Repairing Conversational Misunderstandings and Non-Understandings[1]

Graeme Hirst, Susan McRoy,[2] Peter Heeman,[3] Philip Edmonds, and Diane Horton

Department of Computer Science
University of Toronto
Toronto, Ontario M5S 1A4
Canada

June 1994

Abstract

Participants in a discourse sometimes fail to understand one another, but, when aware of the problem, collaborate upon or negotiate the meaning of a problematic utterance. To address nonunderstanding, we have developed two plan-based models of collaboration in identifying the correct referent of a description: one covers situations where both conversants know of the referent, and the other covers situations, such as direction-giving, where the recipient does not. In the models, conversants use the mechanisms of refashioning, suggestion, and elaboration, to collaboratively refine a referring expression until it is successful. To address misunderstanding, we have developed a model that combines intentional and social accounts of discourse to support the negotiation of meaning. The approach extends intentional accounts by using expectations deriving from social conventions in order to guide interpretation. Reflecting the inherent symmetry of the negotiation of meaning, all our models can act as both speaker and hearer, and can play both the role of the conversant who is not understood or misunderstood and the role of the conversant who fails to understand.

1 Introduction

A common attitude in artificial intelligence research is that the tasks that are so difficult for our computers to perform—often, seemingly impossible—are trivial for people. Language, for example, seems effortless for people, and yet after 40 years of research in computational linguistics, we are still far from a complete solution.

But in fact, language understanding is often difficult for people too. Much that is written or said is not understood. Sometimes this is due to inattentiveness or difficulty in hearing. But sometimes the causes are more fundamental. A scientific or technical paper that expresses complex ideas can be hard to understand simply because of the difficulty of the material. And even simple ideas can be hard to understand if they are poorly written or expressed; high-quality language generation is very difficult for people.

Nevertheless, people are, in general, quite successful in their use of language. That's because they have strategies for coping with their linguistic limitations. If they can't understand what is being said to them, they seek clarification and try to work things out. And people are very good at noticing when a conversation has gone awry as a result of one party misunderstanding the other, figuring out the exact nature of the problem, and saying the right thing to recover from the error.

It would be unreasonable, then, to expect computers to always understand language perfectly. But what we should expect is that computers have the flexibility to deal with the consequences of this imperfection. Like people, they should have the ability to recognize problems in understanding and to correct them.

2 Not understanding and misunderstanding

Participants in a dialogue bring to it different beliefs and goals. These differences can lead them to make different assumptions about one another's actions, construct different interpretations of discourse objects, or produce utterances that are either too specific or too vague for others to interpret as intended. This may lead to not understanding or to misunderstanding.

By *not understanding*, we mean a participant's failure to find any complete and unique interpretation of an utterance. This could mean finding no interpretation at all for some or all of the utterance or it could mean finding more than one interpretation and not being able to choose between the alternatives. An important aspect of not understanding is that the participant is *aware* that it has happened.

By contrast, the participant is not aware, at least initially, when *misunderstanding* has occurred. In misunderstanding, the participant obtains an interpretation that she believes is complete and correct, but which is, however, not the one that the other speaker intended her to obtain.[4] It is possible that a misunderstanding will remain unnoticed in a conversation and the participants continue to talk at cross-purposes. Alternatively, the conversation might break down, leading one participant or the other to determine that a misunderstanding has occurred.

It is thus useful to divide recognized misunderstandings into two types: *self-misunderstandings* are those that are both made and detected by the same participant, and *other-misunderstandings* are those that are made by one participant but detected by another. Self-misunderstandings arise when a participant finds that he cannot incorporate the other's new utterance into the discourse coherently, unless he interprets one of the other's earlier utterances differently. Other-misunderstandings occur when a participant recognizes that if one of his own acts had been interpreted differently, the other's utterance would have been the expected response to it. The participant might then attempt to change the other's interpretation. For example, he might restate his message, or explicitly tell the other that she has misunderstood; or he might do nothing (Blum-Kulka and Weizman, 1988), perhaps in order to avoid social awkwardness.

In our research, we have considered both misunderstanding and not understanding, trying a somewhat different approach to each.

3 Not understanding a referring expression

3.1 Referring as collaboration

The linguistic task of referring to some object or idea can involve a *collaboration* between the speaker and the hearer. The speaker has the goal of having the hearer identify the object that the speaker has in mind. The speaker attempts to achieve this goal by constructing a description of the object that she thinks will enable the hearer to identify it. But since the speaker and the hearer will inevitably have somewhat different beliefs about the world, the hearer might not be able to identify the object from this description—that is, not understand the reference. Often, in such cases, the speaker and hearer will collaborate in making a new referring expression that accomplishes the reference.

In an important series of experiments, Clark and his colleagues—especially Wilkes-Gibbs—have shown that conversants will often engage in a kind of *negotiation* in order for one of them to understand a reference that the other wishes to make (Clark, 1993; Clark and Wilkes-Gibbs, 1986). In their fundamental experiment, Clark and Wilkes-Gibbs gave pairs of subjects each a copy of a set of hard-to-describe tangram figures. The subjects' task was to arrange their sets in the same order, and to do so by conversation alone; neither could see the other's set. The subjects were thus obliged to construct descriptions for each tangram that they hoped the other could interpret correctly; for example, *the one that looks like an angel with a stick.*

Clark and Wilkes-Gibbs found that typically the participant trying to describe a tangram pattern would present an initial referring expression. The other participant would then pass judgment on it, either *accepting* it, *rejecting* it, or *postponing* his decision. If it was rejected or the decision postponed, then one participant or the other would *refashion* the referring expression. This would take the form of either *expanding* the expression by adding further qualifications or *replacing* the original expression with a new expression. The referring expression that resulted from this was then judged, and the process continued until the referring expression was acceptable enough to the participants for their current purposes. The final expression then becomes part of the participants' common ground.

This excerpt from Clark and Wilkes-Gibbs's data illustrates rejection, replacement, and acceptance:

Example 1

1 **A:** Okay, and the next one is the person that looks like they're carrying something and it's sticking out to the left. It looks like a hat that's upside down.

2 **B:** The guy that's pointing to the left again?

3 **A:** Yeah, pointing to the left, that's it! *[laughs]*

4 **B:** Okay.

In this dialogue, B implicitly rejects A's initial presentation by replacing it with a new referring expression in line 2, *the guy that's pointing to the left again*. A then accepts the refashioned referring expression in line 3.

This kind of reference negotiation is not found only in laboratory settings. A particularly clear instance can be seen in the following example from the London–Lund *Corpus of English conversation* (Svartvik and Quirk, 1980, S.2.4a:1–8), in which the conversants collaborate simultaneously on the phrases *that weird creature* and *over there*.

Example 2

1 **A:** What's that weird creature over there?

2 **B:** In the corner?

3 **A:** *[affirmative noise]*

4 **B:** It's just a fern plant.

5 **A:** No, the one to the left of it.

6 **B:** That's the television aerial. It pulls out.

3.2 A model of collaboration on referring

Heeman and Hirst have presented a computational model of this kind of collaboration; it covers the agent who makes the initial referring expression, that is, the *initiator*, and the agent who is to understand the expression, the *recipient*. In this model, the initiator has the goal of referring to something, and constructs a plan, in the form of a sequence of linguistic actions, to try to achieve it, given a set of beliefs about what the recipient believes. The recipient, seeing only the surface linguistic actions, tries to infer the plan in order to understand the reference. Thus, referring expressions are represented by plan derivations, and an unsuccessful referring expression is an invalid plan in whose repair the agents collaborate. This collaboration takes place through the use of plans that judge and refashion the expression.

In addition to the planning component, we also model the change in state of the participant through belief- and goal-adoption rules. So, we are able to model a conversational agent throughout the collaborative activity, both in its role as a speaker and its role as a hearer. In fact, two copies of the model can converse with one another.[5] Acting as a hearer, the system performs plan inference on each set of actions that it observes, and then updates the state of the collaboration. It then takes its turn as speaker. As the new speaker, the system looks for a goal that it can adopt, and then constructs a plan to achieve it. Next, presupposing the other participant's acceptance of the plan, it updates the state of the collaboration. It repeats this until there are no more goals to adopt. The surface actions of the constructed plans form the response of the system. The system then takes its turn as hearer and waits for a response from the other copy.

In the rest of this section, we give a brief overview of the model; details are given by Heeman (1991) and Heeman and Hirst (1994).

3.3 Plans for referring

We extend the earlier approaches of Cohen (1981) and Appelt (1985b) in planning not only the occurrence of a referring expression but also, at the same level, its *content*. So, we use surface-level linguistic actions to account for each part of a description. Table 1 lists these.

INSERT TABLE 1 NEAR HERE

These linguistic actions are the building blocks that referring expressions are made from. Acting as the mortar are intermediate plans, which through their preconditions encode the knowledge of how a description can allow a hearer to identify an object. The constraints express the conditions under which an attribute can be used to refer to an object; for instance, that it be mutually believed that the object has a certain attribute (Clark and Marshall, 1981; Perrault and Cohen, 1981; Nadathur and Joshi, 1983). In addition, the intermediate plans have constraints that keep track of the potential referents, so as to ensure that the referring expression includes sufficient descriptors so that the hearer can (in the speaker's opinion) identify the referent (Dale, 1989; Reiter, 1990). These constraints are not only useful in plan construction but also in plan inference, for they allow the referent to be determined in a manner analogous to constraint satisfaction.

Following Pollack (1990), our plan inference process can infer plans in which, in the hearer's view, a constraint does not hold. In inferring a plan derivation, the system as hearer first finds the set of plan derivations that account for the primitive actions that were observed, without regard to whether the constraints hold. Second, it evaluates each derivation by attempting to find an instantiation for the variables such that all of the constraints hold with respect to the hearer's beliefs about the speaker's beliefs. If the hearer is able to satisfy the constraints, then he will have understood the plan and be able to identify the referent, since a term corresponding to it will have been instantiated in the inferred plan. Otherwise, he has a constraint that is unsatisfiable, making this derivation ill-formed and the referent unresolvable. After all derivations have been evaluated, if there is just one valid plan, then the hearer will believe that he has understood (and will have identified the referent). Otherwise, if there is one invalid derivation, the action containing the constraint that is the source of the invalidity is noted, and will be used by the hearer in his attempt to fix the non-understanding.[6]

3.4 Plans for collaborating

If the initial referring expression is not understood, then the conversants will collaborate in its repair. We have formalized the conversational moves of Clark and Wilkes-Gibbs as discourse actions, which take as a parameter the current referring expression. These actions, given in table 2, make use of plan repair techniques to refashion the expression, and make use of surface linguistic actions, given in table 3, to communicate the parts of the referring plan that need to be removed and parts that need to be added.

INSERT TABLE 2 NEAR HERE

INSERT TABLE 3 NEAR HERE

The first step in repairing an ill-formed plan is to communicate the source of the error to the other participant. The referring expression might include surface linguistic actions that are inappropriate for identifying the referent, thus precluding any object from matching the description, or it might not include enough, and so too many objects would match. In the first case, the error would be in the part of the plan that decomposes into these inappropriate actions, and so the hearer would construct an instance of `reject-plan` that, through its surface linguistic action `s-reject`, would reject them. For instance, if the referring expression were *the weird creature*, and the hearer couldn't identify anything that he thought weird, he might say *What weird thing?*, thus rejecting the surface linguistic action corresponding to *weird*. In the second case, in which too many objects match the description, the error would be in the referring plan that terminates the addition of modifiers. So, the hearer would construct an instance of `postpone-plan`, which, through its surface linguistic action `s-postpone`, might be expressed by a tentatively voiced *Okay*.

In either case, the hearer or the other will *refashion* the expression in the context of the rejection or postponement. In keeping with Clark and Wilkes-Gibbs, we use two discourse plans for refashioning: `replace-plan` and `expand-plan`. The first is used to replace some of the actions in the referring expression plan with new ones (as in line 2 of Example 1 above), and the second is to add more actions. Replacements can be used if the referring expression was either rejected or postponed, while an expansion can be used only in the latter case.

The decomposition of the refashioning plans encodes how a new referring expression can be constructed from the old one. This involves three tasks: first, a single candidate referent is chosen; second, the referring expression is refashioned; and third, this is communicated to the hearer by means of `s-actions`.[7] The first step, choosing a candidate, is obviously pre-determined if the speaker of the refashioning is the person who initiated the referring expression. Otherwise, the speaker must choose a possible candidate. Goodman (1985) has addressed this problem for the case of when the referring expression overconstrains the choice of referent. He uses heuristics to relax the constraints of the description and to pick one that *nearly* fits it. This problem is beyond the scope of this research, and so we simply choose one of the referents arbitrarily (but see Heeman (1991) for how a simplified version of Goodman's algorithm that only relaxes a single constraint can be incorporated into the planning paradigm).

The second step is to refashion the referring expression so that it identifies the candidate chosen in the first step. This is done with plan repair techniques (Hayes, 1975; Wilensky, 1983; Wilkens, 1985), in which we remove the part of the plan that is in error and replan with a more appropriate action inserted in its place. As this technique has been encoded into our refashioning plans, it can be used for both constructing repairs and inferring how another agent has repaired a plan.

Once the refashioning plan is accepted, the common ground of the participants is updated with the new referring expression. So the effect of the refashioning plan is that the hearer will believe that the speaker wants the new referring expression plan to replace the current one. This replacement is done regardless of whether the referring expression plan is in fact valid. If it is, however, valid, then the referent can be understood, prompting the hearer to adopt the goal to communicate this to the speaker, leading to the discourse action of `accept-plan`. Otherwise, the process will repeat, but this time with the new referring expression.

Table 4 shows two copies of the system engaging in a (simplified!) version of Example 2.

INSERT TABLE 4 NEAR HERE

4 Collaboration on referring to objects that are not mutually known

4.1 Referring in direction giving

A crucial assumption of Clark and Wilkes-Gibbs's work—and of Heeman and Hirst's model—is that the hearer of the initial referring expression already has some knowledge of the referent in question. In Clark and Wilkes-Gibbs's experiments, it is one of the given tangram figures; in Example 2, it is an object in the room that both speakers can see. However, a speaker sometimes has to refer to an object that is not previously known to the hearer. One particular situation in which this arises is in giving directions. For example, the speaker might give a direction like the following:

Example 3

1 **A:** Go straight ahead until you get to *a funny-looking building*.

The recipient has to understand the reference well enough that when he later reaches the building, he will recognize it as the intended referent. Although this type of reference is different from the kind of referring action that Heeman and Hirst modeled, conversants can nevertheless collaborate to achieve an understanding of them. This can be seen in the following portion of a telephone conversation recorded by Psathas (1991, p. 196).

Example 4

1 **A:** You just stay on 2A, until you get to Lowell Street.

2 **B:** Is it marked?

3 **A:** Yeah, I think there's a street sign there, it's an intersection with lights.

4 **B:** Okay.

In this dialogue, speaker B has not understood the reference to the intersection at Lowell Street, and so suggests that the intersection might be marked. Speaker A replies with an elaboration of the initial expression.

Edmonds (1993; 1994) has presented a computational model of this type of collaboration that draws from Heeman and Hirst's model. The domain is that of giving directions for someone unfamiliar with an area to get to a particular place. In this section, we give an overview of Edmonds's model.

The basis of the model is that the hearer can accept a referring expression plan if (1) the plan contains a description that is *useful* for making an *identification plan* that the hearer can execute to identify the referent, and (2) the hearer is *confident* that the identification plan is *adequate*.

The first condition, originally described by Appelt (1985a), is important because the success of the referring action depends on the hearer formulating a useful identification plan. We can take the referring expression plan itself to be the identification plan because its intermediate plans encode only useful descriptions. For the second condition to hold, the hearer must believe that the identification plan is good enough to uniquely identify the referent when it becomes visible. This involves giving enough information by using the most visually prominent or *salient* attributes of the referent.

Each agent associates a numeric *confidence value* with each of the attributes in the referring expression, and by composing these,[8] computes a level of confidence in the adequacy of the complete referring expression plan that can be interpreted as ranging from low confidence to high confidence. If the overall confidence value exceeds some set value, the agent's *confidence threshold*, then the agent believes the plan to be adequate. That is, if the agent is the initiator, she believes that the other will be able to understand the reference; if the agent is the other, he believes that he has understood the reference.

Now, the confidence value of each attribute is equivalent to its *salience* within the context of the referring expression. Salience, for our purposes in direction-giving, is primarily visual prominence, but can also involve identifiability, familiarity, and functional importance (Devlin, 1976; Lynch, 1960). One approach is to encode the salient properties in a static hierarchy as Davis (1989), and Reiter and Dale (1992) have done. But, ideally, salience should depend on the context surrounding the referent. For example, the height of a tall building would normally be salient, but not if it were surrounded by other tall buildings. This computation would be quite complex, so we have adopted a middle ground between the simple context-independent approaches, and a full-blown contextual analysis. The middle ground involves taking the type of object into account when choosing attributes and landmarks that relate to it. For example, height and architectural style can be very salient features for describing a building, but not for describing an intersection for which having a sign or traffic lights is important. This approach still allows us to encode salience in a hierarchy, but it is dependent on the referent.

Table 5 is an example of a simple salience hierarchy that an agent might have. The hierarchy is actually a set of partial orderings of attributes, represented by lambda expressions, indexed by object type. In the table, the confidence value of using architectural style to describe a building is 4. The confidence value of a tall building is 3, and so this attribute is less salient than architectural style. The other rows (for describing intersections) follow similarly. Given information about salience, we could construct such a hierarchy, but we do not presume that it would be easy to know what is salient.

INSERT TABLE 5 NEAR HERE

Each agent has his own beliefs about salience. It is the difference in their beliefs that leads to the necessity for collaboration on reference. Ideally, the initiator should construct referring expressions with the recipients' (believed) beliefs about salience in mind, but we have chosen to avoid this complexity by making the simplifying assumption that the initiator is an expert (and thus knows best what is salient).

4.2 Construction and inference of referring plans

The salience hierarchy is used in both plan construction and plan inference.

In plan construction, salience is used for constructing initial referring expression plans, elaborating on inadequate plans, and for suggesting possible elaborations to plans by allowing an agent to choose the most salient properties of the referent first. The agent constructs an initial referring expression plan in almost the same way as in Heeman and Hirst's system. Actions in the intermediate plans of a referring expression plan allow the speaker to choose the most salient attributes that have not yet been chosen, and constraints in the surface linguistic actions make sure the speaker believes that each attribute is true.[9] For example, to construct the reference to the *building* in Example 3, the speaker consulted her salience hierarchy (in Table 5) and determined that architectural style is salient. Hence, she described the building as *funny-looking*. This single attribute was enough to exceed her confidence threshold.

During plan inference, the salience hierarchy is used when judging a recognized plan. Actions in the intermediate plans determine the confidence values of each attribute, and add them up. A final constraint in the plan makes sure the overall confidence sum exceeds the confidence threshold of the agent. This means that judging the adequacy of a referring expression plan falls out of the regular plan evaluation process. For example, after recognizing the reference in Example 3, the hearer evaluates the plan. Assuming he believes the salience information in Table 5, he computes the confidence value of 4. If this value exceeds his confidence threshold, then he will accept the plan. If not, he will believe that there is an error at the constraint that checks his confidence threshold.

4.3 Suggestion and elaboration

If the hearer is not confident in the adequacy of the plan, he uses an instance of `postpone-plan` to inform the initiator that he is not confident of its adequacy (which causes the initiator to raise her own confidence threshold). Now, although he cannot refashion the expression himself, he does have the ability to help the initiator by *suggesting* a good way to expand it; *suggestion* is a conversational move in which an agent suggests a new attribute that he deems would increase his confidence in the expression's adequacy if the expression were expanded to include the attribute. Continuing with the example, if the hearer were not confident about the adequacy of *the funny-looking building*, he might suggest that the initiator use height (as well as architectural style), by asking *Is it tall?*. From this suggestion the initiator might expand her expression to *the tall funny-looking building*. So, in our sense, a suggestion is an illocutionary act of questioning; along with actually suggesting a way to expand a plan, the agent is asking whether or not the referent has the suggested attribute.

To decide what suggestion to make, the agent uses an instance of `suggest-expand-plan`, which has an action in its decomposition that chooses the most salient attribute that has not been used already.

However, only the initiator of the referring expression can actually *elaborate* a referring expression, because only she has the knowledge to do so. Depending on whether the hearer of the expression makes a suggestion or not, the initiator has two options when elaborating a plan. If no suggestion was made, then she can expand the plan according to her own beliefs about the referent's attributes and their salience. On the other hand, if a suggestion was made, she could instead attempt to expand the plan by using the attribute suggested.

In constructing an instance of `expand-plan`, the plan constructor attempts to find an expansion whose results include the surface linguistic actions for the attribute, but this might not be possible. In any case, the speaker constructs an expansion that will make the plan adequate according to her beliefs.[10]

The response to a suggestion depends, obviously, on whether or not the suggestion was used to expand the plan. The speaker can (1) affirm that the plan was expanded with the suggestion by using the `s-affirm` linguistic action; (2) affirm that the suggestion was used, along with additional attributes that weren't suggested, by using `s-affirm` and `s-actions`; or (3) deny the suggestion with `s-deny`, and inform the other by `s-actions` as to how the plan was expanded.

By using the expansion and suggestion moves, the two agents collaborate on refashioning the referring expression until the recipient of the directions is confident that it is adequate.

4.4 Example

We have implemented the model in Prolog. Table 6 shows two copies of the system engaging in a simplified version of Example 4.

INSERT TABLE 6 NEAR HERE

5 Misunderstanding

5.1 Misunderstanding and repair

Participants in a conversation rely in part on their expectations to determine whether they have understood each other. If a participant doesn't notice anything unusual, she may assume that the conversation is proceeding smoothly. But if she hears something that seems inconsistent with her expectations, she may hypothesize that there has been a misunderstanding, either by herself or the other, and produce a *repair*—an utterance that attempts to correct the problem.

One common type of repair involves correcting another speaker's interpretation of the discourse. In the simplest case, a speaker makes an utterance displaying her misunderstanding in the turn immediately following the one she misunderstood. If the other speaker then recognizes the misunderstanding and initiates a sequence to resolve the misunderstanding, this is a *third-turn* (or *third-position*) repair, so called because the repair is initiated in the third turn of the top-level sequence, counting from the misunderstood utterance. Consider Example 5 from Coulthard and Brazil (1984). In this example, B has responded to line 1 with an *acknowledgement*, interpreting line 1 as an *inform*.

Example 5

1 **A:** So the meeting's on Friday.

2 **B:** Thanks.

3 **A:** No, I'm *asking* you.

However, A intended line 1 to be a *yes–no* question (presumably with an *inform* as the expected reply). Recognizing B's misunderstanding, A produces a third-turn repair in line 3, telling B what action A had intended in line 1. Speaker A could have also told B the intended goal (*e.g.,* "No, I want you to tell me.")

Another type of repair involves producing a new reply to a turn that one has apparently misunderstood. If a conversant hears an utterance that seems inconsistent with her expectations (perhaps because she has misunderstood some previous utterance) and the inconsistency leads her to reinterpret an earlier utterance and produce a new response to it, this is a *fourth-turn* (or *fourth-position*) repair (Schegloff, 1987). Such repairs not only display the alternative interpretations, but also indicate some of the information that may underlie a participant's decision to favor one of them over another. Consider the fragment of conversation between a mother and her child (named Russ), shown in Example 6 (Terasaki, 1976).

Example 6

1 **Mother:** Do you know who's going to that meeting?

2 **Russ:** Who?

3 **Mother:** I don't know.

4 **Russ:** Oh. Probably Mrs McOwen and probably Mrs Cadry and some of the teachers.

In this dialogue, Russ initially interprets line 1 as expressing Mother's desire to tell, that is, as a *pretelling* or *preannouncement*, but finds this interpretation inconsistent with her next utterance. In line 3, instead of telling him who's going (as one would expect after pretelling), Mother claims that she does not know (and therefore could not tell). Russ recovers by reinterpreting line 1 as an indirect request, which his line 4 attempts to satisfy. This example also demonstrates agents' reluctance to repair the problems in the utterances of others (Schegloff *et al*, 1977); although Mother might have produced a third-turn repair at line 3, the manifestation of a misunderstanding provided her with an expectable option that allowed her to avoid having to produce an explicit repair.

5.2 The need for both intentional and social information

Any dialogue system must account for the detection and repair of misunderstandings. But to consider possible misunderstandings in addition to intended interpretations would explode the number of alternatives that an interpreter would need to consider, unless there were adequate constraints. And predominant computational approaches to dialogue, which are based on intention, already have difficulty constraining the interpretation process. Proposed sociological accounts are more constrained, but none are computational. Some synthesis of intentional and social accounts of discourse is required.

In intentional accounts, speakers use their beliefs and goals to decide what to say; when hearers interpret an utterance, they try to identify goals that might account for it. This sort of reasoning is difficult to constrain because, although beliefs can narrow the search for an interpretation, there is no principled way of constraining the depth of the search. For each motivation that a hearer considers, he must also consider any higher-level motivations that it might support. To make such an approach workable, many simplifying assumptions have to be made, including the assumption that previous parts of the conversation have been understood correctly. However, there is another way to address misunderstanding that avoids this unconstrained inference of goals: use expectations deriving from social conventions (rather than intention) to guide interpretation.

In sociological accounts provided by Ethnomethodology, both coherent discourse interactions and repairs of misunderstandings are normal activities guided by social conventions (Garfinkel, 1967; Schegloff, 1992). There are conventions regarding the expected range of responses to every action, for example. People then can assume that others are behaving as expected, unless they have reason to believe otherwise. In this way, the conventions give speakers a guide to possible interpretations. Reasoning is also limited, because conventions do not depend on the psychological characteristics of particular participants. What these accounts lack that computational accounts provide is an explanation of how people can identify the convention that is relevant, especially when there is no active expectation.

5.3 A synthesis

In our work (described more fully by McRoy (1993) and McRoy and Hirst (1993a; 1994)), we have developed a model of interaction that addresses the possibility that the participants might differ about the speech act that is performed by some utterance, without requiring extended reasoning about the speaker's goals. According to the model, speakers form expectations on the basis of what they hear, and thus monitor for differences in understanding. If necessary, they also reinterpret utterances in response to new information and generate repairs. Beliefs about the discourse context and conventions for interaction are used to select speech acts that are appropriate for accomplishing the speakers' goals. Interpretation and repair attempt to retrace this selection process abductively—when speakers attempt to interpret an observed utterance, they try to identify the goal, expectation, or misunderstanding that might have led the other agent to produce it.

The model uses both intentional and social sources of knowledge. Intentional information is captured by two relations: one between utterances (input forms) and speech acts, and one between utterances and the attitudes that they express. These relations are the basis for deciding whether a set of utterances is consistent. To capture socially-derived expectations, the theory includes a relation on the speech acts—for each act, which acts are expected to follow. It also contains an axiomatization of speakers' knowledge for generating appropriate utterances and for detecting and repairing misunderstandings. The model demonstrates how these decisions depend on interactions among discourse participants' beliefs, intentions, previously expressed attitudes, and knowledge of social conventions.

The key features of the model that distinguish it from previous ones are the following:

- An account of the detection and repair of speech act misunderstandings and its relation to generation and interpretation. Although there has been work on identifying *potential* sources of misunderstanding, none of it addresses the problem of identifying and repairing *actual* misunderstandings. Also, unifying these tasks requires that linguistic knowledge and processing knowledge be kept distinct, improving the clarity of the model and permitting general knowledge about language to be reused.

- An integration of the socially-determined, structural conventions that have been identified by Ethnomethodology with the use of belief and intention that has been popularized within Artificial Intelligence. As a result, the model does not do extended inference about goals when it is not necessary.

- An account of the nonmonotonicity of discourse reasoning. In particular, the interpretation of utterances and the detection of misunderstandings are both characterized as abduction problems; speech act generation is characterized as default reasoning. As a result, all three processes can be specified within a single theory of communicative interaction.

- A reification of expectation. According to the model, agents form expectations on the basis of social conventions. They filter these expectations by considering the consistency of the Gricean intentions that they have expressed. By contrast, previous models of discourse attempt to eliminate interpretations by using some (necessarily incomplete) set of felicity conditions.

- An axiomatization in Prioritized Theorist (Poole *et al*, 1987). Theorist is a declarative framework for default and abductive reasoning. Thus, linguistic knowledge and processing knowledge are kept distinct.

5.4 The architecture of our model

In the architecture that we have formulated, producing an utterance is a default, deductive process of choosing both a speech act that satisfies an agent's communicative and interactional goals and an utterance that will be interpretable as this act in the current context. Utterance interpretation is the complementary (abductive) process of attributing communicative and interactional goals to the speaker by attributing to him or her a discourse-level form that provides a reasonable explanation for an observed utterance in the current context. Expectations derived from social norms delimit the range of responses that can occur without additional explanation. The attitudes that speakers express provide additional constraints, because speakers are expected not to contradict themselves. We therefore attribute to each agent:

- A set $\mathcal{B}$ of prior assumptions about the beliefs and goals that were expressed by the speakers (including assumptions about misunderstanding), organized into a sequence of turns, *TS*.
- A set $\mathcal{M}$ of potential assumptions about misunderstandings and meta-planning decisions that agents can make to select among coherent alternatives.
- A theory $\mathcal{T}$ describing his or her linguistic knowledge, including principles of interaction and facts relating linguistic acts.

DEFINITION **1:** An *interpretation* of an utterance **u** to hearer **h** by speaker **s** in discourse context **ts** is a set M of instances of elements of $\mathcal{M}$, such that

1. $\mathcal{T} \cup \mathcal{B} \cup M$ is consistent
2. $\mathcal{T} \cup \mathcal{B} \cup M \models utter(\mathbf{s}, \mathbf{h}, \mathbf{u}, \mathbf{ts})$
3. $\mathcal{T} \cup \mathcal{B} \cup M$ is not in conflict with any stronger defaults that might apply.[11]

DEFINITION **2:** It would be *coherent* for **s** to utter **u** in discourse context **ts** if the utterance can be derived from an agent's linguistic knowledge, assuming some set M^{meta} of meta-planning decisions, such that

1. $\mathcal{T} \cup \mathcal{B} \cup M^{meta}$ is consistent
2. $\mathcal{T} \cup \mathcal{B} \cup M^{meta} \models utter(\mathbf{s}, \mathbf{h}, \mathbf{u}, \mathbf{ts})$
3. $\mathcal{T} \cup \mathcal{B} \cup M^{meta}$ satisfies the priority constraints.

That is, **u** is a solution to the following default reasoning problem:

$$\mathcal{T} \cup \mathcal{B} \cup M^{meta} \vdash (\exists u)\ utter(\mathbf{s}, \mathbf{h}, u, \mathbf{ts})$$

DEFINITION **3:** A speaker *S1* is *expected* to do action *R* in dialogue *TS* whenever there is an action *A* that is active in *TS* (because it was performed earlier), *R* as the normal expected reply to *A*, and the linguistic intentions of *R* are compatible with *TS*.

In addition, acts of interpretation and generation update *TS*, the set of beliefs and goals assumed to be expressed during the discourse. The current formalization focuses on the problems of identifying how an utterance relates to a context and whether it has been understood. The update of expressed beliefs is handled in the implementation, but outside the formal language.[12]

The following formulates our characterization of misunderstanding and repair:

Other-misunderstanding

Speaker s_1 might be attempting action a_{new} in discourse ts if:

1. Earlier, speaker s_2 performed act $a_{intended}$;
2. Actions $a_{intended}$ and $a_{similar}$ can be performed using a similar surface form;
3. If s_2 had performed $a_{similar}$, then a_{new} would be expected; and
4. s_1 may have mistaken $a_{intended}$ for $a_{similar}$.

Self-misunderstanding

Speaker s_1 might be attempting action a_{new} in discourse ts if:

1. s_1 has performed action $a_{observed}$;
2. But, the linguistic intentions of a_{new} are inconsistent with the linguistic intentions of $a_{observed}$;
3. $a_{observed}$ and action $a_{intended}$ can be performed using a similar surface-level speech act; and
4. s_2 may have mistaken $a_{intended}$ for $a_{observed}$.

Third-turn repair

Speaker s_1 should tell speaker s_2 that she intended to perform $a_{intended}$ in discourse ts if:

1. s_2 has apparently mistaken an instance of act $a_{intended}$ for act $a_{observed}$; and
2. s_1 may perform a third-turn repair (*i.e.,* the act that s_1 would expect to follow her intended action has not already been done and it would be consistent for s_1 to tell s_2 that she intended to perform $a_{intended}$).

Fourth-turn repair

Speaker s_1 should do action a_{reply} in discourse ts when:

1. s_1 has mistaken an instance of act $a_{intended}$ as an instance of act $a_{observed}$;
2. A reconstruction of the discourse is possible;
3. s_1 would expect to do a_{reply} in this reconstruction; and
4. And, s may perform a fourth-turn repair.

5.5 Example

We have implemented the model in Prolog with the Theorist framework for abduction with prioritized defaults (Poole *et al*, 1987; van Arragon, 1990). Table 7 shows two copies of the system engaging in Example 6.

INSERT TABLE 7 NEAR HERE

5.6 Other kinds of misunderstanding

Although the framework that we have developed is very general, we have developed within it the knowledge for dealing with only certain types of misunderstanding at the level of sentences and speech acts. Conversants may also misunderstand each other because they have come to different beliefs about the structure of the discourse (Horton and Hirst, 1991).

Example 7, in which the conversants are planning a party, shows how discrepancies concerning the structure of discourse can arise. Speaker B misunderstands A's ellipsis in line 7, (as *do you want to ask Karin to come to the party?* instead of *do you want to ask Karin for the recipe?*) and thinks that a topic shift has occurred. At that point, there is a discrepancy between how A and B believe the discourse is segmented.

Example 7

1 **A:** I wonder if we've forgotten anyone.

2 **B:** Did you invite the Mooneys?

3 **A:** Mm hm.

4 **B:** Can you get me the cake recipe?

5 **A:** It's on the shelf above the stove.

6 **B:** It's not there.

7 **A:** Do you want to ask Karin?

8 **B:** Isn't she going to be out of town?

9 **A:** No, I mean ask her for the recipe.

Such discrepancies arise when one conversant attributes the wrong discourse plan to the other. In the above example, after utterance 7 speaker B thinks that speaker A has a plan to resume an earlier topic in the conversation (concerning party invitees), while A actually plans to simply continue the current topic (concerning the cake recipe). Litman and Allen (1990) introduced the notion of discourse plans and used plan recognition techniques to infer them. However, to handle conversations such as example 7, where a discourse plan is wrongly inferred, a plan recognition scheme must be able to both detect such errors (its own or those of the other conversant) and make any necessary revisions to its beliefs or repairs to the conversation.

Such abilities would also enable the modeling of miscommunication arising from discrepancies in the *domain* plans conversants attribute to each other. In the following exchange, a discrepancy arises concerning what domain plan A's question is intended to help her fulfill:

Example 8

1 **A:** Where's CSC104 taught?

2 **B:** Sidney Smith Building, room 2118, but the class is full.

3 **A:** No, I teach it.

Speaker B assumes that A's domain plan is to take CSC104; in fact, her plan is to teach it.

Unfortunately, most existing plan recognition schemes cannot be employed as the foundation of a model of these sorts of miscommunication. Although techniques have been developed to handle situations in which agents' plan libraries differ in some respects (*e.g.*, Pollack 1990; Calistri-Yeh 1991), most current schemes have no mechanism for detecting discrepancies in, or for revising, the plans inferred. Appelt and Pollack (1992) suggested the use of weighted abduction to model the nonmonotonic aspects of plan inference. Although the weighted axioms they define provide only limited coverage, the method itself is interesting. Eller and Carberry (1992) proposed another mechanism that performs detection and revision; it is based on the insight that dialogues requiring plan inference revision are analogous to semantically or syntactically ill-formed input. However, their means of detecting a need to revise[13] is limited; it does not make use of clues in the conversation itself and revision initiated by such clues would appear to require another mechanism.

6 Conclusion

Participants in a dialogue are necessarily limited in the amount of information that they can make explicit. Discourse participants compensate for this limitation by using the evidence provided by their utterances to verify each other's understanding of the conversation as it progresses. To show his understanding and acceptance of an utterance, a hearer may reply with an utterance that is consistent with the speaker's expectations. Alternatively, if he disagrees with the speaker's displayed interpretation, he can initiate a repair. In this way, participants negotiate the meaning of utterances. Reflecting the inherent symmetry of this negotiation, all our models can act as both speaker and hearer, and can play both the role of the conversant who is not understood or is misunderstood and the role of the conversant who fails to understand.

To address nonunderstanding, we have developed two plan-based models of collaboration in identifying the correct referent of a description: one covers situations where both conversants know of the referent, and the other covers situations, such as direction-giving, where the recipient does not. In the models, conversants use the mechanisms of refashioning, suggestion, and elaboration, to collaboratively refine a referring expression until it is successful.

To address misunderstanding, we have developed a model that combines intentional and social accounts of discourse to support the negotiation of meaning. The approach extends intentional accounts by using expectations deriving from social conventions in order to guide interpretation.

Acknowledgements

Our work is supported by the Natural Sciences and Engineering Research Council of Canada and the University of Toronto. For discussions that influenced the course of this work, we are grateful to James Allen, Hector Levesque, Ray Reiter, Jan Wiebe, Mike Gruninger, Sheila McIlraith, Paul van Arragon, Linda Peto, Javier Pinto, Steven Shapiro, Stephen Green, Suzanne Stevenson, Jeff Siskind, and David Traum.

References

Douglas E. Appelt (1985a), "Some pragmatic issues in the planning of definite and indefinite noun phrases", *Proceedings of the 23rd annual meeting of the Association for Computational Linguistics*, Chicago, pp. 198–203.

Douglas E. Appelt (1985b), "Planning English referring expressions", *Artificial Intelligence*, Vol. 26, No. 1, pp. 1–33.

Douglas E. Appelt and Martha E. Pollack (1992), "Weighted abduction for plan ascription", *User Modeling and User-Adapted Interaction*, Vol. 2, No. 1–2, pp. 1–25.

S. Blum-Kulka and E. Weizman (1988), "The inevitability of misunderstandings: Discourse ambiguities", *Text*, Vol. 8, No. 3, pp. 219–241.

Randall J. Calistri-Yeh (1991), "Utilizing user models to handle ambiguity and misconceptions in robust plan recognition", *User Modelling and User-Adapted Interaction*, Vol. 1, No. 4, pp. 289–322.

Herbert H. Clark (1993), *Arenas of language use* (The University of Chicago Press, Chicago, and Center for the Study of Language and Information, Stanford).

Herbert H. Clark and Catherine R. Marshall (1981), "Definite reference and mutual knowledge", in Joshi, Webber, and Sag, 1981, pp. 10–62, and Clark, 1993, pp. 9–59.

Herbert H. Clark and Deanna Wilkes-Gibbs (1986), "Referring as a collaborative process", *Cognition*, Vol. 22, pp. 1–39; reprinted in Cohen *et al*, 1990, pp. 463–493, and in Clark, 1993, pp. 107–143.

Philip R. Cohen (1981), "The need for referent identification as a planned action", *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI-81)*, Vancouver, pp. 31–36.

Philip R. Cohen, Jerry Morgan, and Martha E. Pollack (editors) (1990), *Intentions in communication* (The MIT Press, Cambridge, MA).

Malcolm Coulthard and David Brazil (1984), "Exchange structure", in *Studies in discourse analysis*, edited by Malcolm Coulthard and Martin Montgomery (Routledge & Kegan Paul, London), pp. 82–106.

Robert Dale (1989), "Cooking up referring expressions", *Proceedings of the 27th annual meeting of the Association for Computational Linguistics*, Vancouver, pp. 68–75.

J. R. Davis (1989), *Back Seat Driver: Voice assisted automobile navigation*, PhD thesis, Massachusetts Institute of Technology.

A. S. Devlin (1976), "The 'small town' cognitive map: Adjusting to a new environment", in *Environmental knowing: Theories, research and methods*, edited by G. Moore, and R. Golledge (Dowden, Hutchinson and Ross, Stroudsburg, PA).

Philip G. Edmonds (1993), *A computational model of collaboration on reference in direction-giving dialogues.* MSc thesis, published as technical report CSRI-289, Department of Computer Science, University of Toronto.

Philip G. Edmonds (1994), "Collaboration on reference to objects that are not mutually known", *Proceedings of the 15th International Conference on Computational Linguistics (COLING-94)*, Kyoto, pp. ???–???.

Rhonda Eller and Sandra Carberry (1992), "A meta-rule approach to flexible plan recognition in dialogue", *User Modeling and User-Adapted Interaction*, Vol. 2, No. 1–2, pp. 27–53.

Harold Garfinkel (1967), *Studies in Ethnomethodology* (Prentice Hall, Englewood Cliffs, NJ; reprinted by Polity Press, Cambridge, England, 1984).

Bradley A. Goodman (1985), "Repairing reference identification failures by relaxation", *Proceedings of the 23rd annual meeting of the Association for Computational Linguistics*, Chicago, pp. 204–217.

Philip J. Hayes (1975), "A representation for robot plans", *Proceedings of the 4th International Joint Conference on Artificial Intelligence (IJCAI-75)*, Tblisi, pp. 181–188.

Peter A. Heeman (1991), *A computational model of collaboration on referring expressions.* MSc thesis, published as technical report CSRI-251, Department of Computer Science, University of Toronto.

Peter A. Heeman and Graeme Hirst (1994), "Collaborating on referring expressions", MS, submitted for publication (available on request from the authors).

Diane Horton and Graeme Hirst (1991), "Discrepancies in discourse models and miscommunication in conversation", *Working notes of the AAAI fall symposium: Discourse structure in natural language understanding and generation*, Asilomar, pp. 31–32.

Aravind K. Joshi, Bonnie L. Webber, and Ivan Sag (editors) (1981). *Elements of discourse understanding* (Cambridge University Press, Cambridge, England).

Lynn Lambert and Sandra Carberry (1991), "A tripartite plan-based model for dialogue", *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, Berkeley, pp. 47–54.

Diane J. Litman and James F. Allen (1990), "Discourse processing and commonsense plans", in Cohen *et al*, 1990, pp. 365–388.

K. Lynch (1960), *The image of the city* (The MIT Press, Cambridge, MA).

Kathleen Filliben McCoy (1989), "Generating context-sensitive responses to object misconceptions", *Artificial Intelligence*, Vol. 41, No. 2, pp. 157–195.

Susan W. McRoy (1993), *Abductive interpretation and reinterpretation of natural language utterances*, PhD thesis, published as technical report CSRI-288, Department of Computer Science, University of Toronto.

Susan W. McRoy and Graeme Hirst (1993a), "Abductive explanation of dialogue misunderstandings", *Proceedings of the 6th conference of the European chapter of the Association for Computational Linguistics*, Utrecht, The Netherlands, pp. 277–286.

Susan W. McRoy and Graeme Hirst (1993b), "Misunderstanding and the negotiation of meaning", *Working notes of the AAAI fall symposium: Human–computer collaboration: Reconciling theory, synthesizing practice*, Raleigh, NC, pp. 57–62.

Susan W. McRoy and Graeme Hirst (1994), "The repair of speech act misunderstandings by abductive inference", MS, submitted for publication (available on request from the authors).

Gopalan Nadathur and Aravind K. Joshi (1983), "Mutual beliefs in conversational systems: Their role in referring expressions", *Proceedings of the 8th International Joint Conference on Artificial Intelligence (IJCAI-83)*, Karlsruhe, pp. 603–605.

C. Raymond Perrault and Philip R. Cohen (1981), "It's for your own good: A note on inaccurate reference", in Joshi, Webber, and Sag, 1981, pp. 217–230.

C. Raymond Perrault (1990), "An application of default logic to speech act theory", in Cohen *et al*, 1990, pp. 161–186.

Martha E. Pollack (1986), "A model of plan inference that distinguishes between the beliefs of actors and observers", *Proceedings of the 24th annual meeting of the Association for Computational Linguistics*, New York, pp. 207–214.

Martha E. Pollack (1990), "Plans as complex mental attitudes", in Cohen *et al*, 1990, pp. 77–103.

David Poole, Randy Goebel, and Romas Aleliunas (1987), "Theorist: A logical reasoning system for defaults and diagnosis", in *The knowledge frontier: Essays in the representation of knowledge*, edited by Nick Cercone, and Gordon McCalla (Springer-Verlag, New York), pp. 331–352.

George Psathas (1991), "The structure of direction-giving in interaction", in *Talk and social structure: Studies in ethnomethodology and conversation analysis*, edited by Deirdre Boden, and Don H. Zimmerman (Polity Press, Cambridge, England), pp. 195–216.

Ehud Reiter (1990), "The computational complexity of avoiding conversational implicature", *Proceedings of the 28th annual meeting of the Association for Computational Linguistics*, Pittsburgh, pp. 97–104.

Ehud Reiter and Robert Dale (1992), "A fast algorithm for the generation of referring expressions", *Proceedings of the 14th International Conference on Computational Linguistics (COLING-92)*, Nantes, pp. 232–238.

Emanuel A. Schegloff (1987), "Some sources of misunderstanding in talk-in-interaction", *Linguistics*, Vol. 25, pp. 201–218.

Emanuel A. Schegloff (1992), "Repair after next turn: The last structurally provided defense of intersubjectivity in conversation", *American Journal of Sociology*, Vol. 97, No. 5, pp. 1295–1345.

Emanuel A. Schegloff, Gail Jefferson, and Harvey Sacks (1977). "The preference for self-correction in the organization of repair in conversation", *Language*, Vol. 53, pp. 361–382.

Jan Svartvik and Randolph Quirk (1980), *A corpus of English conversation*, Lund Studies in English 56 (C.W.K. Gleerup, Lund).

A. Terasaki (1976), Pre-announcement sequences in conversation", Social Science Working Paper 99, School of Social Science, University of California, Irvine.

Paul Van Arragon (1990), *Nested default reasoning for user modeling*, PhD thesis, Department of Computer Science, University of Waterloo, published as Research Report CS-90-25.

Robert Wilensky (1983), *Planning and understanding: A computational approach to human reasoning* (Addison Wesley, Reading, MA).

David E. Wilkens (1985), "Recovering from execution errors in SIPE", *Computational Intelligence*, Vol. 1, No. 1, pp. 33–45.

Footnotes

[1]This paper summarizes work reported in greater detail by Edmonds (1993; 1994), Heeman (1991), Heeman and Hirst (1994), Horton and Hirst (1991), McRoy (1993), and McRoy and Hirst (1993a; 1994). Correspondence and requests for reprints should be sent to the first author.

[2]Susan McRoy is now at the Department of Electrical Engineering and Computer Science, University of Wisconsin–Milwaukee, Milwaukee, WI 53201, U.S.A.

[3]Peter Heeman is now at the Department of Computer Science, University of Rochester, Rochester, NY 14627, U.S.A.

[4]Misunderstanding should not be confused with *misconception*. A misconception is an error in the prior knowledge of an participant. McCoy (1989), Calistri-Yeh (1991), Pollack (1986; 1990), and others have studied the problem of how one participant can, in conversation, determine the misconceptions of another in order to correct them.

[5]The model is implemented in Prolog. Input and output are in the form of surface-level descriptions of speech acts; for example, `s-accept` represents the acceptance of a referring expression, and might be realized in a complete natural language system as *Okay*.

[6]We have not explored ambiguous situations: those in which more than one valid derivation remains, or, in the absence of validity, more than one invalid derivation.

[7]Another approach would have been to separate the communicative task from the first two (Lambert and Carberry, 1991).

[8]The present composition function is simple addition. One could envision more complex systems to compute confidence such as an algebra of confidence or a non-numeric system.

[9]In Heeman and Hirst's model, an attribute has to be mutually believed to be used. Here, mutual belief is not possible because the hearer has no knowledge of the referent, but mutual belief is an intended effect of using this plan.

[10]Recall that she raised her confidence threshold as a result of the hearer's postponement move, so now she must meet the new threshold.

[11]More precisely, $\mathcal{T} \cup \mathcal{B} \cup M$ satisfies the priority constraints of Prioritized Theorist.

[12] A related concern is how an agent's beliefs might change after an utterance has been understood as an act of a particular type. Although we have nothing new to add here, Perrault (1990) shows how default logic might be used to address this problem.

[13]Specifically, revision occurs when their plan inference scheme is unable to accommodate the next utterance without violating one of their constraints on well-formed dialogue.

Table 1: Surface linguistic actions for referring expressions.

`s-refer(Ent)`: Expresses the speaker's intention to refer to the object corresponding to the discourse entity `Ent`.

`s-attrib(Ent,Pred)` Used to describe the referent corresponding to the discourse entity `Ent` by means of the predicate `Pred`.

`s-attrib-rel(Ent,OEnt,Rel)`: Used to describe the object corresponding to `Ent` by means of its relationship `Rel` to the object corresponding to `OEnt`.

Table 2: Top-level actions for collaborating

`accept-plan(Plan)`: Expresses the speaker's judgment of acceptance of the referring plan `Plan`. Decomposes into the surface linguistic action `s-accept(Plan)`.

`reject-plan(Plan)`: Expresses the speaker's judgment that the referring plan includes inappropriate descriptors. Decomposes into the surface linguistic action `s-reject(Plan,Acts)`, where `Acts` is the list of inappropriate surface linguistic actions.

`postpone-plan(Plan)`: Expresses the speaker's judgment that the referring plan does not include enough descriptors. Decomposes into the surface linguistic action `s-postpone(Plan)`.

`replace-plan(Plan)`: Used by the speaker to replace part of the current referring plan. Decomposes into `s-actions(Plan,Acts)`, where `Acts` is the list of surface linguistic actions that are to be added to the referring plan (replacing the inappropriate ones).

`expand-plan(Plan)`: Used by the speaker to expand the current referring plan. Decomposes into `s-actions(Plan,Acts)`, where `Acts` is the list of surface linguistic actions that are to be added to the referring plan.

Table 3: Surface linguistic actions for collaborating.

`s-accept(Plan)`: Communicates acceptance of a plan. Could be realized as *Yes* or an emphatic *Okay*.

`s-reject(Plan,Acts)` Communicates rejection of components of a plan. Could be realized, for example, as *What weird thing?*.

`s-postpone(Plan)`: Communicates postponement of judgment on a plan. Could be realized as a tentative *Okay*.

`s-actions(Plan,Acts)`: Communicates an addition to components of a plan. Could be realized, for example, as *in the corner*.

Table 4: Example of referent negotiation.

1 **A:** See the weird creature.

```
s-refer(Entity)
s-attrib(Entity,λX·assessment(X,weird))
s-attrib(Entity,λX·category(X,creature))
```

2 **B:** In the corner?

```
s-postpone(p1)
s-actions(p1,[s-attrib-rel(Entity,Entity1,λX·λY·in(X,Y)),
              s-refer(Entity1),
              s-attrib(Entity1,λX·category(X,corner))])
```

3 **A:** No, on the television.

```
s-reject(p56,[s-attrib-rel(Entity,Entity1,λX·λY·in(X,Y))
              s-refer(Entity1),
              s-attrib(Entity1,λX·category(X,corner))])
s-actions(p56,[s-attrib-rel(Entity,Entity2,λX·λY·on(X,Y)),
               s-refer(Entity2),
               s-attrib(Entity2,λX·category(X,television))])
```

4 **B:** Okay.

```
s-accept(p142)
```

Table 5: A salience hierarchy. Higher values indicate greater salience.

```
salient-attribute(4, building,                     λX·architectural-style(X,Style)).
salient-attribute(3, building,                     λX·height(X,tall)).
salient-attribute(3, intersection,                 λX·called(X,Name)).
salient-attribute(2, intersection,sign,            λX·λY·has(X,Y)).
salient-attribute(2, intersection,traffic-lights,  λX·λY·has(X,Y)).
```

Table 6: Example of suggestion and elaboration.

1 **A:** Go to the Lowell Street intersection.

```
s-goto(Entity)
s-refer(Entity)
s-attrib(Entity,λX·category(X,intersection))
s-attrib(Entity,λX·called(X,'Lowell Street'))
```

2 **B:** Does it have a sign?

```
s-postpone(p1)
s-suggest(p1,[s-attrib-rel(Entity,Entity2,λX·λY·has(X,Y)),
              s-refer(Entity2),
              s-attrib(Entity2,λX·category(X,sign))])
```

3 **A:** Yes, it does, and it also has traffic lights.

```
s-affirm(p1,[s-attrib-rel(Entity,Entity2,λX·λY·has(X,Y)),
             s-refer(Entity2),
             s-attrib(Entity2,λX·category(X,sign))])
s-actions(p1,[s-attrib-rel(Entity,Entity3,λX·λY·has(X,Y)),
              s-refer(Entity3),
              s-attrib(Entity3,λX·category(X,traffic-lights))])
```

4 **B:** Okay.

```
s-accept(p123)
```

Table 7: Example of repair of misunderstanding.

1	**'Mother':**	Do you know who's going to that meeting? `surface-request(m, r, informif(r, m, knowref(r, whoIsGoing)))`
2	**'Russ':**	Who? `surface-request(r, m, informref(m, r, whoIsGoing))`
3	**'Mother':**	I don't know. `surface-inform(m, r, not knowref(m, whoIsGoing))`
4	**'Russ':**	Oh. Probably Mrs McOwen and probably Mrs Cadry and some of the teachers. `surface-informref(r, m, whoIsGoing)`