Agent-Based Knowledge Acquisition

Rose DIENG, Olivier CORBY, Sofiane LABIDI

ACACIA Project, INRIA
2004 Route des Lucioles, BP 93
06902 SOPHIA-ANTIPOLIS CEDEX, FRANCE
E-mail: {dieng, corby, labidi}@sophia.inria.fr
Tel: (33) 93 65 78 10, Fax: (33) 93 65 77 83

Abstract. This paper presents our approach for knowledge acquisition from multiple experts. In order to build a cooperative KBS, representing the knowledge of several experts and intended to multiple users inside an organization, we propose a model of cognitive agent for guiding the process of knowledge acquisition. This model of agent can serve as a basis for specifying the future KBS to be integrated in the organization. An agent-based knowledge acquisition is then seen as the process of identifying the adequate agents and of filling them (both their individual characteristics such as their expertise model or their knowledge graphs, and their social features such as their integration in an organization or their cooperation capabilities).

1 Introduction

The involvement of multiple experts in the development of a knowledge-based system (KBS) can influence problem definition, knowledge elicitation, knowledge modeling, operationalization of the expertise models and validation of the final KBS [28, 21, 27].

A cooperative, explanatory KBS must be able to cooperate with its ends-users for problem solving and to provide them with explanations about its knowledge and its reasoning. The building of such a system involves collaborative knowledge acquisition and collaborative design, as well as the construction of a KBS based on multiple experts.

In both cases (i.e multiple experts and cooperative KBS), models or techniques stemming from the field of distributed artificial intelligence (DAI) seem useful: (a) for modelling the organization where the experts and the intended users of the future KBS work, (b) for modelling the experts, their cooperation and their conflicts, (c) for modelling the knowledge acquisition process and specially the cooperation during knowledge acquisition from a group of experts, (d) for modelling the cooperation and explanation processes between the humans (experts or non-experts) in the effective organization and between the future KBS and the future end-users.

We think that a model of cognitive agent can guide the design of such a KBS, specially when this KBS is aimed at relying on some human, social behaviour and at offering a collaboration between an artificial agent and human agents.

Therefore, our knowledge acquisition method relies on a model of cognitive agent aimed at modelling the experts and the intended users in their organization.

In this paper, after studying the link with composite systems, we present our model of cognitive agent. Then, we describe the method of agent-based knowledge acquisition, relying on this model, and we present an application of our method to traffic accident analysis.

2 Composite systems

The notion of composite systems [17], made of heterogeneous (human, software and hardware) agents seems quite relevant for our work. In a previous research [15], we viewed the development of a multi-expert system as a combination of several human and software components, interacting with one another, depending on the phase of development:

- The analysis of the actual organization made of experts (whose tasks will be partially simulated by the future KBS) and of potential users that need assistance from this future KBS helps to identify the human agents that may then take part in the knowledge acquisition phase.
- The knowledge acquisition phase consists of knowledge elicitation and modelling and involves human agents such as experts, knowledge engineers, potential users. The knowledge acquisition tool - that plays the role of an assistant of the knowledge engineer - is an artificial agent, that may itself rely on a multi-agent architecture.
- During the design phase, progressively, the knowledge engineer designs the artificial agents of the final multi-expert system, with the help of human agents such as the experts, the potential users and perhaps specialists in ergonomics or in user interfaces.
- After effective achievement of the KBS, during validation phase, the knowledge engineer and the validating experts are human agents, while the KBS is an artificial agent.
- In the final phase of effective use of the KBS, the introduction of the KBS into the organization transforms this organization. The artificial agent is the KBS that may be composed of a multi-agent system, comprise an explainer and cooperate with the human agents constituted by the end-users.

This vision allows to model the development of a multi-expert application, (in particular the knowledge acquisition phase) as the behavior of a specific society of (human or artificial) agents: it helps to characterize the roles and tasks of each agent and to emphasize the main relations between such agents (knowledge transfer, explanation, validation, assistance to problem solving...) and to analyse the cooperation underlying the process of knowledge acquisition. We also stressed that the notion of agent should allow to model the end-user as an agent and ease the description of a knowledge acquisition methodology involving several human agents. In [14], we had proposed elements of a model of cognitive agent in order to take into account knowledge acquisition from multiple experts.

This analysis confirms the remark made in [32] where the author notices that the process of design of an object multi-agent system relies on several multi-agent systems obtained by composition of the object multi-agent system and of some of the various human agents that may take part in the design (user, application programmer, system programmer, designer...)

Notice that, in our case, two object multi-agent systems can be considered: the knowledge acquisition tool and the KBS can both be based on a multi-agent architecture. For the knowledge acquisition tool, we need a model of the human agents inside the human organization and of the knowledge acquisition process. For the KBS, the model of agent must allow to specify the intended behaviour of the group constituted by the KBS and the users, and how several such agents will cooperate for problem solving and for explanation. In our past research [12], we had proposed a language for describing different relations concerning one agent or linking several agents (e.g. specialists inside one KBS, or several different cooperating KBS). Inspiration can be taken from this idea in order to propose a language for specifying not only the final multi-agent system but the relations of the human / artificial agents in the organization.

The couple "final KBS - user" can be considered as a couple of two agents that must cooperate in order to perform for example cooperative problem solving. It is then interesting to model the group constituted by the KBS and by the user as a multi-agent system compound of two agents: decomposition of the global task between both agents, distribution of subtasks, planning of job among them, possible interactions among them, their interaction points, possible communication language they use, possible conflicts and way such conflicts will be solved [5]. This vision can be extended to several users: in this case, a multi-agent system is obtained, with an artificial agent (the cooperative system) and several human agents (the users), such human agents can interact among themselves or with the KBS. The interaction among the users or with the KBS will be influenced by the organization to which the users belong.

3 A model of Cognitive Agent

A model of agent, intended to model both experts and users involved in the KBS design, must include *individual aspects* (concerning the agent himself independently of the organization in which he is inserted and independently of the other agents) and *social aspects* related to the agent's insertion in an organization and to his interactions with the other agents. The individual aspects include *general features* not linked to the particular problem to be solved (competence domain, high-level goals, tasks, expertise knowledge consisting of domain knowledge, of general problem solving methods used, of possible strategies) and *problem-specific features* that may depend on the phase of the considered problem solving (intentions, plans, actions, commitments and state).

The social aspects include the organizational structure (that can concern either the way a compound agent is organized or the way an agent is included in a given compound agent), the cooperation modes, the communication languages.

We distinguish *simple agents* that are not made of other agents and *compound agents* (also called *organizations*) that are constituted by subagents gathered through an organizational structure, such subagents being themselves simple or compound. This notion of compound agent allows to model for example a group of cooperating experts or a group of collaborating users. The individual (resp. social) aspects of an agent exist, whether this agent is simple or compound.

The next sections describe more thoroughly such features.

3.1 Individual Characteristics

Long-term, Problem-independent Characteristics. Some individual characteristics of an agent are generally long-term and problem-independent. Among such characteristics, let us cite:

- *identity*: an agent is identified by a name characterizing this agent without any ambiguity and known by the other agents that can call him explicitly by his name. Each time a new agent is created, he must at least own a name, different from the already existing agent names. An agent may have several names, but he must not share any name with any other agent. The problems raised by the notion of identity of an aggregate agent (specially when this identity cannot be considered as permanent and can evolve through time) were studied in [22].
- *role, competence domain, high-level goals* and *tasks* : such characteristics summarize what the agent can do. The role of an agent in an organization he belong to must be indicated. If an agent belongs to several organizations, he can have a different role in each of them. The competence domain indicates the special field of expertise of the agent. When he solves a problem in his discipline, the expert aims at one or a few high-level goals, associated to his main task. This main task can then be decomposed into subtasks: this decomposition can be represented through a task structure [6, 8, 33, 30]. Such a task structure may be static or dynamic: in this last case, the task structure should rather be considered as a problem-dependent characteristics instead of a problem-independent feature. Several levels of abstraction can be adopted for this task structure: either it may be close to the expertise domain [16] or generic [6, 8, 33].

  For example, psychology, vehicle engineering and road infrastructure engineering are examples of competence domains of the experts involved in the task of road accident analysis. When analysing an accident, in addition to the comprehension of the accident scenario, cognitive modelling of the drivers and diagnosis of the drivers' errors are the high level goals of the psychologist, while diagnosis of a possible vehicle malfunctioning and proposal of advices for vehicle design and manufacturing are the high-level goals of the car engineer. Extraction of data from the drivers involved in the accident, as well as analysis of all extracted data (on the drivers, on the vehicles and on the road infrastructure) from a psychology viewpoint are examples of (sub)tasks of the psychologist. The task of data analysis can itself be decomposed into other subtasks.

- *resources*: such characteristics summarize the various resources used by the agent. In the framework of artificial agents, in [7], the authors distinguish sensing resources, sending-receiving resources, acting resources and cognitive resources. For human agents such as experts or users, the relevant resources are rather instruments, software, tools, etc. For example, in traffic accident analysis, the infrastructure engineer and the vehicle engineer may make use of a camera in order to take photographs of the road infrastructure and of the vehicles involved in the accident, while the psychologist uses a cassette recorder in order to record the interviews of the involved drivers, and the vehicle engineer uses a program of kinematics reconstitution, in order to reconstitute the trajectories and speeds of the vehicles.
  In addition to classic interviews, analysis of observations or of thinking-aloud protocols can give complementary information on such resources (and specially on the context where the expert makes use of a given resource).

  Remark: Knowledge can be considered as a cognitive resource, so significant that we prefer to distinguish a specific characteristics, called *expertise model*, in the model of agent.
- *expertise model*: it helps to model the agent's static knowledge on the domain, his problem solving methods, and his strategies. We adopt the decomposition of expertise knowledge in four layers (domain, inference, task and strategy) proposed in KADS-I [6, 33] [1]. An agent's vision of the domain comprises the concepts or entities he knows, his vision of their structure and of the relations linking them, the parts of them he can access or act upon. If a KADS-I expertise model is associated to the agent, his vision of the domain can be described in the domain layer of this expertise model. Moreover, his strategy can be based on an *evaluation function* allowing him to assess the situations, in order to guide his choices according to the context. The *general preferences* of the agent may also be part of this strategy. For example, the order in which an expert prefer to examine some data, the kinds of models / methods / tools he prefers when several ones are available, are parts of his general preferences.

Remark: We consider that competence domain, high-level goals, resources, knowledge are generally independent of the problem to be solved. It seems to be a reasonable hypothesis. But, in some cases, such features may evolve : for example, a particular problem solving may help an agent (for example, the user of the final KBS) to increase his knowledge and to improve his capabilities. Likewise, the resources used for a given problem solving may depend on the considered problem.

[1] We tried to respect the philosophy of KADS-I and have not yet adapted our work to COMMONKADS framework.

When a feature is contextual and not permanent, it is important to acquire all the adequate contextual information indicating the conditions of modification of this feature.

Problem-dependent Characteristics. Other characteristics depend on the problem to be solved and can evolve throughout the problem solving.

- *intentions* or *low-level goals*: during a problem solving, the agent has several simultaneous or successive intentions, that can be explicitly expressed to other agents or remain implicit. A hierarchy may link the different intentions of an agent, as he can handle simultaneously several low-level goals of various levels of abstraction. Of course, such low-level goals can generally be linked (and associated) to the subtasks appearing in the decomposition of the main task of the expert. In [7], the authors distinguish strategic intentions (that are long-term) and tactical intentions (that are short or mid-term).
  So, the knowledge engineer must elicit the possible intentions of the different experts during a problem solving, how such intentions are linked to one another, etc. For example, the analysis of individual thinking-aloud protocols or of case studies can be useful to reveal such intentions, and help the expert to make them explicit and to explain them.
- *plans*: in order to achieve his goals, the agent may make individual plans, perhaps taking into account the common goal(s) of the organization(s) he is included in. A plan is a succession of actions. Of course, a plan should be naturally associated to a task: it can be seen as a detailed description of a possible realization of this task.
  So, the knowledge engineer must elicit the different possible plans each expert has available for a given problem solving, the elementary actions such plans are composed of, the criteria of choice between such plans, the conditions of execution of a plan, its conditions of success, how to repair it in case of need. The way the individual plans of the experts can be integrated in collective plans of the organizations to which the experts belong can also be elicited, through interviews or analysis of observations, case studies or collective thinking-aloud protocols.
- *actions*: we distinguish observable actions, actions having observable consequences, and internal actions without any observable consequence (such as the cognitive actions that are distinct from the effectoric actions, as stressed in [7]). The actions can have different levels of abstraction and different natures (e.g. physical actions versus steps of problem-solving).
  For example, the observable actions of the vehicle engineer are to record the tracks of the vehicles involved in the accident, to examine the photographs of the vehicles and of the road infrastructure, to execute the program of kinematics reconstitution, while his cognitive actions correspond to his internal reasoning and can be simultaneous with the previously described observable actions.
- *state*: several features of the agent can evolve through time, due to the evolution of the problem solving or to the interactions of the agent with the other

agents: such features can be considered as parts of the agent's state. The agent has an *information state* (what he knows about data, about the state of the problem solving), an *intentional state* (at a given instant, both his long-term and short-term intentions), a *strategic state* (the chosen strategy among several possible alternatives), an *evaluative state* (at a given instant, the values assigned to the possible choices) [31].
If we consider that the agent's knowledge (i.e. his model of expertise) can evolve, it may also be considered as a part of the agent's state. The model the agent has about the world and in particular, about the other agents and their intentions, may also be considered as a part of the agent's information state.
The state of a compound agent depends on the individual states of his subagents, but not exclusively.
- *individual history*: at a given moment of the problem solving, it is constituted by the successive past actions and past states of the agent. Such information on the individual history of an agent during a problem solving could be useful, for example for enabling explanatory capabilities in the final KBS. It can be elicited by analysing case studies, observations or thinking-aloud protocols.

3.2 Social Aspects

When an agent belongs to a multi-agent world, his social characteristics may be described through the features described below.

- *interaction points* [32]: they can be used as an external interface of the agent with the outside world. The requests sent by the other agents to this agent rely on what they know about the official competences of this agent, and the kind of requests he can accept. We distinguish *entry interaction points* (cf the demands the agent can receive) and *exit interaction points* (cf the requests he can send). Requests are linked to particular subtasks the agent is known to be able to perform. For example, an entry interaction point of the vehicle engineer agent is *"extract data on the state of the vehicles involved in the road accident"* since other agents such as the psychologist can need some data on the vehicle, for their own tasks. An agent must be able to link a request received in one of his interaction points, to a task he is able to perform.
In the case of a compound agent, there may be several kinds of relations between his interaction points and the interaction points of his subagents: his interaction points may correspond exactly to the set of all the interaction points of the subagents, or to only a subset or even comprise new interaction points not appearing in any of the subagents.
So, during knowledge acquisition phase, the knowledge engineer must elicit information on what can constitute the interaction points of each expert.
- *model of the world*: the external world of an agent consists of other intelligent agents, and of entities such as facts or data. His interactions with the outside

world, and in particular with the other agents help the agent to adapt and modify his model of the world. A *model of the other agents* is part of this model of the world: we can consider that some features of an agent such as domain competence, high-level goals, interaction points can be visible from the other agents. But must they be visible from all agents or not ? Some flexibility on this notion of visibility is needed. For example, the agent psychologist may know the competence domain of the vehicle engineer agent without knowing his intentions or the individual plans he uses for a particular problem solving. An agent may know only the external aspects of another compound agent (competence domain, high-level goals, interaction points) but without knowing his internal structure and the subagents composing this compound agent.

So, the knowledge engineer must elicit what vision each expert has upon the external world and upon the other experts.

– *organizational structure*: a *compound agent* (i.e. an *organization*) is made of several subagents (that may be simple or compound themselves), gathered into an adequate organizational structure and, so, linked by different relations such as cooperative problem solving... This notion of compound agent is close to the notions of *social agent* proposed in [26] or of *aggregate agent* studied in [22]. Some agents can aggregate in order to form a compound agent (seen as a unique agent by the outside world), for a common goal. Such an aggregation is not a simple "concatenation". It must respect an organizational structure. There may be several models of organizational structures, according to which the different agents may aggregate. For example, such agents can be gathered in a horizontal, non hierarchical structure or in a vertical, hierarchical structure. The organizational structure may then influence their types of cooperation: for example, sharing of tasks and of results in a non hierarchical organization, commands, bids and competition in a hierarchical organization... Studies on the roles and responsibilities assigned to distinct agents inside a human organization, on the norms which govern interactions between such agents, on the delegation of tasks to agents can be exploited.

  Two different compound agents can be made of exactly the same subagents but gathered through two different organizational structures. For example, there may be a compound agent made of a psychologist, a vehicle engineer and an infrastructure engineer working with a non hierarchical structure, and another compound agent, made of the three same subagents organized hierarchically with the infrastructure engineer coordinating the group work. Both compound agents will exist at different moments of the problem solving; they may also exist simultaneously, with different high-level goals.

  The interaction points of a compound agent are not necessarily obtained by the union of the interaction points of the internal agents. When an external agent asks a request to the whole organization, he uses the interaction points of the compound agent. The way this request is then forwarded to all or some subagents of the organization depends on the organizational structure of the compound agent.

The recognized competence of the whole organization may be wider (or even smaller) than the simple union of the individual competences of the agents. The list of the tasks the compound agent can perform can also be wider than the union of the subagent task lists.

An organization can be fixed or evolve through time as new agents can be inserted into it or old agents removed. A given agent may belong simultaneously or successively to several different organizations. Even if he belongs to an organization, an agent can receive individual requests (concerning the agent himself and not the whole organization), on his own individual interaction points.

We will suppose that the high-level goals of the compound agent are common to all his subagents. They may have been common before the aggregation of the agents into the compound agent, or, on the contrary, have become common only because of this aggregation. Likewise, they may disappear after the dissociation of the compound agent or, on the contrary, still exist for the separate agents even after their separation. The distribution of tasks among the different agents can be independent of the problem, or fixed for a given problem or dynamically evolving throughout problem solving. Some aspects of the organization can be global and visible to all internal and external agents, while others can be local to the agent or partially known by only some particular agents. Several studies on organizational models and on science of organizations can be exploited [4, 25, 34]. In [4], the author describes a model for human organizations represented by a set of interacting projects. The agents performing such activities correspond to specialist departments. The model helps to represent goals and tasks. In [25], the authors study the integration of men and computers in firms dispersed geographically. They introduce the notion of computational assistants (or intelligent agents), helping for a task and resorting to the services of other agents in case of need. In [34], a framework is proposed in order to model complex organizations, with three levels of modelling: social, logical and physical. All such research gives indications on the possible organizational models that could be useful.

So, the knowledge engineer must elicit the information on the possible organizational structures of the compound agents that may appear during the problem solving, on the way such organizational structures can evolve, on the contexts where a given compound agent will change his organizational structure. Organizational models proposed in the current state of the art could be gathered in a library of predefined organizational models, that could guide the knowledge engineer: he would have to recognize the adequate model corresponding to his application and exploit this model for eliciting the needed information. In case of need, an adaptation of a predefined model or a combination of existing models would be necessary.

- *cooperation modes*: the agent's mode of cooperation may depend or not on the problem to be solved. For example, the class of problems studied can influence the kind of cooperation: in [3], the study of cooperation among designers influenced the multi-agent architecture implementing such a collaboration. In [10], cooperation relies on the notion of goal adoption, which implies a goal common to the different agents. Different types of cooperation are presented: accidental cooperation, unilaterally intended cooperation, mutual cooperation. Cooperation is considered as a function of mutual dependency among the agents. In [31], the author distinguishes negative cooperation (where the agents avoid to do the same task simultaneously) and positive cooperation (where the agents need one another to perform a task). We can adopt all these types of cooperation as possible modes of cooperation of the agents. They can be permanent (independent of the considered problem) or temporary (dependent on the problem to be solved). They may be independent of the other agents playing the role of partners or, on the contrary, depend on the considered partners.

In [24], different coordination modes inside an organization are described: by mutual adjusting, by direct supervision by standardization of procedures, of results or of qualifications. In some coordination modes such as standardization of procedures, the allocation of roles and tasks is rigid. Studies on the cooperation modes in norm-governed human organizations can be also exploited.

More generally, all this previous work on cooperation (either in the science of organizations or in the field of multi-agent systems) can be useful to propose some predefined models of cooperation [2]. Then, during the knowledge acquisition phase, the knowledge engineer can be guided by such models and try to recognize which predefined models can help describe the actual modes of cooperation used effectively by the experts. For example, collective study cases can be useful to elicit such information. As an agent can change his cooperation modes according to the context of problem solving or according to the other agents, the information on such evolution of cooperation modes and on the conditions where a given cooperation mode is preferable to another, must be elicited.

If the proposed predefined models of cooperation are not convenient, the knowledge engineer can adapt one of them or combine them or propose a new one, more adapted to the way the experts cooperate. This use of a library of cooperation models is very similar to the use of KADS library of interpretation models.

- *communication languages and protocols*: the agent may use various languages and protocols in order to communicate with the other agents, and, in particular, to send them requests, results, explanations... The communication languages may vary according to the considered agent. For example, a vehicle engineer may use a specific language, based on very technical terms, equations and graphics, when he works with another vehicle engineer and a simpler, non technical language when he works with experts of the other disciplines.

[2] The meaning we give here to the expression "model of cooperation" is different from KADS [11].

So, the knowledge engineer must elicit the communication languages and protocols of the experts and the contexts in which one expert uses a given language / protocol rather than another.

- *visibility from the other agents*: an agent has visible aspects (for example, observable actions, actions having observable consequences, messages sent to other agents) and internal aspects (for example, knowledge or cognitive actions that have no observable consequences). Some of the internal aspects (such as intentions) can be either explicitly expressed or when they are implicit, can be deduced by the other agents thanks to interpretations of the observable actions of the agent. A compound agent can be sometimes seen as a "blackbox" for the external world that can neither see his subagents nor deal with them directly. But it may be interesting to let the other agents know the structure of the compound agent, even if they cannot directly address requests to such subagents without using the organization interaction points.
- *joint plans*, *joint actions* and *commitments*: in order to achieve collectively a common goal of a compound agent, his subagents may make joint plans, that may comprise joint actions to be performed collectively by such subagents. The subagents can make commitments on the future.
  Analysis of collective studies of cases, or of thinking-aloud protocols during a collective problem-solving by several experts can help to reveal such joint plans and, in particular, indicate the explicit commitments that can be made within a group of experts and the links between the individual plans of the agents and the collective decisions / plans.
- *collective history*: the history of an agent's past interactions with the other agents may be useful, specially for cooperative problem solving or for explanation. His successive integrations in several compound agents, as well as his participation to joint actions inside such compound agents, are part of this collective history. As the agents' individual histories, this collective history can be exploited for explanations.

3.3 Operators on the Agents

Various operators among the agents can be thought out, such as:

- *creation of a new agent* at least identified by his name,
- *sending a message to one or several agents*: request, result, explanation... The addressees of the message must be identified either explicitly by their names or implicitly by the properties they must satisfy.
- *aggregation of several agents*, according to an organizational structure, for a common goal and for a given time duration. This aggregation of agents is a particular case of creation of an agent.
- *decomposition of a task among several agents*: it may be reserved to agents linked through an organizational structure in a compound agent or it may be possible even for separate, independent agents.

- *modification of a compound agent*: he may integrate new subagents, remove some subagents, reorganize himself (i.e. modify his organizational structure).
- *death of an agent*: the dead agent (and in particular, his identity) no longer exists. He is no longer known by any other agent. So, all the compound agents to which the dead agent previously belonged to, must reorganize themselves in order to take this death into account.
- *dissociation of a compound agent*: his subagents become separate, independent agents, and, in particular, the organizational structure of the compound agent disappears, as well as his identity. This dissociation of agents is a particular case of death of an agent.

The knowledge engineer can acquire information on the contexts where such information can be used and exploit them for the design of the final KBS if it is based on a multi-agent architecture.

4 Agent-based Knowledge Acquisition

Knowledge acquisition consists of knowledge elicitation and knowledge modelling. Once the previous model of agent is available, knowledge acquisition can be seen as the process of *identifying the adequate involved agents* and then *building the corresponding artificial agents* in the knowledge acquisition tool and *filling them progressively*.

The identification phase consists of identifying the different kinds of humans involved in knowledge acquisition phase and specially the experts and potential end-users in a given organization. Then the artificial agents that will represent them in the knowledge acquisition tool are progressively identified and built: according to the case, such artificial agents may correspond to one expert (resp. user), to a group of experts (resp. users) or to a combination of subparts of experts (resp. users). So, an agent may represent an expert totally or partially. More precisely, an expert may be represented by a compound agent, made of: a) an agent common to all experts, b) agents common to this expert and to some other experts, c) and an agent representing the specificities of this expert.

The filling of the agents consists of eliciting and modelling knowledge from the adequate human agents (knowledge on the human organization, on the varied expertises, on the cooperation modes, etc.) in order to be able to fill the different individual and social features that must characterize the associated artificial agents. The long-term, problem-independent characteristics such as competence domain, tasks, high-level goals can generally be elicited through individual elicitation techniques. Individual observations or thinking-aloud protocols on case studies by a single expert can reveal also information on problem-dependent, individual features. The expertise model of an individual agent can be built using KADS method and its domain layer structured using our formalism of knowledge graphs. Collective elicitation techniques involving several experts are particularly helpful for refining the previous individual features, but mainly for eliciting the social features of the agents: interaction points, cooperation modes, adaptation of such cooperation modes and of communication language to the context and to the partners, conditions of creation of temporary compound agents for solving a subproblem, etc.

The knowledge engineer can adopt a bottom-up approach (construction of the simple agents and then of the compound agents) or a top-down approach (construction of the compound agents and then of their subagents). The comparison of expertise models can lead to the "decomposition of a given expert" in several subagents ot to the "gathering of several experts" into a single artificial, compound agent.

Identification and filling of the agents are in fact interleaved as, throughout the knowledge acquisition process, the need to split an agent into several ones or to gather several agents into a compound one may appear.

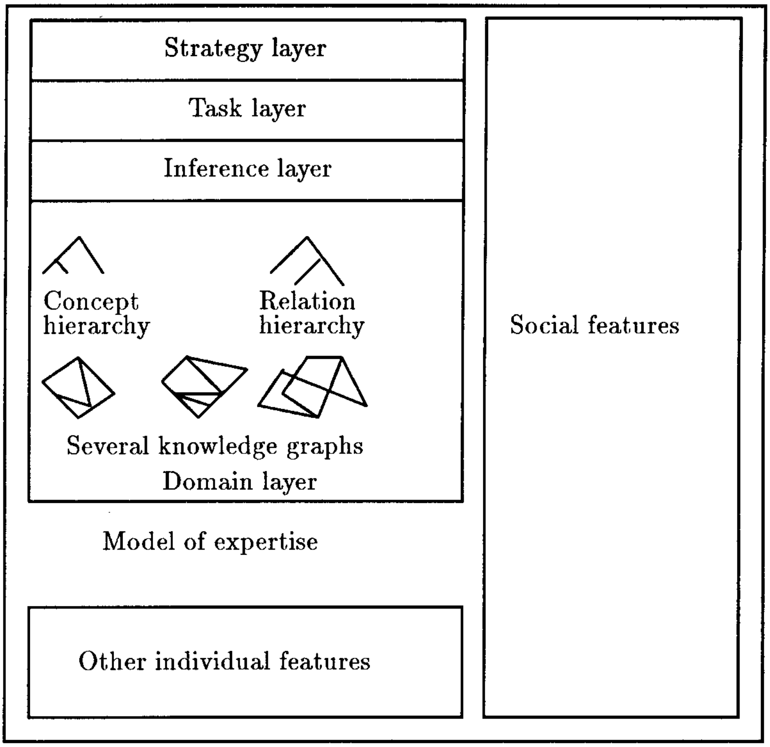

Figure 1: An agent in KATEMES.

The help offered by a knowledge acquisition method / tool can consist of offering a list of predefined possible values, for each characteristics of the model of agent. Such possible values can stem from the state of the art in DAI (in particular, multi-agent systems) or in science of organizations or from experimental results. In the section 3, we studied some of the possible values for the features of our model of agent, by relying on the current state of the art in DAI. Clearly, this work needs to be studied more thoroughly, for example, in order to make choices between several approaches of DAI. We do not intend to offer new DAI concepts or a new DAI tool, but rather to rely on what is currently known and accepted in DAI in order to offer a help during knowledge acquisition phase.

A knowledge acquisition method / tool can offer a library of organizational structures, of cooperation modes, of communication languages based on the current state of the art in DAI, with information on the contexts where each element of the library should be preferably chosen instead of another, and the adequate techniques for eliciting the information needed by this element. Such libraries can then guide the knowledge engineer in a top-down way, in order to elicit knowledge, once he has recognized the model convenient for his application. As for the exploitation of KADS library of interpretation models, he can adapt a model that seems interesting but not perfect, or combine several models, or build a completely new model if none of the predefined models could be adapted or combined.

At present, the proposed agent-based knowledge acquisition method is supported by KATEMES, a knowledge acquisition tool aimed at tackling knowledge acquisition from multiple experts [14]. Notice that KATEMES is not aimed at being a workbench for simulating agents: as it is intended to be used only during knowledge acquisition phase, it only helps to acquire knowledge on the way the human agents behave for problem-solving and for their interaction but it does not simulate such behaviour, and it does not offer any operationalization of the agent model.

For the detection of conflicts among several expertises, KATEMES focuses on the expertise knowledge of the agents. This expertise knowledge is modeled through a KADS-like model of expertise: the reasoning is described through the inference, task and strategy layers. In addition, in KATEMES, the concepts and relations of the domain layer are represented through knowledge graphs. An agent has a hierarchy of concepts, a hierarchy of relations and different knowledge graphs, as shown in Figure 1. Therefore, for the comparison of expertise models, KATEMES focuses on the comparison of such knowledge graphs. The description of our method of comparison of knowledge graphs is out of the scope of this paper and is detailed in [13].
If successful, the algorithm of comparison of the different knowledge graphs of the same nature must lead to unified knowledge graphs, according to the strategy of fusion adopted. Then a new agent must be built, having as expertise knowledge the knowledge common to both experts, i.e. described in the domain layer by the unified knowledge graphs. The remaining parts of each expert (for example, when he owned a knowledge graph for a viewpoint that did not exist for the other expert) constitute his specific knowledge and can be gathered in a new agent corresponding to this expert's specificities (and called agent specific to this expert). For example, for the house building, after comparison of the experts in electricity and in mechanics, the "electrical viewpoint knowledge graph" of the expert in electricity will not be part of the expertise of the common agent and will remain in the electrician's specific agent. Each of the two experts that were compared becomes then a compound agent, made of the common agent and the specific agent.

5 Application

We are applying this agent-based knowledge acquisition method to the design of an expert system in road safety. This application is described thoroughly in [2]. Psychology, vehicle engineering and road infrastructure engineering are the competence domains of the experts involved in traffic accident analysis. We elicited knowledge from two psychologists, three engineers in road infrastructure and two vehicle engineers. Each of the experts, being himself a driver, had at least a common-sense model of the driver, of the vehicle and of the road. But the psychologists had a deeper, finer model of the driver, the vehicle engineers had a finer model of the vehicle and the infrastructure engineers had a finer model of the road infrastructure. Moreover, within a given discipline, there were differences between the specialists: such differences were detailed in [13] and are summed up in figure 2.

For the knowledge elicitation techniques, we used both individual techniques (interview of one expert; thinking aloud protocols of one expert solving a case) and collective techniques (collective solving of a case by several experts either from the same discipline of from different ones; individual solving of a given case by several experts separately, followed by a meeting where they compared their respective solutions and discussed about them, etc.). The individual sessions were helpful to determine the individual characteristics of each expert (competence domain, tasks, high-level goals, resources, expertise model) as well as some social aspects such as the way each expert explicitly described his model of the other experts or his cooperation modes. The collective sessions helped us to refine the individual characteristics and to elicit the social aspects (the experts' interaction points, the way some experts could gather and reorganize temporarily during a given problem solving, their cooperation modes, the way they adapted their communication language to the interlocutor, etc.). The analysis of the case studies gave information on the problem-dependent characteristics of the experts (in particular, their low-level goals, their plans and their successive actions).

The various examples that illustrated the previous section on our model of agent were based on the characteristics obtained for the agents associated to the experts in traffic accident analysis. We will not recapitulate each of such features.

Lastly, the comparison between the experts allowed to progressively determine which artificial agents would represent them. The following comparison procedure seemed interesting: (1) compare both psychologists, (2) compare both vehicle engineers, (3) compare the three infrastructure engineers, either directly, or progressively by first comparing the two infrastructure engineers that know well the region, and had been in the field, and then comparing the obtained common agent with the third infrastructure engineer that has a theoretical background and is not living in the region, (4) compare the three possible common agents obtained.

When the common agent represents only what is really common to the different agents to be compared, the agents obtained are described in figure 2.

Remark: This work is still in progress, as our representation of all the agents and knowledge graphs associated to the different experts is not yet complete.

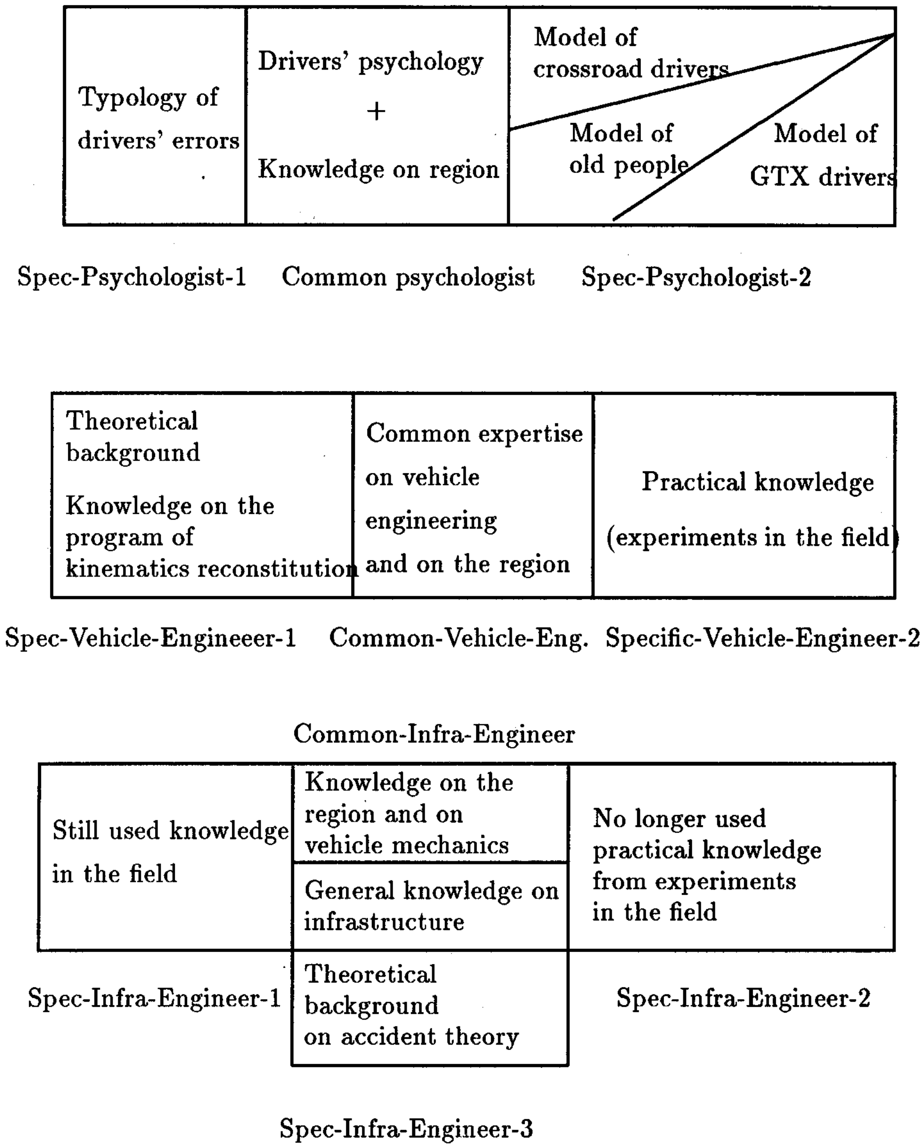

Figure 2: The agents involved in traffic accident analysis.

6 Conclusions

This paper proposed elements of an informal model of agent, allowing to represent the different kinds of humans involved in knowledge acquisition phase and specially the experts and potential end-users in a given organization. This model of agent can then serve as a guide for the design of the object multi-agent system. We applied our agent-based knowledge acquisition method to traffic accident analysis.

Related Work

In [19] the link between distributed knowledge acquisition and multi-agent systems is emphasized. The author adopts a blackboard architecture in order to elicit software requirements from multiple clients (approach similar to knowlege acquisition from multiple experts), and in order to solve conflicts thanks to the splitting of agents. Techniques for comparing several viewpoints and solving conflicts among them are described in [18]. We took inspiration of some aspects of such research for the splitting of agents for solving conflicts, but our model of agent and our agent-based knowledge acquisition method offers new features, different from this approach.

The choice of a multi-agent architecture serves as a methodological guide for knowledge acquisition in [20] and the *EMA* knowledge acquisition method handles a notion of agent [29], but, in both cases, the authors do not evoke the possible exploitation of this notion for knowledge acquisition from multiple experts.

CommonKads [1] offers a model of agent that, as the authors indicate, "serves as a link between the task model, the communication model and the expertise model, by modelling the capabilities and constraints that the experts have, which are involved in solving a task". Clearly, our model of agent aims at the same purpose as the one of CommonKads but, in addition, it is actually the central model in our knowledge acquisition method, which is entirely guided by it: moreover, it contains aspects concerning task, expertise, communication and organization, it aims at modelling cooperative problem solving by several agents and it is strongly inspired of work from DAI field.

Our research is also linked to research analysing software design as an agent-based activity, since a KBS is a particular case of software. In [23], the conflicts among design agents (that may be human or artificial agents) can be detected and solved, in a tool aimed at supporting cooperative design. It would be interesting to study whether such techniques, aimed at classic software, can be extended or adapted for cooperative design of a KBS.

6.1 Discussion and Further Work

As indicated earlier, we studied the current state of the art in DAI in order to indicate some of the possible values for the characteristics of our model of agent. In our study of DAI research, we naturally focused on the aspects that could be applied to human agents such as experts and users, since some characteristics of artificial agents in multi-agent systems cannot, of course, be adequate for modelling the capabilities and behaviour of human agents. But our work must not be seen as a workbench for simulating agents: it only aims at helping to acquire knowledge on the way the human agents in the organization behave for problem-solving and for their interaction, but without simulating such behaviour. In particular, it offers nothing for the operationalization of the agent model. It can be considered as a specification of the capabilities and behaviour of the agents that will appear in the final KBS if it based on a multi-agent architecture.

As an extension of this work, we will study more thoroughly how to be able to offer a guide for the construction of the final KBS: how to exploit our model of agent, in particular, if a DAI tool can be used for the generation of the operational KBS ? How to guide this operationalization while respecting the specifications obtained after knowledge acquisition phase on the behaviour of the agents ?

As a further work, we will try to formalize our model of agent: the present description of the agent's characteristics is informal and based on a synthesis of the properties most often evoked in the field of science of organizations or of multi-agent systems; but we intend to offer a more formal basis for such a description. We will try to take benefit of research on formal models of agents, that is carried out in various areas such as software design, requirements engineering, organizational models or composite systems, within the BRA Working Group MODELAGE.

We will also study methods of comparison of the different KADS layers, so that the splitting of an agent into several agents no longer relies only on the comparison of knowledge graphs performed in the domain layer. Last, we will adapt our work to COMMONKADS instead of relying on KADS-I only.

For the validation of our method through a concrete application, we will complete the representation of the different agents for traffic accident analysis.

Acknowledgements

We deeply thank the French "Ministère de la Recherche et de l'Espace" (contract n. 92 C 0757), "Ministère de l'Equipement, des Transports et du Tourisme" (contract n. 2 93 E 403) and the European Community (contract EP:8319) that funded this research. We also thank very much our colleagues of the French Working Groups MULTI-EXP and COOP, the members of the European KARMA Group and our partners of the Esprit Working Group MODELAGE for our fruitful discussions on knowledge acquisition from multiple experts, on models of agents or on cooperative systems. We thank very much our colleagues from the ACACIA project, from INRETS and from Paris V University that took part in the knowledge elicitation phase upon the traffic accident analysis.

References

1. *The Common KADS Agent Model, KADS-II/M4/TR/SICS/002/V1.1.* May 1993.
2. Ch. Amergé, O. Corby, R. Dieng, D. Fleury, A. Giboin, S. Labidi, and S. Lapalut. Acquisition et modélisation des connaissances dans le cadre d'une coopération entre plusieurs experts : Application à un système d'aide à l'analyse de l'accident de la route. Rapport intermédiaire du contrat MRE n° 92 C 0757, décembre 1993.
3. A. H. Bond. *The Cooperation of Experts in Engineering Design*, chapter 18, pages 463–483. Pitman / Morgan Kaufmann Publ. Inc., November 1989.
4. A. H. Bond. A normative model of collaboration in organizations. In *Proc. of 1990 Workshop on DAI*, Bandera, Texas, October 1990.
5. A. H. Bond and L. Gasser. An analysis of problems and research in distributed artificial intelligence. In *Readings in DAI.* Morgan Kaufman Publ., 1988.
6. J. Breuker, B. Wielinga, M. van Someren, R. de Hoog, G. Schreiber, P. de Greef, B. Bredeweg, J. Wielemaker, and J.-P. Billault. *Model Driven Knowledge Acquisition: Interpretation Models*, 1987. Deliverable A1, Esprit Proj. 1098, Memo 87.
7. B. Burmeister and K. Sundermeyer. Cooperative problem-solving guided by intentions and perception. In Werner and Demazeau, editors, *Decentralized AI, 3*, pages 77–92. Els. Sc. Publ., June 1992.
8. B. Chandrasekaran, T. R. Johnson, and J. W. Smith. Task-Structure Analysis for Knowledge Modeling. *Comm. of the ACM*, 35(9):124–137, September 1992.
9. W. J. Clancey. Heuristic classification. *Art. Int.*, 27(3):289–350, Dec. 1985.
10. R. Conte, M. Miceli, and C. Castelfranchi. Limits and levels of cooperation : disentangling various types of prosocial interaction. In Demazeau and Muller, editors, *Decentralized AI, 2*, pages 147–157. Els. Sc. Publ. B.V., 1991.
11. P. de Greef and J. Breuker. Analysing system-user cooperation. *Knowledge Acquisition*, 4:89–108, 1992.
12. R. Dieng. Relations linking Cooperating Agents. In Demazeau and Muller, editors, *Decentralized AI, 2*, pages 51–69. Els. Sc. Publ. B.V., 1991.
13. R. Dieng, O. Corby, and S. Labidi. Expertise conflicts in knowledge acquisition. In *Proc. of the 8th KAW*, vol. 2, pages 23.1–23.19, Banff, Canada, 1994.
14. R. Dieng, A. Giboin, P.-A. Tourtier, and O. Corby. Knowledge Acquisition for Explainable, Multi-Expert Knowledge-Based Design Systems. In Wetter & al editors, *Current Developments in Knowledge Acquisition : EKAW-92*, pages 298–317, Berlin/Heidelberg, May 1992. Springer-Verlag.
15. R. Dieng and P.A. Tourtier. A Composite System for Knowledge Acquisition and User Assistance. In *Proc. of AAAI Spring Symposium on Design of Composite Systems*, pages 1–5, Stanford, CA, March 1991.
16. R. Dieng and B. Trousse. 3DKAT, a Dependency-Driven Dynamic-Knowledge Acquisition Tool. In *Proc. of the 3rd Int. Symp. on Knowledge Engineering*, pages 85–93, Madrid, Spain, October 1988.
17. E. Doerry, S. Fickas, R. Helm, and M.S. Feather. A Model for Composite System Design. In *Proc. of AAAI Spring Symposium on Design of Composite Systems*, pages 6–10, Stanford, CA, March 1991.
18. S. Easterbrook. Handling conflict between domain descriptions with computer-supported negotiation. *Knowledge Acquisition*, 3(3):255–289, September 1991.
19. S. M. Easterbrook. Distributed Knowledge Acquisition as a Model for Requirements Elicitation. In *Proc. of the 3rd EKAW*, pages 530–543, Paris, July 1989.
20. C. Ferraris. *Acquisition des connaissances et raisonnement dans un univers multi-agents : application à la prise de décision en génie civil urbain*, Février 1992. Thèse de Doctorat en Informatique, Université de Nancy I.
21. B. R. Gaines and M. L. G. Shaw. Comparing the Conceptual Systems of Experts. In *Proc. of the 9th IJCAI (IJCAI-89)*, pages 633–638, Detroit, 1989.

22. L. Gasser. Boundaries, Identity and Aggregation: Plurality Issues in Multi-Agent Systems. In Werner and Demazeau, editors, *Decentralized AI, 3*, pages 199–213. Els. Sc. Publ., June 1992.
23. M. Klein. Detecting and resolving conflicts among cooperating human and machine-based design agents. *The Int. Journ. for AI in Engin.*, 7:93–104, 1992.
24. H. Mintzberg, editor. *The Structuring of Organizations*. Prentice Hall Inc., 1979.
25. J. Y. C. Pan and J. M. Tenebaum. An Intelligent Agent Framework for Enterprise Integration. In *Proc. of AAAI Spring Symposium on Design of Composite Systems*, pages 69–96, Stanford University, CA, March 1991.
26. A. S. Rao, M. P. Georgeff, and E. A. Sonenberg. Social Plans: a Preliminary Report. In Werner and Demazeau, editors, *Decentralized AI, 3*, pages 57–76. Els. Sc. Publ., June 1992.
27. M. L. G. Shaw and B. R. Gaines. A methodology for recognizing conflict, correspondence, consensus and contrast in a knowledge acquisition system. *Knowledge Acquisition*, 1(4):341–363, December 1989.
28. M. L. G. Shaw and J. B. Woodward. Validation in a knowledge support system: construing consistency with multiple experts. *International Journal of Man-machine Studies*, 29:329–350, 1988.
29. S. Spirgi and D. Wenger. Generic techniques in EMA: A model-based approach for Knowledge Acquisition. In *Proc. of the 5th EKAW*, Crieff, UK, May 1991.
30. L. Steels. *COMMET: a componential methodology for knowledge engineering.* 1991. Esprit Project CONSTRUCT DWP/2/3/4.
31. E. Werner. Social Intentions. In *ECAI-88*, pages 719–723, Munich, 1988.
32. E. Werner. The design of multi-agent systems. In Werner and Demazeau, editors, *Decentralized AI, 3*. Els. Sc. Publ., June 1992.
33. B. J. Wielinga, A. T. Schreiber, and J. A. Breuker. KADS: a modelling approach to knowledge engineering. *Knowledge Acquisition*, 4:5–53, 1992.
34. E. Yu. Organization Modelling and Composite Systems. In *Proc. of AAAI Spring Symposium on Design of Composite Systems*, pages 122–126, Stanford University, CA, March 1991.