OK or not OK? Commitments in acknowledgments and corrections

Antoine Venant, Nicholas Asher

OK or not OK? Commitments in acknowledgments and corrections *

Antoine Venant
Université Toulouse 3/IRIT

Nicholas Asher
CNRS/IRIT

Abstract When saying something in a conversation, an agent publicly commits to some interpretation of what she said, which others might dispute. But how does making a commitment affects the commitments of others? We provide an answer to this question for the case of acknowledgments. Commitment semantics poses a dilemma for acknowledgments: either the semantics is strong, entailing that agent i's committing to φ entails other agents commit to i's committing to φ. This makes acknowledgments vacuous. Or further acknowledgments are required to produce commitments about others' commitments, but then grounding seems never achievable in finite time. Building on Venant & Asher (2015) we provide a model of these strong and weak semantics where we show how the two are linked *via* a semantics for inductive, synchronous acknowledgments.

Keywords: dialogue, semantics, commitments, acknowledgments, grounding, dynamics

1 Introduction

Hamblin (1987); Traum & Allen (1994); Traum (1994); Asher & Fernando (1997); Lascarides & Asher (2009) inter alia argue that while a semantics without differing "points of view" of different agents is a good first hypothesis for the analysis of the content of monologue, dialogues typically involve differing points of view from different agents. In particular one agent may not agree with what another agent asserts, or may have a different interpretation of an utterance from that of its author. An adequate semantics for dialogue should proceed by attributing to different dialogue agents separate views of the contents of their conversation. We model this, following others, by assigning each agent her own commitment slate. In this paper we bring out a complication with this approach that has gone so far unnoticed in formal semantics and the prior work we just mentioned, although it is well-known from epistemic game theory: commitment slates interact; agents typically commit to the fact that other agents make certain commitments. We thus formulate the semantics of dialogue moves and conversational goals in terms of nested, public commitments.

What is central in putting this notion of meaning to work is the dynamics of such commitments. Working out this dynamics is the central contribution of the paper. We develop two

semantics for nested commitments, one for a simple propositional language, the other for a full description language for discourse structures of dialogues; we show how one is an approximation of the other. We apply our dynamics to three different sorts of issues: the problem of ambiguity, the semantics of acknowledgments and the semantics of corrections. Before delving into technical issues, we first sketch the three issues our semantics will address.

2   Three phenomena requiring a dynamic semantics of nested commitments

Ambiguity A particular discourse move *m* typically presupposes a particular commitment on the part of *m*'s agent concerning the commitments that *other* agents have made on a prior move *n*. This commitment may not be what the the author of *n* intended for innocuous or strategic reasons. Here is an example (from Venant, Asher & Degremont 2014).

(1)   a. C: N. isn't coming to the meeting. It's been cancelled.

        b. A: That's not why N. isn't coming. He's sick.

        c. C: I didn't say that N. wasn't coming because the meeting was cancelled. The meeting is cancelled because N. isn't coming.

C's initial contribution contains a discourse ambiguity. A has taken C to be committed to one of its possible disambiguations when C turns out to have committed to the other. But A is not unreasonable to take C to be committed to what he takes him to commit to, and the ensuing exchange is about who was committed to what. To represent its meaning, we need nested commitments combined with a means for representing potentially ambiguous discourse moves.

Acknowledgments and corrections For many researchers, including Clark (1996); Traum (1994); Traum & Allen (1994) inter alia, an acknowledgment as in (2c) by 0 of a discourse move *m* by 1 can signal that 0 has understood what 1 has said, or that 0 has committed that 1 has committed to a content *p* with *m*, and serve to "ground" or to establish a mutual belief that 1 has committed to *p*. For Clark, grounding by the other conversational participants is a necessary condition for the content of utterances to enter the common ground. Corrections, and self-corrections, as in (2d), on the other hand, serve to remove commitments.

(2)   a. 0: Did you have a bank account in this bank?

        b. 1: No sir.

        c. 0: OK. So you're saying that you did not have a bank account at Credit Suisse?

        d. 1: No. sorry, in fact, I had an account there.

        e. 0: OK thank you.

The problem is that grounding doesn't follow just from the simple gloss above. With Traum and Allen, we argue that grounding that *i* committed to *p* should require reaching a state

where a **common commitment** over some content K holds (e.g., typically over the fact that i performed m and the content of m is p). A common commitment by a group of agents G means that every agent in G is committed to K and to the fact that every agent is committed to K, and to the fact that every agent is committed that every agent is committed to K, and so on. But if an acknowledgment performed by agent i only brings additional levels of commitments by i to some given content, it is far from straightforward to see how and why grounding would be possible at all in finite time. Clark indeed mentions such a problem and proposes that the solution lies in a continuous exchange of instantaneous and concurrent, unspoken signals. Others like Lascarides & Asher (2009) assume implicit acknowledgments in the absence of explicit corrections, which might provide another possible solution, for cooperative conversations at least. But there has been no logically precise semantics of acknowledgments that logically entail common commitments and grounding. We provide this.

To get a clearer picture of the problem for acknowledgments, we need some assumptions. We take public commitment to be an operator with a weak modal logic (K); an agent commits to a proposition φ ($C_i\varphi$) given a discourse move m, when m entails φ or when i asserts φ. In general commitments do not validate type 4 axioms of modal logics; saying φ is not the same as saying *I commit to* φ, and asserting φ does not entail asserting *I assert that* φ. Analogously to common knowledge, we define common commitments for a group G, $C_G^*\varphi$, as $C_G\varphi \land C_G C_G\varphi \land \dots C_G(C_G)^n\varphi \land \dots$. Common commitments would follow from assuming a perfect communication channel and a view of semantic competence implying perfect knowledge of speaker commitments of unambiguous discourse moves.

These assumptions lead to a very strong view of assertions and other discourse moves. If discourse move m entails p, then i's making m entails $C_G^*C_i p$. But then grounding acknowledgments are semantically superfluous. The only informative contribution of i's acknowledgment of j's move then is that i agrees with the content of j's move. Nevertheless, we can imagine 1 in (2c) acknowledging 0's response even if he patently does not believe its content. Such acknowledgments are often present in legal questioning but in many other conversations too.

We might opt for a much simpler semantics for discourse moves. In particular, (a) an assertion of φ by speaker i only entails that $C_i\varphi$. A similarly weak semantics for acknowledgments would mean that (b) an acknowledgment by j of the content of an assertion by i entails $C_j C_i\varphi$. But this makes grounding impossible in finite conversations: if a discourse move m by i entails only $C_i p$, (a) and (b) entail that all the conversational participants believe $C_i p$ (Traum & Allen 1994; Ginzburg 2012). Then j's acknowledgment of m would entail $C_j C_i p \land Bel_G C_j C_i p$. and $Bel_G C$. Let conversations be strings of discourse moves V^* by players i and j that may be finite or infinite (Asher, Paul & Venant 2014). That is every conversation c is such that $c \in V^\infty$ where $V^\infty = V^* \cup V^\omega$. It follows by induction that

Proposition 1 *A conversational sequence σ of assertions and acknowledgments verifies a common commitment to the fact that i commits to φ, $C^*C_i\varphi$ only if $\sigma \in V^\omega$.*

That is, common commitments are achieved only after an infinite sequence of acknowledgment moves between i and j.

Some other options are logically possible. For instance, we could keep the simple semantics for discourse moves, but assign acknowledgments a very strong semantics. On such a semantics, j's acknowledging a discourse move by i that entails $C_i\varphi$ would imply a common commitment by i and j to $C_i\varphi$. But this implausibly imputes to j the ability to force commitments on i that quickly leads to absurdities. If i says, for instance, *I don't want to go to the meeting*, j can "acknowledge" i's move by saying, *OK, thank you very much for agreeing to go to the meeting*, thus forcing a common commitment to C_i *i goes to meeting*. But clearly i didn't commit to going to the meeting.

Can we do without common commitments? We think not. Common commitments are essential (see also Clark 1996) for strategic reasons and can be present even when mutual beliefs about a shared task are not. Suppose, for instance, that i's goal is that $C_j\varphi$ and that j cannot consistently deny the commitment. If i only extracts from j a move m implying $C_j\varphi$, j has a winning strategy for denying i victory. She simply denies committing to φ (*I never said that*), since $C_j\neg C_j\varphi$ is consistent with $C_j\varphi$, even if $Bel_j C_j\varphi$. Player j lies, but she is consistent. If i manages to achieve $C_j C_j\varphi$, j can still similarly counter i's goal while maintaining consistency. j can assert something to the effect that $C_j\neg C_J C_j\varphi$, which means that there are worlds compatible with her commitments where $\neg\varphi$. However, if i achieves the *common commitment* $C_G^* C_j\varphi$, with G the group of conversational participants, j does not have a way of denying her commitment without becoming inconsistent, as $C^* C_j\varphi \rightarrow (C_j C_j\varphi \wedge C_j C_j C_j\varphi \wedge \cdots)$, for any finite depth of nesting of C_j operators. And only common commitments rule out other, more elaborate ways of defeating conversational goals. For instance, on a weaker semantics j could deny that i had committed to what j had committed to at some level of embedding. Thus, if i's goal depends on j's committing to one of i's commitments, a lack of common commitment will allow j to deny i's achieving her conversational goals.

We could have formulated the problem of grounding and common commitments in doxastic terms: j's acknowledgment of a move by i should lead to a mutual belief by i and j in some content. Problems for the semantics of assertion and acknowledgments analogous to those we have just sketched in terms of commitments will surface for an account of grounding in terms of mutual belief. However, we believe that a semantics for dialogue in terms of commitments is preferable for several reasons. First, commitments are public and directly verifiable by the dialogue participants; beliefs on the other hand are private and hidden. In addition and as a consequence, the link between the content of an utterance and commitments is in principle clear and deductive, whereas the link from the content of an utterance to a belief about that content only follows under strong assumptions like Gricean Sincerity that cannot be assumed to hold of conversational participants in general. People lie and exaggerate for various reasons, which makes the link between contents of utterances and belief uncertain. The link between contents and commitments is not subject to such uncertainty.

We now briefly turn to the semantics of self-corrections. While Lascarides & Asher (2009)

give a general semantics for corrections in terms of simple commitments, according to which one speaker commits to the negation of what another speaker committed to with a prior move, they do not look at self-corrections, which we focus on here. In self-corrections, speakers can not only deny prior commitments but also "undo" or "erase" them with *self-corrections*. For instance, if in (2b) 1 commits to not having a bank account; in (2d) 1 no longer has this commitment. Self-corrections thus entail a revision of commitments. No one has proposed a logical analysis of self-corrections. We provide such an analysis, showing that these moves have an essential, strategic role to play in dialogue, even if we assume a perfect communication channel and unambiguous commitments in dialogue moves.

3   Simple semantics for nested commitments

As a first step toward a full dynamic semantics for commitments, we consider the simple setting of Venant & Asher (2015) on which the representation of a conversational agent's contributions is restricted to a propositional language with a commitment modality for each agent. We then discuss and refine their semantics of acknowledgments, providing a more systematic understanding of the problems sketched in section 2.

Assume a set of agents I. Agents can perform the following kinds of dialogue moves: uttering a proposition φ, which we will write as φ!, acknowledge a previous action α, which we will write as $ack(\alpha)$, or perform an *ambiguous* move $\alpha \sim \beta$, which represents a discourse move whose performance produces an effect ambiguous between those of actions α and β. Ambiguity will play an important role in our formal model of grounding and this final construction.

The dynamic language $\mathcal{L}_D$ is a classical propositional language closed under $\sim$, and the operators ack and C_i. The semantics of $\mathcal{L}_D$ exploits the following standard definitions: a *frame* is a tuple $\langle W, (R_i)_{i \in X} \rangle$ with W a set of so-called *possible worlds* and for each $i \in I$, R_i is a binary relation over W. A *model* $\mathcal{M}$ is a pair $\langle \mathcal{F}, v \rangle$ with $\mathcal{F}$ a Kripke frame and $v : W \mapsto \wp(\text{PROP})$ an assignment at each world w of propositional variables true at w.

The semantics interprets actions of $\mathcal{L}_D$ as action-structures of the logic of public and private announcements (Baltag, Moss & Solecki 1998). The truth of static formulae φ relative to a pointed model ($\langle \mathcal{M}, w_0 \rangle \vDash \varphi$) is classical, with C_i a normal K-modality ($\langle \mathcal{M}, w_0 \rangle \vDash C_i \varphi$ iff φ holds in every world accessible via R_i from w_0). The definition of action-structures needed for the semantics of dynamic formulae $[\alpha^i]\varphi$ is as follows:

Definition 1 (Action-structures (Baltag et al. 1998)) An action-structure is a tuple $\langle \mathcal{F}, k, pre \rangle$ where $\mathcal{F}$ is a frame, $k \in \mathcal{F}$ and $pre : W^{\mathcal{F}} \mapsto \mathcal{L}_D$ associates to each world in $\mathcal{F}$ a formula, called the *precondition* of this world.

An action-structure defines how an action affects each agent's commitments, including his commitment over the action's effect on others' commitments. The following definition of

a model update does this. In informal terms, for each world k of the action structure, the update makes a new copy W^k of the model's set of worlds, deletes from this copy every world failing the precondition $pre(k)$, and finally makes worlds from different copies W^k and $W^{k'}$ accessible to one another if and only if the action structure has k accessible from k' and the model to be updated has corresponding worlds accessible to one another.

Definition 2 (Model update) Let $K = \langle \mathcal{F}, k_0, pre \rangle$ be an action structure. Let $\langle \mathcal{M}, w_0 \rangle$ be a pointed model. Let $|\varphi|^{\mathcal{M}} = \{ w \in \mathcal{M} \mid \mathcal{M}, w \vDash \varphi \}$. If $w_0 \notin pre(k_0)$ the update $\langle \mathcal{M}, w_0 \rangle \star \langle K, k \rangle$ *fails*. Otherwise, the updated model, $\langle \mathcal{M}^K, (w_0, k) \rangle$, is defined as $\langle \mathcal{M}^K, (w_0, k) \rangle$ with

$$W^K = \bigcup_{l \in K} |pre(l)|^{\mathcal{M}} \times l$$

as the set of worlds, accessibility relations are defined as $R_i^{\mathcal{M}^K}((w,l),(w',l'))$ iff (i) $R_i^{\mathcal{M}}(w,w')$ and (ii) $R_i^K(l,l')$, and valuations left unchanged i.e., $v((w,l)) = v(w)$.

Since we want to model changes in commitments, the actions we use only affect agents' commitments, and never affect actual facts. In other words, we always have $pre(k_0) = \top)$. Therefore actions never fail, and updated models always exist.

We interpret pairs of actions α and agents i, α^i as action-structures $[\![\alpha^i]\!]$. Then, the truth of a formula $[\alpha^i]\varphi$ isdefined as:

$$\langle \mathcal{M}, w \rangle \vDash [\alpha^i]\varphi \text{ iff } \langle \mathcal{M}, w \rangle \star [\![\alpha^i]\!] \vDash \varphi$$

Venant & Asher 2015 provides two different ways of interpreting declarative actions $\varphi!$ as action structures, yielding two distinct dynamics of commitments. The first, the *strong* semantics for assertions, systematically produces a common commitment by each agent on $C_i\varphi$ upon i's performing φ. Under this interpretation of $\varphi!^i$, acknowledgments are easily shown to be semantically superfluous. A second interpretation, the "*weak*" semantics, only ensures that $C_i\varphi$ holds after i has performed $\varphi!$; it leaves higher-order commitments as well as other agents' commitments unchanged. This semantics makes common commitments unattainable by any finite sequence of acknowledgments.
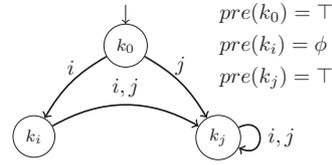
We depict the action structures for the weak and strong interpretation of assertions $[\![\varphi!^i]\!]$ graphically; each circular node represents a world k of the action structure; a directed edge leading from node k to node k' means that $R_i(k,k')$:

Definition 3 (Strong and weak semantics for assertions)

$\llbracket\varphi!^i\rrbracket$ in the strong semantics $\llbracket\varphi!^i\rrbracket$ in the weak semantics

$pre(k_0) = \top$
$pre(k_i) = \phi$
$pre(k_j) = \top$
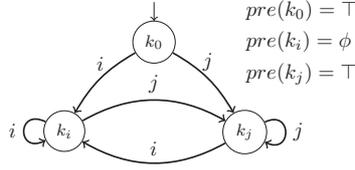
This propositional setting naturally reflects our intuitions for acknowledgments: i's performing $ack(\alpha^j)$ should introduce a commitment by i over the fact that action α has been performed by j, in other words, it should make i apply α^j on i's set of commitments. The following semantics for $ack(\alpha^j)^i$, together with the preceding definitions, yield such a desired behavior:

Definition 4 (Semantics for acknowledgments) Let $\alpha \in \mathcal{A}$ be a linguistic action. Let $\langle K^\alpha, k_0^\alpha, pre^\alpha \rangle = \llbracket\alpha^j\rrbracket$. Let k_0 and k_{-i} be "fresh" symbols not appearing in K^α.

$$\llbracket ack(\alpha^j)^i \rrbracket = \langle \{k_0, k_{-i}\} \cup K_\alpha, k_0, pre \rangle$$

Accessibility relations are set to: $R_i(k_0, k_0^\alpha)$, for $x \neq i$, $R_x(k_0, k_{-i})$, $\forall x \in I\ R_x(k_{-i}, k_{-i})$, $\forall k, k' \in K^\alpha, \forall x \in I\ R_x(k, k')$ iff $R_x^{K^\alpha}(k, k')$ with no other transitions. $pre(k_0) = pre(k_j) = \top$ and pre coincide with pre^α on K^α.

Thus, $\llbracket ack(\alpha^j)^i \rrbracket$ is an action structure whose actual world is such that i is committed to a copy of $\llbracket\alpha^j\rrbracket$ while any other agent x is committed to a world with no precondition (and thus x's commitments won't be affected when updating α^i). To give a more concrete view on acknowledgments we provide in addition on figure 1 the graphical representations of $ack(\varphi!^j)^i$ in the strong and weak semantics for utterances. In both cases, the recursive mechanism for interpreting acknowledgments stays the same; only the interpretation of assertions it combines with varies.

Going beyond Venant & Asher 2015, we now provide a link between the strong and weak semantics, on which the strong semantics corresponds to an idealization of an infinite sequence of acknowledgments. We first define such sequences:

Definition 5 (Iterated acknowledgments) Let $\alpha^x \in A \times I$ be an $\langle$action, agent$\rangle$ pair, let $G \subseteq I$ be a group of agents. Define inductively the set of $\langle$action, agent$\rangle$ pairs $\uparrow_G^n (\alpha^x)$ as follows:

- $\uparrow_G^0 (\alpha^x) = \{\alpha^x\}$
- $\uparrow_G^{n+1} (\alpha^x) = \bigcup_{i \in G} \bigcup_{\beta \in \uparrow_G^n} ack(\beta)^i.$
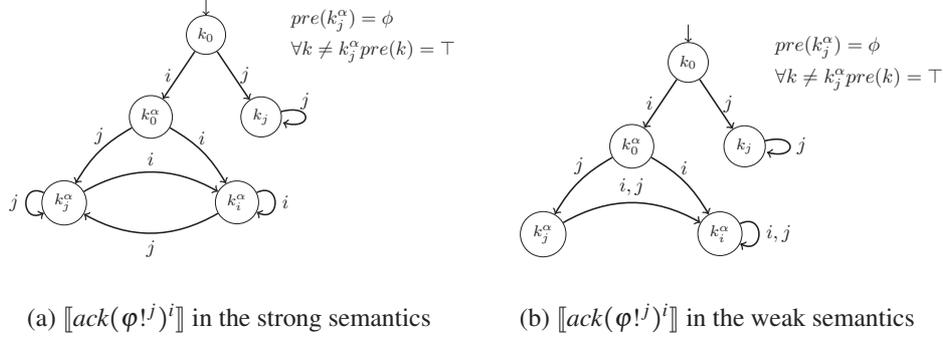
(a) $[\![ack(\varphi!^j)^i]\!]$ in the strong semantics

(b) $[\![ack(\varphi!^j)^i]\!]$ in the weak semantics

Figure 1   Acknowledgments in the strong and weak semantics

Level n in the above hierarchy consists of every possible acknowledgment by every agent in G of actions at the lower level. Note that $|\uparrow_G^n(\alpha^x)| = |G|^n$. An easy induction shows that at each level n, the order in which actions in $\uparrow_G^n(\alpha^x)$ are applied to a model $\mathcal{M}$ do not change the final result; for any complete orderings $\beta_0, \beta_1 \ldots \beta_{|G|^n}$ and $\beta_0', \beta_1', \ldots \beta_{|G|^n}'$ of $\uparrow_G^n(\alpha^x)$, and for every model $\mathcal{M}$, $(\ldots(((\mathcal{M},w) \star \beta_0) \star \beta_1) \star \ldots) \star \beta_{|G|^n})$ and $(\ldots(((\mathcal{M},w) \star \beta_0') \star \beta_1') \star \ldots) \star \beta_{|G|^n}')$ are isomorphic. Our choice of ordering of $\uparrow_G^n(\alpha^x)$ will thus not affect our results.

Let $<_I$ be any total order on the set of agents I. We inductively extend $<_I$ to $\uparrow_G^{n+1}(\alpha^x)$, by defining $ack(\beta)^i <_I ack(\beta')^j$ iff $i <_I j$ or $i = j$ and $\beta <_I \beta'$ (thus taking a lexicographic ordering at each step). We write $\uparrow_G^{n+1}(\alpha^x)_k$ as the kth action in $\uparrow_G^{n+1}(\alpha^x)$ according to $<_I$. We can now define the infinite sequence of actions $!_G^\omega(\alpha^x) = \langle \uparrow_G^0(\alpha^x)_1, \uparrow_G^1(\alpha^x)_1, \uparrow_G^1(\alpha^x)_2, \ldots, \uparrow_G^n(\alpha^x)_1, \ldots, \uparrow_G^n(\alpha^x)_{|G|^n}, \ldots \rangle$ and the infinite sequence of actions $ack_G^\omega(\alpha^x) = \langle !_G^\omega(\alpha^x) \rangle_{n\in\omega}$. We have defined here two infinite sequences of actions: $!_G^\omega(\alpha^x)$ corresponds to the sequence starting with action α executed by x followed by its acknowledgment by every agent in G, again followed by acknowledgments of these acknowledgments by every agent in G and so on, while $ack_G^\omega(\alpha^x)$ corresponds to the same sequence of acknowledgments without the initial performance of α by x.

The last ingredient needed to complete the correspondence between the strong and weak semantics of acknowledgments is a semantics for such infinite sequences:

Assume in what follows that for any $G \subseteq I$, C_G^* is **not** part of the language $\mathcal{L}_D$[1].

Definition 6 Let $\mathcal{U}$ be a pointed model and $\langle \mathcal{U}_n \rangle_{n\in\omega}$ be an infinite sequence of models. $\mathcal{U}$ is a *limit model* of $\langle \mathcal{U}_n \rangle_{n\in\omega}$ iff $\forall k \in \omega \exists n_0 \in \omega \forall n \geq n_0\ \mathcal{U}$ is k-bisimilar to $\mathcal{U}_n$. Notice that existence of a limit model requires $\forall k \exists n_0 \forall n > n_0\ \mathcal{U}_n$ is k-bisimilar to $\mathcal{U}_{n-1}$. Let $\equiv_\omega$ denote k-bisimilarity for every $k \in \omega$. By transitivity of finite bisimulation, two limit models are $\equiv_\omega$

---

1 which does not prevent us to study whether actions might yield a model where it holds

equivalent, and a $\equiv_\omega$ equivalent to a limit model is a limit model.

Definition 7 Let $\langle \alpha_k^{x_k} \rangle_{k \in \omega}$ be an infinite sequence of $\langle$action, agent$\rangle$ pairs. Let $\varphi \in \mathcal{L}_D \cup C_G^* \varphi \mid G \subseteq I$, and $\langle \mathcal{M}, w \rangle$ be a pointed model. Define inductively $\mathcal{U}_0 = \langle \mathcal{M}, w \rangle$ and $\mathcal{U}_{k+1} = (\mathcal{U}_k \star [\![\alpha_k]\!]^{x_k})$. The effects of an infinite sequence of actions are defined as follows:

- If $\langle \mathcal{U}_k \rangle_{k \in \omega}$ has no limit model, the effects are undefined.
- Otherwise $\langle \mathcal{M}, w \rangle \models [\langle \alpha_k^{x_k} \rangle_{k \in \omega}] \varphi$ iff $\forall \mathcal{U}$ if $\mathcal{U}$ is a limit model of $\langle \mathcal{U}_k \rangle_{k \in \omega}$ then $\mathcal{U} \models \varphi$

The following proposition links the strong and the weak semantics:

Proposition 2 *For an action α, the effects of $ack^\omega(\alpha)$ are always defined, moreover, let $\langle \mathcal{M}, w \rangle$ be a pointed model, $\langle \mathcal{M}, w \rangle \models [\varphi!^x] \psi$ in the strong semantics for utterances iff $\langle \mathcal{M}, w \rangle \models [!_I^\omega(\varphi!^x)] \psi$ in the weak semantics.*

We now have a formally precise statement of the intuition that the strong interpretation of $\varphi!^x$ is logically equivalent to $[\varphi!^x][ack_I^\omega(\varphi!^x)]$, i.e., that the strong semantics logically behaves as if utterances were always followed by every possible acknowledgment. It also shows something non-trivial: such infinite iterations of acknowledgments are regular enough to be finitely representable as an update of the model. This means that we can nest such constructions without further effort. Following the same kind of reasoning we can then define the action $ack^i(!_I^\omega(\varphi!^x))$, to represent i's acknowledgment of an infinite sequence of acknowledgment yielding a common commitment over $C_x \varphi$, and shows that this action does the same as the "finite" one $ack^i(\varphi!^x)$ under the strong semantics. Our analysis regarding infinite constructions supports our formal treatment of grounding.

Existing theories, without proposing a fully worked out semantics, solve the problem of grounding by assuming that, at least by default, infinitely many weak acknowledgments can synchronously be performed by conversational agents. The previous discussion establishes that the strong semantics implements precisely such a behavior. Our proposed solution thus complements existing ones.

4 Semantics with nested commitments for richer languages

In the preceding section, our semantics assumed a propositional language in which discourse moves like acknowledgments were analyzed as action operators on propositions. While this approach works for acknowledgments, it is difficult to adapt this to truly relational moves like Correction or most discourse relations. In this section we offer a full dynamic semantics

for a language with nested commitments in which we can refer to conversational moves and make assertions about relations between conversational moves.

We build on the mechanisms of SDRT (Asher & Lascarides 2003), with some modifications and extensions: we deal with nested commitments by allowing the informative content of discourse labels to talk about the content assigned to other discourse labels. As in SDRT, we will assume given lower-level language, typically, Dynamic Predicate Logic (DPL, Groenendijk & Stokhof 1991), or DRT (Kamp & Reyle 1993), and extend this language with discourse labels and relations.

Let I be a finite set of agents. Let $\Pi = \biguplus_{i \in I} \Pi_i$ be a disjoint union of $|I|$ countable sets of symbols (discourse labels). By definition, agent i is the *speaker* of any label $\pi \in \Pi_i$, which we write $spk(\pi) = i$. Let $\mathcal{R}$ be a finite set of relation symbols.

We assume a basic language $\mathcal{L}_0$ with the following: $\mathcal{L}_0$ has a binary constructor $\wedge$ and a unary constructor $\neg$, i.e., $\forall \varphi, \varphi' \in \mathcal{L}_0 \ \varphi \wedge \varphi', \neg\varphi \in \mathcal{L}_0$. We assume a class of models for $\mathcal{L}_0$, and every model $\mathcal{M}$ defines a set of states $X^{\mathcal{M}}$ and an interpretation $[\![\cdot]\!]_0^{\mathcal{M}}$ such that $[\![\varphi]\!]_0^{\mathcal{M}} \subseteq X \times X$, $[\![\neg\varphi]\!]_0^{\mathcal{M}} = \{(x,x) \mid \neg\exists x'(x,x') \in [\![\varphi]\!]_0^{\mathcal{M}}\}$ and $[\![\varphi \wedge \psi]\!]_0^{\mathcal{M}} = [\![\varphi]\!]_0^{\mathcal{M}} \circ [\![\psi]\!]_0^{\mathcal{M}} = \{(x,y) \mid \exists z (x,z) \in [\![\varphi]\!]_0^{\mathcal{M}} \text{ and } (z,y) \in [\![\psi]\!]_0^{\mathcal{M}}\}$. The latter requirements ensure that the interpretation of $\neg$ and $\wedge$ coincide at the lower and upper levels.

Our intended base-level language is DPL, as it yields a rich and expressive discourse semantics that can handle questions. However, in order to make clear the link between the present section's semantics and the dynamic logic of the previous section, we will also occasionally use a simpler propositional test logic $\mathcal{L}_{\text{test}}$ (which can be seen as a fragment of DPL). Simply define $\mathcal{L}_{\text{test}}$ as to contain the same formulae as the propositional logic over signature PROP, states X_{test} are valuations PROP $\mapsto \{0,1\}$, and for two valuations $v, v' \in X_{\text{test}}$, define $[\![\varphi]\!]_{\text{test}} = \{\langle v,v \rangle \mid v \vDash \varphi\}$, where $\vDash$ is the classical truth-maker of the (static) propositional logic over PROP.

The language $\mathcal{L}_{\Pi,\mathcal{R}}$ extends $\mathcal{L}_0$ and is defined by structural induction:

Definition 8 $\mathcal{L}_{\Pi,\mathcal{R}}$ is the smallest language such that:

$$\forall \varphi \in \mathcal{L}_0 \ \varphi \in \mathcal{L}_{\Pi,\mathcal{R}}$$
$$\forall \pi_1, \pi_2, \pi_3 \in \Pi \ \forall R \in \mathcal{R} \ R(\pi_1, \pi_2, \pi_3) \in \mathcal{L}_{\Pi,\mathcal{R}}$$
$$\forall \alpha \in \mathcal{L}_\pi \ \forall i \in I \ C_i \alpha \in \mathcal{L}_{\Pi,\mathcal{R}}$$
$$\forall \alpha \in \mathcal{L}_\pi \ \forall \pi \in \Pi \ \pi : \alpha \in \mathcal{L}_{\Pi,\mathcal{R}}$$
$$\forall \alpha, \alpha' \in \mathcal{L}_\pi \ \alpha \wedge \alpha' \in \mathcal{L}_{\Pi,\mathcal{R}}$$
$$\forall \pi, \pi_1, \pi_2 \in \Pi \ \pi_1 \overset{\pi}{\sim} \pi_2 \in \mathcal{L}_{\Pi,\mathcal{R}}$$
$$\forall \alpha \in \mathcal{L}_\pi \ \neg\alpha \in \mathcal{L}_{\Pi,\mathcal{R}}$$

An SDRS (segmented discourse representation structure) is a formula of $\mathcal{L}_{\Pi,\mathcal{R}}$ such that every occurrence of a formula of $\mathcal{L}_0$, or a formula of the form $C_i \alpha$ is guarded by some label $\pi \in \Pi$. In addition, define $\varphi \to \psi$ as $\neg(\varphi \wedge \neg\psi)$.

$\mathcal{L}_{\Pi,\mathcal{R}}$ adds several constructions to $\mathcal{L}_0$. First, there are statements of the form $\pi : \varphi$ that "store" the proposition φ in the content of a discourse label π. It implements the linguistic action of the speaker of π's saying φ. φ may itself be a structured discursive object, as in SDRT. Relational propositions $R(\pi_1, \pi_2, \pi_3)$ update the content of π_3 to express a relational move, recursively computed through the meaning of R and the content assigned to π_1 and π_2. Commitment operators are added, as well as a construction $\pi_1 \overset{\pi}{\sim} \pi_2$, which will implement ambiguity between the discourse moves stored under π_1 and π_2. We now turn to the semantics of $\mathcal{L}_{\Pi,\mathcal{R}}$.

A model for $\mathcal{L}_{\Pi,\mathcal{R}}$ is simply a model $\mathcal{M}$ for $\mathcal{L}_0$. Let $C = \{c \in (\Pi)^I \mid \forall i \in I \; c(i) \in \Pi_i\}$, the set of function associating to each agent i a label $c(i)$ whose speaker is i. The set of ($\mathcal{L}_{\Pi,\mathcal{R}}$-)states $S^{\mathcal{M}}$ for $\mathcal{M}$ is the product $X^{\mathcal{M}} \times C \times F^{\mathcal{M}}$ where $F^{\mathcal{M}}$ is the set of assignments to labels (also called label-assignments in the remainder of the paper). A state for the extended language $\mathcal{L}_{\Pi,\mathcal{R}}$ is thus a triple, consisting of a state x of the base language, a label $c(i)$ for each agent representing her current commitments[2], and a label-assignment f. We define label-assignments below. The purpose of a label-assignment f is to assign a (semantic) proposition $f(\pi)$ to a given discourse label π—i.e., in our dynamic logic, a set of transitions between states. Such a transition is technically a pair $\langle (x, c, f), (x', c', f') \rangle$, and itself involves assignment functions. We must therefore be careful to avoid circularity in the definition of $F^{\mathcal{M}}$, as is familiar from the work of Frank (1996) on modal subordination. Our problem and our solution are, however, slightly different than those for modal subordination. We will proceed in two steps. We will first define a set of bounded label-assignments, bounded in the sense that it only allows for a finite nesting of labels with a defined content into other labels. We give a semantics relative to such assignments of content to labels. The boundedness of assignments has an intuitive semantic counterpart: in any state, regardless of the linguistic actions executed so far, there is a finite bound n, for each agent i, on the maximal nesting of commitments $C_i C_i \ldots C_i \varphi$ that may evaluate to true for a non-tautological φ. In other words, there is always a maximal n such that agent's i commitments of order n are minimal (contain nothing besides logical tautologies). This is analogous to what happens in the "weak" dynamic commitment logic (without infinite sequences of acknowledgments) if we assume assume an initial model with minimal commitments. An immediate consequence is that it's impossible to reach common commitment on any content. We make this correspondence explicit below.

We will then extend our label-assignment functions to unbounded ones. Avoiding circularity in this case will require a quite tricky mathematical construction; however, the complexity of this construction will fade away thanks to two propositions stating that we can manipulate unbounded assignments exactly as bounded ones (namely, that a label assignment can evaluate any label into a set of state transitions, and that we can construct a new label-assignment from any label assignment σ, set of transitions T and label π by setting $\sigma(\pi) = T$). The semantics with unbounded assignments will be exactly the same as for bounded

---

2 i.e., the label which is currently maximal w.r.t discourse subordination, this allows to dynamically exclude labels which introduce only intermediary contributions that do not constitute a commitment of their own, as e.g., antecedents to conditionals.

assignments.

In what follows, every definition is relative to a given model $\mathcal{M}$; in order to simplify notation we drop the $\cdot^{\mathcal{M}}$ when referring to the sets $F^{\mathcal{M}}, X^{\mathcal{M}}, S^{\mathcal{M}}$ and interpretations $[\![\cdot]\!]_0^{\mathcal{M}}$ and $[\![\cdot]\!]^{\mathcal{M}}$.

We define by simultaneous induction the sets F^n of rank-n label-assignments, and S^n of rank n states.

Definition 9 (Rank of assignments)
(i) $F^0 = \varepsilon$, $S^0 = X \times C \times F^0$
(ii) $F^{n+1} = (\wp(S^n \times S^n))^{\Pi}$(functions from labels to set of transitions between rank-n states),

$$S^{n+1} = X \times C \times (\bigcup_{k \le n} F^n)$$

(iii) $F = \bigcup_{n < \omega} F^n$, $S = X \times C \times F$

Informally, a label-assignment of rank n, is either the special symbol ε, or a function from discourse labels to a set of transitions (of rank $n-1$).

Our semantic requires only two operations on label-assignments: the evaluation of a label-assignment f at a given label π, and the substitution of a new set of transitions T for the set of transitions previously assigned to π yielding a new label-assignment denoted as $f[T/\pi]$. To avoid circularity, however, we cannot define substitution by any set of transitions T; it suffices for our purposes to restrict the definition to *bounded* sets of transitions, bounded in the sense that there must be some $k < \omega$ such that $T \subseteq S^k \times S^k$. $\Delta_B \subsetneq \wp(S \times S)$ is the set of such bounded sets of transitions ($\Delta_B = \bigcup_{n \in \omega} \wp(S^n \times S^n)$). The two operations (substitution and evaluation) are defined below.

Definition 10 $f(\pi)$ is already defined for $f \ne \varepsilon$, since by construction, f is a function. We define in addition $\varepsilon(\pi) = \bigcup_{x \in X, c \in C} \langle (x,c,\varepsilon)(x,c,\varepsilon) \rangle$ (this choice of definition makes ε represents a state with minimal commitments over logical tautologies only).

For $T \in \Delta_B$, $\pi \in \Pi$, $f[T/\pi]$ is defined as the function f' such that:

$$\begin{cases} f'(\pi) = T \\ \forall \pi' \ne \pi \, f'(\pi') = f(\pi) \end{cases}$$

Note that, $T \in \Delta_B$ is sufficient to ensure that $f[T/\pi]$ stays in F.

Since the semantics depends on the considered set of relations and their semantics, we will assume that for every symbol $R \in \mathcal{R}$, assignment f and top labels c, a semantics $|R(\pi_1, \pi_2)|_c^f \in \Delta_B$ is defined, which, given an assignment to π_1 and π_2 provides a new proposition, that is to say, an element of Δ_B. $[\![\cdot]\!]$ interprets formulae into sets of transitions, and is defined inductively as follows (using infix notation $x \, \delta \, y$ as a syntactic sugar for $\langle x, y \rangle \in \delta$):

Definition 11 (Semantics for $\mathcal{L}_{\Pi,\mathcal{R}}$)

- For $\varphi \in \mathcal{L}_0$ $s[\![\varphi]\!]s'$ iff $s = (x,c,f)$, $s' = (x',c,f)$ and $x[\![\varphi]\!]_0 x'$
- $(x,c,f)[\![R(\pi_1,\pi_2,\pi_3)]\!](x',c',f')$ iff $c' = c[\pi_3/spk(\pi_3)]$ and $f' = f[(|R(\pi_1,\pi_2)|_c^f)/\pi_3]$
- $s[\![C_i\alpha]\!]s'$ iff $s = s' = \langle x,c,f\rangle$ and $\forall \langle s,s'\rangle \in f(c(i)) \exists s''\ s'[\![\alpha]\!]s''$
- $(x,c,f)[\![\pi:\alpha]\!](x',c',f')$ iff $f' = f[(f(\pi)\circ[\![\alpha]\!])/\pi]$
- $(x,c,f)[\![\pi_1 \overset{\pi}{\sim} \pi_2]\!](x',c',f')$ iff $x = x'$ and $c' = c[\pi/spk(\pi)]$ and $f' = f[f(\pi_1)\cup f(\pi_2)/\pi]$
- $s[\![\neg\alpha]\!]s'$ iff $s = s'$ and there exists no $s'' \in S$ such that $s[\![\alpha]\!]s''$

With $c[\pi/i]$ defined as the unique c' such that $c'(i) = \pi$ and $\forall j \neq i, c'(j) = c(j)$

Since a bounded label-assignment f assigns a bounded set of transitions to any label, a easy induction on α show that $f(\pi)\circ[\![\alpha]\!]$ is always a bounded set of transitions as well, and thus $f[(f(\pi)\circ[\![\alpha]\!])/\pi]$ is always well defined and so is the semantics.

We provide in addition the semantics of two basic discourse relations: `Continuation` and `Ackn`:

Definition 12

$$|\texttt{Continuation}(\pi_1,\pi_2)|_c^f = f(\pi_1)\circ f(\pi_2)$$
$$|\texttt{Ackn}(\pi_1,\pi_2)|_c^f = \{(\langle x,d,g\rangle,\langle y,e(\pi_1/spk(\pi_1)),h[f(\pi_1)/\pi_1]\rangle) \mid$$
$$(\langle x,d,g\rangle,\langle y,e,h\rangle) \in f(c(spk(\pi_2)))\}$$

We can now link the propositional dynamic logic of the previous section and our dynamic semantics. We define a translation from $\mathcal{L}_D$ into $\mathcal{L}_{\Pi,\mathcal{R}}$ taking $\mathcal{L}_{\texttt{test}}$ as lower-level language. Let $M^0 = \langle X,(R_i^0)_{i\in I}\rangle$ be the model of $\mathcal{L}_D$ whose set of world X is the set of valuations over signature PROP and $R_i^0 = X \times X$ (thus M^0 represents minimal commitments only to logical truths by each agent in I).

Proposition 3 (translation) *Algorithm 1, given as input a formula $\varphi \in \mathcal{L}_D$ and an initial "top" label $c \in C$ for each agent, yields a proposition $t(\varphi,c) = \chi \in \mathcal{L}_{\Pi,\mathcal{R}}$ such that*

$$\langle M^0,x\rangle \vDash \varphi \text{ iff } \exists f \in F\ (x,c,\varepsilon)[\![\chi]\!](x,c,f)$$

Consider $\varphi = [(p!\sim q!)^i]([C_i p!^j]C_j\perp \wedge [ack(p!^j)\sim ack(q!^j)][(C_i p)!^j]\neg C_j\perp)$. This formula of $\mathcal{L}_D$ states that after i said something ambiguous between p and q, j is inconsistent in

---

Algorithm 1 translation of $\mathcal{L}_D$ into $\mathcal{L}_{\Pi,\mathcal{R}}$

Assume a function $\texttt{fresh} : I \mapsto \Pi$ which enumerates each Π_i, i.e., such that each call $\texttt{fresh}(i)$ returns a label in Π_i and successive calls never return twice the same label.

1: **function** $t(\varphi, c)$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \triangleright \varphi \in \mathcal{L}_D \cup \mathcal{A}, c \in C$
2: $\quad$ **Case** $\varphi = p \in \text{PROP}$:
3: $\qquad$ Return p
4: $\quad$ **Case** $\varphi = C_i \varphi$:
5: $\qquad$ Return $C_i t(\varphi, c)$
6: $\quad$ **Case** $\varphi \wedge \psi$:
7: $\qquad$ Return $t(\varphi, c) \wedge t(\psi, c)$
8: $\quad$ **Case** $\varphi = [\alpha^i] \psi$:
9: $\qquad$ Let $\chi_\alpha = t(\alpha^i, c)$ and $\chi_\psi = t(\psi, c)$
10: $\qquad$ Return $\chi_\alpha \rightarrow \chi_\psi$
11: $\quad$ **Case** $\varphi = \psi!^i$:
12: $\qquad$ Return $c(i) : t(\psi, c)$
13: $\quad$ **Case** $\varphi = ack(\beta^x)^i$:
14: $\qquad$ Let $\pi_\beta = \texttt{fresh}(x)$
15: $\qquad$ Let $\chi_\beta = t(\beta^x, c[\pi_\beta/x])$
16: $\qquad$ Return $c(i) : \chi_\beta$
17: $\quad$ **Case** $\varphi = (\alpha \sim \beta)^i$:
18: $\qquad$ Let $\pi_\alpha = \texttt{fresh}(i)$ and $\pi_\beta = \texttt{fresh}(i)$
19: $\qquad$ Let $\chi_\alpha = t(\alpha, c[\pi_\alpha/i])$ and $\chi_\beta = t(\beta, c[\pi_\beta/i])$
20: $\qquad$ Return $\chi_\alpha \wedge \chi_\beta \wedge \pi_\alpha \overset{c(i)}{\sim} \pi_\beta$
21: $\quad$ **end Case**
22: **end function**

---

saying that i is committed to p, unless he first (ambiguously) acknowledges one of the two readings. φ is true in (any world of) $\mathcal{M}^0$. The translation of φ is the formula χ below (where symbols with a $\cdot^x$ exponent denote labels in Π^x):

$$(\pi_1^i : p \wedge \pi_2^i : q \wedge \pi_1^i \overset{\pi^i}{\sim} \pi_2^i) \rightarrow$$
$$\left((\pi^j : C_i p) \rightarrow C_j \bot) \wedge \left((\pi_1^j : \pi^i : p \wedge \pi_2^j : \pi^i : q \wedge \pi_1^j \overset{\pi^j}{\sim} \pi_2^j) \rightarrow ((\pi^j : C_i p) \rightarrow \neg C_j \bot) \right) \right)$$

χ first updates the content of labels π_1^i and π_2^i, set the content of π^i to be the union of the content of those two labels, then set π^i to be i's top label. Since this first state leaves the assignment to label π^j unchanged, If the initial assigment is ε, π^j contains no transition making $C_i p$ hold, hence executing the action $\pi^j : C_i p$ would empty π^j and commits j to the absurdum. The ambiguous action $(\pi_1^j : (\pi^i : p) \wedge \pi_2^j : (\pi^i : q) \wedge \pi_1^j \overset{\pi^j}{\sim} \pi_2^j)$ on the other hand upates π^j to contain at least one transition placing p in the content of π^i, which can then

be selected through $\pi^j : C_i p$, and won't empties j's commitments. Thus, a state $\langle x, c, \varepsilon \rangle$ is accepted by χ.

We now extend our semantics with assignments interpreting infinite descending chains of labels with defined content. This enables us to do two things: 1) deal with infinite sequences of Ackn and 2) have states in which agents are commonly committed to some non-tautological content. We build unbounded label-assignments over bounded ones. by exploiting a notion of bounded bisimulation over bounded label-assignments. In the following, $\bar{x}$ will generally denote a tuple $\langle x, c \rangle \in X \times C$.

Definition 13 We define n-bisimilarity ($\cong_n$) between bounded label-assignments inductively:

  - $\forall f, f' \in F\ f \cong_0 f'$.
  - $f_0 \cong_{n+1} f_1$ iff the two following conditions hold:
  (forth) $\forall \pi \forall \langle (\bar{x}, g_0), (\bar{x}', g_0') \rangle \in f(\pi) \exists \langle (\bar{x}, g_1)(\bar{x}', g_1') \rangle \in f'(\pi)$ such that $g_1 \cong_n g_1'$ and $g_0 \cong_n g_0'$.
  (back) $\forall \pi \forall \langle (\bar{x}, g_1), (\bar{x}', g_1') \rangle \in f'(\pi) \exists \langle (\bar{x}, g_0)(\bar{x}', g_0') \rangle \in f(\pi)$ such that $g_1 \cong_n g_1'$ and $g_0 \cong_n g_0'$.

To ease notational clutter, we extend the notation $\cong_n$ to pairs of transitions, by writing $\langle (\bar{x}, f), (\bar{y}, g) \rangle \cong_n \langle (\bar{x}', f'), (\bar{y}', g') \rangle$ as a shortcut for $\bar{x} = \bar{y}$, $\bar{x}' = \bar{y}'$ and $f \cong_n f'$ and $g \cong_n g'$. We are now ready to define unbounded assignments:

Definition 14 (Unbounded label-assignments) The set of unbounded label-assignments F_ω is defined as the set of infinite sequences $\sigma = \langle \sigma_0, \sigma_1, \ldots, \sigma_i, \ldots \rangle$ of bounded label-assignments, such that: (i) $\forall i\ \sigma_i \cong_{i+1} \sigma_{i+1}$ and (ii) $\forall i\ \sigma_i \in F^{i+1}$.

This construction allows us, at each index, to find a label-assignment that is compatible with and refines the assignment of the preceding indices, allowing us to evaluate arbitrarily deep nestings of discourse labels by picking an assignment at a sufficiently high index.

Let $S_\omega = X \times C \times F_\omega$ denote the new set of states, now based on unbounded assignments (the analog to S). Corresponding transitions are thus in $S_\omega \times S_\omega$.

Definition 15 let $\sigma, \sigma' \in F_\omega$, $\sigma \cong \sigma'$ iff $\forall i\ \sigma_i \cong_{i+1} \sigma_i'$. $\cong$ is an equivalence relation.

We extend the notation $\cong$ to pairs of transitions, as we did for finite bisimulation. Finally, we introduce the notion of a diagonal, which will prove useful later on:

Definition 16 (Diagonal) Let $\langle s_i = (\bar{x}, \sigma^i) \rangle_{i \in \omega}$ be a sequence of states with constant first components $\bar{x} = \langle x, c \rangle$, and such that $\sigma_0^0 \cong_1 \sigma_1^1 \cong_2 \sigma_2^2 \ldots \sigma_n^n \cong_{n+1} \sigma_{n+1}^{n+1}$ we call the diagonal of such a sequence, the state $\delta(\langle s_i \rangle_{i \in \omega}) = \langle \bar{x}, \langle \sigma_i^i \rangle_{i \in \omega} \rangle$

The last ingredients needed for our dynamic semantics with unbounded assignments are evaluation and substitution, which we now define:

Definition 17 (Evaluation) Let $\sigma \in F_\omega$. Define $\sigma(\pi)$ as:

$$\sigma(\pi) = \{ \langle (\bar{x}, \alpha), (\bar{y}, \beta) \rangle \mid \forall i \langle (\bar{x}, \alpha_i), (\bar{y}, \beta_i) \rangle \in \sigma_{i+1}(\pi) \text{ and } \alpha, \beta \in F^\omega \}$$

For any equivalence relation $\bowtie$, let $x = y[\bowtie]$ denote equality modulo $\bowtie$.

Proposition 4 *if* $\sigma \cong \sigma'$ *then* $\sigma(\pi) = \sigma'(\pi)[\cong]$

An unbounded assignment always assigns to a label a set of transitions closed under diagonal; i.e., whenever $\forall k \in \omega \; \langle s_k, s_k' \rangle \in \sigma(\pi)$, such that both $\delta_s = (\langle s_k \rangle_{k \in \omega})$ and $\delta_{s'} = \delta(\langle s_k' \rangle_{k \in \omega})$ are defined, we have $\langle (\delta_s, \delta_{s'}) \rangle \in \sigma(\pi)$.

We now define the substitution operation. As in the bounded case, we must ensure that the set of transitions T we substitute is of the same kind that an unbounded label assignment can have as output. This time the restriction is not one of boundedness, but instead, following the previous remark, that T be closed under diagonal.

Definition 18 (Substitution) Let $T \subseteq S_\omega \times S_\omega$ such that T is closed under diagonal. Define $T_i = \{ \langle (\bar{x}, \alpha_i), (\bar{y}, \beta_i) \rangle \mid \langle (\bar{x}, \alpha), (\bar{y}, \beta) \rangle \in T \}$. For $\pi \in \Pi$ and $\sigma \in F_\omega$ define $\sigma[T/\pi]$ as

$$\sigma' = \sigma_0[T_0/\pi] \cdot \sigma_1[T_1/\pi] \cdot \ldots \sigma_n[T_n/\pi] \cdot \ldots$$

Note that since by definition of F_ω, $T_i \in S^{i+1} \times S^{i+1} \subseteq \Delta_B$, $\sigma_i[T_i/\pi]$ is well defined.

The final proposition we need is:

Proposition 5 $\sigma[T/\pi] \in F_\omega$. *Furthermore* $\sigma[T/\pi](\pi) = T[\cong]$.

Propositions 4 and 5 together imply that we can define our semantics exactly as we did for bounded assignments in definition 11, provided that the interpretation of discourse relations

satisfies the following constraint: for every relation symbol R and assigment f, $|R(\pi_1, \pi_2)|^f$ is closed under diagonalization. This holds for our definition of `Continuation` and `Ackn`.

Diagonalization captures precisely the regularity required for an infinite sequence of actions to be representable as a simple model update, which, as we have seen, is the case for acknowledgments. The very reason that allowed us, in the previous section, to capture an infinite sequence of iterated acknowledgment in a finite action-structure, was that level n actions of the hierarchy of acknowledgments only affects commitments of order n. The model obtained after any acknowledgment of order 1 is 0-bisimilar to the initial model (actual facts are not modified). Applying second-order acknowledgments modifies only second order commitments, and the model after this second step is 1-bisimilar to the model at step one. Applying the successive levels of the acknowledgments hierarchy thus yields a sequence of models such that the nth is $n-$ bisimilar to the $(n+1)^{\text{th}}$. The finite action-structure semantically equivalent to iterated acknowledgments, because it yields through update a single model, which is, for every integer n, $n-$bisimilar to the nth model in the infinite sequence of updates. As the exact same relationship holds between an infinite sequence $\langle \sigma_k \rangle_{k \in \omega}$ of assignment which admits a diagonal and its diagonal, we can use infinite iterated acknowledgments in our dynamic semantics. To that end we introduce a new relation $\texttt{Ackn}^\omega$.

Just as with the propositional infinite sequence of actions of section 3, we define first an infinite sequence of actions. Let $\pi \in \Pi, c \in C$ and for $i, j \in I$ let $\langle \pi^k_{ok_i^j} \rangle_{k \in \omega}$ be a sequence of pairwise disjoint labels in Π_i. Define level-n iterated acknowledgments of π, χ_n^{ack}, with $\chi_0 = \bigwedge_{i,j \in I} \texttt{Ackn}(\pi, \pi^0_{ok_i^j}, c(i))$ and $\forall n\; \chi_{n+1} = \chi_n \wedge \bigwedge_{i,j \in I} \texttt{Ackn}(c(j), \pi^k_{ok_i^j}, c(i))$. Let $x \in X$ and $f \in F$ and define $\langle x, c_n, f_n \rangle$ as the unique state such that $\langle x, c, f \rangle [\![\chi_n]\!] \langle x, c_n, f_n \rangle$. $\forall n\; c_n = c$, moreover the sequence $\langle x, c, f_n \rangle_{n \in \omega}$ admits a diagonal. Let $\delta^{ack}_{\langle x, c, f \rangle}$ denote this diagonal. We can finally add a construction $\texttt{Ackn}^\omega(\pi, (\pi^i)_{i \in I})$ to the language, which , first, copy each agent i's commitment into π^i (in order to not erase the content of the current top labels and keep track of previous states), and then to proceed to iterate acknowledgments. Define therefore for a pair $\langle c, f \rangle$, the pair $\bar{c} = c[(\pi^i/i)_{i \in I}]$ and $\bar{f} = f[(f(c(i))/\pi^i)_{i \in I}]$, this performs the "copy" of each agent's commitment into the new labels. Define:

$$s [\![\texttt{Ackn}^\omega(\pi, (\pi^i)_{i \in I})]\!] s' \text{ iff } s = \langle x, c, f \rangle \text{ and } s' = \delta^{ack}_{\langle x, \bar{c}, \bar{f} \rangle}$$

5 Examples revisited

In previous sections, we defined a propositional dynamics logic describing the evolution of commitments, and a full dialog semantics with relational moves with the same logical mechanisms. Furthermore, these systems describe both possibilities of a "strong"or "weak" semantics, linking them through iterated acknowledgments. But we still have to solve our dilemma: neither a systematic "strong" interpretation of dialog moves, nor a "weak" one is satisfactory, as it either yields meaningless acknowledging moves or impossibility of grounding. However, as both kind of interpretations are now part of a common semantic vocabulary, we are no longer committed to a systematic use of one or the other. The

remaining question is thus: what is exactly the condition at which an agreement is reached that synchronous iterated acknowledgements indeed happened?

Our answer to this question is: one has to ask. Our solution builds on our treatment of ambiguity: acknowledgments, such as the one that 1 performs in (2c) of example (2), are considered as ambiguous in their strength and only a confirming question and answer might raise this ambiguity, and reach a non-deniable common-commitment. Considering again 1's acknowledgment in (2c), we represent 1's contribution (OK), as ambiguous between a simple acknowledgment by 1 and an acknowledgment of a common acknowledgment of 0's "no" answer. Representing 0's "no" as the action $\neg \texttt{bank!}^0$, (2c) is thus modeled as $ack(\texttt{bank!}^1) \sim ack(!^{\omega}_{0,1}(\texttt{bank!}^0))^1)$. In the relational semantics, we have (2b) as $\pi^0 : \neg\texttt{bank}$, followed by

$$\texttt{Ackn}^{\omega}(\pi^0, (\pi^0_{strong}, \pi^1_{strong})) \wedge \pi^0 \overset{\pi'^0}{\sim} \pi^0_{strong} \wedge \texttt{Ackn}(\pi_0, \pi^1_{OK}, \pi^1_{weak}) \wedge \pi^1_{weak} \overset{\pi^1}{\sim} \pi^1_{strong}.$$

This move commits 1 (*via* π^1) to an ambiguous proposition. The subsequent confirming question in 2c, is then modeled as asking 0 to raise the ambiguity. We did not discuss the semantics of questions, but questions in dynamics semantics have been discussed at length in the litterature (see e.g., Groenendijk 2003). We could modifies states as to include issue partitions in order to represent questions. The recent account in dynamic epistemic logic of van Benthem & Minica (2012) could also be integrated into section 3's dynamics. What is crucial independently of this choice, is that 1's commitments indeed licences a polar question $C^*C_0\neg bank$?, to which 0 answering yes[3] brings a commitment $C_0C^*C_0\neg bank$. After such a move, whatever 0 may say, he cannot deny, at any level, that he committed to $\neg bank$. A simple acknowldegment of 0's answer yield the common commitment.

Our relational semantics also allows us to deal with self-corrections. Corrections need the full relational semantics, because one content is revised by another. Consider example (2) again and 0's correction in (2d). When *B* says *No. sorry, in fact, I had an account there*, the move attaches to B's affirmative answer to P's initial question in (2a). But the effects of this discourse move change the surrounding discourse structure. The semantics of Correction replaces the content of the original affirmative answer with the negative answer in the corrective move. Because of this, the follow up question of P and B's second affirmative answer are now moot. As in Lascarides & Asher 2009, this revision requires that our states have a copy of the SDRS constructed for the dialogue and the revision is calculated on that structure. B and P's commitments are recomputed on the revised SDRS. We thus model self-corrections as a revision of one's commitments.

Self-corrections thus erase the commitments of the corrected action and possibly also the commitments ensuing from subsequent dependent actions like its acknowledgment. An immediate consequence is that self-corrections make commitments, even common commitments, unstable (non-monotonic).

---

3 which will disambiguate his own commitments in π'_0 by selecting the content of π^0_{strong}

References

Asher, Nicholas & Tim Fernando. 1997. Nonincrementality and revision in dialogue. In Anton Benz & Gerhard Jäger (eds.), *MunDial Workshop. University of München*, University of München.

Asher, Nicholas & Alex Lascarides. 2003. *Logics of Conversation*. Cambridge University Press.

Asher, Nicholas, Soumya Paul & Antoine Venant. 2014. Message exchange games in strategic conversation. Forthcoming in *Journal of Philosophical Logic*.

Baltag, Alexandru, Lawrence S. Moss & Slawomir Solecki. 1998. The logic of public announcements, common knowledge, and private suspicions. In *7th Conference on Theoretical Aspects of Rationality and Knowledge* TARK '98, 43–56. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. http://dl.acm.org/citation.cfm?id=645876.671885.

van Benthem, Johan & Stefan Minica. 2012. Toward a dynamic logic of questions. *J. Philosophical Logic* 41(4). 633–669. doi:10.1007/s10992-012-9233-7. http://dx.doi.org/10.1007/s10992-012-9233-7.

Clark, Herbert. 1996. *Using Language*. Cambridge, England: Cambridge University Press.

Frank, Anette. 1996. *Context dependence in modal constructions*. Stuttgart, Germany: IMS, University of Stuttgart Ph.d. dissertation.

Ginzburg, Jonathan. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford: Oxford University Press.

Groenendijk, Jeroen. 2003. Questions and answers: Semantics and logic. In *2nd CologNET-ElsET Symposium. Questions and Answers: Theoretical and Applied Perspectives*, 16–23.

Groenendijk, Jeroen & Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14. 39–100.

Hamblin, Charles. 1987. *Imperatives*. Oxford: Blackwells.

Kamp, Hans & Uwe Reyle. 1993. *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Lascarides, Alex & Nicholas Asher. 2009. Agreement, disputes and commitment in dialogue. *Journal of Semantics* 26(2). 109–158.

Traum, David. 1994. *A Computational Theory of Grounding in Natural Language Conversation*: Computer Science Department, University of Rochester dissertation.

Traum, David & James Allen. 1994. Discourse obligations in dialogue processing. In *32nd Annual Meeting of the Association for Computational Linguistics (ACL94)*, 1–8. Las Cruces, New Mexico.

Venant, Antoine & Nicholas Asher. 2015. Dynamics of public commitments in dialogue. In *11th International Conference on Computational Semantics*, 272–282. London, UK: Association for Computational Linguistics. http://www.aclweb.org/anthology/W/W15/W15-0131.pdf.

Venant, Antoine, Nicholas Asher & Cedric Degremont. 2014. Credibility and its attacks. In

Antoine Venant
IRIT
118 Route de Narbonne,
F-31062 TOULOUSE CEDEX 9
FRANCE
antoine.venant@irit.fr

Nicholas Asher
IRIT
118 Route de Narbonne,
F-31062 TOULOUSE CEDEX 9
FRANCE
asher@irit.fr