Belief Dynamics in Cooperative Dialogues

Article *in* Journal of Semantics · May 2000

DOI: 10.1093/jos/17.2.91

Dominique Longin
Institut de Recherche en Inf…

Journal of Semantics 17: 91–118

Belief Dynamics in Cooperative Dialogues

ANDREAS HERZIG and DOMINIQUE LONGIN
Institut de Recherche en Informatique de Toulouse

Abstract

We investigate how belief change in cooperative dialogues can be handled within a modal logic of action, belief, and intention. We first review the main approaches of the literature, and point out some of their shortcomings. We then propose a new framework for belief change. Our basic notion is that of a contextual topic: we suppose that we can associate a set of topics with every agent, speech act, and formula. This allows us to talk about an agent's competence, belief adoption, and belief preservation. Based on these principles we analyse the agents' belief states after a speech act. We illustrate our theory by a running example.

1 INTRODUCTION

Participants in task-oriented dialogues have a common goal, to achieve the task under consideration. Each of the participants has some information necessary to achieve the goal, but none of them can achieve it alone. Consider e.g. a system delivering train tickets to users. The system cannot do that without user input about destination and transport class. The other way round, the user needs the system to get his ticket.

Each of the participants is supposed to be *cooperative*. This is a fundamental and useful hypothesis. Informally, a person is cooperative with regard to another one if the former helps the latter to achieve his goals (cf. Grice's cooperation principles, as well as his conversation maxims (Grice 1975)). For example, if the system learns that user wants a train ticket, then the system will intend to give it to him. The other way round, if the system asks for some piece of information it needs to print the ticket, then the user answers the questions asked by the system.

Each participant is supposed to be *sincere*: his utterances faithfully mirror his mental state. If a participant says 'the sky is blue', then he indeed believes that the sky is blue. Such a hypothesis means that contradictions between the presuppositions of a speech act and the hearer's beliefs about the speaker cannot be explained in terms of lies. Note that our sincerity assumption is much weaker than in other approaches, where sincerity is sometimes viewed as the criterion of input adoption (Cohen & Levesque 1990c).

Under these hypotheses, how should the mental state of a rational agent participating in a conversation evolve? In the sequel we call *belief change* the process leading an agent from a mental state to a new one.

The following dialogue is our running example to highlight different problems and our solutions. There are only two agents, the system *s* and the user *u*:

s_1: *Hello. What do you want?*
u_1: *A first class train ticket to Paris, please.*
s_2: *150 €, please.*
u_2: *Ouups! A second-class train ticket, please.*
s_3: *100 €, please.*
u_3: *Can I pay the 80 € by credit card?*
s_4: *The price isn't 80 €. The price is 100 €. Yes, you can pay the 100 € by credit card.*
u_4: ...

This illustrates that in a conversation agents might change their mind, make mistakes, understand wrongly, etc. Since by our cooperation hypothesis the agents interact with each other in order to achieve the dialogue goal, they are the victims of such phenomena. They must consequently be taken into account when modelling the evolution of mental states.

In our example, the system

- accepts some information (e.g. information about destination and class—cf. u_1);
- derives supplementary information not directly contained in the utterance by using laws about the world (e.g. to derive the price if the user informs about his destination and class—cf. s_2);
- sometimes accepts information contradicting its own beliefs, in particular when the user changes his mind (e.g. switching from a first-class ticket to a second-class ticket—cf. u_2);
- preserves some information it believed before the utterance (e.g. the system preserves the destination even when the class changes—cf. u_2);
- may refuse to take over some information, in particular if the user tries to inform the system about facts the user isn't competent at (e.g. prices of train tickets—cf. s_4).

To sum up, *s* has two complementary tasks: (1) dealing with contradictions between his mental state and consequences of the input, and (2) preserving his old beliefs that do not contradict this input.

We consider each participant to be a *rational agent* having *mental states* represented by different mental attitudes such as belief, choice, goal, intention, etc. Belief change takes place within a formal *rational balance theory* and a formal *rational interaction theory à la* Cohen & Levesque (1990a, 1990c). These approaches analyse linguistic activity within a theory of actions: this is the base of so-called *BDI-architectures* (for Belief, Desire, and Intention). Each utterance is represented by a (set of) speech act(s) (Austin 1962; Searle 1969), in a way similar to Sadek (2000).[1] Belief change triggered by these speech acts is analysed in terms of consequences of these speech acts.

From an objective point of view, a dialogue is a sequence of sets of speech acts $(\alpha_1, \ldots, \alpha_n)$, where each α_{k+1} maps a state S_k to a new state S_{k+1}:

$$S_0 \xrightarrow{\alpha_1} S_1 \xrightarrow{\alpha_2} \ldots \xrightarrow{\alpha_n} S_n.$$

S_0 is the initial state (before the dialogue starts). Given S_k and α_{k+1}, our task is to construct the new state S_{k+1}.

The background of our work is an effective generic real-time cooperative dialogue system that has been specified and developed by the France Télécom R&D Center. This approach consists in first describing the system's behaviour within a logical theory of rational interaction (Sadek 1991, 2000, 1992), and second implementing this theory within an inference system called ARTIMIS (Bretier & Sadek 1997; Sadek *et al.* 1996, 1997). For a fixed set of domains, this system is able to accept nearly unconstrained spontaneous language as input, and react in a cooperative way. The activities of the dialogue system are twofold: to take into account the speaker's utterances, and to generate appropriate reactions. The latter reactive part is completely defined in the current state of both the theory and the implementation. On the other hand, the acceptance of an utterance is handled only partially, in particular its belief change part.

In our approach, building on previous work in Fariñas del Cerro *et al.* (1998), we implement belief change by an axiom of *belief adoption* and one of *belief preservation*. Both of them are based on our key concept of topic of information. We refine our previous work by contextualizing topics by mental attitudes of the agents.

We aim at a logic having both a complete axiomatization and proof procedure, and an effective implementation. This has motivated several choices, in particular a *Sahlqvist-type modal logic* (for which general completeness results exist) that is *monotonic* (contrarily to many approaches in the literature) and which has a notion of intention that is *primitive* (contrarily to the complex constructions in the literature).

In the next section we discuss the failure of the existing approaches to correctly handle belief change (section 2). Then we present an original

[1] We use 'set of speech acts' rather than 'a speech act', because a (literal) speech act may entail indirect speech acts. We develop this question in Herzig *et al.* (2000).

approach based on topics (section 3). This is embedded in a BDI framework (section 4). Finally we illustrate the approach by a complete treatment of our running example (section 5).

2 EXISTING APPROACHES

The most prominent formal analysis of belief change has been done in the AGM (Alchourrón *et al.* 1985) and the KM (Katsuno & Mendelzon 1992) frameworks. There, a belief change operator $\circ$ is used to define the new state $S \circ A$ from the previous state S and the input A.[2]

There are two difficulties if we want to use such a framework. First, until now, update operators have only been studied for classical propositional logic, and not for epistemic or doxastic logic.[3] But an appropriate theory of dialogues should precisely be about the change of beliefs about other agents' beliefs: an agent i believing that p and that another agent j believes p must be able to switch to believing that j believes $\neg p$, while maintaining his belief that p. Nothing is said about that in the current theories of revision and updates.

The second difficulty is that both revision and update have several common properties that must be refined or rejected. For example, the postulate $(S \circ A) \rightarrow A$ (input A has always priority) is problematic: in some approaches the new information may be rejected (as in Sadek's); in our approach, the new information is always accepted, but not all its consequences. We reject the postulate $(S \circ A) \leftrightarrow S$ if $S \rightarrow A$ because it neglects the over-informing nature of some information: our agents may have different behaviour in the cases of over-information.

In the rest of this section we review the logical analyses of belief change in dialogues that have been proposed in the literature. Due to the above difficulties to formalize belief change within the existing frameworks for revision or update, belief change is integrated into a formal theory of rational behaviour.

2.1 *Cohen & Levesque*

Cohen & Levesque (1990a, 1990c) have defined a formal theory of rational interaction where an agent may accept new pieces of information ('inputs'

[2] We view S as not closed under logical consequence. Therefore it can be confused with the conjunction of its elements. Just like Katsuno & Mendelzon (1992), we view $\circ$ as a (metalanguage) operator mapping the formulas S and A to a formula.

[3] Nevertheless, it is known in the belief revision literature that the AGM revision postulates must be considerably weakened if the language contains modalities (Fuhrmann's impossibility theorem (Fuhrmann 1989), (Hansson 1999: section 5.1)).

for short). In this approach, the input corresponds to the speaker's intention to obtain some effect rather than to the speech act itself. The hearer's belief adoption is conditioned by the speaker's sincerity. Their theory allows the agent both to change his beliefs and to reject the input (if the speaker is believed to be insincere).

However, as Sadek notes (Sadek 1991), even lies might generate some effects (for example, the hearer adds to his beliefs that the speaker is insincere). Thus even if the input is rejected, the mental state of the hearer evolves.

Finally, in Cohen & Levesque's approach beliefs not undermined by the act are never preserved from the preceding mental state to the new one (cf. the frame problem in Artificial Intelligence (McCarthy & Hayes 1969)). Thus inconsistency of the newly acquired beliefs with old ones is never the case, simply because old beliefs are given up by the agent. (Such a behaviour corresponds to what has been called the trivial belief change operation in the AGM and KM literature.)

2.2 *Perrault*

Perrault's system is based on Reiter's default logic (Reiter 1980). $A \Rightarrow B$ denotes a normal default. $Do_{\alpha,t}\top$ means that action α is performed at time t, $observe_{j,t}A$ means that agent j observes A at time t, and $\langle \mathsf{Assert}_{i,j}P \rangle$ means that agent i communicates propositional content P to agent j. The main axioms and default rules of Perrault's approach are as follows:

(1) memory: $Bel_{i,t}A \rightarrow Bel_{i,t+1}Bel_{i,t}A$

(2) persistence: $Bel_{i,t+1}Bel_{i,t}A \rightarrow Bel_{i,t+1}A$

(3) observability: $Do_{\alpha,t}\top \wedge Observe_{j,t}Do_{\alpha,t}\top \rightarrow Bel_{j,t+1}Do_{\alpha,t}\top$, where α is performed by the agent i

(4) belief transfer: $Bel_{i,t}Bel_{j,t}A \Rightarrow Bel_{i,t}A$

(5) assertion rule: $Do_{\langle \mathsf{Assert}_{i,j}A \rangle,t}\top \Rightarrow Bel_{i,t}A$

Moreover there is a default schema saying that if $A \Rightarrow B$ is a default then $Bel_{i,t}A \Rightarrow Bel_{i,t}B$ is also a default, for every agent i and timepoint t.

Here sincerity is not required in order to admit an act (as illustrated by axiom (3)). But an agent consumes its effects only if he doesn't yet believe the converse of this effect (in terms of defaults: if the effect is consistent with his current beliefs, cf. (5)). Thus the speaker does not have the right to lie, to make mistakes or to change his mind; otherwise the effect of his act will never be consumed (in technical terms, the default will be blocked).

This is at the origin of an even more radical behaviour: as highlighted in Appelt & Konolige (1989), Perrault's agents never question old beliefs, but only expand their mental state (in the sense of the AGM framework). Indeed, it follows from axioms (1) and (2) that $Bel_{i,t}A \rightarrow Bel_{i,t+1}A$. Consequently if a belief stemming from memory conflicts with a belief stemming from the act, then the default (5) will never been applied, and the effect will never be consumed.

Perrault is aware of that and suggests to achieve persistence by a default rule:

(6) Persistence (bis): $Bel_{i,t+1} Bel_{i,t} A \Rightarrow Bel_{i,t+1} A$

But as he notes himself, in this case there are always two extensions: one where the agent preserves his (old) beliefs and then adopts the input if it is consistent with these beliefs, and another one where the agent adopts the input and then preserves those old beliefs that are consistent with the new information. But there seems to be no way of determining which choice the agent should make.

Perrault's approach has some other problems that we do not discuss here (for example, if the speaker ignores whether A is the case, then he starts to believe it as soon as he utters that A, cf. Appelt & Konolige (1989)).

2.3 *Sadek*

Sadek defines a theory of rationality similar to Cohen & Levesque's, enriching it with two new mental attitudes, uncertainty and need (Sadek 1991, 1992). In his *belief reconstruction* (Sadek 1994), he presents an alternative to Perrault's approach. He enriches the latter's theory by an *axiom of admission*, and orders the application of his axioms of memory, admission, effects acceptance, and preservation. His axiom of admission describes the behaviours that can be adopted by an agent, but it does not specify the way the agent chooses between different possible behaviours. In particular he enables the hearer to reject an act. The latter point seems problematic to us, given that hearers do not reject an act that has been performed, but rather (hypothetically) accept it in order to conclude that it was not this one that has been performed.

2.4 *Rao & Georgeff*

In several papers, Rao & Georgeff have proposed theories and architectures for rational agents (Rao & Georgeff 1991). Such a theory can in principle be applied to dialogues. In Rao & Georgeff (1992), in a way similar to STRIPS, actions and plans are represented by their preconditions together with add- and delete-lists. The latter lists are restricted to sets of atomic formulas. In such a framework, one can a priori neither represent non deterministic actions nor actions with indirect effects (obtained through integrity constraints). Even more importantly, actions can only have effects that are factual: this excludes the handling of speech acts, whose effects are epistemic, and are typically represented by means of nested intensional operators (such as intentions to bring about mutual belief). Recently, they have defined a tableau proof procedure for their logic (Rao & Georgeff 1998).

2.5 *Appelt & Konolige*

Appelt & Konolige highlight the problems of Perrault's approach (Appelt & Konolige 1989). They propose to use hierarchic auto-epistemic logic (HAEL) as a framework. Basically, what one gains from this is that application of default rules can be ordered in a hierarchy. This can be used to fine-tune default application and thus avoid unwanted extensions.

Apart from the relatively complex HAEL technology, it appears that Appelt & Konolige's belief adoption criterion encounters problems similar to Perrault's. Suppose the hearer has no opinion about p. Now if the speaker informs the hearer that p, then under otherwise favourable circumstances the hearer adopts p. But if the speaker informs the hearer that *the hearer believes p* (or that he believes the hearer believes p), then it is clearly at odds with our intuitions that the hearer should accept such an assertion about his mental state. The only means to avoid the latter behaviour is to shift the hearer's ignorance about p to the level of the HAEL hierarchy that has priority (level 0). But in this case the acceptation of the assertion that p would be blocked as well.

3 TOPIC-BASED BELIEF CHANGE

3.1 *The modal language*

Like the previously cited authors, we work in a multimodal framework, with modal operators of belief, mutual belief, intention, and action. Our language is that of first-order multimodal logic without equality and without function symbols (Chellas 1980; Hughes & Cresswell 1972;

Popkorn 1994). We suppose that $\wedge$, $\neg$, $\top$ and $\forall$ are primitive, and that $\vee$, $\rightarrow$, $\perp$ and $\exists$ are defined as abbreviations in the usual way. Let AGT be the set of agents. For $i,j \in AGT$, the belief operators Bel_i and $Bel_{i,j}$ respectively stand for 'the agent i believes that' and 'it is mutual belief of i and j that'. For each $i \in AGT$, the intention operator $Intend_i$ stands for 'the agent i intends that'. In our running example, we use two particular agents, s and u, which stand for the *system* and the *user*.

Speech acts are represented by tuples of the form $\langle \mathsf{FORCE}_{i,j}A \rangle$ where FORCE is the illocutionary force of the act, i, $j \in AGT$, and A is the propositional content of the act. For example $\langle \mathsf{Inform}_{u,s} Dest(\text{Paris}) \rangle$ represents a declarative utterance of the user informing the system that the destination of his ticket is Paris. Let ACT be the set of all speech acts.

With every speech act $\alpha \in ACT$ we associate two modal operators $Done_\alpha$ and $Feasible_\alpha$. $Done_\alpha A$ is read 'speech act α has just been performed, before which A was true'; $Feasible_\alpha A$ is read 'speech act α is feasible, after which A will be true'.[4] In particular, $Done_\alpha \top$ and $Feasible_\alpha \top$ are respectively read 'α has just been performed' and 'α is feasible' (or 'can be performed'). Using the $Done_\alpha$ operator, the beliefs of the system at the state S_k can be kept in *memory* at state S_{k+1}: if B is the conjunction of all beliefs of the agent i at the (mental) state k, and α has just been done, then $Bel_i Done_\alpha B$ is the memory of i in the state $k+1$.

To express temporal properties, we define the *Always* operator, and its dual operator *Sometimes*. *Always A* means 'A always holds' and *Sometimes A* means 'A sometimes holds'. The operator *Always* will enable us in particular to preserve the domain laws in all states.

Formally, acts and formulas are defined by mutual recursion. This enables speech acts where the propositional contents is a non-classical formula. For example:

$Bel_s Done_{\langle \mathsf{Inform}_{u,s} Bel_u Bel_s p \rangle} Bel_s Bel_u Bel_s \neg p$ is a formula.

3.2 *The problem of belief change*

In our approach, unlike Sadek's, we always accept[5] speech acts, but we proceed in two steps: the agent accepts the indirect and intentional effects, but only adopts the speaker's beliefs if he believes the speaker to be competent at these beliefs. Thus, *speaker competence* is our criterion to determine which part of the input must be accepted by the hearer and which part must be rejected. For example, s accepts input about the new

[4] $Done_\alpha A$ et $Feasible_\alpha A$ are just as $\langle \alpha^{-1} \rangle A$ and $\langle \alpha \rangle A$ of dynamic logic (Harel 1984).

[5] 'Accepting' an act means that we admit that it has been performed.

class (after u_2) but rejects input about the price (after u_3), the reason being that he considers u to be competent at classes but not at ticket prices.

Which beliefs of the hearer can be preserved after the performance of a speech act? Our key concept here is that of the *influence of a speech act on beliefs*. If there exists a relation of influence between the speech act and a belief, this belief cannot be preserved in the new state. In our example, the old transport class cannot be preserved through u_2, because the act of informing about classes influences the hearer's beliefs about classes.

How can we determine the competence of an agent at beliefs and the influence of a speech acts on beliefs? The foundation of both notions will be provided by the concept of a *topic*: we start from the idea that with every agent, speech act, and formula some set of topics can be associated. Thus, an agent i is competent at a formula A if and only if the set of topics associated with A is a subset of the set of topics associated with i—the set of topics at which i is competent. And a formula A is preserved after the performance of a speech act α if A and α have no common topic, i.e. occurring both in the set of topics associated with A and in the set of topics associated with α. We give the formal apparatus in the rest of the section.

3.3 *Topic structures*

The concept of topic has been investigated both in linguistics and philosophical logic. For example, in Büring (1995) a semantical value related to the topics is associated with each English sentence. Van Kuppevelt has developed a notion of topic based on questions and has applied it to phenomena of intonation (van Kuppevelt 1991, 1995). In Ginzburg (1995), some sets of topics play a decisive role in the coherence of dialogues.

Several approaches to the notion of topic exist in the philosophical logic literature, in particular those of Lewis (1972) and Goodman (1961). Goodman's notion of 'absolute aboutness' is defined purely extensionally. Hence for him logically equivalent formulas are about the same topics, while this is not the case for us. Moreover, as he focuses on the 'informative aspect' of propositions, the subject of a tautology is the empty set.

Epstein's (1990) notion is quite different from Lewis's and Goodman's. He defines the *relatedness relation* $\mathcal{R}$ as a primitive relation between propositions because 'the subject matter of a proposition isn't so much a property of it as a relationship it has to other propositions' (Epstein 1990: 62). Thus, he does not represent topics explicitly. Then he defines the *subject matter of a proposition* A as $s(A) = \{\{A, B\} : \mathcal{R}(A, B)\}$. More precisely, s is called the *subject matter set-assignment associated with* $\mathcal{R}$ (Epstein 1990: 68). Epstein shows that we can also define s as primitive, and that we can then define two propositions as being related if they have some subject matter in common. Our subject function can be seen as an extension of this function to a multi-modal language.

For us, topics are themes in context, where the set of themes is an arbitrary set and contexts correspond to mental attitudes of agents. We define three functions associating topics to formulas, agents, and speech acts.

3.3.1 Themes, contexts, and topics

A *theme* is what something is about. For example, information on the destination is about the destination but not about the transport class.

Let $\mathcal{T} \neq \emptyset$ be a set that we call the set of themes. In our running example, we suppose that $\mathcal{T}$ contains destinations, classes, prices, and payment.

***Definition* 1** *Let $i \in AGT$. Then ma_i is called an* atomic context. *A context is a possibly empty sequence of atomic contexts. The empty context is noted λ. C is the set of all contexts.*

ma_i stands for 'the mental attitude of agent i'.

***Definition* 2** *A* topic of information (*or* contextual thematic structure) *is a theme together with a context, denoted by $c{:}t$, where $t \in \mathcal{T}$ and $c \in C$.*

For example, ma_u : *price* is a topic consisting in the user's mental attitude at prices, and ma_s : ma_u : *price* is a topic consisting in the system's mental attitude at the user's mental attitude at prices.

For the empty context λ, we have

(7) $\lambda : c = c : \lambda = c.$

By convention, we identify $\lambda : t$ with t. In order to take into account introspection, we postulate

(8) $ma_i{:}ma_i = ma_i.$

Given a set of themes and a set of agents we note $\mathbb{T}$ the associated set of topics. $\mathbb{T}_n$ is the set of topics whose contexts have length at most n. As we have identified $\lambda : t$ with t, $\mathbb{T}_0$ is the set of themes. In this paper, for reasons of representational economy we shall suppose that the length of each context is at most 2. Hence we restrict $\mathbb{T}$ to $\mathbb{T}_2$.[6]

Note that we have overloaded the operator ':'. As we only use λ, c,

[6] We did not find examples requiring length 3. Nevertheless, this restriction can be relaxed easily.

$ma_i, \ldots$ for contexts and only $t, t', \ldots$ for themes, there should be no confusion.

3.3.2 The subject of a formula

***Definition* 3** *The* subject of a formula A *is a set of topics associated with* A *(the topics* A *is about). This notion is formalised by a function* subject *mapping every formula to a set of topics from* $\mathbb{T}$.

We give the following axioms.

***Axiom* 1** $\mathtt{subject}(p) \subseteq \mathcal{T}$ *and* $\mathtt{subject}(p) \neq \emptyset$ *where* p *is atomic.*

An intuition that might be helpful is to think of the subject of p as the predicate name of p.

***Axiom* 2** $\mathtt{subject}(\top) = \emptyset$.

Note that this slightly differs from Epstein's account.[7]

***Axiom* 3** $\mathtt{subject}(\neg A) = \mathtt{subject}(A)$.

***Axiom* 4** $\mathtt{subject}(A \wedge B) = \mathtt{subject}(A) \cup \mathtt{subject}(B)$.

***Axiom* 5** $\mathtt{subject}(Bel_i A) = \{ma_i : c : t \mid c : t \in \mathtt{subject}(A)\}$.

Note that c might be the empty context here. Thus, in our running example:

$\mathtt{subject}(Class(1\mathrm{st})) = \{class\}$

$\mathtt{subject}(Dest(\mathrm{Paris})) = \{destination\}$

$\mathtt{subject}(Bel_s Bel_u Price(80\,€) \wedge Bel_s Price(100\,€))$

$= \{ma_s : ma_u : price\} \cup \{ma_s : price\}$.

***Axiom* 6**

$$\begin{aligned}\mathtt{subject}(Bel_{i,j} A) = {} & \mathtt{subject}(Bel_i A) \cup \\ & \mathtt{subject}(Bel_j A) \cup \\ & \mathtt{subject}(Bel_i Bel_{i,j} A) \cup \\ & \mathtt{subject}(Bel_j Bel_{i,j} A).\end{aligned}$$

***Axiom* 7** $\mathtt{subject}(Intend_i A) = \mathtt{subject}(Bel_i A)$.

[7] Indeed, Epstein stipulates that $\mathcal{R}(\mathcal{A}, \mathcal{A})$ for every formula A. On the contrary, the present axiom makes that $not(\mathcal{R}(\top, \top))$. More generally, we have $\mathcal{R}(\mathcal{A}, \mathcal{A})$ iff the set of atoms of A is nonempty. Due to the logical operators $\top$ and $\bot$ in the language we had to modify that.

Axiom 8 $\mathtt{subject}(Done_{\alpha}A) = \mathtt{subject}(A) \cup \mathtt{subject}(A')$ where A' is the propositional content of α.[8]

Axiom 9 $\mathtt{subject}(\forall x A) = \mathtt{subject}(A)$.

Axiom 10 $\mathtt{subject}(A[t/x]) \subseteq \mathtt{subject}(A)$, *where* $A[t/x]$ *is the formula resulting from the substitution of the variable x by the term t.*

This expresses that if an instance of A is about some topic, then A is about that topic as well.

Due to our restriction to contexts of length 2 we suppose that contexts of the form $ma_i : ma_j : c$ are reduced to $ma_i : ma_j$. The corresponding subject function can be obtained by first simplifying the topics by (7), (8), and the above equation $\lambda : t = t$; and then by reducing those of length greater than 2 to topics of length 2. For example,

$\mathtt{subject}(Bel_u Bel_u Price(150\,€)) = \{ma_u : price\}$,

and $\mathtt{subject}(Bel_u Bel_s Bel_k Price(150\,€)) = \{ma_u : ma_s : price\}$, for any agent k.

3.3.3 The competence of an agent

Definition 4 *The* competence of an agent *i is a set of topics associated with i (the competence of i). This notion is formalised by a function* $\mathtt{competence}$ *mapping every agent to a set of topics from* $\mathbb{T}$.

We assume every agent is competent at his mental states.

Axiom 11 $\mathtt{competence}(i) \supseteq \{ma_i : t \mid t \in \mathcal{T}\}$.

An agent may be competent at some facts. For example, $\mathtt{competence}(u)$ contains destinations and classes, but not prices.[9]

Competence will allow us to formulate in the next section our belief adoption axiom which basically says: 'an agent j adopts the belief of another agent i about a formula A if j considers that i is competent at the subject of A'.[10]

3.3.4 The scope of an act

Definition 5 *The* scope of a speech act α *is a set of topics associated with* α *(the scope of* α*). This notion is formalised by a function* $\mathtt{scope}$ *mapping every speech act to a set of topics from* $\mathbb{T}$.

[8] Another choice would have been $\mathtt{subject}(Done_{\alpha}A) = \mathtt{subject}(A) \cup \mathtt{scope}(\alpha)$. But this would too much mix up the reading of the subject function of 'being about something' with that of the scope function of 'modifying the truth value'.

[9] Note that an agent might be competent at mental attitudes of some other agent. This means that the former agent controls the latter. We do not exploit this further here.

[10] Hence competence should be a 2-argument function. As we only have two participants in our examples, we have dropped the second argument for the sake of simplicity.

Suppose the user informs the system about his destination. As the user is competent at destinations, this speech act influences the system's factual beliefs about the destination. It also influences its beliefs about prices, because a destination change possibly entails a price change. Hence $\mathsf{scope}(\langle \mathsf{Inform}_{u,s} Dest(\text{Paris})\rangle)$ contains the topics *destination*, *price*, ma_s : *destination* and ma_s : *price*.

The scope of a speech act determines which mental attitudes of an agent might be changed by this act.

In the formalization of speech acts the illocutionary force determines a set of formula schemes (the preconditions and the effects of the act) instantiated by the propositional content. The scope of a speech act is the set of topics associated with this act, and must depend on its illocutionary force and its propositional content.

Roughly speaking, the themes of a speech act are determined by its propositional content, and the context by its illocutionary force. Thus, contexts tell us which mental attitudes might change. We propose some axioms in order to compute the scope of a speech act.

The performance of a speech act always influences some mental attitudes of the hearer. In particular:

Axiom 12 $\mathsf{scope}(\langle \mathsf{FORCE}_{i,j} A\rangle) \supseteq \{ma_j : ma_i : t \mid t \in \mathsf{subject}(A)\}$, *for every illocutionary force* FORCE.

For example, consider the speech act where the user informs the system about the ticket price. This speech act influences the system's belief about the user's belief about prices.

Now consider the case where α is a request act. We postulate that the type of mental attitudes $ma_j : ma_i : t$ is the only one that is in the scope of α.

Axiom 13 $\mathsf{scope}(\langle \mathsf{Request}_{i,j} A\rangle) \subseteq \{ma_j : ma_i : t \mid t \in \mathcal{T}\}$.

3.3.5 Topic structures

We have thus defined three functions mapping the different types of expressions in our language to topics.

Definition 6 *Given a set of themes and a set of agents, a* topic structure *consists of the associated set of topics* $\mathbb{T}$ *together with the* subject, scope, *and* competence *functions.*

Is there an interaction between these functions? Consider the speech act $\alpha = \langle \mathsf{Inform}_{u,s} Class(\text{2nd})\rangle$. It follows from the axiom we gave for the scope function that $ma_s : ma_u : class \in \mathsf{scope}(\alpha)$. *Given that the user is* *competent* at classes, α also influences s's factual beliefs about the class, i.e. $ma_s : class \in \text{scope}(\alpha)$.

We propose the following constraint for acts of the informative type.

***Axiom* 14** *If A contains no modal operator, $\alpha = \langle \text{Inform}_{i,j} A \rangle$, and t is a theme such that $t \in \text{subject}(A) \cap \text{competence}(i)$ then $t \in \text{scope}(\alpha)$ and $ma_j : t \in \text{scope}(\alpha)$.*

Note that if this axiom is violated, then the mental state of the hearer might become inconsistent: suppose $ma_s : class \notin \text{scope}(\alpha)$. Then $Bel_s \neg Class(\text{2nd})$ would be preserved after α, while the indirect effect $Bel_s Bel_u Class(\text{2nd})$ of α would entail $Bel_s Class(\text{2nd})$ by the belief adoption axiom.

A given topic structure will allow us to compute the new state by means of two principles: belief adoption and preservation. In the next section we shall present these principles.

3.4 *Axioms for belief change*

Our axioms for belief change are based on a given topic structure. The first one allows one to preserve beliefs:

Axiom Schema of Belief Preservation

$$Done_\alpha A \rightarrow A \text{ if } \begin{cases} \text{scope}(\alpha) \cap \text{subject}(A) = \emptyset \text{ and} \\ A \text{ contains no } Done_\beta \text{ operator, for any } \beta. \end{cases}$$

The restriction to formulas without $Done_\beta$ operators is necessary because our reading of $Done_\beta$ is that β has just been performed (and not at some arbitrary time point in the past).

The second axiom schema allows one to adopt beliefs.

Axiom Schema of Belief Adoption

$Bel_i A \rightarrow A$ if $\text{subject}(A) \subseteq \text{competence}(i)$

The schema expresses that if agent i both believes that A and is competent at A, then A is true.

For example the formula $Bel_s Bel_u Dest(\text{Paris}) \rightarrow Bel_s Dest(\text{Paris})$ can be proved from the instance $Bel_u Dest(\text{Paris}) \rightarrow Dest(\text{Paris})$ of the belief adoption axiom. Indeed, the belief adoption axiom applies because

$\mathsf{subject}(Dest(\text{Paris})) \subseteq \mathsf{competence}(u)$, and we can then use the standard modal necessitation and K-principles for Bel_s. On the contrary, $Bel_u Price(80\,€) \to Price(80\,€)$ is not an instance of our axiom schema, because $\mathsf{subject}(Price(80\,€)) \not\subseteq \mathsf{competence}(u)$.[11]

3.5 *Discussion*

Our $\mathsf{subject}$ function is not extensional: logically equivalent formulas may have different topics. In particular, $\mathsf{subject}(p \vee \neg p) \neq \mathsf{subject}(\top)$. Indeed, $p \vee \neg p$ being an abbreviation of $\neg(\neg p \wedge \neg\neg p)$, we have $\mathsf{subject}(p \vee \neg p) = \mathsf{subject}(p) \neq \emptyset$, while $\mathsf{subject}(\top) = \emptyset$.[12]

It follows from our axioms that the subject of an arbitrary formula is completely determined by the subjects of its atomic formulas. This is representationally interesting, but it is certainly a debatable choice. Notwithstanding, the way we use the $\mathsf{subject}$ function is sound: suppose e.g. $\mathsf{subject}(p) = \{t\}$, $\mathsf{subject}(q) = \{t'\}$, and $\mathsf{scope}(\alpha) = \{t'\}$. Hence p and $p \wedge (q \vee \neg q)$ do not have the same subject, and $Done_\alpha p \to p$ is an instance of the preservation axiom, while $Done_\alpha(p \wedge (q \vee \neg q)) \to (p \wedge (q \vee \neg q))$ is not. But the latter formula can nevertheless be deduced from the former by standard modal logic principles: as $p \leftrightarrow p \wedge (q \vee \neg q)$ we have $Done_\alpha p \leftrightarrow Done_\alpha(p \wedge (q \vee \neg q))$. Hence the theorem $Done_\alpha p \to p$ enables us to deduce $Done_\alpha(p \wedge (q \vee \neg q)) \to (p \wedge (q \vee \neg q))$.

We did not formulate such strong compositionality axioms for the scope function. The reason is that a speech act might influence more than the topics of its propositional contents. For example, the scope of $\langle \mathsf{Inform}_{u,s} Class(1\text{st}) \rangle$ contains not only $ma_u : ma_s : class$ but also $ma_u : ma_s : price$. Our hypothesis here is that the scope of a speech act is determined by the subject of its propositional contents together with the integrity constraints (for example, linking destinations, classes, and prices). This is subject of ongoing research.

Finally, as we have mentioned, competence can be generalised in order to involve an agent j believing i to be competent at some topic. Then

[11] In our preceding approach (Fariñas del Cerro *et al.* 1998) we used non-contextualized topics to formulate axioms for belief change. This turned out to be too weak. Suppose the system believes p, and believes that the user believes p: $Bel_s p \wedge Bel_s Bel_u p$. Now suppose the user informs the system that he does not know whether p. Then the belief $Bel_s Bel_u p$ should go away, while $Bel_s p$ can be expected to be preserved. Hence the scope of this speech act should contain the system's attitudes towards the user's attitudes towards p, but not the system's attitudes towards p. We were not able to distinguish that before.

[12] Note also that this is the reason why we did not state, as is usually done in textbooks '$\top$ abbreviates $p \vee \neg p$, for some p', and instead added $\top$ to the primitive operators $\wedge$ and $\neg$ of our language.

our axiom schema would take the form $Bel_j(Bel_i A \rightarrow A)$ if $\mathrm{subject}(A) \subseteq \mathrm{competence}(j, i)$.

4 THE MULTIMODAL FRAMEWORK

4.1 *Axiomatics*

In this section we give the logical axiom and inference rule schemas. They are those of a normal modal logic of the Sahlqvist type (Sahlqvist 1975), for which general completeness results exist.

Just as in Cohen & Levesque (1990b), Perrault (1990), and Sadek (1991), with each belief operator we associate the (normal) modal logic KD45 (Halpern & Moses 1985). Thus we have the following schemas:

(RNBel) $\dfrac{A}{Bel_i A}$

(KBel) $Bel_i A \wedge Bel_i(A \rightarrow B) \rightarrow Bel_i B$

(DBel) $Bel_i A \rightarrow \neg Bel_i \neg A$

(4Bel) $Bel_i A \rightarrow Bel_i Bel_i A$

(5Bel) $\neg Bel_i A \rightarrow Bel_i \neg Bel_i A$

The rule schema of necessitation (RNBel) and the axiom schema (KBel) are in every normal modal logic, (DBel) is the 'axiom of rationality' (if i believes A then he does not believe $\neg A$), (4Bel) is the axiom of positive introspection (if i believes A then he believes that he believes A), and (5Bel) is the axiom of negative introspection (if i does not believe A then he believes that he does not believe A).

With each operator of mutual belief we associate the normal modal logic KD45, whose logical axioms are similar to these of belief operators. We suppose that mutual belief of i and j implies belief of both i and j, i.e. we have the logical axiom

(9) $Bel_{i,j}A \rightarrow (Bel_i A \wedge Bel_j A)$

To keep things simple we suppose that the logic of each operator of intention is the normal modal logic KD. (The inference rule (RNIntend) and the axioms (KIntend) and (DIntend) are just as (RNBel), (KBel), and (DBel).)

Obviously, our notions of mutual belief and intention are oversimplified: first, our condition (9), linking belief and mutual belief, is weaker than usual, where mutual belief $Bel_{i,j}A$ is identified with the infinite formula

$Bel_iA \wedge Bel_jA \wedge Bel_iBel_jA \wedge Bel_jBel_iA \wedge \ldots$ We argue that such an inductive construction is not necessary at least in a first approach: like, Cohen & Levesque, we suppose that mutual belief directly comes as the indirect effect of a speech act. (This is different e.g. from Perrault's view, where mutual belief is constructed via default rules. See Traum (1999, section 7.2.1) for a discussion of these issues.) Second, we offer no particular principle for intentions. We did this because the existing analyses of intention vary a lot, and the systems that have been put forward in the literature are rather complex. A normal modal logic for intention is too strong: for example, (KIntend) is not a theorem of Cohen & Levesque's logic (and neither is its converse).[13]

All $Done_\alpha$ and $Feasible_\alpha$ operators obey the principles of the (normal) modal logic K. As they are modal operators of 'possible' type, the rule of necessitation and the K-axiom take the form:

$$(\text{RNDone}) \quad \frac{\neg A}{\neg Done_\alpha A}$$

$$(\text{KDone}) \quad (\neg Done_\alpha A \wedge Done_\alpha B) \rightarrow Done_\alpha(\neg A \wedge B)$$

$$(\text{RNFeasible}) \quad \frac{\neg A}{\neg Feasible_\alpha A}$$

$$(\text{KFeasible}) \quad (\neg Feasible_\alpha A \wedge Feasible_\alpha B) \rightarrow Feasible_\alpha(\neg A \wedge B)$$

For example, the first rule means 'it is never the case that inconsistent formulas hold before action α'.

We suppose speech acts to be deterministic: their performance should lead to a single state. This is expressed by the converse (DC) of the modal axiom (D).[14]

$$(\text{DCDone}) \quad Done_\alpha A \rightarrow \neg Done_\alpha \neg A$$

$$(\text{DCFeasible}) \quad Feasible_\alpha A \rightarrow \neg Feasible_\alpha \neg A$$

For example, the last axiom says that there is only one way of executing α (and not one where A holds afterwards, and another one where $\neg A$ holds afterwards). The following conversion axioms (Van Benthem 1991) account for the interaction between the $Done_\alpha$ and $Feasible_\alpha$ operators:

$$(10) \quad Feasible_\alpha \neg Done_\alpha A \rightarrow \neg A$$

$$(11) \quad Done_\alpha \neg Feasible_\alpha A \rightarrow \neg A$$

[13] However, one can define intention operators in a minimal models semantics *à la* Chellas (1980: Ch. 7). This has been undertaken in Herzig & Longin (2000b) & Herzig *et al.* (2000).

[14] We recall that $Done_\alpha$ and $Feasible_\alpha$ are modal operators of type 'possible' (and not 'necessary').

The logic of the *Always* operator is the normal modal logic KT4. ($\mathrm{K_{Time}}$) and ($\mathrm{4_{Time}}$) are just as ($\mathrm{K_{Bel}}$) and ($\mathrm{4_{Bel}}$).

($\mathrm{T_{Time}}$) $AlwaysA \rightarrow A$

The dual to *Always* is *Sometimes*:

($\mathrm{Def_{Sometimes}}$) $SometimesA \overset{\mathrm{def}}{=} \neg Always \neg A$

In order to describe some interactions between the different mental attitudes (Cohen & Levesque 1990b), we propose the following logical axioms.

(12) $Intend_i A \rightarrow Intend_i Bel_i A$

(13) $Bel_i Intend_i A \leftrightarrow Intend_i A$

(14) $Bel_i \neg Intend_i A \leftrightarrow \neg Intend_i A$

(15) $Intend_i Bel_j A \rightarrow Bel_i A \vee Intend_i Bel_i A$

(16) $Bel_i Done_{\langle \mathsf{FORCE}_{i,j} A \rangle} \top \leftrightarrow Done_{\langle \mathsf{FORCE}_{i,j} A \rangle} \top$

The semantics of each of these logical axioms is given in Longin (1999) and Herzig & Longin (2000b).

4.2 *Laws*

Laws are non-logical axioms. We suppose that laws cannot be modified by the belief change process in a dialogue. We use the *Always* operator to preserve them in every state. We note *laws* the set of all laws (which might also be called our non-logical theory).

There are three kinds of laws: *static laws* (alias domain laws, similar to integrity constraints in data bases); *laws governing speech acts* (to describe the different preconditions and effects of the speech acts); *reactive laws* (to describe some reactive behaviours generating intentions).

4.2.1 Static laws

Some of the static laws are believed only by the system, such as those relating destinations, classes, and ticket prices:

(17) $AlwaysBel_s(Dest(\text{Paris}) \wedge Class(\text{1st}) \rightarrow Price(150\,€))$

(18) $AlwaysBel_s(Dest(\text{Paris}) \wedge Class(\text{2nd}) \rightarrow Price(100\,€))$

...

Some static laws are known both by the system and the user. More precisely, they are mutual beliefs:

(19) $AlwaysBel_{i,j}\neg(Class(1st) \wedge Class(2nd))$

(20) $AlwaysBel_{i,j}\neg(Dest(\text{Paris}) \wedge Dest(\text{New-York}))$

...

(There is only one class for a particular ticket, etc.)

4.2.2 Laws governing speech acts

Following Sadek (2000), we associate with each speech act

- a precondition;
- an indirect effect (the persistence of preconditions after the performance of the speech act);
- an intentional effect (in the Gricean sense (Grice 1967));
- a perlocutionary effect (expected effect).

Preconditions take the form $AlwaysBel_k\neg Done_\alpha\neg A'$ where A' is a precondition of α, and k an agent. Note that there is no constraint on k: k may be the speaker or some hearer (mutual belief). For example the precondition of an informative act is:

$$AlwaysBel_k\neg Done_{\langle \mathsf{Inform}_{i,j}A\rangle}\neg(Bel_iA \wedge \neg Bel_iBel_j\neg Bel_iA \wedge \neg Bel_iBelIf_jA \wedge \neg Bel_iBel_jBelIf_jA) \quad (21)$$

where $BelIf_jA$ is an abbreviation of $Bel_jA \vee Bel_j\neg A$.[15] (Preconditions and effects of our speech acts follow from (Sadek 1992, 2000).) The precondition means:

- the agent i believes A;
- i doesn't believe that j believes that he doesn't believe A (sincerity condition);[16]
- i doesn't believe that j knows if A holds or not;

[15] If we suppose that p must be either true or false (in the real world), and if $BelIf_jp$ holds, then j knows necessarily what is true in the real world (but *we* do not knows whether p is true or false). Then, $BelIf_jA$ is read 'j *knows* if A is true or not'. In KD45, Bel_jBelIf_jA is equivalent to $BelIf_jA$. In (21), we keep Bel_jBelIf_j because the precondition is a simplification of an infinite conjunction in the original precondition (Sadek 2000).

[16] The second term is an abbreviation of Sadek's infinite conjunction:
$\neg Bel_iBel_j\neg Bel_iA \wedge \neg Bel_iBel_jBel_i\neg Bel_iA \wedge \neg Bel_iBel_jBel_iBel_j\neg Bel_iA \wedge \ldots$

- i doesn't believe that j believes that j knows if A holds or not (condition of context relevance).[17]

From this law and the standard principles of normal modal logics we can prove formulas of the form $AlwaysBel_k(Done_\alpha \top \rightarrow Done_\alpha A')$, where A' is a precondition of α. For informative acts we have:

$$AlwaysBel_k(Done_{\langle \text{Inform}_{i,j} A \rangle} \top \rightarrow Done_{\langle \text{Inform}_{i,j} A \rangle}(Bel_i A \wedge$$

$$\neg Bel_i Bel_j \neg Bel_i A \wedge \neg Bel_i BelIf_j A)) \tag{22}$$

Suppose the user informs the system he wants a first class ticket. Then we have:

1. $Bel_s Done_{\langle \text{Inform}_{u,s} Class(1st) \rangle} \top$ (s believes the act has just been performed);
2. $Bel_s Done_{\langle \text{Inform}_{u,s} Class(1st) \rangle}(Bel_u Class(1st) \wedge \neg Bel_u Bel_s \neg Bel_u Class(1st) \wedge \neg Bel_u BelIf_s Class(1st))$
   (from 1, (22) with $k = j = s$, $i = u$, and principles of normal modal logics);
3. $Bel_u Done_{\langle \text{Inform}_{u,s} Class(1st) \rangle} \top$ (u believes the act has just been performed);
4. $Bel_u Done_{\langle \text{Inform}_{u,s} Class(1st) \rangle}(Bel_u Class(1st) \wedge \neg Bel_u Bel_s \neg Bel_u Class(1st) \wedge \neg Bel_u BelIf_s Class(1st))$
   (from 3, (22) with $k = i = u$, $j = s$, and principles of normal modal logics).

The formulas 2 and 4 are what we call *presuppositions*: immediately after a speech act α its observers believe that the preconditions of α were true just before the performance of this act. As illustrated by 4, the speaker is also viewed as an observer of his act. In this case, presuppositions are part of his memory (he remembers what he believed just before the performance of the speech act).

The indirect effect is the preservation of preconditions, and must be derived from presuppositions (cf. 2nd and 4th items in the above example) by formulas of the form $AlwaysBel_k(Done_\alpha A' \rightarrow A')$ where A' is a precondition of α: this will follow from our axiom schema of belief preservation (cf. section 3.4).[18]

The intentional effect is always accepted by the hearer and corresponds to the hearer's recognition of the speaker's intention (in Grice's sense).

[17] The second term is an abbreviation of Sadek's infinite conjunction:
$\neg Bel_i Bel_j BelIf_j A \wedge \neg Bel_i Bel_j Bel_i BelIf_j A \wedge \neg Bel_i Bel_j Bel_i Bel_j BelIf_j A \wedge \ldots$

[18] While the preservation of sincerity preconditions seems to be intuitively correct, it seems that the preservation of context relevance preconditions is an a priori choice of the agent, supposing that his act failed. In Herzig & Longin (2000b) we have proposed introducing a transient state of ignorance to overcome that.

The acceptance of this effect is expressed by formulas of the form $AlwaysBel_k(Done_\alpha \top \rightarrow A'')$, where A'' is the intentional effect of α. For an informative speech act the instance of this schema is:

(23) $AlwaysBel_k(Done_{\langle \mathsf{Inform}_{i,j} A \rangle} \top \rightarrow Intend_i Bel_j Intend_i Bel_j A)$

The perlocutionary effect does not obtain systematically: our agents being autonomous, the expected effect of an act does not obtain systematically. Hence the propositional content is not necessarily added to the hearer's belief state. In the case where the new state (obtained by the admission of a speech act and the acceptance of its indirect and intentional effects) entails the perlocutionary effect, we say that the latter has been accepted.

4.2.3 Reactive laws

The reactive laws allow us to generate some intentions:

(24) $AlwaysBel_i(A \wedge Bel_j \neg A \rightarrow Intend_i Bel_j A)$

(25) $AlwaysBel_i(A \wedge Done_{\langle \mathsf{Inform}_{j,i} A \rangle} Bel_i \neg A \rightarrow Intend_i Bel_j Bel_i A)$

(26) $AlwaysBel_i(Done_\alpha(Done_\gamma \top \wedge Bel_i Done_\beta \top)$

$$\rightarrow Intend_i Bel_j Bel_i Done_\alpha Done_\gamma \top) \qquad (26)$$

...

For example, (24) is used for the first part of the utterance s_4 in our running example: the system invalidates the price of 80€, and informs the user that the price is 100€.

Formally:

1. $Bel_s Price(100\,€)$	(hypothesis)
2. $Bel_s Bel_u Price(80\,€)$	(hypothesis)
3. $Bel_{s,u} \neg(Price(100\,€) \wedge Price(80\,€))$	(static law)
4. $Bel_s \neg Price(80\,€)$	(by 1 and 3)
5. $Intend_s Bel_u \neg Price(80\,€)$	(by (24), 4 and 2)
6. $Bel_s Bel_u \neg Price(100\,€)$	(by 2 and 3)
7. $Intend_s Bel_u Price(100\,€)$	(by (24), 1 and 6)

(We didn't give the logical axioms we use; s and u are the agents i and j in the law (24), respectively.) The intentions in 5 and 7 are associated with a denying speech act (the price isn't 80 €) and an informative act (the price is 100 €), respectively.

5 EXAMPLE

We illustrate our analysis of the belief change process by means of our running example. To each utterance we associate a speech act (for example, α_{u_i} corresponds to utterance u_i). We describe parts of the different states S_{s_i} during our example. These parts correspond to the mental state of the system after the different speech acts of the user.

The set of themes is $\mathcal{T} = \{class, destination, price, payment\}$.

The speech acts are:

$$\alpha_{u_1} = \langle \mathsf{Inform}_{u,s} Class(1\text{st}) \wedge Dest(\text{Paris}) \rangle$$

$$\alpha_{u_2} = \langle \mathsf{Inform}_{u,s} Class(2\text{nd}) \rangle$$

$$\alpha_{u_3} = \langle \mathsf{Inform}_{u,s} Price(80\,€) \rangle$$

$$\alpha_{u_4} = \langle \mathsf{ReqInformIf}_{u,s} Payment(\text{credit\_card}) \rangle.$$

The subjects of the atomic formulas are the predicate name, e.g. $\mathtt{subject}(Class(1\text{st})) = \{class\}$.

The scopes of the speech acts are:

- $\mathtt{scope}\, \alpha_{u_1} = \{ma_s : ma_u t, ma_s : t, t \mid t \in \{class, destination, price\}$;
- $\mathtt{scope}\, \alpha_{u_2} = \{ma_s : ma_u t, ma_s : t, t \mid t \in \{class, price\}$;
- $\mathtt{scope}\, \alpha_{u_3} = \{ma_s : ma_u : price\}$;
- $\mathtt{scope}\, \alpha_{u_4} = \{ma_s : ma_u : payment\}$.

The competence of the user and the system is:

- $\mathtt{competence}(u) = \{ma_u : t \mid t \in \mathcal{T}\} \cup \{destination, class\}$;
- $\mathtt{competence}(s) = \{ma_s : t \mid t \in \mathcal{T}\} \cup \{price, payment\}$.

We use the following abbreviations:

- $C1$ and $C2$ are $Class(1\text{st})$ and $Class(2\text{nd})$, respectively;
- $P1$, $P2$ and $P3$ are $Price(150\,€)$, $Price(100\,€)$ and $Price(80\,€)$, respectively;

We have simplified the preconditions of the speech acts.

α_{u_1} has the following effects.

1. performance of the act: $Bel_s Done_{\alpha_{u_1}} \top$
2. presuppositions: $Bel_s Done_{\alpha_{u_1}} (Bel_u(C1 \wedge Dest(\text{Paris})) \wedge \neg Bel_u BelIf_s(C1 \wedge Dest(\text{Paris})))$
3. indirect effect: $Bel_s Bel_u(C1 \wedge Dest(\text{Paris})) \wedge Bel_s \neg Bel_u BelIf_s(C1 \wedge Dest(\text{Paris}))$
4. intentional effect: $Bel_s Intend_u Bel_s Intend_u Bel_s(C1 \wedge Dest(\text{Paris}))$

5. reduction of intention:[19] $Bel_s Intend_u Bel_s(C1 \land Dest(\text{Paris}))$
6. adoption: $Bel_s(C1 \land Dest(\text{Paris}))$
7. application of static laws: $Bel_s P1$

α_{u_2} has the following effects.

1. performance of the act: $Bel_s Done_{\alpha_{u_2}} \top$
2. memory: $Bel_s Done_{\alpha_{u_2}}(Bel_s C1 \land Bel_s Dest(\text{Paris}) \land \ldots)$
3. presuppositions: $Bel_s Done_{\alpha_{u_2}}(Bel_u C2 \land \neg Bel_u BelIf_s C2)$
4. indirect effect: $Bel_s Bel_u C2 \land Bel_s \neg Bel_u BelIf_s C2$
5. intentional effect: $Bel_s Intend_u Bel_s Intend_u Bel_s C2$
6. reduction of intention: $Bel_s Intend_u Bel_s C2$
7. preservation: $Bel_s Bel_u Dest(\text{Paris}) \land$
   $Bel_s Intend_u Bel_s Intend_u Bel_s Dest(\text{Paris}) \land$
   $Bel_s Intend_u Bel_s Dest(\text{Paris}) \land Bel_s Dest(\text{Paris})$
8. adoption: $Bel_s C2$
9. application of static laws: $Bel_s P2$

α_{u_3} has the following effects.

1. performance of the act: $Bel_s Done_{\alpha_{u_3}} \top$
2. memory: $Bel_s Done_{\alpha_{u_3}}(Dest(\text{Paris}) \land C2 \land P2 \land Done_{\alpha_{u_2}}(Dest(\text{Paris}) \land$
   $C1 \land P1 \land \ldots) \land \ldots)$
3. presuppositions: $Bel_s Done_{\alpha_{u_3}}(Bel_u P3 \land \neg Bel_u BelIf_s P3)$
4. indirect effect: $Bel_s Bel_u P3 \land Bel_s \neg Bel_u BelIf_s P3$
5. intentional effect: $Bel_s Intend_u Bel_s Intend_u Bel_s P3$
6. reduction of intention: $Bel_s Intend_u Bel_s P3$
7. preservation: $Bel_s Bel_u C2 \land Bel_s \neg Bel_u BelIf_s C2 \land$
   $Bel_s Intend_u Bel_s Intend_u Bel_s C2 \land Bel_s Intend_u Bel_s C2 \land$
   $Bel_s Bel_u Dest(\text{Paris}) \land Bel_s Intend_u Bel_s Intend_u Bel_s Dest(\text{Paris}) \land$
   $Bel_s Intend_u Bel_s Dest(\text{Paris}) \land Bel_s Dest(\text{Paris}) \land Bel_s C2$
8. application of static laws: $Bel_s P2$
9. application of reactive laws: $Intend_s Bel_u \neg P3 \land Intend_s Bel_u P2$

α_{u_4} has the following effects.

1. performance of the act: $Bel_s Done_{\alpha_{u_4}} \top$
2. memory: $Bel_s Done_{\alpha_{u_4}}(Dest(\text{Paris}) \land C2 \land P2 \land \ldots)$
3. presuppositions: $Bel_s Done_{\alpha_{u_4}}(\neg BelIf_u Payment(\text{credit\_card}) \land$
   $Bel_u \neg Intend_s Done_{\langle \text{InformIf}_{s,u} Payment(\text{credit\_card})\rangle} \top)$
4. indirect effect: $Bel_s \neg BelIf_u Payment(\text{credit\_card}) \land$
   $Bel_s Bel_u \neg Intend_s Done_{\langle \text{InformIf}_{s,u} Payment(\text{credit\_card})\rangle} \top$

[19] This law is due to Sadek (1992), and has been reformulated in Longin (1999) as follows: $Bel_i Intend_j Bel_i A \rightarrow Bel_i A$ where $\text{subject}(A) = \{ma_j : c : t \mid c \in \mathcal{C}, t \in \mathcal{T}\}$.

5. intentional effect: $Bel_s Intend_u Bel_s Intend_u Done_{\langle \mathrm{InformIf}_{s,u} Payment(\mathrm{credit\_card})\rangle} \top$
6. reduction of intention: $Bel_s Intend_u Done_{\langle \mathrm{InformIf}_{s,u} Payment(\mathrm{credit\_card})\rangle} \top$
7. preservation: $Dest(\mathrm{Paris}) \wedge C2 \wedge P2 \wedge \ldots$
8. application of reactive laws: $Intend_s Done_{\langle \mathrm{InformIf}_{s,u} Payment(\mathrm{credit\_card})\rangle} \top$

6 DISCUSSION

We have sketched a theory of change in the context of dialogues. Our framework is based on the notion of topic of information, which is exploited through topic-based axioms of belief adoption and preservation. We think that our concepts are natural and appealing. It is intuitively clear that these two mechanisms permit to implement all possible evolutions of belief.

The framework can be augmented by other concepts such as that of sincerity can be added. The latter could be implemented in a way similar to competence.

Beyond the example dialogue given in the paper, we have tested our approach on a list of toy dialogues provided by France Télécom R&D Center.

In our running example, the propositional contents of the speech acts is rather simple. However, in Sadek (1991) and Longin (1999) there have been defined laws permitting to treat more complex propositional contents.

Note that in some applications it might be necessary to revise part of the competence of an agent. This happens in particular when it turns out that an agent has forgotten information he is competent at. Suppose for example, in u_4 the user says 'Hum, finally I'll pay cash that first-class ticket'. If we do not modify the competence function this case is handled as if u changed his mind about the train class: as u is competent at classes, s starts to believe that he now wants a first-class ticket again. What is needed here is to dynamically modify the competence function during the dialogue. This is possible in our framework. (As competence is a parameter of our logic, it amounts to modifying the logic.)

Perrault and Appelt & Konolige have argued that defaults are crucial elements in a theory of speech acts because they allow the transformation of absence of knowledge into knowledge. In a sense, what we do is to transfer that task to the metalinguistic relations of competence and scope. This permits to keep a monotonic framework, whose behaviour is considerably simpler and predictable than the nonmonotonic approaches of the literature.

We note that a possible worlds semantics for our multimodal logic can be given by adapting the one presented in Fariñas del Cerro *et al.* (1998) (see Herzig and Longin 2000b). Completeness can be proven in a fairly standard way. Indeed, all the semantical conditions are in a particular class that has been investigated in mathematical logic, and for which general completeness results exist (Sahlqvist 1975; Catach 1989; Gasquet 1994). The only difference here is that the preservation and adoption conditions depend on topics. It has been shown in Castilho *et al.* (1999) that nevertheless the standard Henkin proof technique applies straightforwardly.

We are currently implementing a tableau theorem prover for our logic. In previous work we have extended the standard tableaux method in order to deal with dependence information in reasoning about actions (Castilho *et al.* 1997, 1999). The extension of our approach to the present topic-based framework is straightforward.

Acknowledgements

This work has been supported by the France Télécom R&D Center, scientific area 'Intelligent Interactions and Dialogue', in the framework of contract 97 1B 046. Thanks to Renata Wassermann, David Sadek, and Jérôme Lang for relevant discussions. The detailed comments and suggestions of an anonymous referee of Amstelogue'99 and of the reviewers of the present special issue have (hopefully) contributed to improve the paper both in form and contents.

ANDREAS HERZIG and DOMINIQUE LONGIN
Institut de Recherche en Informatique de Toulouse (IRIT)
Université Paul Sabatier
118 Route de Narbonne, F-31062 Toulouse Cedex 4
France
e-mail: Andreas.Herzig@irit.fr, Dominique.Longin@irit.fr
http://www.irit.fr/ACTIVITES/LILaC/

Received: 01.09.1999
Final version received: 21.07.2000

REFERENCES

Alchourrón, C. E., Gärdenfors, P., & Makinson, D. (1985), 'On the logic of theory change: partial meet contraction and revision functions', *Journal of Symbolic Logic*, 50, 2.

Appelt, D. and Konolige, K. (1989), 'A nonmonotonic logic for reasoning about speech acts and belief revision', in M. Reinfrank, J. de Kleer, M. Ginsberg & E. Sandewall (eds), *Proc. of Second Int. Workshop on Non-Monotonic Reasoning*, volume 346 of *LNAI*, Springer-Verlag, 164–175.

Austin, J. L. (1962), *How To Do Things With Words*, Oxford University Press.

Benthem, J. van (1991), *The Logic of Time.* D. Reidel Publishing Company (1st edn, 1983).

Bretier, P. & Sadek, D. (1997), 'A rational agent as the kernel of a cooperative spoken dialogue system: Implementing a logical theory of interaction', in J. P. Müller, M. J. Wooldridge & N. R. Jennings (eds), *Proc. of the ECAI-96, Third Int. Workshop on Agent Theories, Architectures, and Languages (ATAL-96)*, LNAI: Intelligent Agents III, Springer-Verlag.

Büring, D. (1995), 'Topic', in P. Bosch & R. van der Sandt (eds), *The Focus Book*, Cambridge University Press.

Castilho, M. A., Gasquet, O., & Herzig, A. (1997), 'Modal tableaux for reasoning about actions and plans', in S. Steel & R. Alami (eds), *Proc. of European Conf. on Planning (ECP'97)*, number 1348 in LNAI, Springer-Verlag.

Castilho, M. A., Gasquet, O., & Herzig, A. (1999), 'Formalizing action and change in modal logic I: the frame problem', *Journal of Logic and Computation*.

Catach, L. (1989), 'Les logiques multimodales', Ph.D. thesis, Université Paris VI, France.

Chellas, B. F. (1980), *Modal Logic: an introduction*. Cambridge University Press.

Cohen, P. R. & Levesque, H. J. (1990a), 'Intention is choice with commitment', *Artificial Intelligence Journal*, 42, 2–3.

Cohen, P. R. & Levesque, H. J. (1990b), 'Persistence, intentions, and commitment', in P. R. Cohen, J. Morgan & M. E. Pollack (eds), *Intentions in Communication*, MIT Press.

Cohen, P. R. & Levesque, H. J. (1990c), 'Rational interaction as the basis for communication', in P. R. Cohen, J. Morgan, & M. E. Pollack (eds), *Intentions in Communication*, MIT Press.

Epstein, R. L. (1990), *The Semantic Foundations of Logic Volume 1: Propositional Logic*, Kluwer Academic Publishers.

Fariñas del Cerro, L., Herzig, A., Longin, D., & Rifi, O. (1998), 'Belief reconstruction in cooperative dialogues', in F. Giunchiglia (ed.), *Proc. of the 8th Int. Conf. on Artificial Intelligence: Methodology, Systems, and Applications (AIMSA '98)*, volume 1480 of *LNAI*, Springer-Verlag.

Fuhrmann, A. (1989), 'Reflective modalities and theory change', *Synthese*, 81, 115–34.

Gasquet, O. (1994), 'Déduction automatique en logique multi-modale par traduction, Ph.D. thesis, IRIT, Université Paul Sabatier, Toulouse, France.

Ginzburg, J. (1995), 'Resolving questions I, II', *Linguistics and Philosophy*, 18, 567–609.

Goodman, N. (1961), 'About', *Mind*, **LXX**, 277.

Grice, H. P. (1967), 'Utterer's meaning and intentions', *Philosophical Review*, 78, 2.

Grice, H. P. (1975), 'Logic and conversation', in J. P. Cole & J. L. Morgan (eds), *Syntax and Semantics: Speech Acts*, Academic Press.

Halpern, J. Y. & Moses, Y. (1985), 'A guide to the modal logics of knowledge and belief: preliminary draft', in *Proc. of Int. Joint Conf. on Artificial Intelligence (IJCAI '85)*, Morgan Kaufmann Publishers.

Hansson, S. O. (1999), *A Textbook of Belief Dynamics*, Kluwer Academic Publishers.

Harel, D. (1984), 'Dynamic logic', in D. Gabbay & F. Guenthner (eds), *Handbook of Philosophical Logic, Vol. II*, D. Reidel Publishing Company.

Herzig, A. & Longin, D. (2000b), 'Beliefs, intentions, speech acts and topics', Technical Report 00-08-R, IRIT, http://www.irit.fr/ACTIVITES/LILaC.

Herzig, A., Longin, D., & Virbel, J. (2000), 'Towards an analysis of dialogue acts and indirect speech acts in a BDI framework', in M. Poesio & D. Traum (eds), *Proc. of the Fourth Int. Workshop on the Semantics and Pragmatics of Dialogue (Götelog-2000)*, Department of Linguistics, Göteborg University, Sweden, 93–8.

Hughes, G. & Cresswell, M. (1972), *An Introduction to Modal Logic*, 2nd edn, Methuen.

Katsuno, H. & Mendelzon, A. O. (1992), 'On the difference between updating a knowledge base and revising it', in P. Gärdenfors (ed.), *Belief Revision*, Cambridge University Press.

Kuppevelt, J. van (1991), 'Topic en comment', Ph.D. thesis, University of Nijmegen.
Kuppevelt, J. van (1995), 'Discourse structure, topicality and questioning', *Linguistics*, 31, 109–47.
Lewis, D. (1972), 'General semantics', in D. Davidson & G. Harman (eds), *Semantics of Natural Language*, D. Reidel Publishing Company.
Longin, D. (1999), 'Interaction rationnelle et évolution des croyances dans le dialogue: une logique basée sur la notion de topique, Ph.D. thesis, IRIT, Université Paul Sabatier, Toulouse, France, http://www.irit.fr/ACTIVITES/LILaC/.
McCarthy, J. & Hayes, P. (1969), 'Some philisophical problems from the standpoint of artificial intelligence', in B. Meltzer & D. Michie (eds), *Machine Intelligence, volume 4*, Edinburgh University Press, 463–502.
Perrault, C. R. (1990), 'An application of default logic to speech act theory', in P. R. Cohen, J. Morgan & M. E. Pollack (eds), *Intentions in Communication*, MIT Press.
Popkorn, S. (1994), *First Steps in Modal Logic*, Cambridge University Press.
Rao, A. S. & Georgeff, M. P. (1991), 'Modeling rational agents within a BDI-architecture', in J. A. Allen, R. Fikes, & E. Sandewall (eds), *Proc. of the Second Int. Conf. on Principles of Knowledge Representation and Reasoning (KR '91)*, Morgan Kaufmann Publishers, 473–84.
Rao, A. S. & Georgeff, M. P. (1992), 'An abstract architecture for rational agents', in B. Nebel, C. Rich & W. Swartout (eds), *Proc. of the Third Int. Conf. on Principles of Knowledge Representation and Reasoning (KR '92)*, Morgan Kaufmann Publishers, 439–49.
Rao, A. S. & M. P. Georgeff, (1998), 'Decision procedures for BDI logics', *Journal of Logic and Computation*, 8, 3, 293–343. *Special Issue: Computational and Logical Aspects of Multiagent Systems*, Oxford University Press.
Reiter, R. (1980), 'A logic for default reasoning', *Artificial Intelligence Journal*, 13, 81–132.
Sadek, D., Bretier, P., & Panaget, F. (1997), 'ARTIMIS: Natural dialogue meets rational agency', in M. E. Pollack (ed.), *Proc. of the 15th Int. Joint Conf. on Artificial Intelligence (IJCAI '97)*, Morgan Kaufmann Publishers, 1030–5.
Sadek, M. D. (1991), 'Attitudes mentales et interaction rationnelle: vers une théorie formelle de la communication', Ph.D. thesis, Université de Rennes I, Rennes, France.
Sadek, M. D. (1992), 'A study in the logic of intention', in B. Nebel, C. Rich & W. Swartout (eds), *Proc. of the Third Int. Conf. on Principles of Knowledge Representation and Reasoning (KR '92)*, Morgan Kaufmann Publishers, 462–73.
Sadek, M. D. (1994), 'Towards a theory of belief reconstruction: application to communication', *Speech Communication Journal '94, Special Issue on Spoken Dialogue*, 15, 3–4.
Sadek, M. D. (2000). 'Dialogue acts are rational plans', in M. Taylor, F. Néel & D. Bouwhuis (eds), *The Structure of Mutimodal Dialogue*, John Benjamins Publishing Company, Philadelphia/Amsterdam, 167–88. From ESCA/ETRW, Workshop on The Structure of Multimodal Dialogue (Venaco II) (1991).
Sadek, M. D., Ferrieux, A., Cozannet, A., Bretier, P., Panaget, F., & Simonin, J. (1996), 'Effective human-computer cooperative spoken dialogue: the AGS demonstrator', in *Proc. of Int. Conf. on Spoken Language Processing (ICSLP '96)*, Philadelphia, USA.
Sahlqvist, H. (1975), 'Completeness and correspondence in the first and second order semantics for modal logics', in S. Kanger (ed.), *Proc. of 3rd Scandinavian Logic Symposium, Vol. 82, Studies in Logic*.

Searle, J. R. (1969), *Speech Acts: An Essay in the Philosophy of Language*, Cambridge University Press.

Traum, David R. (1999), 'Speech acts for dialogue agents', in M. Wooldridge and A. Rao (eds), *Foundations of Rational Agency*, Kluwer Academic Publishers, 169–201.