KNOWLEDGE ACQUISITION MODELING THROUGH DIALOGUE BETWEEN COGNITIVE AGENTS

Mehdi Yousfi-Monod, Violaine Prince
LIRMM, UMR 5506
161 rue Ada, 34392 Montpellier Cedex 5, France

Keywords: Intelligent agents dialogue, tutoring systems , logical reasoning.

Abstract: The work described in this paper tackles learning and communication between cognitive artificial agents. Focus is on dialogue as the only way for agents to acquire knowledge, as it often happens in natural situations. Since this restriction has scarcely been studied as such in artificial intelligence (AI), until now, this research aims at providing a dialogue model devoted to knowledge acquisition. It allows two agents, in a 'teacher' - 'student' relationship, to exchange information with a learning incentive (on behalf of the 'student'). The article first defines the nature of the addressed agents, the types of relation they maintain, and the structure and contents of their knowledge base. It continues by describing the different aims of learning, their realization and the solutions provided for problems encountered by agents. A general architecture is then established and a comment on an a part of the theory implementation is given. Conclusion is about the achievements carried out and the potential improvement of this work.

1 INTRODUCTION

This research aims at defining a set of algorithms for knowledge acquisition through dialogue between artificial cognitive agents. By *cognitive agents* we mean entities possessing knowledge as well as acquisition and derivation modes. In other words, they are able to capture knowledge externally, and to process and modify it through reasoning. Moreover, agents are characterized by one or several goals. As an artificial intelligence (AI) entity, each agent owns a knowledge base and attempts to make it evolve either by environment observation (reactivity) or by derivation modes (inductive or deductive reasoning). However, human beings as natural cognitive agents favor dialogue as another mean for knowledge revision. This mean tends to consider another agent as a knowledge source, and then to proceed to derivation (by reasoning). Further, the knowledge source could be addressed in order to test wether the acquisition process has succeeded. In a nutshell, this is what happens in *tutored learning*.

To simulate learning we have chosen a socratic dialogue ('teacher' - 'student') where knowledge is presented exclusively by means of a question-answer mode of interaction. The 'student' agent owns belief revision mechanisms and all axioms allowing formal reasoning. Being in the modeling process, we wished not to overload dialogue difficulties with language intrinsic ambiguity (i.e. pure "natural language" problems) and thus specified a "keleton protocol": message data will be exchanged in first-order logic. However, this has been decided only to be able to focus on dialogue particular features; we attempt, the best we can, to make the dialogue situation as close to a natural human-human dialogue between a 'teacher' and a 'student', as possible. We assume that agents use a common formalism concerning terms, predicates and functions. Nevertheless, the 'student' agent may not have predicates (or functions) given by the 'teacher' agent and so can question it on this subject before revising its base. In this paper, we try to show how dialogue initiates reasoning, which leads to an increase as well as a revision of the 'student' knowledge base according to hypotheses we simulate and revision mechanisms we define.

Yousfi-Monod M. and Prince V. (2005).
KNOWLEDGE ACQUISITION MODELING THROUGH DIALOGUE BETWEEN COGNITIVE AGENTS.
In *Proceedings of the Seventh International Conference on Enterprise Information Systems*, pages 201-206
DOI: 10.5220/0002512002010206

2 DIALOGUE AND LEARNING: A BRIEF DESCRIPTION OF RELATED LITERATURE

Several papers deal with human learning via dialogue (DA 91). Those related to computer devices usually rely on human-machine dialogue models (Bak 94; Coo 00). However, for artificial agents only, the very few papers about communication as an acquisition mode are in the framework of noncognitive environment like robots (AMH 96) or noncognitive software agents. It seems that, in artificial systems, learning is often realized without dialogue.

Learning without Dialogue. There are many kind of learning methods for symbolic agents like reinforcement learning, supervised learning (sometimes using communication as in (Mat 97)), without speaking about neural networks models that are very far from our domain. This type of learning prepares agents for typical situations, whereas, a natural situation in which dialogue influences knowledge acquisition, has a great chance to be unique and not very predictable (RP 00).

Dialogue Models. Most dialogue models in computer science (namely in AI) are based on *intentions* (AP 80; CL 92), rely on the Speech Act Theory (Aus 62; Sea 69), to define dialogue as a succession of planned comunicative actions modifying implicated agents' mental state, thus emphasizing the importance of *plans* (Pol 98). When agents are in a knowledge acquisition or transfer situation, they have goals: teach or learn a set of knowledge chunks. However, they do not have predetermined plans: they react *step by step*, according to the interlocutor's answer. This is why an *opportunistic* model of linguistic actions is better than a planning model. Clearly, a tutored learning situation implies a *finalized* dialogue (aiming at carrying out a task) as well as *secondary* exchanges (precision, explanation, confirmation and reformulation requests can take place to validate a question or an answer). We have chosen to assign functional roles (FR) to speech acts since this method, described in (SFP 98), allows unpredictable situations modelling, and tries to compute an exchange as an adjusment between locutors mental states. We have adapted this method, originally designed for human-machine dialogue, to artificial agents.

Reasoning. Reasoning, from a learning point of view, is a knowledge derivation mode, included in agent functionalities, or offered by the 'teacher' agent. Reasoning modifies the recipient agent state, through a set of reasoning steps. *Learning* is considered as the result of a reasoning procedure over new facts or predicates, that ends up in engulfing them in the agent knowledge base. Thus, inspired from human behavior, the described model acknowledges for three types of reasoning: deduction, induction and abduction. Currently, our system uses inductive and deductive mechanisms. Abduction is not investigated as such, since we consider *dialogue as an abductive bootstrap technique* which, by presenting new knowledge, enables knowledge addition or retraction and therefore leads to knowledge revision (JJ 94; Pag 96).

Last, although our system is heavily inspired from dialogue between humans and from human-machine dialogue systems, it differs from them with respect to the following items :

- Natural language is not used as such and a formal-based language is prefered, in the tradition of languages such as KIF, that are thoroughly employed in artificial agents communication. These formal languages prevent problems that rise from the ambiguity intrinsic to natural language.
- When one of the agents is human, then his/her knowledge is opaque not only to his/her interlocutor (here, the system) but also to the designer of the system. Therefore, the designer must build, in his system, a series of "guessing" strategies, that do not necessarily fathom the interlocutor's state of mind, and might lead to failure in dialogue. Whereas, when both agents are artificial, they are both transparent to the designer, if not to each other. Thus, the designer embeds, in both, tools for communication that are adapted to their knowledge level. The designer might check, at any moment, the state variables of both agents, a thing he cannot do with a human.

These two restrictions tend to simplify the problem, and more, to stick to the real core of the task, i.e., controlling acquisition through interaction.

3 THE THEORETICAL FRAMEWORK

3.1 Agents Frame

Our environment focuses on a situation where two cognitive artificial agents are present, and their sole interaction is through dialogue. During this relationship, an agent will play the role of a 'teacher' and the other will momentarily act as a 'student'. We assume they will keep this status during the dialogue session. Nevertheless, role assignation is temporary because it depends on the task to achieve and on each agent's skills. The 'teacher' agent must have the required skill to teach to the 'student' agent, i.e., *to offer unknown and true knowledge, necessary for the 'student' to perform a given task*. Conventionally, 'student' and 'teacher' terms will be used to refer, respectively, to the agents acting as such. The 'teacher' aims at providing a set of predetermined knowledge to the ’student’. This, naturally subsumes that agents cooperate.

3.2 Knowledge Base Properties

Each agent owns a knowledge base (KB), structured in first-order logic, so the knowledge unit is a formula. The ’student’ can make mistakes, i.e., possess wrong knowledge. From an axiomatic point of view, if an agent acts as a ’teacher’ in relation to a given knowledge set, then the ’student’ will consider as true every item provided by the ’teacher’.

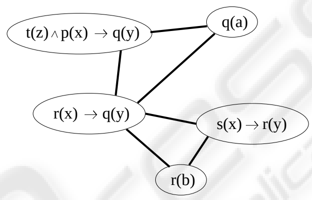
Each KB is manually initiated, however, its update will be automatic, thanks to ’learning’ and reasoning abilities. In order to simplify modelling, we only use formulas such sa (P), $(P \rightarrow Q)$ and $(P \leftrightarrow Q)$. (P) and (Q) are predicates conjunctions (or their negation) of type $(p(A))$ or $(p(X))$ (or $(not(p(A)))$ or $(not(p(X)))$), where $A = \{a_1, a_2, \dots, a_n\}$ is a set of terms and $X = \{x_1, x_2, \dots, x_n\}$ a set of variables. For simplification sake, we note P and Q such predicates conjunctions. Universal quantification is implicit for each formula having at least one variable. We consider that, to initiate learning (from the ’student’ position), the ’teacher’ has to rely on the ’student’s previous knowledge. This constraint imitates humans’ learning methods. Therefore, before performing a tutored learning dialogue, agents must have a part of their knowledge identical (called *basic common knowledge*). The ’teacher’ will be able to teach new knowledge by using the ’student”s already known one. However, our agents do not ’physically’ share any knowledge (their KBs are independent).

Connexity as a KB Fundamental Property. During learning, each agent will attempt to make its KB as “connex” as possible.

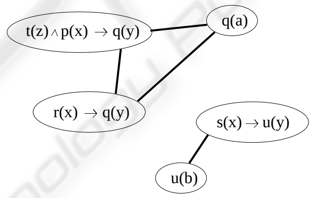
Definition 1. A KB is *connex* when its associated graph is connex. A graph G_Γ is associated to a KB Γ as such:
Each formula is a node. An edge is created between each couple of formulas having the same premise or the same conclusion or when the premise of one equals the conclusion of the other. For the moment, variables and terms are not took into account in premise or conclusion comparison [1]. Thus, in a connex KB, every knowledge element is linked to every other, the path between them being more or less long. As the dialogic situation must be as close as possible to a natural situation, **agents’ KBs are not totally connex**: a human agent can often, but not always, link two items of knowledge, haphazardly taken.

Examples:
A connex KB: $\Gamma_1 = \{t(z) \wedge p(x) \rightarrow q(y), r(x) \rightarrow q(y), s(x) \rightarrow r(y), q(a), r(b)\}$
A non connex KB: $\Gamma_2 = \{t(z) \wedge p(x) \rightarrow q(y), r(x) \rightarrow q(y), s(x) \rightarrow u(y), q(a), u(b)\}$

Definition 2. A connex component (or just component) is a connex subset of formulas in a KB.



(a) Γ_1 Associated Graph



(b) Γ_2 Associated Graph

Figure 1: KB Associated Graphs

Figures 1(a) and 1(b) respectively represent graphs associated to Γ_1 and Γ_2.

Theorem 1. Let A, B and C be three connex formulas sets. If $A \cup B$ and $B \cup C$ are connex then $A \cup B \cup C$ is connex.

Proof. Let us assume that $A \cup B$ and $B \cup C$ is connex and G_A, G_B and G_C are graphs respectively associated to A, B and C. According to definition 1: $A \cup B$ connex is equivalent to $G_A \cup G_B$ connex. Also, $B \cup C$ connex is equivalent to $G_B \cup G_C$ connex. And according to connex graph properties: $G_A \cup G_B$ connex and $G_B \cup G_C$ connex implies $G_A \cup G_B \cup G_C$ connex. So $A \cup B \cup C$ is connex.

Agents will not attempt to increase the number of their connex components. However, there will be some cases where the ’student’ will be forced to do so. Fortunately, in some other cases, learning new knowledge may link two connex components into a new larger one, decreasing the components total number (according to theorem 1).

[1]An abductive reasoning mechanism is contemplated as a possible mean to compare a constant fact $q(a)$ with a predicate with a variable $q(y)$. We only consider the result of a succeeding abduction.

3.3 Dialogue: Using Functional Roles (FR)

A dialogue session is the image of a lesson. The 'teacher' must know what knowledge to teach to the 'student': therefore, a lesson is composed of several elements, each of them contained in a logic formula. The teaching agent provides each formula to the 'student'. However, before that, the teacher waits for the 'student''s understanding (or misunderstanding) message of the last formula. If the 'student' doesn't understand or is not at ease, it can just inform its interlocutor of the misunderstanding or, requests a particular information bit. The FR theory, that models exchanges in this dialogue, allows the attachment of a role to each utterance. Both agents, when receiving a message, know its role and can provide an adequate answer. We assign the type *knowledge* to universal or existential general logical formulas and the type *information* to constant-uttering answer (or question): i.e, which value is 'true', 'false', or 'unknown'. Here are the main FR types used in the our tutored learning dialogue.

1. give-knowledge. Used to teach a knowledge and introduce an exchange, example:
   $give - knowledge(cat(x) \rightarrow mortal(x))$:
   "Cats are mortal."
2. askfor/give-information (boolean evaluation case) :
   askfor-information. Example:
   $askfor - information(cat(Folley))$:
   "Is Folley a cat?"
   give-information. Example:
   $give - information(true)$: "Yes."
3. give-explanation (predicate case). Example:
   $give - explanation(cat(x) \leftrightarrow (animal(x)\, and\, pet(x)\,))$: "A cat is a pet animal."
4. say-(dis)satisfaction: tells the other agent that the last provided data has (has not) been well understood.

There are some FR we do not use ($askfor - knowledge, askfor-explanation, askfor/give-example, askfor/give-precision, askfor/give-reformulation$) likewise some specific uses like the type $askfor/give - information$ in the case of an evaluation by a function. So FR are dialogic clauses allowing the interpretation of exchanged formulas. A functional role of the"$ask - for$" kind implies one or a series of clauses of the "$give$" type, with the possibility of using another "$ask - for$" type if there is a misunderstanding. This case will bring about a clause without argument: "$say - dissatisfaction$". Only "$ask - for$" type roles will lead to interpretative axioms. Other ones are behavioral startings.

3.4 Tutored Learning

3.4.1 Axioms

Our reasoning system is hypothetical-deductive, so it allows belief revision and dialogue is the mean by which this revision is performed. Two groups of axioms are defined: fundamental axioms of the system and those corresponding to the FR interpretation in the system. Each knowledge chunk of each agent is seen as an assumption.

Fundamental axioms. Our system revision axioms include the hypothetisation axiom, hypothesis addition and retraction, implication addition, implication retraction or modus ponens and the *reductio ad absurdum rule.* These are extensively described in (Man 74).

FR interpretation axioms. Interpretation axioms are not in the first order since they introduce clauses and multiple values (like the "unknown" one). Our syntax will be in the first order, but the interpretation is not monotonous.

- $give - knowledge(A) \Rightarrow A \vdash T$;
  any knowledge supplied by the teacher is considered as true.
- $give - information(A) \equiv A\epsilon[T, F, U]$;
  any supplied information is a formula interpretable in a multi-valued space.
- $give - explanation(A) \equiv$
  $(give - information(P), A \leftrightarrow P)$;
  any explanation consists in supplying a right formula, equivalent to the formula A that has to be explained.

With T for True, F for False and U for Unknown.

3.4.2 Tutored Learning Situations

Learning can have several goals like enriching the KB with new data, increasing the KB connexity, widening the predicates base, understanding why some formulas imply others. We mainly focus on the first one because of its importance. In order to learn, the 'student' must first understand received data. By *understanding,* we mean "not increasing the KB components number": the 'student' understands a data that is linked to at least one component of its KB. By definition, we consider that a 'student' *knows* a predicate if it owns it.

3.5 Dialogue Strategy

There are several dialogue strategies depending on the goals chosen by the learner. In this paper, being limited in scope, we consider one goal: enriching the KB with new data (while maintaining connexity as best as

possible), because it is the commonest. Thus we suggest the appropriate common strategy: solving a misunderstanding problem by choosing adequate questions and answers. We have adopted a technique inspired from the socratic teaching method.
For each predicate p_i to be taught, the 'teacher' knows another one p_j linked with p_i by an implication or an equivalence F. Therefore, to ensure that the 'student' understands p_i thanks to p_j, he will have to ask the 'student' if the latter knows p_j. If the 'student' knows it, then the 'teacher' only have to give F to the 'student'. Otherwise, the 'teacher' will find another formula that explains p_j and so on. Once data is understood, the 'student' may realize that some bits are contradictory with its KB, leading to a *conflict*.
Conflict Management. We have studied several types of conflict, those related to implications as well as those related to facts. In this paper, we will only present the first one, which typically takes place when the 'student' has a formula $(P \rightarrow Q)$ and attempts to learn a formula $(P \rightarrow not(Q))$. The solution, for the 'student', is removing $(P \rightarrow Q)$ from its KB and adding $(P \rightarrow not(Q))$. It acts so because we consider that this is **'teacher''s knowledge** (thus true) and so it gets the upper hand on the 'student' one (first axiom). However, the conflict could be hidden if the 'student' has the next formulas: $(P_1 \rightarrow P_2)$, $(P_2 \rightarrow P_3)$, ... , $(P_{n-1} \rightarrow P_n)$ and attempts to learn $(P_1 \rightarrow not(P_n))$: the 'student' has an equivalent to the formula $(P_1 \rightarrow P_n)$. Instead of using a baseline solution consisting in removing all the series of implications, we opted for a more flexible one which attempts to look for a wrong implication and only remove this one. Indeed, removing one implication is sufficient to solve the conflict. The 'student' will then attempt to validate each implication to the 'teacher' through an "$askfor - information$" request. As soon as a wrong implication is found, the 'student' removes it and safely adds the new one. However, if none of the implications is neither validated nor rejected by the 'teacher', the 'student' will be forced to remove all the series before adding the new one to be sure to end up the conflict. We have studied other implication conflict types that are even less easily detectable, but we have not a sufficient space to detail them.

4 SYSTEM ARCHITECTURE AND IMPLEMENTATION

The theoretical approach of section 3 has been specified and partially implemented. The specification is general, the implementation contains some of its elements presented in section 4.2.

4.1 Architecture

The figure 2 displays the main architecture elements of our tutored learning system. It is composed of five main structures: the 'teacher', the 'student', the FR, the strategies and the 'World'. Each agent has a KB, a model of itself and of its interlocutor. It can freely

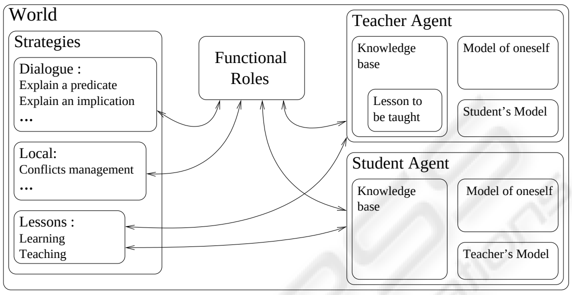


Figure 2: The tutored learning system through dialogue between artificial cognitive agents

update them in order to make them evolve. It has an access to strategies, for learning and teaching lessons and to all the FR rules (seen in section 3.3).

4.2 Implementation

We have implemented a Java program to test conflict solving. This programm is a basic prototype aiming at getting experimental results of a part of our theory. Each agent is an instance of the Thread Java class and has a name, a knowledge base (KB) and a pointer to

- a World class (the environment at which it belongs),
- its possible teacher, student and interlocutor agents,
- the strategies class,
- the fonctional roles (FR) class.

The KB is made of two types of simple objects: facts (a predicate name and a term) and implications (two predicate names, two variables and a direction). An exchanged message is one entry of the KB (a fact or an implication) plus a FR type. Strategies are methods defining sequential actions to perform in order to accomplish the specific asked strategy. They can be used directly by the agent (for lesson teaching for example) or they can by called by their FR component (for predicate explanation implied by a "say-dissatisfaction" for example). The FR component is a switch that routes the agent message to the good FR method according to its FR type. Each FR method uses the adequate strategies to satisfy agents.

Running the process shows that the 'student' has detected a conflict between its KB and a new data provided by the 'teacher'. He then asks the 'teacher' to validate some potentially conflictual knowledge and finally removes the wrong implication[2].

5 CONCLUSION

Our system allows artificial cognitive agents, in a tutored learning situation between a 'teacher' and a 'student', to acquire new knowledge through sole dialogue. The study of such a constraint has led us to define a notion of *connexity* for a knowledge base (KB), allowing to assess the connection level between each element of knowledge of an agent and so to give it a new goal: increasing its KB connexity. As the dialogue situation in highly impredictable and may follow no previous plan, we have adopted the functional role theory to easily model dialogical exchanges. Agents use strategies to learn new knowledge and solve conflicts between external and internal data. (Ang 88) tackles the problem of identifying an unknown subset of hypothesis among a set by queries to an oracle. Our work differs mainly in the communication mean: we use imbricated dialogues instead of queries; and in the learning's aim: our agents aim at learning new formulas and increasing their KB connexity instead of identifying hypothesis.

This work is a first approach in learning by dialogue for cognitive artificial agents. Its aim is to define a set of requirements for an advanced communication. Some paths could be explored like enriching the KB content by new formula types or defining new dialogue strategies. Last, this type of learning could be used in complement with others that rely on interaction with environment, in order to multiply knowledge sources.

REFERENCES

D. Angluin: *Queries and Concept Learning*, Machine Learning, Volume 2, Issue 4, ISSN:0885-6125, Kluwer Academic Publishers, Hingham, MA, USA, 319-342, 1988.

H. Asoh, Y. Motomura, I. Hara, S. Akaho, S. Hayamizu, T. Matsui: *Acquiring a probabilistic map with dialogue-based learning*, In Hexmoor, H. and Meeden, L., editors, ROBOLEARN '96: An Intern. Workshop on Learning for Autonomous Robots, 11-18, 1996.

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

J.L. Austin: *How to Do Things with Words*, ed. J. O. Urmson and Marina Sbis, Cambridge, Mass: Harvard University Press, 1975.

[2]Implementation details are reachable at $http://www.lirmm.fr/\sim yousfi$

M. Baker: *A Model for Negotiation in Teaching-Learning Dialogues*, Journal of Artificial Intelligence in Education, 5, (2), 199-254, 1994.

P. Cohen, H. Levesque: *Intentions in Communication*, Rational Interaction as the Basis for Communication. Bradford Books, MIT Press, seconde dition, chap. 12, 1992.

J. Cook: *Cooperative Problem-Seeking Dialogues in Learning*, In G. Gauthier, C. Frasson and K. VanLehn (Eds.) Intelligent Tutoring Systems: 5th International Conference, ITS 2000 Montreal, Canada, June 2000 Proceedings, Berlin Heidelberg New York: Springer-Verlag, 615-624, 2000.

S.W. Draper, A. Anderson: *The Significance of Dialogue in Learning and Observing Learning*, Computers and Education, 17(1), 93-107, 1991.

J.R. Josephson, S.G. Josephson: *Abductive Inference, Computation, Philosophy, Technology*, New York, Cambridge University Press, 1994.

Z. Manna: *Mathematical Theory of Computation*, International 'student' Edition. McGraw Hill Computer Science Series, 1974.

M. Mataric: *Using Communication to Reduce Locality in Multi-Robot Learning*, AAAI-97, Menlo Park, CA: AAAI Press, 643-648, 1997.

M. Pagnucco: *The Role of Abductive Reasoning Within the Process of Belief Revision*, Ph.D. thesis, University of Sydney, 1996.

M. E. Pollack: *Plan Generation, Plan Management, and the Design of Computational Agents* Proceedings of the 3rd International Conference on Multi-Agent Systems, Paris, France, July, 1998.

A. Ravenscroft, R.M. Pilkington: *Investigation by Design: Developing Dialogue Models to Support Reasoning and Conceptual Change*, International Journal of Artificial Intelligence in Education, 11, (1), 273-298, 2000.

J. Searle: *Speech Acts: An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press, 1969.

G. Sabah, O. Ferret, V. Prince, A. Vilnat, S. Vosniadou, A. Dimitrakopoulo, E. Papademetriou: *What Dialogue Analysis Can Tell About Teacher Strategies Related to Representational Change*, in Modelling Changes in Understanding: Case Studies in Physical Reasoning, Elsevier, Oxford., 1998.