Pauchet, Chaignaud, El Fallah Seghrouchni, *BDIggy: a Cognitive Architecture for Cooperative Problem Solving*, IEEE IAT, Halifax, Canada, pp. 523-526, 2003.

BDIGGY: a Cognitive Architecture for Cooperative Problem Solving

A. Pauchet - LIPN
Avenue J.B. Clment
93430 Villetaneuse
pauchet@insa-rouen.fr

N. Chaignaud - PSI
BP 08 - Place mile Blondel
76130 Mont-Saint-Aignan
chaignaud@insa-rouen.fr

A. El Fallah-Seghrouchni - LIP6
8 rue du Capitaine Scott
75015 Paris
Amal.Elfallah@lip6.fr

Abstract

This article deals with a human model of cooperative problem solving based on a psychological experiment. Our main goal is to improve the design of open multi-agent systems able to interact with humans. We aim at analyzing, modelling and simulating human capabilities of planning and interacting in a cooperative problem solving context. This paper goes on to present an original architecture called BDIGGY *where interaction and planning are linked homogeneously.*

Keywords: *Cognitive modelling, cooperative problem solving, human planning, human interaction, BDI agents.*

1. Introduction

In this paper we argue that the study of human reasoning and interaction is useful to build systems able to well interact with human beings. In fact we share the idea that computers should be adapted to human communication and reasoning processes. Our goal is twofold: build a model of human planning with coordination between agents and design an interaction language that allows artificial and human agents to communicate.

The study presented here is narrowed to a particular application: human subjects have to solve a planning problem where information is incomplete and thus they have to interact with the others to obtain the missing information.

We have set up a psychological experiment to obtain experimental protocols which have been analyzed focusing on human planning [4] and human interactions.

This article describes our human interaction model and shows how planning and interaction can be integrated in an homogeneous agent architecture called BDIGGY. It is organized as follows: Section 2 describes the application we chose and the psychological experiment we have set up. Section 3 details our model of human interaction. Section 4 sketches the BDIGGY architecture including planning and interaction. Section 5 discusses related work.

2. The Experimental Framework

2.1. The Problem to be Solved

Our psychological experiment is related to a travel-agency application. Three salesmen work in a travel agency and have different skills: the first one is specialized in planes, the second in railways and the last one in taxis and coaches. Each of them has to organize a journey for his client. These journeys are characterized by a departure and an arrival point in France, a date and a time of departure, a date and a time of arrival, a number of travellers to be booked and a budget. None of the journeys can be arranged using a single means of transport. Each agent has, therefore, an individual problem to solve but they all participate in solving the other two problems. To communicate, the subjects use e-mails written in natural language.

2.2. The experiment

The subjects are isolated in different rooms and the three problems have to be solved simultaneously. They are provided with a user-interface we designed to be convenient.

All the subjects' actions are recorded by the simulation program and written in text files. The subjects were asked to solve their problem by thinking aloud and these simultaneous verbal utterances called *verbalizations* are recorded by an experimenter. The text files and the verbalizations are called *experimental protocols*. Our cognitive model is based on the analysis of these protocols.

This experiment was carried out with 12 groups of 3 students. These 36 protocols were then split into 2 classes: 8 groups of 3 were analyzed to build the model and 4 groups of 3 for the validation of the cognitive model.

3. A Model of Human Interaction

The experimental protocols have been analyzed from the interaction point of view. Our focus is the sending and receiving of messages.

3.1. The Utterance Level

The first results obtained were presented in [5], where a list of primitive performatives was proposed. These performatives were selected among either KQML [8] or FIPA-ACL [7] performatives before being adapted to remain faithful to the experimental protocols.

If we refer to Searle's classification [11] from the Speech Acts theory, the performatives observed come from the three following classes: the *descriptives* (also called *assertives* or *representatives*), the *directives* and the *commissives*. There are no *declaratives*, which is not surprising according to the studied problem. Moreover, we consider that the *expressives* are not necessary, because an expressive is no more than a particular descriptive: the description of one agent's feelings is the description of one part of the world.

All the performatives are issued from the protocols:

Descriptives	
inform	A gives a piece of information to B
reply	A answers B
error	A does not understand a message
Directives	
query	A asks B for a piece of information
reply-proposal	A accepts or refuses an information proposal
Commissives	
propose-information	A proposes to give information to B
reply-later	A warns B that it will answer later

This list is exhaustive in sofar as we are interested only in information-search dialogs.

In our model, a message content is represented by a mental state applied to a predicate and consequently an utterance is represented by a performative applied to a mental state the scope of which is a predicate:

- A *descriptive* is applied to a *believe*: it describes the way the speaker perceives the world. It corresponds to its *beliefs* which fits the world if the descriptive is satisfied. For example, if agent A wants to inform agent B that there is a train at 9 from Angers to Paris, it can send the following message: inform(A, B, Bel(A, step(train, Angers, Paris, 9h))).

- A *directive* is applied to a *desire*: it is used when the speaker agent wants the hearer agent to do something. It has the *desire* that the world turns into a particular state and he transmits this desire as a abstract plan. For instance, if agent A wants agent B to send it timetables about trains from Angers to Paris, it can write the following message: query(A, B, Des(A, step(train, Angers, Paris, ?X))).

- A *commissive* is applied to an *intention*: A speaker agent uses a commissive to tell the hearer agent that it intends to carry out an action. For example, if agent A proposes to agent B to send it train timetables for Angers-Paris trains, it can send the following message: propose-information(A, B, Int(A, step(train, Angers, Paris, ?X))).

These propositions are not sufficient. A complete semantics of the selected performatives in terms of beliefs, desires and intentions has to be provided.

3.2. The Discourse Level

Interactions could not be considered as a simple series of queries and answers following a rigid scheme. Indeed, each locution conducts to a huge variety of behavior from the interlocutor. In this section, we are interested in modelling this dynamics of conversations.

Our discourse analysis is based on Vanderveken's works [12] which extends the Speech Act theory to discourse.

We divided the agent dialogs into interventions. Each of these interventions is guided by the discourse intention of the initiator subject, according to the first performative he sent. We classified the interventions into the three following categories: *information query*, *information proposal* and *spontaneous sending of information*. The discourse intention is directive for the first one, commissive for the second one and descriptive for the last one.

The way interventions are terminated defines their satisfaction. An intervention can be either *satisfied* (the interlocutor's goal happens) or not.

The starting performative enables us to classify the intervention and the ending performative, if it exists, defines the intervention satisfaction. Nevertheless, an intervention can be considered as terminated by the interlocutors even without explicit emission of an ending performative. In this case, we consider that the intervention has failed and so cannot be satisfied. Spontaneous sending of information needs only one main act: the starting performative.

From our point of view, when the interlocutors do not meet any interpretation problem for an utterance (a non-expected event), they conduct the current intervention as a governing dialog. Otherwise, they initiate an incidental dialog to solve the problem, looking for a common interpretation. When this common interpretation appears, they reactivate the suspended governing dialog.

An example of a sequence of interventions is given fig. 1: intervention 1 and 4 are governing interventions whereas intervention 2 and 3 are incidental ones. Intervention 1 can be classified as an information query because of its starting performative. Its ending performative indicates that the intervention is satisfied. Intervention 2 and 4 are information-proposals and intervention 3 is an information-query.

In the particular case of our application, where messages are e-mails, time is essential to take into account re-queries and to terminate an intervention. For example, when an agent needs a piece of information from another agent, if it still has not received any answer after a certain time, it

Messages	Performatives	
[08:23:48] From: Air-agent To: Railway-agent I have just learnt that it is possible to travel from Paris to Montpellier by train! Could you please give me some timetables?	query	
[08:28:45] From: Railway-agent To: Air-agent Yes, it's possible, there are 7 different departures from 8:12 until 18:28 (departure time). Are you interested in them? Could you be more precise on the time departure? Thanks.	reply + propose-information + query	2 3
[08:29:52] From: Air-agent To: Railway-agent Yes, from 10:15.	reply	1
[08:33:18] From: Railway-agent To: Air-agent Here is the first =======Time table======== Paris-Montpellier (train) Date: Tomorrow -- 1 person 10:30/14:39 - 590 F	reply	
[08:33:49] From: Railway-agent To: Air-agent Now number 2, do you want any others? =======Time table======== Paris-Montpellier (train) Date: Tomorrow -- 1 person 12:06/16:21 - 590 F	reply + propose-information	4

Figure 1. An intervention sequence

will re-ask for its information. Similarly, intervention without any explicit ending speech act should be considered as terminated after a certain time. We model these exchanges of messages and the temporality as well, thanks to timed automata [2] which helps the agent

- to generate a message: it is produced following an automaton.
- to interpret a message: an automaton describes the expected messages. Two interlocutors can manage many interventions simultaneously so this expected event helps to know if a message belongs to an intervention.

We have constructed a pair of automata (an automaton for each interlocutor) for each type of intervention. In this article, we present only the automaton A of the initiator S of an information query (see fig. 2). The bold circles represent the terminal states.

To manage with time, A contains a clock t and a deadline $tsync$ before S considers the intervention is terminated. A also contains a counter m to count how many times S has to re-ask for a piece of information before perhaps receiving an answer from the hearer agent (called H).

In A, S sends an information query (state e1). If S receives an answer in one or more messages (state e2), if this answer is satisfactory, it can either do nothing or thanks H (state e3). The current intervention is therefore satisfied, even without any explicit end message. If S waits during $tsync$ without any answer from H (state e5), it increments m and re-asks H for information. If it happens too frequently (state e6), it can leave the intervention telling H (state e1) or not. If H asks S for more information about its

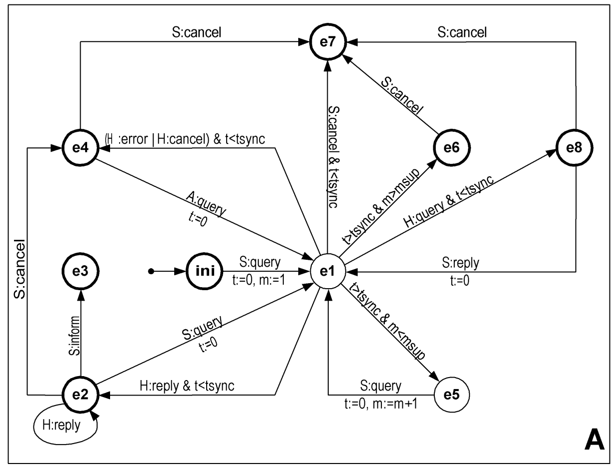

Figure 2. The information-query automaton

query (state e8), an incidental dialog of information query is created, which is similar to the current one but H becomes S and *vice-versa*. The incidental dialog can be satisfied and the governing intervention is re-activated (state e1), or it can failed and S cancels its query explicitly (state e7) or not.

Thus, timed automata are a powerful formalism to take into account the interleaving of the dialogs (state e8 in A_1) and time management in the conversation representation.

4. The BDIGGY Architecture

Our goal is to design a system which simulates the experimental protocols we obtained. In [4], we proposed a cognitive model for planning based on the notions of observation, phase, state of mind, strategy and tactic. This model is implemented by the IGGY system. The agent architecture called BDIGGY we propose to compute our planning model and our interaction models is based both on our IGGY system and on a BDI architecture extended in a multi-agents framework. It includes (see figure 3):

- a perception module as interface between the IGGY system and the environment that generates a set of observations,
- an IGGY model which generates an episode described by some observations, a phase, a state of mind, some strategies and some tactics,
- a desire generator which interprets an episode as desires (goals as abstract plans); this module embodies a representation of the current plans,
- an intention generator which refines a desire into intentions,
- an execution module which carries out necessary actions of an intention,

- and a communication module allowing the agent to interact.

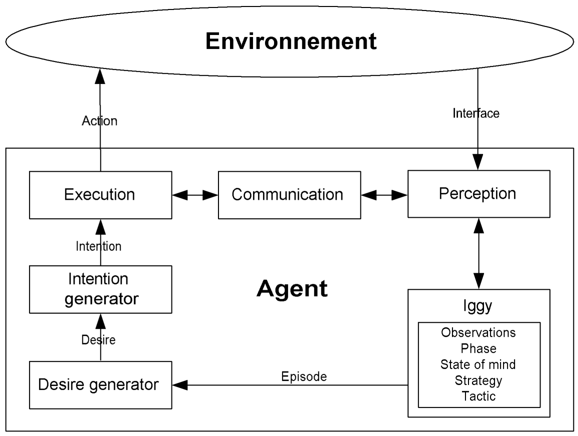

Figure 3. Architecture BDIGGY

In comparison with a classical BDI system such as dMARS [6], our architecture includes a communication module to cooperate with other agents and a planning module that generates plans dynamically according to the environment changes. Moreover, communication and planning are represented in a homogeneous way (*i.e* BDI) in the same system.

5. Related Work

Our study deals with three main issues: cognitive modelling, human cooperative planning and human interaction. To our knowledge, there is very few work integrating all these three aspects.

Allen, Blaylock and Ferguson propose in [1] a model for collaborative agents which integrates an individual problem-solving model, a collaborative problem-solving model and an interaction model in natural language. Even if it does not include explicitly a BDI model, the notions used (*situations*, *objectives* and *recipes*) are very similar.

The Bouzouba and Moulin's point of view adopted in [3] is similar to ours but they are only concerned with a communication model. They propose to extend KQML to KQML+ in order to suit better to the Speech Act theory.

From the computational point of view and as far as communications are concerned, KQML [8] and FIPA-ACL [7] propose some communication protocols which cannot take into account the dynamics of human communications. The Dooley Graphs presented by Van Dyke Parunak in [9] also contain information about the situation and the protagonists of a conversation. But they do not take into account the time dimension for the conversational dynamics.

6. Conclusion and Perspectives

The models we have proposed in this article are based on the analysis of the protocols issued from a psychological experiment. They describe human planning and human interaction as faithfully to the protocols as possible.

These models represented homogeneously are integrated in a same architecture called BDIGGY.

Moreover, it improves classical BDI architecture by including a communication module to generate and interpret the messages and is able to construct dynamically plans.

Work in progress aims at providing a complete semantics in terms of beliefs, desires and intentions for the performatives we have selected (see Sadek [10]), at implementing our system and at validating it.

References

[1] J. Allen, N. Blaylock, and G. Ferguson. A problem solving model for collaborative agents. In *AAMAS'02*, volume 2, pages 774–781, 2002.

[2] R. Alur and D. L. Dill. A theory of timed automata. *Theoretical computer science*, 126:183–235, 1994.

[3] K. Bouzouba and B. Moulin. Les interactions distribuées : Approche basée sur le point de vue personnel des agents. In *MFI'01*, volume 1, pages 15–30, 2001.

[4] N. Chaignaud, I. Cheikhrouhou, and A. Nguyen-Xuan. Generalisation of a cognitive model for problem solving with incomplete information. In *ICCM'00*, pages 58–69, 2000.

[5] N. Chaignaud and A. E. Fallah-Seghrouchni. Apport de la modélisation cognitive aux langages de communication dans les sma. In Hermès, editor, *JFIADSMA'01*, 2001.

[6] M. d'Inverno, D. Kinny, M. Luck, and M. Wooldridge. A formal specification of dmars. Tn 72, Australian Artificial Intelligence Institute, 1997.

[7] FIPA. Specification: Acl. Technical report, Foundation for Intelligent Physical Agents, 1999.

[8] Y. Labrou and T. Finin. A proposal for a new kqml specification. Tr cs-97-03, UMBC, 1997.

[9] H. V. D. Parunak. Visualizing agent conversations: using enhanced dooley graphs for agent design and analysis. In *ICMAS'96*, pages 275–282, 1996.

[10] M. Sadek. A study in the logic of intention. In *KR'92*, pages 462–473, 1992.

[11] J. Searle. *Les actes de langage*. Hermann, 1972.

[12] D. Vanderveken. *Analyse et simulation de conversations. De la théorie des discours aux systèmes multiagents*, chapter 2 : La structure logique des dialogues intelligents, pages 61–100. L'interdisciplinaire, 1999.