Paper 1

Resolving Underspecification using Discourse Information

DAVID SCHLANGEN, ALEX LASCARIDES, ANN COPESTAKE
UNIVERSITY OF EDINBURGH; UNIVERSITY OF CAMBRIDGE AND STANFORD UNIVERSITY
{das|alex}@cogsci.ed.ac.uk, aac@cl.cam.ac.uk
http://www.cogsci.ed.ac.uk/~das

Abstract

This paper describes RUDI ("Resolving Underspecification with Discourse Information"), a dialogue system component which computes automatically some aspects of the content of scheduling dialogues, particularly the intended denotation of the temporal expressions, the speech acts performed and the underlying goals. RUDI has a number of nice features: it is a principled approximation of a logically precise and linguistically motivated framework for representing semantics and implicatures; it has a particularly simple architecture; and it records how reasoning with a combination of goals, semantics and speech acts serves to resolve underspecification that's generated by the grammar.

1 Introduction

Our aim in this work is to investigate formally the interaction between compositional semantics, goals, and discourse structure in task-oriented dialogues. Specifically, we look at how an underspecified semantic representation may be instantiated by discourse information, and we investigate the extent to which we can preserve principled approximations of a general theory of dialogue semantics in a practical implementation for a restricted domain. To this end, we designed an experimental software dialogue system, RUDI.

As a testbed for this dialogue modelling, we chose the domain of fixing appointments, because we had access to a range of realistic dialogues that had been collected as part of the Verbmobil project (Wahlster 2000) and to a parser which was capable of producing semantic representations for them (see below). We concentrate on dialogues that deal with the subtask of identifying a mutually agreed time to meet, ignoring other subtasks such as agreeing on a *place* to meet. The particular kind of underspecification we are investigating arises from the use of definite temporal descriptions in such dialogues. Example (1.1) shows an excerpt from such a dialogue:

(1.1) A: Can we meet on Friday?
B: How about 4pm?

We analyse such definite descriptions as requiring a bridging relation to an antecedent in the context.[1] Neither the bridging relation nor the antecedent are determined by the compositional

[1] We are interested here only in this one class of definite descriptions. (For a general classification cf. (Hawkins 1978).) The term *bridging* was introduced by Clark (1975) for definite descriptions which lack a unique antecedent

BI-DIALOG 2001 (P. Kühnlein, H. Rieser, & H. Zeevat (eds)).

University of Edinburgh; University of Cambridge and Stanford University
{das|alex}@cogsci.ed.ac.uk, aac@cl.cam.ac.uk
http://www.cogsci.ed.ac.uk/~das .

semantics of the utterance, however. Thus, we take the semantic representation of such expressions to contain an underspecified relation between an underspecified antecedent and the referent for the expression.

A task that's co-dependent on resolving this underspecification is computing how the utterance contributes to a *coherent* dialogue. Following Segmented Discourse Representation Theory (SDRT, cf. e.g. (Asher 1993, Lascarides & Asher 1993)), we assume that a dialogue is coherent just in case every proposition (and question and request) is *rhetorically connected* to another proposition (or question or request) in the dialogue, and all anaphoric expressions can be resolved. The rhetorical relations can be viewed as *speech act types* (see (Asher & Lascarides 2001) for details), and they constrain both the semantic content of the utterances they relate, and what we call *speech act related goals* or SARGS.

Our thesis is that information can flow either from resolving the semantic underspecification to computing the rhetorical relation, or *vice versa* (and hence we're claiming rhetorical relations are an essential source of information for resolving semantic underspecification that's generated by the grammar). For example, the rhetorical relation which connects the utterances in (1.1) is inferred on the basis of the sentence moods (justification for this is given shortly), and the semantics of this rhetorical relation constrains the interpretation of *4pm* to be 4pm on Friday (as opposed to the alternative, which is the next 4pm to now).

The inference from linguistic form to the rhetorical relation (or equivalently, the speech act) is a *default* inference, however. Although the sequence of sentence moods in (1.3) is the same as in (1.1), the speech act's semantics is incompatible with all the possible resolutions of the temporal underspecification in (1.3).

(1.3) A: Let's meet next Saturday.
B: How about Sunday?

In such cases, RUDI has the capacity to explore whether an *indirect* speech act (ISA) has taken place; in this case, it will correctly predict that the illocutionary contribution of B's utterance is not simply that of the question, but it also conveys a *rejection* of A's proposal (to meet next Saturday).[2] So in this case, information flows from resolving the underspecification to inferring the type of speech act that B has performed (or equivalently, the rhetorical relation which connects his utterance to A's).

Dialogue (1.4) shows another example where the resolution of anaphoric expressions yields inferences about the speech acts. If *now* is Monday 12th February 2001, then *next week* is the interval from the 19th to the 25th, and from this we conclude B's speech act is to reject A's SARG. If, however, *now* is the 7th February 2001, so next week is the 12th to the 18th, then B's speech act narrows the temporal parameter in A's SARG, viz. the 12th to the 15th. Inferring these different speech acts thus requires knowledge of the times denoted (and the relationship between them).

(1.4) A: Can we meet next week?
B: I'm busy from the 16th to the 25th.

that is present on the basis of what has been explicitly said, and where thus the interpreter "...is forced to construct an antecedent, by a series of inferences, from something he already knows. [...] The listener must therefore bridge the gap from what he knows to the intended antecedent." (Clark 1975, p.413)

In unrestricted domains, these bridging inferences can be quite involved and the reasoning is thus difficult to formalise generally (for an overview, see (Vieira & Poesio 2000), but see also (Asher & Lascarides 1998)). We chose this domain partly so that we can exploit conventional constraints on the possible bridging relations among temporal expressions; for unlike other domains, the possibilities are finite. For example, a complicated nonce-relation like *the first interval in the antecedent that satisfies the description in the anaphoric expression* doesn't seem to be a possible bridging relation, even though it would provide us with a unique antecedent for the example below:

(1.2) A: How about meeting in May?
B: #The Monday is good for me.

In fact, we believe that using only *temporal inclusion* and *next* as candidates for bridging relations is sufficient.

[2]We treat this as an ISA because rejections and questions are incompatible at the level of *semantic value* (see (Asher & Lascarides 2001) for details): a rejection is conveyed via a proposition, whereas a question denotes a set of propositions (ie. its direct answers; see (Groenendijk & Stokhof 1984)). We give more details of this analysis in section 3.2. Note that we abstract away from intonational clues that a contrast is intended here (stress on Sunday), which presumably would be present if B's utterance were *spoken*.

This work is part of a larger project, whose aim is to provide a computationally tractable and formally precise theory of how non-sentential fragments (e.g., *Not Tuesday*) are interpreted and generated. Therefore, we also need to *predict* when one can leave content implicit and when one can't. E.g., in (1.5), B's second utterance is odd. On the one hand, linguistic constraints on antecedents to anaphora stipulate that *4pm* should be resolved to Saturday 4pm (Kamp & Reyle 1993). But on the other hand, one cannot infer any of the candidate rhetorical relations to attach this resulting interpretation of the question to the context. Details are given shortly, but roughly speaking, no rhetorical relation can be computed in this case because the semantics of the relations capture the intuition that B should not ask whether A can meet him on Saturday afternoon, when he knows (because A has told him already) that he can't meet him then.

(1.5) a) A: Can we meet next weekend?
b) B: How about Saturday afternoon?
c) A: I am busy then.
d) B: ??How about 4pm?

This contrasts with the question *Even at 4pm?*, which ameliorates the incoherence in (1.5). In contrast to *How about 4pm*, *Even at 4pm?* can be interpreted as a question which addresses the communicative goal of 'belief transfer' that underlies A's prior utterance; namely, the goal that B believe that A is busy on Saturday afternoon. This shows that reasoning about the linguistic constraints on the interpretation of anaphora, rhetorical relations and communicative goals are all necessary for an adequate account of the coherent interpretation of temporal expressions.

RUDI adopts a dynamic semantic approach to dialogue interpretation: First, a compositional semantic representation of the current clause is constructed via a large HPSG (the English Resource Grammar built in the LinGO project, as parsed by the LKB).[3] This representation is then used to update the semantic representation of the discourse context. The co-dependent tasks of computing speech acts and goals and resolving semantic underspecification are a byproduct of computing this update. For this, we approximate SDRT. In the next section, we will briefly introduce the relevant bits of this theory, and then explain in section 2.2 how we can derive a body of simpler domain-specific rules from this theory in a principled way. Section 3 describes the implementation of these rules. We close with a brief discussion of related work and some conclusions.

2 Theoretical Background

2.1 SDRT

SDRT represents discourse content as an SDRS, which is a recursive structure of labelled DRSs, with rhetorical relations between the labels. In contrast to traditional dynamic semantics (e.g., DRT, (Kamp & Reyle 1993)), SDRT attempts to represent the *pragmatically preferred* interpretation of a discourse. Discourse update is formulated within a precise nonmonotonic logic, in which one computes the rhetorical relation (or equivalently, the speech act type) which connects the new information to some antecedent utterance. As mentioned in the introduction, this speech act places constraints on content and the speech act related goals or SARGs; these in turn serve to resolve semantic underspecification. Note that SARGs are goals that are either conventionally associated with a particular type of utterance or are recoverable by the interpreter from the discourse context; this distinguishes the goals that interact with linguistic knowledge from goals in general.

The rhetorical relations which are relevant to us here are the following:

Q-Elab(α, β) (Question Elaboration): β is a question where any possible answer to it elaborates a plan for achieving one of the SARGS of α. Eg. *A: Let's meet on Monday. How about 2pm?*

[3]Henceforth, we will refer to this grammar/parser combination as ERG/LKB. The LinGO project is described on `http://www-csli.stanford.edu/hpsg/lingo.html`, the LKB on `http://www-csli.stanford.edu/~aac/lkb.html`. See also (Copestake & Flickinger 2000).

IQAP(α, β) (Indirect Question Answer Pair): α is a question and β conveys information from which the questioner can infer a direct answer to α. Eg. *A: Can we meet next week? B: I'm free on Monday.*.

Plan-Correction(α, β)**:** the speaker of β rejects the SARG of α. Eg. (1.4) in the first setting above.

Plan-Elaboration(α, β)**:** β elaborates a plan to achieve a SARG of α. Eg. (1.4) in the second setting.

Note that these speech act types are *relations* (cf. (Searle 1967)), to reflect that the successful performance of the current speech act is logically dependent on the content of an antecedent utterance (e.g., successfully performing the speech act *IQAP*, as with any type of answering, depends on the content of the question α).

The default rules for computing speech acts have the form (1.6) ($A > B$ means *If A then normally B*):

(1.6) $(\langle\tau, \alpha, \beta\rangle \wedge \mathit{Info}(\tau, \beta)) > R(\alpha, \beta)$

$\langle\tau, \alpha, \beta\rangle$ means β is to be attached to α with a rhetorical relation (α and β label bits of content) where α is part of the discourse context τ; $\mathit{Info}(\tau, \beta)$ is a gloss for information about the content that τ and β label; and R is a rhetorical relation. This rule schema contrasts with the plan-recognition approach to computing speech acts (e.g. (Lochbaum 1998)), which uses *only* the goals of the antecedent utterance, rather than its compositional and lexical semantics directly, to constrain the recognition of the current speech act.

There are a number of advantages to allowing direct access to the content of τ in these inferences. For example, the successful performance of the current speech act is often dependent on the *logical structure* of the antecedent utterances, and goals don't reflect this logical structure; rather compositional semantics does (following DRT, (Kamp & Reyle 1993)). In fact, dialogue (1.5) demonstrates this. Given the context, a SARG for (1.5d) is to find a time to meet that's next weekend but not on Saturday afternoon. So computing the speech act solely on the basis of the prior goals and the current linguistic form would predict that *4pm* successfully refers to 4pm on Sunday and the speech act $\mathit{Q\text{-}Elab}(1.5c, 1.5d)$ is performed. The fact that (1.5d) is odd indicates that recognising its speech act is constrained by something else. On our approach, the logical and rhetorical structure of (1.5a-c) plays a central role, for according to *linguistic* constraints defined within dynamic semantics (e.g., (Kamp & Reyle 1993)), (1.5a-c) make Sunday inaccessible, thereby forcing *4pm* to denote 4pm on Saturday.

Some of the axioms of the form (1.6) are in fact derived via a formally precise model of cognitive reasoning, which encapsulates general principles of rationality and cooperativity (see (Lascarides & Asher 1999) for details). For example, such cognitive modelling validates `Q-Elab` and `IQAP` (where α :? means that α is an interrogative):

- `Q-Elab`: $(\langle\tau, \alpha, \beta\rangle \wedge \beta :?) > \mathit{Q\text{-}Elab}(\alpha, \beta)$
  `IQAP`: $(\langle\tau, \alpha, \beta\rangle \wedge \alpha :?) > \mathit{IQAP}(\alpha, \beta)$

`Q-Elab` stipulates that the default role of a question is to help achieve a SARG of a prior utterance. `IQAP` stipulates that the default contribution of a response to a question is to supply information from which the questioner can infer an answer. Thus inferences about speech acts, and hence about implicit content and goals, can be triggered (by default) purely on the basis of sentence moods.[4] This justifies our analysis of (1.1) we gave above. Per default we take B's utterance to attach via *Q-Elab* to A's because it is a question. The semantics of this relation, viz. that the utterance helps elaborating a plan, is only met in this domain if it is true that the time β specifies is temporally included in the time α proposes. We add this information in discourse update so as to ensure that the updated logical form is consistent; and this thereby resolves the underspecification.

[4]Since `IQAP` and `Q-Elab` are derived from axioms which model dialogue participants as rational and cooperative agents, one can view these rules as *short-circuiting* calculable implicatures about the content that the speakers intended to convey (Morgan 1975).

In an attempt to do justice to the complexity of interaction between the different information sources that contribute to dialogue interpretation—both conventional and non-conventional—many researchers have assumed a radically unmodular framework, so that a single reasoning process can access the different kinds of information at any time (eg. (Hobbs, Stickel, Appelt & Martin 1993)). In contrast, SDRT assumes a highly modular framework: reasoning about beliefs and goals is separate from, but interacts with, reasoning about content and speech acts. We will exploit this modularity so as to gain a particularly simple architecture to the implemented system.

Of course, world knowledge (WK) also affects interpretation. In this domain, relevant WK includes knowledge of which plans/actions when performed at time t are (in)compatible with meeting at t, and temporal reasoning with intervals and calendar terms. We'll discuss the former knowledge in the next section, and the latter in section 3.

2.2 Approximation

As we said in the introduction, our aim is to investigate the extent to which we can preserve principled approximations of the underlying theory, while maintaining a relatively good degree of robustness and precision. To this end, we make the assumption that the dialogue participants (DPs) don't digress from trying to reach their main goal, which is to meet at a time t.[5] This means that we assume that all utterances address this goal, so that we can say that the main SARG of all utterances is to provide information about available times for a meeting.[6] The domain-level plan to reach this goal now can be specified as follows: the DPs have to "zero in" on a time, by narrowing down the range of times that are available for a meeting.

Having made this assumption, we can make approximations to the general theory on two levels. First, we approximate knowledge of which events permit meeting at time t and which don't via postprocessing the underspecified semantic form (the MRS[7]) generated by the ERG/LKB. The result is an expression in a discourse input language (DIL), that preserves information about the temporal description of the time variable t, the sentence mood, and whether t was a good time or a bad time. Hence we abstract over information which is irrelevant to the task at hand, such as, for example, whether the utterance was about going to the dentist or going on vacation; they both generate *bad-time*(t).

This kind of postprocessing rule simply encapsulates knowledge of actions in the domain. Others are derived logically "off line" (ie. manually) in SDRT: for example, in this domain, SDRT validates the inference that asking a question about a time t implicates that it's a *good-time*(t) for the speaker to meet. The reasoning goes as follows. By default, a question attaches as *Q-Elab*. The semantics of this relation, namely that the question helps achieve a SARG of a prior utterance, is only met, given our additional assumption, if the utterance serves as a suggestion of a good time. This reasoning is 'hard-wired' into the post-processing rules, and thus we 'short-circuit' some SDRT inferences in the translation from MRS to DIL.

Approximation also occurs at the discourse level. First, we assume that the dialogue participants always believe the content of the other participants' utterances (i.e., the SARG of belief transfer that's conventionally associated with assertions is always successful). This means that questions which attach with *Q-Elab* to prior utterances are never interpreted as questions which elaborate a plan for achieving the SARG of belief transfer. In essence, this means that we assume that B won't utter *Even at 4pm?* in response to A's utterance (1.5d). Of course, this approximation is unjustified in general, but is acceptable in the restricted Verbmobil domain, since it is indeed the case that a dialogue agent assumes that the other agent is competent with respect to his assertions about when he can and can't meet.

Secondly, we utilize the assumptions about the overall purpose of these dialogues and the above approximations manually within SDRT, to yield the valid inferences that follow. In particular, the

[5] This non-digression assumption is of course unfounded in the general case, but can be justified in our simple restricted domain.

[6] In the following, we will simply talk of these SARGs *being* a time t, which is to mean that the goal is to meet at a time within t.

[7] MRSs (Copestake, Flickinger, Sag & Pollard 1999) are similar to Reyle's (1993) UDRSs.

default rules of the form (1.6) yield monotonic rules of a similar form, since proviso the 'non-digression' assumption, exceptions to the defaults can be exhaustively enumerated. By turning default rules into monotonic rules, we avoid computationally expensive consistency checks. Also, fixing the main goal allows us to specify the semantics of the relations for this domain as follows (cf. the general rules above in Section 2.1 and the actual update rules the system uses in Fig. 1.3):

Q-Elab(α, β)**:** β is a question (which means it proposes a *good_time*, see above) and t_β at least overlaps with SARG$_\alpha$, which makes sure that any possible answer addresses α's SARG.

IQAP(α, β)**:** α is a question and β talks about a time that overlaps with SARG$_\alpha$.

Plan-Correction(α, β)**:** the speaker of β rejects the SARG of α, by marking a time as *bad_time* that includes SARG$_\alpha$.

Plan-Elaboration(α, β)**:** β elaborates a plan to achieve a SARG of α, either by marking a time which overlaps with SARG$_\alpha$ as *good_time*, or by marking only parts of SARG$_\alpha$ as *bad_time*.

Another valid SDRT inference is a default rule for attaching to the previous utterance, because otherwise SARGs are left unaddressed, contrary to the cooperativity assumption (see (Lascarides & Asher 1999)); we'll exploit this in RUDI when choosing the site to which the new information connects. Overall, then, we hope that this method of system development will ensure that all rules encoded in the software are logically and linguistically principled.

3 The System

3.1 Overview

RUDI's information state is shown in Fig. 1.1. Its main components are CONTEXT, which holds all information about the discourse context, and CUR-UTT, which represents the current utterance with which the context is to be updated.

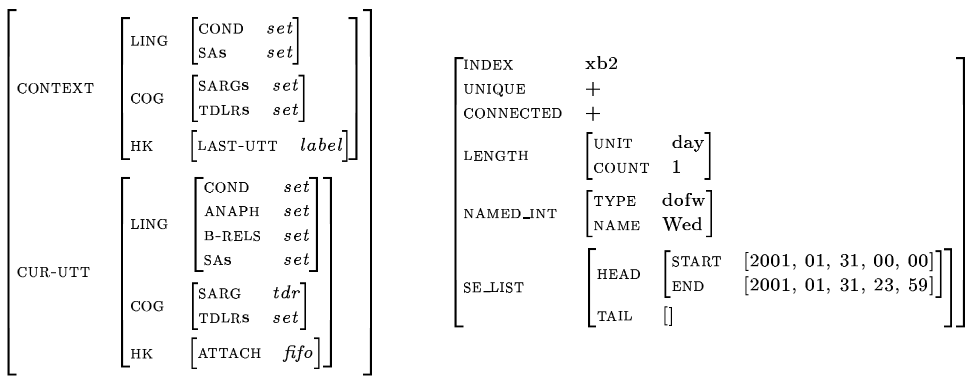

Figure 1.1: RUDI's information state (left) and a TDL-representation (right)

Both representations consist of a linguistic part (LING) and a cognitive part (COG). The linguistic part contains a set of conditions, ie. labelled predicates, and a list of the speech acts performed.[8] CUR-UTT additionally has fields to keep track of the anaphora and possible resolutions. The cognitive part represents information about cognitive states, viz. the SARGs and the intended denotations in the domain, in the form of TDLRs. These are representations in a domain specific language, the temporal domain language (TDL). We encapsulate all knowledge about calendars, durations and intervals in this language; all domain specific reasoning takes places on these structures. Fig. 1.1 shows as an example a TDL representation for Wednesday 31st January 2001. The start- and end-points of the interval are specified in a list-structure, so as to allow representation of non-connected intervals. The feature HK in both parts finally holds information that is strictly

[8]This can be seen as being a notational variant of SDRT-style labelled boxes.

speaking not part of a semantic representation; it is rather "housekeeping" information needed during the update process.

The modular architecture of the update process in RUDI reflects the high degree of modularity within SDRT. In particular, the update process is divided into different stages at which different classes of update rules are applied, as is shown schematically in Figure 1.2 below.

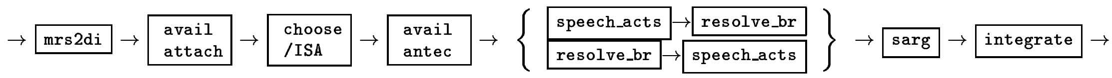

Figure 1.2: The algorithm

The initial stage translates the MRS of the chosen parse into the DIL semantic representation, which abstracts away from certain semantic details, as described above in Section 2.2.[9] At the next stage, an utterance in the context is chosen to which the current utterance can be attached via a rhetorical relation, and this in turn determines which antecedents are available. The preference is to attach to the prior utterance, as explained above. Under certain circumstances, the system tries to add content of an indirect speech act; we'll show how this works in the analysis of example (1.3) in the next section. The following two modules, `speech_acts` and `resolve_bridging`, interact in a special way: the system first tries to infer the speech-act and then uses this information to constrain the temporal bridging relation; if this doesn't succeed, RUDI tries to resolve the bridging relation first, using this additional information to then infer the speech act. Fig. 1.3 shows some of the update rules RUDI uses.[10] *qelab* and *iqap* are rules which don't need access to the intended model (as computed in the TDL), while *plan-correction* and *plan-elaboration* do. In the two latter rules, the function *resolve* relates the anaphor to an antecedent, and *tdr* tries to extend the model built so far (as represented by the TDLRS) so that it satisfies the new set of conditions. The speech act is inferred only if the model can be extended this way.

Including information about anaphora resolution in the antecedent to the rule for inferring *Plan-Elaboration* contrasts with the default rule in SDRT for inferring this speech act, which lacks this information. Adding this information to the antecedent of the monotonic rule is necessary for ensuring that all monotonically derived inferences about speech acts are mutually consistent; the temporal information we've included in the antecedent of the rule ensures that exceptions to inferring *Plan-Elaboration* are stipulated.

Finally, the goal behind the utterance is constructed from the resolved content, including the speech acts (cf. the rules in Fig. 1.6). For example: $R(\alpha, \beta)$ (where R isn't plan-correction) and *plan-correction*(β, γ) entail that the SARG behind γ is to meet at a time which is: the time in the SARG of α minus the time specified in γ. The discourse update is coherent only if temporal information generated by linguistic content (e.g., `avail_antec`, `speech_acts` and `resolve_br`) is *consistent* with these 'cognitive' inferences in `sarg`. This consistency check fails for (1.5), since `avail_antec` constrains *4pm* to be on Saturday, and `speech_acts` constrains the speech act to be *Q-Elab*, but this speech act triggers an inference in `sarg` that *4pm* is 4pm on Sunday. Such inconsistency triggers backtracking, which may ultimately mean choosing an alternative parse for the current utterance ((1.5) is incoherent because no alternative parse is available): One of the goals of this system is to provide a principled symbolic way of choosing a sentence parse that, statistically, may be dispreferred.

[9]At the moment, we have only implemented a few postprocessing rules that deal with our small test corpus. We expect to need lots more of these domain specific rules to extend coverage. Note, however, that the core *logical* rules in the other modules described here are complete as they are.

[10]Note that the rules are *monotonic*, as justified in section 2.2 above. The rule for infering *Plan-Elaboration* deals with the case where β expresses a strict interval of SARG_α as a bad time; a further rule for inferring *Plan-Elaboration* deals with the case where β expresses a good time.

Name:	**Question-Elaboration**
Preconditions:	CUR-UTT.HK.ATTACH $= \langle \alpha, \ldots \rangle$
	CUR-UTT.LING.COND $\supseteq int(\beta)$
Effects:	CUR-UTT.LING.SAS $\supseteq qelab(\alpha, \beta)$
	CUR-UTT.LING.COND $\supseteq temp\_overlap(\text{SARG}_\alpha, t_\beta)$
Name:	**Indirect Question-Answer-Pair**
Preconditions:	CUR-UTT.HK.ATTACH $= \langle \alpha, \ldots \rangle$
	CONTEXT.LING.COND $\supseteq int(\alpha)$
Effects:	CUR-UTT.LING.SAS $\supseteq iqap(\alpha, \beta)$
	CUR-UTT.LING.COND $\supseteq temp\_overlap(\text{SARG}_\alpha, t_\beta)$
Name:	**Plan-Correction**
Preconditions:	CUR-UTT.HK.ATTACH $= \langle \alpha, \ldots \rangle$
	CONTEXT.LING.COND $\supseteq good\_time(\alpha)$
	CUR-UTT.LING.COND $\supseteq bad\_time(\beta)$
	CUR-UTT.LING.COND $\supseteq prpstn(\beta)$
	CUR-UTT.LING.COND $\supseteq temp\_inc(t_\beta, \text{SARG}_\alpha)$
	$resolve$(CUR-UTT.LING)
	$\Gamma =$ CONTEXT.LING.COND $\cup$ CUR-UTT.LING.COND
	tdr(CONTEXT.COG.TDLRS, Γ, CUR-UTT.COG.TDLRS)
	CUR-UTT.COG.TDLRS $\neq \perp$
Effects:	CUR-UTT.LING.SAS $\supseteq plan - correct(\alpha, \beta)$
Name:	**Plan-Elaboration**
Preconditions:	CUR-UTT.HK.ATTACH $= \langle \alpha, \ldots \rangle$
	CONTEXT.LING.COND $\supseteq good\_time(\alpha)$
	CUR-UTT.LING.COND $\supseteq bad\_time(\beta)$
	CUR-UTT.LING.COND $\supseteq prpstn(\beta)$
	CUR-UTT.LING.COND $\supseteq temp\_overlap(\text{SARG}_\alpha, t_\beta)$
	$resolve$(CUR-UTT.LING)
	$\Gamma =$ CONTEXT.LING.COND $\cup$ CUR-UTT.LING.COND
	tdr(CONTEXT.COG.TDLRS, Γ, CUR-UTT.COG.TDLRS)
	CUR-UTT.COG.TDLRS $\neq \perp$
Effects:	CUR-UTT.LING.SAS $\supseteq plan - elab(\alpha, \beta)$

Figure 1.3: The `speech_act`-update rules

3.2 Highlights of a few worked examples

This section shows RUDI at work for a few examples. The first is (1.7) below. The labels of the utterances are given in brackets, and the name of the main temporal referent and what it eventually resolves to according to the rules given in Figure 1.3 is also given:

(1.7)	($h1$)	A:	Can we meet next week?	$x1$	*next week*
	($h2$)	B:	How about Tuesday?	$x2$	*Tuesday of next week*
	($h3$)	A:	Two pm is good for me.	$x3$	*2pm on Tuesday of next week*
	($h4$)	B:	I'm busy then.	$x4$	*2pm on Tuesday of next week*

We will show here how the context is updated by $h4$. At the point of processing $h4$, RUDI has computed the interpretations $x1$–$x3$ indicated above (which in the system are represented by TDLRS). It attached $h2$ to $h1$ via *iqap* and *qelab*, which means it computed that $h2$ gives an (implicit) positive answer to A's question and at the same time elaborates A's proposal (cf. the analysis of (1.1) we gave in section 1). Similarly, $h3$ is attached via *iqap* to $h2$. The SARG of $h3$ is to meet at 2pm on Tuesday of next week. Fig. 1.4 now shows the MRS representation of the compositional semantics of sentence $h4$ that is fed into the system, while Fig. 1.5 shows RUDI's information state after applying the update rules in `mrs2di`, `avail_attach`, `choose/ISA` and `avail_antec`.

The 'previous utterance' attachment rule means RUDI attempts to attach $h4$ to $h3$, making $x3$ the available antecedent to $x4$. The lexical semantics of *then* constrains the bridging relation to be identity; thus the rule `Plan-Correction` applies, for $h4$ specifies a bad time that includes the

$$
\begin{bmatrix}
\text{TOP} & \text{ha1} \\
\text{INDEX} & \text{e2} \\
\text{LISZT} & \left\langle
\begin{bmatrix}\text{pron\_rel} & \\ \text{HANDEL} & \text{ha3} \\ \text{INST} & \text{x4}\end{bmatrix},
\begin{bmatrix}\text{\_busy\_adj\_rel} & \\ \text{HANDEL} & \text{ha3} \\ \text{EVENT} & \text{e2} \\ \text{ARG} & \text{x4}\end{bmatrix},
\begin{bmatrix}\text{time\_rel} & \\ \text{HANDEL} & \text{ha6} \\ \text{INST} & \text{x7}\end{bmatrix}
\begin{bmatrix}\text{def\_rel} & \\ \text{HANDEL} & \text{ha8} \\ \text{BV} & \text{x7} \\ \text{RESTR} & \text{ha9} \\ \text{SCOPE} & \text{ha10}\end{bmatrix},
\begin{bmatrix}\text{\_then\_temp\_rel} & \\ \text{HANDEL} & \text{ha6} \\ \text{ARG} & \text{x7}\end{bmatrix},
\begin{bmatrix}\text{loc\_rel} & \\ \text{HANDEL} & \text{ha3} \\ \text{EVENT} & \text{e13} \\ \text{ARG} & \text{e2} \\ \text{ARG3} & \text{x7}\end{bmatrix},
\begin{bmatrix}\text{prpstn\_rel} & \\ \text{HANDEL} & \text{ha1} \\ \text{SOA} & \text{ha15}\end{bmatrix}
\right\rangle \\
\text{HCONS} & \left\langle \text{ha9 QEQ ha6, ha15 QEQ ha3} \right\rangle
\end{bmatrix}
$$

Figure 1.4: The MRS for "I'm busy then" (simplified)

$$
\begin{bmatrix}
\text{CONTEXT} & \begin{bmatrix}
\text{LING} & \begin{bmatrix}
\text{COND} & \left\{\begin{matrix}
\text{(h1,int),} & \text{(h1,good\_time(x1)),} & \text{(h1,temp\_rel(next,now,x1)),} \\
\text{(h1,interval(x1,week,1)),} & & \text{(h1,unique(x1)),} \\
\text{(h2,int),} & \text{(h2,good\_time(x2)),} & \text{(h2,temp\_rel(temp\_inc,x1,x2)),} \\
\text{(h2,dofw(x2,TUE),} & \text{(h2,unique(x2)),} & \text{(h2,temp\_rel(temp\_overlap,x1,x2)),} \\
\text{(h3,prpstn),} & \text{(h3,good\_time(x3)),} & \text{(h3,numb\_h(x3,14,00,pm)),} \\
\text{(h3, temp\_rel(temp\_overlap(x2,x3))),} & & \text{(h3,temp\_rel(temp\_inc(x2,x3))}
\end{matrix}\right\} \\
\text{SAs} & \left\{\text{iqap(h1, h2), qelab(h1, h2), iqap(h2,h3)}\right\}
\end{bmatrix} \\
\text{COG} & \begin{bmatrix}
\text{SARGs} & \left\{\text{(h1, x1), (h2, x2), (h3, x3)}\right\} \\
\text{TDLRs} & \left\{ now,\ x1,\ x2,\ x3 \right\}
\end{bmatrix}
\end{bmatrix} \\
\text{CUR-UTT} & \begin{bmatrix}
\text{LING} & \begin{bmatrix}
\text{COND} & \left\{\begin{matrix}\text{(h4,prpstn), (h4,bad\_time(x4)),} \\ \text{(h4,temp\_rel(temp\_ident,z0,x4)), (h4,time(x4))}\end{matrix}\right\} \\
\text{ANAPH} & \left\{\text{(z0,}\left\{\text{x3, now}\right\}\text{)}\right\} \\
\text{B-RELS} & \{\} \\
\text{SAs} & \{\}
\end{bmatrix} \\
\text{COG} & \begin{bmatrix}\text{SARG} & \text{--}\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
$$

Figure 1.5: Adding "I'm busy then" to the information state

good time from $h3$ (cf. Fig. 1.3 above). This updated content triggers inferences in `sarg` (cf. Fig. 1.6): $h4$'s SARG is $h2$'s SARG (since $h3$ was attached to this) minus $x4$; i.e., the SARG of $h4$ is to meet on Tuesday of next week, but not at 2pm (cf. Fig. 1.6). The interpretation of (1.4) is similar in the situation *now* is Monday 12th February 2001 (so *next week* is 19th–25th):

(1.4) A: Can we meet next week?
B: I'm busy from the 16th to the 25th.

Let's now look at an example where the content of an indirect speech act has to be computed explicitly. We proposed earlier that B's response in (1.3) should be analysed as an implicit plan-correction; ie. B's question tacitly rejects A's SARG.

(1.3) A: Let's meet next Saturday.
B: How about Sunday? (↖ ISA: *That is bad for me.*)

To account for this analysis, we have to compute the content of this implicit speech act. This is done as follows. Suppose we have labelled A's utterance $h1$ and B's $h2$, and their temporal referents $x1$ and $x2$. Based on the linguistic clue "sentence mood", `speech_acts` infers *q-elab* (as it does for example (1.1)). However, this speech act has as monotonic consequence that the temporal relation *temp_overlap*$(x1, x2)$ must hold (see above section 2.2), and this is rejected by the `tdr`. None of the other rules fire, and therefore we have to backtrack. On entering `choose/ISA` via backtracking, we call the indirect speech act module. In this module we have update rules specifying that two *good_time*s that do not temporally overlap trigger the introduction of a new label, which we will here call $h2'$, with the content that an explicit *That is bad for me* would

Name:	**Q-Elab**
Preconditions:	$\textsc{cur-utt.ling.sas} \supseteq qelab(\alpha, \beta)$
Effects:	$t_\nu = \textsc{sarg}_\alpha \cap t_\beta$
	$\textsc{cur-utt.cog.sargs} \supseteq \langle \beta, t_\nu \rangle$
Name:	**IQAP-good**
Preconditions:	$\textsc{cur-utt.ling.sas} \supseteq iqap(\alpha, \beta)$
	$\textsc{context.ling.cond} \supseteq good\_time(\beta)$
Effects:	$t_\nu = \textsc{sarg}_\alpha \cap t_\beta$
	$\textsc{cur-utt.cog.sargs} \supseteq \langle \beta, t_\nu \rangle$
Name:	**IQAP-bad**
Preconditions:	$\textsc{cur-utt.ling.sas} \supseteq iqap(\alpha, \beta)$
	$\textsc{context.ling.cond} \supseteq bad\_time(\beta)$
Effects:	$\textsc{cur-utt.cog.sargs} \supseteq \langle \beta, t_\nu \rangle$
	$t_\nu = \textsc{sarg}_\alpha \cap \overline{t_\beta}$
Name:	**Plan-Correction**
Preconditions:	$\textsc{cur-utt.ling.sas} \supseteq plan\text{-}correct(\alpha, \beta)$
	$\textsc{context.ling.sas} \supseteq R(\gamma, \alpha)$
	$R \neq plan\text{-}correct$
Effects:	$t_\nu = \textsc{sarg}_\gamma \cap \overline{t_\beta}$
Name:	**Plan-Elab good**
Preconditions:	$\textsc{cur-utt.ling.sas} \supseteq plan\text{-}elab(\alpha, \beta)$
	$\textsc{context.ling.cond} \supseteq good\_time(\beta)$
Effects:	$t_\nu = \textsc{sarg}_\alpha \cap t_\beta$
	$\textsc{cur-utt.cog.sargs} \supseteq \langle \beta, t_\nu \rangle$
Name:	**Plan-Elab bad**
Preconditions:	$\textsc{cur-utt.ling.sas} \supseteq plan\text{-}elab(\alpha, \beta)$
	$\textsc{context.ling.cond} \supseteq good\_time(\beta)$
Effects:	$t_\nu = \textsc{sarg}_\alpha \cap \overline{t_\beta}$
	$\textsc{cur-utt.cog.sargs} \supseteq \langle \beta, t_\nu \rangle$

Figure 1.6: The `sarg`-update rules

get.[11] The dialogue is then processed with this additional content, which means that we infer *plan-correct*$(h1, h2')$ and *q-elab*$(h2', h2)$.

The reason that we compute the (labelled) content of the indirect speech act explicitly in this case is because plan-corrections are constrained to take propositions as their second arguments (since they are a kind of assertion); see (Asher & Lascarides 2001) for details. Generating this content explicitly allows us to capture rhetorical relations between the indirect speech act and the 'direct' one that could not be captured otherwise. This contrasts with the case of indirect answers, which can be of any sentence type (e.g., a question or a request can entail content from which the interpreter can compute a direct answer, as demanded by the semantics of *IQAP*).

B's question in (1.3) contrasts with the question (1.5d), which cannot be interpreted as an indirect speech act of plan-correction. This is because A has already stipulated that Saturday afternoon is a bad time (for him). And, informally, the module `choose/ISA` fails to generate a coherent interpretation in this case, to reflect the fact that when B wants A to revise his assessment of t as a bad time, he needs to do this explicitly (we forego stipulating the formal rule here). So, for example, inserting the plan-corrective move *B: But I would much prefer to meet you on Saturday afternoon* between (1.5c) and (1.5d) ameliorates the incoherence (note that (1.5d) would attach to this explicit *plan-correction* with *q-elab*). `choose/ISA` failing to provide a discourse update triggers further backtracking; an alternative parse of the sentence must be chosen, but there isn't one, thereby yielding discourse incoherence.

[11] These ISA-rules have to be constrained carefully, since there seem to be strong conventional constraints on how such an indirect plan-correction can be conveyed. For example, there must be contrasting elements present, which explains why substituting B's utterance in (1.3) with "How about the 15th?" would make the dialogue sound a lot worse. Investigating the exact nature of these constraints remains as future work.

4 Related Work

(Stede et al. 1998) compute the temporal content of scheduling dialogues in German. Their approach to representing the temporal domain is similar, but they don't offer principled constraints for resolving anaphora. (Wiebe et al. 1998) adopt a data-intensive approach to interpreting temporal expressions. We are, however, also interested in predicting when a definite description is coherent and when it's not, which this approach doesn't do.

Interpreting questions and their answers is crucial in this domain. (Traum et al. 1999) analyse questions and answers by implementing the QUD-model within the TRINDI dialogue management system. The QUD model constructs an ordered stack of *questions under discussion*, which determines what utterances would be (currently) felicitous. (Cooper et al. 2000) develop a method of "question accommodation" to deal with cases where felicitous, indicative utterances provide information that doesn't answer any question on the stack. It seems, however, that even in our domain, additional mechanisms to this are needed to account for some implicatures. Even for a simple exchange like (1.1), the QUD model as it stands predicts that two questions are on the stack; however, it fails to model that B's intention in (1.1) was not simply to ask a question, but also to implicate an *answer* to A's question (in the positive). It fails to detect this because the QUD model doesn't reason about the second question's rhetorical function in the context of the first question. This gap in the theory also means that the rule for accommodating questions overgenerates. Because the accommodated question need not be rhetorically linked to the existing QUDs, B's utterance in (1.8) can trigger the accommodation of a question like "On which day can we meet?", thereby predicting (1.8) is acceptable, contrary to intuitions:

(1.8) A: [said on the 1st] Can we meet next week?
B: ??The 20th is fine.

It is quite likely that the QUD-model could be extended to overcome these problems. However, we hope that by allowing access to a richer discourse structure than a stack of questions, we will constrain the necessary inferences in a more effective manner.

5 Conclusion

We have developed a system which explores the information flow between recognising speech acts, inferring the underlying goals of utterances and resolving semantic underspecification that's generated by the grammar within the domain of scheduling dialogues. The main feature of the system was to approximate a logically precise theory of the semantic and pragmatic interpretation of discourse, by making assumptions that DPs don't digress from the main goal, that they always believe each other, and by 'short-circuiting' reasoning about domain-level plans to meet (e.g., that you can't meet and go to the dentist at the same time) within a post-processing module. This allowed us to encode within the system the simpler and more computationally tractable axioms that are derived (manually) from these assumptions within the underlying logical theory. We aim eventually to test the extent to which the nonmonotonic reasoning that generally underpins computing implicatures can be made monotonic in relatively restricted domains, and to apply the result to the processing of fragments. We actually believe that the monotonic approximation of the theory will be pushed to its boundaries even in the very simple domain we've chosen here, thereby demonstrating default reasoning is an essential component to any realistic, rule-based dialogue system.

Acknowledgements

We would like to thank the Dialogue Systems Group Edinburgh for helpful discussion. This research was partially supported by the National Science Foundation, grant number IRI-9612682 to Stanford University. Alex Lascarides is supported by an ESRC (UK) research fellowship.

Bibliography

Asher, N. (1993), *Reference to Abstract Objects in Discourse*, Studies in Linguistics and Philosophy, Kluwer Academic Publisher, Dordrecht.

Asher, N. & Lascarides, A. (1998), 'Bridging', *Journal of Semantics* **15**(1), 83–113.

Asher, N. & Lascarides, A. (2001), 'Indirect speech acts', *Synthese* . to appear.

Clark, H. (1975), Bridging, *in* R. Schank & B. Nash-Webber, eds, 'Theoretical Issues in Natural Language Processing', MIT Press, Cambridge, Mass.

Cooper, R., Engdahl, E., Larsson, S. & Ericsson, S. (2000), Accommodating questions and the nature of QUD, *in* 'Proceedings of Götalog 2000', Gotheburg.

Copestake, A. & Flickinger, D. (2000), An open-source grammar development environment and broad-coverage english grammar using HPSG, *in* 'Proceedings of the 2nd Linguistic Resources and Evaluation Conference', Athens, Greece, pp. 591–600.

Copestake, A., Flickinger, D., Sag, I. & Pollard, C. (1999), Minimal recursion semantics: An introduction. Stanford University, Stanford, CA.
*http://www-csli.stanford.edu/~aac/papers/newmrs.ps

Groenendijk, J. & Stokhof, M. (1984), Studies on the Semantics and Pragmatics of Questions, PhD thesis, Centrale Interfaculteit, Amsterdam.

Hawkins, J. (1978), *Definiteness and Indefiniteness*, Croom Helm.

Hobbs, J. R., Stickel, M., Appelt, D. & Martin, P. (1993), 'Interpretation as abudction', *Artificial Intelligence* **63**, 69–142.

Kamp, H. & Reyle, U. (1993), *From Discourse to Logic: Introduction to Model-theoretic Semantics, Logic and Discourse Representation Theory*, Kluwer Academic Publishers.

Lascarides, A. & Asher, N. (1993), 'Temporal interpretation, discourse relations and commonsense entailment', *Linguistics and Philosophy* **16**(5), 437–493.

Lascarides, A. & Asher, N. (1999), Cognitive states, discourse structure and the content of dialogue, *in* 'Proceedings to Amstelogue 1999'.

Lochbaum, K. (1998), 'A collaborative planning model of intentional structure', *Computational Linguistics* **24**(4), 525–572.

Morgan, J. L. (1975), Some interactions of syntax and pragmatics, *in* P. Cole, ed., 'Syntax and Semantics Volume 9: Pragmatics', Academic Press, pp. 261–280.

Reyle, U. (1993), 'Dealing with ambiguities by underspecification: Construction, representation and deduction', *Journal of Semantics* **10**, 123–179.

Searle, J. (1967), *Speech Acts*, CUP.

Stede, M., Haas, S. & Küssner, U. (1998), Tracking and understanding temporal descriptions in dialogue, Verbmobil Report 232, Technische Universität Berlin.

Traum, D., Bos, J., Cooper, R., Larsson, S., Lewin, I., Matheson, C. & Poesio, M. (1999), A model of dialogue moves and information state revision, Trindi Deliverable D2.1, University of Gothenburg.

Vieira, R. & Poesio, M. (2000), 'An empirically-based system for processing definite descriptions', *Computational Linguistics* **26**(4).

Wahlster, W., ed. (2000), *Verbmobil: Foundations of Speech-to-Speech Translation*, Artificial Intelligence, Springer, Berlin, Heidelberg.

Wiebe, J. M., O'Hara, T. P., Öhrström-Sandgren, T. & McKeever, K. J. (1998), 'An empirical approach to temporal reference resolution', *Journal of Artificial Intelligence Research* **9**, 247–293.