 3 has the casting vote. Agent 1 broadcasts `suggestion(null)`; agent 2 broadcasts `suggestion(null)`; and agent 3 broadcasts `suggestion(null)`. SVM terminates. L = {`3.a`, `2.a`, `3.b`, `1.b`, `2.b`}.

Here, suppose that agent 2 is responsible for access to an information source. The two acts introduced by this agent, `2.a` and `2.b`, allow it to permit and forbid access. Although agent 3 is not in control of this information source, it may give up its licence to access the information source. Agent 2, the manager agent is interested in issuing commands and allowing agents to commit to activities, hence its interest in `1.a` and `1.b`. It does, however, accept the inclusion of `3.a` rather than `1.a`—it accepts that it should not interfere with agents to whom it has given commands. This simple example illustrates the sorts of communicative actions that can be included in a common language and how the simple voting mechanism may be used to construct such a language. This language can be seen as a subset of a more complete language for managing the activities of agents within an organization. Certain relevant features of such a language are illustrated including different types of commands, the management of information/resource access, etc.

6. DISCUSSION AND RELATED WORK

Further analysis of Lindahl's (1977) theory indicates that the individualistic two-agent positions discussed above can be generalized to n-agent normative positions. Thus, in a communicative exchange, an act could specify the positions prior to and following the utterance of the speaker, the hearer and, for example, an agent that overhears the utterance. Collectivistic two-agent positions—those that cannot be reduced to a conjunction of one-agent types—are also of interest. These allow acts such as "one of you two must take out the garbage" to be specified: $O(E_i p \vee E_j p)$. Such positions are common if a group of agents is being addressed; see Kumar et al. (2000) for some work on communicating with groups of agents within the KAoS framework, and Norman and Reed (2002) for a view on the logic of group-directed imperatives.

Normative systems specification has been shown to be a useful method in the specification of multi-agent systems—for some recent work see Krogh and Herrestad (1999), Jones and Sergot (1996), Santos et al. (1997), and Singh (1999). These authors and others emphasize different requirements of such specifications. For example, Krogh and Herrestad (1999) and Singh (1999) emphasize the "directedness" of norms; i.e., that certain norms are relations between two (or more) specific agents. Sergot (1999) further analyzes Lindahl's theory and presents a system, Norman-G, that exploits the structure of the space or normative positions to aid in the disambiguation of legal texts. Santos et al. (1997) investigate the possibility of considering action that is unsuccessful (the logic of E is that of successful action). Norman and Reed (2001) discuss, following von Wright (1968), the distinction between bringing about a state of affairs and bringing about that an action is done. Each of these models, or ontologies of norms, may be relevant to the task of negotiating the semantics of a shared communication language. The important issue, however, is that, for the agents to communicate in the first place they must share such an ontology. The mechanism proposed here simply allows the agents to agree upon a number of communication primitives that capture the meaning of the messages the exchange.

Using the agent language 3APL, Hindriks et al. (2000) initially provide a parsimonious set of communicative primitives on which to build their model. They describe two primitives, tell and ask, for information exchange, and two more, req and offer, for the exchange of a request. The focus of this work is squarely upon the hearer and the deductive (for information) and abductive (for requesting) logical reasoning that the hearer is able to perform after the receipt of a message. Nevertheless, they do comment (Hindriks et al. 2000, Sec. 3) that more complete primitive definitions can be defined by introducing further constraints. Their example is

$$\mathsf{inform}(i, p) =_{df} p?; \mathsf{tell}(i, p)$$

which defines the inform by introducing a sincerity condition on the hearer in addition to the tell primitive. They also discuss how more complex exchanges—which might in some systems be characterized as single primitives—can be composed from the four basic acts, taking as their example various exchanges in meeting organization.

Thus, this facet of the Hindriks et al. work can be seen as concentrating on certain dimensions of semantic space—those characterizing the hearer's state after the receipt of a primitive. Each of the four basic primitive acts represent a single point within that space, and then, by composition and by introducing further restrictions on the speaker, gives access to a subset of the rest of the space. The precise semantic definition of the four acts, which is one of the attractions of the work from the point of view of formal ACL design, necessarily restricts the space that can be addressed. So for example, tell leads the hearer to update (that is, to introduce while maintaining consistency) its belief set with the propositional content. This is therefore a more 'optimistic' primitive in the inform cluster, closer to FIPA inform than it is to KQML tell or the KAoS INF. The 3APL tell thus lies at a precise point on the 'level of success' axis, which precludes using it to define any of the other members of the cluster (such as FIPA, KQML, and Cohen and Levesque), which lie at different points on that axis.

Hindriks et al. (2000, Sec. 6) comment: "Instead of defining a set of communicative primitives corresponding to all kinds of subtle distinctions related to speech acts, we propose to derive what kind of speech act has been performed by an agent from the characteristics of the content proposition and the mental state of the agent." We thoroughly agree that there are indeed all kinds of subtle distinctions, but argue that these distinctions have real operational value, which can be exploited by the agents themselves. To do so demands not defining a set of primitives *a priori*, but allowing agents to develop that set for themselves.

Finally, the work of Steels and Kaplan (1999) tackles the problem of language acquisition by a group of agents by using environmental stimuli to give a grounded semantics to new primitives. Agents employ video cameras to view a common environment (a whiteboard with variously colored and positioned geometric shapes), and then play a 'guessing game' which implements a simple form of reinforcement learning. One agent uses a primitive to describe a feature in the environment; a hearer agent then uses past knowledge of the primitive to guess the feature being described. Success reinforces the inferred meaning of the primitive; failure decreases confidence in a particular interpretation. The work shows that a language evolves which has primitives of varying abstraction (e.g., "in the bottom half," "red and right in the top left corner"), and also demonstrates polysemy and synonymy.

Steels and Kaplan (1999, p. 863) claim that "agents start with no prior designer-supplied ontology nor lexicon." However, an implicit ontology is provided, at least in terms of its broad structure and bounds: the capabilities of the camera enforce one set of restrictions, and the designer's assumptions of salient features enforce another. Thus, for example, color and position dominate the meanings of all the primitives examined, though it is the agents themselves which do indeed determine for specific primitives the values—or, rather, sets of values—along these axes.

Thus Steels and Kaplan (1999) focus on a specific semantic space, having axes of color and position. The individual primitives that they discuss have either specific values on one or more axes ("red," "blue and on the edge"), or have ranges of values on one or more axes ("toward the center," "close to the left and toward the top"). There are several important differences in the way that primitive meanings are determined in their work and that described in this article. First, the Steels and Kaplan approach competently handles axes with continuous values. As described above, this is something which a technique like SVM would be poor at handling, but at which an argumentation-based negotiation might be much better. In addition, however, there are also several reasons why the Steels and Kaplan work would be ill suited to the problems described in this article. First, when the meaning of primitives is defined in terms of the state of the real, observable world, feedback is a straightforward matter (of pointing to an area of the whiteboard). When such meanings are based on mentalistic states, normative positions, and the myriad other abstract notions which characterize much current work in ACLs, it becomes difficult to see how such direct feedback can be effected. Secondly, Steel and Kaplan's multi-agent system has the luxury of being able to get things wrong for quite a long time—their graphs showing the evolution of the language have an x axis measured in thousands of communicative exchanges. The current approach, in contrast, offers a potential means for determining a language with good fit for purpose quickly and efficiently.

7. CONCLUSIONS AND FUTURE WORK

This article has advocated a new approach to agent communication languages. Rather than viewing the specification of ACLs as an off-line, design-time process, we have argued that in open multi-agent systems it should be a dynamic, run-time process. The reason for this is that it enables agents that have not been designed to interoperate to come to a mutual and explicit understanding about the semantics of their communication primitives and because it enables agents to tailor the meaning of these primitives to the context in which they find themselves. The method by which agreements about the use and meaning of these primitives can be obtained is one of negotiation. This is because the agents are autonomous stake-holders and, if they are to interact successfully, they need to come to a mutually acceptable agreement both on the primitives that should be used and on their semantics.

To this end, we introduced the concept of semantic space which provides, for the first time, a structured framework for analyzing the potential properties of agent communication languages. This framework can be used by designers of such languages in order to understand the full range of possibilities for their choice of communication primitives. Presently, such design is fairly ad hoc in nature. However, the existence of a systematic framework also opens up similar opportunities for the agents themselves to complete this process at run-time. Thus, agents can use their knowledge of the dialog type, their communication objectives, and their social relationships with one another to tailor the communication language to their prevailing circumstances. Specifically, we argued that a simple voting mechanism is the most appropriate method for organizing this semantic fixing activity and we detailed an appropriate protocol that could be used for this endeavor. The proposed approach demands a common understanding of (at least part of) semantic space shared between agents, but, we argue, such an assumption of common understanding is both inevitable at some level, and also most appropriately made at the level of the ontology of possible communication semantics, rather than at the level of the semantics of a particular language.

For the future, a number of important avenues require further investigation. First and foremost, the structure of semantic space—the fragments of an ontology of communication—need to be further developed. Section 3 of this article has demonstrated the approach, but it needs to be applied to other dimensions, and, particularly to dimension clusters. Thus the dimensions important to contract-related primitives, including the Lindahlian and sincerity dimensions have been developed here. Other clusters of dimensions, for example, for information exchange, competitive bidding, and so on, need to be identified, teased apart and added to semantic space in a coherent fashion. This enterprise will inevitably both offer up interesting research questions, invigorating the area, and also lead to pragmatic solutions in the standardization task. This mapping will invariably extend the dimensionality of the framework and it will be interesting to observe the extent of a fully expanded space. Second, additional types of communication primitives need to be mapped into the semantic space framework. Third, it will be interesting to see whether this approach can be applied to extended communication exchanges (i.e., to conversations and protocols) where many of the same requirements about flexibility, dynamism, and context awareness also apply. Finally, the framework needs to be made operational by applying it to open communities of agents.

ACKNOWLEDGMENTS

REFERENCES

CHELLAS, B. F. 1980. Modal logic: An introduction. Cambridge University Press.

COHEN, P. R., and H. J. LEVESQUE. 1990. Intention is choice with commitment. Artificial Intelligence, **42** (2–3):213–262.

FIPA. 2000. Agent communication language specification. *http://www.fipa.org*.

GMYTRASIEWICZ, P., and M. N. HUHNS. 2000. The emergence of language among autonomous agents. IEEE Internet Computing. pp. 90–92.

GROSZ, B., and S. KRAUS. 1996. Collaborative plans for complex group action. Artificial Intelligence, **86**:267–357.

HINDRIKS, K. V., F. S. DE BOER, W. VAN DER HOEK, and J.-J. CH. MEYER. 2000. Semantics of Communicating Agents Based on Deduction and Abduction. *In* Issues in Agent Communication, Lecture Notes in Artificial Intelligence, vol. 1916. Springer-Verlag, Berlin, pp. 80–95.

JENNINGS, N. R., P. FARATIN, A. R. LOMUSCIO, S. PARSONS, C. SIERRA, and M. WOOLDRIDGE. 2001. Automated negotiation: prospects, methods and challenges. International Journal of Group Decision and Negotiation, **10**(2):199–215.

JENNINGS, N. R. 2000. On Agent-Based Software Engineering. Artificial Intelligence, **117**(2):277–296.

JONES, A. I. J., and M. J. SERGOT. 1996. A formal characterisation of institutionalised power. Journal of the IGPL, **4**(3):429–445.

KRAUS, S., K. SYCARA, and A. EVENCHIK. 1999. Reaching agreements through argumentation: A logical model and implementation. Artificial Intelligence, **104**:1–69.

KROGH, C., and H. HERRESTAD. 1999. 'Hohfeld in cyberspace' and other applications of normative reasoning in agent technology. Artificial Intelligence and Law, **7**(1):81–96.

KUMAR, S. M. J. HUBER, D. R. MCGEE, P. R. COHEN, and H. J. LEVESQUE. 2000. Semantics of agent communication languages for group interaction. *In* Proceeding of the Seventeenth National Conference on Artificial Intelligence. pp. 42–47.

LABROU, Y., and T. FININ. 1997. Semantics and conversations for an agent communication language. *In* Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence. pp. 584–591.

LINDAHL, L. 1977. Position and Change: A Study in Law and Logic. D. Reidel Publishing Company.

MCBURNEY, P., and S. D. PARSONS. 2001. Agent ludens: Games for agent dialogs. *In* Proceedings of the Third Workshop on Game-Theoretic and Decision-Theoretic Agents. *Edited by* P. Gmytrasiewicz and S. Parsons. AAAI Press.

NORMAN, T. J., and C. A. REED. 2001. Delegation and responsibility. *In* Proceedings of the Seventh International Workshop on Agent Theories, Architectures, and Languages. *Edited by* C. Castelfranchi and Y. Lespérance. Springer-Verlag, Berlin, pp. 136–149.

NORMAN, T. J., and C. A. REED. 2002. Group delegation and responsibility. *In* Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems.

NORMAN, T. J., C. SIERRA, and N. R. JENNINGS. 1998. Rights and commitment in multi-agent agreements. *In* Proceedings of the Third International Conference on Multi-Agent Systems. pp. 222–229.

PANZARASA, P., T. J. NORMAN, and N. R. JENNINGS. 2001. Social mental shaping: modelling the impact of sociality on autonomous agents' mental states. Computational Intelligence, **17**(4):738–782.

PARSONS, S. D., and N. R. JENNINGS. 1996. Negotiation through argumentation—a preliminary report. *In* Proceedings of the Second International Conference on Multi-Agent Systems. pp. 267–274.

PITT, J. V., and E. H. MAMDANI. 2000. Some remarks on the semantics of FIPA's agent communication language. Autonomous Agents in Mult-Agent Systems, **2**:333–356.

REED, C. A. 1998. Dialogue frames in agent communication. *In* Proceedings of the Third International Conference on Multi-Agent Systems. pp. 246–253.

SANDHOLM, T. W. 1999. Distributed rational decision making. *In* Multi-Agent Systems. *Edited by* G. Weiss. The MIT Press, Cambridge, MA, pp. 201–258.

SANTOS, F., A. I. J. JONES, and J. CARMO. 1997. Responsibility for action in organisations: A formal model. *In* Contemporary action theory: Social action, vol. 2. *Edited by* G. Holmström-Hintikka and R. Tuomela. Kluwer Academic Publishers, pp. 333–350.

SERGOT, M. 1999. Normative Positions. *In* Norms, Logics and Information Systems. *Edited by* P. McNamara and H. Prakken. IOS Press, pp. 289–308.

SINGH, M. P. 1999. An ontology for commitments in multi-agent systems. Artificial Intelligence and Law, **7**(1):97–113.

SMITH, I. A., P. R. COHEN, J. M. BRADSHAW, M. GREAVES, and H. HOLMBACK. 1998. Designing conversation policies using joint intention theory. *In* Proceedings of the Third International Conference on Multi-Agent Systems. pp. 269–276.

STEELS, L., and F. KAPLAN. 1999. Situated grounded word semantics. *In* Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence. pp. 862–867.

VON WRIGHT, G. H. 1968. An essay in deontic logic and the general theory of action. *In* Acta Philosophica Fennica, vol. 21. North Holland.

WALTON, D. N., and E. C. W. KRABBE. 1995. *Commitment in dialog: Basic conceptions of interpersonal reasoning.* SUNY.

WOOLDRIDGE, M. J. 2000. Semantic issues in the verification of agent communication languages. Autonomous Agents and Multi-Agent Systems, **2**(1):9–31.