Generating Indirect Answers to Yes-No Questions

Nancy Green
Department of Computer Science
University of Delaware
Newark, DE 19716, USA
Internet: green@udel.edu

Sandra Carberry
Department of Computer Science
University of Delaware
Visitor: Inst. for Research in Cognitive Science
University of Pennsylvania
Internet: carberry@udel.edu

Abstract

An indirect answer to a Yes-No question conversationally implicates the speaker's evaluation of the truth of the questioned proposition. We present the approach to generation used in our implemented system for generating and interpreting indirect answers to Yes-No questions in English. Generation of a discourse plan is performed in two phases: content planning and plan pruning. During content planning, stimulus conditions are used to trigger speaker goals to include appropriate extra information with the direct answer. Plan pruning determines what parts of this full response do not need to be stated explicitly – resulting in, in appropriate discourse contexts, the generation of an indirect answer.

1. Introduction

Imagine a discourse context for (1) in which R's use of just (1d) is intended to convey a No, i.e., that R is *not* going shopping tonight. (By convention, square brackets indicate that the enclosed text was not explicitly stated.) The part of R's response consisting of (1d) - (1e) is what we call an *indirect answer* to a Yes-No question, and if (1c) had been uttered, (1c) would have been called a *direct answer*.

```
1.a. Q: I need a ride to the mall.
  b.     Are you going shopping tonight?
  c. R: [no]
  d.     My car's not running.
  e.     The timing belt is broken.
```

According to one study of spoken English [Ste84], 13 percent of responses to Yes-No questions were indirect answers. Thus, the ability to interpret indirect answers is required for robust dialogue systems. Furthermore, there are good reasons for generating indirect answers in a dialogue system. First, a direct *yes* or *no* alone may be misleading if extra information is needed to qualify the answer. Second, an indirect answer may contribute to a more efficient dialogue. For example, in addition to providing the requested information, it may anticipate a follow-up question from

Q, or it may allow R to respond immediately without asking for clarification of the question. That is, the increased efficiency is more the result of avoiding the follow-up question or clarification subdialogue than the result of omitting the direct answer. Third, an indirect answer may be preferable for politeness considerations, as in (1).

We have developed a computational model for the interpretation and generation of indirect answers to Yes-No questions in English. More precisely, by a *Yes-No question* we mean one or more utterances used as a request by Q (the questioner) that R (the responder) convey R's evaluation of the truth of a proposition p. An indirect answer implicitly conveys via one or more utterances R's evaluation of the truth of the questioned proposition p, i.e. that p is true, that p is false, that there is some truth to p, that p may be true, or that p may be false. Our model presupposes that Q's question has been understood by R as intended by Q, that Q's request was appropriate, and that Q and R are engaged in a cooperative goal-directed dialogue. Both the interpretation and generation components have been implemented in Common Lisp on a Sun SPARCstation.

A language user's *pragmatic knowledge* of how language is used to answer Yes-No questions in English can be used to constrain the problem of generating and interpreting indirect answers. This knowledge is encoded in our model as a set of domain-independent discourse plan operators and a set of coherence rules, described in section 2.[1] Our system is *reversible* in that the same pragmatic knowledge is used by the generation and interpretation modules. Generation is modeled as discourse plan construction and interpretation as discourse plan inference. The output of generation is a discourse plan which can be realized by a tactical generation component [McK85], and the output of interpretation is the speaker's inferred discourse plan. We found that while the above pragmatic knowledge is sufficient for interpretation, it is not sufficient for the

---

[1] Our main sources of data were previous studies [Hir85, Ste84], transcripts of naturally occurring two-person dialogue [Ame92], and constructed examples.

problem of content-planning during generation. This paper describes our approach to generation, including our solution to this problem.

Generation of a discourse plan, described in section 4, is performed in two phases: content planning and plan pruning. During content planning, *stimulus conditions*, described in section 3, are used to trigger new speaker goals to include appropriate extra information with the direct answer. Plan pruning determines what parts of this full response do not need to be stated explicitly. In appropriate discourse contexts, a plan for an indirect answer is generated. In other contexts, i.e. those in which the direct answer must be given explicitly, a plan for a direct answer with appropriate extra information is generated. Note that, according to the study mentioned earlier [Ste84], 85 percent of direct answers are accompanied by such information. Thus, it is important to model this type of response as well. Section 5 discusses other factors in generating indirect answers, and section 6 surveys related work.

2. Pragmatic Knowledge

A language user's knowledge of how to answer a Yes-No question appropriately can be described by a set of discourse plan operators representing *full answers*. A full answer consists of a direct answer (which we call the *nucleus*) and, possibly, extra relevant information (*satellites*). An indirect answer can be modeled as the result of R's providing one or more satellites without expressing the nucleus explicitly. Due to coherence constraints holding between nucleus and satellite, Q may infer R's discourse plan by inferring which coherence relation between an indirect answer and a candidate direct answer is plausible. The coherence relations are similar to the subject-matter relations of Rhetorical Structure Theory (RST) [MT87].[2]

The discourse plan operators used to generate and interpret R's response in (1) are given in Figures 1 and 2. Figure 1 depicts the top-level operator for constructing a No answer. To explain our notation, s and h are constants denoting speaker (R) and hearer (Q), respectively. Symbols prefixed with "?" denote propositional variables. The variable in the header of *Answer-no* will be instantiated with the questioned proposition, e.g. in (1) that R is going shopping. Applicability conditions are necessary conditions for appropriate use of the operator. For example, for a No answer to be appropriate, R must believe that the questioned proposition p is false. Also, an answer is not appropriate unless s and h share the expectation

---

[2]The terms *nucleus* and *satellite* have been borrowed from RST. Note that according to RST, a property of the nucleus is that its removal results in incoherence. However, in our model, a direct answer may be removed without causing incoherence, provided that it is inferrable from the rest of the response.

```
(Answer-no s h ?p):
  Applicability conditions:
    (discourse-expectation (informif s h ?p))
    (believe s (not ?p))
  Nucleus:
    (inform s h (not ?p))
  Satellites:
    (Use-otherwise s h (not ?p))
    (Use-obstacle s h (not ?p))
    (Use-contrast s h (not ?p))
  Primary goals:
    (BMB h s (not ?p))
```

Figure 1: Discourse plan operator for No answer

```
(Use-obstacle s h ?p):
;; s tells h of an obstacle explaining
;;   the failure ?p
  Existential variable: ?q
  Stimulus conditions:
    (explanation-indicated s h ?p ?q)
    (excuse-indicated s h ?p ?q)
  Applicability conditions:
    (believe s (cr-obstacle ?q ?p))
    (Plausible (cr-obstacle ?q ?p))
  Nucleus:
    (inform s h ?q)
  Satellites:
    (Use-elaboration s h ?q)
    (Use-obstacle s h ?q)
    (Use-cause s h ?q)
  Primary goals:
    (BMB h s (cr-obstacle ?q ?p))
```

Figure 2: Discourse plan operator for Obstacle

that s will provide s's evaluation of the truth of p, which is denoted as *(discourse-expectation (informif s h p))*. Primary goals describe the intended effects of the operator. We use *(BMB h s p)* to denote that h believes it to be mutually believed with s that p [CM81].

In general, the nucleus and satellites of a discourse plan operator describe primitive or non-primitive communicative acts. Our plan formalism allows zero, one, or more occurrences of a satellite in a full answer. The expected (but not required) order of nucleus and satellites is the order they are listed in an operator. *(inform s h p)* denotes the primitive act of s informing h that p. The satellites in Figure 1 refer to non-primitive acts, described by discourse plan operators which we have defined (one for each coherence relation used in a full answer). For example, *Use-obstacle*, a satellite of *Answer-no*, is defined in Figure 2.

To explain the additional notation in Figure 2, *(cr-obstacle q p)* denotes that the coherence relation

190

```
Operator:         Satellites:
-------------     -----------
Answer-yes        Use-condition
                  Use-elaboration
                  Use-cause
Answer-no         Use-otherwise
                  Use-obstacle
                  Use-contrast
Answer-hedged     Use-contrast
Answer-maybe      Use-result
                  Use-usually
                  Use-possible-cause
Answer-maybe-not  Use-result
                  Use-usually
                  Use-possible-obstacle
```

Table 1: Satellites of top-level operators

cr-$obstacle$ holds between q and p.[3] Thus, the first applicability condition can be glossed as requiring that s believe that the coherence relation holds. In the second applicability condition, *(Plausible (cr-obstacle q p))* denotes that, given what s believes to be mutually believed with h, the coherence relation *(cr-obstacle q p)* is plausible. In our model, this sort of condition is evaluated using a set of coherence rules based on the relation definitions of RST.

In summary, we provide our system with a set of discourse plan operators representing shared domain-independent knowledge of the informational content of a full answer. (Operators whose names begin with *Answer* are top-level operators.) Table 1 shows the satellites of each top-level discourse plan operator. In addition, we provide a set of coherence rules defining the plausibility of the coherence relations generated by the satellite discourse plan operators.[4]

3. Stimulus Conditions

Applicability conditions prevent inappropriate use of a discourse plan operator. However, they do not model a speaker's motivation for choosing to provide extra information. A full answer might provide too much information if every satellite whose applicability conditions held were included in the full answer. On the other hand, at the time when he is asked a question, R may not have the primary goals of a potential satellite operator. Thus, a goal-driven approach to selecting satellites would provide insufficient information. To overcome these limitations, we have incorporated *stimulus conditions* into the discourse plan operators in our

---

[3]Although this is not one of the original relations of RST, it is similar to other causal relations defined in RST.

[4]In addition to RST subject-matter relations, we have identified the following coherence relations in our corpus: cr-obstacle, cr-possible-cause, cr-possible-obstacle, and cr-usually.

model. Playing a key role in content determination, stimulus conditions describe conditions motivating a speaker to include a satellite during plan construction. They can be thought of as triggers which give rise to new speaker goals. In order for a satellite to be included, all of its applicability conditions and at least one of its stimulus conditions must be true. While stimulus conditions may be derivative of deeper principles of cooperativity [Gri75a] or politeness [BL78], they provide a level of precompiled knowledge which reduces the amount of reasoning required for content determination.

Our methodology for identifying stimulus conditions was to survey linguistic studies (described below), as well as to analyze the possible motivation of the speaker in the examples in our corpus. According to Stenström [Ste84], the typical motivation for providing extra information is to answer an implicit wh-question. Other reasons cited are to provide an explanation justifying a negative answer, to qualify the answer, or "social reasons". Levinson [Lev83] claims that the presence of an explanation is a structural characteristic of dispreferred responses, such as refusals and unexpected answers. Brown and Levinson [BL78] show how various uses of language are motivated by politeness. For example, use of conversational implicature is one strategy for a speaker to avoid doing a face-threatening act,[5] and as we discuss in section 6, an indirect answer conversationally implicates the direct answer. Hirschberg [Hir85] claims that an indirect scalar response, rather than a Yes or No answer, is necessary in certain cases to avoid misleading the hearer.

In the rest of this section, we describe each of the stimulus conditions used in our system. For each stimulus condition, we give one or more axioms defining sufficient conditions for it. Axioms are encoded as Horn clauses, where symbols prefixed with "?" are variables, all variables are implicitly universally quantified, and the antecedent conditions are implicitly conjoined. The head, i.e. the consequent, of an axiom is a stimulus condition appearing in one or more of our discourse plan operators. Table 2 summarizes which stimulus conditions appear in which operators. For example, the stimulus condition *(explanation-indicated s h ?p ?q)* occurs in the discourse plan operator shown in Figure 2. Keep in mind that additional constraints on *?p* and *?q* arise from the applicability conditions of the operator containing the stimulus condition.

3.1 Definitions

Explanation-indicated

This stimulus condition appears in all of the oper-

---

[5]They claim that another strategy is to give the direct answer last, i.e. in our terminology, to give the satellite(s) before the nucleus.

Operator:	Stimulus conditions:
Use-cause	Explanation-indicated
Use-condition	Clarify-condition-indicated
Use-contrast	Appeasement-indicated
	Answer-ref-indicated
	Clarify-extent-indicated
	Substitute-indicated
Use-elaboration	Answer-ref-indicated
	Clarify-concept-indicated
Use-obstacle	Explanation-indicated
	Excuse-indicated
Use-otherwise	Explanation-indicated
	Excuse-indicated
Use-possible-cause	Explanation-indicated
Use-possible-obstacle	Explanation-indicated
	Excuse-indicated
Use-result	Explanation-indicated
Use-usually	Explanation-indicated

Table 2: Stimulus conditions of discourse plan operators

ators for providing causal explanations. For example in (2),[6] which indirectly conveys a No, R gives an explanation of why R won't get a car.

2.a. Q: actually you'll probably get a car
   won't you as soon as you get there
  b. R: can't drive

Note that this stimulus condition may contribute to greater dialogue efficiency by anticipating a follow-up request for an explanation. Formally, the condition is defined by the following axiom.

```
((explanation-indicated s h ?p ?q)
<-
(wbel s (wbel h (not ?p)))
(unless (wbel s (accepts-authority h s)))))
```

This may be glossed as, s is motivated to give h an explanation q for p, if s suspects that h suspects that p is not true, unless it is the case that s has reason to believe that h will accept p on s's authority. *(wbel agent p)*, which we usually gloss as *agent* suspects that p, denotes that *agent* has some confidence in the belief that p. Note that R may acquire the suspicion that Q doubts that p is true by means of syntactic clues from the Yes-No question, e.g. the form of the question in (2a).

Excuse-indicated

Although this stimulus condition appears in some of the same causal operators as *Explanation-indicated*,

---

[6]Stenström's (110)

it represents a different kind of motivation. A Yes-No question may be interpreted as a prerequest [Lev83] for a request, i.e. as an utterance used as a preface to another request.[7] Prerequests are often used to check whether a related request is likely to succeed, or to avoid having to make the other request directly. Thus, a negative answer to a Yes-No question used as a prerequest may be interpreted as a refusal. To soften the refusal, the speaker may give an explanation of the negative answer, as illustrated in (1). Formally, the condition is defined by the following axiom.

```
((excuse-indicated s h (not ?p) ?q)
<-
 (wbel s (prerequest h s (informif s h ?p)))))
```

This may be glossed as, s is motivated to give h an excuse q for *(not p)*, if s suspects that h's request, *(informif s h p)*, is a prerequest. Techniques for interpreting indirect speech acts [PA80, Hin89] can be used to determine whether the antecedent holds.

Answer-ref-indicated

This condition appears in *Use-Elaboration*, illustrated by (3),[8] and in *Use-Contrast*, illustrated by (4).[9]

3.a. Q: Did you have a hotel in mind?
  b.   [What hotel did you have in mind?]
  c. R: [yes]
  d.   There's a Holiday Inn right near
        where I'm working.

4.a. Q: You're on that?
  b.   [Who's on that?]
  c. R: no no no
  d.   Dave is.

In (3), R has interpreted the question in (a) as a prerequest for the wh-question shown in (b). Thus, (d) not only answers the question in (a) but also the anticipated wh-question in (b). Similarly in (4), R may interpret the question in (a) as a prerequest for the wh-question in (b), and so gives (d) to provide an answer to both (a) and (b). Formally, the condition is defined by the following axiom.

```
((Answer-ref-indicated s h ?p ?q)
<-
 (wbel s (want h (knowref h ?t ?q)))
```

This may be glossed as, s is motivated to provide h with q, if s suspects that h wants to know the referent

---

[7]In analyses of discourse based on speech act theory, e.g. [PA80], a question such as (1b) might be described as an indirect request. However, since it may be responded to appropriately by a Yes or No, it is included in the class of Yes-No questions addressed in our model.

[8]American Express tape 1

[9]Stenström's (102). Note that a No answer may be conversationally implicated by use of (4d) alone.

of a term t in q. As in *Excuse-indicated*, techniques for interpreting indirect speech acts can be used to determine if the antecedent holds.

Substitute-indicated

This condition appears in *Use-contrast*, illustrated by (5).

```
5.a.Q: Do you have Verdi's Otello or Aida?
  b.R: [no]
  c.    We have Rigoletto.
```

Although Q may not have intended to use (5a) as a prerequest for the question *What Verdi operas do you have?*, R suspects that the answer to this wh-question might be helpful to Q. Formally, the condition is defined by the following axiom.

```
((Substitute-indicated s h ?p ?q)
<-
 (wbel s (need h (knowref h ?t ?q))))
```

This may be glossed as, s is motivated to provide h with q, if s suspects that it would be *helpful* for h to know the referent of a term t in q. The antecedent would hold whenever obstacle detection techniques [AP80] determine that h's not knowing the referent of t is an obstacle to an inferred plan of h's. However, note that not all helpful responses, in the sense described in [AP80], can be used as indirect answers. For example, even if the clerk (R) at the music store believes that Q's not knowing the closing time could be an obstacle to Q's buying a recording, the response (6), given instead of (5c), would not convey (5b) since it cannot be coherently related to (5b).

```
6. R: We close at 5:30 tonight.
```

Clarify-concept-indicated

This stimulus condition appears in *Use-elaboration*, as illustrated by (7).[10]

```
7. Q: Do you have a pet?
   R: We have a turtle.
```

In (7), R was motivated to elaborate on the type of pet they have, since turtles are not prototypical pets. Formally, the condition is defined by the following axiom.

```
((Clarify-concept-indicated s h ?p ?q)
<-
 (concept ?p ?c)
 (has-atypical-instance ?q ?c))
```

This may be glossed as, s is motivated to clarify p to h with q, if p has a concept c, and q provides an atypical instance of c. Stereotypical knowledge is used to evaluate the second antecedent.

[10] Hirschberg's (177).

Clarify-condition-indicated

This stimulus condition appears in the operator *Use-condition*, as illustrated by (8).[11]

```
8.a. Q: Um let me can I make the reservation
        and change it by tomorrow
  b. R: [yes]
  c.    if it's still available.
```

In (8), a truthful Yes answer depends on the truth of (c). Formally, the stimulus condition is defined by the following axioms.

```
((Clarify-condition-indicated s h ?p ?q)
<-
 (ignorant s ?q))
```

```
((Clarify-condition-indicated s h ?p ?q)
<-
 (wbel s (not ?q)))
```

These may be glossed as, s is motivated to clarify a condition q for p to h 1) if s doesn't know if q holds, or 2) if s suspects that q does not hold.

Clarify-extent-indicated

This stimulus condition appears in *Use-contrast*, as illustrated by (9).[12]

```
9.a. Q: Have you gotten the letters yet?
  b. R: I've gotten the letter from X.
```

On the strict interpretation of (9a), Q is asking whether R has gotten *all* of the letters, but on a looser interpretation, Q is asking if R has gotten *any* of the letters. Then, if R has gotten some but not all of the letters, a Yes would be untruthful. However, if Q is speaking loosely, then a No might lead Q to erroneously conclude that R has not gotten any of the letters. R's answer circumvents this problem, by conveying the extent to which the questioned proposition (on the strict interpretation) is true. Formally, the condition is defined by the following axioms.

```
((clarify-extent-indicated s h
          (some-truth ?p) ?q)
<-
 (wbel s (ignorant h ?q))
 (believe s (highest-true-exp ?q ?p)))
```

```
((clarify-extent-indicated s h (not ?p) ?q)
<-
 (wbel s (ignorant h ?q))
 (believe s (highest-true-exp ?q ?p)))
```

These may be glossed as, s is motivated to clarify to h the extent q to which p is true, or the alternative q to p which is true, if s suspects that h does not know if q

[11] American Express tape 10ab
[12] Hirschberg's (59)

holds, and s believes that q is the highest expression alternative to p that does hold. According to Hirschberg [Hir85] (following Gazdar), sentences p_i and p_j (representing the propositional content of two utterances) are expression alternatives if they are the same except for having comparable components e_i and e_j, respectively. Further, Hirschberg claims that in a particular discourse context, there may be a partial ordering of values which the discourse participants mutually believe to be salient, and that the ranking of e_i and e_j in this ordering can be used to describe the ranking of p_i and p_j. In the above example, (9b) is a realization of the highest true expression alternative to the questioned proposition, p, i.e. the proposition that R has gotten all the letters.[13]

Appeasement-indicated

This stimulus condition appears in *Use-contrast*, as illustrated by (10).[14]

```
10.a. Q: Did you manage to read that section
         I gave you?
    b. R: I've read the first couple of pages.
```

In (10), R conveys that there is some (though not much) truth to the questioned proposition in an effort to soften his answer. Note that more than one stimulus condition may motivate R to include the same satellite. E.g., in (10), R may have been motivated also by *clarify-extent-indicated*, which was described above. However, it is possible to provide a context for (9) where *appeasement-indicated* holds but not *clarify-extent-indicated*, or a context where the converse is true. Formally, the condition is defined by the following axiom.

```
((appeasement-indicated s h (not ?p) ?q)
<-
 (wbel s (undesirable h (not ?p)))
 (wbel s (desirable h ?q)))

((appeasement-indicated s h
            (some-truth ?p) ?q)
<-
 (wbel s (undesirable h (not ?p)))
 (wbel s (desirable h ?q)))
```

This may be glossed as, s is motivated to appease h with q for p not holding or only being partly true, if s suspects that *(not p)* is undesirable to h but that q is desirable to h. The antecedents to this rule would be evaluated using heuristic rules and stereotypical and specific knowledge about h's desires. For example, two

heuristics of rational agency that might lead to beliefs about h's desires are 1) if an agent wants you to perform an action A, then your failure to perform A may be undesirable to the agent, and 2) if an agent wants you to do A, then it is desirable to the agent that you perform a part of A.

4. Generation Algorithm

The inputs to generation in our model consist of 1) a set of discourse plan operators augmented with stimulus conditions, 2) a set of coherence rules, 3) R's beliefs, and 4) the discourse expectation that R will provide R's evaluation of the truth of the questioned proposition p. The output of the generation algorithm is a discourse plan which can be realized by a tactical generation component [McK85]. We assume that when answer generation begins, the speaker's only goal is to satisfy the above discourse expectation.[15]

Our answer generation algorithm has two phases. In the first phase, *content planning*, the generator creates a discourse plan for a full answer, i.e., a direct answer and extra appropriate information, e.g. (1c) given explicitly, followed by (1d) - (1e). In the second phase, *plan pruning*, the generator determines which propositions of the planned full answer do not need to be explicitly stated. For example, given an appropriate model of R's beliefs, our system generates a plan for asserting only the proposition conveyed in (1e) as an answer to (1a) - (1b).[16] Note that an advantage of our approach is that, even when it is not possible to omit the direct answer, a full answer is generated.

4.1 Content planning

Content planning is performed by top-down expansion of an answer discourse plan operator. The process begins by instantiating each of the operators with the questioned proposition until one is found such that its applicability conditions hold. Next, this (top-level) answer operator is expanded. A discourse plan operator is expanded by deciding which of its satellites to include in the full answer and expanding each of them (recursively). A satellite (e.g. *Use-obstacle* in Figure 2) is selected if, for some instantiation of its existential variable, all of the applicability conditions and *at least one* of the stimulus conditions of the instantiated operator hold.

Note that testing of applicability conditions and stimulus conditions (including searching for an appro-

---

[13]Recall that additional constraints on p and q arise from the applicability conditions of operators containing this stimulus condition, namely *Use-contrast* in this case. Thus, another constraint is that it is plausible that *cr-contrast* holds. Note that *cr-contrast* also is defined in terms of such an ordering.

[14]Hirschberg's (56)

[15]The plan that is output by our algorithm specifies an ordering of discourse acts based upon the ordering of coherence relations specified in the discourse plan operators. However, to model a speaker who has more than the above initial goal, reordering may be required.

[16]The tactical component must choose an appropriate referring expression to refer to R's car's timing belt, depending on whether (1d) is omitted.

priate instantiation of the existential variable) is handled by a theorem prover. Our model of R's beliefs, represented as a set of Horn clauses, includes 1) general world knowledge presumably shared with Q, 2) knowledge about the preceding discourse, and 3) R's beliefs (including "weaker beliefs" encoded using the *wbel* operator) about Q's beliefs. Much of the shared world knowledge needed to evaluate our coherence rules consists of knowledge from domain plan operators.

4.2 Plan pruning

The output of the content planning phase, an expanded discourse plan representing a full answer, is the input to the plan pruning phase of generation. The expanded plan is represented as a tree of discourse acts. The goal of this phase is to make the response more concise,[17] i.e., to determine which of the planned acts can be omitted while still allowing Q to infer the full discourse plan. To do this, the generator considers each of the acts in the frontier of the tree from right to left. (This ensures that a satellite is considered before its related nucleus.) The generator creates a trial plan consisting of the original plan minus the nodes pruned so far and minus the current node. Then, using the interpretation module, the generator simulates Q's interpretation of a response containing the information that would be given explicitly according to the trial plan. If Q could infer the full plan (as the most preferred interpretation), then the current node can be pruned. Otherwise, it is left in the plan and the next node is considered.

For example, consider Figure 3 as we illustrate the possible effect of pruning on a full discourse plan. The leaf nodes, representing discourse acts, are numbered 1 - 8. Arcs labelled *N* and *S* lead to a nucleus or satellite, respectively. Node 8 corresponds to the direct answer. Plan pruning would process the nodes in order from 1 to 8. The maximal set of nodes which could be pruned in Figure 3 is the set containing 2, 3, 4, 7, and 8. That is, nodes 2 - 4 might be inferrable from 1, node 7 from 5 or 6, and node 8 from 4 or 7. In the event that it is determined that no node can be pruned, the full plan would be output. Note that our interpretation algorithm (described in [GC94]) performs a subprocess, *hypothesis generation*, to recognize missing propositions other than the direct answer, i.e. the propositions at nodes 2, 3, 4, and 7.

4.3 Example

This section gives an example of how the discourse plan shown in Figure 4 would be generated, corresponding to the answer shown in (11), where only (11d) is explicitly stated.
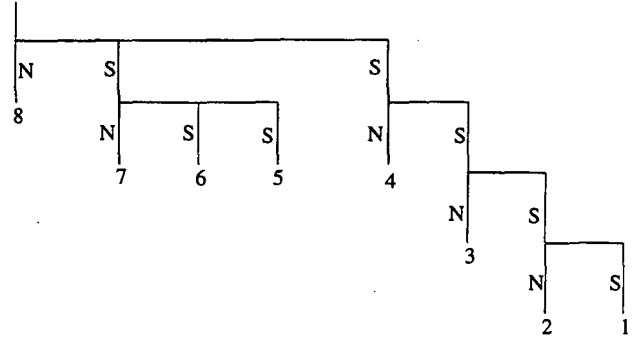

Figure 3: Example of full discourse plan before pruning

```
11. a.  Q: You went to the party, didn't you?
    b.  R: [no]
    c.      [The baby sitter could not sit.]
    d.      The baby sitter was sick.
```

In phase one, the generator creates a discourse plan for a full answer. First, it must decide which top-level discourse plan operator to expand. Each of the top-level operators would be instantiated with the questioned proposition, *(occur (go-party s) past)*. The top-level operators would be instantiated and tested until one was found whose applicability conditions held, namely, *Answer-no*. Next, the generator must decide which of the satellites of *Answer-no* to include. So, it would search for all propositions *q* such that, when the existential variable *?q* of a satellite is instantiated with *q*, then all of the satellite's applicability conditions and at least one of its stimulus conditions hold. In this example, the proposition *(not (in-state (can-sit sitter) past))* would be found to satisfy the applicability conditions and the *explanation-indicated* stimulus condition of *Use-obstacle*. (Note that this proposition might be chosen because the model of R's beliefs includes a domain plan operator for going to a party, with a precondition that the agent's sitter can baby sit.)

Next, this instantiation of *Use-obstacle* must be expanded, i.e., the generator must decide which of the satellites of *Use-obstacle* to include. By a procedure similar to that described above, it would decide to include (another) *Use-obstacle*, where the existential variable is instantiated with the proposition *(in-state (sick sitter) past)*. (Note that this proposition might be chosen because the model of R's beliefs includes the shared belief that being sick typically renders a sitter unfit to perform his or her duties.) The stimulus condition *explanation-indicated* might be the motivation for including this satellite. (For example, suppose it is shared knowledge that R's baby sitter is rarely unavailable.) If further attempts to expand the plan were unsuccessful, then the full plan would contain the discourse acts shown in Figure 4. (The acts in the plan are labelled (b) - (d) corresponding to the part of (11) they represent.) Note that if this plan were output

```
(Answer-no s h (occur (go-party s) past)):
   Nucleus:
b.   (inform s h (not (occur (go-party s)
                               past)))
     Satellites:
     (Use-obstacle s h
           (not (occur (go-party s) past)))
     Nucleus:
c.      (inform s h
            (not (in-state (can-sit sitter)
                           past)))
        Satellites:
        (Use-obstacle s h
            (not (in-state (can-sit sitter)
                           past)))
        Nucleus:
d.          (inform s h (in-state (sick sitter)
                               past))
```

Figure 4: Discourse plan for (11)

without undergoing plan-pruning, it would be realized as a direct answer (11b), and extra information (11c) - (11d).

In phase two, plan-pruning, the generator's overall goal is to make the answer as concise as possible. Therefore, it would consider the primitive acts specified in the full plan in the order (d), (c), (b). Since (d) is not inferrable from any other act, it would be automatically retained in the plan. Next, a trial plan with (c) pruned from it would be considered. Interpretation of the response consisting of (b) and (d) is simulated. Since the full plan could be inferred, (c) would be pruned. Next, another trial plan would be considered, where, in addition to (c), (b) has been pruned. Since the full plan could be inferred from the resulting trial response consisting of just (d), (b) would be pruned too. Thus, the output of phase two would be the plan shown in Figure 4 with (b) and (c) marked as pruned.

5. Other factors in generation

The primary focus of our research has been on 1) use of pragmatic knowledge which is common to both the interpretation and generation of indirect answers, i.e. the so-called reversible knowledge, and 2) identifying stimulus conditions. In this section, we give a brief description of some other factors in generating indirect answers.

First, note that we made the simplifying assumption that R's only initial goal is to answer the question (accurately, efficiently, and politely). However, it is possible for R to construct an indirect answer which satisfies multiple initial goals of R. For example, R might have decided to give the answer in (12c) not only as an explanation, but also as background for R's

request in (12d).

```
12.a. Q: Are you going to the lecture?
   b. R: [no]
   c.    I have to get my brakes fixed.
   d.    Can you recommend a garage near
         campus?
```

Second, stylistic goals may affect the generation of indirect answers. In addition to affecting syntactic and lexical aspects of discourse [Hov90, MG91, DH93], stylistic goals may affect which and how much information is given. For example, elaboration can be used to make an answer more lively, as shown in (13).

```
13.a. Q: Do you have a car?
   b. R: [yes]
   c.    I bought a British-racing-green
         Austin-Healey 3000 last week.
```

Also, extra information that entails the direct answer (e.g. repetitions and generalizations) may be given for emphasis. Further, a stylistic goal of terseness may override the stimulus conditions we have provided for indirect answers. For example, depending upon the degree of Q's interest in R's affairs, the indirect answer in (14c) - (14f) may provide Q with more than Q cares to know.

```
14.a. Q: Are you going to the movies tonight?
   b. R: [no]
   c.    I can't afford to.
   d.    I spent $900 on my car last week.
   e.    It needed a new transmission.
   f.    The car is 15 years old.
```

Third, certain syntactic forms for expressing Yes-No questions require more than just *yes/no* to make clear whether R is confirming or disconfirming that a proposition p holds [Ste84]. For example,[18] (15) shows the use of amplification (i.e. supplying an auxiliary verb) in answer to a negative-polarity, positive-bias question.

```
15.a. Q: Didn't Ann get the letter?
   b. R: Yes she did.
   c. R: No she didn't.
```

(A question is said to have positive or negative polarity depending on the presence or absence of negation, respectively. Bias describes the questioner's presumed belief in the truth of the questioned proposition.) In answer to (16a), a negative-polarity, negative-bias question,[19] R provides an indirect answer, giving the extent to which the proposition that R has been up-

---

[18](15) and (16) are Stenström's (105) and (111), respectively.

[19]Although it is expressed as a declarative sentence, Stenström classifies (16a) as a request for confirmation, which we treat as a type of Yes-No question.

stairs is true, in order to avoid the same type of confusion.

16.a. Q: I don't think you've been upstairs yet.
   b. R: Um only just to the loo.

6. Related research

Our work differs from most previous work in cooperative response generation in that, in our model, the information given in an indirect answer conversationally implicates [Gri75a] the direct answer. Hirschberg [Hir85], who provided general rules for identifying when a scalar conversational implicature is licensed, claimed that speakers may give indirect responses to Yes-No questions in order to block potential incorrect scalar implicatures of a simple Yes or No. She implemented a system which determines whether the Yes/No[20] licenses any unwanted scalar implicatures, and if so, proposes alternative scalar responses which do not. This type of response is similar, in our model, to a speaker's use of *Use-contrast* motivated by *clarify-extent-indicated*, as illustrated in (9). (As noted earlier, our coherence rules for the relation *cr-contrast* as well as our axioms for *clarify-extent-indicated* make use of the notion of a salient partial ordering which was elucidated by Hirschberg.) However, Hirschberg's model does not account for quite a variety of types of indirect answers which can be generated using the other operators in our model, nor for other motives for using *Use-contrast*.

Rhetorical or coherence relations [Gri75b, Hal76, MT87] have been used in several text-generation systems to aid in ordering parts of a text (e.g.[Hov88]) as well as in content-planning (e.g. [McK85, MP93]). The discourse plan operators based on coherence relations in our model (i.e. the operators used as satellites of top-level operators) play a similar role in content-planning of an indirect answer. When coherence relations are used for content-planning, it is necessary to constrain the information which thereby may be selected. McKeown [McK85] uses discourse focus constraints. In [Moo89], plan selection heuristics are used that maximize the hearer's presumed familiarity with the concepts in the text, prefer general-purpose to less general operators, and minimize verbosity. Maybury's [May92] system uses "desirable" preconditions, preconditions that are not necessary preconditions, to prioritize alternative operators. In contrast to the above, by the use of stimulus conditions, our model is able to trigger new, opportunistic speaker goals.

Moore and Pollack [MP92] show the need to distinguish the intentional and informational structure of discourse, where the latter is characterized by the

---

[20]Hirschberg did not addresss other possible types of direct answers, represented by the Answer-Hedged, Answer-Maybe, and Answer-Maybe-Not operators in our model.

---

sort of relations classified as subject-matter relations in RST. Note that in our interpretation component [GC94], informational relations (i.e. plausible coherence relations holding between indirect answers and candidate direct answers) are used to infer the speaker's goal to convey a particular answer. Moore and Paris [MP93] argue that it is necessary for generation systems to represent not only the speaker's top-level intentional goal, but also the intentional subgoals that a speaker hoped to achieve by use of a rhetorical relation so that, if a subgoal is not achieved, then an alternative rhetorical means can be tried. In our model, the operators used as satellites of top-level answer discourse plan operators are based on RST's subject-matter relations. The primary goals of these operators are similar to the effect fields of the corresponding RST relation definitions. As Moore and Paris argue, goals based on these relation definitions are rhetorical goals, not intentional goals. However, in our model, it is possible to determine the intentional subgoals by checking which stimulus conditions motivated the speaker to include a particular satellite operator. For example, if *explanation-indicated* motivates R to include a *Use-cause* satellite, then one can view R as having the implicit intentional goal of giving an explanation. Thus, our model provides the information necessary to recover from the possible failure of (all or part of) an indirect answer.

Finally, our use of simulated interpretation during plan pruning has some similarity to previous work. In Horacek's approach to generating concise explanations [Hor91], a set of propositions representing the full explanation is pruned by eliminating propositions which can be derived from the remaining ones by contextual rules. Jameson and Wahlster [JW82] use an anticipation feedback loop algorithm to generate elliptical utterances.

7. Conclusion

An indirect answer to a Yes-No question conversationally implicates the speaker's evaluation of the truth of a questioned proposition. The generation of indirect answers is important if a dialogue system is to respond efficiently, accurately, and politely. This paper has presented the approach to generation used in our implemented system for generating and interpreting indirect answers to Yes-No questions in English. Ours is the first system to generate a wide range of types of indirect answers (as well as full answers). The system is reversible in that the same pragmatic knowledge is used in generation and interpretation. Generation is performed in two phases: content planning and plan pruning. During content planning, stimulus conditions are used to trigger speaker goals to include appropriate extra information with the direct answer. Plan pruning determines what parts of this full response do not need to be stated explicitly – resulting in, in appropriate discourse contexts, the generation of an indirect

answer.

References

[AP80] James F. Allen and C. Raymond Perrault. Analyzing intention in utterances. *Artificial Intelligence*, 15:143–178, 1980.

[Ame92] American Express tapes. Transcripts of audiotape conversations made at SRI International, Menlo Park, California. Prepared by Jacqueline Kowto under the direction of Patti Price.

[BL78] Penelope Brown and Stephen Levinson. Universals in language usage: Politeness phenomena. In Esther N. Goody, editor, *Questions and politeness: Strategies in social interaction*, pages 56–289. Cambridge University Press, Cambridge, 1978.

[CM81] H. Clark and C. Marshall. Definite reference and mutual knowledge. In A. K. Joshi, B. Webber, and I. Sag, editors, *Elements of discourse understanding*. Cambridge University Press, Cambridge, 1981.

[DH93] Chrysanne DiMarco and Graeme Hirst. A computational theory of goal-directed style in syntax. *Computational Linguistics*, 19(3), 1993.

[GC94] Nancy Green and Sandra Carberry. A hybrid reasoning model for indirect answers. To appear in *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, 1994.

[Gri75a] H. Paul Grice. Logic and conversation. In P. Cole and J. L. Morgan, editors, *Syntax and Semantics III: Speech Acts*, pages 41–58, New York, 1975. Academic Press.

[Gri75b] J. E. Grimes. *The Thread of Discourse*. Mouton, The Hague, 1975.

[Hal76] M. Halliday. *Cohesion in English*. Longman, London, 1976.

[Hin89] Elizabeth Ann Hinkelman. *Linguistic and Pragmatic Constraints on Utterance Interpretation*. PhD thesis, University of Rochester, 1989.

[Hir85] Julia Bell Hirschberg. *A Theory of Scalar Implicature*. PhD thesis, University of Pennsylvania, 1985.

[Hor91] Helmut Horacek. Exploiting conversational implicature for generating concise explanations. In *Proceedings*. European Association for Computational Linguistics, 1991.

[Hov88] Eduard H. Hovy. Planning coherent multisentential text. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, pages 163–169, 1988.

[Hov90] Eduard H. Hovy. Pragmatics and natural language generation. *Artificial Intelligence*, 43:153–197, 1990.

[JW82] Anthony Jameson and Wolfgang Wahlster. User modelling in anaphora generation: ellipsis and definite description. In *Proceedings of the European Conference on Artificial Intelligence*, 1982.

[Lev83] S. Levinson. *Pragmatics*. Cambridge University Press, Cambridge, 1983.

[May92] Mark T Maybury. Communicative acts for explanation generation. *International Journal of Man-Machine Studies*, 37:135–172, 1992.

[McK85] Kathleen R. McKeown. *Text Generation*. Cambridge University Press, 1985.

[MG91] Marzena Makuta-Giluk. A computational rhetoric for syntactic aspects of text. Technical Report CS-91-56, Master's thesis, Department of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, December 1991.

[Moo89] Johanna D. Moore. *A Reactive Approach to Explanation in Expert and Advice-Giving Systems*. PhD thesis, University of California at Los Angeles, 1989.

[MP92] Johanna D. Moore and Martha E. Pollack. A problem for rst: The need for multi-level discourse analysis. *Computational Linguistics*, 18(4):537–544, December 1992.

[MP93] Johanna D. Moore and Cecile Paris. Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational Linguistics*, 19(4):651-694, December 1993.

[MT87] W. C. Mann and S. A. Thompson. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):167–182, 1987.

[PA80] R. Perrault and J. Allen. A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics*, 6(3-4):167–182, 1980.

[Ste84] Anna-Brita Stenström. Questions and responses in english conversation. In Claes Schaar and Jan Svartvik, editors, *Lund Studies in English 68*. CWK Gleerup, Malmö, Sweden, 1984.