The cognitive coherence approach for agent communication pragmatic

Philippe Pasquier[*]
Laboratoire DAMAS
Université Laval, Département d'informatique
G1K 7P4, Québec, Canada
pasquier@iad.ift.ulaval.ca

Brahim Chaib-draa
Laboratoire DAMAS
Université Laval, Département d'informatique
G1K 7P4, Québec, Canada
chaib@iad.ift.ulaval.ca

ABSTRACT

Different approaches have investigated the syntax and semantic of agent communication. However, all these approaches (including : agent communication languages (ACLs), conversation policies and dialogue games) have not indicated how agents should dynamically use communications. In fact, most of these approaches have mainly focused on "structure" of dialogues although developers are more interested in agents' capabilities of having "useful" conversations in respect to their goals rather than in their abilities of structuring dialogues. This leads us to propose a theory of use of conversations between agents. This pragmatic theory extends and adapts the cognitive dissonance theory (a major theory of social psychology) to multi-agent systems. In this paper, we show how this theory allows us to provide generic conceptual tools for the automation of both agent communicational behavior and attitude change processes. The cognitive coherence that we propose is formulated in terms of constraints and elements of cognition and allows us to define cognitive incoherences and dialogue utility measures. We show how these measures could be used to solve common problems and answer some critical questions concerning agent communication frameworks use. Finally, the theory is illustrated with an example of dialogue games automatic use.

Categories and Subject Descriptors

I.2.0 [**Artificial Intelligence**]: General—*Philosophical foundations*; I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multi-agent systems*

Keywords

pragmatic, coherence, dissonance, dialogue utility

[*]Ph.D. student

1. INTRODUCTION

Agents and multi-agents technologies allow the conception and development of complex applications. In the current distributed data processing paradigm, the fundamental characteristic of these systems is the agents skill to communicate among them in a useful way regarding to their individual and collective goals. If numerous works aimed to define agents communication languages (noted ACL for Agent Communication Language hereafter), few concentrated on their dynamic and automatic use by agents. This last task is left to the system designers, which specifies manually, by means of rules, the agent communicational behavior. In this paper, we introduce a theoretical framework for the automation of this behavior as a first step to fill this gap.

After having explained our problematic in detail (section 2), we present a new approach for agent communications pragmatic, the cognitive coherence theory. This conceptual framework is based on the cognitive dissonance theory which is one of main motivational theories in social psychology (section 3). Then, we indicate how this approach supplied generic answers to automate different aspects of conversations in multi-agents systems (section 4). Finally, we give an example using dialogue games as an interactional framework to illustrate our computational theory (section 5).

2. MOTIVATIONS

2.1 Structural versus cognitive coherence

Among the theories of communication, one distinguishes *cognitive theories* from *interactional theories* [20]. Interactional theories articulate around the notion of *structural coherence* (often called conversational coherence [7]) and deal with the shape of communication: What are the dialogue units ? What are the structural regularities of conversations or which are the conventional aspects constraining the dialogue shape/structure ?

On the other hand, cognitive theories deal with message production as well as message perception and reception. Central questions are then: what to communicate, when to communicate it and to whom ? How to understand and interpret incoming messages and how to react ? Those theories articulate around the notion of *cognitive coherence* and address the functional aspect of communication at both internal and external level. Which are the elements which urge an agent to introduce such a type of dialogue rather than another one ? At the internal level, how does an agent cognitively react to a statement in terms of mental states update ? At the external, public level (toward the environment), which are the commitments the agent wants to obtain ? Why ? What is the conversation utility ? Is the agent or the group of conversing agents satisfied by the conversation ?

In those two approaches, coherence is a central notion. But it is necessary to guard against confusing the structural coherence of the dialogue - is it allowed to pursue the dialogue in this way ? - with its cognitive coherence. Does the message content suit in regard to previous messages and agents mental states ? Is the message content coherent with the agent's internal state ? Do agents hold relevant dialogues in regard to their goals ? Do agents take advantage of the conversation ? One should make a difference between the respect of dialogue structural constraints (for example, to satisfy a dialogue game by respecting its rules) and the agents cognitive satisfaction. Even if these two coherence dimensions are different, they are often connected and working on a cognitive theory does not allow denying the need for an interactional theory. Indeed, when one has determined what to say, when and to whom, how conversation can take place remains to be known. On the other hand, working on cognitive coherence allows us to exceed this level and ideas advanced in the following sections are valid for any rich enough interactional and conventional communication framework.

2.2 Interactional frameworks in MAS

Regarding communication, the MAS community has been concentrating for some years on building a standard interactional framework. Main current ACLs, KQML [12] and FIPA-ACL [15], are both based on speech acts theory. Semantics of messages is formulated in terms of mental states, private aspects of agents [13]. Dialogue is supposed to emerge from the chaining of produced speech acts stemming from agents' intentions by way of recognition and reasoning on others' intentions. This "mentalistic" approach has been criticized [22, 25]. It raises the semantic verification problem: agents should be able to verify that the others act according to held dialogues[1]. For messages semantics to be verifiable, it would be necessary to have access to agents' private mental states which is generally not possible. A second major problem raised by this formulation is the sincerity assumption. This hypothesis, necessary for the definition of ACL mentalistic semantic, is considered too restrictive by the MAS community. It forbids certain dialogue types in domains where such an hypothesis would not hold, as it is the case for negotiation dialogues in electronic business [10].

More recently, some authors have proposed social approaches for agent communication introducing a public conventional layer expressed in term of social commitments [26, 6, 14]. These approaches allow (1) resolving the semantic verification problem,(2) getting rid of the sincerity hypothesis and (3) facilitating the treatment of the social aspects of communication. Among these approaches which we shall consider as conventional, dialogues games [24, 8, 21] seem to appear as a good alternative between strictly "mentalistic" or "social" approaches which do not specify anything about dialogue structure (which is supposed to emerge) and protocols which reduce the searching space for possible continuations to its strict minimum, causing the loss of the flexibility and adaptability of conversations. For those reasons, we retain dialogue games as our interactional framework.

2.3 Problem and objectives

The interactional tools proposed by conventional approaches such as dialogue games do not include the necessary elements for their automatic use by cognitive agents. The introduction of the commitments public layer requires rethinking a pragmatic theory, i.e. a theory of use, widened to these new frameworks. Indeed, agents do not directly have to reason anymore about others' private intentions but rather on taken and to be taken social commitments. These commitments are those obtained by the agent or by the other agents and stemming from held conversations as well as those imposed by the interactional framework constraints or by the system conventions. Besides, conventional interactional frameworks do not supply any guarantee about the utility of held conversations. Nevertheless, we are more interested in agents' capabilities of having "useful" conversations in respect to their individual and collective goals rather than in their abilities of structuring dialogues. Our objective is twofold: (1) to complete the dialogue games approach by defining a cognitive pragmatic theory suited to the public layer treatment and (2) to introduce a metric to allow the agents to consider the utility of held conversations as well as to guide them in the choice of conversations to hold. The following sections present our contributions concerning these objectives.

3. THE COGNITIVE COHERENCE FRAMEWORK

In cognitive sciences, cognitions gather all cognitive elements: perceptions, propositional attitudes such as beliefs, desires and intentions, feelings and emotional constituents as well as social commitments. From the set of all cognitions result attitudes which are positive or negative psychological dispositions towards a concrete or abstract object or behavior. All attitudes theories, also called cognitive coherence theories appeal to the concept of homeostasis, i.e. the human faculty to maintain or restore some physiological or psychological constants despite the outside environment variations. All these theories share as a premise the *coherence principle* which puts coherence as the main organizing mechanism: *the individual is more satisfied with coherence than with incoherence.* The individual forms an opened system whose purpose is to maintain coherence as much as possible (one also speaks about balance or about equilibrium). Attitude changes result from this principle in incoherence cases.

The cognitive dissonance theory, initially presented in 1957 by Festinger [11] is one of the most important theories of social psychology. It generated hundreds of studies and extrapolations on human attitudes, behaviors, beliefs, values, decision-taking consequences, inter-personal discords and others important psychological phenomena [17]. This is partially explained by the very general and abstract formulation of this theory which makes it easy to manipulate. In communication theories [20] it appears as one of the main cognitive theories for messages reception and treat-

[1] This semantics verification should not be mistaken with the formal semantics checking: agents are implemented in accordance with the ACL mathematical or logical semantics.

ment. Numerous formalizations and models of cognitive dissonance were produced [18]. Let propose ours which is explicitly adapted for AI and MAS.

Our formulation is inspired by the coherence theory of the computational philosopher Thagard [28] which allows us to directly link the cognitive dissonance theory with notions, common in AI and MAS, of elements and constraints. In our theory, elements are both private and public agent's cognitions: beliefs, desires, intentions, social commitments. Elements are divided in two sets: the set A of accepted elements (which are interpreted as true, activated or valid according to the elements type) and the set R of rejected elements (which are interpreted as false, inactivated or not valid according to the type of elements). Every non-explicitly accepted element is rejected. Two types of non-ordered binary constraints on these elements are inferred from the pre-existing relations that hold between them in the agent's cognitive model:

- *Positive constraints*: positive constraints are inferred from coherence or consonance relations which can be: explanation relations, deduction relations, facilitation relations and all other positive associations considered.
- *Negative constraints*: negative constraints are inferred from incoherence or dissonance relations: mutual exclusion, incompatibility, inconsistent and all the negative relations considered.

For each of these constraints a weight reflecting the importance and validity degree for the underlying relation is attributed. These constraints can be satisfied or not: a positive constraint is satisfied if and only if the two elements that it binds are both accepted or both rejected. On the contrary, a negative constraint is satisfied if and only if one of the two elements that it binds is accepted and the other one rejected. So, two elements are said to be *coherent* if they are connected by a relation to which a satisfied constraint corresponds. And conversely, two elements are said to be *incoherent* if and only if they are connected by a relation to which a non-satisfied constraint corresponds. Given an elements partition among A and R, one can measure the *coherence degree* of a single element by calculating the sum of the satisfied incoming constraints weights divided by the number of concerned constraints. And symmetrically, one can measure the *incoherence degree* of a single element as the sum of the weights of unsatisfied constraints divided by the number of concerned constraints. In the same way, one can measure the coherence degree of a set of elements by adding the weights of constraints connected to this set (the constraints of which at least a pole is an element of the considered set) which are satisfied divided by the total number of concerned constraints. Symmetrically, the incoherence of a set of cognitions can be measured by adding the weights of non-satisfied constraints concerned with this set and dividing by the total number of concerned constraints.

In this frame, the basic hypothesis of the cognitive dissonance theory is that incoherence (what Festinger names dissonance) produces for the agent a tension which incites him to change. The more intense the incoherence, the stronger are the insatisfaction and the motivation to reduce it. A cognition incoherence degree can be reduced by: (1) abolishing or reducing the importance of incoherent cognitions (2) adding or increasing the importance of coherent cognitions.

Festinger's second hypothesis is that in case of incoherence, the individual is not only going to change his cognitions or to try to change those of the others to try to reduce it, he is also going to avoid all the situations which risk increasing it. Those two hypotheses were verified by a large amount of cognitive and social psychology studies and experiences [31].

One of the major interests of the cognitive dissonance theory captured by our formulation is to supply incoherence measures, i.e. a metric for cognitive coherence. These measures match exactly the dissonance intensity measures first defined by Festinger. One can wonder in which circumstances incoherence arises. In fact, there are various situations in which incoherence can appear:

- *Initial direct contact with a situation*: a new situation can introduce new elements incoherent with preexisting cognitions;
- *A change in the situation*: a change in the situation can lead coherent cognitions to become incoherent;
- *Communication*: communication with others can introduce cognition elements which are incoherent with those of the agent;
- *Simultaneous existence of various cognitions*: in the general case, a cognition is connected with several others among which some are coherent and others dissonant.

3.1 Incoherence, social influence and attitude change

We link private and public cognitions with the following:

- According to practical reasoning, private cognitions finally end in intentions and we make the classical distinction between *intention to* (do something or make someone doing something) and *intention that* (a proposition holds) [3];
- Regarding public cognitions, we distinguish *commitments in action* from *propositional commitments* [30];
- A commitment is the socially accepted counterpart of an intention, commitments in action are the counterparts of "intentions of" and propositional commitments are the counterparts of "intentions that".

Those relations are not completely new since many authors have already considered individual intentions as a special kind of commitment [3, 29].

In MAS, knowing when an agent should try to modify the environment (the public social commitments layer, among others) to satisfy his intentions, and when the agent has to modify his mental states to be coherent with his environment is a crucial question. In our model, *any agent tries to maximize his coherence*, i.e. tries to reduce his incoherences beginning with the most intense one. To reduce an incoherence, the agent has to accept or reject cognitions to better satisfy the constraints which connect them. These cognitions can be private or public. But all the cognitions are not equally modifiable. This is what Festinger names the resistance to change of cognitions. The resistance to change of a cognition is a function of the number and the importance of the elements with which it is coherent, also depending on its type, age as well as the way by which it was acquired: perception, reasoning, communication. To be able to integrate communication into our model, it is now necessary to introduce the fundamental link which exists between our formulation of the cognitive dissonance theory and the notion of social commitment.

Social commitments are particular cognitions which are not individually modifiable but must be socially established and dialogue games are tools for attempting to establish collectively accepted commitments. That is, in order to modify, reject or accept a social commitment an agent has to have a dialogue. Dialogues are the only means for agents to try to establish social commitments coherent with their private cognitions. However, after those dialogues, some commitments can remain incoherent without being modifiable anymore. They are then social obligations and fixe one of the poles of the constraints which are connected to them. To reduce possible incoherence while conforming to taken commitments, agents should then change their private cognitions to restore the coherence. This is the spring of the attitude change in our system and it formalizes the vision of the psychologists Brehm and Cohen on this subject [4], supported by a great number of experiments. An example of this attitude change mechanism is supplied in section 5.

4. AGENT COMMUNICATION AS COHERENCE SEEKING

4.1 Incoherence Typology

This section presents a typology of incoherences which aims to introduce a simple but useful vocabulary to handle coherence problems in the explicitly distributed frame of MAS. Incoherence being conceptually close to the notion of conflict, the following typology is borrowed from works on conflicts [23]:

- *Internal and external incoherences*: an incoherence is internal when all the involved cognitions are relative to the same agent and external when incoherent cognitions involve at least two agents. More concretely, an incoherence is external for an agent if it is an incoherence between his cognitions and some others or social cognitions. *Shared internal incoherence* is a special case arising when several agents have the common knowledge that they experience the same internal incoherence.

- *Explicit and implicit dissonances*: we define explicit by the fact of being in "the state of having knowledge of", and implicit by that of being in "the state of not having knowledge of " [2]. An incoherence is explicit for an agent if all the involved cognitions are explicit for that agent. A dissonance is implicit for an agent if at least one of the incoherent cognitions is implicit for him. An implicit incoherence is a potential explicit incoherence. Notice that in MAS internal incoherence will be doubtless always explicit since we do not consider any implicit internal level.

4.2 Link coherence - initiative, topic and relevance

In AI, dialogue initiative usually raises particularly delicate problems. When should an agent initiate a dialogue and why ? The answer supplied by our coherence frame is that an agent takes the dialogue initiative if he experiences an incoherence he can not reduce alone. Either because he knows that it is an external incoherence which involve other agents, or because it is an internal incoherence he has no capacities to reduce alone, he has then to count on the other agents cooperation. Among the potentially multiple incoherent elements, the agent will choose the most incoherent one as the conversation initial subject (topic).

With the relevance theory, Sperber and Wilson [27] advanced the idea that the speaker chooses what he is going to say by dynamically estimating the relevance of his ideas. Every cognition element relevance varies during the conversation. The speaker undertakes a speech act only when its relevance is maximal. With our approach, an agent who takes initiative is going to attack the incoherence which has the biggest magnitude, because it is the most cognitively relevant choice for him. The following section indicates how the coherence frame allows agents to choose which type of dialogue to engage.

4.3 Link with dialogue types

In this section, we analyze how dialogue types observed in dialectic can be bound to cognitive coherence. Some recent works use a dialogue typology due to Walton and Krabbe [30]. These authors distinguish six dialogue types defined by their first purpose (to which interlocutors subscribe) and appropriate private goals of each agent (which can be incompatible, i.e. incoherent):

1. *Persuasion*: the initial situation is an external incoherence of point of view and the global purpose is to resolve it. Every participant tries not to change his private cognitions (according to their resistance to change) and to change those of the others. To do this, agents typically resort to argumentation [19]. Thus, it is an external incoherence reduction technique.

2. *Negotiation*: starting from a conflict of interest (a type of external incoherence), the global purpose is to conclude a contract, to come to an agreement. Every agent has his own purpose and wants to maximize his profit or interests. The conflict resolution is usually made by an exchange of offers and of counter offers. It is frequent that dialogues of persuasion are nested in a negotiation, offers being thus argued. It is a technique of external incoherence reduction.

3. *Inquiry*: participants of this type of dialogues are in an initial situation of shared internal incoherence. They all suffer the same internal incoherence and they want to inquire together to increase the efficiency of the reduction. Common purpose coincides with individual purposes. It is a shared internal incoherence reduction technique.

4. *Deliberation*: Each agent has his own preferences and all agents have to choose together among the offers of each one. The participants have as a common purpose to take a decision (to choose a plan or an action). Their individual purpose is to influence decision in their interest (which could match public interest). It is a reduction technique for explicit external incoherence.

[2]One can have knowledge of something without being in "the state of having knowledge of" as it is the case with forgetfulness. For example, one can have the knowledge that for driving by night, it is necessary to turn on the lights, but it can happen that one forgets.

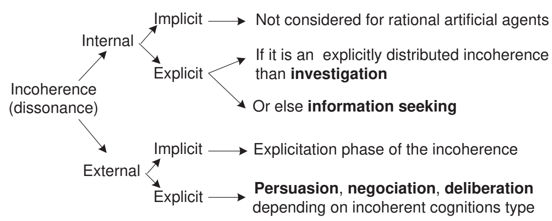

Figure 1: Typology of cognitive dissonances and link with dialogue types.

5. *Information seeking*: It is the only dialogue type which is always asymmetrical. An agent tries to obtain information from others. It is a technique of internal incoherence reduction. With this dialogue type only the information applicant agent is in an incoherent state. The reduction is asymmetrical but to facilitate it, it is frequent that the applicant clarifies his incoherence to the other agents by indicating why he looks for such information, i.e. by making his incoherence explicit to others. This reduction can be made through a dialogue, but it can also take form in other actions (for example: reading a reference book, searching the internet) as far as incoherence is reduced that is information found.

6. *Eristic*: Eristic dialogue has an highly conflicting and incoherent initial situation. Contrary to the other dialogue types it generally involves feelings and emotions more than reason rationality and coherence. This is why we will not detail it here.

As one can notice, all dialogue types arise from an incoherent initial situation. Notice that the reverse is false, i.e. all incoherences are not handled by dialogue. Our point is that two agents communicate if an incoherence forces them to do so. From then on, conversation might be seen, on top of its already known characteristics, as a generic procedure for attempting to reduce incoherence. Due to the conceptual nearness between dissonance, incoherence and conflict notions, this hypothesis is close to the classic position of dialectic: any dialogue arises from a conflict [16]. For some years, several authors have been insisting again on the role of conflicts in the communicative process:

- For Walton and Krabbe [30], the question "is there a conflict ?" is the base of their dialogues initial situations analysis;
- For Dessalles [9], "a great number of dialogues find their origins in cognitive conflicts between desires or beliefs";
- For Baker [1], "dialogues result from an opposition between conflicting goals".

Finally, figure 1 summarizes the various incoherence types as well as dialogue types which could be used to reduce them. The example in section 5 reconsiders this dialogue type choice in the more specific and formal framework of dialogue games.

4.4 Link coherence - utility and dialogue dynamics

4.4.1 Dialogue utility

Mathematical utility is generally defined in terms of degrees of satisfaction and usually takes form in a difference of values (which should be economical or system dependant) taken before and after some actions. In the cognitive coherence theory, according to the coherence principle (section 3) satisfaction is equivalent to coherence. So action utility can be computed as coherence measures differences. More precisely, expected utility for a conversation is equal to the difference between the intensity degree of the incoherence which dialogue attacks and the expected incoherence degree after this dialogue if this one made a success in favor of the agent. Agents can also calculate the utility of a conversation dynamically by working out again the incoherence degrees during the dialogue. When a dialogue unity is ended, either incoherence is reduced and the dialogue ends, or the agent can keep on trying to reduce it (as long as possible, because there must be dialogues that cannot be questioned when they failed, or else the system would not finish). Before presenting a detailed example using those utility measures, let's see why they are useful for dialogue control automation.

4.4.2 Intra-dialogue dynamic

An agent selects a dialogue type according to the incoherence type which he wishes to reduce, i.e. the type of problem he wishes to settle. But during this resolution, other incoherences can appear, these sometimes can have to be reduced so that the main reduction can continue. This is what brings agents to embed sub-dialogues to reduce new incoherences before resuming the main dialogue concerning the initial incoherence. Most dialogue games interactional frameworks provide syntactic facilities to do so. In other cases incoherence can move, leading the conversing agents to chain two dialogue units. Within the cognitive coherence theory, the dialogue structuration is determined by the incoherence reduction chaining.

4.4.3 Inter-dialogue dynamic

Dialogues are attempts to reduce incoherence which can fail. The dialogue utility measure defined within the coherence framework is useful to guide the agent in his communicational behavior. Following a useless or not very useful dialogue, i.e. incoherence is not reduced, the agent has to decide how to act. The agent will probably persevere in his reduction attempt by taking into account this failure: if it is still possible, he will propose a different dialogue type or a different proposition of the same dialogue type or else he will update his mental states as described in section 3.1. But in all cases he should take note of this failure which can be useful to guide him for the following dialogues, i.e. reduction attempts.

In particular, in open and heterogenous MAS, an agent is led to communicate with unknown agents, it is then necessary for him to form an idea of dialogues held with these. The agent will be able to take into account the utility of previously held dialogues to select these interlocutors. It will be in the agent's interest to strengthen exchanges with agents with whom dialogues are useful and numerous incoherences (i.e. problems) are resolved and on the contrary, he will be able to take into account useless dialogues by weakening his social links with the involved interlocutors. Dialogue utility measures supply precious information which can be used by a social relations management tool.

4.5 Link coherence - mood, intensity

Recently, needs to integrate emotions into artificial agents have appeared [2]. The cognitive coherence model allows making a direct link between coherence measures and agent's mood. Our theory supplies a value system in which a coherence state is a comfortable state and coherence gains are satisfactions and reassurance who could lead to happiness, smile... On the contrary, an incoherent state is a discomfort state and agent can be worried or afraid of potentially future incoherences or disappointed by a failed attempt of reduction...

Besides, some interactional frameworks allow using various intensity degrees for speech/dialogue acts illocutionary forces. Nevertheless, no agent theory indicates how this selection should be made. Quantitative measures defined by the cognitive coherence theory supply means to guide the agent in the choice of the suitable intensity degree.

Since a conversation is engaged in as an attempt to reduce an incoherence, its magnitude gives the importance of the resulting conversation. This incoherence intensity influences the choice of intensity degrees of used speech acts in a direct way. For example, an agent who needs information to reduce an internal incoherence is going to enter an information seeking dialogue which include directive act(s). The intensity degree of the illocutionary force is then going to depend on the intensity of the aforementioned incoherence: (1) an invitation, an advice if the incoherence is very light,(2) a recommendation, a demand if it is a little more intense and (3) a plea, an order or an entreaty if the incoherence magnitude is very high and its reduction crucial.

If these parameters of emotions, mood and dialogue acts intensity seem less important for completely artificial MAS, this track is interesting for human machines interfaces and intelligent tutorial systems, among others. Obviously, this intensity degree selection factor is not unique. Other factors can intervene in this choice: social agreements (it is generally forbidden to give an order to a superior in the hierarchy), relations among agents (nearness, confidence, trust, past of the relation) are also important for selecting those intensity degrees.

5. DETAILED EXAMPLE

The theory presented in the previous sections must be envisaged as a new layer above the existing agents architectures. Its integration will be realized by reformulation of the networks of cognitions in terms of elements and constraints so that the various coherence and utility measures defined above can apply.

For our example, we suppose that in the considered system, DIAGAL dialogue games language defined by Chaib-draa and Maudet [5] is used as the interactional framework. In DIAGAL, accepted commitments are not modifiable, they constitute social obligations and bring penalties if they are not respected[3]. An accepted commitment of x to y on the proposition p (or action α) is noted: $C(x, y, p)$. A rejected commitment is noted: $\neg C(x, y, p)$. In DIAGAL, four dialogue games are defined: $offer$, $request$, $inform$, ask. They respectively aim to lead to the acceptance of: a commitment in action from the initiator to the partner, a commitment in action of the partner to the initiator, a propositional commitment from the initiator to the partner and a propositional commitment of the partner to the initiator. The games compositions possibilities (sequencing, choice, pre-sequencing and embedding) are made possible by way of a special control meta-game that Maudet and Chaib-draa named contextualisation game and which allows proposing the entry or the exit of a game at any moment.

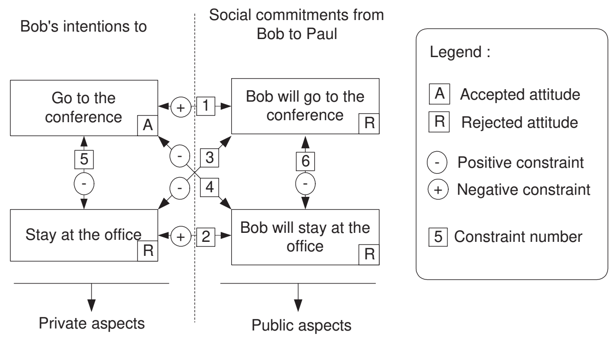

Figure 2: Network of constraints in the initial situation for Bob.

In our example, an agent Bob works to plan his next day. He has choice between two mutually exclusive intentions to: (1) go to the conference or (2) stay at the office. Two social commitments in action can correspond at the social level to these two intentions to (we established those links in section 3.1): the commitment of Bob to his boss Paul to go to the conference and the commitment of Bob to Paul to stay at the office. These commitments are connected to the corresponding intentions by correspondence relations and are connected among them by a relation of mutual exclusivity. Besides, the intention to stay at the office and the commitment to go to the conference as well as the intention to go to the conference and the commitment to stay at the office are connected by incompatibility relations.

One can infer constraints corresponding for all these relations (as indicated in section 3). Positive constraints came from correspondance relations and negative constraints came from mutual exclusion and incompatibility relations. Besides, Bob initially has the intention to go to the conference, an intention that he deduced from his beliefs and desires like in classic BDI systems. So, this intention belongs to the set of accepted elements A and all the other elements are in the set of rejected elements R (which means that Bob is not yet committed either to go to the conference or to stay at the office and that he does not intend to stay at the office). Figure 2 shows this initial constraints network.

By allocating a unitarian weight to each constraint defined above, one can calculate the network coherence by adding weights of satisfied constraints and divide the result by the total number of constraints. We obtain a 3/6 coherence (numbered constraints 5 , 3 and 2 are satisfied). We can calculate the coherence of each element of this network as well. Bob's intention to go to the conference has a coherence of 2/3 (among the three constraints concerned to this cognition, only the numbered 4 and 5 are satisfied). The rejected intention to stay at the office has a coherence of 2/3. The rejected commitment to go to the conference, has a coherence of 0/3 and the rejected commitment to stay at the office a coherence of 2/3.

[3]We let the penalties aside for the sake of simplicity.

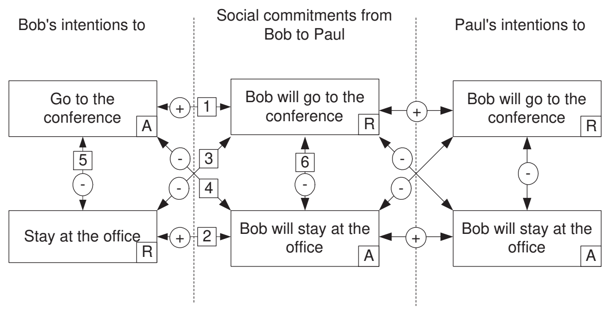

Figure 3: Network of constraints in the final situation for Bob and Paul.

In order to maximize the coherence, Bob should then modify the state of acceptance of the least coherent element (that is to try to reduce the local incoherence of the most important magnitude). Thus it would be a question for him of succeeding in accepting the commitment to go to the conference. As it is a social commitment which must be socially accepted, Bob cannot reduce this explicit external dissonance solely. Since this incoherent commitment implies his boss Paul, Bob is therefore going to begin a dialogue with Paul about his intention to go to the conference and more precisely to attempt to get the corresponding commitment accepted. One can calculate the expected utility of the various dialogues games he can use. If Bob tries and succeeds in negotiating with his boss the commitment to stay at the office, the coherence of the network will be 2/6. If, he tries and succeeds to commit to go to the conference, coherence will be 6/6 and finally, if the agent Bob keeps silent it will stay 3/6.

Among the various available dialogues games, Bob is going to choose the one that allows a commitment in action of the initiator to his interlocutor to be accepted. This dialogue game should have an entry condition which can unify with $\neg C(bob, paul, go\_to\_the\_conference)$ meaning that considered commitment is not already established and a success condition which can unify with $C(bob, paul, go\_to\_the\_conference)$ meaning that in case of success of the dialogue, the corresponding commitment will be accepted. The only game which satisfies these constraints is the DIAGAL $offer$ game. So Bob will suggest entering a game of this type to Paul. Under the dialogical cooperation principle (meaning that agents agreed at least to communicate), one supposes that Paul agrees to enter this game. The $offer$ game is thus going to be played and Bob is going to respect his commitment to make an offer for Paul, by suggesting to him the commitment to go to the conference. According to the $offer$ game rules, it is then for Paul to react.

Figure 3 presents Paul's private cognitions and the related constraints. To avoid losing coherence, Paul should refuse Bob's offer. If this commitment was accepted, Paul's coherence would crosses 3/6 in 1/6 resulting in a negative dialogue utility for him as one can calculate it (notice that no commitment is accepted yet). Furthermore, to maximize his coherence, Paul has better things to do before. Indeed, the most dissonant element for Paul is the rejected commitment from Bob to him to stay at the office (coherence of 0/3). To try making it accepted, Paul will open an embedded dialogue by proposing the game which allows establishing a type of commitment unifiable with this last one that is the DIAGAL *request* game. Bob agrees to play this sub-game. Paul asks then Bob to accept the commitment $C(Bob, Paul, stay\_at\_the\_office)$. The fact that Paul is Bob's superior forbids Bob to refuse (without which, one would have entered an argumentation and our example would have to be much bigger). The success condition of the *request* game is fulfilled and its ending is proposed by Paul. Bob accepts. We are then in the position indicated in figure 3. Paul, whose coherence is now maximal, that is 6/6, does not wish to change anything else. He then refuses the offer stemming from the initial $offer$ game opened by Bob and they close it.

The coherence of Bob, finally fell by this dialogue to 2/6. To restore this coherence, it is not possible anymore for Bob to try to negotiate acceptance or rejection of commitments to Paul, because they have already been discussed and one can not return above accepted commitments (with DIAGAL). So, the only solution for Bob, allowing him to restore a 6/6 coherence is to change its private cognitions. Bob is therefore going to reject his first intention to go to the conference and to accept the intention to stay at the office. This is the attitude change process described in section 3.1. As long as an agent can change the outside world to maximize his coherence, he tries and since the outside world is not modifiable any more, he changes his own private cognitions to conform to it.

Finally, our two agents held a conversation to which could correspond the following human dialogue (however, human dialogue games grounding moves are usually more implicit):

> *Bob: I had thought of something for tomorrow...* (Proposition to enter an $offer$ game)
>
> *Paul: yes, go ahead.* (Proposition acceptance)
>
> *Bob: I would like to go to the conference, is it possible ?* (Offer, a kind of commissive act, conditional to the hearer's acceptance)
>
> *Paul: Actually, I had a request for you on that matter.* (Proposition to embed a *request* game)
>
> *Bob: Really ?* (Acceptance of the proposition)
>
> *Paul: I request you to stay at the office tomorrow.*(Request)
>
> *Bob: OK, (long sigh).*(Acceptance of the request and closing of the *request* game)
>
> *Paul: This mean that you could not go to the conference.*(Refusal of the initial offer, closing of the $offer$ game and end of the dialogue.)

6. CONCLUSION

In this article, we presented an agent communication pragmatic theory: the cognitive coherence theory. It is well-founded on cognitive sciences work (namely computational philosophy and social psychology) while using reasonings and calculations on elements and their associated constraints. Proposed as a new layer above classical cognitive agent architecture it supplies theoretical and practical elements for automating agent communication. The incoherence and utility measures defined within the cognitive coherence framework provide the necessary mechanisms to answer (even partially) the following questions which are usually poorly treated in the MAS literature:

- When should an agent takes a dialogue initiative, on which subject, with whom and why (section 4.2) ?

- By which type of dialogue (section 4.3) ?
- Which intensity to give to illocutionary forces of dialogue acts (section 4.5) ?
- How to define and measure the utility of a conversation (section 4.4.1) ?
- When to stop dialogue or if not how to pursue it (see section 4.4.2) ?
- What are the impacts of the dialogue on agents' attitudes (see section 3.1) ?
- What are the impacts of the dialogue on agents' mood (see section 4.5) ?
- What are the consequences of the dialogue on social relations between agents (see section 4.4.3) ?

Obviously, each of these problems could not be profoundly discussed here, but our purpose is rather to give an overview of our approach for agent communication pragmatic emphasizing its wide coverage. Because there is much more to say about it, future publications will be dedicated to more specific aspects of the cognitive coherence theory sketched here. Although those ideas already have been implemented for validation purposes using hybrid connectionist-symbolyc formalism, a richer prototype system will be presented.

7. REFERENCES

[1] M. Baker. *Knowledge Acquisition in Physics and Learning Environments*, chapter An Analysis of Cooperation and Conflict in students' collaborative explanations for Phenomena in Mechanics. Springer-Verlag: Heidelberg, Germany, 1991.

[2] J. Bates. The role of emotion in believable agents. *Communications of the ACM*, 37(7):122–125, July 1994. utilis.

[3] M. E. Bratman. What is intention? In P. R. Cohen, J. L. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 15–32. The MIT Press: Cambridge, MA, 1990.

[4] J. Brehm and A. Cohen. *Explorations in Cognitive Dissonance*. John Wiley and Sons, inc, 1962.

[5] B. Chaib-draa, N. Maudet, and M.A. Labrie. Request for action reconsidered as dialogue game based on commitments. In *Workshop on Agent Communication Language (AAMAS02)*, 2002.

[6] M. Colombeti. Commitment-based semantic for agent communication languages. In *1st Workshop on the History and Philosophy of Logic, Mathematics and Computation*, 2000.

[7] R.T. Craig. *Conversationnal coherence: form, Structure and strategy*. CA:Sage, Beverly Hills, 1983.

[8] M. Dastani, J. Hulstijn, and L.V. der Torre. Negociation protocols and dialogue games. In *proceedings of the BNAIC*, 2000.

[9] J.-L. Dessalles. *Formal Semantics and Pragmatics of Dialogue*, volume TWLT-13, chapter The Interplay of Desire and Necessity in Dialogue, pages 89–97. Enschede : University of Twente, 1998.

[10] F. Dignum. Issues in agent communication : An introduction. In F Dignum and M. Greaves, editors, *Issues in Agent Communication*, Lecture Notes in Artificial Intelligence, pages 1–16, 2000.

[11] L. Festinger. *A Theory of Cognitive Dissonance*. CA : Stanford University Press, 1957.

[12] T Finin and R. Fritzson. Specification of the KQML agent communication language plus example agent policies and architectures. Technical report, DARPA Knowledge Sharing Initiative External Interface Working Group, 1994.

[13] T. Finin and Y. Labrou. Semantics for an ACL. In *Agents, Architecture and Languages*, 4. Springer-Verlag, Berlin, 1998.

[14] R. Flores and R. Kremer. Bringing coherence to agent conversation. In M. Wooldridge, P. Ciancarini, and G. Weiss, editors, *Agent-Oriented Software Engineering II*, Lecture Notes in Computer Science, pages 50–67. Springer-Verlag, 2001.

[15] FIPA[Fondation for Intelligent Physical Agents]. FIPA ACL message structure specification. http://www.FIPA.org, 2002.

[16] C.L. Hamblin. *Fallacies*. Methuen, London, UK, 1970.

[17] E. Harmon-Jones and J Mills, editors. *Cognitive Dissonance : Progress on a Pivotal Theory in Social Psychology*. American Psychological Association, 1999.

[18] E. Harmon-Jones and J Mills, editors. *Cognitive Dissonance : Progress on a Pivotal Theory in Social Psychology*, chapter Mathematical Models of Dissonance, pages 235–297. American Psychological Association, 1999.

[19] J.O. Keefe. *Cognitive Dissonance*, chapter Persuasion : Theory and research, pages 61–78. Sage, Newbury Park, California, 1991.

[20] S.W. Littlejohn. *Theories of Human Communication*. Waldsworth Publishing Company, 1992.

[21] N. Maudet. *Modéliser les Conventions des Interactions Langagières : la Contribution des Jeux de Dialogue*. PhD thesis, ENSEIHT-UPS, 2001.

[22] B. Moulin. *Agent and Multi-agent Systems*, chapter The social dimension of interacions in multi-agent systems. LNAI 1441. Springer, Berlin, 1997.

[23] P. Pasquier and F. Dehais. Approche Générique du Conflit. In D.L. Scapin and E. Vergisson, editors, *ErgoIHM 2000*, Biarritz, France, 2000. ESTIA.

[24] C. Reed. Dialogue frames in agent communication. In *Proceedings of the Third Internationnal Conference on MultiAgent Systems (ICMAS)*, 1998.

[25] M. P. Singh. Agent communication languages: rethinking the principles. *IEEE Computer*, 12(31):40–47, 1998.

[26] M.P. Singh. *A social semantics for agent communication language*, pages 31–45. Springer-Verlag, 2000.

[27] D. Sperber and D. Wilson. *Relevance*. Harvard University Press, Cambridge MA, 1986.

[28] P. Thagard and K. Verbeurgt. Coherence as constraint satisfaction. *Cognitive Science*, 22:1–24, 1998.

[29] G.H. von Wright. *Freedom and determination*. North Holland Publishing Co., 1980.

[30] D. Walton and E. Krabbe. *Commitment in Dialogue*. Suny Press, 1995.

[31] R. Wickland and J. Brehm. *Perspectives on Cognitive Dissonance*. NY: Halsted Press, 1976.