Micro-analysis of the belief transfer in information dialogues

Roser Morante and Harry Bunt
Tilburg University
The Netherlands
{r.morante|harry.bunt}@uvt.nl

Abstract

This paper describes work in progress, aimed at providing detailed empirical evidence about the processes of creating and updating information states in dialogue participants as the result of the utterances they exchange.

1 Introduction

Formal and computational work on dialogue modelling much of the time relies on the modelling of beliefs, goals, and intentions, following the 'BDI' paradigm that goes back to the work of Perrault, Allen and Cohen (Allen and Perrault, 1980; Cohen and Perrault, 1979; Perrault and Allen, 1980). More recently, this approach has taken a new form known as the 'information-state' or 'context-change approach', which uses the representation of agent's states of beliefs and other information in relation to the systematic (often plan-based) use of dialogue acts, as e.g. by Allen & Schubert (1994); Bunt (1996); Larsson & Traum (2000); Traum & Hinkelman (1992). With very few exceptions,[1] this more recent work does not involve truly formal modelling of information states, nor is it based on much empirical research into the details of how dialogue acts create and update the information states of dialogue participants. This paper describes some of our ongoing research that is aimed at providing an empirical basis for modelling the dynamics of dialogue agents' information states, by studying examples of recorded dialogue fragments under a formal microscope, trying to indicate for each dialogue utterance in detail which information it creates, strengthens, or cancels.

In doing this analysis, we include a first stage which is intended to be largely theory-neutral, by simply looking at the utterances, deciding exactly what they mean, and trying to make explicit what information they convey to the addressee. In a later stage we will perform another analysis of the same material using the system of dialogue acts defined in Dynamic Interpretation Theory (DIT) to annotate utterances and apply the definitions of the dialogue act types. We then compare the two analyses. Already at this stage it seems evident that the two analyses will *not* give the same results, since some of the dynamics of agents' information states is determined by global properties of stretches of dialogue, rather than purely locally by the effects of individual dialogue acts. This is an interesting first result. Second, the analysis is throwing new light on the phenomenon of grounding, which we believe can be analysed fruitfully by applying a formal notion of *mutual belief* and

---

[1] An exception is the work by Poesio & Traum, (1998).

investigating how mutual beliefs about weak beliefs may get strengthened to strong mutual beliefs.

In section 2 of this paper we briefly introduce DIT; in section 3 we describe the analysis method we applied, and we provide an example. Finally, in section 4 we summarize our initial findings and indicate directions for future research.

2 Information states in DIT

The background of our work is the theoretical framework of Dynamic Interpretation Theory (DIT; Bunt 2000), which gives a central position to the notions of *dialogue context* and *dialogue act*. A dialogue participant's beliefs about the domain and about the dialogue partner form a crucial part of his information state which in DIT is also called his *context*; there is no objective notion of dialogue context in DIT, but only the contexts (information states) of each participant. Dialogue acts are defined in DIT as semantic operations, used by dialogue participants to influence each other's context.

This means that dialogue acts are situated conceptually between utterances and context-changing operations; utterances are assumed to encode multiple dialogue acts, and their context-changing effects are defined through these dialogue acts. In this paper, however, we analyze the beliefs created and/or changed by dialogue utterances *directly*, without the intervention of dialogue acts. This has several potential advantages.

First, an analysis of context change based on dialogue acts is in danger of paying too little attention to changes that are not due to local effects of individual dialogue acts. We will see examples of that, relating to the phenomenon that grounding often occurs through implicit positive feedback. Second, by relating on the one hand context changes directly to utterances and in a later stage on the other hand assigning dialogue acts to utterances

(on the basis of utterance and context features), we will obtain evidence on the validity and limitations of the modelling of a dialogue in terms of dialogue acts. Third, for the same reason, such an analysis can provide detailed insight into the semantics of dialogue acts whose meaning is not so easily defined in terms of changing beliefs, such as positive and negative feedback acts or time management acts. Fourth, and finally, the analysis of how the participants' information states change during a dialogue can help us to identify (sub-)types of dialogue acts that have not been noted before.

In this paper we focus on the analysis in terms of beliefs and goals. We analyze the processes involved in the creation and maintenance of the agent's beliefs about the partner's current beliefs and goals. The development of this analysis should make explicit how the flow of information in a dialogue affects the beliefs of both participants, in particular in relation to grounding and to persistence (and strength) of beliefs and goals. The data we analyze is a collection of information seeking dialogues, in which a user interacts with a simulated interactive help assistant for a fax machine.

3 Analysis method

For every utterance we represent several types of effects of the utterance on the cognitive state of the speaker and the hearer: effects of understanding, effects of expectations of being understood, and effects of processing the information which is being transferred. We represent the effects by means of some operators.

We define four types of beliefs (weak belief ($\Vdash$), strong belief ($\Vdash$), knowing the value of ($\vdash$), strong mutual belief ($\Vdash^*$)), and a notion of goal ($\vdash\!\!\!\sim$).

- Weak belief: a belief that an agent is not certain about and that requires confirma-

n.	op	beliefs system		n.	op.	beliefs user
		USER: Waar moet ik het te kopiëren papier invoeren? (Where should I feed the paper to be copied?)				
	prec	$\forall x.\varphi(x)S \vdash \psi(x)$		gu1	goal	$U \vdash\!\!\sim \forall x.\varphi(x)U \vdash \psi(x)$
	prec	$S \Vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$				
s1		$S \Vdash gu1$				
s2		$S \Vdash^{\pm} U \Vdash \cdot\; S \Vdash gu1$		u1		$U \Vdash^{\pm} U \Vdash \cdot\; S \Vdash gu1$
		SYSTEM: Wilt U een kopie maken? (Do you want to make a copie?)				
gs1	goal	$S \vdash\!\!\sim S \vdash p$			prec	$U \vdash p$
					prec	$U \Vdash p$
s3		$S \Vdash^{\pm} S \Vdash \cdot\; U \Vdash gs1$		u2		$U \Vdash gs1$
				u3		$U \Vdash^{\pm} S \Vdash \cdot\; U \Vdash gs1$
		USER: Ja (Yes)				
s4		$S \Vdash U \Vdash p$				
s5		$S \Vdash U \Vdash gs1$				
s6	ad:s4	$S \Vdash p$				
s7		$S \Vdash^{\pm} U \Vdash \cdot\; S \Vdash U \Vdash gs1$		u4		$U \Vdash^{\pm} U \Vdash \cdot\; S \Vdash U \Vdash gs1$
s8		$S \Vdash^{\pm} U \Vdash \cdot\; S \Vdash U \Vdash p$		u5		$U \Vdash^{\pm} U \Vdash \cdot\; S \Vdash U \Vdash p$
s9	ca:gs1	$S \Vdash^{\pm} U \Vdash \cdot\; S \Vdash p$		u6		$U \Vdash^{\pm} U \Vdash \cdot\; S \Vdash p$
		SYSTEM: In de invoersleuf (In the paper tray)				
s10	st:s2	$S \Vdash^{\pm} U \Vdash S \Vdash gu1$		u7	st:u1	$U \Vdash^{\pm} U \Vdash S \Vdash gu1$
s11	st:s7	$S \Vdash^{\pm} U \Vdash S \Vdash U \Vdash gs1$		u8	st:u4	$U \Vdash^{\pm} U \Vdash S \Vdash U \Vdash gs1$
s12	st:s8	$S \Vdash^{\pm} U \Vdash S \Vdash U \Vdash p$		u9	st:u5	$U \Vdash^{\pm} U \Vdash S \Vdash U \Vdash p$
s13	st:s9	$S \Vdash^{\pm} U \Vdash S \Vdash p$		u10	st:u6	$U \Vdash^{\pm} U \Vdash S \Vdash p$
				u11		$U \Vdash S \Vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$
				u12	ad:11	$U \vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$
s14		$S \Vdash^{\pm} S \Vdash \cdot\; U \Vdash S \Vdash gu1$		u13		$U \Vdash S \Vdash \cdot\; U \Vdash S \Vdash gu1$
s15		$S \Vdash^{\pm} S \Vdash \cdot\; U \Vdash S \Vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$		u14		$U \Vdash^{\pm} S \Vdash \cdot\; U \Vdash S \Vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$
s16		$S \Vdash^{\pm} S \Vdash \cdot\; U \Vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$		u15	ca:gu1	$U \Vdash^{\pm} S \Vdash \cdot\; U \Vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$
		USER: Dank u (Thank you)				
s17	st:14	$S \Vdash^{\pm} S \Vdash U \Vdash S \Vdash gu1$		u16	st:u13	$U \Vdash^{\pm} S \Vdash U \Vdash S \Vdash gu1$
s18	st:15	$S \Vdash^{\pm} S \Vdash U \Vdash S \Vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$		u17	st:u14	$U \Vdash^{\pm} S \Vdash U \Vdash S \Vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$
s19	st:16	$S \Vdash^{\pm} S \Vdash U \Vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$		u18	st:u15	$U \Vdash^{\pm} S \Vdash U \Vdash \forall x : \varphi(x) \wedge \mu(x) \rightarrow \psi(x)$

Figure 1: Simplified analysis of a dialogue fragment

tion to become strong belief. We assume that the hearer has a weak belief about the effects of his utterances, as long as he does not receive any feedback.

- Strong belief: a belief that the agent has no doubt about. We start from the assumption that the addressee of an utterance has no doubt about the appropriateness of the utterance. This is why the effect of an utterance in the addressee is represented by a strong belief.

- Knows value of: is formally defined as an abbreviation of a combination of strong beliefs.

- Mutual beliefs: these are the beliefs that both agents have about what is mutually believed (recursively).

We describe how beliefs change through operations on previous beliefs. As the dialogue evolves new beliefs are created and existing beliefs might change, or get cancelled. In order to model the changes we define the following operations that update beliefs and goals:

- Strengthening (st): A belief of S as an effect of U's utterance-1 will be strengthened when U emits another utterance related to utterance-1 that allows S to think that his belief was right. If there is negative evidence the belief is cancelled.

- Cancellation (ca): A belief is cancelled when it is disconfirmed or negated, or when new beliefs have been created that cancel its persistence. A belief can be cancelled if there is negative evidence about the belief. A goal to obtain information is cancelled when that information is provided.

- Adoption (ad): A belief is adopted when the agents incorporate it in their knowledge of the world. When adoption takes place it is often the case that a goal is accomplished, and thus cancelled.

We present a simplified example of analysis in Figure 1. Columns 1 to 3 contain the information related to the system's beliefs, and columns 4 to 6 contain the information related to the user's beliefs. In columns 1 and 4 the beliefs are numbered, in columns 2 and 5 the operations on beliefs are indicated, and in columns 3 and 6 the beliefs are formalized.

For every utterance we indicate the most important preconditions and goals, and the effects it causes in the hearer and speaker. Preconditions are conditions in the cognitive state of the speaker, that trigger or enable the emission of the utterance.

In this analysis grounding is interpreted as the coincidence of the same beliefs mutual beliefs in both participants.

4 Discussion and future work

The analysis shows that during a dialogue both participants for a while build up the state of beliefs about each other's beliefs and goals, as well as about the domain of discourse, and that after some time, when sufficiently much implicit or explicit positive feedback has occurred, the belief states become simpler and converge to a state when some common ground is established.

Interestingly this process is determined to some extent by nonlocal effects of sequences of utterances rather than by purely local effects of individual utterances.

Future research will involve an analysis that first, and independently, annotates all dialogue acts. Subsequently, the detailed local effects of dialogue acts, as predicted by DIT, will be compared with the present analysis.

The result will be feeded back into the theory and will be used to further investigate the interaction between dialogue acts and dialogue (belief) content.

References

James F. Allen and C. R. Perrault. 1980. Analyzing intention in dialogues. *Artificial Intelligence* 15(3): 143-178.

James F. Allen and L. Schubert. 1994. The TRAINS project: A case study in defining a conversational planning agent. Technical Report TR 532, URCSD.

Harry Bunt. 1996. Interaction management functions and context representation requirements. In S. LuperFoy, A. Nijholt, and G. Veldhuizen van Zanten (eds.). *Dialogue Management in Natural Language Systems. Proc. of 11th Twente Workshop on Language Technology*. University of Twente, Enschede, pp. 187-198.

Harry Bunt. 2000. Dialogue pragmatics and context specification. In H. Bunt and W. Black, (eds.). *Abduction, Belief and Context in Dialogue*. John Benjamins, Amsterdam, pp. 81-150.

P.R. Cohen and C. R. Perrault. 1979. Elements of a plan-based theory of speech acts. *Cognitive Science* 3: 177-212.

Staffan Larsson and David R. Traum. 2000. Information state and dialogue management in the Trindi dialogue move engine toolkit. *Natural Language Engineering* 6(3-4): 323-340.

R.C. Perrault and J.F. Allen. 1980. A plan-based analysis of indirect of speech acts. *American Journal of Computational Linguistics* 6: 167-182.

M. Poesio and D. Traum. 1998. Towards an axiomatization of dialogue acts. *Proceedings of the Twente Workshop on the Formal Semantics and Pragmatics of Dialogues (13th Twente Workshop on Language Technology)*, pp. 207-222.

D.R. Traum and E.A. Hinkelman. 1992. Conversation acts in task-oriented spoken dialogue. *Computational Intelligence Special Issue: Computational Approaches to Non-Literal Language* 8(3): 575-599.