Belief Dynamics in Cooperative Dialogues*

Andreas Herzig and Dominique Longin†

Institut de Recherche en Informatique de Toulouse

Abstract

The context of this work is man-machine dialogues where each actor can interact with his environment, with the aim of communicating or acquiring information. We suppose that during the dialogue each participant can mistake a piece of information, forget it or simply change his mind.

We begin with some hypotheses about the actors, and present some approaches. We point out some shortcomings of the latter and introduce a new framework for belief change. We present our logical and non-logical theory; our basic notion is contextual topic: we suppose that we can associate a set of topics to every agent, speech act and formula. This allows to talk about the competence of an agent, the belief adoption and the belief preservation. We show how an agent's belief state can be define after a speech act and give a basic example.

1 Introduction

We focus on task-oriented dialogues. Participants in such dialogues have only one common goal, viz. to achieve the task under concern. Each of the participants has some information necessary to achieve the goal, but none of them can achieve it alone. For example suppose there is a system that delivers train tickets to users. The system cannot do that without user information 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 hypothesis. Informally, a person is cooperative w.r.t. 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]). Cooperation is a useful hypothesis in our context. For example, if the user wants a train ticket, then the system will intend to give it to him. If the system needs some piece of information to print the ticket, then the user answers the questions asked by the system.

Each participant is supposed to be sincere, i.e. 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 cannot be explained in terms of lies. Such an assumption is much weaker than in other approaches, where sincerity is

*This work has been supported in part by the Centre National d'Études des Télécommunications (CNET) of France Telecom, in the framework of their Audiotel project.

†Thanks to Renata Wassermann, David Sadek and Jerôme Lang for relevant discussions. We also wish to thank an anonymous referee of Amstelogue'99 for his comments.

sometimes viewed as the criterion of adoption of the incoming information [Cohen and Levesque, 1990c, for example].

Under these hypotheses we aim at describing the evolution of the mental state of a rational agent participating in a conversation[1] . As we illustrate by an example below, in a conversation agents might change their mind, make mistakes, understand wrongly, ... Since the agents must interact with each other in order to achieve the goal of the dialogue (by our cooperation hypothesis), they are the victims of such phenomena. Consequently these phenomena must be taken into account when modelling the evolution of mental states. In the sequel we call *belief change* the process leading an agent from a mental state to a new one.

Probably, the most prominent formal analysis of belief change has been done in the AGM [Alchourrón *et al.*, 1985] and the KM [Katsuno and Mendelzon, 1992] frameworks. There, a belief change operator is used to define the new state from the previous one and a new piece of information (carried by an utterance). The distinction between these two is respectively based on different hypotheses on the nature of the incoming information: in the case of revision, the real world does not change (the agent just gets new information on a static world), while in the case of update it is supposed to correspond to an event that occurred in the real world. This distinction is reflected by different postulates and different semantics.

There are two reasons not to use such a framework. The first is that revision and update are incompatible [Herzig, 1998, Theorem 22], forcing us to (perhaps dynamically) choose between them. Such a choice supposes that the agent is able to recognize whether the incoming information corresponds to real world[2] change or not. Thus, if the speaker changes his mind then the hearer should update his beliefs, and if the speaker informs the hearer about his intentions (e.g. destination and transportation class in the train ticket scenario) then the hearer should revise his beliefs: he discovers the intentions of the speaker in the course of the dialogue, and he may have made wrong inferences due to misunderstanding. But it seems that he is not able in general to distinguish between these two cases.

The second reason is that both revision and update have several common properties that are not suitable in dialogues: the new information is always prioritary, the over-informing nature of some information is neglected, no distinction is made between different levels of belief (factual, introspective, alternating, ...).

Some other approaches consider each participant as a *rational agent* having *mental states* represented by different mental attitudes such as belief, knowledge, choice, goal, intention... There, belief change takes place within a formal *rational balance theory* and a formal *rational interaction theory* [Cohen and Levesque, 1990a, 1990c]. These approaches are based on a generalization of the linguistic activity in a theory of action: they are the basis 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], and belief change (entailed by these speech acts on the agent's mental state) must be understood and handeld as a consequence of these speech acts.

We start from the latter framework. Each participant is viewed as an autonomous agent interacting with his environment and entertaining some mental attitudes which in turn influence its behavior and its decisions. Utterances are represented by speech acts [Searle, 1969, Searle and Vanderveken, 1985]. From the agent's point of view, the dialogue is a sequence of speech acts $(\alpha_1, \ldots, \alpha_n)$, where each α_{k+1} maps a mental state S_k to a new mental state S_{k+1}: nd

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

S_0 is the agent's initial belief state (before the dialogue starts). Given S_k and α_{k+1}, our task is to construct the new mental state S_{k+1}. 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 topics of information. We refine our previous work by contextualizing topics by mental attitudes of the agents.

The following dialogue is our running example to illustrate different problems and our solutions. There are only two agents, the system s and the user u. This example is relevant only if we take into account some laws about the world (e.g., a particular train ticket has only one destination, one class and one price, ...). In our approach, these laws are called *static laws* (Sect. 3.3):

s_1 : *Hello. What do you want?*
u_1 : *A first class train ticket to Amsterdam, 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 : ...

Within such dialogues we are particularly interested in the point of view of the machine, i.e. we study the evolution of the system's beliefs given an utterance of the user. In this case, the system:

- generally accepts some type of information (e.g. information about destination and class in train ticket selling scenarios – cf. u_1 below);
- derives supplementary information that was not directly contained in the utterance (e.g. the price if the user informs about his destination and class – s_2);
- may accept information contradicting its own beliefs such as in the case where the user changes his mind (e.g. switching from a first-class ticket to a second class ticket – u_2);
- generally preserves information it believed before the utterance (e.g. the system preserves the destination even when the class changes – from s_2 to s_3);
- 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 – s_4).

We focus on the evolution of the mental state of the system s, from a hearer's point of view. s has two complementary tasks: (1) dealing with contradictions between his mental state and the new piece of information (or *input* for short), and (2) preserving his old beliefs that do not contradict this input.

In the next section we discuss the failure of the existing approaches to correctly handle such changes of mental states (Sect. 2). Then we introduce our modal framework (Sect. 3). In this framework we present an original approach based on topics (Sect. 4). Finally we illustrate the approach by a complete treatment of our running example (Sect. 5).

2 Existing approaches

In this section we review the logical analyses of belief change in dialogues that have been proposed in the literature.

2.1 Cohen and Levesque

Cohen and Levesque have defined in [Cohen and Levesque, 1990a, 1990c] a formal theory of rational interaction where an agent may accept new pieces of information. 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 unsincere).

In our view sincerity is not a relevant criterion for belief change. While such a hypothesis permits to explain contradictions (in terms of lies), it may lead to very naive agents: whatever the speaker utters, the hearer will believe it as soon as the speaker is believed to be sincere by the hearer.

Moreover and as Sadek notes [Sadek, 1991a], even lies might generate some effects (e.g. the hearer adds to his beliefs that the speaker is unsincere). Thus even if the input is rejected the mental state of the hearer evolves.

Finally their system says nothing about the way the beliefs not undermined by the act can be preserved from the preceding mental state to the new one. (This is nothing else the frame problem, well-known in Artificial Intelligence [McCarthy and Hayes, 1969].) Thus inconsistency of the newly acquired beliefs with old ones is never the case, just because the latter are simply 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, $Observe_{j,t}$ means that agent i observes at time t, and $\langle \mathsf{Assert}_{i,j}\ P\rangle$ means that agent i communicates propositional content P to agent j.

$$\text{memory: } Bel_{i,t}A \rightarrow Bel_{i,t+1}Bel_{i,t}A \tag{1}$$

$$\text{persistence: } Bel_{i,t+1}Bel_{i,t}A \rightarrow Bel_{i,t+1}A \tag{2}$$

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

$$\text{belief transfer: } Bel_{i,t}Bel_{j,t}A \Rightarrow Bel_{i,t}A \tag{4}$$

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

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 the axiom (3)). But an agent consumes its effects only if he doesn't believe the converse of this effect yet (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, else 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 and Konolige, 1989], Perrault's agents never question old beliefs and expand their mental state (in the sense of the AGM framework). Indeed, it follows from the 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 :

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

As he notes, 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 determine which choice the agent should make.

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

2.3 Other approaches

Several authors have proposed similar solutions. Some use monotonic logics (Sadek, Rao and Georgeff), and some nonmonotonic (Appelt and Konolige).

Sadek defines a theory of rationality similar to that of Cohen and Levesque, enriching it with two new mental attitudes, viz. uncertainty and need [Sadek, 1991a, 1992]. In his *belief reconstruction* [Sadek, 1994], he presents an alternative to Perrault's approach, enriching the latter's theory by an *axiom of admission*, and ordering the application of his axioms of memory, admission, consumption and preservation. His axiom of admission describes the behaviours that can be adopted by an agent, but 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 accept it in order to derive that it was not this one that has been performed. An implementation has been done at CNET (France Telecom), and make up the kernel of several applications [Sadek *et al.*, 1997].

In several papers, Rao and Georgeff have proposed theories and architectures for rational agents [Rao and Georgeff, 1991]. Such a theory can in principle be applied to dialogues. In [Rao and 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 nondeterministic 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 typically represented by means of nested intensional operators (such as intentions to bring about mutual belief). Recently, they defined a tableau proof procedure for their logic [Rao and Georgeff, 1998].

Appelt and Konolige highlight the problems of Perrault's approach [Appelt and Konolige, 1989]. They propose to use hierarchic autoepistemic logic (HAEL) as a framework. Basically, what one gains there is that defaults can be stratified. This can be used to fine-tune default application and thus avoid unwanted extensions.

Apart from the relatively complex HAEL technology, it seems that Appelt and Konolige's belief adoption criterion encounters problems similar to Perrault's. Suppose the hearer has no opinion about p. 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 seems to be to shift the hearer's ignorance about p to the prioritary level 0 of the HAEL hierarchy. But in this case the acceptation of the assertion p would be blocked as well.

3 The multi-modal framework

3.1 Language

Like the previously cited authors, we work in a multi-modal framework, with modal operators for belief, mutual belief, intention and action. Our language is that of first order multi-modal logic without equality and without function symbols [Chellas, 1980, Hughes and Cresswell, 1972, Popkorn, 1994]. We suppose that $\wedge$, $\neg$, $\top$ and $\forall$ are primitive, and that $\vee$, $\rightarrow$, $\bot$ 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 "the agents i and j mutually believe 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 [Austin, 1962, Searle, 1969] 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 $\langle \mathsf{Inform}_{u,s}\ Dest(\text{Amsterdam}) \rangle$ represents a declarative utterance of the user informing the system that the destination of his ticket is Amsterdam. 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".[3] 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 mental 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 $k+1$.) 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.

To express temporal properties, we define the *Always* operator, and his dual operator *Sometimes*. $Always A$ means "A always holds" and $Sometimes A$ means "A sometimes holds". The operator *Always* allows in particular to preserve the laws in all mental states.

3.2 Axioms

Just like in [Cohen and Levesque, 1990b, Perrault, 1990, Sadek, 1991a], with each belief operator we associate the (normal) modal logic KD45 [Halpern and Moses, 1985]. Thus, for all agent $i \in ACT$ and all formulas A and B:

$$\frac{A}{Bel_i A} \tag{RN$_{\text{Bel}}$}$$

$$Bel_i A \wedge Bel_i(A \to B) \to Bel_i B \tag{K$_{\text{Bel}}$}$$

$$Bel_i A \to \neg Bel_i \neg A \tag{D$_{\text{Bel}}$}$$

$$Bel_i A \to Bel_i Bel_i A \tag{4$_{\text{Bel}}$}$$

$$\neg Bel_i A \to Bel_i \neg Bel_i A \tag{5$_{\text{Bel}}$}$$

are logical axioms: (RN$_{\text{Bel}}$) and (K$_{\text{Bel}}$) are axioms of every normal modal logic, (D$_{\text{Bel}}$) is the "axiom of rationality" (if i believes A then he doesn't believes $\neg A$), (4$_{\text{Bel}}$) is the positive introspection (if i believes A then he believes that he believes A), and (5$_{\text{Bel}}$) is the negative introspection (if i doesn't believe A then he believes that he doesn't believes 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 common belief is related to belief by the logical axiom

$$Bel_{i,j} A \to Bel_i A \tag{7}$$

To keep things simple we suppose that the logic of each operator of intention is the normal modal logic KD. (The axioms (RN$_{\text{Intend}}$), (K$_{\text{Intend}}$) and (D$_{\text{Intend}}$) are similar to (RN$_{\text{Bel}}$), (K$_{\text{Bel}}$) and (D$_{\text{Bel}}$).)

Obviously, our notions of common belief and intention are oversimplified: first, our condition linking belief and common belief is weaker than the usual induction axiom. We argue that such an inductive principle is not necessary at least in a first approach: as Cohen and Levesque, we suppose that common belief directly comes as the indirect effect of a speech act. (This is different from Perrault's view, where mutual belief is constructed via default rules.) 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, (K$_{\text{Intend}}$) is not a theorem of Cohen and Levesque's logic (and neither is its converse). However, we think we are able to define intention operators in a minimal models semantics [Chellas, 1980, Chap. 7].

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:

$$\frac{\neg A}{\neg Done_\alpha A} \tag{RN$_{\text{Done}}$}$$

$$(\neg Done_\alpha A \wedge Done_\alpha B) \to Done_\alpha(\neg A \wedge B) \tag{K$_{\text{Done}}$}$$

$$\frac{\neg A}{\neg Feasible_\alpha A} \tag{RN$_{\text{Feasible}}$}$$

$$(\neg Feasible_\alpha A \wedge Feasible_\alpha B) \to Feasible_\alpha(\neg A \wedge B) \tag{K$_{\text{Feasible}}$}$$

We want speech acts to be deterministic in order that performing it should lead to a single mental state. It is classical to represent such a determinism by the converse of the modal axiom (D). Thus, for all $\alpha \in ACT$ and every formula A:

$$Done_\alpha A \rightarrow \neg Done_\alpha \neg A \qquad (\text{DC}_{\text{Done}})$$

$$Feasible_\alpha A \rightarrow \neg Feasible_\alpha \neg A \qquad (\text{DC}_{\text{Feasible}})$$

The following logical axioms illustrate the interaction between the $Done_\alpha$ and $Feasible_\alpha$ operators:

$$Done_\alpha \neg Feasible_\alpha A \rightarrow \neg Feasible_\alpha Done_\alpha A \qquad (8)$$

$$Feasible_\alpha \neg Done_\alpha A \rightarrow \neg Done_\alpha Feasible_\alpha A \qquad (9)$$

The logic of the *Always* operator is the normal modal logic KT4. (K_{Time}) and (4_{Time}) are similar to (K_{Bel}) and (4_{Bel}):

$$Always A \rightarrow A \qquad (\text{T}_{\text{Time}})$$

and

$$Sometimes A \stackrel{\text{def}}{=} \neg Always \neg A \qquad (\text{Def}_{\text{Sometimes}})$$

In order to describe some interactions between the different mental attitudes [Cohen and Levesque, 1990b], we propose the following logical axioms. For all agent $i, j \in AGT$ and every formula A:

$$Intend_i A \rightarrow Intend_i Bel_i A \qquad (10)$$

$$Bel_i Intend_i A \leftrightarrow Intend_i A \qquad (11)$$

$$Bel_i \neg Intend_i A \leftrightarrow \neg Intend_i A \qquad (12)$$

$$Intend_i Bel_j A \rightarrow Bel_i A \vee Intend_i Bel_i A \qquad (13)$$

$$Bel_i Done_{\langle \mathsf{F}_{i,j}\ A \rangle} \top \leftrightarrow Done_{\langle \mathsf{F}_{i,j}\ A \rangle} \top \qquad (14)$$

We give the semantics of every logical axiom in [Longin, 1999].

3.3 Non-logical laws

The laws are non-logical axioms. These axioms musn't be modified by the belief change process. They are preserved in every mental state by the *Always* operator. Let *LAWS* be the set of all laws.

Our non-logical theory contains tree kinds of laws: *static 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).

Static laws. Some of the static laws are believed only by the system. It's the case for the prices of tickets according to the destination and the class, for example:

$$Always Bel_s (Dest(\text{Amsterdam}) \wedge Class(\text{1st}) \rightarrow Price(150\,€)) \qquad (15)$$

$$Always Bel_s (Dest(\text{Amsterdam}) \wedge Class(\text{2nd}) \rightarrow Price(100\,€)) \qquad (16)$$

...

Some others are knew both by the system and the user:

$$Always Bel_{i,j} \neg (Class(\text{1st}) \wedge Class(\text{2nd})) \qquad (17)$$

$$Always Bel_{i,j} \neg (Dest(\text{Toulouse}) \wedge Dest(\text{Amsterdam})) \qquad (18)$$

...

Laws governing speech acts. Following Sadek [Sadek, 1991b], we associate with each speech act

- a precondition;
- an indirect effect (viz. 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 are represented by formulas of the form $AlwaysBel_k \neg Done_\alpha \neg A'$ where A' is a precondition of α, and k an agent (there is no constraint on k: it may be the speaker, or a hearer). For example, such a law for informative acts is (preconditions and effects of our speech acts follows from [Sadek, 1991a, 1991b]):

$$AlwaysBel_k \neg Done_{\langle \mathsf{Inform}_{i,j}\ A\rangle} \neg (Bel_i A \wedge \neg Bel_i Bel_j \neg Bel_i A \wedge \neg Bel_i BelIf_j A \wedge \neg Bel_i Bel_j BelIf_j A) \tag{19}$$

where $BelIf_j A$ is an abbreviation for $Bel_j A \vee Bel_j \neg A$. The precondition means: the agent i believes A, and he doesn't believes that j believes that he doesn't believes A (sincerity condition[4]), and he doesn't believes that j knows if A holds or not, and he doesn't believes that j believes that j knows if A holds or not (condition of relevance to the context[5]).

From this law and the standard principles of our logic, it follows formulas of the form $AlwaysBel_k(Done_\alpha \top \rightarrow Done_\alpha A')$ where A' is a precondition of α. For example:

$$AlwaysBel_k(Done_{\langle \mathsf{Inform}_{i,j}\ A\rangle} \top \rightarrow Done_{\langle \mathsf{Inform}_{i,j}\ A\rangle}(Bel_i A \wedge \neg Bel_i Bel_j \neg Bel_i A \wedge \neg Bel_i BelIf_j A \wedge \neg Bel_i Bel_j BelIf_j A)) \tag{20}$$

Suppose that the user informs the system he would like a first class ticket. Then:

1. $Bel_s Done_{\langle \mathsf{Inform}_{u,s}\ Class(1\text{st})\rangle} \top$ (the act has just been performed) ;
2. $Bel_s Done_{\langle \mathsf{Inform}_{u,s}\ Class(1\text{st})\rangle}(Bel_u Class(1\text{st}) \wedge \neg Bel_u Bel_s \neg Bel_u Class(1\text{st}) \wedge \neg Bel_u BelIf_s Class(1\text{st}) \wedge \neg Bel_u Bel_s BelIf_s Class(1\text{st}))$
   (by (20) where $k = s$, 1., and principles of our logic);
3. $Bel_s Done_{\langle \mathsf{Inform}_{u,s}\ Class(1\text{st})\rangle}(Bel_u Class(1\text{st}) \wedge \neg Bel_u Bel_s \neg Bel_u Class(1\text{st}) \wedge \neg Bel_u BelIf_s Class(1\text{st}) \wedge \neg Bel_u Bel_s BelIf_s Class(1\text{st}))$
   (by (20) where $k = u$, 1., and principles of our logic).

The formulas 2. and 3. are what we call *presuppositions*: when a speech act has just been performed, an agent believes now that the preconditions of this act were true for the speaker just before the performance of this act.

THE INDIRECT EFFECT is the preservation of preconditions, and must be derived from presuppositions by formulas of the form $AlwaysBel_k(Done_\alpha A' \rightarrow A')$ where A' is the preconditions of α: this is a particular instance of our axiom schema of belief preservation (cf. Sect. 4.2).

THE INTENTIONAL EFFECT is the consumption by the hearer of the intention of the speaker (in the sens of Grice). Such a consumption is represented by formulas of the form

$AlwaysBel_k(Done_\alpha\top \to A'')$ where A'' is the intentional effect of α. For example, an instance of this schema for an informative speech act is:

$$AlwaysBel_k(Done_{\langle \mathsf{Inform}_{i,j}\ A\rangle}\top \to Intend_i Bel_j Intend_i Bel_j A) \tag{21}$$

The perlocutionary effect is not consumed (there is no schema to add it to the mental state of an agent): our agents are autonomous (the expected effect of an act doesn't happen systematically). If the new mental state, obtained by the admission of a speech act and the consumption of his indirect and intentional effects, entails the perlocutionary effect, we said that the latter has been consumed (but it is a misuse of language).

Reactive laws. The reactive laws allow us to generate some intentions:

$$AlwaysBel_i(A \wedge Bel_j\neg A \to Intend_i Bel_j A) \tag{22}$$
$$AlwaysBel_i(A \wedge Done_{\langle \mathsf{Inform}_{j,i}\ A\rangle} Bel_i\neg A \to Intend_i Bel_j Bel_i A) \tag{23}$$
$$AlwaysBel_i(Done_\alpha(Done_\gamma\top \wedge Bel_i Done_\beta\top) \to Intend_i Bel_j Bel_i Done_\alpha Done_\gamma\top) \tag{24}$$

...

For example, (22) is used for the first part of the utterance s_4 of our running example: the system invalidates the price of 80 €, and informs the user that the price is 100 €. Formally (we don't give the logical axioms we use, s and u are respectively the agents i and j in the law (22)):

1. $Bel_s Price(80\,€)$ (hypothesis)
2. $Bel_s Bel_u Price(100\,€)$ (hypothesis)
3. $Bel_{s,u}\neg(Price(80\,€) \wedge Price(100\,€))$ (static law)
4. $Bel_s\neg Price(100\,€)$ (by 1. and 3.)
5. $Intend_s Bel_u\neg Price(100\,€)$ (by (22), 4. and 2.)
6. $Bel_s Bel_u\neg Price(80\,€)$ (by 2. and 3.)
7. $Intend_s Bel_u Price(80\,€)$ (by (22), 1. and 6.)

The intentions in 5. and 7. are respectively associated with a deny speech act (the price isn't 80 €) and an informational speech act (the price is 100 €).

3.4 The problem of belief change

In our approach, unlike to Sadek we always accept[6] 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 about these beliefs. Thus, the speaker's 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 the input about the new class (after u_2) but rejects the input about the price (after u_3), the reason being that he considers u to be competent in classes but not in 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 mental 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.)

All this presupposes that we are able to determine the competence of an agent and the influence relationship between speech acts and beliefs. The foundation for both notions will be provided here 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 is competent about a formula A if and only if the set of topics associated with A is a subset of the set of topics associated with this agent – the set of topics the agent is competent in. And a formula A is preserved after the performance of a speech act α if there is no topic that occurs both in the set of topics associated to A and in the set of topics associated to the speech act. The next section gives the formal apparatus.

4 Topic-based belief change

To describe the relation between a formula and a speech act, and the relation between a formula and an agent, we use the concept of topic. This notion is very important from a linguistic and a logical point of view.

For example, in [Büring, 1995], a semantical value related to the topics is associeted 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.

Epstein (1990) 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, page 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, page 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 $\mathfrak{subject}$ function can be seen as an extension of this function to a multi-modal language.

Other studies of the notion of topic exist in the literature, in particular those of Lewis [Lewis, 1972] and Goodman [Goodman, 1961]. Both are quite different from Epstein's. 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 focusses on the "informative aspect" of propositions, the subject of a tautology is the empty set.

4.1 Topic structures

4.1.1 Themes and topics

A *theme* is what something is about. For example, an 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 destinations, classes, and prices are among the themes.

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*.* $\mathcal{C}$ *is the set of all the contexts.*

ma_i stands for "the mental attitude of agent i". In this paper we shall suppose that the length of each context is at most 2.

Definition 2 *A* topic of information *(or* contextual thematic structure*) is a theme together with a context denoted by* $c{:}t$ *where* $t \in \mathcal{T}$*,* $c \in \mathcal{C}$*, and* c *is of length at most 2. Given a set of themes* $\mathcal{T}$ *we note* $\mathbb{T}$ *the associated set of topics.*

$\lambda{:}c = c{:}\lambda = c$. By convention, $\lambda{:}t$ is abbreviated by t. Hence $\mathcal{T} \subseteq \mathbb{T}$. 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.

4.1.2 The subject of a formula

Definition 3 *We call* subject of a formula *a set of topics the formula is about. This notion is formalized by a function* $\mathfrak{subject}$ *mapping a formula to a set of topics from* $\mathbb{T}$*.*

The spirit of our $\mathfrak{subject}$ function is that of Epstein. We give the following axioms.

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

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

(viz. The truth is about nothing.)[7]

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

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

Axiom 5

$$\begin{aligned}\mathfrak{subject}(Bel_i A) = \{ma_i{:}c{:}t \mid\ & c \neq ma_i \textit{ and} \\ & (c{:}c'{:}t \in \mathfrak{subject}(A) \textit{ or } ma_i{:}c{:}t \in \mathfrak{subject}(A))\}.\end{aligned}$$

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

$\mathfrak{subject}(Class(\text{1st})) = \{classe\}$
$\mathfrak{subject}(Dest(\text{Nice})) = \{destination\}$
$\mathfrak{subject}(Bel_s Bel_u Price(80\,€) \wedge Bel_s Price(100\,€)) = \{ma_s{:}ma_u{:}prix\} \cup \{ma_s{:}prix\}$.

Theorem 1 *If* A *is a formula that contains no modal operator, then:*

$$\begin{aligned}\mathfrak{subject}(Bel_i A) &= \mathfrak{subject}(Bel_i ... Bel_i A) \\ &= \{ma_i{:}t \mid t \in \mathfrak{subject}(A)\}.\end{aligned} \tag{25}$$

$$\begin{aligned}\mathfrak{subject}(Bel_i Bel_j A) &= \mathfrak{subject}(Bel_i ... Bel_i Bel_j ... Bel_j Bel_k ... A) \\ &= \{ma_i{:}ma_j{:}t \mid i \neq j, t \in \mathfrak{subject}(A)\}.\end{aligned} \tag{26}$$

Axiom 6 $\mathfrak{subject}(Intend_i A) = \mathfrak{subject}(Bel_i A)$*.*

Axiom 7 $\mathfrak{subject}(Done_\alpha A) = \mathfrak{subject}(A) \cup \mathfrak{subject}(A')$ *where A' is the propositional content of α.*

Axiom 8 $\mathfrak{subject}(\forall x A) = \mathfrak{subject}(A)$.

Axiom 9 $\mathfrak{subject}(A[t/x]) \subseteq \mathfrak{subject}(A)$, *where $A[t/x]$ is the formula resulting from the substitution of the variable x by the term t.*

Our $\mathfrak{subject}$ function is not extensional: logically equivalent formulas may have different topics. (An intuition that might be helpful is to think of the subject of A as the set of predicates names occurring in A).

4.1.3 The competence of an agent

Now we associate topics to agents.

Definition 4 *We call* competence of an agent *the set of topics in which this agent is competent. This notion is formalized by a function* $\mathfrak{competence}$ *mapping an agent to a set of topics from* $\mathbb{T}$.

We assume every agent is competent in his mental states.

Axiom 10 $\mathfrak{competence}(i) \supseteq \{ma_i{:}t \mid t \in \mathcal{T}\}$.

An agent may be competent about some facts. For example, $\mathfrak{competence}(u)$ contains destinations and classes, but not prices.[8]

Competence will allow us to formulate in the next section our belief adoption axiom which basically says: "an agent j adopts the belief of an other agent i about a formula A if j considers that i is competent in the subject of A. [9] "

4.1.4 The scope of an 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.

Definition 5 *We call* scope of a speech act *a set of topics associated to this speech act. This notion is formalized by a function* $\mathfrak{scope}$ *mapping a speech act to a set of topics from* $\mathbb{T}$.

The scope of a speech act determines what mental attitudes of an agent are questioned by this act.

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 11 $\mathfrak{scope}(\langle \mathsf{FORCE}_{i,j}\ A\rangle) \supseteq \{ma_j{:}ma_i{:}t \mid t \in \mathfrak{subject}(A)\}$, *for every illocutionary force* FORCE.

In the case of request, we postulate that these mental attitudes are the only one questioned.

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.

4.1.5 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, a* topic structure *consists of the associated set of topics together with the* $\mathfrak{subject}$, $\mathfrak{scope}$, *and* $\mathfrak{competence}$ *functions.*

Is there an interaction between these functions? We propose the following axiom for acts of the informative type.

Axiom 12 *If A is factual, $\alpha = \langle \mathsf{Inform}_{i,j}\, A \rangle$, and t is a theme such that $t \in \mathfrak{subject}(A) \cap \mathfrak{competence}(i)$ then $t \in \mathfrak{scope}(\alpha)$ and ma_j:$t \in \mathfrak{scope}(\alpha)$.*

For example, if the user informs the system about his destination, as the user is competent at destinations, then this speech act influences the system's factual beliefs about the destination, and also about prices, because a change in the destination possibly entails a change in the price: $\mathfrak{scope}(\langle \mathsf{Inform}_{u,s}\, Dest(\text{Amsterdam}) \rangle)$ contains the topics *destination*, *price*, ma_s:*destination* and ma_s:*price*.

A given topic structure will allow us to reconstruct beliefs by means of two principles: competence and preservation. In the next section we shall present these principles.

4.2 Axioms for belief change

Our axioms for belief change are based on a given topic structure. The first one is the axiom schema of belief preservation:

Axiom Schema of Belief Preservation.

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

The restriction to formulas without $Done_\beta$ operator 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 is about the adoption of beliefs.

Axiom Schema of Belief Adoption.

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

The schema expresses that: if the agent i believes that A, and if he's comptent about A, then A is true.

For example the formula $Bel_s Bel_u Dest(\text{Amsterdam}) \rightarrow Bel_s Dest(\text{Amsterdam})$ can be proved from the instance $Bel_u Dest(\text{Amsterdam}) \rightarrow Dest(\text{Amsterdam})$ of the belief adoption axiom, because $\mathfrak{subject}(Dest(\text{Amsterdam})) \subseteq \mathfrak{competence}(u)$. (By using 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 $\mathfrak{subject}(Price(80€)) \not\subseteq \mathfrak{competence}(u)$.

Remark 1 *In our preceding approach [Fariñas del Cerro* et al.*, 1998] we had 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 about the users's attitudes about p, but not the system's attitudes about p. We were not able to distinguish that before.*

5 Example

We illustrate our belief change process by our running example. To each uterrance number, we associate a speech act number (e.g. α_{u_i} points at the u_i utterance). We describe the different mental states S_{s_i} of the system after the different speech acts of the user. S_{s0} is the initial state (supposed empty). Thus, the different mental states that we want to describe are the following:

$$S_{s_0} \xrightarrow{\alpha_{s1}} S_{s_1} \xrightarrow{\alpha_{s2}} \ldots \xrightarrow{\alpha_{sn}} S_{s_n}$$

The different speech acts performed are respectively:
$\alpha_{u_1} = \langle \mathsf{Inform}_{u,s}\ Class(\text{1st}) \wedge Dest(\text{Nice})\rangle$
$\alpha_{u_2} = \langle \mathsf{Inform}_{u,s}\ Class(\text{2nd})\rangle$
$\alpha_{u_3} = \langle \mathsf{Inform}_{u,s}\ Price(80€)\rangle; \langle \mathsf{ReqInformIf}_{u,s}\ Payment(\text{carte de crédit})\rangle$
The different scopes corresponding to these speech acts are:

- $\mathfrak{scope}(\alpha_{u_1}) \supseteq \{ma_s{:}ma_u{:}t, ma_s{:}t\}$ where t is a theme of the propositional content of α_{u_1} ($t \in \{class, destination, \ldots\}$);
- $\mathfrak{scope}(\alpha_{u_2}) \supseteq \{ma_s{:}ma_u{:}t, ma_s{:}t\}$ where t is a theme of the propositional content of α_{u_2} ($t \in \{class, \ldots\}$);
- $\mathfrak{scope}(\alpha_{u_3}) \supseteq \{ma_s{:}ma_u{:}t\}$ where t is a theme of the propositional content of each speech act of α_{u_3} ($t \in \{price, payment, \ldots\}$);
- $\mathfrak{competence}(u) \supseteq \{ma_u{:}t \mid t \in \mathcal{T}\} \cup \{destination, class\}$;
- $\mathfrak{competence}(s) \supseteq \{ma_s{:}t \mid t \in \mathcal{T}\} \cup \{price, payment\}$;.

We use the following abbreviations:

- $S_{s_i}^{-Done}$ is the conjunction of formulas of the state S_{s_i} which contain no *Done*-operator;
- $C1$ and $C2$ are respectively $Class(\text{1st})$ and $Class(\text{2nd})$;
- D is $Dest(\text{Nice})$;
- $P1$, $P2$ and $P3$ are respectively $Price(150€)$, $Price(100€)$ and $Price(80€)$ (bad price);

- CC is $\mathit{Payment}$(credit card).

We do not give every formula produced by our logic. The preconditions of the speech acts are simplifyed. Thus, the different mental states are the following:

$S_{s_0} = \emptyset$

S_{s_1} contains the following formulas:

1. performance of the act: $Bel_s Done_{\alpha_{u_1}} \top$
2. presuppositions: $Bel_s Done_{\alpha_{u_1}} (Bel_u(C1 \wedge D) \wedge \neg Bel_u BelIf_s(C1 \wedge D))$
3. indirect effect: $Bel_s Bel_u C1 \wedge Bel_s Bel_u D \wedge Bel_s \neg Bel_u BelIf_s(C1 \wedge D)$
4. intensional effect: $Bel_s Intend_u Bel_s Intend_u Bel_s C1 \wedge$ $Bel_s Intend_u Bel_s Intend_u Bel_s D$
5. reduction of intention: $Bel_s Intend_u Bel_s C1 \wedge Bel_s Intend_u Bel_s D$
6. adoption: $Bel_s C1 \wedge Bel_s D$
7. static laws: $Bel_s P1$

S_{s_2} contains the following formulas:

1. performance of the act: $Bel_s Done_{\alpha_{u_2}} \top$
2. memory: $Bel_s Done_{\alpha_{u_2}} S_{s_1}^{-Done}$
3. presuppositions: $Bel_s Done_{\alpha_{u_2}} (Bel_u C2 \wedge \neg Bel_u BelIf_s C2)$
4. indirect effect: $Bel_s Bel_u C2 \wedge Bel_s \neg Bel_u BelIf_s C2$
5. intensional 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: $Bel_s Bel_u D \wedge Bel_s Intend_u Bel_s Intend_u Bel_s D \wedge$ $Bel_s Intend_u Bel_s D \wedge Bel_s D$
8. adoption: $Bel_s C2$
9. static laws: $Bel_s P2$

S_{s_3} contains the following formulas:

1. performance of the act: $Bel_s Done_{\alpha_{u_3}} \top$
2. memory: $Bel_s Done_{\alpha_{u_3}} S_{s_2}^{-Done}$
3. presuppositions: $Bel_s Done_{\alpha_{u_3}} (Bel_u P3 \wedge \neg Bel_u BelIf_s P3 \wedge \neg BelIf_u CC \wedge$ $Bel_u \neg Intend_s Done_{\langle \mathsf{InformIf}_{s,u}\, CC \rangle} \top)$
4. indirect effect: $Bel_s Bel_u P3 \wedge Bel_s \neg Bel_u BelIf_s P3 \wedge Bel_s \neg BelIf_u CC \wedge$ $Bel_s Bel_u \neg Intend_s Done_{\langle \mathsf{InformIf}_{s,u}\, CC \rangle} \top)$

5. intensional effect: $Bel_s Intend_u Bel_s Intend_u Bel_s P3 \wedge$
   $Bel_s Intend_u Bel_s Intend_u Done_{\langle \mathsf{InformIf}_{s,u}\ CC \rangle} \top$

6. reduction of intention: $Bel_s Intend_u Bel_s P3 \wedge$
   $Bel_s Intend_u Done_{\langle \mathsf{InformIf}_{s,u}\ CC \rangle} \top$

7. preservation: $Bel_s Bel_u C2 \wedge Bel_s \neg Bel_u BelIf_s C2$
   $\wedge Bel_s Intend_u Bel_s Intend_u Bel_s C2 \wedge Bel_s Intend_u Bel_s C2 \wedge Bel_s Bel_u D \wedge$
   $Bel_s Intend_u Bel_s Intend_u Bel_s D \wedge Bel_s Intend_u Bel_s D \wedge Bel_s D \wedge Bel_s C2$

8. static laws: $Bel_s P2$

9. reactive laws: $Intend_s Bel_u \neg P3 \wedge Intend_s Bel_u P2$

The aim of our approach is always to generating a consistent mental state. Formally, we can describe how the new mental state S_{k+1} is constructed from the state S_k:

$$S_{k+1} = \{Bel_i Done_{\alpha_{k+1}} A : A \in S_k\} \cup LAWS.$$

Hence, $S_{k+1} \rightarrow C$ ssi $LAWS \vdash (Bel_i Done_{\alpha_{k+1}} Bel_i Done_{\alpha_k} \ldots Bel_i Done_{\alpha_1} S_0) \rightarrow C$.

In the below example, our propositional contents are relatively simple: however, there exists some laws which permit to treat more complex propositional contents [Sadek, 1991a, Longin, 1999].

6 Conclusion

We have sketched a theory of change in the context of dialogues. It is based on the notion of topic of information, which is exploited through topic-based axioms of belief adoption and preservation.

Perrault and Appelt & Konolige have argued that defaults are crucial elements in a theory of speech acts because in a way they permit to transform absence of knowledge into knowledge. In a sense, what we do is to transferring that task to the metalinguistic relations of competence and scope. This permits to keep a monotonic framework.

We have supposed that the set of topics associated with a formula is determined by those of the atomic formulas occurring in it. This is certainly a debatable choice. It was mainly motivated by representational economy. Notwithstanding, the way we use the $\mathfrak{subject}$ function is sound: suppose e.g. $\mathfrak{subject}(p) = \{t\}$, $\mathfrak{subject}(q) = \{t'\}$, and $\mathfrak{scope}(\alpha) = \{t'\}$. Hence p and $p \wedge (q \vee \neg q)$ do not have the same subject, and $Done_\alpha p \rightarrow p$ is an instance of the preservation axiom, while $Done_\alpha(p \wedge (q \vee \neg q)) \rightarrow (p \wedge (q \vee \neg q))$ is not. Nevertheless, the latter formula can be deduced from the former standard modal logic principle: as $p \leftrightarrow p \wedge (q \vee \neg q)$ we have $Done_\alpha p \leftrightarrow Done_\alpha(p \wedge (q \vee \neg q))$. Hence $Done_\alpha p \rightarrow p$ is equivalent to $Done_\alpha(p \wedge (q \vee \neg q)) \rightarrow (p \wedge (q \vee \neg q))$.

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

Last but not least, we note that a possible worlds semantics can be given to our logic by adapting the one presented in [Fariñas del Cerro *et al.*, 1998].

Notes

1 Note that the conversation may be in natural language; however it isn't necessary.

2 We consider that the real world includes everything that is external to the agent (including speaker intention).

3 $Done_\alpha A$ et $Feasible_\alpha A$ are respectively just as $\langle \alpha^{-1} \rangle A$ and $\langle \alpha \rangle A$ of dynamic logic [Harel, 1984].

4 The second term is an abbreviation of the infinite conjunction of Sadek :
$\neg Bel_i Bel_j \neg Bel_i A \wedge \neg Bel_i Bel_j Bel_i \neg Bel_i A \wedge \neg Bel_i Bel_j Bel_i Bel_j \neg Bel_i A \wedge \ldots$

5 The second term is an abbreviation of the infinite conjunction of Sadek :
$\neg Bel_i Bel_j BelIf_j A \wedge \neg Bel_i Bel_j Bel_i BelIf_j A \wedge \neg Bel_i Bel_j Bel_i Bel_j BelIf_j A \wedge \ldots$

6 "Accepting" an act means that we admit that it has been performed.

7 Note that here we differ slightly from Epstein's view. Indeed, Epstein stipulates that $\mathcal{R}(A, A)$ for every formula A. On the contrary, the present axiom makes that $not(\mathcal{R}(\top, \top))$. More generally, we have $\mathcal{R}(A, A)$ iff the set of atoms of A is nonempty. Clearly, this is due to the fact that Epstein does not have the logical operators $\top$ and $\bot$ in his language.

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

9 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.

References

[Alchourrón *et al.*, 1985] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(2):510–530, June 1985.

[Appelt and Konolige, 1989] Douglas Appelt and Kurt Konolige. A nonmonotonic logic for reasoning about speech acts and belief revision. In M. Reinfrank, J. de Kleer, M.L. Ginsberg, and E. Sandewall, editors, *Proceedings of Second International Workshop on Non-Monotonic Reasoning*, volume 346 of *LNAI*, pages 164–175, Allemagne, June 1989. Springer-Verlag.

[Austin, 1962] John L. Austin. *How To Do Things With Words.* Oxford University Press, 1962.

[Büring, 1995] D. Büring. Topic. In P. Bosch and R. v. d. Sandt, editors, *The Focus Book.* 1995.

[Chellas, 1980] B. F. Chellas. *Modal Logic: an introduction.* Cambridge University Press, 1980.

[Cohen and Levesque, 1990a] Philip R. Cohen and Hector J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42(2–3):213–261, 1990.

[Cohen and Levesque, 1990b] Philip R. Cohen and Hector J. Levesque. Persistence, intentions, and commitment. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, editors, *Intentions in Communication*, chapter 3, pages 33–69. MIT Press, Cambridge, MA, 1990.

[Cohen and Levesque, 1990c] Philip R. Cohen and Hector J. Levesque. Rational interaction as the basis for communication. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, editors, *Intentions in Communication*, chapter 12, pages 221–255. MIT Press, Cambridge, MA, 1990.

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

[Fariñas del Cerro *et al.*, 1998] Luis Fariñas del Cerro, Andreas Herzig, Dominique Longin, and Omar Rifi. Belief reconstruction in cooperative dialogues. In Fausto Giunchiglia, editor, *Proceedings of 8th International Conference on Artificial Intelligence: Methodology, Systems, and Applications (AIMSA'98)*, volume 1480 of *LNAI*, pages 254–266, Bulgaria, September 1998. Springer-Verlag.

[Ginzburg, 1995] J. Ginzburg. Resolving questions i,ii. *Linguistics and Philosophy*, 18:567–609, 1995.

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

[Grice, 1967] H. Paul Grice. Utterer's meaning and intentions. *Philosophical Review*, 78(2):147–177, April 1967. Also available in [Grice, 1989, Chapter 5].

[Grice, 1975] H. Paul Grice. Logic and conversation. In J. P. Cole and J. L. Morgan, editors, *Syntaxe and Semantics : Speech acts.* Academic Press, New York, 1975. Also available in [Grice, 1989, Chapter 2].

[Grice, 1989] H. Paul Grice. *Studies in the way of words.* Harvard University Press, USA, 3rd edition, 1989.

[Halpern and Moses, 1985] Joseph Y. Halpern and Yoram Moses. A guide to the modal logics of knowledge and belief: preliminary draft. In *Proceedings of IJCAI'85*, pages 480–490, Los Angeles, 1985.

[Harel, 1984] David Harel. Dynamic logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume II. D. Reidel Publishing Company, 1984.

[Herzig, 1998] Andreas Herzig. Logics for belief base updating. In Didier Dubois, Dov Gabbay, Henri Prade, and Philippe Smets, editors, *Handbook of defeasible reasoning and uncertainty management*, volume 3 - Belief Change, pages 189–231. Kluwer Academic Publishers, 1998.

[Hughes and Cresswell, 1972] G.E. Hughes and M.J. Cresswell. *An Introduction to Modal Logic.* Methuen, 2nd edition, 1972.

[Katsuno and Mendelzon, 1992] Hirofumi Katsuno and Alberto O. Mendelzon. On the difference between updating a knowledge base and revising it. In Peter Gärdenfors, editor, *Belief Revision*, number 29 in Cambridge Tracts in Theorical Computer Science, pages 183–203. Cambridge University Press, 1992.

[Lewis, 1972] D.K. Lewis. General semantics. In D. Davidson and G. Harman, editors, *Semantics of natural language.* D. Reidel Publishing Company, 1972.

[Longin, 1999] Dominique Longin. *Interaction rationnelle et évolution des croyances dans le dialogue : une logique basée sur la notion de topique.* PhD thesis, Université Paul Sabatier, Institut de Recherche en Informatique de Toulouse (IRIT), Toulouse, France, to appear 1999.

[McCarthy and Hayes, 1969] J. McCarthy and P. Hayes. Some philisophical problems from the standpoint of artificial intelligence. In B. Meltzer and D. Michie, editors, *Machine Intelligence*, volume 4, pages 463–502. Edinburgh University Press, 1969.

[Perrault, 1990] C. Raymond Perrault. An application of default logic to speech act theory. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, editors, *Intentions in Communication*, chapter 9, pages 161–185. MIT Press, Cambridge, MA, 1990.

[Popkorn, 1994] Sally Popkorn. *First Steps in Modal Logic.* Cambridge University Press, Cambridge, 1994.

[Rao and Georgeff, 1991] Anand S. Rao and Michael P. Georgeff. Modeling rational agents within a bdi-architecture. In J. A. Allen, R. Fikes, and E. Sandewall, editors, *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning (KR & R'91)*, pages 473–484, San Mateo, CA, 1991. Morgan Kaufmann Publishers.

[Rao and Georgeff, 1992] Anand S. Rao and Michael P. Georgeff. An abstract architecture for rational agents. In Bernhard Nebel, Charles Rich, and William Swartout, editors, *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, pages 439–449, Cambridge, Massachusetts, 1992.

[Rao and Georgeff, 1998] Anand S. Rao and Michael P. Georgeff. Decision procedures for bdi logics. *Journal of Logic and Computation*, 8(3):293–343, June 1998. Special Issue: Computational and Logical Aspects of Multiagent Systems, Oxford univerty Press.

[Reiter, 1980] Ray Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[Sadek *et al.*, 1997] David Sadek, Philippe Bretier, and Franck Panaget. Artimis: Natural dialogue meets rational agency. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI'97).* Morgan Kaufmann Publishers, August 1997.

[Sadek, 1991a] M. D. Sadek. *Attitudes mentales et interaction rationnelle : vers une théorie* *formelle de la communication.* PhD thesis in Computer Sciences, Université de Rennes I, Rennes, France, June 1991.

[Sadek, 1991b] M. D. Sadek. Dialogue acts are rational plans. In *Proceedings of ESCA/ETRW, Workshop on The Structure of Multimodal Dialogue (Venaco II)*, pages 1–29, Maratea, Italie, September 1991.

[Sadek, 1992] M. D. Sadek. A study in the logic of intention. In Bernhard Nebel, Charles Rich, and William Swartout, editors, *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, pages 462–473, Cambridge, Massachusetts, October 1992.

[Sadek, 1994] M. D. Sadek. Towards a theory of belief reconstruction: Application to communication. *Speech Communication Journal'94, special issue on Spoken Dialogue*, 15(3–4):251–263, 1994.

[Searle and Vanderveken, 1985] J. R. Searle and D. Vanderveken. *Foundation of illocutionary logic.* Cambridge University Press, 1985.

[Searle, 1969] John R. Searle. *Speech acts: An essay in the philosophy of language.* Cambridge University Press, New York, 1969.

[van Kuppevelt, 1991] J. van Kuppevelt. *Topic en Comment.* PhD thesis, University of Nijmegen, 1991.

[van Kuppevelt, 1995] J. van Kuppevelt. Discourse structure, topicality and questioning. *Linguistics*, 31:109–147, 1995.

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
`mailto: {herzig,longin}@irit.fr`
`http://www.irit.fr/ACTIVITES/EQ_ALG/`