Preventing False Temporal Implicatures: Interactive Defaults for Text Generation*

Jon Oberlander and **Alex Lascarides**
Centre for Cognitive Science *and* Human Communication Research Centre
University of Edinburgh
2 Buccleuch Place, Edinburgh EH8 9LW Scotland

Introduction

Given the causal and temporal relations between events in a knowledge base, what are the ways they can be described in text?

Elsewhere, we have argued that during interpretation, the reader-hearer H must infer certain temporal information from knowledge about the world, language use and pragmatics. It is generally agreed that processes of Gricean implicature help determine the interpretation of text in context. But without a notion of logical consequence to underwrite them, the inferences—often defeasible in nature—will appear arbitrary, and unprincipled. Hence, we have explored the requirements on a formal model of temporal implicature, and outlined one possible nonmonotonic framework for discourse interpretation (Lascarides & Asher [1991], Lascarides & Oberlander [1992a]).

Here, we argue that if the writer-speaker S is to tailor text to H, then discourse generation can be informed by a similar formal model of implicature. We suggest two ways to do it: a version of Hobbs et al's [1988, 1990] Generation as Abduction; and the Interactive Defaults strategy introduced by Joshi et al [1984a, 1984b, 1986]. In investigating the latter strategy, the basic goal is to determine how notions of temporal reliability, precision and coherence can be used by a nonmonotonic logic to constrain the space of possible utterances. We explore a defeasible reasoning framework in which the interactions between the relative knowledge bases of S and H helps do this. Finally, we briefly discuss limitations of the strategy: in particular, its apparent marginalisation of discourse structure.

The paper focuses very specifically on implicatures of a temporal nature. To examine the relevant examples in sufficient detail, we have had to exclude discussion of many closely related issues in the theory of discourse structure. To motivate this restriction, let us therefore consider first why we might want to generate discourses with structures which lead to temporal complexities.

*The authors gratefully acknowledge the support of the Science and Engineering Research Council through project number GR/G22077. HCRC is supported by the Economic and Social Research Council. We thank our anonymous reviewer for their helpful comments. Email contact: jon@cogsci.ed.ac.uk

Getting Things Out of Order

Consider the following suggestion for generating textual descriptions of causal-temporal structures. Describe things in exactly the order in which they happened. If textual order is made to match eventual order, then perhaps little can go wrong; for the hearer can safely assume that all the texts she hears are narrative. Under these circumstances, the problem of selecting adequate regions in the space of utterances pretty much dissolves. We do not believe that this suggestion will work, in general, and consider here two arguments against it.

Hovy's argument

Basically, the generation strategy suggested above fails to emphasise the force of some eventualities over others (cf. the nucleus-satellite distinction in RST). A useful device for emphasis is the topic-comment structure: we mention the important event first, and then the others, which fill out or give further detail about that important event. These 'comments' on the 'topic' may be effects, but they could also be the cause of the topic. If the latter, then textual order and temporal order mismatch; the text is a causal explanation in such cases, and having only narrative discourse structure available would preclude its generation. Compare (1) and (2), modified from Hovy [1990].

(1) First, Jim bumped Mike once and hurt him. Then they fought. Eventually, Mike stabbed him. As a result, Jim died.

(2) Jim died in a fight with Mike. After Jim bumped Mike once, they fought, and eventually Mike stabbed him.

The textual order in (1) matches temporal order, whereas in (2) there is mismatch. And yet (2) is much better than (1). This is because the 'important' event is Jim's death. Everything mentioned in (1) leads up to this. But because the events are mentioned in their temporal order, the text obscures the fact that *all* the events led to Jim's death, even though syntactic markers like *and then* and *as a result* are used.

The causal groupings are clearer in (2) because it's clear during incremental processing that the text following the mention of Jim's death is a description of how it came about. This is so even though *no*

syntactic markers indicate this causal structure. By contrast, in (1) the reader realises what's going on only at the last sentence. The discourse structure is therefore unclear until the whole text is heard, for the narrative requires a common topic which is only stated at the end.

So (2)'s a better discourse than (1); but we would never generate it, if textual order had to mirror eventual order. If a generation system were permitted to generate (2), however, a price must be paid. The proper interpretation of (2) relies on the recruitment of certain causal information, left implicit by the utterance. The generator thus has some responsibility for ensuring that the interpreter accomplishes the required inferences. A formal model of implicature must be folded into the generation process, so that the appropriate reasoning can proceed.

States interact with causal information

In Lascarides and Oberlander [1992], we considered in detail the following pair of examples:

(3) Max opened the door. The room was pitch dark.

(4) Max switched off the light. The room was pitch dark.

Now, no-one would want to say that (3) involved a room *becoming* pitch dark immediately after a door was opened. Rather, most accounts (such as those based in or around DRT, such as Hinrichs [1986]) will take the state of darkness to overlap the event of door-opening. That's how one might say states are dealt with in a narrative: events move things along; states leave them where they are. But if we have a piece of causal information to hand, things are rather different. In (4), it seems that the state *doesn't* overlap the previously mentioned event.

If one wishes to preserve the assumption about the role of states in narrative, it would have to be weakened to the constraint that states either leave things where they are, or move them along. This is not a very convincing move. An alternative is to formalise the role of the additional causal knowledge. Informally, the basis for the distinct interpretations of (3) and (4) is that the interpretation of (4) is informed by a causal preference which is lacking in the case of (3): if there is a switching off of the light and a room's being dark that are connected by a causal, part/whole or overlap relation, then normally one infers that the former caused the latter. This knowledge is defeasible, of course. In generation, such knowledge will constrain the space of adequate utterances; if H lacks the defeasible causal knowledge that switching off lights cause darkness, then (4) won't be adequate for H, who will interpret (4) in the same way as (3), contrary to S's intentions. Given this, S must contain a defeasible reasoning component to compute over such knowledge.

The important point for now is that even if we describe things in the order in which they are assumed to happen, this doesn't necessarily make the candidate utterance a good one. If the speaker and the hearer possess differing world knowledge, there may be problems in retrieving the correct causal-temporal structure.

Two Methods of Generating with Defeasible Knowledge

Generation by Defeasible Reasoning

There is a very general way in which we might view interpretation and generation in terms of defeasible reasoning. Consider the process of discourse interpretation as one of KB extension. The KB contains an utterance-interpretation, and a set of knowledge resources; the latter may include general knowledge of the world, knowledge of linguistic facts, knowledge about the discourse so far, and about the speaker's knowledge state. We then try to extend the KB so as to include the discourse interpretation. Consider now the process of generation; it too can be thought of as KB extension. This time, the KB contains a temporal-causal structure, and a set of knowledge resources, perhaps identical to that used in interpretation. We now try to extend the KB so as to include the realization of a linguistic structure's semantic features (with predicates, arguments, connectives, orderings), where these features ensure that the final linguistic string describes the causal structure in the KB. This view might be described as *generation by defeasible reasoning*.

Modulo more minor differences, these notions are close to the ideas of *interpretation as abduction* (Hobbs et al [1988]) and *generation as abduction* (Hobbs et al [1990:26–28]), where we take abduction, in the former case for instance, to be a process returning a temporal-causal structure which can explain the utterance in context. Correspondences between a defeasible deduction approach and an abductive approach have been established by Konolige [1991]; he shows that the two are nearly equivalent, the consistency-based approach being slightly more powerful [1991:15–16], once closure axioms are added to the background theory. Lascarides & Oberlander [1992b] discuss in detail how such a generation process produces temporally adequate utterances.

Interactive defaults

Here, we turn to another, less powerful but simpler, method of applying defeasible reasoning: the Interactive Defaults (ID) strategy introduced by Joshi, Webber and Weischedel [1984a, 1984b, 1986]. Rather than considering the defeasible process as applying directly to the KB's causal network, we instead consider its role as constraining or debugging candidate linearised utterances, generated by some other process; here we will remain neutral on the nature of that originating process.

A speaker S and a hearer H interact through a dialogue; a writer S and a reader H interact through a text. Joshi et al argue that it is inevitable that both S and H infer more from utterances than is explicitly contained within them. Taking Grice's [1975] Maxim of Quality seriously, they argue that since both S and H know this is going to happen, it is incumbent upon S to take into account the implicatures H is likely to make on the basis of a candidate utterance. If S detects that something S believes to be false will be among H's implicatures, S must block that inference somehow. The basic way to block it is for S to use

a different utterance; one which S does not believe will mislead H.

In terms of defeasible reasoning, the point is that S must use it to calculate the consequences of the candidate utterance; if the process allows the derivation of something S believes to be false, the utterance should not be used in its current form. Joshi et al illustrate with the following example; given the KB in (5), and the question in (6), they want the process to show why the answer in (7b) is preferred to that in (7a):

(5) Sam is an associate professor; most associate professors are tenured; Sam is not tenured.

(6) Is Sam an associate professor?

(7) a. Yes.

b. Yes, but he is not tenured.

We wish to elaborate this interactive defaults strategy (ID), and consider in greater formal detail the defeasible reasoning about causal-temporal structures that S and H are assumed by S to indulge in; and to consider which candidate utterances are eliminated on this basis.

Discourse Structure and Temporal Constraints

ID requires a theory of implicature in terms of defaults, and an underlying logical notion of nonmonotonic or defeasible inference. We also require a formal characterisation of the properties an adequate candidate utterance must possess; we define these below in terms of *temporal coherence, reliability and precision*. Furthermore, we assume a model of discourse structure is required. For certain discourse relations, such as *Narration* and *Explanation*, are implicated from candidate utterances (cf. texts (1) and (2)), and these impose certain temporal relations on the events described. We turn to this latter issue first.

Discourse Structure and Inference

The basic model in which we embed ID assumes that candidate discourses possess hierarchical structure, with units linked by discourse relations modelled after those proposed by Hobbs [1985]. Lascarides & Asher [1991] use *Narration, Explanation, Background, Result* and *Elaboration*. They provide a logical theory for determining the discourse relations between sentences in a text, and the temporal relations between the events they describe. The logic used is the nonmonotonic logic Common Sense Entailment (CE) proposed by Asher & Morreau [1991].

Implicatures are calculated via default rules. For example, they motivate the following rules as manifestations of Gricean-style pragmatic maxims and world knowledge, where the clauses α and β appear in that order in the text. Informally:

- **Narration**
  If clauses α and β are discourse-related, then normally $Narration(\alpha, \beta)$ holds.
- **Axiom on Narration**
  If $Narration(\alpha, \beta)$ holds, and α and β describe events e_1 and e_2 respectively, then e_1 occurs before e_2.
- **Explanation**
  If clauses α and β are discourse-related, and the event described in β caused that described in α, then normally $Explanation(\alpha, \beta)$ holds.
- **Axiom on Explanation**
  If $Explanation(\alpha, \beta)$ holds, then event e_1 described by α does not occur before event e_2 described by β.
- **Causal Law**
  If clauses α and β are discourse-related, and α describes the event e_1 of x falling and β the event e_2 of y pushing x, then normally e_2 causes e_1.
- **Causes Precede Effects**
  If event e_2 causes e_1, then e_1 doesn't occur before e_2.

The rules for Narration and Explanation constitute defeasible linguistic knowledge, and the Axioms on them, indefeasible linguistic knowledge. The Causal Law is a mixture defeasible linguistic knowledge and world knowledge: given that the clauses are discourse-related somehow, the events they describe must be connected in a causal, part/whole or overlap relation; here, given the events in question, they must stand in a causal relation, if things are normal. That Causes Precede their Effects is indefeasible world knowledge. These rules are used under the CE inference regime to infer the discourse structures of candidate texts. Two patterns of inference are particularly relevant: Defeasible Modus Ponens (birds normally fly, Tweety is a bird; so Tweety flies); and the Penguin Principle (all penguins are birds, birds normally fly, penguins normally don't fly, Tweety is a penguin; so Tweety doesn't fly).

For example, in the absence of information to the contrary, the only one of the rules whose antecedent is satisfied in interpreting text (8) is Narration.

(8) Max stood up. John greeted him.

Other things being equal, we infer via Defeasible Modus Ponens that the *Narration* relation holds between (8)'s clauses, thus yielding, assuming logical omniscience, an interpretation where the descriptive order of events matches their temporal order. On the other hand, in interpreting text (9), in the absence of further information, *two* default laws have their antecedents satisfied: Narration and the Causal Law.

(9) Max fell. John pushed him.

The consequents of these default laws cannot both hold in a consistent KB. By the Penguin Principle, the law with the more specific antecedent wins: the Causal Law, because its antecedent logically entails that of Narration. Hence, (9) is interpreted as a case where the pushing caused the falling. In turn, this entails that the antecedent to Explanation is verified; and whilst conflicting with Narration, it's more specific, and hence its consequent—*Explanation*—follows by the Penguin Principle. Compare this with (8): similar logical forms, but different discourse structures, and different temporal structures.[1]

[1] The formal details of how the logic CE models these

Temporal Constraints

So against this background, what are the properties we require of candidate utterances? We concentrate on those constraints that are central to temporal import. Following Bach [1986], we take 'eventualities' to cover both events and states. We define *temporal coherence*, *temporal reliability* and *temporal precision*—the notions that will characterise the adequacy of an utterance—in terms of a set C of relations between eventualities. This set intuitively describes when two eventualities are connected. The relations in C are: causation, the part/whole relation,[2] temporal overlap, and the immediately precedes relation (where 'e_1 immediately precedes e_2' means that e_1 and e_2 stand in a causal or part/whole relation that is compatible with e_1 temporally preceding e_2).[3] The definitions are:

- **Temporal Coherence**
  A text is *temporally coherent* if the reader can infer that at least one of the relations in C holds between the eventualities described in the sentences.
- **Temporal Reliability**
  A text is *temporally reliable* if one of the relations in C which the reader infers to hold does in fact hold between the eventualities described in the sentences.
- **Temporal Precision**
  A text is *temporally precise* if whenever the reader infers that one of a proper subset of the relations in C holds between the eventualities described in the sentences, then she is also able to infer *which*.

A text is temporally *incoherent* if the natural interpretation of the text is such that there are no inferrable relations between the events. A text is temporally *unreliable* if the natural interpretation of the text is such that the inferred relations between the events differ from their actual relations in the world. In addition, a text is temporally imprecise, or as we shall say, *ambiguous*, if the natural interpretation of the text is such that the reader knows that one of a proper subset of relations in C holds between the eventualities, but the reader can't infer *which* of this proper subset holds.

It follows from the above definitions that a text can be coherent but unreliable. On the other hand, there may be no question about reliability simply because we cannot establish a temporal or causal relation between the two eventualities. At any rate, a generated utterance is adequate only if it is temporally coherent, reliable and precise. We intend to apply the ID strategy to eliminate candidate utterances that are inadequate in this sense.

interpretations, and those of (3) versus (4), are given in Lascarides & Asher [1991]. Note that although double applications of the Penguin Principle, as in (9), are not valid in general, they show that for the particular case considered here, OE validates the double application.

[2]We think of 'e_1 is part of e_2' in terms of Moens and Steedman's [1988] event terminology, as 'e_1 is part of the preparatory phase or consequent phase of e_2'.

[3]We assume that an event e_1 precedes an event e_2 if e_1's *culmination* occurs before e_2's. So there are part/whole relations between e_1 and e_2 that are compatible with e_1 temporally preceding e_2.

Applying the ID strategy

Before applying ID with temporal constraints, we must consider the possible relations between the knowledge of speaker S and that which speaker S has about hearer H's knowledge state. Notice, incidentally, that Joshi et al explicitly adopt the view that ID is for debugging candidate utterances. In principle, their framework, however, is more general. Although the idea of debugging is intuitive, we shall sometimes talk in terms of constraining the space of possible utterances, rather than of debugging specific utterances. The definitions of temporal constraints are relevant either way.

Relative KBs

Let $B(S)$ be S's beliefs about the KB, linguistic knowledge (LK) and world knowledge (WK). Let $B^+(H)$ be S's beliefs about what H believes about the KB, LK and WK. And let $B^-(H)$ be S's beliefs about what H doesn't know about the KB, LK and WK (so $B^+(H)$ and $B^-(H)$ are mutually exclusive). Problems concerning reliability and precision arise when $B(S)$ and $B^+(H)$ are different, and when S's knowledge of what H believes is partial (i.e. for some p, $p \notin B^+(H)$ and $p \notin B^-(H)$). Suppose that S's goal is to convey the content of a proposition contained in his KB, say q. Suppose also that a WFF p is relevant to generating a particular utterance describing q. Then there are several possible relations between $B(S)$, $B^+(H)$ and $B^-(H)$ that concern p:

- **Case 1**
  S knows p and also knows that H does not:
  $p \in B(S)$ and $p \in B^-(H)$
- **Case 2**
  S knows p and isn't sure whether H does or not:
  $p \in B(S)$ and $p \notin B^+(H)$ and $p \notin B^-(H)$
- **Case 3**
  H potentially knows more about p than S does:
  $p \notin B(S)$ and $p \notin B^+(H)$ and $p \notin B^-(H)$
- **Case 4**
  S thinks H is mistaken in believing p:
  $p \notin B(S)$ and $p \in B^+(H)$

Of course, the cases where both S and H both believe p ($p \in B(S)$ and $p \in B^+(H)$) and where neither do ($p \notin B(S)$ and $p \in B^-(H)$) are unproblematic, and so glossed over here. We look at each of the above cases in turn, considering the extent to which the definitions of reliability, coherence and precision help us constrain the utterance space (or alternately, debug candidate utterances).

Case 1: S knows more about p than H

We now examine the problems concerning reliability that arise when $p \in B(S)$ and $p \in B^-(H)$. There are two possibilities: either p represents defeasible knowledge of the language or the world, or p is some fact in the KB. We investigate these in turn.

p is defeasible knowledge Let p be a defeasible law that represents knowledge that S has and which S knows H lacks. To illustrate, take the case where p is the causal preference introduced earlier:

- **Causal Law**
  If the clauses α and β are discourse related, and α and β describe respectively the events e_1 of x falling and e_2 of y pushing x, then normally e_2 caused e_1.

Consider the case where S intends to convey the proposition that John's pushing Max caused the latter to fall. Suppose S has a KB which will allow her to generate the description in (9), among others.

(9) Max fell. John pushed him.

We have argued that this text is coherent, precise and reliable for S because the causal law (about the usual causal relation between pushings and fallings) is more specific than the linguistic rule (*Narration*). But since H lacks the causal law, (9) will trigger a different inference pattern in H; one in which *Narration* wins after all. S must block this pattern by changing the utterance; she has essentially two options. If clause order is kept fixed, then S could shift tense into the pluperfect as in (10); or else S can insert a *clue word*, such as *because*, into the surface form, to generate (11):

(10) Max fell. John had pushed him.

(11) Max fell because John pushed him.

The success of the latter tactic requires S and H to mutually know a new linguistic rule, more specific than *Narration*, such as the following:[4]

- **Non-evidential 'Because'**
  If α and β are discourse-related, and the text segment is α *because* β, then normally the event described in α caused that described in β.

On the other hand, if clause order is not taken to be fixed, then S can simply reorder (9):

(12) John pushed Max. Max fell.

So, when S believes H lacks the relevant causal law, S can simply reorder, and let *Narration* do the rest. However, recalling the above discussion, in some cases a discourse structure that invokes *Explanation* is better than one that invokes *Narration*. So simply reordering events and letting the rule for *Narration* achieve the correct inferences won't work successfully in all cases. Furthermore, recalling the discussion about states and causation above, it becomes apparent that this tactic of always letting *Narration* do the work will lead to problems with texts like (3) and (4).

(3) Max opened the door. The room was pitch dark.

(4) Max switched off the light. The room was pitch dark.

The reason is that, in the absence of the causal law which relates light switching to darkness, (4) will be analysed exactly as (3), giving the wrong result. A solution would be to replace the state expression with an event expression:

(4′) Max switched off the light. The room *went* pitch dark.

[4]This is a pragmatic, rather than semantic rule; it's not obvious that this is the best choice of representation.

An obvious alternative is to introduce further clue words, and appropriate linguistic rules for reasoning about them. This means exploiting linguistic knowledge to overcome the gaps in H's world knowledge. This helps explain the observation that texts which describe events in reverse to temporal order, without marking the reverse, may be quite rare. It's easy enough to interpret such texts, when we have the appropriate WK. But if a considerate speaker or writer has reason to believe that some or all of her audience lacks that WK, then she will either avoid such descriptive reversals, or mark them with the type of clues we have discussed.

p is a fact in the KB We now turn to the case where p is a fact about the KB which S knows and which S knows H lacks. Suppose that p asserts a causal relation between two events that does not represent an exception to any defeasible causal preferences, and that S wishes to convey the information that p. Then S can simply state p by exploiting H's available LK. Clue words may not be needed. For example, if p is the fact that Max stood up and then John greeted him, S can tell H this by uttering (8); *Narration* will make (8) reliable and precise for H.

(8) Max stood up. John greeted him.

Similarly, if p is the fact that Max opened the door, and while this was going on the room was pitch dark, then (3) is reliable and precise for H:

(3) Max opened the door. The room was pitch dark.

But what if p asserts a causal relation between two events that violates a defeasible causal preference that H has? Suppose p asserts that Max's fall immediately preceded John's pushing him. And suppose that S knows that H has the defeasible causal law mentioned above, but lacks p. Then neither (9) nor (12) are reliable for H, indicating that S cannot generate an atomic text to assert p.

(9) Max fell. John pushed him.

(12) John pushed Max. He fell.

H would interpret (9) as an explanation; and (12) as a narrative, for nothing will conflict with Narration in that case: the causal preference for pushings causing fallings would simply reinforce the temporal structure imposed by *Narration*. The obvious option is to move from (9) to (13); another option is to recruit the pluperfect, as in (14); note that (15) is *not* a solution, since *so* can be read evidentially.

(13) Max fell. And then John pushed him.

(14) John pushed Max. He had fallen.

(15) Max fell. So John pushed him.

The need to utter (13) rather than (9) explains why it can be necessary to use *and then*, even though the full-stop is always available and, by *Narration*, has the default effect of temporal progression. So, in general, one might wish to paraphrase Joshi et al: if a relation can be defeasibly inferred to hold between two eventualities, and S wants something different, it is essential to *mark* the desired relation with something stronger.

Case 2: S knows p but isn't sure if H does

In general, S will have only partial knowledge about H's beliefs. This has its drawbacks.

p is a defeasible causal preference Suppose that S isn't sure whether or not H believes the defeasible causal law relating falling and pushing. Then there are at least two ways in which S's model of H's knowledge can be expanded to a complete statement of H's knowledge. The first, B_1, contains the causal law. The second, B_2, does not. Now uppose that S wishes to convey the proposition that John's pushing Max caused Max to fall. Then, if S assumes H's knowledge corresponds to B_1, then H will find a reliable interpretation for (9).

(9) Max fell. John pushed him.

On the other hand, if S assumes that H's knowledge corresponds to B_2, then H will interpret (9) in an undesirable way, with the falling preceding the pushing; as we said before, Narration would win.

Under this model, S isn't sure how H will interpret (9), because S doesn't know if H's knowledge corresponds to B_1 or B_2. Hence the ambiguity of (9) manifests itself to the generator S, if not to the hearer H, because S doesn't have sufficient information about H to predict which of the two alternative temporal structures H will infer for (9). This is slightly different to the previous case where S actually *knows* H lacks the causal law, making (9) unreliable.

To avoid uttering unreliable text, S will have to utter something other than (9). Indeed, it may be possible for S not to worry about the ambiguity of (9) at all, if some 'safe' strategy can be found that would guide S's expansion of H's knowledge in a way that would ensure the generation of reliable text for H. A plausible strategy for S's reasoning about H would be the following: if S isn't sure whether or not H knows p, then assume H doesn't know p. On the face of it this seems plausible. But just how safe is it?

We state it in terms of $B^+(H)$ and $B^-(H)$:

- If $p \notin B^+(H)$ and $p \notin B^-(H)$, assume $p \in B^-(H)$ and generate-and-test under this assumption.

But this won't work in general. If S wants to convey a violation of the causal law p, but H actually believes p, then the strategy will suggest the use of (9), which will actually be unreliable for H.

In fact, there is no safe strategy, save the one where S considers several alternative expansions of H's knowledge. As a result, ambiguity of text will manifest itself to S in certain cases, because of her partial knowledge of H. This is perhaps somewhat surprising. Nonmonotonic reasoning is designed as a medium for reasoning with partial knowledge. And yet here we have shown S cannot maintain textual reliability on the basis of a partial statement of H's KB, even if nonmonotonic inference is exploited.

p is a fact about the KB: Ambiguity Suppose that S wants to convey the information that Max's fall immediately preceded John pushing him, and suppose S knows that H knows the causal law, but S doesn't know for sure if H knows already that Max fell *before* John pushed him. Then, for similar reasons as those mentioned earlier, S isn't sure if (9) is reliable or not.

(9) Max fell. John pushed him.

To be sure that text is reliable in this case, S will again have to exploit linguistic knowledge; for example, by uttering (13) instead of (9).

(13) Max fell and then John pushed him.

Case 3: H as advisor, S as pupil

Suppose that for a certain proposition p, $p \notin B(S)$, $p \notin B^+(H)$ and $p \notin B^-(H)$. This corresponds to H potentially knowing more about p than S, but S not knowing what more. That's pretty much the position of the tutee in a tutorial dialogue, and the advice-taker in an advisory dialogue.

Case 4: S thinks that H is mistaken

Suppose that $p \notin B(S)$ and $p \in B^+(H)$. Then S doesn't believe p even though she's aware H does. This implies that S thinks H is mistaken in believing p.

The fact that $p \notin B(S)$ and $p \in B^+(H)$ could entail that a text that's reliable for S isn't for H. For example, suppose that H believes, by some weird perception of social convention, that there is a defeasible causal preference that greetings cause standing ups. Suppose that S wants to describe the situation where Max stood up and then John greeted him (i.e. an exception to H's causal preference). Then this is like the exception case above concerning falling and pushing: (16) is reliable for S but not for H.

(16) Max stood up. John greeted him.

Again, S could compensate for this by explicitly marking the temporal relation. Alternatively, the fact that $p \notin B(S)$ and $p \in B^+(H)$ could entail that a text that's *unreliable* for S is reliable for H. Again, let p be the causal law that says that greetings cause standing ups. But this time suppose that S wants to describe the situation where John's greeting Max caused him to stand up. So this time, S wants to describe an instance of the causal law. Then both (16) and (17) are reliable for H, but only the latter is reliable for S.

(17) John greeted Max. He stood up.

(16) is unreliable for S. Arguably, it wouldn't be in the set of possible linguistic realisations, but only if this set is assumed to be characterised by what S finds reliable. But we have no argument for this assumption, and so we don't make it.

Conclusions

Here, we summarise the current state of the model, and briefly discuss two of its limitations.

We admitted that that job of defeasible reasoning in generation could be very general; but that we were going to look at it in the context of the Interactive Defaults strategy. ID applies to the candidate utterances (or the space of utterances), and criticises the utterances (or the space), producing better utterances, or a smaller space. The notion of logical

consequence supported by CE was used to make precise how utterances are constrainted by ID. Crucially, we used Defeasible Modus Ponens and the Penguin Principle. The grounds for criticism were the temporal ramifications of the utterance; if it was incoherent for H, unreliable for H or dangerously ambiguous (for S), it was bad.

One limitation of the model is that, although it permits reasoning about the knowledge or beliefs of interlocutors, it neglects their goals and intentions to do actions. ID does not deal with the phenomena which motivate the work following Cohen and Perrault [1979] and Allen and Perrault [1980], (cf. Cohen, Morgan and Pollack [1990]). In particular, ID does not let S take into account those inferences H will make in attempting to ascribe a plan to S. Hobbs et al [1990:44–45] argue that inferences leading to plan recognition are less significant in interpreting long written texts or monologues. Hence, it might be argued that the generation of such discourses need not give H's plan recognition particular weight. Nonetheless, ID is incomplete, to the extent that such inferences influence discourse generation.

Secondly, discourse structure and temporal structure have become somewhat detached. Sometimes, it's only the causal-temporal structure derivable from the candidate that is being criticized. It may therefore be thought that the discourse structure is an idle wheel as things stand, and should be either eliminated (cf. Sibun [1992]), or be trusted with a greater share of the work, enriching the discourse with useful clue words (cf. Scott and Souza [1990]). Our tentative view is that the latter view is plausible, and anyway is closer to the idea of generation by defeasible reasoning, canvassed early on.

The ID strategy examined here seems to involve a lot of hard work generating simple candidates which almost always require debugging. It would be preferable if we could do this work in advance, by default. The alternative is explored in Lascarides and Oberlander [1992b], in which we abduce discourse structures from event structures, and then interleave deduction and abduction to derive linguistic realisations. But in turning to the more global approach, we should not lose sight of the fact that simple texts are sometimes best. (2) illustrates this point: the rhetorical relations inferred aren't syntactically marked, and yet the text is more natural than (1), where the relations are marked. As might be expected, there seems to be a trade-off between the naturalness of the output and its computational cost.

References

Allen, J. F. & Perrault, C. R. [1980] Analyzing Intention in Dialogues. *Artificial Intelligence*, **15**, 143–178.

Asher, N. & Morreau, M. [1991] Common Sense Entailment: A Modal Theory of Nonmonotonic Reasoning. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, Sydney, Australia, August 1991.

Bach, E. [1986] The algebra of events. *Linguistics and Philosophy*, **9**, 5–16.

Cohen, P. R. & Perrault, C. R. [1979] Elements of a Plan-Based Theory of Speech Acts. *Cognitive Science*, **3**, 177–212.

Cohen, P. R., Morgan, J. & Pollack, M. E. [1990] *Intentions in Communication*. Cambridge, MA: MIT press.

Grice, H. P. [1975] Logic and Conversation. In Cole, P. and Morgan, J. L. (eds.) *Syntax and Semantics*, Volume 3: *Speech Acts*, pp41–58. New York: Academic Press.

Hinrichs, E. [1986] Temporal Anaphora in Discourses of English. *Linguistics and Philosophy*, **9**, 63–82.

Hobbs, J. R. [1985] On the Coherence and Structure of Discourse. Report CSLI–85–37, Center for the Study of Language and Information, Stanford, Ca., October, 1985.

Hobbs, J., Stickel, M., Martin, P. & Edwards, D. [1988] Interpretation as Abduction. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, SUNY, Buffalo, N.Y., June, 1988, pp95–103.

Hobbs, J., Stickel, M., Appelt, D. & Martin, P. [1990] Interpretation as Abduction. Technical Note No. 499, SRI International, Menlo Park, Ca., December 1990.

Hovy, E. [1990] Pragmatics and Natural Language Generation. *Artificial Intelligence*, **43**, 153–197.

Joshi, A., Webber, B. & Weischedel, R. M. [1984a] Preventing False Inferences. In *Proceedings of the 10th International Conference on Computational Linguistics and the 22nd Annual Meeting of the Association for Computational Linguistics*, Stanford University, Stanford, Ca., 2–6 July, 1984, pp134–138.

Joshi, A., Webber, B. & Weischedel, R. [1984b] Default reasoning in interaction. In *Proceedings of the Non-Monotonic Reasoning Workshop*, AAAI, N.Y., October, 1984, pp144–150.

Joshi, A., Webber, B. & Weischedel, R. [1986] Some Aspects of Default Reasoning in Interactive Discourse. Report MS-CIS-86-27, University of Pennsylvania.

Konolige, K. [1991] Abduction vs Closure in Causal Theories. Forthcoming Research Note in *Artificial Intelligence*. Page references to ms.

Lascarides, A. & Asher, N. [1991] Discourse Relations and Common Sense Entailment. Submitted to *Journal of Logic, Language and Information*. DYANA deliverable 2.5b, Centre for Cognitive Science, University of Edinburgh.

Lascarides, A. & Oberlander, J. [1992a] Temporal Coherence and Defeasible Knowledge. *Theoretical Linguistics*, **18**.

Lascarides, A. & Oberlander, J. [1992b] Abducing Temporal Discourse. In Dale, R. Hovy, E. Rösner, D. and Stock, O. (eds.) *Aspects of Automated Natural Language Generation*. Berlin: Springer-Verlag.

Scott, D. R. & Souza, C. S. [1990] Getting the Message Across in RST-based Text Generation. In R. Dale, C. Mellish and M. Zock (eds.) *Current Research in Natural Language Generation*. London: Academic Press.

Sibun, P. [1992] Generating Text without Trees. To appear in *Computational Intelligence: Special Issue on Natural Language Generation*, 8.