Toward a Theory of Communication and Cooperation for Multiagent Planning

Eric Werner

University of Hamburg
Department of Computer Science, Project, WISBER, P.O. Box 302762
Jungiusstrasse 6, 2000 Hamburg 36, West Germany; and
Department of Computer Science
Bowdoin College, Brunswick, Maine 04011

Abstract

In this paper we develop a formal computational theory of high-level linguistic communication that serves as a foundation for understanding cooperative action in groups of autonomous agents. We do so by examining and describing how messages affect the planning process and thereby relating communication to the intentions of the agents. We start by developing an abstract formal theory of knowledge representation based on the concept of information. We distinguish two types of information: state information, which describes the agent's knowledge about its world (knowing that) and process information, which describes the agent's knowledge of how to achieve some goal (knowing how). These two types of information are then used to formally define the agent's representation of knowledge states including the agent's intentional states. We then show how situations and actions are related to the knowledge states. Using these relations we define a formal situation semantics for a propositional language. Based on this semantics, a formal pragmatic interpretation of the language is defined that formally describes how any given knowledge representational state is modified by a given message. Finally, using this theory of meaning of messages or speech acts, a theory of cooperation by means of communication is described.

Keywords: Automated Reasoning, Communication, Distributed Artificial Intelligence, Information, Intention, Knowledge Representation, Planning

1. Introduction

One of the most important and fruitful areas of research in Artificial Intelligence has been the planning of sequences of action and reasoning about action [Fikes & Nilsson 71, Sacerdoti 77, Moore 80, Pednault 85]. Recently attempts have been made to extend the theory and techniques evolved for single agent planning to multiagent planning [Konolige 80, Georgeff 83]. This work has given birth to an important new area: Distributed Artificial Intelligence (DAI) where the central problem is the cooperation of multiple intelligent agents to achieve a common goal [Genesereth et al. 86, Rosenschein 86].

However, little attention has been paid to the role of high-level communication in cooperative planning and reasoning [Rosenschein 86]. We will argue that communication must play a central role in multiagent planning and cooperative action since without communication the achievement of complex multiagent goals and actions is computationally unfeasible.

In this paper we develop a formal computational theory of high-level linguistic communication that serves as a foundation for understanding cooperative action in groups of autonomous agents. We do so by examining and describing how messages affect the planning process and thereby relating communication to the intentions of the agents. We start by developing an abstract formal theory of knowledge representation based on the concept of information. We distinguish two types of information: state information, which describes the agent's knowledge about its world (knowing that) and process information, which describes the agent's knowledge of how to achieve some goal (knowing how). These two types of information are then used to formally define the agent's representation of knowledge states including the agent's intentional states. We then show how situations and actions are related to the knowledge states. Using these relations we define a formal situation semantics for a language fragment. Based on this semantics, a formal pragmatic interpretation of the language is defined that formally describes how any given knowledge representational state is modified by a given message. Finally, using this theory of meaning of messages or speech acts, a theory of cooperation by means of communication is described.

After describing the problem in §2 and looking at previous approaches in §3, we develop a formal theory of knowledge representation in §4 - §5. We then develop the communication theory in §6. A pragmatics of speech acts is sketched in §7. A theory of social cooperation is outlined in §8.

2. The Problem of Social Action in DAI

How is it possible for a group of independent agents, such as humans, robots or processes in a distributed environment to achieve a social goal? By a **social goal** we mean a goal that is not achievable by any single agent alone but is achievable by a group of agents. Note that the coordination of sequential processes [Dijkstra 68] and the problem of multi-robot control [Lozano-Pérez 83] are special cases of this more general problem.

The key element that distinguishes social goals from other goals is that they require cooperation; social goals are not, in general, decomposable into separate subgoals that are achievable independently of the other agent's activities. In other words, one agent cannot simply proceed to perform its action without considering what the other agents are doing. Examples include the operation of a factory, the construction of a ship, or lifting a couch.

Complex social goals will require many levels of cooperation. How does a group of agents achieve the cooperation that is necessary to accomplish social goals?

The possible solutions to our problem range between two poles: From those involving no communication to those involving high-level, sophisticated communication. The solutions implicit in previous research fall somewhere in between. However, none of the previous approaches develop the solution adopted by human agents, namely, that of using high-level linguistic communication to achieve complex social action. This is the solution we will investigate. First, we look more specifically at previous approaches.

3. Previous Approaches

Previous research in computer science on multiagent action, e.g., in operating systems theory, distributed systems, parallel processing and distributed artificial intelligence DAI, has implicitly or explicitly taken a position with regard to the problem of how cooperative social action is to be achieved. They have been limited to the following kinds of communication:

3.1. No Communication

The agent rationally infers the other agent's intentions (plans) [Genesereth et al. 86, Rosenschein 86]. However, there are difficulties inherent in this approach: First, the solution fails to work when there are several optimal paths to the same goal. For then there is by definition no general rational way of deciding which choice to make, and communication is necessary to resolve the uncertainty. Second, rationally inferring the decisions of the other agents requires knowledge of the other agent's beliefs. How does the agent get that knowledge except by some form of communication? Third, if the other agents are themselves speculating on what the others are going to do, we get potentially infinite nestings of belief. Finally, irrespective of the above difficulties even if cooperation were possible by pure mutual rational deduction, the computational cost of rationally deducing the other agent's intentions would be enormous for cooperative activity of even mild complexity. We are not saying rational deduction is not used in cooperative behaviour. Indeed, often it is necessary: see related work on helpful responses [Allen 79, Allen and Perrault 80]. Our claim is that it is inadequate for achieving sophisticated cooperative action.

3.2. Primitive Communication

In this case, communication is restricted to some finite set of fixed signals (usually two) with fixed interpretations [Dijkstra 68, Hoare 78]. Georgeff [83] has applied this work to multiagent planning, to achieve avoidance of conflict between plans for more than one agent. It has also been applied in robotics to coordinate parallel activity [for a review see Lozano-Pérez 83].

The coordination made possible by these means is limited, being primarily used to avoid conflicts between sequential processes. Sophisticated cooperative action is virtually impossible.

The reason is that the direct reference to one of a large repertoire of actions is not possible due to the limited number and types of signals available. Arbitrarily complex actions cannot be formed since there is no syntax of signals to build up complex actions. Hence, arbitrarily complex commands, requests and intentions cannot be expressed. It is somewhat analogous to the distinction between machine-level and task-level robot programming [Lozano-Pérez 83].

3.3. Plan and Information Passing

The agent A communicates his total plan to B and B communicates her total plan to A. Whichever plan arrives first is accepted [Rosenschein 86]. While this method can achieve cooperative action, it has several problems: First, total plan passing is computationally expensive. Second, there is no guarantee that the resulting plan will be warranted by the recipient's database [Rosenschein 86]. In addition to Rosenschein's criticisms, there are general problems with any form of total plan passing: First, total plan passing as a communication strategy is unfeasible. In any real world application there is a great deal of uncertainty about the present state of the world as well as its future. Hence, for real life situations total plans cannot be formulated in advance, let alone be communicated. At best, general strategies are communicable to the agent with more specific choices being computed with contextual information. Similar difficulties arise with preformulated linguistic intentions [see Grosz 85].

Second, a given agent will usually have additional goals distinct from the sender. The sender must somehow guess the additional goals that the recipient wants if he is to choose the correct plan. A mutually satisfactory plan is guaranteed only if abstract goals and not just total plans can be communicated. Finally, and most importantly, how the plan is passed is left open, i.e., there is no theory of communication given.

As for information passing in isolation [Rosenschein 86], it suffers from all the problems mentioned in §3.1, except the second; since there is no explicit communication of intentions these must be deduced.

3.4. Message Passing

Hewitt [77] has, we believe, the fundamentally correct intuition that control of multiagent environments is best looked at in terms of communication structures. However, he gives no formal syntax, semantics, or pragmatics for such communication structures. Thus no systematic account or theory of communication for message passing between agents is given.

3.5. High-level Communication

A great deal of good work has been done on speech act planning [Cohen and Perrault 79, Allen and Perrault 80, Appelt 85]. It would seem this work would be ideal for our purposes. What is lacking is that those works are restricted to the planning by a single agent of some communicative act to another agent. They do not give an explicit formal theory of how complex intentional states are formed by the process of communication. The reason is that they do not explicate the conventional meaning of the speech act and how that is related to planning and intention formation. No systematic theory of the semantics or pragmatics of a language fragment is developed.

Appelt [85] does implicitly describe the information state I by Know and Belief operators. Similarly, the intentional state S , described below, is implicitly described by an Intends operator. However, there is no explicit formal theory of these structures given. Grosz [85] takes an important step in this direction when she clearly reconizes these structures for discourse theory. She does not make any attempt at formalization.

To sum up, in none of the above studies is a formal computational theory given as to how it is possible to communicate incrementally, to tailor and adjust plan communication to fit an uncertain world of changing circumstances. Therefore, no complex communication of strategic information is possible. In this paper we extend the investigation to complex communication between agents in a high-level language. This makes possible the coordination of arbitrarily complex social activity. We begin with some conceptual preliminaries.

4. Situations And Actions

Let IND be the set of **individuals**, R the set of n-ary relations on IND, for $n \geq 0$. Let T be the set of all times ordered by a linear relation $<$. Let TP be the set of **time periods** over T [see Allen 84]. In context, we will use t to represent either instants or time periods. Let s be a **situation** at a given instant. A situation is a partial description of the state of the world. Situations are defined in terms of IND and R [Barwise and Perry 83, McCarthy and Hayes 69]. Let **Sit** be the set of all possible situations. An **event** e is a partial function from the set of times into the set of possible situations, $e : T \Rightarrow Sit$. Let EVENTS be the set of all possible events.

Let a **world state** σ be a total description of the state of the world at a given instant. Hence a world state will be a totally defined situation. Σ is the set of all possible world states. Let H be a **possible history** of the world over time T. H will be a total function from the set of times T into the set of possible states Σ, $H : T \Rightarrow \Sigma$. Let Ω be the set of all possible histories. A given history **realizes** an event e over period $\iota \varepsilon TP$ iff Domain (e) $= \tau$ and for each time t ε Domain (e), $e_t \subseteq H_t$. Let ParHist(Ω) be the set of all possible **partial histories** H^t where $H \varepsilon \Omega$.

Actions will be special kinds of events. A simple action a has special roles associated with it, namely, that of agent and object. Let ACT $\subseteq$ EVENTS be the set of all possible actions. An **action** may be viewed as an ordered pair $a = < p, e >$, where p is in the role of agent and e is an event generated by that agent. An action a is **realized** in a world history H if the event e is realized in H and p performs e in H. Note that our formalism allows simultaneous actions because actions and events are not functions on possible states. Rather events are realized in relation to a sequence of world states, i.e., a world history. An event e thus generates a class e* of all world histories that realize e [compare Georgeff 86].

5. Knowledge Representation

5.1. Two Kinds of Uncertainty

To motivate the development that follows we distinguish two kinds of uncertainty. Normal human action as well as robot action occurs in the context of the agent being uncertain about the exact state of the world [Brooks 82]. For example, a robot may not know exactly where an object is. We will call this **state uncertainty**. An agent may also be uncertain about how to do something or about what some other agent will do. For example a robot may not know how to open a bottle, or robot A may not know exactly where robot B will go. We will call this **process**

uncertainty. This distinction is an epistemological categorization of the nature of knowledge. State uncertainty is reduced by perception [Brooks 82] and by the communication of state information. Process uncertainty is reduced by search and by the communication of process information [Werner].

5.2. Information States

Agents act in the context of having knowledge about their world. Without sufficient knowledge of the state of the world, action would be impossible. In fact, strategies for action only exist given sufficient state information. Actions have informational preconditions [Moore 80].

Formally, we represent the agents' state information by an information set $I \subseteq ParHist(\Omega)$. I is, thus, a set of partial histories H^t. If $H^t \varepsilon I$ then it means relative to the information available to the agent, H^t is a possible history at time t. We will refer to I_A as the agent A's **information state**. An information state I is the set-theoretic analogue of 'world conditions' in the situated-automate approach [see Rosenschein S.J. 86]. Let I^* be the set of all $H \varepsilon \Omega$ such that there is an $H^t \varepsilon I$ and H^t is a partial history of H. I^* is the set of histories allowed by the information I. With each information set I we associate a set of **alternatives** Alt (I). Alternatives are the choices available to the agent given the information I. The greater the information the more refined the alternatives and the greater is the number of strategies that force specific goals.

5.3. Intentional States

A **strategy** π is a function from information states I to the alternatives at I. With any given strategy π we associate a set π^*, called the **potential** of π, of all worlds $H \varepsilon \Omega$ where H is a possible outcome of π. Intuitively, the Set π^* is a set of all world histories that are consistent with the strategy π. Thus $H \varepsilon \pi^*$ if H is a possible history given π. An **intentional state** S_A of an agent A is a set of strategies $\pi \varepsilon S_A$ consisting of all those strategies that are consistent with A's plans and intentions S represents total intentional state of the agent. These are the strategies actually governing the agents actions. Which strategies actually apply depends on the actual information I that is available to the agent. Some of the strategies in S will be information gathering strategies. Intentional states will include action strategies, linguistic strategies LS, as well as cognitive strategies.

5.4. Representational States

The **representational state** of an agent can thus be characterized by $R = <I, S, V>$. We include V for the sake of completeness. It represents the agent's **evaluation** of situations. The representational state R_A may include the agent A's representation of B's representation, $R^B{}_A$. It may also include the agent A's representation of B's representation of A's representation, $R^{BA}{}_A$. Thus we can represent arbitrary levels of nesting of representations.

6. Communication

We assume our agents communicate in a high-level language such as English. Let L be a fragment of some high-level language. We distinguish two basic types of speech acts in L, directives and informatives. **Directives** are used to change the intentional state of another agent. **Informatives** are used to change the information state of another agent. Directives will include commands, demands, and requests, including requests for linguistic action, e.g., questions. Informatives will include assertions about the state of the world.

To illustrate our theory, we will now provide a very simple formal language and give the syntax, semantics, and pragmatics for that fragment. We will then show how it can be used by agents in order to communicate state information and process information so that they can cooperate.

6.1. Syntax

The language L_{p_t} will include logical and temporal connectives: $\wedge$ (=and), $\vee$ (= or), $\neg$ (=not), $\wedge^{\Rightarrow}$ (= and then), while (= while)

1. **Atomic Formulas** We distinguish two kinds of atomic sentences, **pure informatives** p, q and **pure directives** p!, q!. Directives will be indicated by adding an exclamation point. Note, in general, p and p! are distinct formulas that need have no relationship to one another.

2. **Complex Formulas** The formulas of our language will contain all and only those informatives or directives that satisfy the following conditions:

(a) **Informatives** If α and β are informatives then $\alpha \wedge \beta$, $\alpha \vee \beta$, $\neg \alpha$, $\alpha \wedge^{\Rightarrow} \beta$, and α while β are informatives.

(b) **Directives** If $\alpha!$ and $\beta!$ are directives then the following are also directives: $\alpha! \wedge \beta!$, $\alpha! \vee \beta!$, $\neg \alpha!$, $\alpha! \wedge^{\Rightarrow} \beta!$, and $\alpha!$ while $\beta!$.

6.2. Referential Semantics

With each atomic formula p we associate a referential component of its meaning, REF(p).

For Informatives: REF(p) $\subseteq$ EVENTS. The referential component of an informative will be an event type or class of events.

For Directives: REF(p!) $\subseteq$ ACT. The referential component of a directive will be an action type or a class of actions.

Below we will make use of a meta-predicate Holds, it is defined in the usual way by induction on the structure of the formulas of the language L . For atomic formulas, Holds (p, H, t) iff exists an e ε REF(p) such that e is realized in H at t where t ε TP. Given that Θ is either an informative or directive then **Holds** (Θ, H , t) means that the event referred to by Θ is realized in the history H relative to the time period t.

Intuitively, the referential component is the set of events or actions of some type to which the sentences α or $\alpha!$ refer. For example, "Open the door!" refers to the action of opening the door as does "Bill opened the door". In the former case, no agent is specified so the class of possible situations where there is an opening of the door, includes all possible agents who can stand in the role of opening the door. Contextual information in the discourse situation will, in general, reduce this set.

6.3. Pragmatics

While the referential semantics formally describes the type of situation the sentence is about, the pragmatics, in our sense of the term, formally describes how the sentence affects the knowledge representational state of the agent given the sentence is accepted as valid and appropriate. To avoid possible confusion, we use the term **pragmatics** in the original sense that Morris [38] used it when he first made the distinction between syntax , semantics,. and pragmatics. According to Morris semantics describes the relationships between language and the world, while pragmatics includes the relation of language to the speaking and understanding subject. Since our theory of meaning includes the state of information and intention of the agent, we use the term pragmatics to refer to this theory of meaning. The **pragmatic interpretation**, Prag, thereby is an account of the conventional meaning of the sentence as understood by an agent in terms of his representation of his and others' intentions and world knowledge. Informatives will transform the information state I of the agent while directives will transform the intentional state S of the agent. The more complex the language fragment the more complex the structure of Prag will be.

I. **For Informatives**

1. For atomic formulas p, Prag(p) : INF $\Rightarrow$ INF

Prag(p) is an operator on the class of possible information states INF of the agent. It takes an information state and gives another information state where Prag(p)(I) = $\{H^t : H^t \varepsilon$ I and exists an e ε REF(p) such that e is realized in H$\}$. Using standard operator notation we abreviate this to Prag(p)(I) = Prag(p) I.

2. Given Prag is defined for the formulas α and β:

$\mathrm{Prag}(\alpha \wedge \beta)\ I = \mathrm{Prag}(\alpha)\ I \cap \mathrm{Prag}(\beta)\ I$

$\mathrm{Prag}(\alpha \vee \beta)\ I = \mathrm{Prag}(\alpha)\ I \cup \mathrm{Prag}(\beta)\ I$

$\mathrm{Prag}(\neg \alpha)\ I = I - \mathrm{Prag}(\alpha)\ I$

Prag ($\alpha \wedge\!\!\Rightarrow \beta$) I = $\{H^t : H^t \varepsilon$ I and there exist times t_0, t' ε TP where Holds(α, H, t_0) and Holds(β, H, t') and $t_0 < t'\}$

Prag (α while β) I = $\{H^t : H^t \varepsilon$ I and for all t_0, t' ε TP , if t_0 contains t' then if Holds(β, H, t') then Holds(α, H, t_0)$\}$

For example, the pragmatic interpretation of the sentence α = 'Jon opened the door' is arrived at as follows: REF(α) is the event of Jon opening the door. Prag(α) is an operator on the hearer's information state I such that Prag(α)I is the reduction of the set I to those histories where the event referred to by α occurred. The hearer A knows α if α holds in all the worlds in I . Thus, A comes to know that α as a result of receiving and interpreting the message α . This semantics of know is similar to that used by Appelt and Hintikka, however, the idea is much older. It goes back at least as far as Boltzmann in his work on the statistical foundations of the second law of thermodynamics and is later used by von Neumann in his mathematical theory of games.

Note that Prag describes the **pragmatic competence** of an ideal speaker and not the actual performance. He may for various reasons not accept the message. But for him to understand the assertion or directive, the conversational participant must know what the effect of the message is supposed to be if he were to accept it. Thus, a participant will not just have an actual informational and intentional state I and S but also hypothetical representational states HI and HS that are used to compute the pragmatic effect of a given message. If the participant then accepts the message, HI or HS will become a part of the actual representational state $R = (I, S, V)$.

II. **For Directives**

1. For atomic directives p!, Prag(p!) : $INT \Rightarrow INT$

The pragmatic interpretation Prag for directives is an operator on the class of all possible intentional states INT of an agent. Prag(p!) transforms intentional states to produce a new intentional state where the agent intends to do p!. Specifically, given intentional state S, Prag(p!) S = { π : forces(π , p!)}. Where **forces**(π , p!) iff for all H ε π*, there is an action a ε REF (p!) such that a is realized in H. For a more complex language forces would be relativized to the discourse situation d and the information state I_d of the agent in the discourse situation.

2. Given Prag is defined for directives α! and β!,

$\text{Prag}(\alpha! \wedge \beta!)\ S = \text{Prag}(\alpha!)\ S \cap \text{Prag}(\beta!)\ S$

$\text{Prag}(\alpha! \vee \beta!)\ S = \text{Prag}(\alpha!)\ S \cup \text{Prag}(\beta!)\ S$

$\text{Prag}(\neg \alpha!)\ S = S - \text{Prag}(\alpha!)\ S$

Prag $(\alpha! \wedge\!\Rightarrow \beta!)$ S = { π : for all H ε π* and there exist times t_0, t' ε TP where Holds(α!, H, t_0) and Holds(β!, H, t') and $t_0 < t'$}

Prag (α! while β!) = { π : for all H ε π*, exists, t' ε TP such that Holds(α!, H, t) and Holds(β!, H, t') and t' contains t}. Note that 'while' is surrounded by two directives. It means that the actions are to be done in parallel.

For example, if α! = 'Open the door !' , REF(α!) refers to the situation of the addressee A opening the door. Prag(α!) operates on A's intentional state S_A such that A opens the door. Prag does this by removing all those possible plans of A that do not force α!. And those are the plans π that have some world H ε π* where the situation referred to by α! is not realized in H. The result is that the agent performs the directive no matter what other goals he may have. Again, we are talking about the ideal pragmatic competence. Another example: "I am going to the bank". This informs the hearer of the intentions of the speaker. It is either a self-directive that updates the intentional state of the speaker (see § 7) as it is being said, or it reports the speaker's existing intentions. In either case, it fits within our semantic theory since our theory of meaning directly quantifies over intentional states. Note, even if it is a report it still updates the hearer's representation $S_h{}^s$ of the speaker's intentions.

The pragmatic theory of meaning is compositional in that the meaning of the whole is systematically related to the meaning of the parts. Prag distributes over the propositional structure of the sentence. This is as it should be since it is the key property that allows us to interpret abitrarily complex directives. For example, the command "Get the money from the bank and then go to the airport!" should be given a meaning that is composed of the meanings of the individual conjuncts and the interpretation of "and then" should relate these two meanings. This is precisely what our theory does.

One might object and say that Prag is too abstract. It seems to say no more than 'a message transforms the representational state of a conversational participant in a way that is compatible with the compositional meaning of the message.' A similar objection would apply to the semantic efforts of Tarski, Montague [74], and Barwise [83]. One might object and say that they are saying no more than 'a sentence is true if it is true' or 'the meaning is given by the truth conditions' or 'the meaning of a message is what it refers to'. The point is that these statements are conditions on the semantic enterprise. The detailed compositional structure of these theories is what gives them their power and their potential usefulness in the design and building of complex programs that understand dialogue. Previous semantic theories have been restricted to assertions. Prag extends the semantics to include nonassertive speech acts (e.g., commands, requests, statements of intention, ect.). They, as it turns out, are the speech acts most useful for understanding social action.

7. Speech Act Theory

7.1 Overview

Our work provides a theoretical framework that gives a systematic account of the conventional meaning of speech acts. The conventional meaning of the speech act consists of two parts: its referential component and its force [Searle 69]. The referential component is given by a situation semantics. The force is defined in terms of the pragmatic interpretation. The force of the speech act determines which type of representation is to be transformed by the operator Prag(α). Thus we are able to give an explicit theory of the force of speech acts.

7.2. The Force of Illocutionary Acts

What distinguishes a request from an assertion? One answer is that their force is different. But what is force? According to Searle, when humans communicate they are engaged in an activity. Each portion of the communication is an action. An utterance, according to Searle, can be broken down into two basic components, the **illocutionary force** F and the **propositional content** p. The utterance is symbolized as F(p). In order to classify the different types of force F. Searle and Vanderveken attempt to reduce the force of a speech act to more primitive features (e.g., the point, the direction of fit, ect.). The force and the propositional content is then used to divide speech acts into six general classes. For details see Searle and Vanderveken [85].

Searle's attempt to define force is inadequate because some of the dimensions are redundant. The point, for example, has no classificatory function since there is a one to one correspondence between the speech act type and the 'point' feature. The features are also vague and of questionable computational usefulness. Behind these problems lies a more devastating problem: As Searle and Vanderveken admit, they have no semantics for the two most central features in the definition of force, namely, the point and direction of fit of the speech act. Instead, they leave these notions primitive and unanalyzed. That, however, amounts to leaving the notion of force an unanalyzed concept. A proper theory of force requires a theory of intention. Since we have outlined such a theory, we can use it to formally define the force of a speech act.

7.3 Speech Acts in Communication

When people use language to communicate they do so to get things done. That is why utterances have the effect of actions. But the reason that utterances have the effect they do is because they influence the cognitive state of the conversants. It is the harmony of the cognitive states of agents that makes possible cooperative social action and forms the basis of society.

On our view the meaning of the speech act is best understood if we understand how the speech act is meant to influence the cognitive states of the conversants. The force of a speech act lies in its unique distribution of effect on the cognitive substates of the conversants. A directive, for example, is meant to change the intentional state of the recipient in such a way that the recipient will perform the actions referred to by the propositional content of the directive. The assertive is meant to influence the informational state of the addressee.

One objection to our view may be that the theory of how a speech act effects the hearer is the study of perlocutionary effect. The perlocutionary effect is subject to the idiosyncrasies of individual performance and understanding and, therefore, cannot be the meaning of the speech act. We think differently. One must make a distinction between the ideal cognitive competence of the understanding subject (i.e., the ability of the subject to understand the speech act) and the actual cognitive performance. The meaning of a speech act is described by how it is to effect the ideal cognitive state of the conversants, given that the message is accepted. (see Perrault [87] for a similar view)

7.4. A Pragmatics of Speech Acts

We now give a semantic, pragmatic description of some of the speech acts in Searle's taxonomy. First some needed definitions. In what follows we abstract from pure informatives and pure directives and allow that any given utterance α will have both directive and informative content. Prag(α) will thus be defined on I, S, and V Let $<s, \alpha, h>$ be a speech act where s is the speaker, α is the sentence expressed and h is the hearer in the discourse situation d. Let the speaker s have representational state $R_s = (I_s, S_s, V_s)$ and the hearer h have representational state $R_h = (I_h, S_h, V_h)$. The different kinds of speech acts can be differentiated by how they effect the cognitive state of the conversants. Specifically, the **force of a speech act** is the set of subrepresentations in R that are to be transformed by the speech act. An information state I **forces** α, in symbols, $I \Vdash\!\Rightarrow \alpha$ iff for all $H \varepsilon I^*$, α holds in H. An intentional state S **forces** α, in symbols, $S \Vdash\!\Rightarrow \alpha$ iff for all $\pi \varepsilon S$, π forces α, i.e., iff for all $H \varepsilon \pi^*$, α holds in H. Below we will use the shorthand notation of αI for Prag(α)(I).

1. **Assertives**: Example: "Bill opened the door."

   1.1. $I_h \ {}_{\alpha}\!\Rightarrow \alpha I_h$ 1 2. $I_s{}^h \ {}_{\alpha}\!\Rightarrow \alpha I_s{}^h$

Remark: I_h transforms to αI_h. Assertives effect the informational state of the hearer. They also effect the hearer's representation of the speaker's beliefs..

2. **Directives**: Example: "Open the door!"

   2.1. $S_h \ {}_{\alpha}\!\Rightarrow \alpha S_h$ 2.2. $S_s{}^h \ {}_{\alpha}\!\Rightarrow \alpha S_s{}^h$

Remark: The command updates the hearer's intentions to αS_h where h does the action α. $\alpha S_s{}^h$ describes the speaker's representation of the hearer's new intentions.

3. **Commissives**: Example: "I will open the door."

   3.1. $S_s \ {}_{\alpha}\!\Rightarrow \alpha S_s$ 3.2. $S_h{}^s \ {}_{\alpha}\!\Rightarrow \alpha S_h{}^s$

Remark: The speaker commits himself to following those strategies that insure the propositional content of α, i.e., all the worlds in each π^* realize the action referred to by α. $\alpha S_h{}^s$ represents the hearer's resulting representation of the speaker's modified intentions.

4. **Declarations**: Example: "I resign.", "You're fired."

   4.1. $I_h \ {}_{\alpha}\!\Rightarrow \alpha I_h$ 4.2. $I_s \ {}_{\alpha}\!\Rightarrow \alpha I_s$ 4.3. $S_s \ {}_{\alpha}\!\Rightarrow \alpha S_s$
   4.4. $S_h \ {}_{\alpha}\!\Rightarrow \alpha S_h$ 4.5. $S_{institution} \ {}_{\alpha}\!\Rightarrow \alpha S_{institution}$

Remark: Both the hearer and speaker update their information states to αI_h and αI_s, respectively, where they know the resulting state brought on by the declaration. Furthermore, a declaration such as "you're fired" has specific intentional consequences such as no longer being paid. $\alpha S_{institution}$ indicates that the declaration also has institutional effects. Namely, it effects the composite intentions of all those with roles involved in the employment relationship.

5. **Representative Declaratives:** Example: "I find you guilty"

5.1.	$I_h \ {}_\alpha\Rightarrow \ \alpha I_h$	5.2.	$I_s \ {}_\alpha\Rightarrow \ \alpha I_s$
5.3.	$S_s \ {}_\alpha\Rightarrow \ \alpha S_s$	5.4.	$S_h \ {}_\alpha\Rightarrow \ \alpha S_h$
5.5.	$S_{institution} \ {}_\alpha\Rightarrow \ \alpha S_{institution}$	5.6.	$I_s \ \Vert\!\Rightarrow \ \alpha$

Remark: The representative declarative differs from the declaration in that the former must be based on certain facts obtaining. $I_s \ \Vert\!\Rightarrow \ \alpha$ expresses this condition. Again we see how social roles in an institution are affected by a declaration. The judge's declaration of guilt and sentencing has very specific intentional consequences for the police and parole board, etc. These complex intentions are packed into the composite institutional role structure $\alpha S_{institution}$. What is so interesting is that our theory allows us to talk about such complex social processes. It takes a small step toward a better understanding of the relationship between linguistic communication and social structure. It is this property of our theory that makes it a promising candidate for the design of the complex systems being contemplated in distributed artificial intelligence.

We have developed the outlines of a formal theory of meaning (semantics and pragmatics) of speech acts. We have used this theory to give a definition of illocutionary force in terms of the specific subrepresentations that the speech act is to modify. The subrepresentations are only sketched. But the point of the approach is quite clear. The cognitive states of the conversational participants, for example, system and user, play a dominant role in the theory of meaning and force of speech acts. An actual implementation of an algorithm for Prag and an actual knowledge representation scheme to describe the information, intentional, and evaluative states requires making significantly more detailed system design decisions. We have aimed at providing a general theoretical framework for designing systems with a communicative competence using natural language.

Our work, while distinct in its aims, is compatible with the work in speech act planning [Cohen 78, Cohen and Perrault 78, Allen 79], and discourse [Grosz 85], as that work is at the level of describing the speaker's linguistic intentional states. In fact, our work provides unifying theoretical context for that work in that the planning of speech acts and, more generally, discourse is part of the intentional component of the representational knowledge state.

7.5. Conversational Strategies

We generalize the notion of a speech act to a conversational act or strategy which once learned can be invoked to achieve certain categories of goals. We hypothesize that real conversation is planned using whole linguistic action strategies or linguistic modules rather than individual speech acts. Within a module or a communicational frame individual surface speech acts may involve further planning. The available linguistic strategies are a specialized sub-representation: The linguistic intentions LS within the overall intentional state S. It should be noted that scripts are just a particular instance of our more general conversational action strategy, for a script is just a well-structured plan. And a plan is simply a partial strategy.

8. Social Cooperation and Communication

Social action is made possible by the communication of state information and process information. State information is relayed by informative speech acts. Process information is relayed by directive speech acts. The social act, abstractly viewed, results from the composition of the agents' strategies.

Intuitively, at the lowest level, the use of directives by an agent to control another can be viewed as a form of incremental plan passing. The plan is passed by messages that in effect are a coding for the construction of a plan or more generally a strategy. The recipient if he understands the conventional meaning of the message interprets the directive $\alpha!$ as a partial strategy. We can view the pragmatic effect of $\alpha!$ as either a reduction of the possible plans that guide actions of the agent, i.e., the set S or, equivalently, as building up the intentional state of the agent.

Informatives are a way to pass state information and help to achieve a goal by either fulfilling the informational preconditions of an action required by a strategy or by acting as a form of indirect speech act [Allen 79] where the sender gives information that the recipient uses to rationally deduce what the sender wants [see §3]. Once interpreted the indirect speech act pragmatically acts like a directive that sets up the intentional state of the recipient. We now present a slightly more formal account of cooperation.

Social action demands different levels of communicational complexity and structure. The simplest case is a **master-slave relationship** with one-way communication. One agent A uses a directive $\alpha!$ to control the actions of the recipient B. It works because the high-level message is given a pragmatic interpretation $Prag(\alpha!)$, which operates on the intentional state S_B in such a way that $Prag(\alpha!)S_B$ forces the desired goal, i.e., $Prag(\alpha!)S_B \Vdash\!\Rightarrow g_A$. An intentional state S **forces a goal** g, in symbols, $S \Vdash\!\Rightarrow g$ iff for all $\pi \in S$, π forces g, i.e., iff for all $H \in \pi^*$, g is realized in H. A may also communicate state information β to B to fulfil informational preconditions required by a strategy or to perform an indirect speech act.

More complex is the case of **one way cooperation** where A communicates $\alpha!$ to B so that $Prag(\alpha!)S_B + S_A \Vdash\!\Rightarrow g_A$. By definition the composite S + S' of two intentional states S, S' together force a goal g, in symbols, $S + S' \Vdash\!\Rightarrow g$ iff for all $\pi \in S$, $\pi_O \in S'$, and for all $H \in \pi^* \cap \pi_O^*$, g is realized in H. In other words, A sets up B's intentions so that when combined with A's intentions, their actions together achieve A's goal g_A.

Still more complex is the case of **mutual cooperation** where A and B have a mutual exchange of directives and informatives before proceeding to act. The mutual exchange results in a conversational history $h_\Theta = \Theta_1, \ldots \Theta_n$ where each Θ_i is either a directive or an informative speech act that includes information about the speaker and addressee in the discourse situation d. The pragmatically interpreted conversation $Prag(h_\Theta) = Prag(\Theta_1)Prag(\Theta_2) \ldots Prag(\Theta_n)$ then results in the mutual goal $g_{A,B}$, i.e., $Prag(h_\Theta)R_B + Prag(h_\Theta)R_A \Vdash\!\Rightarrow g_{A,B}$.

Sophisticated and permanent **societal cooperation** is made possible by the formation of social structures. A **social structure** can be viewed as a set of social roles $\mathbf{rol_1}, \ldots, \mathbf{rol_n}$, in a given environment Ω. Roughly, each **social role**, **rol**, is an abstract description of an agent $R_{\mathbf{rol}} = \langle I_{\mathbf{rol}}, S_{\mathbf{rol}}, V_{\mathbf{rol}} \rangle$ that defines the state information, permissions, responsibilities, and values of that agent role. When an actual agent A assumes a role **rol**, he internalizes that role by

constraining his representational state R_A to $R_A + R_{rol}$. A social structure may have implicit and codified **laws** that further define the intentional states of the agents as well as the roles of the social structure. These laws have their effect by acting on the intentional states of the agents. The society generated by the social structure functions because its agents take on the social roles that achieve the societal goals $g_{society}$. The roles and laws are such that
$R_{rol_1} + \ldots + R_{rol_n} \| \Rightarrow g_{society}$.

9. Conclusion

We have developed a theory of linguistic communication that explains social cooperation. We did this by developing a formal account of the agent's knowledge states, specifically his or her intentional states. The pragmatic interpretation also enabled us to give an account of the force of the speech act. The pragmatic interpretation links the linguistic message with its effect on the planning process as defined by the intentional state. It becomes possible to build up intentional states of unlimited complexity. This allowed us to give an account of social cooperative action because the intentional states of the agents are mutually modified by a communicative exchange, i.e., a conversation or discourse. The intentional states are thereby set up in such a way so that the social goal is achievable. We hinted at a clarification of the complex relationship between language and society made possible by our communication theory.

References

Allen, J. F., "Towards a General Theory of Action and Time," ARTIFICIAL INTELLIGENCE, 23, pp. 123 - 154, 1984.

Allen, J. F., "A Plan-Based Approach to Speech Act Recognition," Thesis, Department of Computer Science, University of Toronto, 1979.

Allen, J. F. and Perrault, C. R., "Analyzing Intention in Utterances," ARTIFICIAL INTELLIGENCE, 15, pp. 143 - 178, 1980.

Appelt, D. E., PLANNING ENGLISH SENTENCES, Cambridge University Press, New York, 1985.

Barwise, J., and Perry, J., SITUATIONS AND ATTITUDES, Bradford Books/MIT Press, 1983.

Brooks, R.A., "Symbolic Error Analysis and Robot Planning," INTERNATIONAL JOURNAL OF ROBOTICS RESEARCH,1, No. 4, pp. 29 - 68, 1982.

Cohen, P. R., "On Knowing What to Say: Planning Speech Acts," Techn. Rep. 118, Department of Computer Science, University of Toronto, 1978.

Cohen, P. R., and Perrault, C. R., "Elements of a Plan-Based Theory of Speech Acts," COGNITIVE SCIENCE, 3, pp. 177 - 212, 1979.

Dijkstra, E.W. "Cooperating SequentiaL Processes," in F. Genuys (ed), PROGRAMMING LANGUAGES. Academic Press, New York, 1968.

Fagin, R., Halpern J. Y., and Moshe, Y. V., "What Can Machines Know? On the Epistemic Properties of Machines," Proc. AAAI-86, pp. 428 - 434, Philadelphia, PA, 1986.

Fikes, R.E., and Nilsson, N. J., "STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving," ARTIFICIAL INTELLIGENCE, 2, pp. 189 - 208, 1971.

Genesereth, M. R., Ginsberg, M. L., and Rosenchein, J. S., "Cooperation without

Communication," Proc. AAAI-86, pp. 561 - 57, 1986.
Georgeff, Michael, "Communication and Interaction in Multi-agent Planning," Proc. AAAI-83, pp. 125 - 129, 1983.
Georgeff, M. P., "The Representation of events in Multiagent Domains," Proc. AAAI-86, pp. 70 - 75, Philadelphia, PA, 1986.
Grosz, B. J., "The Structures of Discourse Structure," Techn. Note 369, Artificial Intelligence Center, SRI International, Menlo Park, California, 1985.
Hewitt, C., "Control Structures as Patterns of Passing Messages,"ARTIFICIAL INTELLIGENCE, 8, pp. 323 - 363, 1977.
Hoare, C. A. R., "Communicating Sequential Processes," Comm. ACM, 21, pp. 666 - 677, 1978.
Lozano-Rérez, T., "Robot Programming," Proc. IEEE, 71, No. 7, pp. 821 - 841, 1983.
Konolige, K., "A First-Order Formalization of Knowledge and Action for a Multiagent Planning System,"Techn. Note 232, Artificial Intelligence Center, SRI International, Menlo Park, California, 1980.
McCarthy, J., and Hayes, P., "Some Philosophical Problems from the Standpoint of Artificial Intelligence," in B. Meltzer and D. Michie (editors), MACHINE INTELLIGENCE; 4, 1969.
McDermott, D., "A Temporal Logic for Reasoning about Processes and Plans," COGNITIVE SCIENCE, 6, pp. 101 - 155, 1982.
Montague, R., "The Proper Treatment of Quantification in Ordinary English", In Thomason, R., (ed.), FORMAL PHILOSOPHY: Selelcted Papers of Richard Montague, NewHaven: Yale University Press, pp. 247-270, 1974.
Moore, R. C., "Reasoning About Knowledge and Action", Tech. Note 191, Artificial Intelligence Center, SRI International, Menlo Park, California, 1980.
Morris, C. W., "Foundations of the theory of Signs", INTERNATIONAL ENCYCLOPAEDIA OF UNIFIED SCIENCES, Neurath, Carnap & Morris, (eds.), pp. 79-137, 1938
Pednault, E. P. D., "Preliminary Report on a Theory of Plan Synthesis,"Techn Note 358, Artificial Intelligence Center, SRI International, Menlo Park, California, 1985.
Perrault, C. R., and Allen, J. F., "A Plan-Based Analysis of Indirect Speech Acts," AMERICAN JOURNAL OF COMPUTATIONAL LINGUISTICS, 6, # 3 - 4, 1980.
Rosenschein, Jeffrey S., "Rational Interaction: Cooperation Among Intelligent Agents," Ph.D. Thesis, Stanford University, 1986.
Rosenschein, Stanley, J., "Formal Theories of Knowledge in AI and Robotics,"Techn. Note 362, Artificial Intelligence Center, SRI International, Menlo Park, California, 1986.
Sacerdoti, E. D., A STRUCTURE FOR PLANS AND BEHAVIOUR, Elsevier North-Holland, Inc., New York, 1977.
Searle, J. R., SPEECH ACTS: AN ESSAY IN THE PHILOSOPHY OF LANGUAGE, Cambridge University Press, London, 1969.
Werner, E., "Uncertainty, Search and Heuristic Information", manuscript, Department of Computer Science, Bowdoin College, Brunswick, Maine