PROPAGATING EPISTEMIC COORDINATION THROUGH MUTUAL DEFAULTS I

Richmond H. Thomason
Intelligent Systems Program
University of Pittsburgh
Pittsburgh, PA 15260
U.S.A.

ABSTRACT

A *mutual default* is a rule, capable of tolerating exceptions, that is mutually supposed by a group G: i.e., the rule is supposed by all members of the group, is supposed by all members of the group to be supposed by all members of the group, etc. A family of propositional attitudes B_i indexed for $i \in G$ (and representing, say, supposition) is *coordinated* for G if B_i applies to the same propositions for all members i of G, and is commonly supposed by the members of G to do so.

This paper is a preliminary exploration of formal postulates that ensure maintenance of coordination of a propositional attitude, representing the common ground of a conversation, in dynamic environments that allow for assertional speech acts. I present results showing that mutually supposed rules of conversation provide a mechanism for preserving coordination. If coordination can be assumed, reasoning about propositional attitudes can be greatly simplified, through collapse of iterated operators.

I also show how coordination maintenance can be secured, at least in unexceptional cases, when rules of conversation are defeasible; this relaxation of the theory is needed because plausible conversational rules are subject to exceptions.

The project of formalizing coordination maintenance using Circumscription Theory raises some interesting technical problems; to provide a finitely axiomatized theory of coordination, it is apparently necessary to quantify over intensional types of at least third order. For this purpose, Richard Montague's Intensional Logic seems to be an appropriate vehicle.

1. SHARED CONVERSATIONAL CONTEXT

Planning is certainly not all there is to discourse. But planning and plan recognition are among the most important reasoning mechanisms that come into play when language is used. Through the work of a research community including James Allen, Philip Cohen, Hector Levesque, C. Raymond Perrault, and others,[1] plan-based models of speech acts and discourse have been developed to a very sophisticated level, achieving improvements in our understanding of the phenomena and in guiding the implementation of natural language understanding systems.

Certain goals—speech act goals—play a central part in this approach, both in generating and in interpreting discourse. Allen, Cohen, Levesque, Perrault, and in fact all computational linguists who have written on this topic and whose work I am familiar with, consider the goal of an act of assertion to be the creation of a kind of belief. (For instance, in telling you that your shoelace is

[1] Some references to this work are provided in the bibliography; I haven't tried to supply a complete list.

untied I am trying to get you to believe that your shoelace is untied.) It is also generally agreed that the goal of an assertive speech act should be not merely a hearer's belief, but a *mutual* belief. (My assertive speech act will not have been successful, for instance, if you came to believe that your shoelace is untied, but I didn't also come to believe that you now believe this.) This conclusion has become generally accepted on its own merit, it seems, and also through the weight of psychological authorities like Clark and Marshall [4], and of philosophical ones like Lewis [7] and Schiffer [11].

However, it is not easy to see how to formalize this goal appropriately, if we assume that conversational agents are reasonably well informed, and are sane in the sense that they reject goals that a reasonable person would consider impracticable. The problem is that *mutual* belief involves iterated combinations of the beliefs of several agents, and even when there are only two agents, these iterations are in principle unbounded, so that a mutual belief will comprise a "conjunction" of infinitely many independent beliefs. Adding a mutual belief operator to the planning language, or even infinitary conjunctions to it, as Allen seems to do in [1], seems problematic even aside from computational considerations, if we want speech act goals to be practicable. How can I reasonably intend to create a mutual belief, any more than I can reasonably intend to write down all the natural numbers?

In Cohen and Levesque [5], this problem is dealt with by relaxing the goal. Assertive acts, according to this paper, aim not at a mutual belief in the asserted proposition, but at a hearer's belief that a certain complex belief is mutually believed. But this only seems to postpone the practicability problem: if mutual belief can't be achieved, for more or less a priori reasons, then it's hard to see why the ideal hearer shouldn't recognize the very same problem that led Cohen and Levesque to settle for less. Also the ideal speaker should realize that this is how the hearer will reason. And so, if the speaker is careful he shouldn't expect that his act will induce the hearer to believe that a genuine mutual belief has been achieved.

Reacting not to this difficulty, but to problems that he traces to the indefeasible character of Cohen and Levesque's rules, Perrault not only developed, in [10], a nonmonotonic formulation of discourse rules, but also provided a reflective mechanism for restoring mutual belief as the goal state. Perrault shows that mutual belief in the asserted proposition can be obtained by strengthening the rules of conversational dynamics so that they themselves will be mutually believed. A later reformulation of Perrault's idea in Hierarchic Autoepistemic Logic was implemented in a natural language understanding system; see Appelt and Konolige [2].

Since, like many other practical rules, discourse rules are defeasible, it is good to see how to state them in one or more of the formalisms for nonmonotonic reasoning. But Perrault's solution to the feasibility of creating mutual beliefs, by shifting the burden to mutual discourse rules and to the underlying logic of belief, seems to me to be potentially even more important, and in this paper I want to concentrate on that idea, extending and developing it.

2. THE IDEA OF COORDINATION

A slightly different model of discourse, originating, I think, in Robert Stalnaker's work, can be found in the philosophical literature; see Stalnaker [12] and later works, such as [13,14,15]. On this model, the data structures of conversation are *public,* or *shared.* Among these will be a belief-like propositional attitude that represents the common ground in a conversation, and that evolves as the conversation develops. On such a view, it is very natural to speak as if the common ground were the beliefs of a kind of transcendent agent, the conversational group.

Obviously, such an assumption has to be justified somehow in terms of underlying characteristics of the members of the group. I wish to propose, as the presupposition of this conception of publicity, a condition of *group coordination* of the relevant conversational information. The members of a coordinated discourse group must not only share the same beliefs, but mutually believe that they share the same beliefs.[2]

I will show here that Perrault's methods can be extended to provide for the maintenance of coordination in conversations; coordination will be preserved after an assertion has been made, as long as it is mutually believed that the assertion was not anomalous. To push this result through, I will have to install a much more powerful underlying logic. But this secures a far-reaching simplification in the formalization of discourse planning, since—as I will show—we can reason about coordinated belief in exactly the same way that we would reason about the beliefs of a single agent. All iterated operators collapse, and there is no need to distinguish the beliefs of one agent from those of another, or for that matter from those of the group itself. It remains to be seen if this idea is useful computationally. But at a theoretical plane, it certainly does seem to provide a considerable improvement in modeling discourse.

3. SUPPOSITION

Since beliefs are private, we can't take the public common ground that is created in the course of a conversation to consist of beliefs. But we can take it to consist of suppositions.

A supposition is like a belief, but may be temporary, ad hoc, and not taken seriously by the supposer. Beliefs themselves are constantly being changed, but it is less easy to change them at will. I can easily suppose (for the sake of argument, say) that the stock market will rise tomorrow, but I can't voluntarily create the corresponding belief as I please. Closely related to this involuntary feature is the action-guiding character of beliefs, which distinguishes them from mere suppositions. I will not want to buy stock because I've supposed for argument's sake that the market will rise; but I will have a motive to buy stock if I believe that it will rise.

Letting supposition replace belief in the theory of speech acts provides a single, well-motivated solution to swarms of objections and counterexamples that have been proposed in the philosophical and computational literature.[3] An early example of this sort is cited by Grice in the William James Lectures: when a student answers a question at an oral examination, he is not trying to get his history teacher to believe that, say, William invaded England in 1066. (Evidently, this and similar examples were much discussed at Oxford during the 1950's.) Appelt and Konolige's objection in [2] to the speech act theory of Perrault [10]—that a speaker should not be able to convince himself of something that he doesn't already believe simply by saying it—is another problem that Stalnaker's idea solves.

Also—a point that is important in relation to the axiomatization of discourse rules that I will offer later in this paper—the idea enables default rules to be formulated so that they can be adopted by reasonable agents. It doesn't seem right to assume—even as a default—that if a speaker asserts p the hearer will believe p. But it *does* seem plausible that if p is asserted, p will then, by default, become part of the things that are taken for granted in later stages in the conversation. Thus, by

[2]The idea of coordination isn't stressed in Stalnaker's work, though I think it is implicit there. Coordination is, of course, a major theme in Lewis' [7].

[3]This was first realized by Stalnaker. See the references in the bibliography of this paper, and especially [12].

using supposition-based defeasible rules, the theory that I will present in the final section of this paper provides a solution (though in a very simplified case), of the impracticability problem that I raised in Section 1.

Just as we can suppose things for the sake of argument, and in the extreme case can even temporarily suppose something that we know to be false in order to refute it, we can suppose things for the purposes of conversation. We want to think of these conversational suppositions as the product of a joint project, as the shared creations of a group. But to justify this, we have to show how conversational suppositions can be coordinated.

4. COORDINATED SUPPOSITION

Let B_x be an epistemic operator indexed by the individual variable x. For definiteness, I'll assume that B_x has **S5**-like properties (but without the alethic property that what is supposed must be true); for an axiomatization, see the logic **KD45** in Chellas [3].

Where G is a one-place predicate (representing a group of communicators) and Γ is a theory, B is *weakly G-coordinated* with respect to Γ if **(Coord Scheme 1)** and each instance of **(Coord Scheme 2)** is deducible from Γ.

(Coord Scheme 1) $$\forall x\, y[[G(x) \wedge G(y)] \supset \forall q[B_x(q) \equiv B_y(q)]]$$

(Coord Scheme 2) $$\forall x_1 \ldots x_n\, y\, z[[G(x_1) \wedge \ldots \wedge G(x_n) \wedge G(y) \wedge G(z)] \supset B_{x_1} \ldots B_{x_n}(\forall q[B_y(q) \equiv B_z(q)])]$$

Provided that an operator B_G is present in the language that can serve to represent common supposition for G, we also have a notion of *G-coordination*, in which **(Coord Scheme 1)** and **(Coord Scheme 2)** are replaced by the following axioms. (Note that **(Coord Scheme 1)** and **(Coord Axiom 1)** are identical.)

(Coord Axiom 1) $$\forall x\, y[[G(x) \wedge G(y)] \supset \forall q[B_x(q) \equiv B_y(q)]]$$

(Coord Axiom 2) $$\forall y\, z[[G(y) \wedge G(z)] \supset B_G \forall q[B_y(q) \equiv B_z(q)]]$$

Later we will be dealing with time-dependent supposition operators $B_{x,t}$ that are doubly indexed. In this case, we will need to speak of weak G-coordination with respect to Γ, and of G-coordination, *at t*.

5. LOGICAL RESOURCES

To deal with supposition, we need a modal logic. To deal with circumscription, we need a second order logic. But to cope with circumscriptive rules involving *mutual* supposition we will require an even more powerful logical engine.

The problem is that Circumscription Theory requires a finite axiomatization of the domain theory. So, since part of our domain is mutual supposition, we'll have to finitely axiomatize mutual supposition. To manage this, I don't see how to avoid quantifying over higher order intensional types—in particular, over sets of sets of propositions. Since I'll need to go this far, I'll simply go the whole distance, and adopt Montague's Intensional Logic as the background logic. This logic has the advantage of being very general and powerful, it has been well described,[4] and the formalism

[4]See Montague [9] and Gallin [6].

is relatively familiar, through its use in natural language semantics. I'll assume some familiarity with Intensional Logic in what follows.

The variable t is designated to range over times, which I assume to be individuals, i.e. to have type e; $p(t)$ is a proposition depending on t; this represents the assertion that is made at t.

Some type conventions: G always has type $\langle \mathsf{e}, \mathsf{t}\rangle$; p has type $\langle \mathsf{e}, \langle \mathsf{s}, \mathsf{t}\rangle\rangle$; x, a, and b all have type e; O has type $\langle\langle \mathsf{s}, \mathsf{t}\rangle, \mathsf{t}\rangle$; and q always has type $\langle \mathsf{s}, \mathsf{t}\rangle$. Other typed terms will be flagged with their type at their first occurrence in a formula. I will use '$O\mathbf{(}\phi\mathbf{)}$' as an abbreviation of Montague's '$O(\hat{}[\phi])$'; and '$O_1 \circ O_2$' as an abbreviation of '$\lambda q O_1(\hat{}O_2(q))$'.

Where η has type $\langle \mathsf{e}, \langle\langle \mathsf{s}, \mathsf{t}\rangle, \mathsf{t}\rangle\rangle$ (i.e., where η represents a propositional attitude), a *finite iteration* of η is any expression having the form $\eta_{\alpha_1} \circ \ldots \circ \eta_{\alpha_n}$, where each α_i is either a or b. (Here, 'η_a' is just an alternative notation for $\eta(a)$, a term of type $\langle\langle \mathsf{s}, \mathsf{t}\rangle, \mathsf{t}\rangle$.)

In everything that follows, I will assume for simplicity a conversational group of just two members: i.e., I'll postulate $\forall x[G(x) \equiv [x = a \vee x = b]]$.

6. WEAK COORDINATION MAINTENANCE IN A MONOTONIC SETTING

The following simple rule for conversational update serves as a good introduction to the technical issues.

(Simple Update) $\qquad \forall x t[G(x) \supset \forall q[B_{x,t+1}(q) \equiv B_{x,t}\mathbf{(}\check{}[p(t)] \supset \check{}q\mathbf{)}]]$

According to this rule, each agent updates his conversational suppositions by adding the proposition that is asserted and its supposed consequences to them.

Obviously, this rule will not maintain coordination. After an update, agents that were coordinated at t will still in fact suppose the same things as one another. But nothing guarantees that they will suppose that they have the same suppositions. The difficulty is that, though all agents share the same update rule, they needn't suppose that they share this rule. To ensure the maintenance of coordination, then, we want to close **(Simple Update)** under iterated supposition operators.[5]

Let ϕ be an instance of the scheme **(Mutual Update Scheme)** if either ϕ has the **(Simple Update)** form

$$\forall x t[G(x) \supset \forall q[B_{x,t+1}(q) \equiv B_{x,t}\mathbf{(}\check{}[p(t)] \supset \check{}q\mathbf{)}]]$$

or the form

$$\forall x t[G(x) \supset O\mathbf{(}\forall q[B_{x,t+1}(q) \equiv B_{x,t}\mathbf{(}\check{}[p(t)] \supset \check{}q\mathbf{)}]\mathbf{)}],$$

where O is a finite iteration of B_t.

Supposing that **(Mutual Update Scheme)** belongs to our background theory Γ, it is easy to show that if B is weakly G-coordinated at t with respect to Γ then B is weakly G-coordinated at $t+1$ with respect to Γ.

7. MUTUAL SUPPOSITION IN INTENSIONAL LOGIC

In this section I will sketch the development of a finite theory of mutual supposition in intensional logic, giving the bare elements only and without proving any theorems. I will take for granted the

[5] This is the idea of Perrault's that I described in Section 1.

logical axioms of Gallin [6]. The formalized theory will have two nonlogical axioms, **(MS1)** and **(MS2)**. We will need many definitions as well: these are numbered **(D1)** to **(D9)**.

The first axiom is our simplifying assumption about the size of the group.

(MS1) $$\forall x[G(x) \equiv [x = a \vee x = b]]$$

With our first definition, we plunge into the higher order aspects that enter into the formalization of mutuality.

(D1) $$\mathcal{C}(P^{\langle\langle\langle s,t\rangle,t\rangle,t\rangle}, B_t) \ =_{df} \ P(B_{a,t}) \wedge P(B_{b,t}) \wedge \forall O[P(O) \supset [P(B_{a,t} \circ O) \wedge P(B_{b,t} \circ O)]]$$

Intuitively, **(D1)** characterizes the sets of propositional attitudes that contain the attitudes expressed by $B_{a,t}$ and $B_{b,t}$, and also are closed under finite iterations of B_t. We now introduce a constant FI of type $\langle\langle\langle s,t\rangle,t\rangle,\langle e,\langle\langle\langle s,t\rangle,t\rangle,t\rangle\rangle\rangle$, which is intended to pick out the *finite* iterations of an agent-indexed propositional attitude. The intention of axiom **(MS2)** is to capture this, using the usual higher order account of finite iterations. In **(MS2)**, '$FI(B_t, O)$' abbreviates '$[FI(O)](B_t)$'.

(MS2) $$\mathcal{C}(FI, B_t) \wedge \forall X^{\langle\langle\langle s,t\rangle,t\rangle,t\rangle}[\mathcal{C}(X, B_t) \supset \forall O[FI(B_t, O) \supset X(O)]]$$

From **(MS2)** we can prove that FI in fact singles out all the finite iterations of $\mathrm{B_t}$; for instance, we can prove $FI(B_{a,t} \circ B_{b,t} \circ B_{a,t}, B_t)$.

Using FI, we now define B_G of type $\langle e, \langle\langle s,t\rangle,t\rangle\rangle$. $B_{G,t}$ represents coordinated supposition at t; $B^T_{G,t}$ represents true coordinated supposition at t.

(D2) $$B_G \ =_{df} \ \lambda t \lambda q \forall O[FI(O, B_t) \supset O(q)]$$

(D3) $$B^T_G \ =_{df} \ \lambda t \lambda q[B_{G,t}(q) \wedge {}^{\vee}q]$$

8. MUTUALLY SUPPOSED SUPPOSITION UPDATE

The next four definitions have to do with update. **(D6)** and **(D7)** characterize two fundamental notions of supposition update, represented by constants $MIASU$ and $GASU$ of type $\langle e, \langle e, t\rangle\rangle$.

(D4) $$\eta^T \ =_{df} \ \lambda q({}^{\vee}q \wedge \eta(q))$$

(D5) $$AU(O, O', t) \ =_{df} \ O' = \lambda q O({}^{\vee}[p(t)] \supset {}^{\vee}q)$$

(D6) $$MIASU(t, t') \ =_{df} \ \lambda t \lambda t' \forall x[G(x) \supset B^T_{x,t}(AU(B_{x,t}, B_{x,t'}, t))]$$

(D7) $$GASU(t, t') \ =_{df} \ B_{G,t'} = \lambda q[B_{G,t}({}^{\vee}[p(t)] \supset {}^{\vee}q)]$$

In **(D4)**, η has type $\langle\langle s,t\rangle,t\rangle$, so, for instance, $B^T_{x,t}$ represents true supposition.

AU is meant to represent a relation of "assertion based supposition update" between O and O$'$: the propositions satisfying O$'$ are the updates relative to $p(t)$ of those satisfying O. $MIASU$ stands for a relation of true, "mutually supposed, individual assertion-based supposition update." This relation holds between two times if the group mutually believes that each of its members follows the update rule in getting from the first time to the second. $GASU$ represents a relation of "group assertion-based supposition update." This relation holds between two times in case group update from the first time to the second follows the group update rule.

One of our main results is that *MIASU* induces a kind of pre-established harmony, so that it's as if there is a group that behaves like an agent, in that its suppositions evolve according to the rule for individuals. That is, we will show that *MIASU* implies *GASU*.

9. COORDINATION IN INTENSIONAL LOGIC

The suppositions of the members of a group at a time are *coordinated* if the group truly mutually believes that all members of the group suppose alike. This provides the definition of a constant *Coord* of type $\langle \mathsf{e}, \mathsf{t} \rangle$.

(D8) $$Coord =_{df} \lambda t \forall x[G(x) \supset B^{T}_{G,t}(\forall q[B_{x,t}(q) \equiv B_{G,t}(q)])]$$

Theorem **(T1)** is an *object-level* result; what is claimed when **(T1)** is stated is that a certain sentence is provable in Gallin's version of Intensional Logic from our axioms. In this presentation, I will simply state major results, without giving proofs. Thus, though if I were proceeding systematically I would have to supply, and prove, a number of lemmas about coordination, I state only one representative and fairly important example, without proof.[6]

(T1) $$Coord(t) \supset \forall O[FI(O, B_t) \supset [\forall x[O = B_{x,t}] \wedge [O = B_{G,t}]]]$$

Results such as **(T1)** are the support for my claim that iterated attitudes will collapse if supposition is coordinated.

10. COORDINATION MAINTENANCE WITH INDEFEASIBLE UPDATE

One of our main theorems about update in a monotonic setting is that mutually supposed individual update implies group update. Note that this is proved as a consequence in Intensional Logic of our axioms, (MS1) and (MS2).

(T2) $$\forall t t'[MIASU(t, t') \supset GASU(t, t')]$$

The proof of this result is fairly complex, and depends on showing that the desired property of mutual supposition after update holds for all finite iterations of supposition after update: that is, we show

$$MIASU(t, t') \supset \forall O[FI(O', B'_t) \supset [O = \lambda q B_{G,t}(\check{}[p(t)] \supset \check{}q)]].$$

Details of the proof will be supplied in Thomason [17].

From **(T2)** it is relatively easy to prove that coordination will be preserved by true, mutually supposed individual assertion-based supposition update. Again, this is an object-level theorem in Intensional Logic.

(T3) $$\forall t t'[MIASU(t, t') \supset [Coord(t) \supset Coord(t')]]$$

[6] The proofs are actually fairly complex. I intend to provide the details in an appendix to a forthcoming paper, Thomason [17].

11. THE NEED FOR DEFEASIBLE RULES

We have shown that an update rule, to the effect that what is asserted is always added to the stock of beliefs, will preserve coordination if it is true and mutually supposed. That is, the following monotonic rule of update will preserve coordination:

(Indefeasible Update) $\forall t\, MIASU(t, t+1)$

In fact, this is just the following special case of **(T2)**.

(T4) $\forall t[MIASU(t, t+1) \supset [Coord(t) \supset Coord(t+1)]]$

The only trouble with such a rule is that it is not in general true (and so in general will not be supposed by reasonable agents). The simplest sort of counterexample is just the case in which the speaker's assertion hasn't been correctly identified, perhaps because it wasn't heard. Or someone may object to it, or a hearer could misidentify the speech act, or the assertion may fail to make sense in a number of ways that make the hearer suspect that conversational repairs of some sort are needed.

The texture of ways in which **(Indefeasible Update)** can fail is entirely open-ended, in that it seems to be a hopeless task to enumerate precisely the all the cases in which it can fail. This sort of open-endedness is a familiar story, one that is commonly retold when people justifiy the need for a nonmonotonic formalization of domains such as planning or puzzle solving. The same sort of point has been made to motivate nonmonotonic rules of conversational update in Perrault [10] and in Appelt and Konolige [2].

12. COORDINATION MAINTENANCE WITH DEFEASIBLE UPDATE

Though there seems to be general agreement that conversational rules are an appropriate domain for the application of theories of nonmonotonic reasoning, there is lack of agreement on which theory to use. Perrault [10] uses Reiter's Default Logic, and Appelt and Konolige [2] use Hierarchic Autoepistemic Logic.

Though choice of framework may be important in applications of the theory of conversation to implemented natural language processing systems, the choice should not be very important, I think, at an abstract level where one is trying to understand the general requirements for a theory of conversational update. This is my goal in the present work—I have not yet tried to implement the ideas.

In what follows I will use McCarthy's Circumscription Theory as a means of formalizing defeasible suppositional update. This decision was based primarily on the flexibility and logical depth of this approach to nonmonotonic reasoning. A secondary motivation of the choice is the natural fit between the logical resources that are required by Circumscription Theory, which uses second order logic, and the requirements of a theory of mutual supposition, which—as we have seen—lead us to adopt a rather powerful higher order modal logic. Though you may be able to go a certain distance with schemes, as in Perrault [10] and in Section 6 of this paper, I think it could be difficult to go very far without a more explicit formalization of mutuality.

I use the presentation of Circumscription Theory in Lifschitz [8], which to my knowledge is the best developed and most sophisticated version of the theory available in print, as well as the presentation that is probably most suitable for logicians. However, the application that I will make here will not make much use of the power of the later versions of Circumscription Theory.[7] In fact, it seems from the work that has been done so far that applications to rules of conversation are relatively straightforward in terms of the demands they make on the nonmonotonic apparatus.[8] Here, I will confine myself to a case that is about as simple as it can be; it is meant only to illustrate the general idea.

We retain the constants B, G, p, and FI, which have types $\langle \mathsf{e}, \langle \mathsf{e}, \langle\langle \mathsf{s}, \mathsf{t}\rangle, \mathsf{t}\rangle\rangle\rangle$, $\langle \mathsf{e}, \mathsf{t}\rangle$, $\langle \mathsf{e}, \langle \mathsf{s}, \mathsf{t}\rangle\rangle$ and $\langle\langle\langle \mathsf{s}, \mathsf{t}\rangle, \mathsf{t}\rangle, \langle \mathsf{e}, \langle\langle\langle \mathsf{s}, \mathsf{t}\rangle, \mathsf{t}\rangle, \mathsf{t}\rangle\rangle\rangle$. We add one abnormality predicate $ab1$, of type $\langle \mathsf{e}, \mathsf{t}\rangle$. We also retain the axioms of the indefeasible theory: **(MS1)** and **(MS2)**.

The intended interpretation is now as follows. The "times" I have been speaking of are really conversational turns. At each turn the speaker means something; this proposition is represented by $p(t)$. I'm indulging in gross oversimplification; of course, when some proposition is meant there usually is a linguistic vehicle, and a process of interpretation that gets from this vehicle to one or more things that are meant. This process of interpretation has to be thoroughly public and mutual.[9] In fact, one of the purposes of the theory that I am developing here is the formalization of this process. However, I leave all that out here; $p(t)$, then, represents a proposition that is meant to be recognized as something to be added to the conversational suppositions. And it is plausible that as a matter of default p(t) *will* be added to the record; that is, the conversants will take it to be added unless some recognizable exception is noted. Moreover, it is a public or mutual matter that this default obtains.

To provide for a few exceptional cases, we allow the hearer to make any of three responses at each time: (1) the null response, which we represent as the empty proposition $\hat{}\forall x[x = x]$, (2) *What,* which can be taken to be something like "I don't understand," and *No,* which can be taken to be something like "I disagree." Either of the last two responses creates an exception to the rule of "assertional uptake." To represent this, we add six conversational axioms to our list.

(D9) $\quad \top =_{df} \forall x[x = x]$

(Limited Responses) $\quad \forall t[r(t) = \hat{}\top \vee r(t) = What \vee r(t) = No]$

(Exclusive Alternatives) $\quad \hat{}\top \neq What \wedge \hat{}\top \neq No \wedge What \neq No$

(Rule for 'What') $\quad \forall t[[r(t) = What] \supset ab1(t)]$

(Rule for 'No') $\quad \forall t[[r(t) = No] \supset ab1(t)]$

(Defeasible Update) $\quad \forall x[G(x) \supset B^{T}_{G,t}(\neg ab1(t) \supset AU(B_{x,t}, B_{x,t+1}, t))]$

(Conventionality of Normalcy) $\quad \forall t[B_{G,t}(\neg ab1(t)) \supset \neg ab1(t)]$

The first four of these axioms are well motivated by what has just been said about the intended interpretation. The fifth, **(Defeasible Update)**, says that it is mutually supposed of every agent

[7] As far as I can see, the generalization of the underlying logic from extensional second order logic to Intensional Logic does not create any logical difficulties, and would be trouble-free even if we wished to minimize higher order constants. However, in this application, we will only need to minimize a single first order constant.

[8] I suspect that extending the theory to take implicature into account would change this, since this phenomenon seems to draw heavily on world knowledge, and to involve many complex defaults, that may well conflict with one another in some cases.

[9] See Thomason [17] for discussion and illustrations of this point.

in the group that the agent will make an appropriate assertional update if nothing unusual occurs. This is actually a fairly straightforward adaptation of the indefeasible update rule, except that—as will be seen in a moment—it is critical that the $B^T_{G,t}$ operator is placed outside the scope of $\neg ab1(t)$. We cannot assume that abnormality will automatically be recognized when it occurs, much less that it will be mutually supposed to have occurred.

The last postulate is plausible enough if you think about the interpretation of normality. Given the public, conventional nature of conversational defaults, it can't possibly happen that a situation is conversationally abnormal if at the same time the group supposes it to be normal.

We now form a circumscriptive theory by minimizing $ab1$ relative to the axioms we have given, allowing all other parameters to vary freely. That is, in terms of Lifschitz [8], we add the following single policy axiom to the rest.

(Minimize $ab1$) $$V_{ab1,ab1}$$

It is easy to see that our axioms yield the following consequence.

(T5) $$\forall t[B_G(\neg ab1(t)) \supset MIASU(t, t+1)]$$

Putting this together with **(T3)**, we have our final theorem.

(T6) $$\forall t[B_G(\neg ab1(t)) \supset Coord(t) \supset Coord(t+1)]$$

Notice that we would not want to strengthen this result by replacing the antecedent $B_G(\neg ab1(t))$ by $ab1(t)$; it is too easy to imagine situations that are abnormal, but are not mutually supposed to be abnormal. (For instance, the hearer responds with "What?" but the speaker doesn't hear the response.) This, of course, raises the question of how things that happen prominently in a shared environment come to be mutually supposed; I have nothing new to say about this problem.

REFERENCES

[1] Allen, J., "Recognizing intentions from natural language utterances." In M. Brady and R. Berwick eds., *Computational models of discourse,* Cambridge, MA: MIT Press, 1983, pp. 107–166.

[2] Appelt, D. and K. Konolige, "A practical nonmonotonic theory for reasoning about speech acts." *26th Annual meeting of the Association for Computational Linguistics: proceedings of the conference.* Association for Computational Linguistics, 1988, pp. 170–178.

[3] Chellas, B., *Modal logic.* Cambridge, England: Cambridge University Press, 1980.

[4] Clark, H., and C. Marshall, "Definite reference and mutual knowledge." In A. Joshi, I. Sag, and B. Weber, eds., *Elements of discourse understanding,* Cambridge, England: Cambridge University Press, 1981.

[5] Cohen, P., and H. Levesque, "Speech acts and rationality." *23rd Annual meeting of the Association for Computational Linguistics: proceedings of the conference.* Association for Computational Linguistics, 1985, pp. 49–59.

[6] Gallin, D., *Intensional and higher-order modal logic.* Amsterdam: North-Holland, 1975.

[7] Lewis, D., *Convention.* Cambridge, MA: Harvard University Press, 1969.

[8] Lifschitz, V., "Circumscriptive theories." In R. Thomason, ed., *Philosophical logic and artificial intelligence,* Reidel, 1989, pp. 109–159.

[9] Montague, R., "Pragmatics and intensional logic." In R. Montague, *Formal Philosophy,* pp. 119–147. New Haven: Yale University Press, 1974.

[10] Perrault, C.R., "An application of default logic to speech act theory." In *Intentions in communication.* P. Cohen, J. Morgan, and M. Pollack, eds., Cambridge, MA: MIT Press, 1990.

[11] Schiffer, S., *Meaning.* London: Oxford University Press, 1972.

[12] Stalnaker, R., "Pragmatics." In G. Harman and D. Davidson, eds., *Semantics of natural language,* pp. 380-397. Dordrecht: D. Reidel, 1972.

[13] Stalnaker, R., "Pragmatic presuppositions." In M. Munitz and P. Unger, eds., *Semantics and philosophy,* pp. 197–213. New York: Academic Press, 1975.

[14] Stalnaker, R., "Indicative conditionals." *Philosophia* 5 (1975), pp. 269–286.

[15] Stalnaker, R., "Assertion." In P. Cole, ed., *Radical pragmatics,* pp. 315–332. New York: Academic Press, 1981.

[16] Thomason, R., "Accommodation, meaning, and implicature: interdisciplinary foundations for pragmatics." In *Intentions in communication.* P. Cohen, J. Morgan, and M. Pollack, eds., Cambridge, MA: MIT Press, 1990.

[17] Thomason, R., "Propagating epistemic coordination through mutual defaults II." Forthcoming.