Chapter 7

Discourse Relations for Dialogue

7.1 Introduction

In the preceding chapters we explored SDRT's analysis of single author texts and monologue. This chapter and the next extend the *logic of information content* to handle dialogue. Chapter 9 will extend the *glue logic* accordingly.

Up to now, we have countenanced rhetorical relations whose semantics involve only the contents of the clauses they relate (e.g., *Elaboration*), or text organisation (e.g., *Parallel*). We call these *domain-level* rhetorical relations. But as Moore and Pollack (1992) show, rhetorical relations whose semantics involve the cognitive states of discourse participants are also important, because these record an interplay between beliefs and intentions on the one hand and discourse content on the other. In this chapter, we will introduce rhetorical relations that take labels of questions and requests as arguments or have a semantics defined in terms of cognitive states. The 'cognitive level' rhetorical relations will encode an information flow between goals, beliefs and discourse content. In line with Grice (1975) and the plan-recognition approaches to discourse interpretation, we develop a theory where general cognitive principles affect discourse interpretation. In fact, we'll show in Chapter 9 that some of the axioms for constructing logical form are derived from a logical model of agent rationality and cooperativity. But just as the glue logic has only restricted access to compositional and lexical semantics (see Chapter 5), it has only restricted access to cognitive states too. This is perhaps the main distinguishing feature of SDRT over the approaches to interpretation of Hobbs *et al.* (1993), Grosz and Sidner (1986, 1990) and others.

7.2 Why Dialogue and Monologue are Similar

We now highlight ways in which SDRT is useful for analysing dialogue.

7.2.1 Availability in Dialogue

In previous chapters and elsewhere (e.g., Asher (1993)), we argued that DRT's notion of accessibility is necessary for resolving anaphora in monologue but not

sufficient; one must combine it with constraints based on discourse structure. The same is true for dialogue (Asher and Lascarides, 1998b). Consider examples (1abc) and (1ab′c) from Chapter 3:

(1) a. A: How can I get to 6th Street?

b. B: You can ask someone Downtown.

b′. B: There's someone Downtown that you could ask.

c. A: What's his name?

Intuitively, the pronominal reference *his* in (1c) is less acceptable in the context (1abc) than it is in (1ab′c). Plan recognition approaches to dialogue interpretation (e.g., Lochbaum (1998)) would have to account for this on the basis that the context (1ab) yields different inferences about B's beliefs and intentions than the context (1ab′) (see Chapter 3 for details). But (1b) and (1b′) have similar linguistic forms, and so predicting the cognitive differences would be problematic.

An extension of the DRT account of anaphora offers an alternative, however, as suggested in Chapter 3. DRT allows scope differences between quantifiers (e.g., *someone*) and modalities (e.g., *can*) to affect the content of pronouns. More precisely, compositional semantics by itself yields the logical form (2b′) (in slightly simplified form) for (1b′); note that *someone* takes semantic scope over *can* because of the *there*-construction (Sag and Wasow, 1999). In contrast, the grammar yields an underspecified logical form for (1b), reflecting a semantic scope ambiguity between *someone* and *can*. The possible resolved logical forms of (1b) correspond (again slightly simplified) to (2b) and (2b′):[1]

(2) b.

you, d
$downtown(d)$,
Can $\begin{array}{\|l\|}\hline q, y \\ \hline you\ ask\ y\ question\ q \\ q = ?,\ location(y, d) \\ \hline \end{array}$

b′.

you, y, d
$downtown(d),\ location(y, d)$
Can $\begin{array}{\|l\|}\hline q \\ \hline you\ ask\ y\ question\ q \\ q = ? \\ \hline \end{array}$

[1]The condition $q = ?$ is generated by *ask*, which takes a question as one of its arguments. SDRT predicts that it must resolve to (1a), since this is the only available candidate.

The difference between these DRSs lies in where in the discourse structures the discourse referent y is introduced: in (2b) it outscopes the modal *Can*, while in (2b′) it doesn't. Now, although both (2b/b′) depict possible readings of (1b), the preferred reading is the one where the scope ordering is the same as the c-commanding one (Chierchia and McConnell-Ginet, 1992), making (2b) the *preferred* reading of (1b) (in the absence of information to the contrary).

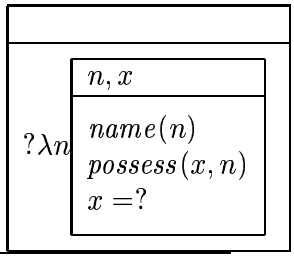
The grammar will produce an anaphoric expression $x =?$ for the pronoun *his* in (1c). The only candidate antecedent for x that has the appropriate case marking is y. But y is not *available* in (2b), assuming that the representation of (1c) is *not* attached to it with *Contrast* or *Parallel* (see the availability definitions in Chapter 4). This is because y is introduced in an embedded DRS in (2b). In contrast, y is available in (2b′). Thus availability in SDRT predicts that the pronoun in (1c) is uninterpretable relative to the preferred reading of (1b), whereas its acceptable relative to the logical form (2b′) of (1b′). In other words, the pronoun in (1c) forces a dispreferred resolution of the scope ambiguity of (1b) in (1abc), but not in (1ab′c).

This analysis makes critical use of the assumption that (1c) attaches to (1b) with a relation other than *Contrast*. Changing (1c) to (1c′), where *but* is a monotonic clue for a *Contrast* relation, seems to improve acceptability:[2]

(1) a. A: How do I get to 6th Street?
b. B: You can ask someone downtown.
c′. A: But what's his name?

We can use SDRT to explain this observation too. Recall from Chapter 4 that *Contrast* can hold only if there's a partial isomorphism between the DRS structures of (1b) and (1c′). And this isomorphic mapping affects anaphoric possibilities: a discourse referent y that's introduced in a sub-DRS D_1 of (1b) can act as an antecedent to an anaphor in the sub-DRS D_2 of (1c′) if the partial isomorphism maps D_1 to D_2. So *Contrast* can allow elements in embedded DRSs to be antecedents, contrary to DRT's predictions. So, how does this affect the interpretation of (1abc′)? Well, the anaphoric condition $x =?$ in (1c′) is in a DRS D_2 that's embedded within the scope of a λ operator that's generated by the *wh*-element *what* (see Chapter 2 for the semantics of *wh*-elements):

(2) c′.



[2] (1c′) may be attached to (1b) with other relations too. In fact, we'll see later that it also attaches with the relation *Q-Elab*, which is discussed in detail in Section 7.6.2.

Contrast induces a total isomorphism between the *preferred* reading (2b) of (1b) and the representation (2c′) of (1c′), with the (sub)-DRS D_1 in (2b) that introduces y mapped to D_2. So *Contrast* makes y in (2b) available to $x =?$ in (2c′). Thus according to SDRT, the pronoun in (1abc′) is acceptable, even with (1b)'s preferred interpretation.

It now only remains to explain why (1c) couldn't attach with *Contrast* to (1b), even though the very similar (1c′) attaches this way. As we observed in Chapter 4, the semantics of *Contrast* must reflect the fact that a *monotonic* clue for the *Contrast* relation, such as the cue phrase *but* or intonation, is necessary when one proposition denies a default consequence of the other. For example, observe that (3c) denies the default consequence of (3a) that John hates football, and (3ac) is odd unless there's a rise-fall-rise intonation on (3c):

(3) a. John hates sport.

b. But he loves football.

c. ??He loves football.

This property of *Contrast* prevails for dialogue too (i.e., when (3a) and (3c) are uttered by different speakers). And because of this, *Contrast* can't be part of the interpretation of (1abc). We'll discuss *Contrast* in dialogue in more detail shortly.

Dialogue (1) shows that small changes to surface form affect anaphora. SDRT makes use of a semantics in which the scope differences in (1b) vs. (1b′) and the rhetorical differences in (1abc) vs. (1abc′) affect anaphoric possibilities. In the above analysis, we didn't reason about A's and B's cognitive states to predict the differences in acceptability; instead we used the compositional semantics of the context *directly*, together with preferences for resolving semantic scope ambiguities based on *linguistic* concepts such as c-command. One could conceivably explain the differences in acceptability while relying only on the underlying communicative goals. But we feel that if we can avoid additional, defeasible and often ill-understood inferences (from what is said to what was intended), we should; the theory that doesn't use these inferences will be both more robust and more empirically testable, as well as more elegant.

7.2.2 Rhetorical Relations in Monologue and Dialogue

As conversational analysts have made clear (e.g. Sacks (1992), Sacks *et al.* (1974), Schegloff (1972)), dialogue turns provide an important clue to discourse structure. We will attempt to refine this: we'll use the rhetorical relations from previous chapters to connect constituents across dialogue turns, thereby enabling their semantics to predict inferences about content of the kind discussed by Sacks and others.

Let's illustrate this by example. In dialogue (4), taken from Sacks (1992), Henry and Mel jointly contribute to an *Elaboration* of Joe's contribution:[3]

(4) a. Joe: We were having an automobile discussion...
b. Henry: discussing the psychological motives for
c. Mel: drag racing in the streets.

Dialogue (5), also from Sack's corpus, features a *Contrast* across turns between (5c) and (5d) (as indicated by *but*); and an *Elaboration* between (5d) and *he starts talking, interacting* in (5a):

(5) a. A: There was this guy. He came to the sessions. He never said anything. Then one day he shows up, and he starts talking, interacting.
b. B: Why didn't he say anything before?
c. A: Dunno. Shy maybe.
d. But anyway he's yammerin away and telling these jokes...

Recognising these rhetorical relations is important for understanding the dialogue, because, just as with monologues, their truth conditional entailments (see Chapter 4) capture some intuitive aspects of the dialogue's meaning. For example, in (4), the *Elaboration* yields an inference that the automobile discussion included a discussion of drag racing—indeed, you learn that drag racing is something to do with automobiles if you didn't know this before. And in (5), the *Elaboration* between the final clause in (5a) and (5d) correctly predicts that the event of telling jokes is part of the event of his talking and interacting.[4]

The constructed example (6) also features rhetorical relations across turns:

(6) a. A: Jones got kicked out of school.
b. B: He was caught buying liquor.
c. C: That was really dumb of him.

Interpreting B's utterance as providing an *Explanation* for (6a) accounts for the inference that Jones' buying liquor caused him to be kicked out of school.

Narration can also hold across conversational turns, though some extra assumptions about the context are required: the *Narration* must be about events that both participants witnessed or agreed to imagine together. This ensures that they are telling the same story. For example, suppose that A and B are recounting to C what happened while they were minding the store:

[3]Note that Henry's and Mel's utterances contribute a *single* proposition. We gloss over this here, however.

[4]This latter constituent (5a) is available to (5d), because the question answer session in (5bc) induces subordination; we'll discuss this shortly. In fact, the cue phrase *anyway* is a clue for this discourse pop.

(7) a. A: A well dressed gentleman came in this morning.
b. B: He asked to see the most expensive suit.
c. C: Did he buy it?

Connecting A's and B's speech acts with *Narration* accounts for the temporal progression of the events, which forms part of the natural interpretation of (7ab).

Relations like *Parallel* and *Contrast* also apply to dialogue (e.g., (5cd)), though sometimes they have additional inferential effects to those that SDRT predicts for monologue. Ducrot (1984) and Moeschler (1989) suggest that there is some nonmonotonic consequence of some generalisation of the first proposition that is contradicted by some generalisation of the second. In (8) (an English translation of an example from (Moeschler, 1989)), passing exams is a nonmonotonic consequence of working hard, but (8b) contradicts this:

(8) a. A: Paul worked really hard this semester.
b. B: But he failed his exams.

We readily read into the exchange in (8) an implicit rebuttal by B of A's claim. While the analysis of contrasts in monologue is compatible with this observation (see Asher (1993), Txurruka (1997)), it indicates that *Contrast* across dialogue turns sometimes invites additional inferences. We will analyse disputes like (8) in Chapter 8.

7.3 Some Differences Between Monologue and Dialogue

Although SDRT's existing mechanisms are useful, we must extend the framework considerably. Dialogue is different (and harder) than monologue, because with the introduction of more than one participant there emerges the possibility of information exchange, cooperation, agreement and disagreement. Discourse structure must also incorporate questions and requests. We highlight here some of dialogue's distinctive properties.

7.3.1 An SDRS for Each Agent?

Each dialogue agent has his or her own interpretation of the dialogue. So we must decide from whose perspective to model it. Agents' interpretations will often coincide; indeed, it seems plausible that when there's no reason to believe the contrary, it's mutually believed that everyone has the same SDRS/interpretation. This default dovetails with the assumption we make in SDRT, that all agents mutually believe all the axioms of SDRT (see Chapter 9 for details); i.e., they abide by the same rules of interpretation. So in particular, they mutually believe the meaning postulates on rhetorical relations in the logic of information content, and also the default axioms for inferring rhetorical relations in the glue logic.

But when are the SDRSs that the dialogue participants build the same and when are they different? Clearly, A and B will build different SDRSs if A mishears or misunderstands the compositional and lexical semantic content of B's utterance. SDRT update already predicts this: A and B will add different new information to their SDRSs, and may compute different rhetorical relations in the glue logic because different axioms will apply.

Agents A and B believe the same *axioms* of interpretation, but they could have different beliefs about contingent propositions. So, what happens when both A and B construct the same compositional and lexical semantics of the new information (which B contributed, say), but they have different (contingent) beliefs (e.g., B believes this new information is true, but A believes it's false)? The answer to this question reveals ways in which SDRT parts company with approaches to discourse interpretation like Hobbs *et al.*'s (1993) and Litman and Allen's (1990), where discourse interpretation amounts to updating a set of beliefs in an unmodular architecture.

An important feature of SDRT is that the reasoning an agent does to compute discourse content and the reasoning he does with his own beliefs about (contingent) states of affairs (e.g., whether he believes this content to be true) are kept separate: the two logics are related, but unlike Hobbs *et al.* (1993) the logic in which A constructs the logical form of the discourse has only restricted access to his own beliefs. This means that even when A doesn't believe what B says, A can still construct the interpretation that B intended, *without* reasoning about what B actually believes or intends at all. Consider dialogue (9) in a context where A believes (9b) to be false.

(9) a. A: Max is in jail.

b. B: Yeah, he was caught embezzling company funds.

Intuitively, B's intention was to provide an *Explanation* for A's utterance. But given the veridicality of *Explanation*, as given by the schema introduced in Chapter 4 and repeated below, this interpretation is inconsistent with A's beliefs about the world (though perhaps not his beliefs about B's beliefs).

- `Satisfaction Schema for Veridical Rhetorical Relations:`

$$(w, f)[\![R(\alpha, \beta)]\!]_M(w', g) \text{ iff } (w, f)[\![K_\alpha \wedge K_\beta \wedge \phi_{R(\alpha,\beta)}]\!]_M(w', g)$$

Elaboration, *Narration*, *Explanation*, *Parallel*, *Contrast*, *Background*, *Result* are all veridical.

So what SDRS does A build? Well, he must compute a rhetorical connection in the glue logic. Let's assume, as intuitions dictate, that there is a default axiom in this logic, whose antecedent is verified by the compositional and lexical semantics of (9a) and (9b), and whose consequent is $\mathit{Explanation}(9a, 9b)$. Now assume that

the glue logic *doesn't* have access to information about A's beliefs, even for A! So A is in a situation where nothing blocks him from inferring $Explanation(9a, 9b)$ in the glue logic (in particular, A's belief that (9b) is false isn't in the glue logic, and so doesn't block this inference). So just like B, A infers the *Explanation* relation. Hence, his interpretation of the discourse content is the same as B's. And he's constructed this interpretation *without* reasoning about B's beliefs and goals (although he may infer, in a separate logic of cognitive modelling, properties of B's cognitive state once he's constructed this logical form; e.g., sincerity axioms in this logic plus the veridicality of *Explanation* would support an inference that B believes (9b); details of this are given in Chapter 9).

The difference between A and B is not in their different interpretations of the content of (9), but rather in their beliefs about whether that content is true. A will detect that he doesn't believe (9) once he has built its SDRS; he won't detect this within the glue logic, but rather via the logic of information content (since this logic contains the premise, at least for A, that (9b) is false).

This contrasts with Hobbs *et al.*'s (1993) approach, where interpreting a discourse amounts to constructing an abductive proof that it's a discourse segment from premises which are a set of beliefs. The abduced propositions are then new beliefs. Hobbs *et al.* suggest that the proof and assumptions must be consistent with the interpreter's beliefs. But undesirable consequences follow from this. One can either assume that A interprets (9) without considering his model of B's beliefs and intentions. In this case, A will compute a meaning for (9) that doesn't feature an *Explanation* relation (since this would be inconsistent with his belief that (9b) is false), and he won't detect that this interpretation is distinct from what B intended (because he won't be reasoning about what B intended). This strategy is undesirable, because it makes A's detecting his disagreement with B logically impossible. The alternative is that A computes the meaning of (9) that B intended (i.e., an *Explanation* relation), but he must base the abductive proof that *Explanation* holds on his beliefs about B's private beliefs and intentions (for *Explanation* is inconsistent with his own private beliefs and hence with A's and B's mutual beliefs too). But reasoning with a model of another agent's cognitive state should be avoided whenever possible, because such models are always fragmentary and incomplete.

SDRT provides a way for A to avoid reasoning about B's beliefs and goals when constructing his interpretation of (9). We can avail ourselves of this alternative strategy thanks to SDRT's separation of the logic in which you construct the interpretation (i.e., the glue logic) from the logic in which you reason about what's privately or mutually believed. In SDRT, A computes the rhetorical relation for (9) via compositional and lexical semantics *and nothing else*. He can *then* infer that B has the requisite beliefs (and goals) for this interpretation (e.g., that B believes (9b)) and may at this point detect disagreement. This strategy for building logical form exploits the fact that agents are helpful, and typically ensure

that their observable actions (i.e., the compositional and lexical semantics of their utterances) are sufficient for conveying their communicative intentions. Thanks to this cooperativity, agents can align their dialogue interpretations while reasoning by and large only with observable facts.

We'll have much more to say about disagreements in dialogue in the next chapter. The main point we want to make here is that A and B will build identical SDRSs (subject to alphabetic variance), even when they disagree about which propositions in the discourse are true. So, the SDRSs being identical is a sign that A and B agree on what the content of the dialogue is, but it's *not* necessarily a sign that they agree that content is true. The SDRSs *may* represent content that's *settled* or agreed to be true. Such an SDRS might represent the *common ground* in Stalnaker's (1978) sense (see also Clark (1992)). On the other hand, if one of the dialogue participants doesn't believe the SDRS he's built for the dialogue, then clearly it doesn't represent common ground. Subsequent dialogue moves may indicate this. Thus, one very important question for us is: when can dialogue agents assume that they agree, thereby increasing their common ground? Given that in SDRT an agent can construct an interpretation that's inconsistent with his beliefs, building identical SDRSs is insufficient (though necessary) for increasing the common ground. But certain discourse moves are good clues for agreement, and we need to model this precisely in order to make the right predictions about agreement. We address this in Chapter 8.

There are clearly times when A will need to reason about B's cognitive state in order to make sense of B's utterance. In such cases, the glue logic and the logic for cognitive modelling interact, even though they are separate. For example, when A cannot construct anything but a very underspecified SDRS for B's contribution (because, for example, its compositional and lexical semantics is insufficient for inferring any rhetorical relation), then A may try to model what B believes or intends, in an attempt to work out what rhetorical connection B intended. We'll provide examples of this in Chapter 9.

7.3.2 Cognitive Constraints on Anaphora

Even though we distinguish between building the logical form of a discourse and the participants believing that logical form, we must account for interactions between discourse content and cognitive states. We saw examples of this interaction in discourses (10) and (11) from Chapter 2, where assumptions about what's desirable affect whether an imperative is commanded.

(10) Come home by 5pm and we will finish the shelves tonight.

(11) Smoke a packet of cigarettes a day and you will die before you're 50.

A complete account of anaphora involves inferences about cognitive states too. Consider dialogue (12): if one ignores the goals behind the utterances, then one can't explain the difference in acceptability between (12a-d,e) vs. (12a-d,e′).

(12) a. A: How about meeting next weekend?
b. B: That sounds good.
c. Shall we meet on Saturday afternoon?
d. A: I'm afraid I'm busy then.
e. ??How about 3pm?
e′. How about 11am?

To see why, consider how *3pm* in (12e) is interpreted. This anaphoric expression must be related to an accessible antecedent. Constraints on antecedents of the kind described in DRT and the previous chapters restrict the possibilities to a subset of the explicit linguistic expressions. So *at most* the possible antecedents are: *then* (which is resolved to Saturday afternoon) in (12d); *Saturday* and *Saturday afternoon* in (12c); and *next weekend* in (12a). Exactly which of these are available will depend on the discourse structure, a matter we return to in detail in Section 7.6.2.

By Clark's (1977) uniqueness constraints on bridging, next weekend cannot be the antecedent, because this won't yield a unique bridging relation (there are two 3pms in next weekend). Saturday, on the other hand, makes *3pm* resolve to 3pm on Saturday (and the acceptability of (12e′) is evidence that Saturday is available in (12a–d)). The important point is that linguistic constraints on anaphora block *3pm* from resolving to 3pm on Sunday. But the goals explain why *3pm can't* be 3pm on Saturday, even though the linguistic constraints rule out the alternatives: why is *A* asking whether they can meet at 3pm on Saturday if he has already implied via (12d) that he doesn't want to meet then? So on the one hand, the goal that's been established when (12e) is uttered must play a role in constraining its content, for otherwise *3pm* would be successfully interpreted as 3pm on Saturday. And on the other hand, if one uses *only* the goals to compute the meaning of *3pm*, then one would predict that *3pm* in (12e) successfully refers to 3pm on Sunday, since one knows when interpreting (12e) that the goal is to meet on Saturday morning or Sunday, making *3pm* refer uniquely (to Sunday).

This gives rise to two important criteria. First, dialogues like (1) and (12) show that anaphora resolution depends on linguistic accessibility and discourse structure. Second, constraints on anaphora resolution must take account of cognitive states in at least some discourses (notably dialogues), but ideally without a detailed model of cognitive states. SDRT as we have developed it so far doesn't meet this second criterion, because the rhetorical relations we've introduced have a semantics defined entirely at the domain-level (e.g., they constrain the temporal relations between the events described) and cognitive states are ignored.

We'll rectify this by adding new discourse relations whose semantics are defined in terms of beliefs and goals. This will reflect the fact that some speech acts reveal things about the cognitive states of the agents, as well as perhaps things about the events and individuals described in the utterances.

One such cognitive-level discourse relation is *Q-Elab* (standing for Question Elaboration); this will connect (12c) to (12b) and (12e′) to (12d). *Q-Elab* will capture an intuitive aspect of cooperativity during conversation: namely, asking a question can be part of a deliberate strategy for achieving a goal which lay behind a prior utterance. Specifically, when a question β attaches to a previous utterance α with *Q-Elab*, then any answer to the question β must elaborate a plan to achieve the goal that lay behind saying α (note that the plan addresses the speaker's *goal* behind α, and not the content of α itself). So the semantics of *Q-Elab* is defined in terms of goals as well as 'domain-level' content.

For example, *Q-Elab*$(12d, 12e')$ is consistent, because all answers to (12e′) (both positive and negative) provide useful information for implementing a plan to achieve the goal behind (12d), which is to find a meeting time within Saturday morning or Sunday. This is because all answers reduce the search space of candidate times (if the answer is positive, the space is reduced to 11am on Saturday; if the answer is negative, it's reduced by excluding 11am on Saturday). *Q-Elab* is not only consistent, but it also seems plausible that it actually holds, because it reflects A's attempts to be cooperative with B (i.e., make dialogue moves that help A and B find a time to meet).

Now in general, if we infer that *Q-Elab*(α, β) holds (and in fact, we'll show in Chapter 9 that one can infer this relation via the sentence moods of α and β), we may learn something about the goals that lay behind α by speculating about which plans answers to β elaborate. Similarly, we may learn things about the content of β because from a known goal behind uttering α we will accommodate additional content to the question β in order to ensure that its answers elaborate a plan to achieve that goal. So the semantics of relations like *Q-Elab* can model some aspects of the information flow between dialogue content and cognitive states.

This helps to explain the incoherence of (12e) in the following way. The only available interpretation of *3pm* (i.e., 3pm on Saturday) is inconsistent with the semantics of *Q-Elab*$(12d, 12e)$, because knowing whether or not B can meet A at 3pm on Saturday doesn't help to implement a plan for finding a time to meet on Saturday morning or Sunday. But the glue-logic axioms will imply *Q-Elab*$(12d, 12e)$, because it won't have full access to information about content and goals (the inference will follow on the basis of sentence moods), and this creates an inconsistency in the logic of information content. Further, revising the glue logic inferences upon discovering this inconsistency will fail to yield an alternative attachment. Thus the discourse is incoherent because we can't construct a consistent interpretation.

Grosz and Sidner (1986) and many others have emphasised the importance of modelling goals and plans. Dialogue (12) shows one can't ignore goals. But equally, one cannot ignore constraints on anaphora of the kind we've developed so far in SDRT. Grosz and Sidner's (1986) model of linguistic structure is quite different from SDRT's, in that it contains elements that *aren't* linguistically explicit. When interpreting (12e), their linguistic structure includes the current discourse segment purpose (DSP) as a prominent element, which in this case is to find a time to meet which falls on Saturday morning or Sunday. It's therefore in danger of overgenerating the possible interpretations of *3pm* in (12e). In what follows we'll show how SDRT can combine information about goals with the other information sources it uses. But in order to show how to incorporate goals into the SDRT framework, it's helpful to look at how SDRT relates to an older and more established pragmatic theory—the theory of speech acts pioneered by (Searle, 1969).

7.4 SDRT and Speech Acts

There's a natural connection between rhetorical relations and Searle's (1969) notion of speech acts. Speech act theorists typically associate utterances with their communicative intentions (e.g., Poesio and Traum (1995), Litman and Allen (1990) and others). In fact, Searle suggests that the *illocutionary point*, which is roughly the intention that prompted the act, is the basis for distinguishing types of speech acts in the taxonomy. Searle (1969) also argues that the connections between an utterance and its illocutionary force are typically a matter of linguistic convention. For instance, they're encoded within sentence mood: interrogatives (e.g., (13)) express questions; declaratives (e.g, (14)) express assertions; and imperatives (e.g., (15)) express requests.

(13) Is your name Anakin?

(14) Your name is Anakin.

(15) Avoid the dark side of the force!

SDRT demonstrates that in many cases, one can use model-theoretic semantics instead of intentions as the basis for distinguishing one speech act type from another. For example, an assertion is a different kind of speech act from a question, because they denote different kinds of semantic objects: a proposition and a set of propositions respectively. Requests are also different, since they denote actions. Of course, there are *speech act related goals* (SARGs) that are either conventionally associated with these different speech acts or recoverable by the interpreter from the discourse context (Asher and Lascarides, 1998b). These SARGs are also by default an intention of the utterer of the speech act. But such SARGs needn't

be the basis for distinguishing the different acts; rather, SDRT uses the truth conditions to do this (for details see Asher and Lascarides (2001a)).

Searle and more recent AI work (e.g., Litman and Allen (1990)) typically take speech acts to be a property of an individual utterance. In contrast, SDRT shows that many types of speech acts must be understood relationally, because successfully performing the speech act is logically dependent on the content of an antecedent utterance. So technically speaking, the type must be (at least) a two place relation (see also Holdcroft (1984)). For example, if one uses an utterance to conclude something, then that conclusion must be relative to some antecedent hypothesis or argument. And this is relational, because successfully performing this speech act is logically dependent on the content of the antecedent. Answering is also inherently relational: an answer is an answer to some prior question, and successfully performing the act of answering is logically dependent on the content of the question. Similarly, providing concessive information (relative to some conclusion) and contrasting the content of one utterance with a prior one are also relational. Indeed, Searle's (1975) own example (16) is one where there's a need for relational speech acts.

(16) a. A: Let's go to the movies tonight.

b. B: I have to study for an exam.

Understanding that (16b) is a rejection is essential to understanding the content of the dialogue and why it's coherent. But a rejection is better thought of as a *relation* between the utterance and some antecedent proposal (in this case, (16a)), than as a property of the utterance (16b) itself, because successful performance of this speech act is logically dependent on this prior contribution.[5]

The rhetorical relations in SDRT and other theories (e.g., Mann and Thompson (1987), Hobbs (1985), Polanyi (1985)) thus constitute distinct types of (relational) illocutionary force. Explanations, elaborations, giving backgrounds or describing results are all things that speakers *do* with utterances. Moreover, in rhetorical theories of discourse, these illocutionary contributions are all defined via not only an individual utterance, but also an antecedent utterance in the discourse context. For example in (17), the speaker doesn't simply assert that John pushed Max, he *explains* the antecedent assertion that Max fell.

(17) Max fell. John pushed him.

(18) A: John failed his exams.
B: No he didn't, he got 60%.
A: I meant John Smith.

[5]This relational view of speech acts is already implicit in Carberry and Lambert (1999); e.g., they represent the speech act of expressing doubt as a relation between utterances.

Providing this explanation must be an intentional act by the speaker, since otherwise one cannot understand why the speaker juxtaposed the sentences or why the discourse is coherent. This explanation is relational: explanations explain some prior contribution, and the constraints on successfully providing an explanation are dependent on the content of that prior contribution. *Correction*s (e.g., those in (18)) are relational for similar reasons. We've seen many other relational speech acts in the course of this book.

In light of this evidence, it is reasonable to conclude that the typology of rhetorical relations includes the typology of those speech acts whose illocutionary point has to do with connecting information in a discourse together. There may be other speech acts that have received attention in the literature, but these too can be added to SDRT. So the rhetorical relations in SDRT are *speech act types*. For example in (17), the second proposition β stands in an *Explanation* relation to the first proposition α (i.e., *Explanation*(α, β) holds). This means that uttering β amounts to performing the speech act of providing an *Explanation* relative to the discourse context α. This has truth conditional effects: the second clause doesn't merely convey that the event of John pushing Max occurred; rather, since it explains why Max fell, it also conveys the *cause* of Max falling. Thus the semantics of discourse relations in SDRT capture illocutionary forces. The link between these discourse relations and the participants' cognitive states—such as what intentions they have—thus links illocutionary forces to intentions and beliefs. We discuss this link in detail in Chapter 9.

A theory of discourse structure like SDRT offers some advantages over traditional speech act theory. First, it yields a richer typology of speech acts, and hence it's better placed to encode the semantic and cognitive effects of successfully performing a particular speech act. For example, the speech act of asserting in traditional speech act theory gets divided into several different (sub)types of speech act in SDRT: explanations, narrations, backgrounds, elaborations, corrections and so on (Asher and Lascarides, 1998b, 2001a). And each of these distinct kinds of assertion have distinct truth conditional effects, amounting to more fine-grained information about the illocutionary force than that given simply by knowing the utterance was an assertion. Examples like (12) demonstrate that the typology of relational speech acts associated with interrogatives is also richer than the traditional speech act view. The relational speech act *Q-Elab*, for example, connects a question to some antecedent utterance, and can be thought of as a *subtype* of the speech act of questioning.

A second advantage of SDRT is that it offers a specific and formal criterion for distinguishing speech acts in the ontology. SDRT distinguishes a rhetorical relation—or equivalently, a speech act type—in the ontology only if it has empirical consequences within truth conditional semantics: the rhetorical relation must affect some aspect of the context change potential of an utterance within dynamic semantics, where this effect can't be explained straightforwardly by other means.

For example, it might impose constraints on antecedents to anaphora that can't be expressed otherwise. So a *Correction* is distinguished from an *Explanation* (although they're both a subtype of asserting), because $Correction(\alpha, \beta)$ doesn't entail that K_α is true while $Explanation(\alpha, \beta)$ does. SDRT is a theory of how such speech acts make a difference to discourse interpretation.

Finally, SDRT's glue logic provides a more detailed and formally precise theory of alignment between speech acts and linguistic form than is usually given (e.g., Searle (1969), Bach and Harnish (1979), Vanderveken (1990)). The problem is that the link between illocutionary force and linguistic form is often defeasible (Hobbs *et al.*, 1993, Lascarides and Asher, 1993b, Perrault, 1990). For instance, an indicative utterance may not be an assertion. In SDRT, rhetorical relations are calculated within a nonmonotonic glue logic as described in Chapter 5, where exceptions to default rules can be predicted. The glue logic is essentially a formal theory of alignment between linguistic form and speech acts. It contrasts with more traditional theories of alignment, in that inferences about speech acts require information about the *prior* utterance that the current utterance is connected to; knowing that the current utterance is indicative, for example, is insufficient. And unlike plan-recognition approaches, the theory of alignment uses the content of the prior utterance, and not just its goals. This richer theory of alignment is a consequence of the richer typology of speech acts.

SDRT's way of linking illocutionary acts and goals is also distinct from the link that's described in traditional speech act theory. We use an extension of the glue logic to reason in a defeasible way about speech act types and speaker's goals (details will come in Chapter 9). This is part of a more general, linguistically directed model of agent rationality and cooperativity. In particular, if the agent who uttered β, which we'll write as $S(\beta)$, performed the speech act $R(\alpha, \beta)$, then one of $S(\beta)$'s SARGs in uttering β was that the interpreter believe $R(\alpha, \beta)$'s semantic effects. For veridical speech act types, this entails the SARG of belief transfer of K_β and hence Searle's connection between the act of asserting and its goal of belief transfer. But our theory will be more general, in that indicatives won't always lead to asserting and a SARG of belief transfer. For example, for a non-veridical speech act, such as $Def\text{-}Consequence(\alpha, \beta)$, the semantic effects do *not* include that K_β is true, and hence one doesn't infer that $S(\beta)$'s SARG in uttering β is that the interpreter believe K_β, even though β is indicative. Rather, the SARG is that the interpreter believe that *if* K_α is true, then by default so is K_β. We'll work out these ideas in Chapter 9.

7.4.1 Indirect Speech Acts

SDRT brings a host of techniques and insights from formal semantics to pragmatics, and these can help with traditional problems of speech act theory like the phenomenon of indirect speech acts. We've observed already that understanding

the motives behind utterances is often crucial to successful communication (e.g., (12)). But as is well known (e.g., Searle (1975)), the relationship between the surface form of an utterance and its underlying purpose isn't always straightforward:

(19) Can you pass the salt?

Sentence (19) is an interrogative and so expresses a question. Usually, the speaker's goal in asking a question is to get an answer. But (19) plausibly has a different purpose: it's a request, where the speaker's goal is for the interpreter to pass the salt. This is an *indirect speech act* (ISA), which Searle defines to be an utterance in which one speech act is performed indirectly by performing another. With (19), requesting the hearer to pass the salt is performed indirectly by performing another communicative act—asking about the hearer's ability to pass the salt. One challenge then is to make sense of this notion of performing "two speech acts in one." We have offered a detailed analysis of ISAs in Asher and Lascarides (2001a); we give brief highlights of that analysis here.

We provide evidence in Asher and Lascarides (2001a) that some ISAs behave as if they are assigned incompatible semantic types. For instance, observe in (20) how modification with *please* and the direct answer (20b) indicates that (19) is behaving linguistically like it's both a request and a question:

(20) a. A: Can you please pass the salt?

b. B: Yes [(*uttered as B passes the salt*)]

And yet requests and questions are of incompatible semantic type: the former are actions and the latter are sets of propositions. Thus (20) is linguistic evidence that the semantic value of (19) is not ambiguous or indeterminate; but rather it's overdetermined and has *both* illocutionary forces at once.

To characterise the semantic value of an utterance like (19), typing it 'simultaneously' as both a question and a request, we resort to a device that is used within the lexicon to represent an item that has an overdetermined semantic type; namely, dot types (see Chapter 6 for a brief discussion of dot types as they're used in the lexicon). A dot type is one that is made up of two or more constituent types which can be mutually incompatible. Sentence (19) will be assigned a *dot type* of semantic object, with a question and a request as its constituents; i.e., *question•request*. The glue logic has access to the fact that the semantic value of (19) is a dot type, and so it can exploit either its question value or its request value. Given the role of dot types in accounting for co-predication (see Chapter 6 for details), we are thus essentially appealing to an analogy between co-predication in grammar—such as (21) where *book* is simultaneously modified by predicates that select for physical objects and for abstract information content—and the 'simultaneous' use of *please* (which modifies requests) and *yes* (which modifies questions) in discourse.

(21) The book is 500 pages long and describes the theories of Freud.

Asher and Pustejovsky (2001) explore the formal properties of these dot types. Their main property for our purposes is that if a term v is typed as $t_1 \bullet t_2$ and v is predicated over by a predicate P that selects an argument of type t_1 (or t_2), then we may introduce a new term w of type t_1 (or type t_2) that's linked to v, and P predicates over w. The rule that formalises this is called `Dot Exploitation` (see Asher and Lascarides (2001a) for details). Unlike other rules for manipulating semantic types, it is *ampliative* not *destructive*: the complex type remains in the context for further predication.

Analysing the ISA (19) as a dot type explains how the two predications in (20) on the same speech act object can have conflicting type requirements. We assume that the grammar assigns (19) the dot type *question•request*, and thus the rule `Dot Exploitation` will then introduce a label of type *request* that is the argument of *please*, and that is linked with *O-Elab* (standing for Dot Elaboration) to the original speech act of complex type. Another application of `Dot Exploitation` will generate the label of type *question* to which B's verbal response (20b) is connected, and this is also linked with *O-Elab* to the original speech act. So for us, the hallmark of conventionalised ISAs is that the grammar assigns to them a dot type involving incompatible constituent types, which allow us to explain their 'dual' linguistic behaviours. This is just like the role of dot types in the analysis of co-predication given in Chapter 6.

While in principle any combination of independent types can form a dot type, natural language does not seem to function that way. There always seems to be some natural relationship between the constituent types. Consider the book: its physical and abstract perspectives are naturally related in that the physical is a realisation of the abstract information content—a relation familiar to metaphysics since the time of Plato. With regard to ISAs, the natural connection between the constituent types in a dot type comes from the Gricean reasoning that Searle remarked on. Reasoning about the speaker's goals from assumed mutual beliefs yields the coherent link between the question and the request in (19), for example (see Gordon and Lakoff (1975) for details).

Gricean reasoning can also make a question into an answer to another question (e.g., (22b)) and a question into a negative commentary or *Correction* of some prior assertion (e.g., (23b)):

(22) a. A: Do you want another piece of chocolate cake?
b. B: Is the pope Catholic?

(23) a. A: Reagan was the best US president of the 20th century.
b. B (to C): Is he really such an idiot as to believe that?
c. C: Yes, he is.
c′. A: Well, maybe you're right. Maybe Reagan was mediocre.

C's response (23c) to B's utterance (23b) indicates that (23b) still functions as a question, in that C can provide a direct answer to it. The alternative 'continuation' of the dialogue (23ab) given in (23c′) indicates that (23b) also functions as a commentary by B on A's view. In this sense, the utterance (23b) contributes to the meaning of the dialogue as both a question and as a correction, even though these types of speech act are incompatible (since corrections are a kind of assertion, which are of an incompatible semantic type to questions).

The hallmark of all ISAs, then, seems to be that Gricean reasoning provides a connection between one kind of speech act and an incompatible kind of speech act. For conventionalised ISAs like (19), a dot type consisting of incompatible speech act types is assigned by the grammar. For what we'll call *unconventionalised* ISAs (e.g., (23b)) the grammar doesn't assign a dot type; rather, Gricean reasoning connects the speech act that's assigned by the grammar (e.g., *question* in the case of (23b)) to the 'implicit' incompatible type of speech act (*Correction*), which is 'accommodated' into the interpretation because of demands on content that are imposed by factors such as discourse coherence and speaker rationality.[6] In this respect, (16b) appears different: Gricean reasoning connects the act of assertion to the act of rejection, but there isn't any *semantic* incompatibility between an assertion and a rejection (because they're both propositions) in the way that there is between questions and assertions, for example. So (16b) doesn't involve two incompatible types of speech acts.

This discussion thus yields an alternative definition of ISAs to those given by speech act theory (e.g. Searle, Bach and Harnish, and the like). An utterance is a conventionalised ISA if (a) the grammar assigns it a complex speech act type of the form $s_1 \bullet s_2$, such that s_1 and s_2 are distinct (incompatible) types of semantic objects; and (b) Gricean-style principles of rationality and cooperativity link the constituent type s_1 to the type s_2. This Gricean link means there's a sense in which the complex type $s_1 \bullet s_2$ is asymmetrical, characterised by the Gricean information flow from s_1 to s_2.[7] For example, (19) is a conventionalised ISA because (a) there's linguistic evidence (e.g., (20)) that the grammar assigns (19) the complex dot type *question•request*, and (b) as Gordon and Lakoff (1975) show, Gricean-style reasoning links the question to the request. Similarly, (22b) is also a conventionalised ISA, with dot type *question* • *answer*.

An utterance is an unconventionalised ISA, if similar Gricean style reasoning leads to the inference of an implicit speech act (Asher and Lascarides (2001a) make precise this notion of *implicit*, but for now it can be taken to mean that this speech act isn't derivable from the theory of alignment), and this implicit

[6]Note that utterance (23b) doesn't behave *linguistically* like a correction, since it cannot be felicitously preceded by *no*, in the way that corrections can; compare *??No, is he such an idiot as to believe that?* with *No, Reagan was mediocre.*

[7]Copestake and Briscoe (1995) show, via data involving quantification, that the complex types assigned to lexical entries like *book* are also asymmetrical.

type is semantically incompatible with that inferred for the utterance itself (by the theory of alignment). So (23b) is an unconventionalised ISA, and thanks to pragmatics, it is assigned the complex type *Q-Elab*•*Correction* (*Q-Elab* because answers to (23b) will help elaborate a plan to achieve (23a)'s SARG of belief transfer or demonstrate the SARG isn't achievable), even though the compositional semantics of (23b) is of the simple semantic type *question*. The assertion (16b) is not an ISA, however, because there is no incompatibility between an assertion and a rejection at the level of semantic value (they're both propositions). This analysis of ISAs makes unconventionalised ISAs a special case of a general pragmatic process during interpretation. Generally, interpreters accommodate content in order to preserve assumptions that the speaker was rational and that the discourse is semantically well-formed (Lewis, 1969). For example in Chapter 4, we modelled many discourses where temporal information is accommodated. An unconventionalised ISA arises when what's accommodated into the semantic representation of the discourse is a speech act of incompatible type to the one that's predicted by the grammar.

In appropriate contexts, Grice's example (24) is interpreted as a request.

(24) I'm out of gas.

While (24) is not grammatically functioning as a request as well as an assertion (e.g., it can't be modified with *please*), its discourse function connects it to an implicit request: the proposition expressed by the utterance in (24) explains *why* the speaker needs help and what sort of help he needs. An interpreter needs to recognise this in order to respond in a competent way (i.e., help the speaker obtain gas). And in order to recognise that (24) explains the request for help, he must reason about the agent's cognitive state (Asher, 1999). This Gricean reasoning therefore links two incompatible types just as it does for (19), except the grammar doesn't assign (24) the dot object. So we would classify (24) as an unconventionalised ISA. We give a more detailed analysis of (24) in Chapter 9.

7.5 Technical Details of SDRT for Dialogue

We've just described how SDRT can extend traditional approaches to dialogue interpretation such as speech act theory. We'll now provide the formal developments that underlie this.

Extending SDRT to dialogue requires an analysis of imperatives and interrogatives. In Chapter 2 we gave a compositional dynamic semantics of them. We investigate now how context supplements this compositional semantics with further content. As before, we exploit the semantics of rhetorical connections, but this time the connections take labels of interrogatives and imperatives as arguments.

In order to express the semantics of these relations, we must extend the underlying language (and logic) in a couple of ways. First, we need to keep track of who said what; this is needed for those rhetorical relations which constrain the beliefs and goals of the dialogue agents. The definition of the language of information content in Chapter 4 abstracted away from speakers. Secondly, we'll need to express information about the dialogue agents' beliefs and intentions. Thus we'll introduce propositional terms into the language since they are the objects of belief.

To represent who said what in a dialogue, we need to rework slightly the definition of a discourse structure. Recall from Definition 11 in Chapter 4 (page 138) that an SDRS is a set of discourse labels A and a function $\mathcal{F}$ which assigns SDRS-formulae to each label π in A. For dialogue, we will assume that in addition to this, an SDRS also consists of a function S from labels to dialogue participants. $S(\pi)$ is the agent/speaker of the speech act marked by the label π, and it is he who conveys the formulae that are, via $\mathcal{F}$, associated with π. As before, we use K_π to represent the formula $\mathcal{F}(\pi)$ that's labelled by π.

To represent the beliefs and intentions of discourse agents, we need to extend the language of Chapter 4 to include a semantics of these attitudes. This will make the model theory for discourse structures a dynamic version of the model theory for higher order modal logic (Gallin, 1975). But we will not investigate this in much detail here, as we're mainly concerned with particular semantic effects of discourse structure.

To express beliefs, we'll use a 2-place predicate $\mathcal{B}$ that takes an individual and a propositional term as its arguments. We turn formulae into propositional terms via the use of $^\wedge$; so an (S)DRS-formula K_α becomes the propositional term $^\wedge K_\alpha$. We will also express propositional terms as $p, p', p_1, p_2, \ldots$; and we'll express the extensions of a propositional term p with the formula $^\vee p$. So if ϕ is a formula, then $^\wedge\phi$ is a propositional term, $^\vee{}^\wedge\phi$ is a formula, and $\mathcal{B}_A(^\wedge\phi)$ means A believes that ϕ.

Now let's define the semantics of these new expressions. Recall from Chapters 2 and 4 that if ϕ is a formula, then $[\![\phi]\!]_M$ is a relation between world-assignment pairs; or equivalently a function from world-assignment pairs to world-assignment pairs. Thus propositional terms in dynamic semantics must denote a set of pairs of world-assignment pairs (where the worlds in a pair are the same). Now, $[\![.]\!]_M^{w,f}$ is by definition the range of $[\![.]\!]_M$ at (w, f). So given the denotations of propositional terms, $[\![^\wedge\phi]\!]_M^{w,f}$ must be a set of pairs of world assignment pairs. With this in mind, the semantics of the intension and extension operators are given below; note that these are simply the dynamic-semantic correlates of their standard Montagovian interpretations:

Definition 24 **The Truth Conditions of $^\wedge$ and $^\vee$**

- Where ϕ is a formula

$[\![^{\wedge}\phi]\!]^{w,f}_M = \{\langle (w,g),(w',h)\rangle : w = w', g \supseteq f, \text{ and } (w,g)[\![\phi]\!]_M(w',h)\}$
(i.e., the intension of ϕ is the proposition it denotes).

- Where p is a propositional term
  $(w,f)[\![^{\vee}p]\!]_M(w',g)$ iff $\langle (w,f),(w',g)\rangle \in [\![p]\!]^{w,f}_M$.

The above definitions ensure that $^{\vee\wedge}\phi$ is logically equivalent to ϕ.

The language also includes predicates on labels that tell us whether the associated formula expresses a question, a proposition or a request: these are given by α:?, α: | and α:! respectively. For us, the link between formula type and semantic entity expressed is just a matter of compositional semantics. On our modular approach, an interrogative formula *always* expresses a question. This contrasts with speech act theory and other pragmatic approaches (e.g. Sperber and Wilson (1986)) where questions and requests are defined with respect to the intentions and goals of the dialogue participants. On these other views, the relation between questions and interrogative formulas becomes very complex and often murky; one and the same formula will not always have the same type of semantic value, for such theories make the semantic value of an expression's use depend on the cognitive state of its utterer. As many have argued (e.g. Groenendijk and Stokhof (1984)) these approaches give the wrong semantics to interrogative formulas. For instance, it's not clear what rhetorical questions express, nor what embedded interrogatives express (e.g., *John knew who came to the party*). Given these difficulties and the availability of several elegant semantic theories of questions (Groenendijk and Stokhof, 1984, Ginzburg, 1995a,b, Ginzburg and Sag, 2001) of which we avail ourselves below (and in Chapter 2), we will stick with the view that interrogative formulas always express questions, propositional formulas always express propositions and imperative formulas always express requests.[8]

7.6 Simple Relations for Dialogue

7.6.1 Indirect Question Answer Pairs

A response to a question needn't correspond to a direct answer, according to the compositional semantics of the question (Groenendijk and Stokhof, 1984). For example, (25b) doesn't match the compositional semantics of answers to the yes/no-question (25a):

(25) a. A: Did John fail his exams?

b. B: He got 60%.

[8]The matter of indirect speech acts, when for example an interrogative also expresses a request (e.g., *Can you pass the salt?*), is discussed in detail in Asher and Lascarides (2001a) and in Section 7.4.1.

However, assuming that A knows the pass mark, A can use the information in (25b) to compute an answer: If the pass mark is above 60%, the answer is *yes*; if not, it's *no*.

SDRT represents this rhetorical move with the relation *Indirect Question Answer Pair*, or *IQAP*. Informally, $IQAP(\alpha, \beta)$ holds only if K_α is a question and the questioner $S(\alpha)$ believes that an answer p to his question normally follows from K_β. In other words, although the response itself wasn't a direct answer, the questioner can infer one from it. Note that this 'paraphrase' for $IQAP(\alpha, \beta)$ introduces a condition on the questioner's *beliefs*. The formal truth definition of $IQAP(\alpha, \beta)$ will reflect this. And as such, *IQAP* is a cognitive-level rhetorical relation. We'll show in Chapter 9 that nevertheless, *IQAP* can be inferred during dialogue interpretation *without* using beliefs and intentions as clues.

Note furthermore that according to this informal semantics, indirect answers subsume direct answers (for if K_β is a direct answer, then the questioner can compute a direct answer from it, namely K_β). We mark the relationship between a true *direct* answer and its question with *QAP*, standing for Question Answer Pair. And the formal semantics of *QAP* and *IQAP* will ensure that $QAP(\alpha, \beta) \Rightarrow IQAP(\alpha, \beta)$ is valid.

Before we define the full truth definitions of *IQAP* and *QAP*, let's examine a few discourses which feature these relations so as to observe what these discourses entail. Dialogue (26) is an example where $QAP(\alpha, \beta)$ holds and hence so does *IQAP*:

(26) a. A: Who came to the party?

b. B: John and Mary.

Intuitively, B's successful performance of the speech act of *QAP* only occurs if his answer is true; for otherwise he hasn't answered the question. The semantics of *QAP* must therefore entail that B's contribution is true; that is, $QAP(\alpha, \beta)$ holds only if K_β does too. We call such relations *right-veridical* relations.

Direct answers to *how*-questions are adverbials of manner (Asher and Lascarides, 1998b). Hence the exchange in (27) is not *QAP*:

(27) a. A: How can I get to 6th Street?

b. B: It's just a couple of stops on the LX bus, which stops right around the corner.

c. A: OK, thanks.

However, proviso certain assumptions, A can compute a direct answer from the content of (27b): I can get to 6th Street by going to the bus stop around the corner, catching the LX bus, and getting off (a couple of stops later) at 6th Street. So we would like to ensure that SDRT's glue logic allows us to infer *IQAP* for this example; similarly for (1ab) and (1ab′). Furthermore, *IQAP* is like *QAP* in that

it should clearly be a right-veridical relation. Discourse (27) also features the rhetorical relation *Acknowledgement*, which connects (27c) to (27b). We'll talk about *Acknowledgement* in more detail when we address the issue of common ground in the next chapter.

Since $IQAP$ and QAP are to be right-veridical relations, we must ensure that their truth definitions validate the following:

- `Right-Veridicality of QAP and IQAP`
  $QAP(\alpha, \beta) \Rightarrow K_\beta$
  $IQAP(\alpha, \beta) \Rightarrow K_\beta$

But once again it's important to stress that since the glue logic doesn't have full access to cognitive states, an agent can infer $IQAP(\alpha, \beta)$ in the glue logic even when he believes that K_β is false. So, even if A already believes that John got less than 60% when he hears (25), he will still infer $IQAP$ in the glue logic. And so his interpretation of the discourse will match B's; the difference lies in their beliefs about whether B's response is true. A may detect this difference in the logic of information content by detecting in this logic that his model doesn't satisfy the SDRS. He may then use his next turn to indicate the dispute to B. We'll discuss such corrective moves in the next chapter.

As with all rhetorical relations in SDRT (see Chapter 4), $QAP(\alpha, \beta)$ and $IQAP(\alpha, \beta)$ denote transitions on information states within the model theory. But since their first arguments are questions, this makes a difference to the anaphoric potential of their contributions. Following the work of Groenendijk and Stokhof (1984), the dynamic semantics of questions amounts to a set of dynamic propositions (recall a dynamic proposition is a set of pairs of world assignment pairs). Consequently, a unique semantic value for an indefinite NP that's introduced in a question is not assured. Consider, for example, the *yes/no*-question (28a) when its true direct answer is *no*, thereby amounting to *It's not the case that a man came by this morning*:

(28) a. A: Did a man come by this morning?

b. B: No. ??He was wearing a blue suit.

c. B: He was wearing a blue suit.

In this case, the denotation of the question is expressed by DRSs where the negation outscopes the discourse referent introduced by the indefinite; or to put it another way, the output world-assignment pairs which denote (28a) don't define unique values for the indefinite. This correctly blocks anaphoric binding of pronouns in subsequent discourse, as illustrated in (28b).

On the other hand, the alternative response (28c) to the question implicates a *positive* answer: (28c) attaches to (28a) with $IQAP$, and being a positive answer its content is equivalent to $[\![\phi]\!] \circ [\![\psi]\!]$, where ϕ can be paraphrased as *A man came*

in this morning (i.e., (28a)'s true direct positive answer) and ψ is the compositional semantics of (28c). This makes the discourse referent introduced by the indefinite (in ϕ) available as an antecedent to the pronoun. Indeed, the principle of maximising discourse coherence (MDC) that we discussed in Chapter 5 explains why (28c) is interpreted as implicating a positive answer rather than a negative one: all else being equal, MDC forces interpretations where semantic underspecification is minimised, and in this case we can resolve the (underspecified) meaning of the pronoun only by interpreting it as a positive answer to the question.

Thus in general, a question's semantic values do not pass on bindings of variables to new constituents, although elements of those semantic values can. The semantics of $QAP(\alpha, \beta)$ and $IQAP(\alpha, \beta)$ must make only the discourse referents introduced in K_β available for future anaphoric reference, but not those that are introduced in the question K_α (note that for the dialogue (28ac), the value of K_β includes the content of the positive answer, as mentioned earlier). The question K_α simply serves to constrain the conditions under which the relation $IQAP(\alpha, \beta)$ itself holds.

Having made these observations about anaphora, we're now in a position to formalise the semantics of these relations. To make the definitions easier to read, we use a predicate *Answer* to relate a question to one of its true direct answers. *Answer* takes the intension of a question (which is a set of propositions) and a propositional term as its arguments. It's the counterpart to the relation *QAP*, which takes *labels* as arguments. So semantically, $Answer(^\wedge K_\alpha, p)$ is true just in case p is an answer to the question K_α:

$$(w,f)[\![Answer(^\wedge K_\alpha, p)]\!]_M(w',g) \text{ iff } (w,f) = (w',g) \text{ and } [\![p]\!]_M^{w,f} \in [\![^\wedge K_\alpha]\!]_M^{w,f}$$

The semantics of $QAP(\alpha, \beta)$ given below thus ensures that this relation holds only if (a) K_β holds; and (b) K_β is answer to the question K_α:

- **Semantics for QAP**
  $(w,f)[\![QAP(\alpha,\beta)]\!]_M(w',g)$ iff
  $w = w'$, $(w,f)[\![K_\beta]\!]_M(w',g)$ and $(w,f)[\![Answer(^\wedge K_\alpha, ^\wedge K_\beta)]\!]_M(w,f)$

Notice that *QAP* doesn't shift the world index and in this it's like the relations of Chapter 4.[9] We give examples of relations that shift the world index in Section 7.6.6. The relation *Correction* whose semantics we discuss in Chapter 8 will also shift the world index.

The semantics for *IQAP* is a little more complex because it constrains the questioner's beliefs. The semantics for *IQAP* below says that the relation holds of α and β iff (a) K_β is true; and (b) there is a proposition p that (by (bi)) is an answer to the question, it's true (see (bii)) and the questioner believes that

[9] In fact, the conjunct $w = w'$ in this definition is redundant: since K_β is a proposition $(w,f)[\![K_\beta]\!]_M(w',g)$ will hold only if $w = w'$.

p normally follows from K_β (see (biii)). Note that $^\vee p$ rather than p is in the consequence of the $>$-statement in (biii) for otherwise the result is not well-typed ($>$ connects formulae, and not terms).

- `Semantics for IQAP`
  $(w,f)[\![IQAP(\alpha,\beta)]\!]_M(w',g)$ iff
  (a) $w = w'$, $(w,f)[\![K_\beta]\!]_M(w',g)$, and
  (b) there is a p such that:
    (i) $(w,f)[\![Answer(^\wedge K_\alpha, p)]\!](w,f)$,
    (ii) $(w,f)[\![^\vee p]\!]_M(w',g)$ and
    (iii) $(w,f)[\![\mathcal{B}_{S(\alpha)}(K_\beta > {}^\vee p))]\!]_M(w,f)$

This semantics ensures that the following axiom is valid:

- `Axiom on IQAP`:
  $IQAP(\alpha,\beta) \Rightarrow \exists p(Answer(^\wedge K_\alpha, p) \wedge \mathcal{B}_{S(\alpha)}(K_\beta > {}^\vee p))$

And thus assuming that belief is closed under validity (i.e., $\mathcal{B}_x\phi$ is valid for all valid formulae ϕ), it follows that $QAP(\alpha,\beta) \Rightarrow IQAP(\alpha,\beta)$ is valid, as required. For if $QAP(\alpha,\beta)$ holds then one can instantiate the p in `Semantics for IQAP` with $^\wedge K_\beta$ (for note that $K_\beta > K_\beta$ is valid, and so $\mathcal{B}_{S(\alpha)}(K_\beta > K_\beta)$ is valid).

Furthermore, if belief distributes over $>$ (i.e., $\mathcal{B}_x(\phi > \psi)$) entails $\mathcal{B}_x\phi > \mathcal{B}_x\psi$, then the semantics of $IQAP$ will entail that one of its illocutionary effects is that the questioner believes a direct answer to his question. We'll make use of this when we reason about how to infer $IQAP$ in the glue logic in Chapter 9.

The second argument β to $IQAP$ must label a proposition K_β: the semantics of $IQAP(\alpha,\beta)$ places K_β in the antecedent of a $>$-statement and so if K_β didn't denote a proposition the result wouldn't be well-typed. And yet speakers can use imperatives as answers to questions, particularly *how*-questions, as in (29):

(29) a. A: How does one make onion soup?

b. B: Chop onions and saute them in olive oil or butter until soft, add stock until the onions are well covered, and simmer for two hours at low heat.

The *content* of the (complex) imperative (29b) is sufficient for A to compute a direct answer to his question, which by the compositional semantics of *how*-questions must be an adverbial of manner: one makes onion soup by chopping onions, and then frying the onions in oil, and so on. Note that A doesn't need to *perform* the actions denoted by the imperative to compute this answer; interpreting the content of the imperatives is sufficient. Observe also that in contrast to the 'null' context, the imperative (29b) isn't commanded in the context of (29a). We must capture this in the semantics of the rhetorical relation which connects (29b) to (29a). In particular, this relation should be non-veridical.

Now, we could exploit the analysis of indirect speech acts to represent dialogues like (29), marking (29b) as having the complex type *request*•*IQAP*, where the *IQAP*-part of this complex type would label the propositional component K'_β of the action $\delta K'_\beta$ that represents the content of the imperative (29b). However, it's not clear that Gricean reasoning—i.e., reasoning about beliefs and intentions—connects the request to the speech act type *IQAP*. Rather, the content of the imperative itself seems to be sufficient for inferring its rhetorical role as providing an answer to the question. And yet we argued earlier that Gricean reasoning is a hallmark of indirect speech acts. Further, although (29b) is in imperative form, in this context it's not commanded. So we would assign (29b) incorrect truth conditions if we made it a complex type of speech act whose semantics will make its request component commanded.

In view of this, we don't treat (29b) as an indirect speech act but rather as introducing a new rhetorical relation called $IQAP_r$, whose semantics is very closely related to *IQAP*. Like *IQAP*, the relation $IQAP_r$ requires its first argument to label a question (see clause (b) of the semantics of $IQAP_r$ given below). But unlike *IQAP*, its second argument must label an imperative (see clause (a) below). Like *IQAP*, this imperative has content from which the questioner can compute a true direct answer (see clause (b) below). But unlike *IQAP*, $IQAP_r$ is not veridical; this reflects the fact that the rhetorical role of the imperative is not that it's commanded, but rather that its content convey an answer to the question. All these aspects of the semantics of $IQAP_r$ are captured as follows:

- **Axioms on IQAP$_r$**
  (a) $IQAP_r(\alpha, \beta) \Rightarrow \beta :!$
  (b) $(IQAP_r(\alpha, \beta) \wedge \beta : \delta K'_\beta) \Rightarrow \exists p(Answer(^\wedge K_\alpha, p) \wedge \mathcal{B}_{S(\alpha)}(K'_\beta >^\vee p))$

Note how neither clause (a) nor clause (b) imply the action $\delta K'_\beta$. Hence the transitions from input to output world-assignment pairs that are generated by $IQAP_r(\alpha, \beta)$ don't have the characteristic of a commanded imperative as described in Chapter 2; i.e., the world index isn't updated to one where the action $\delta K'_\beta$ has been performed.[10]

In fact, the command status of an imperative response to a question appears to depend on the semantics of the question. If answers to the question implicate that the questioner is the agent of a deontic attitude, then an imperative response is commanded:

(30) a. A: Where should I go now?
b. B: Turn left at the next set of traffic lights.

[10]Note that connecting the imperatives in (29b) with *Narration* has the desired temporal effects on the content of the imperative, even though by the non-veridicality of $IQAP_r$ this complex imperative isn't commanded.

We could formulate further axioms for $IQAP_r$ that reflect these effects on the command status of the request; see Lascarides (2001) for details.

Other kinds of responses to questions are possible. For example, a response may rule out some true answers, but it's not an *IQAP* because it's not sufficiently informative that the questioner can compute a direct answer from it; e.g., the response in (31) is like this, in a context where A's beliefs about the party are insufficient to infer who did come from knowing that Mary didn't:

(31) a. A: Who came to the party?

b. B: Well, not Mary.

We mark this relationship as a Partial Question Answer Pair (*PQAP*), which again is right-veridical. On the other hand, B's response to A's question may be designed to indicate that B doesn't have sufficient information to answer the question:

(32) a. A: Who came to the party?

b. B: I don't know.

b′. B: I would need to look in the visitor's book to find out.

(32b) doesn't provide information that helps us to eliminate any of the possible answers to (32a), and indeed it explains why this information isn't given. We mark the fact that the proposition (32b) fails to provide information leading to an answer with *NEI*(32b,32a), where *NEI* stands for Not Enough Information. The response (32b′) implicates that whoever signed the visitor's book came to the party. But it doesn't implicate who (in particular) signed the visitor's book, or indeed that anyone signed the visitor's book; thus it doesn't eliminate any of the possible answers to (32a) and it attaches with *NEI* to (32a).

NEI may strike one as a rather odd discourse relation, since its semantics appears to point to a *lack* of a connection between the question and the response. But as we'll see in Chapter 9, it's important that cooperative speakers signal when they can't give answers to questions posed by other dialogue participants; *NEI* codifies this sort of signalling. The formal rendition of the semantics of *NEI* is given below. In words, it says that $NEI(\alpha, \beta)$ holds just in case the response K_β to the question K_α is true and implicates that the respondent doesn't know an answer to K_α:

- `Semantics for NEI:`
  $(w, f)[\![NEI(\alpha, \beta)]\!]_M(w', g)$ iff
  $w = w'$,
  $(w, f)[\![K_\beta > \neg\exists p(\mathcal{B}_{S(\beta)}(p) \wedge Answer(\hat{}K_\alpha, p))]\!]_M(w, f)$ and
  $(w, f)[\![K_\beta]\!]_M(w', g)$

So a response like *I don't know* connects with *NEI* to the question, since it implicates (in fact, it entails) that the speaker doesn't know an answer to the question (at least, it entails this once the antecedent to the anaphoric argument to *know* is resolved to the interrogative and hence to its answers as this is what interrogatives denote). Note that *PQAP* and *NEI* have compatible semantics: one can with a single utterance both supply information that rules out certain propositions from being answers to questions, and also implicate that a true answer isn't known. In fact, using maxims of Quantity one might argue that (31ab) is like this, and hence (31b) attaches to (31a) with *NEI*, as well as the relation *PQAP* mentioned earlier.

7.6.2 Question Elaboration and Other Relations Involving Plans

The semantics of some rhetorical relations are defined in terms of what we will call *speech act related goals* or SARGs: A SARG is a goal that is either conventionally associated with a particular type of utterance, or is recoverable by the interpreter from the discourse context (Asher and Lascarides, 1998b); this reflects the intuition that people say things for a particular purpose while distinguishing the goals that interact with linguistic knowledge from goals in general.

For example, following Searle (1969) one can associate (by default) with the speaker of an interrogative the SARG that he know an answer to his question (we'll specify this formally in Chapter 9, when we introduce the logic of cognitive modelling). The SARG of an imperative is that the action it conveys be performed. However, although SARGs form a part of SDRT's model of the interactions between intentions and dialogue content, SDRT departs from purely pragmatic theories like Searle's (1969) and Sperber and Wilson's (1986), in that the link between the formula type (e.g., interrogative) and its semantic value (or in Searle's terms, illocutionary force; e.g., question) is a matter for *compositional semantics*, rather than the underlying communicative intention or SARG. Unlike these pragmatic theories, SDRT separates the semantic value of an utterance from its underlying intention/SARG, and thus obtains a uniform semantics of indicatives, interrogatives and imperatives (see Chapter 2 for motivation of this view).

SDRT captures the influence of SARGs on content through its inventory of rhetorical relations. For example, SARGs enter into the semantics of the rhetorical relation *Q-Elab* (Question Elaboration) that we mentioned earlier: *Q-Elab*(α, β) holds if β is a question whose possible answers all specify an essential part of a plan to bring about $S(\alpha)$'s SARG in uttering α. *Q-Elab* is a subordinating relation, to reflect the fact that subsequent utterances to β may, like β, help specify a plan for achieving the SARG underlying the speech act α. For example, consider (33):

(33) a. A: Can we meet next week?
b. B: How about Tuesday?
c. A: How about 3pm?

According to the above claim about the SARGs of interrogatives, A's SARG in uttering (33a) is that A know whether B and A can meet next week. Now, all answers to (33b) provide information which elaborates a plan for working out an answer to (33a). If the answer to (33b) is *good*, then A and B can meet on Tuesday; in fact, this holds because (33b) implicates, in this context, that B can meet A on Tuesday (we'll discuss this implicature later). If the answer is *bad*, then at least A and B have limited the set of candidate times for meeting next week. Either way, they know more than they did in a way that will help them implement a plan for achieving the goal that A know an answer to (33a). The relation between (33c) and (33a) is similar to that between (33b) and (33a): i.e., all possible answers to (33c) provide information that elaborates a plan for achieving A's SARG underlying (33a). In fact, *Q-Elab* connects (33c) to (33b) for similar reasons (and this connection makes *3pm* resolve to 3pm on Tuesday as this is required by the semantics of *Q-Elab*). This example shows that it's useful to ensure that (33a) is available to (33c), to reflect its status as a 'follow-up' question. So we need *Q-Elab* to be a *subordinating* relation.

The relation *Q-Elab* can also hold between a request and a question, as illustrated in dialogue (34), where answers to (34b) all elaborate a plan which fulfils the SARGs that lay behind the request (which, as we will specify formally in Chapter 9, is that A and B meet next week):

(34) a. A: Let's meet next week.
b. B: How about Tuesday?

Let's specify the semantics of the relation *Q-Elab* a little more precisely. Obviously, *Q-Elab* isn't right veridical, since its right argument is a question and hence it lacks a truth value (see Chapter 2).[11] But like veridical relations, *Q-Elab* constrains the semantic value of the question in a way that's dependent on the semantics of the first argument of the relation. This helps disambiguate anaphoric expressions. For example, assuming that (35) is uttered on 1st April, there are two linguistically accessible antecedents to which the temporal anaphoric expression *the 15th* can resolve: *now* (which resolves *the 15th* to the 15th of April); and *July* from (35a) (which resolves *the 15th* to the 15th of July):

(35) a. A: Shall we take a day off in July?
b. B: How about the 15th?

[11] *Q-Elab*(α, β) is left veridical if α is indicative or imperative, but not if it's an interrogative. We gloss over this for now though.

However, the fact that (35b) is intended to connect to (35a) with *Q-Elab* (and how A infers this in the glue logic is addressed in Chapter 9) means that *the 15th* must resolve to 15th of July, for otherwise answers to (35b) won't help A devise a plan to know an answer to (35a). Similarly, in (33) resolving *Tuesday* to the Tuesday of next week is not just the only accessible choice, it's also the choice that's demanded by the semantics of *Q-Elab*. We must specify the meaning postulates for *Q-Elab* so that these effects on content are recorded.

Even a partial formalisation of the semantics of *Q-Elab*(α, β) via a meaning postulate is tricky. Questions that connect to the context with *Q-Elab* are part of a dialogue strategy to achieve prior goals that were part of that context: if *Q-Elab*(α, β) holds, then any direct answer to K_β should yield inferences about the plans that $S(\alpha)$ can execute (in the actual world) for achieving a SARG of α. The executable plans should all make true at some point the answer to the question, and it should not be possible to "splice out" the bits of those plans that make the answer true, and still get a plan that is executable and leads to the SARG of α (thus the answer was 'essential' information).

To formalise this, we exploit the representation of actions and plans that we used in the compositional semantics of imperatives in Chapter 2. To this language of dynamic action terms, we'll add an operator *Done* which maps plans into propositions (following Asher and Koons (1993)). Where a is a plan (i.e., an action term), *Done*(a) tells us if a has been accomplished: $(w, f)[\![Done(a)]\!](w', g)$ is true iff $w = w'$ and there is a world w'' such that $(w'', f)[\![a]\!](w', g)$ (note that $[a]Done(a)$ is valid). Thus $Done(a) > Done(a')$ means that executing the plan a normally results in accomplishing the action a' too. And $Done(a) \Rightarrow Done(a')$ means that the action a' is part of or a 'subaction' of the plan a. Although a SARG is a goal and can therefore be thought of as an action, we make the arguments to SARGs propositional terms, because this makes some of the proofs of theorems about cognitive modelling in Chapter 9 easier to follow. Note that we can represent SARGs this way because we can always construct an action from a proposition via δ, and a proposition from an action via *Done*.

The semantics of *Q-Elab* implies that a plan to achieve some goal is derived from a particular bit of information (the answer), together with relevant background assumptions (which include world knowledge). In fact, we will formulate a relationship between all of the following:

1. a SARG p (so p is a propositional term) that underlies a speech act uttered by dialogue participant A;
2. (new) information p';
3. the mutual beliefs $KB_{\tau,A,B}$ of dialogue participants A and B at the point where they have both interpreted the dialogue context τ;

4. the (private) beliefs $KB_{\tau,A}$ of A at the point where he has interpreted the dialogue context τ; and

5. a plan a.

The relation is this: suppose that (a) A augmenting his knowledge base $KB_{\tau,A}$ with the information p' leads to a situation where he can infer that whenever the plan a is accomplished it normally leads to the SARG p being accomplished too (i.e., $Done(a) > {}^{\vee}p$); and moreover (b) this plan a cannot be inferred from $KB_{\tau,A,B}$ (i.e., from what A and B mutually believe, having just interpreted the dialogue context τ). So in a sense, the information p' was essential in helping A generate this plan a for making ${}^{\vee}p$ true. Then we say that $(p, KB_{\tau,A}, p', KB_{\tau,A,B}) >> a$ holds. More formally, the definition of this relationship is given as follows:

- `Definition of` $(p, KB_{\tau,A}, p', KB_{\tau,A,B}) >> a$:
  $(p, KB_{\tau,A}, p', KB_{\tau,A,B}) >> a$ iff
  (a) $((KB_{\tau,A} \wedge {}^{\vee}p') > (Done(a) > {}^{\vee}p)) \wedge$
  (b) $\forall a' \neg (KB_{\tau,A,B} > (Done(a') > {}^{\vee}p))$.

Note that the 'background knowledge' $KB_{\tau,A,B}$ is parameterised with respect to τ. We have to do this, because beliefs evolve as the dialogue proceeds. Moreover, there is an asymmetry in the use of A's and B's beliefs in the definition: the plan a must be derivable from A's knowledge base augmented with p', but not necessarily from B's. There's this asymmetry because it will ultimately be A who utters the speech act with SARG p, and thus it's essential for A to achieve a cognitive state where he can accomplish p. But it's not necessary for B to do this; he must simply ensure that A achieves this state.[12]

This relation helps to specify the constraint we want on the semantics of $Q\text{-}Elab(\alpha, \beta)$:

- `Axiom on Q-Elab`:
  $(Q\text{-}Elab(\alpha, \beta) \wedge \text{SARG}(\alpha, p) \wedge Answer({}^{\wedge}K_{\beta}, p')) \Rightarrow$
  (a) $\exists a((p, KB_{\tau,S(\alpha)}, p', KB_{\tau,S(\alpha),S(\beta)}) >> a$
  (b) $\wedge Executable(a) \wedge \neg({}^{\vee}p > Done(a)))$

In words, the `Axiom on Q-Elab` says the following. If $Q\text{-}Elab(\alpha, \beta)$ holds between an utterance α that A said (where A's SARG is p), and the question β that B said, then any answer p' to β should specify sufficient information such that: (a) adding it to A's beliefs allows him to identify a plan a which, when executed,

[12] Stone (2001) observes this asymmetry for requests too. When a system commands the user *Complete the form by typing your* ID *number*, the content of this request may be insufficient for the system to infer an appropriate executable action (e.g, if the system doesn't know the user's ID number). However, for this request to be felicitous, it's necessary for the user to able to use the content of the request to infer an action that he can execute.

typically achieves the SARG p, and the dialogue agents didn't know a plan for achieving p before; and (b) this plan a is executable and more detailed than p itself.

This is only a partial description of the semantics of *Q-Elab*, but it suffices for our purposes. To see how this axiom helps us specify the meaning of utterances in context, consider the following dialogue.

(36) a. A: Let's meet in two weeks.

b. B: Are you free on the seventeenth?

c. A: I'm afraid I'm busy then.

Let's suppose that A's SARG in uttering (36a) is to meet in two weeks (we show how this follows from SDRT's cognitive model in Chapter 9), and that *Q-Elab*$(36a, 36b)$ holds (we demonstrate in Chapter 9 that this connection is inferred in the glue logic). Let's now focus on *Q-Elab*'s semantic effects. These match our intuitive interpretation of (36b) in this context. First, *Q-Elab*$(36a, 36b)$ entails that B can meet A on the seventeenth;[13] for if he can't, then even if A answers *yes* to the question, A is going to be no nearer knowing an executable plan for achieving his SARG, and so the monotonic constraints on *Q-Elab* holding would be violated.[14] Secondly, *Q-Elab*$(36a, 36b)$ implies that the next two weeks contains an interval of time which satisfies the description *the seventeenth*. For only then do answers to (36b) (and in particular, the answer (36c)) help to specify a plan to reach A's SARG to meet in two weeks. For instance, B can use the answer (36c) to derive the following plan for meeting in two weeks: find out which days in two weeks apart from the seventeenth are days when A is free. Further, B could not derive this plan (as executable and leading to the SARG of meeting in two weeks) from what A and B mutually believed given the discourse context (36a). From a model theoretic perspective, this plan is a set of pairs of states (initial and end states of the plan) that's a distinct subset of the set of state pairs that are the value of the general plan of meeting in two weeks. So both parts of the consequent of `Axiom on Q-Elab` are satisfied (making *Q-Elab* consistent), but only if *the seventeenth* denotes an interval contained within the next two weeks.

[13]Note that this implicature doesn't result from resolving ellipsis; (36b) doesn't involve any ellipsis. The question *How about the seventeenth?* does include an elided construction, however, which would be resolved to *How about we meet on the seventeenth?*, and the use of *we* in this resolved utterance implicates that B can meet A on the seventeenth. Thus this ellipsis resolution would bring about similar inferences to the ones that are generated by inferring the rhetorical relation *Q-Elab* in (36).

[14]In fact, this implicature leads to an interpretation of (36b) as an *indirect speech act* (ISA), since the implicature is of a different semantic type to the question (being a proposition) and the semantics of a distinct speech act type from *Q-Elab* encapsulates its connection to (36a). The relation is in fact *Plan-Elaboration*, to be discussed shortly. So (36b) is an ISA of type *Q-Elab* • *Plan-Elaboration*.

This latter inference, that the seventeenth must fall within the next two weeks, helps to predict incoherence. If (36) is uttered on the 20th of the month, we correctly predict that the rhetorical function of B's question cannot be a *Q-Elab* (for this time of utterance would ensure that the interval in two weeks doesn't contain an interval that's denoted by *the seventeenth*). Now suppose that we replace (36b) with (36b′):

(36) b′. B: Is your wife's hair grey?

This fails to satisfy the constraint on *Q-Elab* too, unless very special circumstances are assumed, since answers to (36b′) don't typically help A devise a plan to meet. So if A attaches (36b′) to (36a) with *Q-Elab*, the resulting SDRS will be unsatisfiable because of `Axiom on Q-Elab`. In fact, the glue-logic axioms will make *Q-Elab* the only candidate relation (see Chapter 9), at least in this context; so adding the information $\neg$*Q-Elab*$(36a, 36b')$ to the premises in the glue logic won't yield an alternative. The fact that the only SDRS that's predicted by the glue logic is unsatisfiable (for A) in the logic of information content is why (36ab) when it's uttered on the 20th of the month and (36ab′) are odd. This is not to say that (36b′) couldn't ever be attached with *Q-Elab*. A and B might have some special beliefs that allow them to use answers (36b′) to fulfill A's SARG for (36a).

The discourse relation *Q-Elab* also contributes to the analysis of (12):

(12) a. A: How about meeting next weekend?

b. B: That sounds good.

c. Shall we meet on Saturday afternoon?

d. A: I'm afraid I'm busy then.

e. ??How about 3pm?

Consider the inferences that arise when trying to attach (12e) to (12d). The extensions to the glue logic in Chapter 9 will ensure that *Q-Elab* is the only candidate relation. For now let's just assume that *Q-Elab*$(12e, 12d)$ holds. By `Axiom on Q-Elab`, answers to (12e) must elaborate a plan to achieve a goal behind (12d).

Now if *3pm* in (12e) resolves to 3pm on Saturday, then as before the *Q-Elab* relation would entail that A is able to meet at 3pm Saturday afternoon, which contradicts his response in (12d).[15] This interpretation of *3pm* also doesn't help A (or B) devise a plan to meet on Saturday morning or on Sunday, which is the SARG of (12d). But there is no other possible content for *3pm* other than 3pm on Saturday, because of SDRT's constraints on antecedents. So if we attach (12e) to (12d) with *Q-Elab*, then we can resolve *3pm* only on pain of inconsistency. Therefore, *Q-Elab* can't hold between (12e) and the context, even though this is

[15]Note that the SDRS representing the dialogue (12a–d) entails (12d) is true, because it connects to (12c) with the right-veridical relation *IQAP*.

the default expectation. Note that (12e) can't attach to (12a) with *Q-Elab* either (in fact, the presupposed content of *Saturday* as a definite is bound to (12a), and so the only accessible antecedent for *3pm* that generates a unique bridging relation is still Saturday). This *Q-Elab* relation between (12a) and (12e) still implies that A can meet B at 3pm on Saturday, for the same reasons this followed when attaching (12e) to (12d). And so attaching (12e) to (12a) with *Q-Elab* would still result in an inconsistent SDRS, thanks to the compositional semantic content of (12d) (that's entailed by the SDRS). Since a consistent interpretation cannot be constructed, SDRT predicts correctly that the dialogue as it stands sounds odd.

Of course, a complete analysis of these dialogues needs precise predictions about how plans evolve during dialogue processing, and this forms the focus of much of the prior work on dialogue interpretation (e.g., Grosz and Sidner (1990), Litman and Allen (1990), Poesio and Traum (1995)). In particular, we must model how SARGs accumulate and are refined through *Q-Elab*s. We will defer this problem to Chapter 9, when we examine in detail how axioms of agent rationality and cooperativity allow us to derive how SARGs and plans evolve. We will give axioms for computing SARGs that use compositional semantics, lexical semantics and discourse relations as clues.

Plan-Elab (standing for Plan Elaboration) is a relation that is just like *Q-Elab*, except that the second argument to the relation labels a proposition rather than a question. *Plan-Elab*(α, β) holds if K_β allows $S(\alpha)$ to compute a plan for achieving his SARG. Dialogue (37) features a *Plan-Elab* relation:

(37) a. I need to get to London by 1pm tomorrow.

b. British Airways has a flight which leaves at 10am.

Like *Q-Elab*, *Plan-Elab* is a subordinating relation, but unlike *Q-Elab*, it's right-veridical; *Plan-Elab*$(\alpha, \beta) \Rightarrow K_\beta$ is valid. The `Axiom on Plan-Elab` below adds additional content beyond right veridicality; it uses plans in the same way that the constraint on *Q-Elab* did:

- `Axiom on Plan-Elab`:
  (*Plan-Elab*$(\alpha, \beta) \wedge$ SARG$(\alpha, p)) \Rightarrow$
  (a) $\exists a((p, KB_{\tau,S(\alpha)}, {}^\wedge K_\beta, KB_{\tau,S(\alpha),S(\beta)}) >> a$
  (b) $\wedge$*Executable*$(a) \wedge \neg({}^\vee p >$ *Done*$(a)))$.

The difference is that this time, ${}^\wedge K_\beta$ itself must provide the information for elaborating the plan a, rather than answers to the question β.

The relation *R-Elab* is similar, save that the second argument to the relation is a request:

(38) a. I want to catch the 10.20 train to London.

b. Go to platform 1.

R-Elab(α, β) holds only if β is a request (and hence K_β is an action term), and this action is part of a plan a (i.e., $Done(a) \Rightarrow Done(K_\beta)$) to achieve the SARG p of α (i.e., $Done(a) >^{\vee} p$), where a is executable and no plan a' for achieving p was mutually known before β was uttered and interpreted. This last clause ensures the informativity of β in this relation:[16]

- `Axiom on R-Elab:`
  $(R\text{-}Elab(\alpha, \beta) \wedge \text{SARG}(\alpha, p)) \Rightarrow$
  (a) $(\exists a((Done(a) \Rightarrow Done(K_\beta)) \wedge executable(a) \wedge (Done(a) >^{\vee} p))$
  (b) $\wedge \neg \exists a'(KB_{\tau, S(\alpha), S(\beta)} > (Done(a') >^{\vee} p)))$

The veridicality of *R-Elab* depends on the semantic type of its arguments. For example, if α and β both label requests, then *R-Elab*(α, β) is veridical; but if α is a question and β is a request (e.g., (32ab$''$)), then *R-Elab*(α, β) is not veridical.

(32) a. A: Who came to the party?
b$''$. B: Look in the visitor's book to find out.

R-Elab connects (32b$''$) to (32a) to reflect the fact that it describes an action which elaborates a plan to achieve the SARG of the question (which is to know an answer). Note that the relationship between the content of this request and the question is distinct from that in (29). In (29), the content of (29b) was sufficient to compute a direct answer to (29a) *without* carrying out the action described by the request. In contrast, the content of (32b$''$) isn't sufficient for A to infer a true direct answer.[17] Rather, A has to *do* the action it describes in order to know an answer. This is reflected in our discourse semantics via the distinct truth conditions of the speech act types $IQAP_r$ and *R-Elab*.

Finally, one agent can reject the SARGs associated with another agent's speech act (a speech act that Searle (1968) defines as rejection):

(16) a. A: Let's go to the movies tonight.
b. B: I have to study for an exam.

We record this move with *Plan-Correction*: *Plan-Correction*(α, β) means that K_β nonmonotonically entails that $S(\beta)$ has intentions or goals that are incompatible with the SARG $S(\alpha)$ had in uttering α. In this case, planning to study for an exam (tonight) is incompatible with going to the movies. Note that recognising B's rhetorical move as *Plan-Correction* is co-dependent on the inference that he intends to study *tonight*. So this is yet another example where the rhetorical relation has semantic effects. *Plan-Correction* is a subordinating relation, because $S(\alpha)$ can then tell $S(\beta)$ why his original utterance α has a SARG that should not be

[16] For *Q-Elab* and *Plan-Elab*, the definition of $>>$ contains the informativity constraint.

[17] In particular, according to the compositional semantics of questions given in Chapter 2, *whoever signed the visitor's book* is not a direct answer to *Who came to the party?*

rejected (as we'll see in Section 7.6.5, this would attach to α with a relation called *Explanation*$*$). So α must still be available, even when it's been 'plan-corrected'. *Plan-Correction* appears to be veridical on the right, which means that it will function semantically like *IQAP*. It may be that one can have anaphoric links across the two elements as in (39):

(39) a. A: Let's take a train to India.
b. B: That's too dangerous.

The pronoun *that* refers back to the event type of taking a train to India. But this is an abstract object, and according to Asher (1993) these are subject to the SDRT-theoretic availability constraint.

With these rhetorical relations in place, let's now examine an example of an **indirect speech act**:

(40) a. A: Let's meet on Saturday.
b. B: Can we meet on Sunday instead?

Let's assume that (40ab) are labelled α and β respectively. Then β labels a question whose semantics is inconsistent with *Q-Elab*(α, β), for its answers fail to provide A with information that helps him to generate a plan to meet B on Saturday. In fact, through asking this question, B appears to be conveying a *Plan-Correction*. But this is an incompatible type of speech act from asking a question, since the second argument to any *Plan-Correction* is a proposition or a request (and it's right-veridical, making it inconsistent with the second argument being a question). So according to the definition of ISAs we gave in Section 7.4.1, (40b) is an ISA of type *question*•*Plan-Correction*. This means that (40b) actually generates *two* labelled bits of content: one label β_1 labels the compositional semantics of the question; the other label β_2 labels the proposition that attaches to α with *Plan-Correction*. So β_2 labels something like *Saturday is not good for me*. We can therefore say something more specific about the type of speech act that β_1 is: the question attaches with *Q-Elab* to β_2 (i.e., the plan-corrective portion of (40b)). So the speech act type of (40b) is in fact *Q-Elab*•*Plan-Correction* (note that *Q-Elab* is a subtype of the speech act type *question*).

With this analysis of questions as ISAs in mind, let's return to (12):

(12) a. A: How about meeting next weekend?
b. B: That sounds good.
c. Shall we meet on Saturday afternoon?
d. A: I'm afraid I'm busy then.
e. ??How about 3pm?

We argued that (12e) is incoherent because it can't be consistently interpreted as a *Q-Elab*. But what stops one from interpreting it as an ISA; in particular, as a *Plan-Correction*? Well, it should be clear that such an interpretation is inconsistent too, since A would then be plan-correcting himself. The analysis for *Plan-Correction* as it stands doesn't stipulate that A can't do this. But observe that (16) and (39) are incoherent if A and B are the same person. We can add such a constraint in SDRT, and we'll assume hence forward that some such constraint obtains.

More interestingly, (12) is still incoherent, *even if* we change the speaker of (12e), as in (12e″):

(12) e″. B: ??How about 3pm?

So why can't (12e″) be interpreted as a *Plan-Correction*? Well, it can't be interpreted this way, because the question portion of the ISA could still not be rhetorically connected to the context. In particular, connecting it with *Q-Elab* to the proposition that expresses the *Plan-Correction* (e.g., *I can only make it on Saturday afternoon*) would be inconsistent, because B already knows the answer to this question, thanks to (12d), and so its answers don't generate new information for B that help him achieve his SARGs. This indicates that the ISA interpretation would have to be one which conveys a *Correction* of (12d) as well (i.e., B would be conveying a proposition which expresses the information that he doesn't believe (12d) is true; see Chapter 8 for details of the semantics of *Correction*). But the linguistic form of B's question doesn't allow this interpretation, and the conditions we specify in Chapter 8 for inferring *Correction* will reflect this. Contrast this case with one where B utters *Are you sure you can't meet at 3pm?*; this is more acceptable than (12e″) because this does allow one to infer that an (indirect) *Correction* speech act has been performed.

7.6.3 Question Coordination

In discourse structures such as those of Grosz and Sidner (1986), goals that can still be attended to form a stack. Once a goal is achieved, it and everything above it is popped off the stack. A right-frontier constraint on discourse structure, as utilised in SDRT for example, makes similar predictions about what can be attended to and what is 'closed off'.

The default goal of a question is that the questioner know an answer. So these constraints on what can be attended to predict that questions must be answered in the reverse order in which they were asked, or not answered at all. This prediction captures the intuition that answering an earlier question can render the most recent question moot. For instance, the answer (41d) to question (41a) makes the question (41c) moot:

(41) a. A: Where are you?

b. B: Let me see. I'm in a village, uh,

c. A: What road did you take leaving Toulouse?

d. B: I'm in Couiza.

This is predicted by Grosz and Sidner (1986)—the goal of finding an answer to (41c) is popped off the stack once (41d) is interpreted as an answer to (41a). So subsequent utterances can't be interpreted as answers to (41c). The right-frontier constraint in SDRT makes a similar prediction: *Q-Elab*(41a, 41b) and *IQAP*(41a, 41d) ensure that (41c) isn't on the right frontier anymore, making it unavailable for subsequent attachment. So utterances subsequent to (41d) can't answer (41c) (unless this question is explicitly reintroduced).

Ginzburg (1995a,b) has observed that these predictions are problematic:

(42) a. A: Where were you on the 15th?

b. B: Uh, let me think.

c. A: Do you remember talking to anyone right after the incident?

d. B: I was at home.

e. I didn't talk to anyone after the incident.

(42d) answers (42a), and (42e) answers (42c). This violates both Grosz and Sidner's stack constraint and the current definition of availability in SDRT, for the 'open' questions are answered in order rather than reverse order. And yet (42) is acceptable. This problem also emerges in the following continuation of (41), where the second clause of (41d′) answers (41c) even though the 'superordinating' question (41a) is already answered by (41d′)'s first clause:

(41) d′. B: I'm in Couiza. I left Toulouse by route A61.

d″. B: Couiza. Route A61.

We have to modify availability in SDRT to reflect these possibilities. (41a-d′) and (42) show that, like the relations *Parallel* and *Contrast*, there are structures involving questions that make a strict right-frontier constraint or stack architecture unworkable. In response to this, we propose that full answers in a *single* conversational turn recapitulate enough of the material in the question that they can attach with *IQAP* to *any* question node that was available at the start of that turn. We restrict this to *full* answers because (41d″) cannot be interpreted this way. In effect, this yields something like the coordinate structures of Ginzburg (1995a,b): one can answer all unanswered questions in any order, so long as it's done in a single turn and one doesn't use fragmentary answers. Thus in (41abcd′), the questions (41a) and (41c) are closed off only when B has *finished* his reply (41d′), but not before. Similarly for (42).

This 'loosens up' availability, but not to the point where there are no constraints at all. Once B has finished his conversational turn (43d), the question (43c) is closed off and so B's contribution in (43f) is correctly predicted to be odd.

(43) a. A: Where are you?

b. B: Let me see. I'm in a village, uh,

c. A: What road did you take leaving Toulouse?

d. B: I'm in Couiza.

e. A: OK I see. I'll come get you.

f. B: ??I left Toulouse by route A61.

Furthermore, constraints in addition to the ones given by availability constrain the order in which questions can be answered. The availability constraint allows one to use a *single* turn to answer the questions in any order. However, if the semantics of the questions demand the same kind of answer (e.g., they are all *yes/no*-questions), then the principle of `Maximise Discourse Coherence` (`MDC`) constrains the order of answerhood further:[18]

(44) a. A: Do you want to come on the picnic, do you have a swimsuit and do you want to go swimming afterward?

b. B: Yes, yes and maybe.

If one assumes that (44) answers the questions in the order in which they were asked, then in addition to the *QAP* relations that hold between each question and its answer, one can also infer *Parallel* relations between the question-answer pairs. If the questions aren't answered in order, however, one cannot support an inference that *Parallel* holds. `MDC` stipulates that we (monotonically) prefer interpretations that maximise the number of relations. And so `MDC` together with availability capture the intuitive interpretation of (44): the questions are answered in the order in which they were asked.

7.6.4 Other Relations Involving Questions

The semantics of *Q-Elab* uses cognitive states. But the content of a question—or its answers—can also be related to the context purely at the 'domain' level; i.e., via the individuals and events that are described in the utterances. For example, consider the following dialogue:

(45) a. A: John arrived at the party at 8pm last night.

b. B: And then what happened?

[18] This example is due to Dan Hardt, personal communication.

Intuitively, coherent answers to B's question (45b) should form a *Narration* with (45a). Thus in a similar manner to previous examples, the rhetorical connection of the question to its context makes the content of the question more than what's linguistically explicit: (45b) is not just a question about any event that occurs after 8pm, but it's a question about which events happened *which are occasioned by John's arrival.* This serves to explain why (45c) is a better continuation of the dialogue than (45c′) (unless subsequent utterances explain how the events of John's arrival and the sun setting are connected), and it also explains why one infers that Mary was at the party in the continuation (45c″) of the dialogue (this follows from the spatio-temporal effects of *Narration*, as defined in Chapter 4):

(45) c. A: He danced with Mary.

c″. A: ??The sun set.

c″. A: Mary made a fool of herself.

The relation between the proposition and question in (45ab) is distinct from *Narration* itself, however. First, *Narration* is veridical, and it doesn't make sense to say that (45b) is true. Second, *Narration* is a coordinating relation, but as (45c) and (45c″) show, subsequent utterances can attach to (45a), and so we should attach (45b) to (45a) with a *subordinating* relation to reflect (45a)'s availability.

Narration doesn't hold between (45a) and (45b), but it *does* hold between (45a) and any answers to (45b). So we mark the relation between the proposition (45a) and the question (45b) as $Narration_q$. Informally, $Narration_q(\alpha, \beta)$ holds just in case α is a proposition and β is a question, and any possible answers K to K_β (according to the compositional semantics of questions and answers) are such that K_α and K satisfy the necessary constraints on *Narration*. So $Narration_q$ is left-veridical (i.e., $Narration_q(\alpha, \beta)$ entails K_α), because K_α satisfies *Narration*'s constraints only if it's true. $Narration_q$ is not right-veridical, however, since K_β is a question.

If one computes $Narration_q(45a, 45b)$ for (45)—indeed, the cue word *then* would help predict this relation—then the correct spatio-temporal implicatures of the question (45b) follow. Now, it's possible that, as well as $Narration_q(45a, 45b)$ holding, $Q\text{-}Elab(45a, 45b)$ may hold too. For suppose that A's SARG in asserting (45a) was one of belief transfer (i.e., that B believe (45a)).[19] And suppose that the content (45a) itself is insufficient for this goal to be achieved. Then answers to B's question (45b) provide further information about events last night, and as such elaborate a plan for achieving the goal that B believe (45a) (or alternatively, the answers will provide B with information from which he infers that the goal that he believe (45a) is unachievable). In this scenario, (45b) is a clarification question (or in other words, a question for justification), of the kind that are discussed in van Kuppevelt (1996). For us, clarification questions are simply a

[19] Indeed, we might take this to be a default goal for assertions. We'll discuss this in Chapter 9.

particular kind of *Q-Elab*: one in which the SARG is belief transfer. Note that the constraints on *Q-Elab*, as described in `Axiom on Q-Elab`, would only be satisfied in (45) because of the domain-level relation $Narration_q$ between the utterances. So the cognitive-level *Q-Elab* relation holding is dependent on the domain-level one holding in this case. This kind of co-dependence between domain-level and cognitive-level rhetorical relations was observed in Moore and Pollack (1992), and we'll return to this issue when we discuss requests in Section 7.6.6.

It's important to recognise the domain-level relationship between (45a) and (45b) in contexts where *Q-Elab* doesn't hold. (45b) is coherent even in contexts where it's mutually known that A's SARG for (45a) has been achieved, making *Q-Elab*(45*a*, 45*b*) unsatisfiable. For example, B can explicitly indicate that belief transfer is successful, and still follow up with the question (45b):

(46) a. A: John arrived at the party at 8pm last night.

b. B: Yes, I know. And then what happened?

If we did not include 'domain-level' relations such as $Narration_q$ in the ontology, then we would be unable to explain this coherence.

As well as *Narration* having its 'question' counterpart, so do *Elaboration*, *Explanation*, *Result* and *Background*, as shown in (47ab), (47cd) (from van Kuppevelt (1996)), (48) and (49):

(47) a. A: A well-known book publisher is searching for manuscripts.

b. B: What kind of manuscripts?

c. A: Fiction will be considered.

d. B: How do you know?

(48) a. A: John failed his degree.

b. B: What effects did that have on his career?

(49) a. A: John failed his degree.

b. B: Was he living in student dorms at the time?

We mark these relations as $Elaboration_q$, $Explanation_q$, $Result_q$ and $Background_q$ respectively. They have a similar semantics to $Narration_q$; we forego giving formal details. As before, relations at *both* the domain level and the cognitive level—in the form of *Q-Elab*—may hold. For example in (47), the roles of the questions may be to help achieve the goals that B believe (47a) and (47c).

7.6.5 Metatalk Relations

Some rhetorical relations connect the content of one utterance to the *performance* of uttering another rather than to its content. Following Polanyi (1985), we call these *metatalk* relations. For example, consider (50) and (51):

(50) a. A: It's getting late.

b. B: Aren't you enjoying yourself?

(51) I'm cold. Please close the window.

In (50), the discourse relation is a metatalk one, because an answer to the question (50b) will explain why the speech act of asserting (50a) was performed; or more specifically, why the *indirect* speech act of requesting to leave was performed. This explanation of the speech act rests on the fact that answers to the question (50b) will explain why A has the goals which led to his uttering (50a). Note that *Q-Elab*(50a,50b) can't hold, because answers to (50b) don't elaborate a plan to achieve A's SARG of (50a) (neither the conventional SARG that B believe it's getting late, nor the SARG underlying the indirect speech act, that A leave). Nor do answers to (50b) provide explanations or elaborations for the *content* itself of (50a), since A's level of enjoyment has nothing to do with lateness. In (51), the proposition that I'm cold results in my making the request to close the window, because it results in my having the SARG associated with this request, of you closing the window (so that I can get warmer).

The relation between the constituents in (51) is *Result**: *Result**(α, β) entails that the content of α results in $S(\beta)$ having the SARG associated with his speech act β. This relation is *-ed to indicate that it's different from normal *Result*s, in that the content of α doesn't result in the *content* of β, rather it results in the performance of the speech act. However, it's still a result of some sort, and hence the name *Result**. Observe that *Result** is veridical, and so it functions semantically like conditions given by other veridical relations as discussed in Chapter 4. Similarly, *Explanation**(α, β) entails that the content of β explains why $S(\alpha)$ performed the speech act α; it's also veridical. *Explanation** connects *I can't right now* to *No* in (54ac), since *I can't right now* explains why B has performed, through uttering *No*, the speech act of *Plan-Correction* to A's request (54a):

(54) a. A: Please go get some bread.

b. B: No, I can't right now.

I can't right now itself may also attach to (54a) with *Plan-Correction* (since its SARG is incompatible with the A's SARG that B get some bread).

One can also express metatalk in a conditional; i.e., if the antecedent is true then the speech act associated with the consequent is performed. This is illustrated in (52a) where the question is asked on condition that the antecedent is true, and in (52b) where the imperative is requested if the antecedent is true.

(52) a. If you failed the test, then why should I listen to you?

b. If you failed the test, then promise me you won't tell anyone.

The constituents in these sentences are connected with *Consequence** to reflect this semantics (and in (52b) they are also connected with *Consequence*).

More generally, metatalk relations have the following characteristic: $R^*(\alpha, \beta)$ holds, where R is the content-level discourse relation (e.g., *Explanation* or *Result*), if and only if the content of one of the arguments (i.e., α or β) stands in the relation R to fact that the speaker of the other utterance has the SARG of that utterance. For example, if β is a request (and so its SARG is that the action described is performed), then $Result^*(\alpha, \beta)$ entails that the content of α results in the speaker having the SARG that this action is performed; this holds for (51), where being cold results in the speaker having the goal of getting someone to close the window. In (50), answers to (50b) will explain why A wants to leave (and hence why he performed the indirect speech act of requesting to leave). So a complex metatalk relation holds between (50a) and the question (50b); a relation we might call $Explanation_q*$. $Explanation_q{}^*(\alpha, \beta)$ holds iff an answer to β would explain why α was performed.

Recognising this indirect speech act interpretation of (50a) is crucial to understanding B's 'metatalk' move (i.e., to seek an explanation for why A is requesting to leave). In (53a), A stipulates the reasons for having the SARG that underlies his indirect request for help, that's associated with the question *Can you help me? Result** holds between these clauses, because being broken down results in A asking for help. *Result** also connects the clauses in (53d):[20]

(53)
a. A: Je suis tombé en panne. Est-ce que tu peux m'aider?
I've broken down. Can you help me?

b. B: Où es tu?
Where are you?

c. A: Je suis devant le refuge qui se trouve à envion un km après Couiza. Il y a là une cabine telephonique.
I'm just in front of the refuge which is about 1km after Couiza. There's a telephone box.

d. B: Il y a plusieurs refuges aux alentours de Couiza. Dans quelle direction es tu parti de Couiza?
There are several refuges near Couiza. In which direction did you leave Couiza?

This is because B is first correcting a presupposition in A's utterance (53c) that is introduced by the definite noun phrase—that there's only one refuge 1km beyond Couiza. But this *Correction* also in turn serves to motivate why he then asks his next question. Hence *Result** connects the clauses.

[20](53) is taken from the Toulouse-Stuttgart Procope corpus, and we'll give a full analysis of this example in Chapter 8.

7.6.6 Other Relations Involving Requests

We have already seen that requests can answer questions (e.g., (30)); they can convey one doesn't have enough information to answer the question (i.e., they can be arguments to *NEI*; e.g., (32ab′)); and requests elaborate plans to achieve prior SARGs (*R-Elab*; e.g., (38)).

There are also a number of possible responses to a request: one can acknowledge it (e.g., (54ab)); or refuse to accede to the request (which we marked with *Plan-Correction*, e.g., (54ac)); or one can follow a request with a *Q-Elab* (as in (54ad), where answers to (54d) provide information that help B perform the action); or one can respond to a request with another request that's connected to it with *R-Elab* (so the second request describes an action that's part of a plan to achieve the action described in the first request, as in (54ae)):

(54) a. A: Please go get some bread.

b. B: Ok.

c. B: No.

d. B: Do you have any money?

e. B: Please get me some money to pay for it.

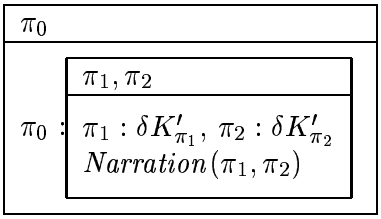
A sequence of imperatives uttered by the same speaker, however, often exhibit similar effects on content as the corresponding indicatives. For example, compare the narrative (55) with the corresponding sequence of imperatives (56) (modified from Webber *et al.* (1995)):

(55) John entered Bill's office. He got a red file folder.

(56) Enter Bill's office. Get a red file folder.

In both cases, one infers that the red file folder is in Bill's office. Indeed, getting a red file folder that's not in Bill's office would not constitute an action which satisfies the request given in (56). We used the semantics of *Narration* to capture these spatio-temporal implicatures for (55). So the intuitive interpretation of (56) is captured in its logical form (56′), which features this relation too (where K'_{π_1} represents the proposition that the addressee is in Bill's office, and K'_{π_2} represents the proposition that the addressee gets a red file folder):

(56′)



First, the veridicality of *Narration* yields the desired effect that the two imperatives are commanded in sequence; i.e., there's a command to go to the office and a command (once the former has been accomplished) to get a red file folder. Secondly, the spatio-temporal consequences of *Narration* (see Chapter 4) entails the red file folder that's introduced in K'_{π_2} is in Bill's office, and the event of going to Bill's office must precede that of getting a red file folder (this also follows from the two commands being commanded in sequence). All of these semantic consequences of *Narration* agree with our intuitive interpretation of (56).

In Chapter 5, we inferred *Narration* when interpreting (55) on the basis of a default rule whose antecedent was verified by the lexical semantics of the constituents. This default axiom would apply in the imperative case (56) as well, given the very similar lexical semantics; in fact, the only thing that's different is the sentence mood, but the axiom for inferring *Narration* was neutral about sentence mood. Thus we explain the similarities between (55) and (56) by (a) using the same default axioms in the glue logic to construct the SDRSs, and hence (b) using the same rhetorical relations in their logical forms, to predict the same spatio-temporal implicatures.

It would be hard to obtain such a uniform analysis of (55) vs. (56) in plan recognition approaches to dialogue interpretation, since in these approaches one would interpret the second clause on the basis of the *goals* of the first clause, but not its content directly. Given that the goals associated with indicatives vs. imperatives are very different (i.e., belief transfer vs. performing an action), the proofs which underly the interpretation of the second clause would be using entirely different premises. Furthermore, such an approach would need to rely on proving that the speaker intended the actions to be performed in the order described from knowledge of his intentions and his knowledge of domain-level plans of action; such a proof is elusive for (56). In contrast, we obtain an interpretation of (56) while bypassing reasoning about intentions, relying instead only on linguistic information; and then we simply assume that the speaker had the requisite intentions for conveying the content that was computed (e.g., because *Narration* is veridical, we presume the speaker intends for the actions to be performed in sequence). We believe that exploiting linguistic content in this way is conceptually more elegant, since one should interpret discourse by using observable information (e.g., lexical semantics), as opposed to unobservable information (e.g., goals), whenever possible.

Elaboration can be used in a similar way to the narrative example (56) to capture the semantics of (58) and its similarities to (57):

(57) John took the train from Paris to Madrid. He changed in Toulouse.

(58) Take the train from Paris to Madrid. Change in Toulouse.

Again, the inferences which yield *Elaboration* are exactly the same for both (57) and (58) (see Chapter 5): e.g., knowledge about the locations of Paris, Madrid

and Toulouse play a role. The speaker's intentions aren't part of the proof that *Elaboration* holds in (58). However, the entailments of *Elaboration* reveal things about the speaker's goals: the part-whole relation between the events constrains the actions that he wants performed.

Now consider two examples where an imperative is followed by a proposition which describes a defeasible consequence of performing the action described by the imperative:

(59) a. Turn left

b. and you'll see the roundabout.

(11) a. Smoke a packet of cigarettes a day

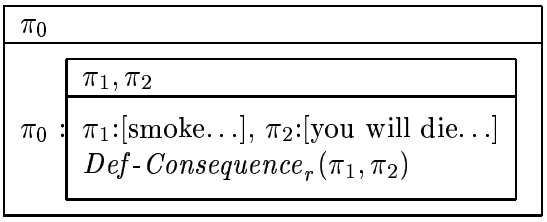
b. and you will die before you're 50.

As we observed in Chapter 2, the imperative (59a) is typically understood as commanded but the one in (11) is not. We are now in a position to represent this difference in content, although *predicting* this difference in the glue logic requires reasoning about goals; something we explore in Chapter 9. But let's focus on the representation of content for now, dealing with (11) first. As we said earlier, when the conjunct *and* connects an imperative and an indicative (in that order), it appears to have a meaning other than its usual truth functional interpretation: *and* in (59) and (11) portrays a *conditional* relation, rather than a conjunctive one. From Chapter 4, the appropriate discourse relation to reflect this conditional relationship is something analogous to *Def-Consequence*: $Def\text{-}Consequence_r$ is like *Def-Consequence* save that K_α is an action term:

$$(w, f)[\![Def\text{-}Consequence_r(\alpha, \beta)]\!]_M(w'g,) \text{ iff } (w, f)[\![[K_\alpha]\top > K_\beta]\!](w', g)$$

That is, doing the action K_α results in a state where normally K_β is true. $Def\text{-}Consequence_r$ isn't veridical. Thus $Def\text{-}Consequence_r(11a, 11b)$ does *not* entail (11a) is commanded. Nor does it entail that (11b) is true, as required. So the SDRS (11′) captures its intuitive interpretation:

(11′)



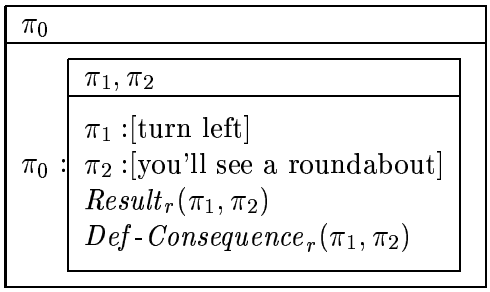
The situation is a bit different for (59), however. Unlike (11), the state of affairs described in the indicative isn't deemed to be something that both the hearer and speaker wish to avoid. Furthermore, the imperative is commanded. But how do we capture this in the logical form of (59)? And how does the glue

logic predict the differences between (59) vs. (11)? Well, we assume that the glue logic supports the following principle: whenever the indicative consequence of an action is not *a priori* desirable nor undesirable (i.e., the SDRS features π_0 : *Def-Consequence*$_r(\pi_1, \pi_2)$ and the cognitive states of the participants feature $\neg wants_{S(\pi_2)} \neg K_{\pi_2}$ and $\neg wants_{S(\pi_2)} K_{\pi_2}$), then π_1 and π_2 are also connected with a further rhetorical relation *Result*$_r$ that has a more specific semantics than *Def-Consequence*$_r$ in that it is left-veridical (note that K_α is an action term):

$$(w, f)[\![Result_r(\alpha, \beta)]\!]_M(w', g) \text{ iff } (w, f)[\![K_\alpha \wedge ([K_\alpha]\top > K_\beta)]\!]_M(w', g)$$

So if we assume a context for (59) where the consequence of the action is not *a priori* either desirable or undesirable, then the compositional semantics of *and* will generate the *Def-Consequence*$_r$ relation, and thus the above principle in the glue logic will generate the logical form (59″) of (59):

(59′)

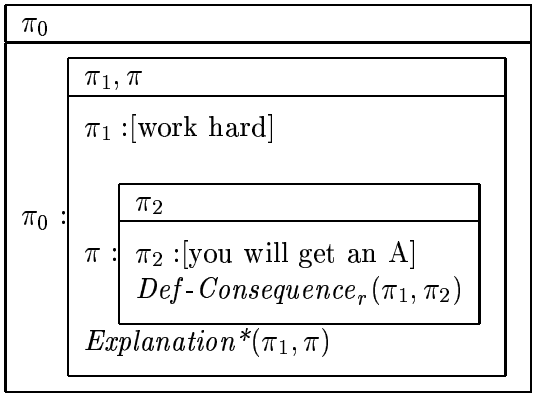


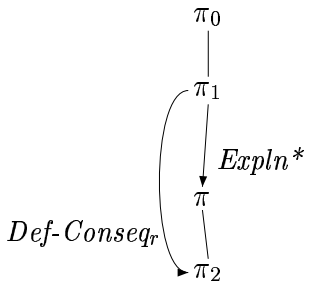
By the left-veridicality of *Result*$_r$, the imperative *turn left* is interpreted as commanded. Now suppose that, in contrast to both (59) and (11), the defeasible consequence of the action is *a priori* desirable: e.g., (60):

(60) Work hard for the next month and you will get an A in the course.

Then the logical dependency between performing the action and the proposition being true (i.e., *Def-Consequence*$_r(\pi_1, \pi_2)$) also *explains why* α is commanded. So the logical form of (60) is (60′) (we've also depicted the structure graphically, and observe this meets the constraint on SDRSs that $Succ_D$ be well-founded):

(60′)





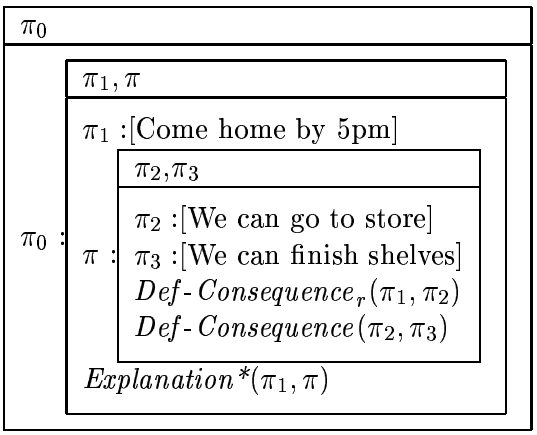
In words, (60′) says: working hard is commanded, and the fact that performing this action normally results in getting As explains why it's commanded. Note that we know working hard is commanded and that the defeasible consequence relation actually holds because *Explanation** is veridical. Note also that the imperative plays a 'dual' rhetorical role: at the content level, it's part of a conditional relationship; and at the cognitive level, it's being *motivated* by this conditional relationship. This is why π_1 is connected to labels at different levels in the structure: i.e., to π and to π_2 (which is outscoped by π).

Thus overall, knowledge of which states are desirable and which not must affect our decisions about the rhetorical connections in (59) vs. (60) vs. (11), which all have very similar surface forms. A similar discourse structure to (60′) conveys the intuitive content of (61) too (taken from Moore and Pollack (1992)):

(61) π_1. Come home by 5pm.
π_2. Then we can go to the hardware store before it closes.
π_3. That way, we can finish the shelves tonight.

Moore and Pollack argue that (61) poses a problem for RST, because the cognitive-level relation MOTIVATION is co-dependent on the domain-level relation CONDITION, but these relations are asymmetric in RST, and this asymmetry means that it's impossible for both these relations to hold simultaneously in a well-formed RST tree. SDRT doesn't assume the asymmetry of rhetorical connections that RST does; it also allows one utterance to be related to another at several levels. We therefore bypass these problems; and indeed we exploited this feature in our analysis of (60). The analysis of (61) is similar. Roughly, *Def-Consequence* relations (this relation is akin to CONDITION in RST) hold between π_1 and π_2, and between π_2 and π_3, and these domain-level connections explain why π_1 is commanded. I.e., there's an *Explanation** relation (*Explanation** being akin to MOTIVATION in RST) between the request and the label for the SDRS which represents the fact that there are these defeasible consequence relations between the imperative and the (desirable) propositions. So the SDRS for (61) is (61′):

(61′)



As in (60′), this SDRS captures the fact that the imperative π_1 plays two distinct roles in this discourse: it's part of an SDRS π which describes the typical outcome of performing the action; which in turn is an argument to an *Explanation** relation which conveys why the imperative is commanded. It also captures the fact that the imperative π_1 is commanded, but it doesn't entail that the propositions π_2 or π_3 are true. And it conveys the fact that the *Explanation** relation at the cognitive-level holds because of the *Def-Consequence* relations at the domain-level.

Presumably, the *Explanation** in (60) and (61) is inferred on the basis of knowledge about desirable states. That is, we would need a glue-logic axiom that goes something like this: if (a) γ labels $Def\text{-}Consequence_r(\alpha, \beta)$ where (b) K_β (or more accurately, its 'shallow' description in the glue language) is desirable, then the SDRS also includes (c) the relation $Explanation^*(\alpha, \gamma)$.

As we mentioned earlier, the glue logic should generate the different relation $Result_r$ if the consequence of the action is neither desirable nor undesirable but somehow 'neutral'. This rhetorical relation captures the intuitive interpretation of (62) (cf. (59) and (61)):

(62) Turn left. Then you will see a roundabout.

The imperative is commanded and an outcome of the action is stated, but it's not clear that this outcome was mentioned so as to explain why the request is commanded. It could simply have been stated so as to provide information that the hearer can use to check that he's performed the action of turning left correctly. The relation $Result_r$ has these semantic effects.

7.7 Conclusions

In this chapter we have extended the logic of information content in SDRT. We introduced some new discourse relations which (a) take labels for interrogatives and imperatives as arguments, and/or (b) reveal information about the beliefs and intentions of the dialogue agents.

We used these new discourse relations to model implicatures in dialogue. We analysed the semantics of questions and their responses (e.g., *IQAP*), and the role of questions and requests as 'information-providers' for achieving prior goals (e.g, *Q-Elab*, *R-Elab*, $IQAP_r$). We also observed that sequences of indicatives often generate similar implicatures to sequences of the corresponding imperatives; we modelled this by using the same rhetorical relations in their logical forms.

We argued that one can reconceive rhetorical relations as (relational) speech act types. This yields a more fine-grained taxonomy of speech acts than is traditionally envisaged (e.g., Searle (1975)). The semantics of different rhetorical relations are then tantamount to statements about illocutionary force.

There is still much work to be done. All of the relations surveyed so far are compatible with a view of a monotonic discourse update such as that given in Chapter 5, where the results of $update_{\text{SDRT}}$ simply add to the content of the discourse context so far. But that isn't always the case, as we'll see in Chapter 8. We need to represent the content of a dialogue when a dispute has taken place; e.g., (63):

(63) A: John got the top prize for his essay.
B: No, it was Sue who got it.

Such dialogues contain utterances with incompatible contents. So the discourse relations which connect them cannot be veridical, for this would render the semantics of the dialogue inconsistent (which it's clearly not, since it's not meaningless). In fact, interpretations of dialogues where disputes take place can be *non-incremental*, in that propositions which were entailed by the content of the dialogue context before the dispute took place may not be entailed by the dialogue after the dispute.

Secondly, we need a theory of how to *construct* logical forms for dialogue. This involves extending the glue logic so as to model complex interactions between the cognitive states of the participants, the speech acts they perform (i.e. the rhetorical relations), and their interpretations of other participants' moves. So we need a general theory of cognitive modelling, which is interfaced to the glue logic in a particular way: principles of rationality of cooperativity should serve to predict which rhetorical relations hold in the dialogue. The glue logic and accompanying theory of cognitive modelling is the topic of Chapter 9.