Agreement and Disputes in Dialogue

Alex Lascarides
School of Informatics,
University of Edinburgh
alex@inf.ed.ac.uk

Nicholas Asher
IRIT
Université Paul Sabatier, Toulouse
asher@irit.fr

Abstract

In this paper we define agreement in terms of shared public commitments, and implicit agreement is conditioned on the semantics of the relational speech acts (e.g., Narration, Explanation) that each agent performs. We provide a consistent interpretation of disputes, and updating a logical form with the current utterance always involves extending it and not revising it, even if the current utterance denies earlier content.

1 Introduction

A semantic theory of dialogue should account for what content dialogue agents agree on. This includes implicit agreement:

(1)  
  a. A: The room went dark.
  b. A: Max turned out the light.
  c. B: And John drew the blinds.

Intuitively, $A$ and $B$ agree that the room went dark, that Max turned out the light, and that the latter is at least part of the reason why the former occurred. Thus, implicatures can be agreed upon (that (1b) is part of the cause of (1a) goes beyond compositional semantics), and agreement can be implicated ($B$ does not repeat (1a) and (1b) nor utter OK to indicate his agreement with $A$).

In principle, the Grounding Acts Model (GAM, Traum (1994), Traum and Allen (1994)) supports implicit agreement. But it demands an acceptance act for agreement to occur, and its current rules don’t predict such an act from (1c). Segmented Discourse Representation Theory (SDRT, Asher and Lascarides (2003)) errs in the opposite direction. It stipulates that lack of disagreement implicates agreement, and so in (1) too much is agreed upon; e.g., (1c). Thus, SDRT needs modification to deal with (1), just as GAM needs supplementation.

Agreement can occur even in the context of corrections or disputes. In (2), $A$ asserts (2a) and $B$ its negation, but a consistent interpretation of (2) overall is a pre-requisite to explaining how $A$ and $B$ end up agreeing on (2b).

(2)  
  a. A: It’s raining.
  b. B: No it’s not.
  c. A: OK.

Since a correction negates content in the discourse context, an obvious strategy for maintaining consistency would be to revise the semantic representation of the context when updating it with a correction. But we want to avoid revision, both at the level of model theory and at the level of composing logical form. This is for two reasons. Firstly, revision means that there is in principle no general way of stating what information is preserved from the previous discourse state to the current one. But if we construct logical form in a monotonic way—in our case, this means that the discourse structure for a conversation at turn $n$ is an elementary substructure of the discourse structure at turn $n + 1$—then standard preservation results from model theory apply. Secondly, monotonicity guarantees that interpretation algorithms can proceed incrementally, combining information from various sources in a nondestructive way (Alshawi and Crouch, 1992).

To our knowledge, there is currently no dynamic semantics for dialogue that yields adequate interpretations of corrections and implicit agreement. We will address this gap here. In Section 2, we re-
view two existing approaches to motivate our basic strategy, which we then describe in Section 3. We will refine SDRT so that it tracks each dialogue participant’s public commitments. Further, while identifying a speech act involves default reasoning, constructing logical form will be monotonic, in the sense that the logical form of an updated discourse always extends that of its discourse context, rather than revising it.

2 Motivation

We will say that a proposition \( p \) is grounded just in case \( p \) is agreed by the dialogue agents to be true. This follows Clark’s terminology, in particular the concept of grounding a joint action at level 4 (Clark, 1996, p388). Clark’s work focusses almost entirely on grounding at the so-called ‘lower’ levels; how agents ground an understanding of what was said, for instance. By contrast, in order to focus on grounding at the higher level, we will assume a highly idealised scenario where dialogue agents understand each other perfectly, resolving ambiguities in the same way. One of Clark’s main claims is that grounding at all levels occurs only when there is positive evidence for it, and we aim to explore in a logically precise manner exactly what amount of positive evidence suffices for grounding a proposition. In future work, we intend to demonstrate that our definition of grounding can model grounding at the lower levels too; this will involve extending the framework to represent misunderstandings.

GAM links the speech acts performed with its effects, including effects on grounding (Traum, 1994). Each conversational participant builds a conversational information state (or CIS). Update effects of particular speech acts (and their preconditions) are specified in terms of changes to (and conditions on) the CIS. For example, Figure 1 is the update rule for the speech act \( e \) where \( B \) asserts \( K \) to \( A \). It updates the common ground \((G)\) to include an event \( e' \) that \( B \) intends \( A \) to believe \( K \) and a conditional event \( e'' \) that should \( A \) accept the assertion, then \( A \) would be socially committed to \( B \) to believe \( K \) (shown via the attitude SCCOE). The update rules form a hierarchy, so that more specific acts inherit effects from more general ones. The speech act in Figure 1 inherits that \( B \) is SCCOE-ed to \( A \) to \( K \), for instance. Decision trees then predict which speech acts have been performed.

While it is possible in principle for GAM to include rules that accurately predict \((1c)\)’s illocutionary effects, the rules that are actually provided only recognise \((1c)\) as an assertion. Consequently, its effects are under-generated: \( B \) is socially committed to \((1c)\), but not to \((1a)\), \((1b)\) or a causal relation between them. GAM needs to be supplemented with rules for inferring that \( B \) was also implicitly accepting parts of \( A \)’s contribution.

Such acceptances, we argue, should be conditioned on relational speech acts. \((1c)\) continues \((1b)\) as a narrative, and the narrative so formed explains \((1a)\). These are relational speech acts (Asher and Lascarides, 2003): they are speech acts because continuing a narrative or explaining something are things that people do with utterances; and they are relational because the successful performance of the speech act Explanation, say, is logically dependent on the content of the utterance (or sequence of utterances) that is being explained (in this case, \((1a)\)). Thus even though the compositional semantics of \((1c)\) does not entail \((1b)\) or \((1a)\), its illocutionary contribution does entail them—or, perhaps more accurately, entails that \( B \) is publicly committed to them. Similarly, through using \((1b)\) as an Explanation of \((1a)\), \( A \) is publicly committed to \((1a)\), \((1b)\) and a causal relationship between them. Thus, what is grounded amounts to the shared semantic entailments of the rhetorical relations—or speech acts—that both \( A \) and \( B \) performed. This explains why positive evidence for grounding is necessary (Clark, 1996): both agents must perform a speech act with appropriate semantic consequences for a proposition to become grounded. An implicit acceptance (or acknowledgement in SDRT terms) is then logically dependent on the formal semantic interpretations of the relational speech acts performed. For instance, \( B \)’s commitments to \((1a)\) and \((1b)\) stem from Narration and Explanation acts he performed in uttering \((1c)\).

Since GAM incorporates relational speech acts, the general principles that we propose here could extend it. However, we have chosen to use SDRT because it defines logical form more abstractly, allowing us to exploit its model theory to determine grounded propositions. In contrast to GAM, we will not explicitly represent what’s grounded (and what’s not) in logical form. Doing so would force us to in-
corporate revision should grounded content get disputed, as can happen in a dynamic setting, where facts and beliefs change as the agents engage in dialogue. We will make grounding a property of the interpretation of a logical form, and not part of its form.

SDRT offers a formal semantics of relational speech acts (Asher and Lascarides, 2003). Furthermore, in contrast to theories of discourse interpretation that equate interpreting a discourse with its effects on the agents’ beliefs (e.g., Hobbs et al. (1993), Grosz and Sidner (1990)), SDRT separates the glue logic (i.e., the logic for constructing a logical form of what was said) from the logic for interpreting the logical form (i.e., reasoning about whether what was said is true, or should be believed). This enables SDRT to maintain a decidable procedure for computing logical form, even though identifying the speech acts performed inherently involves commonsense reasoning, and hence consistency tests. Asher and Lascarides (2003, p78) argue that it must be decidable to explain why, as Lewis (1969) claims, people and large have a common understanding of what was said.

SDRT’s current representation of (1) is (1’), where $\pi_1$, $\pi_2$ and $\pi_3$ label the contents of the clauses (1a–c) respectively, and $\pi_0$ and $\pi$ label the content of the dialogue segments that are created by the rhetorical connections:

$$(1') \quad \pi_0 : \text{Explanation}(\pi_1, \pi) \quad \pi : \text{Narration}(\pi_2, \pi_3)$$

In words, (1’) implies that the room went dark, and this was caused by a combination of Max switching off the light followed by John drawing the blinds. In the absence of speech acts of denial such as Correction, SDRT stipulates that all content is grounded (Asher and Lascarides, 2003, p363). This leads directly to the wrong predictions for (1).

Unlike GAM, SDRT fails to track the different commitments of individual speakers. Simply labeling each speech act with its speaker doesn’t suffice, as dialogue (3) shows.¹

$$(3) \quad \pi_1. \quad A: \text{John went to Harrods.}$$
$$\pi_2. \quad B: \text{He bought a suit.}$$
$$\pi_3. \quad A: \text{He then entered the food halls.}$$
$$\pi_4. \quad B: \text{He looked for foie gras.}$$

Intuitively, A’s utterance $\pi_3$ publicly commits him not only to Narration($\pi_2, \pi_3$), but also to Narration($\pi_1, \pi_2$) (for this latter speech act entails, while the former does not, that John bought the suit at Harrods). And yet B was the speaker who performed the speech act Narration($\pi_1, \pi_2$), for it is B who uttered $\pi_2$. Accordingly, we abandon representing dialogue with a single SDRS, and replace it with a tuple of SDRSS—one SDRS per discourse participant per turn, representing all his commitments up to and including that turn. We define grounding a proposition $p$ in terms of joint entailments from those commitments, and hence grounding becomes a semantic property of the logical form. This solves SDRT’s over-generation problems with grounding. For instance in (1), A’s public commitments are to Explanation($\pi_1, \pi_2$). B, on the other hand, is committed to the content expressed by (1’). The shared public commitments then accurately reflect what A and B agree on. We also avoid the under-generation problems of GAM; grounding need not arise from an acceptance but instead from so-called veridical rhetorical relations (e.g., Explanation and Narration) and the logical relationships among their meanings.

Grounded content is not marked as such in logical form. This makes monotonic construction of logical form feasible, even when grounded propositions get disputed. A further part of our strategy for eschewing revision is to assume that the SDRSS for each turn represent all of A’s and B’s current commitments, although Sacks (1992) attests many similar, naturally occurring dialogues where the agents build a narrative together.

¹For simplicity, we use a contructed example here, although Sacks (1992) attests many similar, naturally occurring dialogues where the agents build a narrative together.
from the beginning of the dialogue to the end of that turn. The alternative, where prior but ongoing commitments from turn $i - 1$ are not shown in the representation of turn $i$, and accordingly the input context for interpreting turn $i$ is the output one from interpreting turn $i - 1$ (e.g., $A$’s utterance (2c)), and without revision the output context from turn $i$ would then be $\bot$. We want to avoid revision while maintaining consistency. Representing all current commitments in each turn avoids revision in the model theory, because one can compute the current commitments of $A$ and $B$ by dynamically interpreting their SDRSs for just the last turn. One can detect how $A$’s commitments have changed during the dialogue, but only by comparing the SDRSs for the relevant turns.\footnote{Prévo et al. (2006) represent dialogue in terms of commitment slates. Their idea inspired our work, but the details differ considerably, particularly on monotonic construction.}

We will model disputes by adding non-truth-preserving operators over relevant segments in the logical form. This avoids the need for downdating and revision in both the construction and the interpretation of logical form.

### 3 Individuating Commitments

The logical form for a dialogue turn proposed in Section 2 generalises to dialogues with more than two agents in the obvious way: the logical form of a dialogue turn is a set $\{S_a : a \in D\}$, where $S_a$ is an SDRS and $D$ is the set of dialogue agents. The logical form of the dialogue overall will be the logical forms of each of its turns (and all dialogue agents build all the SDRSs in the logical form, not just the SDRSs representing their own commitments). We assume an extremely simple notion of turns, where turn boundaries occur whenever the speaker changes (even if this happens mid-clause), and we ignore for now cases where agents speak simultaneously.

This new logical form for dialogue requires a new dynamic interpretation. The context $C_d$ of evaluation for interpreting a dialogue turn is a set of dynamic contexts for interpreting SDRSs—one for each agent $a \in D$:

$$C_d = \{ (C_a^i, C_a^o) : a \in D \}$$

Thus $C_a^i$ and $C_a^o$ are world assignment pairs, given the definitions from Asher and Lascarides (2003). For instance, (4) defines the dynamic interpretation of veridical relations (e.g., *Narration*, *Explanation*), where meaning postulates then stipulate the illocutionary effects $\varphi_{R(\alpha,\beta)}$—e.g., for *Narration* they stipulate the spatio-temporal progression of the events (we gloss the content that’s labelled $\pi$ as $K_\pi$, and $m$ in $[.]_m$ stands for monologue). Equation (5) defines the dynamic interpretation of *Correction*.

1. $(w, f)[R(\alpha, \beta)]_m(w', g)$ if\n   $$\langle w, f \rangle [K_\alpha \land K_\beta \land \varphi_{R(\alpha, \beta)}]_m(w', g)$$
2. $(w, f)[\text{Correction}(\alpha, \beta)]_m(w', g)$ if\n   $$\langle w, f \rangle [\neg K_\alpha \land K_\beta \land \varphi_{\text{Correction}(\alpha, \beta)}]_m(w', g)$$

The context change potential (CCP) of a dialogue turn $T = \{S_a : a \in D\}$ is the product of the CCPs of the individual SDRSs:

$$C_d[T]_dC_d' \text{ iff } C_d' = \{ (C_a^i, C_a^o) : (C_a^i, C_a^o) \in C_d, a \in D \}$$

Accordingly, dialogue entailments can be defined in terms of the entailment relation $\models_m$ for SDRSs afforded by $[.]_m$:

$$T \models_d \phi \text{ iff } \forall a \in D, S_a \models_m \phi$$

This makes $\models_d$ the shared entailment of each agent’s public commitments. And we assume that content $\phi$ is grounded or agreed upon by a dialogue turn $T$ iff $T \models_d \phi$. Finally, given that the SDRSs for a dialogue turn reflect all an agent’s current commitments, the interpretation of the dialogue overall is the CCP of its last turn.

The logical form of (3) is shown in Table 1 (we have omitted the logical forms of the clauses, labelled $\pi_1$ to $\pi_4$). The semantics of the SDRSs for the last turn correctly predict the following proposition to be grounded (for it is entailed by them): John went to Harrods, followed by buying a suit (at Harrods), followed by his entering the food halls.

There is a sharing of labels across the SDRSs in Table 1. This general feature reflects the reality that one speaker may perform a relational speech act whose first argument is part of someone else’s turn,
or part of his own previous turns. Sharing labels captures the intuition that an agent’s speech acts can reveal his commitments (or lack of them) to contextual content, even if this is linguistically implicit.

Including prior but ongoing commitments in the SDRS for the current turn has consequences for the general architecture of the theory: we must stipulate what commitments persist across turns when constructing the SDRSs. Consider the fourth turn of dialogue (3). Intuitively, uttering π₄ commits B to the illocutionary content of Narration(π₃, π₄). But in addition, he is also committed at this point to Narration(π₁, π₂) ∧ Narration(π₂, π₃), as shown. Those commitments persist from prior turns: they are even transferred from one speaker to another. However, we will shortly examine other examples, involving corrections and even explicit acknowledgements (or an acceptance in Traum’s (1994) terminology), where the commitments do not persist. To handle the data, we must make the ‘commitment persistence’ principle sensitive to distinct relational speech acts, and it must support a monotonic construction of logical form.

To motivate our persistence principle, consider how A and B get to the commitments shown in Table 1. A’s SDRS for the first turn is π₁ : Kπ₁, where Kπ₁ stands for the representation of John went to Harrods. Since B hasn’t said anything yet, his SDRS for the first turn is ∅. SDRT’s glue logic uses default axioms to predict the relation that connects π₂ to π₁ (Asher and Lascarides, 2003); here, these defaults should yield that B is committed to π₂B : Narration(π₁, π₂) (we adopt the convention that the root label of the speaker d’s SDRS for turn j is named πjd). A’s SDRS for the second turn is the same as the first turn: he hasn’t spoken since, and so his commitments are unchanged.

In the third turn, the glue logic should predict that A’s utterance π₃ forms a narrative with π₂. But simply adding this to A’s prior SDRS isn’t sufficient. First, the result is not a well-formed SDRS, because it won’t contain a single root label. Secondly, it misses an important interplay between discourse structure and grounding: adding only Narration(π₂, π₃) to A’s existing commitment to Kπ₁ makes A committed to the compositional semantics of π₂, but not to its illocutionary contribution conveyed by B (e.g. that John bought the suit at Harrods). And yet intuitively, uttering π₃ implicates that this (linguistically implicit) content is agreed on.

Dialogues (1) and (3) feature discourse relations that occur in monologue as well. Several agents can use these to build up a narrative together, as noted by Sacks (1992). Sacks’ observations affirm that such discourse relations can be used to perform ‘implicit’ acknowledgements, and what’s more they suggest that the implicit acknowledgement is not only of the prior contribution’s compositional semantics but also its illocutionary effects. These observations lead us to add the following Persistence principle to the glue logic, together with axioms that identify undenied commitments (UC(α) stands for the undenied commitments of the utterance or segment α):

- **Persistence:**
  \[
  \lambda : R(\alpha, \beta) \rightarrow \lambda : UC(\alpha)
  \]

Different glue-logic axioms will then identify the undenied commitments for different speech acts. The present case concerns simple left veridical (slv) relations—those that do not explicitly endorse or criticise any previous commitments. Note φ > ψ means “If φ then normally ψ”, and T(d, j, π) means that label π is a part of agent d’s SDRS for turn j:

- **Undenied Commitments:**
  \[
  (\lambda : R(\alpha, \beta) \land T(d_1, j, \lambda) \land slv(R) \land \\
  \lambda' : R'(\gamma, \alpha) \land T(d_2, j - 1, \lambda')) > \\
  (\lambda : UC(\alpha) \rightarrow \lambda : R'(\gamma, \alpha))
  \]
Undenied Commitments states that if \( d_1 \) commits to \( R(\alpha, \beta) \) where \( R \) is simple left veridical and \( d_2 \) is already committed to \( R'(\gamma, \alpha) \), then normally the undenied commitments of \( \alpha \) include \( R'(\gamma, \alpha) \). Examples of simple left veridical relations include Narration and Explanation but not Acknowledgement (since this explicitly endorses prior content) or Correction (since this denies prior content).

Persistence and Undenied Commitments predict that \( A \)'s SDRS for the third turn of (3) includes \( \pi_{3A} : \text{Narration}(\pi_1, \pi_2) \). This is because default rules yield \( \pi_{3A} : \text{Narration}(\pi_2, \pi_3) \), and \( \text{Narration}(\pi_1, \pi_2) \) is in \( B \)'s SDRS. Persistence and Undenied Commitments likewise predict that \( \text{Narration}(\pi_1, \pi_2) \) and \( \text{Narration}(\pi_2, \pi_3) \) are a part of \( B \)'s SDRS for the fourth turn, as shown in Table 1.

Undenied Commitments is defeasible. This is because if the illocutionary contribution of \( A \)'s (left-veridical) speech act \( R(\alpha, \beta) \) conflicts with some proposition \( p \) that \( B \) conveyed by uttering \( \beta \), then clearly \( A \)'s speech act should not be construed as an implicit acknowledgement of \( p \). This affects the analysis of (1), whose logical form is Table 2. \( B \)'s SDRS after the second turn does not include \( \text{Explanation}(\pi_1, \pi_2) \), even though his utterance \( \pi_3 \) attaches with the veridical relation Narration to \( \pi_2 \), and \( A \)'s SDRS for turn 1 includes \( \text{Explanation}(\pi_1, \pi_2) \). Persistence applies to this example (for label \( \pi_2 \)) and the antecedent to Undenied Commitments is satisfied, but \( \text{Explanation}(\pi_1, \pi_2) \) is not an undenied commitment of \( \pi_2 \) because its (nonmonotonic) semantic consequences conflict with those of \( \text{Explanation}(\pi_1, \pi) \), a speech act that the logic glue must identify as one that \( B \) intended to perform (or, in other words, publicly commit to) as a byproduct of uttering \( \pi_3 \). Explanation(\( \pi_1, \pi_2 \)) conflicts with Explanation(\( \pi_1, \pi \)) because the former nonmonotonically entails, via a scalar implicature, that Max turning out the light was the sole cause of the room going dark, while the latter (monotonically) entails it was a strict part of it. This example illustrates how the default logic rendered by > must be specified in terms of the consistency in what follows nonmonotonically, rather than what follows monotonically.

Undenied Commitments does not apply for the veridical relation Acknowledgement; i.e., utterances of the form OK, I agree, repeating prior content, and the like. In words, Acknowledgement(\( \pi_1, \pi_2 \)) entails \( K_{\pi_1}, K_{\pi_2} \) and that \( K_{\pi_2} \) implies \( K_{\pi_1} \); to use the GAM term, it is an act of explicit acceptance. Dialogue (6) illustrates why Acknowledgement behaves differently from the simple left veridical relations like Narration:

(6) \( \pi_1 \). B: John is not a good speaker
\( \pi_2 \). B: because he’s hard to understand.
\( \pi_3 \). A: I agree he’s hard to understand.

The compositional semantics of \( \pi_3 \) makes \( A \) explicit about what in \( B \)'s turn he acknowledges: \( A \) must be committed to (at least) Acknowledgement(\( \pi_2, \pi_3 \)). What is outside the scope of the acknowledgement—namely, \( B \)'s putative explanation for why John is not a good speaker—is not denied in (6). It would be consistent to add Explanation(\( \pi_1, \pi_2 \)) to \( A \)'s commitments, but it’s simply not warranted. Dialogue (6) shows that when the explicit endorsement conveys sufficiently specific content, it appears to carry a scalar implicature that this precise content is endorsed, and no more.

Another reason for excluding explicit acknowledgements from the set of simple left veridical relations is that such speech acts come with their own grounding requirements. Acknowledgements can have scope over implicatures as well as compositional semantic contents, since the first argument to an Acknowledgement relation can be a label of an arbitrarily complex SDRS. So by acknowledging \( \pi_j \), we do not thereby acknowledge the implicatures of \( \pi_j \) itself; had we wished to do so, we would have included them within the scope of the acknowledgement. That is, we would infer the relation Acknowledgement(\( \pi_j, \pi_i \)), where \( \pi_j \) has semantic scope over \( \pi_j \), making \( \pi_j \) and the rhetorical relations it engages in part of what is (explicitly) endorsed. It is because the discourse function of an acknowledgement is precisely to say what one agent commits to from another agent’s turn—i.e., what are the undenied commitments in this case—that Persistence applies redundantly.

Explicit acknowledgements have been studied by Traum and Hinkelman (1992), among others. Here, we will ignore interpretations of an utterance \( \pi_2 \) (e.g., OK) as an acknowledgement that \( K_{\pi_1} \).
Table 2: The logical form of (1).

| Turn | A’s SDRS                              | B’s SDRS                              |
|------|---------------------------------------|---------------------------------------|
| 1    | $\pi_{1A} : Explanation(\pi_1, \pi_2)$ | $\emptyset$                           |
| 2    | $\pi_{1A} : Explanation(\pi_1, \pi_2)$ | $\pi_{2B} : Explanation(\pi_1, \pi)$  |
|      |                                       | $\pi : Narration(\pi_2, \pi_3)$       |

was said (represented in SDRT with the so-called metatalk relation $\text{Acknowledgement}^*(\pi_1, \pi_2)$), instead focussing entirely on an interpretation of $\pi_2$ using $\text{Acknowledgement}$ (i.e., a commitment to $K_{\pi_1}$, which in turn entails a commitment that $K_{\pi_1}$ was said). But even so there is ambiguity, because linguistic form does not always fully determine what the acknowledgement has scope over. Let’s assume that A’s utterance $\pi_3$ in (7) is an acknowledgement of content and not just of understanding that content:

(7) $\pi_1$. B: John is not a good speaker
$\pi_2$. B: because he is hard to understand.
$\pi_3$. A: OK.

$\text{Acknowledgement}(\pi_2, \pi_3)$ entails $K_{\pi_2}$. Making $\pi_2$ the only label that’s acknowledged leads to an interpretation where the proposition that $\pi_2$ explains $\pi_1$ is not acknowledged. This ‘narrow scope’ attachment permits A to continue by challenging the explanatory link, e.g., by uttering but that’s not why he’s not a good speaker. Another interpretation of (7) is that A commits to all of B’s commitments, including the implicatures: this is expressed by adding $\text{Acknowledgement}(\pi_{1B}, \pi_3)$ to A’s SDRS, where $\pi_{1B} : Explanation(\pi_1, \pi_2)$. Indeed, if OK is all that A says, then one defaults to this wide-scope interpretation. Even if A follows OK with He is hard to understand with high pitch accents and a falling boundary tone, the preferred interpretation contrasts with (6), to be one where OK is an $\text{Acknowledgement}$ of $\pi_{1B}$, and He’s hard to understand is an explanation of that acknowledgement act (marked with the metatalk relation $\text{Explanation}^*$ in SDRT). It is straightforward to add glue-logic axioms for constructing logical form that reflect these principles for identifying the first argument of $\text{Acknowledgement}$.

In dialogue (2), A commits to the negation of his prior commitment. As before, constructing B’s SDRS for the second turn involves using the glue logic to identify how $\pi_2$ connects to $\pi_1$. So long as their semantic incompatibility is transferred, in shallow form, to the glue logic, then the general principle that the necessary semantic consequences of a speech act are normally sufficient for inferring that it was performed will apply, yielding $\pi_{2B} : \text{Correction}(\pi_1, \pi_2)$ (see Table 3). The cue phrase OK is then used by the glue logic to infer $\pi_{3A} : \text{Acknowledgement}(\pi_2, \pi_3)$. This resolves the underspecified content OK to $K_{\pi_2}$; and thus as before the glue logic also yields $\pi_{3A} : \text{Correction}(\pi_1, \pi_3)$, as shown. It’s not raining is entailed by the SDRSs for turn 3. The interpretation of each turn is consistent (i.e., the output state is non-empty), although the SDRSs for turn 2 are mutually inconsistent (A’s SDRS entails that it’s raining and B’s entails it’s not). Finally, the content associated with each label does not change from one turn to the next, making the construction of logical form monotonic.

Clark (1996) doesn’t make precise exactly what counts as sufficient positive evidence for grounding. Similarly, Traum and Allen (1994) don’t provide rules for inferring when a speaker has performed an implicit acceptance. Our framework makes the quantity of positive evidence that’s needed for grounding propositions logically precise, in terms of the relational speech acts that both speakers perform, and the logical relationships between the semantics of those speech acts. Persistence and Undenied Commitments capture a general class of examples involving implicit agreement. Sufficient positive evidence for grounding a proposition through explicit endorsements and challenges rests on the formal semantic interpretation of the relevant speech acts—namely $\text{Acknowledgement}$ and $\text{Correction}$—and the rules by which one determines the first argument of these relations.

4 Conclusion

We have presented a novel treatment of agreements and disputes in which the construction of logical form is monotonic in the subsumptive sense.
(Shieber, 1986); the semantic representation of the discourse context is an elementary substructure of the representation of the dialogue updated with the current utterance, even if the current utterance denies earlier content. However, the logical form remains a product of complex default reasoning, since identifying the speech acts that were performed involves commonsense reasoning with the linguistic and non-linguistic context.

The relationship between the grounded propositions and the interpretation of the dialogue is entirely transparent and is defined in terms of the model theory of the logical forms. It provides a logical basis for exploring Clark’s (1996) notion of positive evidence for grounding. A crucial ingredient in our account was the use of relational speech acts, and the logical relationships among their semantics.

We believe our definition of grounding as shared commitment is capable of modelling Clark’s more central concern—grounding the understanding of what was said. The left-veridical relations that are the hallmark of grounding at level 4 entail grounding at the lower levels thanks to the semantics of DSDRSs. Moreover, SDRT’s metatalk relations—such as Explanation∗(α, β) and Acknowledgement∗(α, β)—commit an agent to the fact that Kα was said without committing him Kα. Thus shared commitments that follow from a representation of the dialogue can ground acts at lower levels without grounding (or denying) acts at level 4. A full model of grounding at lower levels, however, requires us to extend the framework to handle misunderstandings.

This paper presents just some first steps towards a dynamic theory of grounding. For instance, we have not yet modelled the impact of questions and imperatives on public commitments and grounding. We have started to explore links between public commitments and other attitudes, such as beliefs, preferences, and intentions (Asher and Lascarides, 2008), but this also remains a matter of ongoing research.

### References

H. Alshawi and R. Crouch. Monotonic semantic interpretation. In *Proceedings of ACL*, pages 32–39, 1992.

N. Asher and A. Lascarides. *Logics of Conversation*. CUP, 2003.

N. Asher and A. Lascarides. Commitments, beliefs and intentions in dialogue. In *Proceedings of Lendial*, 2008.

H. H. Clark. *Using Language*. CUP, 1996.

B. Grosz and C. Sidner. Plans for discourse. In J. Morgan P. R. Cohen and M. Pollack, editors, *Intentions in Communication*, pages 365–388. MIT Press, 1990.

J. R. Hobbs, M. Stickel, D. Appelt, and P. Martin. Interpretation as abduction. *Artificial Intelligence*, 63(1–2):69–142, 1993.

D. Lewis. *Convention: A Philosophical Study*. Harvard University Press, 1969.

L. Prévot, N. Maudet, and P. Muller. Conversational game-board and discourse structure. In *Proceedings of Constraints in Discourse*, Ireland, 2006.

H. Sacks. *Lectures on Conversation*. Blackwells, 1992.

S. Shieber. *An Introduction to Unification-based Approaches to Grammar*. CSLI Publications, 1986.

D. Traum. *A Computational Theory of Grounding in Natural Language Conversation*. PhD thesis, University of Rochester, 1994.

D. Traum and J. Allen. Discourse obligations in dialogue processing. In *Proceedings of ACL*, pages 1–8, 1994.

D. Traum and E. Hinkelman. Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, 8(3):575–599, 1992.