Design Considerations on Dialogue Systems:
From Theory to Technology
- The Case of Artimis -[1]

David SADEK

France Téleco
CNET– DIH/DIPS
Technopole Anticipa – 2, avenue Pierre Marzin – 22307 Lannion Cedex – France
David.Sadek@cnet.francetelecom.f

ABSTRACT

Significant features of an interactive system that makes the user view it as a user-friendly dialogue system are introduced and discussed. It is argued that it is onl when they are satisfied jointly and independently from dialogue state that they can appreciably affect the degree of system conviviality. It is also claimed that the way that the satisfaction of these features is handled, namely extensionally or as a consequence of a deeper "intelligence" of the system, determines a fundamental division in the approaches of dialogue system design. Classical (structural and plan-based) approaches to dialogue are discussed. Then, a new paradigm fo designing and implementing dialogue systems, that of rational dialogue agent, based on rational interaction approach, is introduced. It is shown how this paradigm and approach have led to Artimis, a technology of rational agency, which provides a generic framework to instantiate effective advanced dialogue systems.

1. INTRODUCTION

The emerging perspective of accessing information through the World Wide Web by a large, international, heterogeneous population brings the problem of the quality of interaction to the forefront. This makes it necessary to consider natural language as an appropriate, if not the most, media for interaction. Furthermore, the possibility of using spoken language is attractive fo many reasons, the main reason being that speech is the natural and fundamental communication vehicle and i the only modality available to most telephone users. Recent efforts show a growing industrial interest along this direction.

In the general case, interaction should be *user-friendly*. Obviously, the user-friendliness (or conviviality) criteria vary not only according to the application context but also according to the category of the potential user population. The reluctance of a user to use a system often depends on a particular non-satisfactory behavior of the system. Conviviality should be viewed and handled as a global feature of a system, and should "emerge" from it *intelligence*; it should not be handled as a set of specific features identified one by one by the system's or the interface's designer.

Most of the time, it is unlikely that a user can access the desired information with just a single query. The query might be imprecise, incomplete, intrinsicall inconsistent, or incoherent with respect to the dialogue history. It might not be completely understood by the system, especially in a context of spoken interaction, given that automatic speech recognition (ASR) and natural language understanding technologies are known to be imperfect. Furthermore, even if the request is well formed, unambiguous, perfectly recognized and understood, it is possible that the size of the answer is so big that the system has to consult the user in order to reach an agreement on the responses the user is really intersted in.

Clarification, completion and negotiation dialogues are underlied by strategies that should guide the user to provide the necessary elements of a "task" description that the system knows and considers practically and economically feasible. The final objective of the dialogue would be that of accomplishing the required "task" by performing the appropriate (communicative and/or non-communicative) actions.

Simple human-computer spoken dialogues usually implement rigid strategies. This makes ASR robust because rigidity requires the speaker to respond to a proposed menu with a small active vocabulary at each dialogue step.

As technology progresses, it becomes possible to reduce rigidity by allowing the user to speak with more freedom. This requires a more sophisticated dialogue component, with more detailed knowledge about the content and structure of the information repository, and a more elaborate capability of reasoning about the user's utterances (or messages, in general). In this view, *interface ergonomy* is largely based on *agent's intelligence* and *knowledge* (see [58] for more details on this consideration).

In user-friendly interaction, the system's intelligence is made complex by the need to interpret (in context) and

---

[1] A large part of this paper is derived from [63]. For the sake of conciseness, several references that appeared in [63] have been removed.

handle a very large variety of utterances (and, more generally, multimedia messages), and by the need to generate a large variety of (possibly multimedia) answers. This is due, in particular, to the fact that the granularity or the degree of completeness of the user's requests can vary significantly. Modern and useful telephone services, even without ASR but just using the telephone keyboard, should require a certain degree of "intelligence" to produce an answer like: "I do not have any subscriber by the last name of André neither in Antibes nor in the closest surrounding area, but I have ten in Nice, one of them being a women. If you want to select one of them press the two first letters of the name; if you want to select the subscriber women press 1; if you want the whole list press #; and if your are not intersted in any of them press *".

Note that even if communication in natural language, especially with speech, has obvious (and non-obvious) advantages compared to other media (such as direct manipulation), one can easily think of intelligent, convivial interaction with mouse input and image output.[2] The phenomena, problems, and interaction mechanisms introduced in this paper are illustrated with examples from natural language interaction, but they can be easily transposed to interaction contexts that involve other communication medias, possibly used jointly with natural language.

In this view, designing a service should not be approached as a twofold operation: (1) the definition of the service content or functionalities and (2) the designing of its user interface. Rather, a dialogue system should *embody* the service, and it is the system's intelligence that will guaranty a "good" ergonomy of the service.

Section 2 discusses some basic dialogue phenomena that need to be taken into account to design more ”intelligent” dialogue systems. Section 3 discusses the impact of user-friendly requirements on system design. Section 4 is about the computational models of dialogue. Section 5 introduces the principles of *rational interaction* to design dialogue agents. Section 6 discusses aspects of the use of natural language in advanced dialogue systems. Section 6 introduces some operational aspects of the Artimis technology. Some remarks conclude the paper

2. BASIC DIALOGUE PHENOMENA

A number of laboratory prototypes for human-computer dialogue systems have been developed in the past,[3] and more recently industrial applications have entered commercial trials. Most of these systems show the capability of chaining simple exchanges with a human user following a stereotypical structure and a limited application framework. In general, their ability to generate cooperative answers does not create the conditions of an interaction perceived as *natural* by humans, even for highly constrained and limited application frameworks. There are reasons for this. The first reason is the difficulty of integrating methodologies for artificial *intelligent* agents and human *natural* communication. This difficulty increases in the context of spoken dialogue because, in this case, communication is strongly affected by ASR errors.

Another reason for the limitation of dialogue systems is that most of the approaches consider dialogue as an isolated phenomenon, in which external manifestations have to be identified and reproduced by machines. This ignores the *cognitive* aspects of dialogue, i.e. the link between external manifestations and the internal *"intelligence"* of the machine.

Among the features of an interactive system that makes the user view it as a user-friendly human-computer dialogue system, the most significant are: negotiation ability, contextual interpretation, flexibility of the input language, flexibility of interaction, capability to produce cooperative reponses, and the adequacy of the response style. These features are now introduced and discussed.

2.1. Negotiation Ability

Negotiation between the user and the dialogue system is made necessary by the fact that the ultimate objective of the enterprise is *user satisfaction*. In fact, the approriate reaction may not correspond to what has been apparently requested by the user (even if all the words uttered by the user were perfectly recognized). In fact, the user might have deliberately formulated an incomplete request, like ”I am looking for a job around here”, or ”I would like to know the departure time of a flight to Paris”, waiting for the system to assist her/him in formulating more specific requirements [51]. The system may face different situations in addition to the just mentioned case of incompleteness.

The user's request may contain sufficient elements to access a data-base and provide an answer, but the number of possible answers might be too large. In this case the dialogue strategy is to negotiate with the user more constraints on her/his request, by signaling that the list of possible answers would be too long.

Another possibility is that there is no answer available for the user's request. In this case, the system has to advise the user and guide her/him to formulate a request that is compatible with the system's knowledge.

The user may also wish to negotiate new requests (on the basis of the answers given by the system), especially once she/he has a better understanding of what the system can provide.

---

[2] For instance, a "lodging" server that offers access to an image data-base can allows the user to select in various menus a set of search criteria. If the selected criteria have not identified the sought lodging precisely, the system will propose to the user images considered close to what is sought; the user can modify her/his request by clicking on the part of the image she/he wants to modify.

[3] For spoken dialogue see, e.g., [1,21,39,41,49].

All these cases may intertwine, thus increasing the complexity of the negotiation process. Sophisticated negotiation capabilities usually make the dialogue system more user-friendly.

2.2. Contextual Interpretation and Language Flexibility

An interactive system that expects the user to formulate her/his message in such a way that it can be understood independently from the context, can, in no way, be viewed as a convivial system. A fortiori, a "question-answer" system (i.e., a system which can only react to completely specified requests) cannot be really considered as a dialogue system.

An effective dialogue system has to be able to interpret a user's utterance with respect to the previously exchanged utterances. This capability is a necessary requirement fo real systems because users frequently use *ellipses* (e.g., once a list of flights for Paris has been required and obtained, the user might then ask: "And for New York"), *anaphoras* (e.g., "Give me the cheapest flight") or *deixis* (e.g., "Is a server of this type available from here?").

Note that in natural communication, contextual interpretation is the general case rather than the exeption, because as noted in [32] the communicated message is always *differential*: a person does not describe a (whole) situation; she/he tends to express the difference between the situation she/he wants to describe and what she/he believes that her/his interlocutor already knows. More specifically, during a dialogue, the interlocutors tend to make reference to an existing *situatio* in order to modify it, rather that making frequent initializations of new situations.

For a speech-based system to handle complex context dependencies, it has to be able to accept spontaneous speech and hence very large vocabularies. Whether or not its ASR component has to correctly recognize all th words might be debatable, but, certainly, the system has to be able to properly identify all the semanticall relevant words embedded in a very large variety of sentences.

In a multimedia system, the flexibility of the input language would involve, for example, that no constraint (such as specific order) be imposed on the way media can be combined, or that a media be not only reserved for a certain type of action.

2.3. Interaction Flexibility

There are certainly situations where the user might find it useful to be closely guided by the system. However, not being constrained to follow a pre-established structure, as is the case of navigating with a hypertext system, notably increases the comfort of a dialogue system. It should be possible for the user to engage into a clarification sub-dialogue before answering a system's question, or to change topic or dialogue objectives even before the completion of a (seemingly) mutually agreed dialogue task.

More importantly, the need for a non pre-established structure of interaction clearly appears wheneve communication troubles arise. Indeed, while it is acceptable that during a "consensual" proceeding of the dialogue, the user need not take the intiative to express her/his requests, she/he should be allowed to deviate from the "regular" course of interaction whenever she/he has identified a communication problem. Note that the possibility of distorsion between the message sent and the message received is common in communication. The problems of erroneous perception or interpretation [23] should be treated as the general case rather that the exception. It is therefore natural that the user can rectif a misunderstanding of the system, or correct her/himself. The more the contestation possibility is independent from the dialogue state, the less the user feels afraid to be engaged in one-way paths that she/he has not choosen. Only a system that allows for flexible interaction (i.e., without any pre-established intearction structure) can offer this possibility in a generic way (i.e., independently from the dialogue state).

In the general case, the possibility of contestation i directly related to the global capability of a system to revise its beliefs. So far, there are only a few systems that provide it. The reason for that is the difficulty to implement it in the approaches where dialogue i constrained to follow pre-established structures.

2.4. Cooperative Reactions

Let us consider the following question that could be asked at a booth of a railway station: "What is the departure time of the next train to Paris?". It is not unnatural to get an answer such as the following: "at 3:30 pm, track number 21". Consider also the following question "Do you know what time is it?" to which one can answer “yes! I do know”, but for which it is natural to expect an answer like "it is 2:30". The common aspect of these examples is that the answer extends beyond the question in a pertinent way. When this happens, the answer is said to be *cooperative*.

In human-human dialogue, cooperative responses are the general case rather than the exception. A friendl dialogue system should attempt to produce them too. There is a large variety of cooperative answers. Major types of *cooperative answers* are introduced in th following (see, e.g., [66] for other types of cooperative answers).

Completion answers

Completion answers (sometimes called over-informative answers) contain more information than what has been explicitly requested by the user (see in the example above the answer to the request about "the departure time of the next train to Paris"). The need to provide additional information and its content have been *inferre* from user requests and dialogue history.

Corrective answers

Corrective answers are given when questions impl certain things to be true, while, in reality, they are not. For example the question: ”At what time does the flight from Montreal to Tucson stop in Chicago?” implies that there is a flight from Montreal to Tucson (which is not true) and that it stops in Chicago (which, consequently, is also not true). Obviously, the answer to a question like this has to be a corrective one, like: ”There are no flights from Montreal to Tucson”.

Suggestive answers

There are questions for which the answer should be "negative", but from which it is possible to infer that the user might be interested in the answer of another question of a related topic even if she/he did not mention this topic explicitly. This is the case of answers of the type ”No, there no train to Grenoble today *but there is a train to Lyon and from there you can take a shuttle bus to Grenoble*”. Suggestive answers may be welcome even in the case of positive answers, like: "Yes, we have single rooms without a TV for 500 Francs, *but you can get a nicer room with a TV for just 100 Francs more*". In certain cases, a reasoning process may infer that th answer to a question related to the one the user has effectively asked may be more informative and appreciated by the user.

Conditional answers

There are questions for which there may be positive answers under conditions that constrain the user request beyond her/his intention. For example: ”Give me the flights to Detroit before 7:00 AM” may receive a conditional answer of the type: ”There is a flight at 6:50 AM in weekdays only”.

Intensional answers

There are cases in which providing a "factorized" answe not only makes it more presentable and understandable to the user (especially in the context of spoken communication) but can also inform her/him about the genericity of its sematic. For example, the question: ”Give me the list of all the non-smoking flights between Canada and the Netherlands”, for which an intensional answer would be: ”All KLM flights”.

It is worth noting that calculating the intensional form of an answer on the basis of an extensional set of solutions may require a relatively complex inference process.

2.5. Adequacy of Response Style

For a dialogue system, not only must it determine the right reaction, but it must also determine the appropriate way to present it to the user. In particular, it is important for the system to be able to choose the media, or some combination, for the answer as well as the appropriate level of verbosity. For example, if the answer to a question is a list of 30 flights, it is better to display it on a screen, if this is possible, rather than using spoken output. If this is not possible, it may be necessary to negotiate with the user in order to arrive at an answer of an acceptable size. In the case of graphic output, colors, font type and size, and layout of the answer may affect user satisfaction.

More generally, "formatting" the answer, "factorizing" it according to semantic criteria, choosing the best verbosity level and the right media (or combination) are operations which contribute to the quality of a dialogue system reaction.

3. IMPACT OF USER FRIENDLY REQUIREMENTS ON SYSTEM DESI

Even though setting a list of conviviality criteria such as those introduced above may be crucial for external specification and for evaluation of dialogue, this cannot be taken as a methodological basis for system design.

The first reason is that the impact on system design of such a list of criteria is limited by the fact that it is potentially open to extensions, such as the need for responses in real time, and/or customization, and ma depend on the application domain.

But more importantly, most of these criteria are interdependent, and it is only when they are satisfied jointly and generically (i.e., independently from dialogue state) that they can significantly affect the degree of system conviviality.

Even if their external manifestations are different, these criteria rely on the same basic mechanisms. Therefore, aiming at satisfying them one by one is neither an optimal nor a generic approach. Thus, the designer should attempt to meet most of them with a global view. Importantly, the way that the satisfaction of these criteria is handled, namely extensionally or as a consequence of a deeper "intelligence" of the system determines a fundamental division in the approaches of dialogue system design.

4. COMPUTATIONAL MODELS FOR DIALOGUE SYSTEMS

4.1. Structural Approaches

Structural approaches are mainly based on finite state diagrams or on dialogue grammars. They can be anchored by a computer science background (see, fo example, [68,9,66,17]) or by a linguistic background (see, for example, [13,40,4,35]) and attempt to model regularities in human-computer or human-human dialogues. Structural approaches are based on the assumption that there exists a regular structure in dialogue, and that this structure can be represented "finitely" (for example by a finite state automaton or a context-free grammar). They consider that dialogue coherence is residing in its structure and are therefore descriptive: they are based on descriptions of observed sequencing of utterances, not on explanations of what is observed. Although effective practical systems can be built with these approaches, these systems tend to appear rigid to the user and limit her/his degree of satisfaction.

The structural approaches concentrate on the *co-text* (i.e., the text that "comes with"), leaving away the *contextual* nature of communication. These limitations rule out these approaches as a basis for computational models of *intelligent interaction*. Other approaches (such as described below) are required. The implementation of these approaches is more complex as is their specification. ecent progress has shown that prototypes of these systems can be built and tested. These systems have inference engines and knowledge representations with which real-time implementations is now possible. This motivates a great interest in their study.

4.2. Classical Plan-Oriented Approaches

Classical plan-based or differential approaches (see, e.g., [15,2,48,33,11,28,12,67]) consider a communication intervention not only as a collection of *signs* (e.g., a sequence of words), but also as a realization of observable communicative actions, also called *speec acts* or *dialogue acts* or *communicative acts*, such as *inform, resquest, confirm, commit, etc.* [4] Language is viewed as a means for indentifying and instantiating a common context within dialogue partners.

These approaches follows the philosophical principle that ”communicating is acting” [3]. They rely on the idea that communicative actions, similarly to (physical) non-communicative actions are oriented toward goal achievement, and are planned with this motivation. In this view, the objective of communication is to change the *mental state* (including beliefs, intentions, etc.) of the interlocutor. Thus, communicative actions can be planned and recognized, as regular actions, on the basis of mental states. Dialogue analysis is considered in the framework of explaining actions based on mental states, relying on general models of actions and mental attitudes.

In practice, it is assumed that persons generally have *goals* and *plans* in mind when they interact with other persons or machines. The purpose of a dialogue is to recognize such goals and plans, and to produce effects corresponding to the *purpose* of the plans.

The plan-based approaches can be enriched with structural models of discourse [26,34].

4.3. Comments On The Classical Plan-Based Approach

Fundamental work about the differential (or plan-based) approach to natural communication [15,2,48,8] has shown that the philosophical theory of speech acts can have a formal foundation in the theory of action. In spite of the large popularity gained by the plan-based approach, it is still difficult to formulate, within this framework, a "coherent", global solution to the problem of user-friendly, ”natural” human-computer communication.

One reason is that the use of logical formulation fo mental attitudes is weak (if not hazardous) if the interpretation of the formalism remains intuitive. For example, the concepts of *belief* and *intention* which are fundamental components in communication and cooperation philosophy [24], cannot be properly used without an adequate model of their semantics. Note that analysing these concepts is not interesting for the stud of communication (and cooperation) only, but more generally for modelling the background common to th so-called intelligent behaviors. In fact, the more general problem is the maintenance of the *rational balance* existing between the different mental attitudes of an agent, and also existing between the agent's attitudes and plans and actions. In the classical plan-based approaches, the relation between mental attitudes (in particular intention) and action, is purely operational. The lack of explicit ationality principles and their (logical) links with action models is an important limitation, not only of this approach, but more generally of the work related to the concept of intelligent agency.

A second reason is that since a communicative act is, in principle, an action, its modelling has to cope with the problems, well-known in AI, related to characterizing action effects and preconditons (such as the fram problem [37], the qualification problem [38], or the ramification problem [21]. In particular, the effect of a communication act on its recipient (and also on it author) depends on her/his mental state before the act is performed. A more general problem underlying this phenomenon is that of *belief reconstruction* [57,52] following the observation of a communicative act (or, more generally, an event). Belief revision (and its couterpart, consistency preserving) following the "consommation" of an event is an aspect of this problem. Thus, characterizing the effect of an act as a function of the mental context in which the act has been performed turns out to be a difficult issue.

The third reason is that the differential approach has, somewhat, misestimate the question of the criteria fo determining the act *types* to model, and, for a given act, the question of the specification of the mental attitudes it encapsulates (more details can be found in [53,15]).

---

[4] Independently from a given characterization and taxonomy of *speech acts*, a *communicative act* can be defined as an act that is produced to be observed by (at least) one other agent, thus aiming at causing a change in his mental state [52,53]. An agent uses this behavior to communicate an intention. If the term "speech" in "speech act theory" is taken to mean "instrument of communication", then the notions of communicative act and of speech act are identical. Notice that non-linguistic acts, such as actions of referring in pointing at objects, are communicative acts.

5. RATIONAL INTERACTION

5.1. Motivations

The rational interaction approach can be viewed as a recasting, in a comprehensive formal framework, of the plan-based approach, and as adopting a "radical" view of communication as a special case of intelligent behavior. It is thus based on the assumption that a system capable of carrying on an ”intelligent” dialogue has to be an *intelligent system*, in which the communication ability is not primitive, but is grounded on a more general competence that characterizes *rational behavior*.[5]

In a simplified way, for an agent to behave rationally is to be permanently driven, at a certain *representation level*, by principles that optimally select actions leading to those future in comformity with a given set of motivations and desires. It is at this (hypothesized) *Knowledge Level* [43] that the concepts of mental attitude and intentional action are relevant.

The first most significant contribution to the rational interaction approach is Cohen and Levesque's work (see, in particular, [14,15]), which provided a robust methodological framework for expressing formal theories of intention and communication. Even though their account suffered from a certain number of theoretical modelization problems (see [52,54]) and handled only some aspects of rational behavior and (cooperative) interaction, it had been the first rigourous, formal analysis of intentional action and communication.

A conversational system is a set of human or artificial agents each one of which is capable of accomplishing communication acts and of interpreting communication acts of other agents.

For each act of a conversational system, three agent types are considered, namely *the author*, the *recipient* and the *observer*. The author as well as the recipient are also observers. The accomplishment of the act also modifies the mental state of the author who, among other things, has to add to her/his *belief system*, the fact that the act has just been performed [57,52].

In regard to the formal framework to couch this approach, the logic representation is adequate for various reasons: its homogeneity, its genericity (due its large coverage), its ability to properly describe mental states (which makes it easy to maintain), and its potential usability as a tool for both modelization and implementation.

A delicate problem concerning the implementation of these systems is the need of methods and procedures for *automated inference* (or *theorem proving*), the heart of the reasoning system, with acceptable time (and space) complexity. Some interesting solutions have been proposed in some of the above mentioned systems.

The rational interaction approach leads to a new paradigm for designing dialogue systems, that of *dialoguing rational agent* . The Artimis technology developed at France Télécom - Cnet is an effective implementation of this paradigm.

As a framework to dialogue system design, the formal approach of rational interaction relies on two basic ideas. The first one is that a dialogue process can be completely justified by rational behavior principles (which are more basic that discourse rules), and does not require, in principle, any structural model; instances of dialogue structure dynamiclly emerges from the dynamics of rationality principles. The second one is that the same logical theory can account for different aspects of rational behavior, in particular in situations of cooperative dialogue.

Due to the genericity of its principles, this approach achieves the *robustness* required by a(n intelligent) dialogue system: to soundly react to complex situations, possibly incompletely specified when the system has been designed.

5.2. Mental attitudes

Two basic notions are at the centre of the intelligent behavior modelling, namely mental attitude (or more generally Intentionality [65]) and action. The approach aiming at designing computational models of communication, which comes within the more general framework of rational action, consequently relies on these two notions.

Mental attitude can be intuitively viewed as the relation between an agent and a situation (identifiable to a proposition) or an object (in a general sense).

In logical terms, a proposition is a *belief* of an agent if the agent considers that the proposition is true. Belief is the mental attitude whereby an agent has a model of the world. An agent's belief system is the mean that allows the agent to maintain and update its representation of the external world.

Uncertainty is a way of representing an approximate perception of the world. In the formal framework for rational interaction [52,53], uncertainty is not represented by any type of degree, not even qualitative; it is handled in a global way and expresses the fact that the agent believes that a proposition is not true, but it is more likely than its negation.

Intention is strongly related to action. It is the mental attitude whereby an agent can determine and control his evolution. It is a composite concept, which cannot be analysed independently from the other mental attitudes [14]. Formally, an agent has the intention to bring about a proposition if and only if :[6] (1) the agent believes that

---

[5] The most consensual achievement of intelligent behavior is *rationality*. An overview of different aspects of the notion of rationality can be found in [50]. See also [19], for example, for an approach to *economic rationality*.

[6] Note that while the object of belief or uncertainty is necessarily a proposition, intention might have an action (that is a term in the logical sense) as object.

the proposition is false; (2) the agent *chooses* to evolve toward futures in which the proposition is true and to act coherently with this choice; and (3) the agent commit himself to maintain this choice until the agent comes to believe that the proposition holds or he comes to believe that the proposition is impossible [7]

It may be debatable to ascribe mental attitudes to machines. Whether or not a machine can have mental attitudes is probably undecidable. What is relevant is th possibility of considering a machin *as if* it has such or such mental attitude, *through the causal role this mental attitude plays in its behavio* . As analysed by Dennet [17], the problem is not to know whether a system is really intentional or not, but to know whether it can be considered as such coherently. As McCarthy [36] pointed out, it is legitimate to consider a machine as having mental attitudes if this is useful.

5.3. Bases of the formalism

Concepts of mental attitude (belief, uncertainty, intention[8]) and action are formalized in the framework of first order modal logic described in detail in [52,54].
For the sake of brevity, only the very few aspects of the formalism, required for the presentation, are introduced in the following.
Symbols $\neg$, $\wedge$, $\vee$ and $\Rightarrow$ represent classical logical connectives of negation, conjunction, disjunction and implication, while $\forall$ and $\exists$, are respectively the universal and existential quantifiers.
Symbol p represents a closed formula denoting a proposition. ϕ, ψ and δ are formula schemas. i, j and h are variable schemas denoting agents. Notation $\models \phi$ means that formula ϕ is valid.
Mental attitudes considered as semantic primitives, i.e., belief, uncertainty and choice (or preference, or, to some extent, goal) are formalized by the modal operators B, U and C respectively. Intention is taken here as a macro-attitude and formalized by the modal operator I. Formulae of type $B(i,p)$, $U(i,p)$, $C(i,p)$, and $I(i,p)$ can be respectively read as: "i believes that p is true", "i is uncertain about the truth of p", "i desires that p be currently true", and "i intends to bring about p". In the following, we only mainly use belief and intention.

The formal model consists of a set of logical axioms (and derived properties) that formalizes action models, basic principles for rational bahavior and introspection, communication and cooperation [52,54,57]. They can be summarized as follows:

5.4. Rationality principles and action models

The components of an action model, in particular, a communicative act (CA) model that are involved in a planning process characterize both the reasons for which the action is selected and the conditions that have to be satisfied for the action to be planned. For a given action, the former is referred to as the *rational effect* (RE),[9] and the latter as the *feasibility preconditions* (FP), or th qualifications of the action.
Two rationality principles relate an agent's intention to her plans and actions. The first principle gives an agent the capability of planning an act whenever the agent intends to achieve its RE. It states that an agent's intention to achieve a given goal generates his intention to be in a situation where one of the acts known to the agent, whose rational effect (RE) corresponds to the agent's goal, has just been done; and that the agent has no reason for not doing them. Formally, this is expressed by the following property:

$$|=I(i,p) => I(Done(a_1|...|a_n))$$

where a_k, k ranging for 1 to n, are all the actions such that: (1) p is the rational effect of a_k (i.e., the reason for which a_k is planned); (2) agent i knows action a_k: $Bref(i,a_k)$; *and (3)*$\neg C(i, \neg Possible(Done(a_k)))$.[10]

The second principle imposes on an agent, whenever the agent selects an action (by virtue of the first rationalit principle), to seek the satisfiability of its FPs.[11] It states that an agent having the intention to be in a situation where some action be done, adopts the intention that the action be feasible, unless he believes that it is alread feasible. This is formally expressed as follows:

$$|=I(i,Done(a))=> B(i,Feasible(a)) \vee I(i,B(i,Feasible(a)))$$

If an agent has the intention that (the illocutionar component of) a communicative act be performed, h necessarily has the intention to bring about the act RE. The following property formalizes this idea:

$$\models I(i,Done(a)) \Rightarrow I(i,RE(a))$$

where $RE(a)$ is the rational effect of act a.
Consider now the opposite aspect of CA planning: the consommation of CA's. When an agent observes a CA, he has to come to believe that the agent performing the act has the intention (to make public his intention) to achieve the act RE. This kind of act effect is called the *intentional effect*. The following property captures this consideration:[12]

$$\models B(i,Done(a) \wedge Agent(j, a) \Rightarrow I(j,RE(a)))$$

---

[7] See [14,54,52] for formal theories of intention relying on the bases just introduced.

[8] Intention is defined in tems of belief and choice.

[9] This effect is also referred to as the *perlocutionary effect* in some of the author's previous work, in analogy with the use of the term in speech acts theory.

[10] In the third condition, for effectiveness needs, an I operator (intention) can be reasonably substituted to the C (choice or desire) operator.

[11] See [53] for a generalised version of this property.

[12] Precisely, this property is as follows: $\models B_i(Done(a) \wedge Agent(j, a) \Rightarrow I(j,B(i,I(jRE(a)))))$.

Some FPs persist after the correponding act has been performed. For the particular case of CAs, this property is valid for all the FPs which do not refer to time. Then, when an agent observes a CA (and "admits" it), he has to come to believe that the persistent FPs holds:

$$|\!=B(i,Done(a) \Rightarrow FP(a))$$

A model of rational action should specify feasibilit preconditions and the rationale of the action. Th expression of such a model is, in general, complex fo two main reasons. The first is that the set of action qualifications is potentially infinite (see [53] for the case of communicative acts). The second reason is that the effect of an action on the world is strongly context-dependent and cannot be formulated in general terms [47,57]; furthermore, it is difficult to "summarize" what an action should leave unchanged.

A solution that goes round the problem of effect specification is directly related to the expression of the rationality principles. In fact, if it is not possible to specify the actual effects of an action, it is still possible to state (in a logically valid way) what is expected from an action, that is what are the reasons for which the action has been selected. This is exactly what i expressed by the first rationality principle. This semantics for action effect, within the framework of a model of rational behavior, allows one to overcome th problem of effect unpredictibility.

The set of feasibility preconditions for a CA can be split into two subsets: the *ability precondition* and the *context-relevance preconditions*. The abilit preconditions characterise the intrinsic ability of an agent to perform a given CA. For instance, to *sincerely assert* some proposition *p*, an agent has to believe that *p*. The context-relevance preconditions characterize the relevance of the act with respect to the context in which it is to be performed. For instance, an agent can be intrinsically able to make a promise while believing that the promised action is not needed by the addressee. The context-relevance preconditions may correspond to the Gricean quantity and relation maxims [24].

As an example, is a simplified model (as far as th expression of the preconditions is concerned) of the comunicative act of an agent *i informing* an agent *j* that a proposition ϕ holds, is as follows:

$$<i, Inform(j,\phi)>$$
$$\text{FP: } B(i,\phi) \land \neg B(i,B(j,\phi))$$
$$\text{RE: } B(j,\phi)$$

This model is directly axiomatized within the logical theory through the above mentioned rationalit principles and the following schema:

$$B(h,Feasible(<i,Inform(j,\phi)>) <=> B(i,\phi) \land \neg B(i,B(j,\phi)))$$

Notice that actions are not handled by a planning process as data structures, as in the case of the classical plan-based approach, but have a logical semantics within the theory itself.

Notice also that the two first rationality principles specify by themselves, without any non-logic artifact, a planning algorithm that deductively generates plans of actions, by allowing the inference of causal chains of intentions.

Coherence relations (that could also be viewed, in some sense, as properties of rational behavior), such as the consistency of an agent's beliefs and the relations between intentions and beliefs, together with rationalit (and cooperation) principles, assess a sound framework for the *rational balance* between the different mental attitudes of an agent on one side, and between mental attitudes and action plans on another side. Here are some examples, that respectively express that an agent maintains the consistency of her beliefs, that an agent will not bring about a situation if she believs that it already holds, and that an agent cannot be uncertain of her own mental attitudes:

$$|\!= B(i,\phi)\Rightarrow\neg B(i,\neg\phi)$$
$$|\!= I(i,\phi) \Rightarrow B(i,\neg\phi)$$
$$|\!= \neg U(i,M(i,\phi))$$

where M represents any modality, and therefore belonging to $\{B,\neg B, C, \neg C, U, \neg U, etc.\}$.

5.5. On Cooperative behavior

Cooperative answers are a significant manifestation of cooperative behavior. However, their role must be relativized with respect to this behavior taken in a generic way.[13] Firstly, they must not be implemented as response schemas to be instantiated systematically: thei production totally depends on the context. For example, it is wrong to write a specific rule in a dialogue system, stating that if there is a request of the departure time of a train, the answer should also specify the number of the track. Such a rule would appear inadequate every time the context would allow one to infer that the user does not intend to take that train, but she/he just wants to check of the validity of the schedule in her/his possession.

Analogously, an agent should not be blindly cooperative: for example, the agent has to take into account the degree of confidentiality of the information she is handling. In conclusion, providing cooperative answer should appear as a global predisposition of an agent to behave cooperatively, that materializes in a way depending on the context of interaction and still governed by the principles of rational behavior.

An agent's predisposition to cooperation is underlyed b minimal commitments that aim at adopting and

---

[13] Grice's cooperation principle and maxims [24,25], though stated infomally, are a reference point for charaterizing cooperative (and "efficient") contribution to a dialogue.

satisfying the partner's intentions, and therefore, firstly at recognizing them.

Intention recognition includes the complex operation of reconstructing communicative acts from linguistic forms. This reconstruction is, at a first step, literal, and reveals the *communicated meaning* only following the inference of causal chains of intentions. This inference process allows, in particular, to understand *indirect speech acts* [5]. The literal meaning of an indirect speech act reveals a ”weaker” intention compared to the real intention of the speaker. This may indicate a lack of information of the speaker about the information the system can provide. For example, the question ”Do you know the telephone numbers of the City Hall?” may express the request of getting such a phone number if available. It can also be just a matter of politeness like in ”It is warm here” meaning in reality ”Please, open the window”. A non-literal reaction to indirect speech acts illustrates an elementary form of cooperation in situation of communication.

The above mentioned commitments can be expressed in terms of more primitive principles of cooperative interaction such as intention adoption, sincerity and pertinence [55]. These principles can be expressed in terms of mental attitudes of belief and intention, and, therefore, be formalized in the theory of rational interaction sketched above.

5.6. Cooperation principles

Let us consider some of the basic principles that express the motivations for agent to behave in a cooperative manner. The *minimal principle of cooperation* (or *the intention adoption principle*) states that agents must not only react when they are addressed but, more than that, they must adopt the interlocutor's intention wheneve they recognise it, and if there have no objection to adopt it. In other words, *if an agent i believes that an agent j intends to achieve property p, and that itself does not have an opposite intention, then i will adopt the intention that j will (sometime) come to believe that p holds.* Formally, such a principle translates into the following schema [52]:

$$(B(i,I(j,\phi)) \wedge\neg I(i,\neg\phi)) => I(i,B(j,\phi))$$

Another cooperation principle is *Belief adjustment* , which enables an agent to produce of corrective answers, for example. Such answers are produced with the intention of correcting a belief that is considered wrong. Such a belief is usually a presupposition (inferred by *implicature* [25]) from the recognized communicaive act. A corrective intention arises in an agent when her belief concerning a proposition, about which she is competent, is in contradiction with that of her interlocutor. Formally this property is expressed by the following schema:

$$B(i,\phi\wedge B(j,\neg\phi) \wedge Comp(i,\phi))=> I(i,B(j,\phi))$$

where competence is formally defined as follows:

$$Comp(i,\phi) \equiv B(i,\phi)=>\phi \wedge B(i,\neg\phi)=>\neg\phi$$

Other properties characterizing cooperation capabilities are similarly handled as basic behavior principles and formalized as logical axioms (or derived properties). Some examples are the following:

Sincerity: An agent *i* cannot have the intention that an agent *j* comes to believe a proposition *p* without believing *p* herself or without having the intention to come to believe it.

Pertinence: if an agent *i* intends to let an agent *j* believe a proposition *p*, then *i* has to believe that *j* does not believe *p* yet.

Harmony with other agents: In a multi-agent context, the behavior adopted by a cooperative agent toward other agents has to be basically a generalization of the behavior of the agent with respect to herself. For example, a cooperative agent should not cause a loss of information to other agents. In particular, she does not have to bring about uncertainty for other agents, unless she considers that uncertainty is the "right" attitude with respect to a propositon; and this supposes that she i adopting this attitude herself too.

5.7. Domain-dependent Constraints

Usually, two levels of cooperative behavior in dialogue are identified, namely the *communication* and the *domain* levels, and therefore, two types of intentions to be recognized. Recognition of intentions at the first level is the basis for the detection of indirect speech acts, while the intentions recognized at the second level, which underly domain plan recognition, are the starting point for the production of the over-informative, corrective or suggestive components of cooperative answers. Indeed, an over-informative answer can be produced to remove an obstacle in the (recognized) plan of the dialogue partner 2]. A corrective answer ma signal the reason for the failure of the partner's plan [42]. A suggestive answer may propose an alternate plan to satisfy the partner's intentions.

In practice, the process of plan recognition can be complex and expensive [5,31], particularly when the application domain is rich. This requires one to have a model of all the domain-dependent actions that may be implicitly or explicitly referred to in the communication, which is relatively a heavy task. In the framework of informative dialogues, in particular for data-base query it is possible to emulate the plan recognition process b introducing specific functions. These functions consider the ”domain” component of a speaker's message as a set of constraints which, in a default situation, have to be augmented, reduced or substituted. Thus, they can perform the following types of operations:

- find the reasons for the failure of a request, when, for example, the answer is empty or negative;
- compute a solution to a request *close* (according to a given distance criterion) to the request that was asked,
- find information to add to what was strictly explicited by the request,

- find the appropriate information to be negotiated to constraint a request for which the set of possible answers is too large.

These procedures can be generic with respect to the computation they perform (see, e.g., [30,27,20,29, 61,58]) but, in general, they "freeze" schemata for plan recognition. In relation to the global process of rational interaction, these functions are *black-boxes* that do not affect the overall logical integrity. They produce intentions as a "regular" plan recognition process, and, especially, they can be integrated, in a natural way (as meta-predicates) into the global logic model.

For example, the production of an intention leading to ”over-information” (i.e., richer information than what was initially required) may result from the following property: *if an agent i has the intention that an agent j believes proposition p and i thinks (through its function of "over-information") that q is pertinent in the ”stream” of p, then i will adopt the intention that j comes to be aware (i.e., to believe) that q.*

The black-boxes functions for domain-dependent constraint management are directly accessible from the logical framework characterizing the behavior of a rational agent [5]. As an example, the access to the *over-information* procedure is made by the following schema, where *OVERINF* is a meta-predicate:

$$B(i,(I(i,B(j,p))) \wedge OVERINF(p,q)) => I(i,B(j,q)))$$

This formula expresses the following property: *if a agent i has the intention that an agent j believes a proposition p and i thinks (because of its over-information function) that proposition q can be a pertinent over-information of p, then i adopts the intention that j comes to be aware (i.e., to believe) that q.*

6. DIALOGUE AND NATURAL LANGUAGE

The understanding and reaction determination process of a dialogue system (and therefore of a rational agent) functions on formal representations of semantico-pragmatic information, such as dialogue acts. Mental states are reconstructed, through inference processes, on the basis of observing and "admitting" dialogue acts. Such acts are obtained by sentence interpretation and formally represented as logical formulae.

Going from language to these representations implies that the system has, in addition to its language perception and production components (e.g., speech recognition and speech synthesis systems), interpretation and generation capabilities for the language(s) used by its interlocutor(s). More than for other media, in the case of natural language, the design and implementation of such mechnisms is a whole issue.

As mentioned above, the flexibility of the input language is a requirement for cooperative dialogue. Obviously, the more flexible the input is, the more sophisticated are th required mechanisms that have to be implemented fo the extraction of the semantic information relevant fo dialogue proceeding [10]. This problem of utterance analysis and interpretation is particularly salient in the context of oral interaction. Indeed, in such a context, not only is it difficult to constraint the user to a speech mode which may be in between a command language and spontaneous speech, but the inescapable presence of speech recongnition errors makes it more difficult to find robust cues to anchor meaning extraction.

Similarly, natural language generation starting from formal representations is an issue per se. In a dialogue context, the difficulty is increased by the fact that the linguistic realization of a given dialogue act strongl depends on the dialogue state. There is no bi-univoc relation between communicative acts and utterances: an act sequence can be realized by a single sentence, and a single act may be verbalized by a complex utterance. Moreover, the naturalness of the produced utterances depends on the system's capability to generate typical linguistic phenomena such as ellipsis, anaphoras, and to take into account the linguitic behavior of the interlocutor [44].

For both interpretation and generation, the efficiency of the implemented mechanisms is generally a function of the degree of genericity of the relationships set up between natural language and "mental representation" [6].

7. THE ARTIMIS TECHNOLOGY OF FRANCE TÉLÉCOM - CNET

Artimis[14] is an agent technology developed by France Télécom - Cnet, that provides a generic framework to instantiate *intelligent dialogue agents*. These agents can interact with human users as well as with other software agents. When instantiated in a human-agent interaction context, Artimis-like agents can engage in mixed-initiative cooperative interaction in natural language with human users. The resulting systems are able to display advanced dialogue functionalities, such as negotiating the user's requests, producing cooperative answers (which involves relevant, possibly non explicitly requested information), performing complex tasks, etc.

Roughly speaking, the Artimis software consists of three main components 62] a rational unit (which constitutes the heart of the technology), a domain knowledge representation and management unit, and a natural language processing unit including two components for understanding and generation.

The rational unit is the decision kernel of the agent. It endows the agent with the capability to reason about knowledge and actions. It performs cooperative rational reaction calculus producing motivated plans of actions, such as plans of (or including) communicative acts. In this framework, communicative acts are modelled a

---

[14] Artimis is a french acronym for " *Agent Rationnel à base d'une Théorie de l'Interaction mise en oeuvre par un Moteur d'Inférence Syntaxique*".

regular rational actions, thus enabling the agent to handle interaction.

The two natural language components [62,44] are essential to use the technology in a context of interaction with humans. They bridge the gap between the communication language (which, in this case, is natural language) and the internal semantical knowledge representation, namely Arcol,[15] the Artimis' interagent communication language, in terms of communicative acts with semantic contents expressed in a powerful language: a first-order modal language.

Without the two natural language components, *the rational unit is an intelligent communicating agent* (in a context of human-agent interaction, the human user is viewed as a particular agent; no assumption is made about the interlocutor's type).

The Artimis model is the formal theory of rational interaction sketched above [52]. The system involves a homogeneous set of generic logical properties, which embodies the core potential of the system. This potential is independent of its specific use in a given application domain. An inference engine, which is a theorem prove [5,7] faithfully executes the theory.

Artimis is a stand-alone software package, currentl integrated in a speech-telephony-computer platform (*i.e.*, a speech recognition software, a speech synthesis software and a telephony board/software). Artimis currently works in lab versions of several real applications. On of these test applications is AGS,[16] a directory of voice servers hosted by France Telecom (Audiotel servers). The resulting system, Artimis/AG [61,60], is a prototype of a cooperative spoken dialogue system, applied to the areas of "employment" and "weather forecasting" of the Audiotel servers directory [62].

Example:

Natural Language Analysis / Interpretation

The sentence analyser/interpreter (i.e., understanding) produces the best coherent interpretation from the most likely word sequence output by the speech recognizer. The goal is to reconstruct, as far as possible, in a logical form (namely, in Arcol), the dialogue act realized by the input utterance. The utterance analysis is based on detecting "small" syntactic structures which potentially activate semantic entities.

For example, let us consider utterance: *"Je voudrai connaître le numéro d'un serveur météo pour la region de Lannion"* ("I'd like to know the number of a server for weather forecasts in the Lannion area"), recognized b the speech recognition component *as "Je voudrais X météo pour X Lannion"* ("I'd like X weather forecast fo X Lannion"), which in turn, activates the concepts of *intention of the user*, *weather forecast server*, and *Lannion* (a city in Brittany, North-West of France).

Starting from the set of the activated concepts (which can be a list of possibilities in the case of nondeterminism due to syntactic overlapping), a *semantic completion process* finds other possible "ingredients" to be added to the detected concepts to build a well-formed logical formula that represents the semantic content of the dialogue act. This process is based on the hypothesis of *semantic connectivity* in the user's utterance, and assumes that the semantic content of the utterance corresponds to a path in a semantic network describing the application domain knowledge.

In the previous example, the inferred semantic complements are that the question deals with a *telephone number* of a *serve* about a *topic* that is weather forecast and whose *geographic domain* is Lannion.

Formally, this leads to the production of the following dialogue act:

$$<u, Inform(s,I(u, Bref(u,\iota x\ numtel(x) \wedge \exists y\ server(y) \wedge number(x,y) \wedge topic(y,weather\text{-}forecast) \wedge domain(y,lannion))>$$

meaning that the user (u) *informs* the system (s) that she/he wants ($I(u,...)$) to know ($Bref(u,\iota x...)$) the telephone number of a weather forecast server for Lannion.

It is worth noting that semantic-island driven analysis and semantic completion ensure a syntactic and semanti robustness to the analysis/interpretation process, particularly required in the context of spontaneous speech.

Attitude Inference and Reaction Planning

The rational unit implements the formal theory of rational intraction. It gives the system its dialogue abilities, which result from explicit reasoning processes. The *inference engine* that supports the rational unit is a theorem prover for first order modal logic, based on a "syntactic" approach (extended modal resolution and schema instantiation by sub-formulae unification) [67].

For instance, let us assume that, after utterance analysis/interpretation, the recognized communicative act is that *the user wants to inform the server that she/he wants to know if p (e.g., if the server 08-36-68-02-22 i operated by Météo France)*. On the basis of the rationality principles, the system infers *the intention of the user to know if p*. The coope ation principles allo the system to adopt *the intention that the user eventually comes to know if p*. Again based on the rationality principles, the system adopts *the intention of informing the user that p or informing her/him that not p*. The system then selects which one of these two actions i currentl *feasible* (for example, if the system believes *p*, it will be the action of *informing the user that p*) and

---

[15] Arcol is an acronym for ARtimis COmmunication Language. Arcol has been adopted in 1997 by the FIPA standardization consortium as the basis of the standard ACL, the inter-agent communication language.

[16] AGS is a French acronym for "Audiotel - Guide des Services".

transmits the selected action (i.e., the dialogue act) to the natural language generator.

Natural Language Generation

The natural language generation subsystem verbalizes the plan of dialogue acts produced by the rational unit, by producing a utterance (a sequence of words) relevant to the current linguistic context, as an answer to the user. Natural language generation follows two highly coupled phases [46,44,45]. The first one determines the surface acts (that specify, in particular, declarative, imperative, interrogative modes) and reference acts (that specify, in particular, the designation modes: nominal groups, pronouns, proper nouns, etc.), that achieve the dialogue act(s) to be sent to the user. In the second phase, the best formulation of the acts specified in the first phase is found, depending on the linguistic resources actually available (e.g., language, lexicon, grammatical structures) and on the current linguistic context (dialogue state, vocabulary and syntax used by the user in the previous dialogue turn, preferences of the user group, etc.).

For example, if the system wants to inform the user that there is a relation of type *provider* between the "Côtes d'Armor" weather forecast server and "Météo France", the system would send the following message to the generator:

$$<s, Inform(u, provider(\iota y\ (server(y) \wedge topic(y, weather\text{-}forecast) \wedge domain(y, cotes\text{-}d\text{-}armor)), meteo\text{-}france) >$$

The generator will produce (in natural language: currently French, English, or Russian), depending on the context, a declarative sentence with a proper noun and a nominal group, or a positive answer with a pronoun and a proper noun, or an elliptic sentence with a proper noun:

"Le fournisseur du serveur de météo générale pour les Côtes d'Armor est Météo France"
The server provider for general weather forecast for the Côtes d'Armor area is Météo France
"Oui, il est fourni par Météo-France".
Yes, it is provided by Météo-France
"Météo-France".
Météo-France.
etc.

The natural language generator integrated to the Artimis/AGS system is based on generic mechanisms that make it easily portable to other applications but also to other languages. For instance, the system has been easily ported from French into English (with the possibility to switch from one language to another one during the same dialogue) with a very moderate effort (about ½ engineer-month).

A sample dialogue:

A sample dialogue carried by the Artimis / AGS system is given in the following. S stands for the system, U for the user, R for the output of the ASR component. XXX represents one or more words that are semantically irrelevant.

S: **Welcome to the Artimis A.G.S. server. How can help you**
Bienvenue sur le serveur Artimis A.G.S. Que puis-je pour vous ?
U: **I want the weather forecasts. Are there any servers?**
R: Je veux la météo. Y a t-il des serveurs ?
R: XXX XXX meteo y-a-t-il XXX
S: **Are you interested in general weather forecasts or in specialised weather forecasts?**
Êtes-vous intéressé par la météo générale ou par la météo spécialisée ?
U: **Specialised weather forecasts.**
La météo specialisée.
R: XXX meteo specialisee
S: **A server for which specialised weather forecasts?**
Un serveur de quelle météo spécialisée ?
U: **What can you propose?**
Que me proposez vous ?
R: XXX proposez XXX
S: **There are eight kinds of specialised weather forecasts. The two first are atmospheric weather forecasts and coastal weather forecasts. Which specialised weather forecasts are you interested in?**
Il y a huit types de météo spécialisées. Les deux premières sont la météo aérienne et la météo côtière. Vous recherchez un serveur de quelle météo spécialisée ?
U: **Coastal, for the Nord-Pas-de-Calais area.**
Cotière pour le Nord-Pas-de-Calais
R: cotiere XXX nord pas calais.
…

8. CONCLUSIONS

It is widely accepted that, alone, (spoken or written) natural language processing, or mere integration of communication media, will not suffice to obtain a truly "natural" dialogue between humans and computers. Moreover, improving system ergonomy can only be a placeholder for the fundamental requirement, which is system intelligence.

Many prototypes developed so far, especially those based on simple finite state diagrams, are too rigid and do not exhibit that degree of ”intelligence” that makes them appear as cooperative as a normal user would expect. An important question that has to be addressed at this point is ”What should be the basis for system intelligence?”.

A new paradigm, that of *rational dialogue agent,* for designing and implementing convivial dialogue systems has been introduced in this paper. An associated formal framework for representing behavior principles (rationality, communication, cooperation, etc.) and fo reasoning about mental attitudes (belief, uncertainty, intention) and communicative acts (and actions, in general) has been provided.

It has been shown how this paradigm and approach have led to Artimis, a rational dialogue agent technology, that provides a generic framework to instantiate effective advanced dialogue systems.

Artimis technology bridges the gap between bridges the gap, in a "clean" way, between fundamental research and real (end-users) applications. It also opens ne scientific and technological perspectives for the stud and development of really "intelligent" interactive agents.

9. REFERENCES

1 Allen J.F., Ferguson G., Miller B., & Ringer E. Trains as an embodied natural language system. *Proceedings of the AAAI'95 Fall Symposium on Embodied Language and Action*, Cambridge, MA, 1995.

2 Allen J.F. & Perrault C.R. Analyzing intention in utterances. *Artificial Intelligence,* 15(3):143--178, December 1980.

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

4 Bilange E. *Dialogue personne-machine : modélisation et réalisation informatique*, Hermès, 1992.

5 Bretier P. La communication orale coopérative : contribution à la modélisation logique et à la mise en oeuvre d'un agent rationnel dialoguant. *Thèse de Doctorat en Informatique, Université de Paris XIII,* France, 1995.

6 Bretier P., Panaget F., & Sadek M.D. Integrating linguistic capabilities into the formal model of a rational agent: Application to cooperative spoken dialogue. *Proceedings of the AAAI'95 Fall Symposium on Rational Agency* , Cambrige, MA, 1995.

7 Bretier P. & Sadek D. A rational agent as the kernel of a cooperative spoken dialogue system: Implementing a logical theory of interaction. *Intelligent Agent III – Proceedings of the third international workshop on Agent Theories, Architectures, and Languages (ATAL),* Müller J.P., Wooldridge M.J., & Jennings N.R. eds., LNAI, 1996.

8 Bruce B.C. Belief systems and language understanding. *TR 2973, Bolt Baranek and Newman Inc (BBN)*, Cambridge, MA, 1975.

9 Bunt H.C. Information dialogues as communicative action in relation to partner modelling and information processing. *The structure of multimodal dialogue*, Taylor M.M., Néel F., & Bouwhuis D.G., eds., North-Holland, 1989.

10 Cadoret, V. Détermination d'actes de dialogue : une approche combinant représentation explicite des connaissances et apprentissage connexionniste. *Thèse de Doctorat en Informatique, Université de Rennes I,* France, 1995.

11 Carberry S. *Plan recognitioon in natural language*. The MIT Press, 1990.

12 Chu-Carrol J. & Carberry S. A plan-based model for response generation in collaborative task-oriented dialogues. *Proceedings of AAAI'94*, Seattle, WA, 1994.

13 Clarke D.D. The linguistic analogy or when is a speech act like a morpheme? In Ginsborg G.P. ed., *Emerging strategies in social psychological research,* John Wiley, Chichester, 1979.

14 Cohen P.R. & Levesque H.J. Intention is choice with commitment. *Artificial Intelligence,* 42(2-3):213--262, 1990.

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

16 Cohen P.R. & Perrault C.R. Elements of a plan-based theory of speech acts. *Cognitive Science,* 3(3):177--212, 1979.

17 Dahlbäck N. & Jönsson A. An empirically based computationally tractable dialogue model. *Proceedings of the 14th Conference of the Cognitive Science Society (COGSCI'92),* Bloomington, Indiana, 1992.

18 Dennett D.C. *The intentional stance*. The MIT Press, 1987.

19 Doyle J. Rationality and its role in reasoning. *Proceedings of AAAI'90*, Boston, MA, 1990.

20 Ferrieux A. & Sadek M.D. An efficient data-driven model for cooperative spoken dialogue. *Proceedings of ICSLP'94*, Yokohama, Japan, 1994.

21 Fraser N.M. & Dalsgaard P. Spoken dialogue systems: a Europeen perspective. *Proceedings of ISSD'96*, Philadelphia.

22 Ginsberg M.L. & Smith D.E. Reasoning about action I: A possible world approach. *Artificial Intelligence,* Vol. 35, pp. 165-195, 1988.

23 Gitton S. Mise en oeuvre d'un système expérimental de dialogue oral et modèle formel de traitement d'erreurs. *Thèse de Doctorat en Informatique,* France, 1995.

24 Grice H.P. Meaning. *Philosophical Review*, 66:377--388, 1957.

25 Grice H.P. Logic and conversation. In Cole P. & Morgan, J. eds., *Syntax and semantics: Speech acts* , volume3, pages 41-48. Academic Press, New York, 1975.

26 Grosz B.J. & Sidner C. Plans for discourse. *Intentions in communication*. Cohen P.R., Morgan J., & Pollack M.E., eds. MIT Press, 1990.

27 Guyomard M. & Siroux J. Suggestive and corrective answers : A single mechanism. *The structure of*

multimodal dialogue, Taylor M.M., Néel F., & Bouwhuis D.G., eds., North-Holland, 1989.

28 Guyomard M., Nerzic P., & Siroux J. Plans, métaplans et dialogue. *Actes de la 4ème école d'été sur les Traitement des langues naturelles*, Lannion, France, 1993.

29 Jokinen, K. Cooperative Response Planning in CDM. *Dialogue Management in Natural Language Processing Systems. Proceedings of the 11th Twente Workshop on Language Technology*, LuperFoy S., Nijholt A. and Veldhuijzen van Zanten G., Eds., pp. 159-168. Twente.

30 Kaplan S.J. Cooperative responses from a protable natural language query system. *Artificial Intelligence*, 19:165--187, 1982.

31 Kautz H.A. A formal theory of plan recognition and its implementation. *Reasoning about plan*, Allen J.F., Kautz H.K., Pelavin R.N., & Tenenberg J.D., eds., Morgan Kaufmann, 1991.

32 Kayser D. Représenter les connaissances : Pourquoi ? Comment ?. *Actes du séminaire GRECO Dialogue homme-machine à composante orale*, Nancy, 1984.

33 Litman D.J. & Allen J.F. Discourse processing and commonsense plans. In [Cohen et al. 90], pages 365-388.

34 Lochbaum. K. Using Collaborative Plans to Model the Intentional Structure of Discourse. *PhD thesis, Harvard University*, available as Tech. Report TR-25-94, 1994.

35 Magadur J.-Y. Représentation structurelle du dialogue oral homme-machine. *Actes des 19ème Journées d'Etudes sur la Parole*, Bruxelles, Belgique, 1992.

36 McCarthy J. Ascribing mental qualities to machines. *TR Stan-CS-79-725, AIM-32* , Stanford University, CA, 1979.

37 McCarthy J. Circumpscription - A form of non-monotic Reasoning. *Artificial Intelligence*, Vol 13 pp. 27-39, 1980.

38 McCarthy J. & Hayes P. Some philosophical problems from the stand point of AI. *Machine Intelligence*, n° 4, 1969.

39 Meng H., Busayapongchai S., Glass J., Goddeau D., Hetherington L., Hurley E., Pao C., Polifroni J., Seneff S., & Zue V. WEELS: A conversational system in the classifieds domain. *Proceedings of ISSD'96* (and also *Proceedings of ICSLP'96*, Philadelphia, USA, 1996).

40 Moeschler J. *Modélisation du dialogue*, Hermès, 1989.

41 Naito M., Kuroiwa S., Takeda K., Seichi Y., & Yato F. A real-time speech dialogue system for voice activated telephone extension service. *Proceedings of the ESCA/ETR Workshop on Spoken Dialogue Systems: Theories and Application* , Vigso, Denmark, 1995.

42 Nerzic P. Erreurs et échec dans le dialogue oral homme-machine : détection et réparation. *Thèse de Doctorat en Informatique*, France, 1993.

43 Newell A. The knowledge level. *Artificial Intelligence*, 18:87--128, 1982.

44 Panaget F. D'un système générique de génération d'énoncés en contexte de dialogue oral à la formalisation logique des capacités linguistiques d'un agent rationnel dialoguant. *Thèse de Doctorat en Informatique*, France, 1996.

45 Panaget F. A unified representation of lexical and grammatical resources. *Proceedings of the sixth Europeen Workshop on Natural Language Generation*, Duisburg, 1997.

46 Panaget F. Le générateur de langage naturel de l'agent dialoguant Artimis. *Revue Traitement Automatique des Langues, Vol.39, n°2*, 1998.

47 Perrault C.R. An application of default logic to speech act theory. *The structure of multimodal dialogue*, Taylor M.M., Néel F., & Bouwhuis D.G., eds., North-Holland, 1989.

48 Perrault C.R. & Allen J.F. A plan based analysis of indirect speech acts. *American Journal of Computational Linguistics*, 6(3):167-182, 1980.

49 Pieraccini R. & Levin E. A spontaneous speech understanding system for database query applications. *Proceedings of the ESCA/ETR Workshop on Spoken Dialogue Systems: Theories and Applications*, Vigso, Denmark, 1995.

50 Russel S. Rationality and intelligence. *Proceeding of IJCAI'95*, Montréal, 1995.

51 Sadek M.D., Logical task modelling for man-machine dialogue. *Proceedings of AAAI'90*: 970-975, Boston, MA, 1990.

52 Sadek M.D. Attitudes mentales et interaction rationnelle: vers une théorie formelle de la communication. *Thèse de Doctorat en Informatique*, France, 1991.

53 Sadek M.D. Dialogue acts are rational plans *Proceedings of the ESCA/ETRW Workshop on the structure of multimodal dialogue*, pages 1-29, Maratea, Italy, 1991.

54 Sadek M.D. A study in the logic of intention. *Proceedings of the 3rd Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, pages 462-473, Cambridge, MA, 1992.

55 Sadek M.D. Attitudes mentales et fondements du comportement coopératif. *Systèmes coopératifs: de la modélisation à la conception* Pavard B., ed., Octares Editions., pages 93-117, 1994.

56 Sadek M.D. Communication theory = Rationalit principles + Communicative act models. *Proceedings of the AAAI'94 Workshop on Planning for Interagent Communication*, Seattle, WA, 1994.

57 Sadek M.D. Towards a theory of belief reconstruction: application to communication. *Speec Communication Journal*, special issue on Spoken Dialogue (from the *International Symposium on*

Spoken Dialogue, Tokyo, Japon, 1993), 15(3-4):251-263, 1994.

58 Sadek M.D. Le dialogue homme-machine : d l'ergonomie des interfaces à l'agent intelligent dialoguant. *Nouvelles Interfaces Homme-Machine*, Lavoisier Editeur Paris, Série Arago n°18, pp. 277-321, 1996.

59 Sadek M.D., Ferrieux A., & Cozannet A.Towards an artificial agent as the kernel of a spoken dialogue system: A progress report. *Proceedings of the AAAI'94 Workshop on Integration of Natural Language and Speech Processing*, Seattle, WA, 1994.

60 Sadek M.D., Bretier P., Cadoret V., Cozannet A., Dupont P., Ferrieux A., & Panaget F. A cooperative spoken dialogue system based on a rational agent model: A first implementation on the AGS application. *Proceedings of the ESCA/ETR Workshop on Spoken Dialogue Systems : Theories and Applications*, Vigso, Denmark, 1995.

61 Sadek M.D., Ferrieux A., Cozannet A., Bretier P., Panaget F., & Simonin J. Effective human-computer cooperative spoken dialogue: The AGS demonstrator. *Proceedings of ICSLP'96,* Philadelphia, 1996.

62 Sadek M.D., Bretier P., & Panaget F. Artimis: natural dialogue meets rational agency. *Proceeding of IJCAI'97,* Nagoya, Japan, 1997.

63 Sadek D. & De Mori R. Dialogue systems. *Spoken Dialogues with Computers,* De Mori R., ed., Academic Press, 1998.

64 Searle J.R. *Speech Act* , Cambridge University Press, 1969.

65 Searle J.R. *Intentionality: An essay of philosophy of mind*, Cambridge University Press, 1983.

66 Siroux J., Gilloux M., Guyomard M., & Sorin C. Le dialogue homme-machine en langue naturel : un défi ? *Annales des Télécommunications,* Vol. 44(1-2), pp. 53-76, 1989.

67 Traum, D. R. Conversational Agency: The TRAINS-96 Dialogue Manager. *Dialogue Management in Natural Language Processing Systems. Proceedings of the 11th Twente Workshop on Language Technology*, LuperFoy S., Nijholt A. and Veldhuijzen van Zanten G., Eds., pp. 1-11. Twente, 1996.

68 Winograd T. & Flores F. *Understanding computers and cognition: A new foundation for design,* Ablex Publishing Co., Northwood, NJ, 1986.