

## Agent Tasks in Collaborative Intermediary Interaction<sup>1</sup>

Wim Van Loo  
University of Ghent<sup>2</sup>

Harry Bego  
Tilburg University<sup>3</sup>

**Abstract** We draw attention to a difference in use of the notion of *domain task* in mainstream AI and European research on dialogue planning. An integrated approach calls for an adaptation of notions of task dialogue and information seeking dialogue. We make a distinction based on the execution state of domain actions and the roles played by the agents involved. It is argued that the observed differential emphasis on user and system task domain is related to preferences for intentional versus conversational perspectives on utterance sequencing.

### 1 Introduction

In explaining the goals of interlocutors, mainstream computational models of information seeking dialogue center on the actions in the task domain of the information *seeker*. In this research tradition, user domain plans were mainly introduced to facilitate recognition of user goals, while dialogue structure is conceived primarily as a side-effect of information seeking. Conversely, European research tends to focus on the actions performed by the information *provider*. Here, the 'task' of the information provider is sometimes viewed as the main source of structure of information seeking dialogue. This is often implicit in European usage of the phrase *task-oriented dialogue*. An integrated view calls for an adaptation of notions of task dialogue and information seeking dialogue. We make a distinction based

<sup>1</sup>This work was supported by Tilburg University in the context of ESPRIT II project P5254 "A Pragmatics Based Language Understanding System" (PLUS), and by the University of Ghent under Grant Nr. 01111192-OZF/RUG

<sup>2</sup>University of Ghent, Department of French, Blandijnberg 2, B-9000 Gent, Belgium, wvlo@flw.rug.ac.be

<sup>3</sup>Institute for Language Technology and AI (ITK), PO box 90153, 5000 LE Tilburg, The Netherlands, hbego@kub.nl

on the execution state of domain actions and the roles played by the agents involved.

We apply these ideas in a discussion of the role of the task notion in views on dialogue management. It is argued that the observed differential emphasis on user and system task domain is related to preferences for intentional versus structural conversational perspectives on utterance sequencing. Furthermore, in some continental work the communicative activity of information seeking is not clearly distinguished from the agents' domain tasks. We point out that such confusion is engendered by a view of the system as a set of database access functions rather than as an agent with communicative competence distinguishable from its task domain. On the basis of a review of several current approaches to dialogue planning, we list a few consequences for the design of the top-level of a dialogue manager.

### 2 Domain Tasks, Planning, and Dialogue Structure

Two types of dialogue seem to emerge from current views on dialogue planning. Dialogue occurring during task *execution* is one kind, with an agent for instance describing actions taking place, or commanding an other agent involved in the execution what to do next. Dialogue then proceeds in parallel to task execution and topic structure will match task structure. This in fact is what has been called *task-oriented dialogue* (e.g. Grosz, 1978).

Dialogue can however also occur during a task *planning* phase, since an agent will often want to collect information in advance, completing a plan before executing it. For instance the value of some parameter is wanted, or data for competing plans, or an entire action scheme is needed. Dialogue taking place under these circumstances has been called *information seeking dialogue* (e.g. Carberry 1988) or *plan exploration dialogue* (Ramshaw 1991). The communicative goals defined by the information seeking agent concern data related to the task at hand. In preparing a plan for execution the information seeker can choose any incomplete part as input to communicative interaction. To be able to activate the appropriate intentional and attentional state, and to be able to respond cooperatively, an

information provider will want to recognize the domain action sequence involved, using its knowledge of domain plans and plan inference rules, and can then provide the data that allow its completion. But intentional and attentional structure, as well as linguistic structure, are then probably only locally based on the inferred plans. The global structure of the dialogue is determined by *communicative* principles mainly.

This entails a view of information seeking dialogue as dialogue *embedded in a suspended domain task*. The general situation however is an *interleaved* one, in which in the course of execution of a domain plan, exploration of a subplan is necessary. The two types of interaction can therefore alternate, and in tasks consisting of relatively closely interspaced sequences of physical actions we can have a swift succession of what can be called *concurrent* and *non-concurrent* phases of dialogue (cf. Bego 1991).

Up to this point we have assumed that just one plan is involved. A difference in expertise is the essential condition for engaging in communication. The role of the expert is to assist the non-expert in planning or executing a domain task. In the work discussed the task notion refers to actions in the information seeker's domain, as in e.g. Grosz & Sidner, 1986 or Litman & Allen, 1987.

### 3 The Task of the Information Provider

In the French literature however it is quite common to speak of task-oriented dialogue whenever the information *provider's* task is involved. Such a view reveals itself most clearly in the context of collaborative human-machine interaction research where the machine is an agent that offers some service, or is applied as a tool to achieve certain domain goals. Pierrel (1987) states explicitly that dialogue occurring in this context is *oriented on* and even *directed by* the task that is to be executed (*orienté par la tâche*). Examples of tasks are to use automated banking facilities, to control the status of processes, as well as to issue commands, and to retrieve information. In communication systems two knowledge sources record properties of such tasks:

a static *task model* and a dynamic database storing *process state*. Depending on the application, the latter records e.g. the status of the environment (in the case of commanding robots or remedial support systems for disabled persons), the status of processing (in the case of process control), or the status of the database query (in the case of inquiries). The aim of an interpreter is then to detect in the user's utterances the relevant goals in terms of database accesses. *Task management* is the process of seeking information that enables a smooth realization of such tasks. It is argued that the task-related and goal-oriented nature of human-computer dialogue restricts possible language acts, language constructions and use of lexical items.

It is clear that Pierrel, when stating that human-computer dialogue is task-oriented, in the first place refers to activities the *system* is involved in. This usage obviously is at variance with Anglo-Saxon literature. Furthermore, such factors as whether the system plays the role of the expert or the non-expert, or whether we are in a planning phase or in an execution phase, are not considered.

### 4 Agent Roles in Collaborative Interaction

What is called for obviously is a view of both information seeker and provider as communicating agents with domain tasks to plan or execute. The general situation is that more than one task may be involved, and that roles and levels of expertise with respect to the tasks may vary. In the context of two-agent interaction generally two tasks will be relevant, one related to the first agent A1 - let us call this task T1 - and one to the second agent A2 - call it T2. We assume that T1 and T2 are independent and, as a consequence, will exclude the situation in which A1 cooperates with A2 *in a common activity*. This means that we do not have to consider two tasks at the same time, and therefore can setup the situation such that we can refer to T2 as the task at hand. An agent A1 can then have three possible roles depending on his level of expertise and the task T2 related to the other agent A2:

1. T2 is being executed by A2 and A1 *guides the actions* of A2. A1 is an expert with regard to

T2, while A2 has less expertise. Dialogue will proceed concurrent to T2.

2. T2 is being planned by A2 and A1 *assists in the planning* of the actions of A2. Here A1 again is the expert with regard to T2, while A2 has less expertise. But no execution is driving the dialogue.
3. A1 is *not involved* in T2. A1 isn't aware of T2 or of the relevance of its planning and/or execution to the dialogue. The dialogue proceeds non-concurrent to T2. A1's level of expertise with regard to T2 does not matter here, hence his intentional and attentional states are not determined by it.

It is clear that the air compressor assembly example matches the first position. A1 is the expert guiding the actions of novice A2 who is executing task T2. In the context of collaborative human-machine interaction, the machine can have two possible roles, viz. A1 or A2. When commanding a robot in a simulated world, the human is the expert A1. But we can also imagine a flight simulator guiding the actions of the human novice A2. Plan exploration dialogues belong to the second type. A1 is the expert planning a task T2 with A2. What interaction situation matches the third case? Consider the situation in which someone asks an operator of an information service for a telephone number without making clear what his motives are for getting to know that number. Here the operator does not know what the caller's domain goals are, and does not need to know in order to satisfy his request. The operator A1 is not involved in T2 which is being planned by caller A2. It is obvious that the caller's plan T2 will not be executed during the dialogue, so dialogue proceeds non-concurrent to T2.

This however does not yet mean that we have a non-concurrent type of dialogue with regard to T1, the activity of the operator. An appeal could be made to the operator's expertise in finding the requested information, and then dialogue could proceed concurrent to this activity T1 in finding information, e.g. by querying a database. The point now is that the two interactants, the operator and the caller, play the same type of role, namely the one described in the third set of circumstances. Neither

of the interactants is involved in the task related to the other agent. None of them needs to be an expert with regard to the task related to the other agent. They even do not have to be aware of the task the other is involved in.

## 5 Task Dialogue versus Information Dialogue

If we assume that the task T1 we abstracted from above consists in database accessing, it seems justifiable to disregard the user's awareness of this task. Consequently, we can disregard the system's database actions in a distinction of no more than three basic kinds of dialogue:

1. *Task dialogue.* Here the system is the expert with regard to both the user's task and the accessing of a certain database. It guides the user's actions. Examples of this situation are to assemble an air compressor, to place an order, to hire a car, to make a reservation, to fill in a form. Looking up whether goods are available when the caller is placing an order, can serve as an example where the expert's task is present.
2. *Planning dialogue.* Here again the system is the expert with regard to both the user's task and its own task. The system assists in the planning of the actions of the user. Examples of this setting are to provide administrative information or information about travel opportunities and itineraries. Looking up train connections by the expert when assisting the caller to plan a journey is an example where the expert's task is present. Two special cases occur, depending on whether alternative plans are considered or not (cf. Ramshaw 1991).
3. *Parameter dialogue.* Here the system is a domain expert with regard to the accessing of a certain database only and it doesn't need to be aware of the user's task. The system provides pieces of data which the user needs to plan his task, without assisting the user in the planning itself. This is the general type of setting when requests are addressed to a telecom information service.

The settings differ in that the system's task of accessing a database is optional in (1) and (2), while it is not optional in (3), for in that case the user typically addresses the system because it is thought of as able to provide the parameters the user needs. Settings (3) and (2) differ from (1) in that in (1) it is the user's task which is directing dialogue, where in the other two cases it is the system's task that directs dialogue, at least to the extent that the system is executing database queries during dialogue and not just planning an access. The main point here is that the *user's* task is not being executed during the dialogue. In using the terms 'task dialogue' and 'information seeking dialogue' then we consider the user's task only.

We will now take a brief look at a few recent approaches to dialogue management in information seeking and task dialogue, and try to clarify some of the questions encountered. It is important to take into account another dimension of views on dialogue management at this point, for next to differential emphasis on user and system task domain, approaches differ in their pragmatic view of utterance sequencing. *Conversational* approaches depart from syntactic notions such as constituency and combinatory principles like adjacency pairing. *Intentional* approaches give priority to communicative goals. Interestingly, in work on dialogue management that takes the conversational point of view dialogue is sometimes structurally modelled on system task, while in the intentional view, which tends to focus more often on user goals, dialogue structure is a side-effect. We will first discuss work which emphasises the system's task and which seems to take a conversational stance, but does not clearly distinguish communicative activity from domain actions.

## 6 Information Seeking versus Domain Tasks

In Pierrel (1987), a view on the communicative activity of information seeking itself as a task is sometimes at the background of references to human-computer interaction as *purposive dialogue* (le dialogue finalisé) and exemplified by typing the activity of making an enquiry as a task. That the use of the term *task* tends to shift from non-

communicative actions belonging to some often physical domain, to aspects of the communicative interaction itself is especially illustrated in Luzzati (1989). Here the activity the user is involved in is completely mixed up with the actions the expert has to undertake in order to help the user to perform his activity. Moreover, domain actions occurring during the interaction with the system are identified with communicative activity itself. Such an explicit widening of the concept of a task to a structured set of both domain and communicative actions in the consultation of a database is also found in Sabah (1989) and Bilange (1991). The preference for a syntactic conversational approach, or more precisely, for dialogue grammar, as expressed in this work, seems to derive at least in part from a unification of dialogue and system task, possibly inspired by the argument of Amalberti et al. (1984) that experts develop special strategies for directing communication when some task is to be performed, which comes close to a view on the cognitive activity of performing a dialogue as executing a task structure.

Obviously, it is advisable not to put the communicative activity of information seeking on a par with the domain problem that an information seeker is trying to solve and for which he or she consults the information system, nor with the actions the information provider has to execute in order to satisfy the request of the information seeker. The information provider must be viewed as an agent with communicative competence distinguished from his domain tasks. We can then maintain a view of dialogue in which some task at hand is the *topic* of the dialogue, rather than one in which ultimately the task at hand *is* the dialogue. Only from the former perspective it makes sense to state that dialogue structure is related to task structure. From the latter vantage point, dialogue structure *is* task structure by definition, and it even is no longer clear what we are supposed to mean by *task dialogue*, all the more if that is somehow meant as opposed to information seeking dialogue. We will now briefly discuss preferences for dialogue management strategies that seem to be related to intentional and conversational standpoints.

## 7 Agent Tasks and Dialogue Management

Interpreting an experiment held in a hospital, Amalberti et al. (1984) argue that a dialogue in which the expert takes the initiative, is most successful in satisfying the customer's request as it forces him to formulate his intentions explicitly. Consequently, Pierrel (1987), Roussanly (1988) and Bilange (1991) assume that when the system gears the dialogue towards task structure, it will in the most efficient way obtain the data needed for task performance. For the purpose of directing dialogue, Roussanly (1988) proposes an operational dialogue model, called the *model of the dialogue*, which is conceived as a structured set of schemata for dialogue phases, where each phase is associated with actions or procedures to be executed by the dialogue manager. The general pattern of the dialogues is described in terms of the stereotypical strategies of the expert and the stereotypical purposes of the user in engaging in the dialogue. Roussanly essentially thinks of dialogue acts as atomic dialogue phases. In his typology, three kinds of acts are discerned: (a) acts related to the task, viz. fact, question and goal, (b) acts related to the dialogue in progress, viz. opening, closing, request to repeat an utterance, satisfaction, contestation, and (c) acts related to the communication channel, such as acts to keep the channel open and acts related to the perturbation of the channel.

Bilange (1991) not only defines a similar typology, but he also tries to model knowledge relating to the hierarchical and functional organization of these acts according to the model of the Geneva School (e.g. Roulet et. al., 1985). To this aim, an adaptation and implementation of the dialogue grammar developed by Moeschler (1989) is presented. Information on the dialogue structure and the functionality of a dialogue act is included in the definition of the preconditions and effects of communicative acts, thus allowing for aspects of partner modeling, task and communication management. Separate rules of conversation implement the notion of *dialogue strategy*, e.g. by invoking repetitions or reintroducing topics. Syntactic conversational approaches to dialogue management such as these that, as we ob-

served above, tend to lend priority to action in the system's task domain, are to be contrasted to intentional approaches, such as the work on speech act planning by Cohen and Perrault (1979) and Allen (1983), where dialogue structure is largely a side-effect. A recent analysis of dialogue planning that takes an intentional stance is Litman & Allen (1987), but in this work there is a meta-level of *discourse plans* that imparts structure to dialogue<sup>4</sup>.

Building on such stack-based models as Litman & Allen (1987), a dialogue manager which does not engage in non-communicative acts will have a top-level in terms of communicative interactions, e.g. using discourse plans, and will *use domain plans as data*, mainly for communicative cooperativity, while a controller for task dialogue will also *execute domain acts at the top-level*. Pure information seeking dialogue is a solely *communicative* process of execution of communicative acts only, i.e. no information is shared by observation or physical interaction.

In terms of dialogue structure, it is clear then that pure information seeking dialogue surface structure is to a lesser extent determined by the information seeker's task, perhaps partially, as in Carberry's (1990) top-down task planning dialogue. A distinction between dialogue segmentation due to the structure of the domain task, and segmentation according to purely communicative principles is therefore required, e.g. in terms of Allwood's (1990) principles of interaction management. Dialogue segments can be described at and viewed as driven by a purely communicative level of conversational exchanges or discourse plans. Pieces of domain data are input to or output of interaction at this level. In order to initiate and terminate a dialogue segment in which an information goal is specified and possibly satisfied, both partners perform low level communicative acts, e.g. relevant to ensuring contact and floor taking, to the specification of semantic content, or to the function of an utterance (cf. Bunt & Allwood, 1992). Both know how functional types relate, what surface properties and what epistemic aspects are relevant. Feedback acts help control what is mutually believed.

While such an account considers communicative

<sup>4</sup>The relationship between concepts of *conversational exchange* and *discourse plan* as seen from different paradigms seems an especially pressing issue.

principles only, in the general case the information provider's task *is* present. The motivation for the communicative interaction both agents are engaged in is to get information from the other agent in order to plan or execute their own task, but the system's task is executed during the dialogue. The expertise of the system consists in finding and retrieving the information the caller needs to complete his domain plan. To be able to provide the requested information, the expert needs to query a database. It is then the system's goal to get sufficient data from the user in order to complete and form this query. In other words, the system's and the user's task *interface* by means of the dialogue. Such dialogues may in fact be characterized by a subtle interference of conflict and cooperation. One concern is the relevance of parts of the other agent's task when both agents have their own task to be planned or executed.

## References

- Allen, J.F., (1983), Recognizing intentions from natural language utterances, in: Brady & Berwick (eds.): *Computational Models of Discourse*, MIT Press, Cambridge, Mas.: 107-166.
- Allwood, J., J. Nivre & E. Ahlsen (1990) Speech Management - on the non-written life of speech, *Nordic Journal of Linguistics*, 13, 3-48.
- Amalberti, R., N. Carbonell & P. Falzon, (1984), Stratégie de contrôle du dialogue en situation d'interrogation téléphonique, in: *Actes du séminaire GRECO sur le dialogue homme-machine*, Nancy, France: 384-402.
- Bego, H. (1991), *Context Models and Rule Systems for Communicative Acts in Information Dialogue*, MMC Research Report 20, ITK, Tilburg.
- Bilange, E. (1991), A task independent oral dialogue model, in: *Proceedings of the 1991 conference of the European ACL*.
- Bunt, H.C. & Allwood, J. (1992) *Pragmatics in PLUS*, Esprit P5254 Working Paper.
- Carberry, S. (1988), Modeling the User's Plans and Goals, in: *Computational Linguistics* 14(3), 23-37.
- Carberry, S.(1990), *Plan recognition in Natural Language Dialogue*, MIT Press, Cambridge.
- Cohen, P.R. and C.R. Perrault (1979) Elements of a Plan Based Theory of Speech Acts, *Cognitive Science* 3, 177-212.
- Grosz, B.J. (1978), Discourse Analysis, in: D. Walker (ed.), *Understanding Spoken Language*, North Holland, New York, 235-268.
- Grosz, B.J. & C.L. Sidner (1986) Attention, Intention and the Structure of Discourse, *Computational Linguistics* 12(3), 175-204.
- Litman, D.J. & Allen, J.F (1987) A Plan Recognition Model for Subdialogues, in: *Cognitive Science* 11, 163-200.
- Luzzati, D. (1989), *Recherches sur le dialogue homme-machine: modèles linguistiques et traitements automatiques*, Ph.D. Thesis Université de la Sorbonne Nouvelle Paris 3, Paris, France.
- Moeschler, J. (1989), *Modélisation du dialogue*, Hermès, Paris, France.
- Pierrel, J.M. (1987), *Dialogue oral homme-machine*, Hermès, Paris, France.
- Ramshaw, L. (1991), A three level model for plan exploration, *Proceedings of the 29th ACL*, Univ. of Berkeley, Cal.
- Roulet, E. et al. (1985), *L'articulation du discours en Français contemporain*, Lang, Berne, Suisse.
- Roussanaly, A. (1988), *DIAL: la composante dialogue d'un système de communication orale homme-machine finalisée en langage naturel*, Ph.D. Thesis Université de Nancy 1, Nancy, France.
- Sabah, G. (1989), *L'intelligence artificielle et le langage, Vol. 2: Processus de compréhension*, Hermès, Paris, France.