LOGICAL TASK MODELLING FOR MAN-MACHINE DIALOGUE

M.D. SADEK
Centre National d'Etudes des Télécommunications
Route de Trégastel - B.P.40
F 22301 Lannion Cedex - FRANCE

Abstract
To design a task-independent dialogue system, we present a task-oriented dialogue analysis in terms of finding the referents of definite descriptions and we show how this analysis leads to a goal-oriented inferential representation of the task. This representation provides a logical generic model of the task, which is compatible with a belief system. Then, we show how this task model, jointly used with the domain-specific user model for which we propose a formalization, enables a dialogue system to plan request negotiation dialogues.

1. Introduction
The problem we deal with in this paper comes within a more general research concerning the logical plan-oriented approach to Man-machine oral communication. This research aims at designing a task-independent dialogue system. Briefly sketched, the adopted approach is based on a logical representation of mental attitudes and on a recast of speech acts theory (Searle 1969) into action theory (in a similar way as in (Cohen & Perrault 1979)). In this view, one main component of a communicating agent is a belief system\(^1\) involving several types of information such as a task model and a user model, all formalized in the same logical language.

Roughly speaking, a task-independent dialogue system is a portable dialogue system. When designing such a system, information about the potential task must appear only as a formal parameter for which it is not necessary to know some specific instance. To reach this purpose, we have to find a formal expression for the information characterizing the task, that is, a logical abstract data structure.

\(^1\) Although other mental attitudes than belief are also dealt with, the use of the term "belief" in "belief system" is justified in that in our approach, we take an autoepistemic point of view. Hence, the agents we model are introspective; they believe all their mental attitudes and adopt all the mental attitudes they believe about themselves.

In this paper, we present a task-oriented dialogue analysis in terms of finding the referents of definite descriptions and we show how it leads to a goal-oriented inferential representation of the task. This representation gives a logical generic model of the task, which is compatible with a belief system. Then, we show how this task model, jointly used with the domain-specific user model for which we propose a formalization, enables a dialogue system to plan request negotiation dialogues.

For our concern here, we restrict our interest to database inquiry dialogues.

2. Logical preliminaries
To refer to objects in the discourse domain, one uses referring expressions such as proper names or definite descriptions (the green box on the table, the departure time of the flight to Lannion, the word I just uttered, ...). A definite description is an expression one can formulate according to the syntactic pattern "the so (and so)". It is intended to refer to one and only one object, precisely that one which is "so (and so)". In English (as in French), a definite description is a noun phrase starting with a definite article.

To handle definite descriptions within a logical formalism, the first order logic language is augmented by the operator \(\tau\) (Russell 1905), which is a term-producing operator: if \(\psi(x)\) is a well formed formula with \(x\) as free variable then \(\tau\psi(x)\) is a term. This term describes a definite description and can be read "the unique object which satisfies the property \(\psi\)" or in a compact form "the \(\psi\)". Intuitively, a formula of the form \(\phi(\tau\psi(x))\) is false if there is no object in the discourse domain satisfying the description \(\psi\) or if there is more than one object. Thus, if equality is provided in the logical language, one has the equivalence

\[\phi(\tau\psi(x)) \iff \exists x(\psi(x) \land \forall y(\psi(y) \implies y = x) \land \phi(x)).\]

2 Strictly speaking, we should say "an agent uses some description to refer to..." instead of "some description refers to...". Though, in this paper, we continue using the latter expression or expressions equivalent to it.
Given this equivalence, the truth conditions of a formula of the form \( \phi(t \psi(x)) \) can be easily stated. Let \((M, \nu)\) be a first order (tarskian) interpretation where \(M\) is a first order assignment and \(\nu\) a variable valuation. Then, \((M, \nu) \models \phi(t \psi(x))\) if and only if (1) the projection of \(\psi|_{M}\) on the component \(x\) is a singleton \(\{d\} \subseteq D\) (where \(D\) is the interpretation domain) and (2) \(d\) belongs to the projection of \(\phi|_{M}\) on the component \(t \psi(x)\), where \(\models\) is the truth relation and \(\phi|_{M}\) and \(\psi|_{M}\) the respective extensions of the formulæ \(\phi\) and \(\psi\) according to the assignment \(M\).

3. On the mental attitudes formalization

We base an agent mental model on three primitive mental attitudes: (autoepistemic) belief, uncertainty\(^3\) and choice (or goal, as called in (Cohen & Levesque 1986)), and we formalize it in a first order modal language. The concept of belief is formalized by the modal operator \(K\) the logical (external) model of which is a KD45 possible-world semantical Kripke structure (Halpern & Moses 1985) with the fixed domain principle (Garson 1984). (A formula as \(K(i, p)\) can be read "\(p\) is a consequence of \(i\)’s beliefs"). Moreover, it is assumed that there is in the language a set of constants which are rigid designators (called in this paper standard names), that is, constants which have the same extensions in all possible worlds. It is also assumed that this set of constants cover the interpretation domain. Equality is interpreted as contingent identity.

Now, let’s examine the way knowing something (or someone) is represented. A formula of the form \(K(i, \phi)\) is not relevant whenever one has to represent an assertion like "\(i\) knows \(j\)’s address" because the expression "\(j\)’s address" doesn’t denote a proposition. It is a referring expression potentially denoting some object in the discourse domain. A plausible formalization of the assertion "\(i\) knows \(j\)’s address" can be \(3xK(i, \text{address}(j) = x)\).

However, there is an information which is conveyed by the previous assertion and which is not expressed in the above formula: the uniqueness of \(j\)’s address. The above formula doesn’t reject the possibility that \(i\) knows more than one address for \(j\). The fact that an agent \(i\) knows a definite description \(t \delta(x)\) is formalized by \(3yK(i, t \delta(x) = y)\), which we note below \(Kref(i, \delta(x))\).

Now, looking at the concept of choice (or goal), it is formalized by the modal operator \(C\) which is interpreted, as the belief operator, in a possible-world semantical Kripke structure. The properties imposed to the semantical model,

\(\phi(t \psi(x))\) can be formally stated. Let \((M, \nu)\) be a first order (tarskian) interpretation where \(M\) is a first order assignment and \(\nu\) a variable valuation. Then, \((M, \nu) \models \phi(t \psi(x))\) if and only if (1) the projection of \(\psi|_{M}\) on the component \(x\) is a singleton \(\{d\} \subseteq D\) (where \(D\) is the interpretation domain) and (2) \(d\) belongs to the projection of \(\phi|_{M}\) on the component \(t \psi(x)\), where \(\models\) is the truth relation and \(\phi|_{M}\) and \(\psi|_{M}\) the respective extensions of the formulæ \(\phi\) and \(\psi\) according to the assignment \(M\).

The formula \(W(i, \phi)\) can be read "\(i\) needs \(p\) to hold". Intuitively, an agent \(i\) needs \(p\) to hold if and only if, seeing that \(p\) is not a consequence of her beliefs, she adopts the intention that it will be the case. Here, all that is useful to know about the concept of need is that both schemas \(W(i, \phi) \land K(i, \phi)\) and \(W(i, \phi) \land K(i, \neg \phi)\) are satisfiable, that the following schemas are valid:

\[\begin{align*}
I(i, \phi) & \Rightarrow W(i, \phi) \\
W(i, \phi) & \Leftrightarrow K(i, W(i, \phi)) \\
W(i, \phi \land \psi) & \Rightarrow W(i, \phi) \land W(i, \psi)
\end{align*}\]

and that the following inference rule is sound:

\[\frac{
W(i, \phi) \land K(i, \neg \phi)
}{I(i, \phi)}\]

4. Dialogue and definite descriptions

To design a task-independent dialogue system, we have to answer the following question: how to characterize the abstraction level at which the topic one is communicating about becomes merely a parameter which is not pertinent in the logical reasoning model? In other words, how to build up a structural model for the task, which is compatible with the belief system?

The idea underlying the answer to this question is that in an informative task-oriented dialogue, the generic problem which agents pose to each other is to identify the referents of definite descriptions.\(^7\) The task-oriented (bi-agent) dialogue "game" can be sketched as follows: an agent \(i\), through a request directed to an agent \(j\), provides

---

3 We don’t make use here of the concept of uncertainty; so, it will not be introduced. (For a logical formalization of this concept, see (Sadek 1990).)
4 This property is proposed in (Cohen & Levesque 1986).
5 For a detailed analysis of the concept of choice, see (Sadek 1990).
6 but fundamentally recast in order (among other things) to take into account the agents introspective and cooperative abilities (see (Sadek 1990)).
7 In (Clark & Marshall 1981), (Cohen & Perrault 1981) and (Nadathur & Joshi 1983), logical conditions (based on mutual beliefs) for the use and understanding of referring expressions are proposed. In (Appelt 1983), the problem of planning referring expressions is dealt with.
(in the primitive case\(^8\)) a definite description, thereby (among other things) expecting \(j\) to identify the referent denoted by the description and, generally, to "return" a standard name (or an identifying description) of that referent. If the agent \(j\) evaluates the description not to be an identifying one, the negotiation "rule" consists in trying to get from \(i\) fuller information in order to construct, starting from the first description, an identifying one. To clarify this idea, let's take an example of a task-oriented dialogue:

(a) \(i\): I'd like to know the telephone number of Mr DuPont at Lannion city
(b) \(j\): It's the 96.00.00.00
(c) \(i\): And those of Tenc and Tassne companies
(d) \(j\): At Lannion city?
(e) \(i\): Yes
(f) \(j\): Do you know the directory name?
(g) \(i\): For Tenc or for Tassne?
(h) \(j\): ...

By (a), \(i\) (indirectly) request \(j\) to identify the referent denoted by the definite description "the telephone number of Mr DuPont at Lannion city" (and to "return" a standard name of this referent). This description can be formalized as follows:

\[
\forall x (\text{Tel}(x) \land \text{Belongs}(x, \forall y (\text{Person}(y) \land y = \text{DuPont}) \land \text{Resides}(y, \forall z (\text{City}(z) \land z = \text{Lannion})))
\]

A slightly different formalization may be:

\[
\forall x (\text{Tel}(x) \land \text{Belongs}(x, \forall y (\text{Person}(y) \land y = \text{DuPont}) \land \text{Resides}(y, z), \forall z (\text{City}(z) \land z = \text{Lannion})))
\]

The optimal formalization depends on the internal organization of the data base and, therefore, from a computational point of view, on the control strategy of term evaluation. Though, one can note that probably the formula

\[
\forall x (\text{Tel}(x) \land \text{Belongs}(x, \forall y (\text{Person}(y) \land y = \text{DuPont}) \land \text{Resides}(y, z), \forall z (\text{City}(z) \land z = \text{Lannion})))
\]

doesn't make an optimal formalization of the description "the telephone number of Mr DuPont at Lannion city" since it ignores the existing relation between "Mr DuPont" and "Lannion city".

Let's return to the dialogue above. The agent \(j\) evaluates the description uttered in (a) to be an identifying one and, thus, provides by (b) a constant which standardly names the referent of the description. By (c), the agent \(i\) makes a new request concerning the identification of the referents of two descriptions: "the telephone number of Tenc company" and "the telephone number of Tassne company". But, this time, \(j\) evaluates the descriptions not to be identifying ones and then starts a negotiation dialogue attempting to construct indentifying descriptions; This is illustrated by the exchanges starting at (d). Note that by (g), it's \(i\) who evaluates the definite description "the directory name" not to be an identifying one.

5. From definite description to goal-oriented inference rule

The main point here is that in a task-oriented dialogue, the framework according to which a task-oriented dialogue system considers some description relating with the application attributes, to be or not an identifying one, is the basis of the task model. This framework is built up according to the two following criteria: (i) The way the dialogue system (designer) believes that the user conceives the application organization schema, and (ii) The integrity constraints which appear in a (fine-grained) conceptual schema describing the application.

Let's see through a generic example how the integrity constraints can be viewed as descriptions which are, then, formalized in a goal-oriented inferential form and can, thus, be naturally included in the belief system. The resulting inference rules gives a logical model of the task.

For simplicity, we don't make account of the criterion (i) and we suppose that the integrity constraints reflect only functional dependencies between the attributes of the data base corresponding to the application. Let \(a, b, c\) and \(r\) be the data base attributes\(^9\) and \(a, b \rightarrow r\) and \(c \rightarrow r\) functional dependencies. These dependencies give rise to definite descriptions which can be formalized as follows:

\[
\forall x (r(x) \land R1(x, \forall y (a(y), \forall z (b(z))));
\]
\[
\forall x (r(x) \land R2(x, \forall y (c(y)))
\]

where \(R1\) and \(R2\) are some real or virtual relations respectively between \(a, b\) and \(r\), and between \(c\) and \(r\)\(^10\). For example, the first expression can be read "the \(r\) given the \(a\) and the \(b\)". From a functional point of view, the expressions above are jointly interpreted as follow: to identify an \(r\)

---

8 By primitive case, we mean the case where the utterance involves only one global description (may be including other descriptions).
9 In practice, taking into account contextual information, this term can be simplified. Indeed, from a computational point of view and under some circumstances, one can merely write \(\forall x (\text{Tel}(x), \text{DuPont}, \text{Lannion})\). Nevertheless, the explicit representation of the unary predicates as \(\text{Tel}(x), \text{Name}(y)\) or \(\text{City}(z)\) (which represent the attributes of the data base and, somehow, call to mind a multi-sorted logic) makes fluent not only the access to the data base but also the exchanges between the properly so called dialogue module and the analysis and generation natural language modules, and even allows a type control over the manipulated objects.
10 From a logical point of view, the attributes can be represented as unary predicates (namely classes).
11 \(R1\) and \(R2\) would, for instance, correspond to the relations "Belongs" or "Resides" of the example above.
differently from standardly naming it, a description must identify either an a and a b, or a c. In terms of goal-oriented inference in a reasoning model, this proposition means that the potential intention of the server-agent (which is considered below to be a dialogue system) to know the referent of the description "the r" can be converted either in her potential intention to know the referents of the descriptions "the a" and "the b", or in her potential intention to know the referent of the description "the c". This is formalized by the following inference rule:

\[
W(s, Kref(s, a(u))) 
\]

\[
W(s, Kref(s, a(x)) \land Kref(s, b(y))) \lor W(s, Kref(s, c(z)))
\]

where s denotes the server-agent and x, y, z and u are individual variables. Expressed in this way, the model we propose, describes the task as a set of potential intention reduction rules the formalization of which comes within the knowledge representation schema required by the logical reasoning (and planning) model. When a particular application is chosen, these rules are extracted from its corresponding conceptual schema taking into account the criteria (i) and (ii) we mentioned above.

6. Task modelling

Russel’s theory of definite description (Russel 1905) imposes (the existence and) the uniqueness of the referent. In the case of data base inquiry dialogues, this constraint is too strong and, therefore, adopting it as such will lead, at least, to a non-friendy behaviour. Indeed, one cannot reasonably expect from the user to utter only descriptions which identify one and only one referent at one and the same time, otherwise one would be committed to draw the conclusion that what was uttered asserts some erroneous proposition. Thus, a user who requests to know "the telephone number of Mr Dupont" when it happens that Mr Dupont has two telephone numbers, must be able to obtain the information he is looking for, and should not be, in anyway, penalized for having uttered a non-identifying definite description according to Russel’s theory.

12 Note the recursive aspect of this assertion, the termination clause being the standard naming. Somehow, this calls to mind a problem dealt with in (Levesque 1984) where Levesque suggests a solution to avoid that a system loops indefinitely in collecting co-referential terms when answering a wh-question.

13 Consequently, it doesn’t take into account contextual information. A description like "the box on the table" (which can be formalized by \(\forall x(\Box x \land \text{On}(x, \text{Table}(y)))\)) is considered to be a non-identifying one whenever there is more than one table, even if there is only one table on which there is a box.

To overcome this problem, we have to broaden the concept of definite description in the framework of data base inquiry dialogues, by relaxing the referent uniqueness constraint and, therefore, by allowing a description to identify more than one referent. However, a maximal threshold must be imposed to the number of referents which can be identified by a given definite description. A description is considered to be a non-identifying one only beyond this degeneration threshold. In doing so, the interpretation of the definite description concept has just to be recast in terms of the size of the (non-empty) set of objects, the elements of which are referred to by the description, instead of a unique individual.

First, taking into account the criterion (i), a degeneration threshold is fixed for the attributes about which the dialogue system can be asked. Then, for each such an attribute and on the basis of the conceptual schema of the application, the sets of attributes, which satisfy the corresponding threshold, are selected. An attribute together with a corresponding set of this form, represents an identifying description. For example, a, b and r being attributes of the application, if the number of r which are identified given some a and some b is inferior to the fixed threshold, the description "the r corresponding to the a and the b" is considered to be an identifying one. The couple \((a, b), r\) is then added to a set which we call the set of dependencies of the task model.

Formally, if we let A be the set of all the attributes of the application and \(\Sigma(A)\) the powerset of A there is a certain subset \(T\) of \(\Sigma(A) \times A\) which describes the set of dependencies of the task model. In terms of the data base, an element \((E, e)\) is in \(T\) only if every concrete instance \(ins(E)\) of E. determines at most \(\lambda\) instances of e, where \(\lambda\) is the maximal threshold above which the description "the e corresponding to \(ins(E)\)" is considered to be a non-identifying one. Now, considering the reasoning model, for every e about which the dialogue system can be asked, the set \((E_1, e), \ldots, (E_n, e)\) is stated as an inference rule of the form

14 This is not to be confused with the interpretation of multiple definite descriptions as in the expression "the telephone numbers of Mr Dupont". In this case, the description is really intended to denote a set of objects, while in our case the description is intended to denote a unique object while, in fact, it may not. In other words, in the first case, the referent is intended to be a set of objects while in the second case, it is the set of the (potential) referents which is dealt with. In practice, in the case of data base inquiry dialogues, the method we propose here can be adapted to deal with multiple definite descriptions and even with indefinite descriptions.

15 Note that it is also this subset which is affected when updating the task model.
Where $K_{ref}(s, E_i)$ is short for $K_{ref}(s, e_1(x_i^1)) \land \ldots \land K_{ref}(s, e_{t(0)}(x_i^{t(0)}))$, $E_i$ being $[e_1, \ldots, e_{t(0)}]$. Precisely, these inference rules are the part of the reasoning model which corresponds to the task model. They enable the dialogue system to set up dialogues which aim to negotiate information to be acquired to satisfy the user's request.

Henceforth, a description is considered to be or not an identifying description according to the task model. Note that in a task-oriented dialogue, some description can be an identifying one according to the task model without being an identifying one according to the data base. This is precisely the case when the set of referents is empty. In a cooperative dialogue, this case should give rise to corrective and suggestive answers (Guyomard & Siroux 1989).

7. Domain-specific user model

To show how the task model is used to plan request negotiation dialogues, we need to introduce another part of the belief system, which concerns the domain-specific user model. In a conversational situation, an agent has a priori assumptions about the beliefs and the behaviour of the other agents. These assumptions include the mutual beliefs about the speech acts and their usage conditions, or about the principles of cooperation. Another kind of assumptions made by an agent pertains to the pragmatics of the dialogue domain. We call them specific-domain presuppositions. Most of them characterize a default knowledge about the interlocutor. For instance, in some particular application, the server-agent has to consider that "generally, when a user requests to be informed about a flight arrival time, she/he knows the flight departure airport". This kind of belief has to be formalized in the same way as the other components of the belief system.

Hence, considering the previous example, we have:

$$K(s, Cont) \land \neg K(s, \neg K_{ref}(u, da(x))) \Rightarrow K_{ref}(u, da(x))$$

We have noted that presuppositions are generally default beliefs about the user and we formalize them as autoepistemic beliefs. Indeed, default reasoning and autoepistemic reasoning don't have the same motivation (Moore 1985). However, as shown in (Konolige 1988), the corresponding logics are formally equivalent. This result justifies our proposition, notably when the homogeneity of the reasoning model of a computer system has to be preserved.

8. Using the task model to plan request negotiation dialogue

We briefly sketch the way the dialogue system uses the task model jointly with the domain-specific user model to plan dialogue in order to complete a user's request, that is, to construct an identifying description according to the task model.

Let's consider a very simplified form of the plan-oriented model of the illocutionary act (Cohen & Perrault 1979, Searle & Vanderveken 1985) $Informerefi,j,reflx))$, denoting the event of an agent $i$ who informs an agent $j$ of an identifying name of the referent $ref$. The applicability precondition of this act is $K_{ref}(i, ref(x))$ and its intended (or, somehow, its perlocutionary) effect is $K_{ref}(j, ref(x))$. Suppose that the task-component of the belief system involves the rule

$$W(s, K_{ref}(s, r(x)))$$

and that among the domain-specific presuppositions, we have

$$K(s, W(u, K_{ref}(u, r(x)))) \land \neg K(s, \neg K_{ref}(u, h(y)))$$

$$\Rightarrow K_{ref}(u, b(y)))$$

Suppose also that we have the cooperation rule

$$K(s, W(u, \phi) \land \neg W(s, \neg \phi))$$

and the (simplified) belief transmissibility rule

$$K(s, K(\phi) \land \neg K(s, \neg \phi))$$

Suppose now that the current model of the world (resulting from a request of the form "What is the $r$ for which $1x a(x) = x_0$") involves

$$K(s, W(u, K_{ref}(u, r(x)))) \land \neg K_{ref}(u, r(x)))$$

$$K(s, u, a(x_0))$$

In the following, we don't mention all the inference steps, especially those related to the KD45 belief inference system. From (6), (1) and (8), in the one hand, and (7) and (9), in the other hand, we respectively have

$$W(s, K_{ref}(u, r(x)))$$

$$K(s, a(x_0))$$

Now, $s$ has a need to satisfy, namely (10). From (10), (3) and (8), $I(s, K_{ref}(s, r(x)))$ is inferred and gives rise to a planning process. Suppose that the control strategy uses a backward chaining. So, the instance $Informerefi,u,s,r(x))$ is selected and gives rise to a planning process. Suppose that the control strategy uses a backward chaining. So, the instance $Informerefi,u,s,r(x))$ is selected and gives rise to a planning process. Suppose that the control strategy uses a backward chaining.
its precondition is not satisfied (because of (6)) and is already a goal currently tracked. Here, the task model comes into play. (2), (4) and (10) allow to infer
\[ W(s, Kref(s, a(x))) \]
(12)
\[ W(s, Kref(s, b(x))) \]
(13)
(12) doesn’t give rise to an intention since \( K(s, a(x_0)) \) holds. But (13) allows to infer \( I(s, Kref(s, b(y))) \). The instance \( Informerref(u, s, b(y)) \) is then selected. Its precondition is satisfied because of (5). This completes the planning process. The resulting plan is the sequence of illocutionary act instances \( <Informerref(u, s, b(y)), Informerref(s, u, r(x))> \). Note that the first instance of the sequence is planned to be performed by the user. So, the server-agent has to make the user wanting its performance. This is brought about in a second planning step where the "want-precondition" (not mentioned here, for simplicity) of the selected act, has to be achieved. The planning step we described in this section is viewed as a task-oriented problem solving step.

9. Conclusion

The task modelling method we propose here is completely transparent to the dialogue system model and independent of a specific application. The result is a set of goal-oriented inference rules. As defined, these rules give a logical generic model of the task. They are syntactically and semantically compatible with the other belief system components and, thus, constitutes a common-place part of the logical reasoning model. Used jointly with the domain-specific presuppositions, which are formalized in an autoepistemic statement form, the proposed task model enables the dialogue system to plan request negotiation dialogues in a natural way. Furthermore, grounded on the set of dependencies, it allows the dialogue system to efficiently inquire the data base, avoiding frequent accesses which leads in most cases to degenerate answers.

References


