COGNITIVA: A Context Independent Cognitive Architecture for Agents Combining Rational and Emotional Behaviours

RICARDO IMBERT, ANGÉLICA DE ANTONIO
Facultad de Informática
Universidad Politécnica de Madrid
Campus de Montegancedo, s/n – 28660 Boadilla del Monte (Madrid)
SPAIN

Abstract: - The traditional view of intelligent behaviours as products of pure rational processes fails when trying to explain most of human behaviours, in which emotion plays a key role. However, emotional factors add an extra complexity to agent architectures, making them, hitherto, either few efficient or few reusable.
This paper presents a context independent cognitive architecture for agents combining rational and emotional behaviours, named COGNITIVA, which bets on *adaptivity* as a weapon to fight against that complexity.

Key-Words: - Cognitive agent architecture, specification process, emotion, personality, virtual characters

1 The Role of Emotions in the Rational Process

1.1 Emotion and Reason, Oil and Water?

Emotions and reason have been traditionally considered as two sides of the same coin and, therefore, antagonistic and non combinable. Emotions are something rather irrational that plays down value to human rationality [1], something "non scientific" [2].

However, recent theories [3] [4] suggest that emotions are an essential part of human intelligence, playing a critical role in processes such as perception, learning, attention, memory, rational decision making and other abilities usually associated to intelligent behaviours.

The initial approach fails, probably, in considering "emotional systems" as systems that lose the desirable rationality and control. However, it is not right considering laws and rational norms as the unique and more important parts when interpreting human behaviour and intelligence. It is also an error considering human behaviour independent of any emotional process. Up to this point, it is worth to remark that, from the neurological perspective, *no polarization, or clean dividing line occurs between thinking and emotions* [2].

1.2 Emotional Architectures

Most of the theoretical models of emotion, coming from Psychology are not appropriate to be applied into computer systems, since they weren't conceived with that purpose. The adaptation of these approaches and the development of new theories, more suitable to the automation of their elements and processes, have reduced the number of theoretical models of emotion present in most of the emotional systems: appraisal models, motivational models, dimensional models…

However, neither these models nor the architectures and systems developed from them incorporate successfully emotions to the general rational process. Some deficiency or drawback is always imputed to each one, although, depending on the contexts and problems, they also prove sometimes to be acceptably adequate.

The empirical results of these approaches show that emotional factors cannot be considered as yet another component in the agent's architecture, but all the architecture must exhibit an emotional orientation.

1.3 Adaptivity vs. Specificity or Generality

Behind these architectures (usually agent-oriented) underlies a very intricate structure. Sometimes, their elements and dynamics are interwoven with the restrictions and particularities of the application context, mingling with them, making these approaches few reusable (cf. [5], [6], [7]).

Other times, architectures are intrinsically very generic, independent from any specific problem (cf. [8], [9], [10]). The lack of an adaptation to the particular needs of the problem originates less-efficient, computationally demanding mechanisms. This forces to reconsider the whole structure of the architecture and to simplify some of their original capabilities, in order to offer viable developments.

In our opinion, nowadays solutions do not provide the level of desired quality/satisfaction because they – just– fail in the "attitudes" with which complexity is faced: instead of looking for *specificity* or *generality*, the key is in *adaptivity*.

A new perspective is needed. A perspective considering the complexity of this kind of systems in an efficient, reusable way; a perspective allowing an adaptation to the specific problem and context needs without loosing the generic purpose of the architecture; a perspective maintaining a structure, components and processes coherent and understandable.

2 Adaptable Cognitive Architecture

This paper proposes the definition of a generic architecture, named COGNITIVA, to develop agents whose behaviours are emotionally influenced. Considering an agent as a continuous *perception-cognition-action* cycle, the scope of this architecture is circumscribed to its *cognitive* activity, although it does not restrict any of the other two modules (*perceptual* and *actuation*).

Different from the generic preceding architectures, COGNITIVA provides mechanisms to be adapted to each specific context and problem, from a double perspective:

- *Adaptation of its structure*: COGNITIVA is a multilayered architecture that covers several kinds of behaviour: reactive, deliberative and social. Besides, it includes a flexible model that allows establishing dependencies and influences among elements such as personality traits, attitudes, physical states, concerns and emotions. Even more, both the behaviour modes and the elements are configurable, according to the particular needs of every situation.
- *Adaptation of its process of application*: from the generic architecture, a progressive specification process has been designed, to apply it to every particular context. This process begins with a *functional specification* of the architecture, which provides a particular design and implementation of each one of the information structures and functions defined in the generic architecture.

  This first specification is yet too context independent to the application context. In fact, the same functional specification may be used as a basis for many different contexts. The approach to each one of them is made in a second specification phase, the *contextual specification*, in which all the particular values and procedures of the application environment are included.

In the following sections, COGNITIVA and its main components are described with a deeper detail, along with some results obtained from its application.

3 Description of COGNITIVA

Internally, COGNITIVA can be considered as a hybrid architecture, combining reactive, deliberative and social skills. Fig. 1 shows a schematic perspective of the three quasi-horizontal levels, namely:

- *Reactive layer*, to provide the agent with immediate responses to the perceived changes in the environment.
- *Deliberative layer*, to provide the agent with goal-directed behaviours, from its individual abilities point of view.
- *Social layer*, to provide the agent with behaviours in which the existence of other agents and the interaction with them is considered.
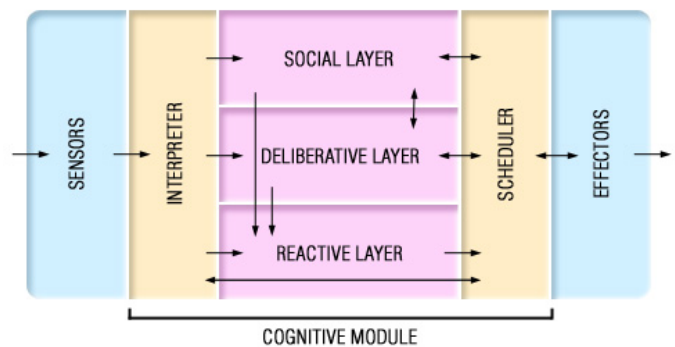

Fig. 1. General Schema of COGNITIVA.

3.1 Management of the Current State of the Agent: Beliefs

With independence of the decision making process carried out in each one of the layers, it is very difficult for an agent to exhibit coherent behaviours exclusively from the perceptual input coming from the sensors (perceptual module). It is necessary to consider other information sources, such as its knowledge about the environment, about other agents and, even more, about itself. All this information is represented internally as a *beliefs* set.

To manage the beliefs, a taxonomy has been defined. In a first level, the taxonomy differentiates the object of the belief: *places* –physical, conceptual or virtual–, *objects*, *individuals* and the *current situation*.

Besides, the agent beliefs related to places, objects and individuals are classified into:

- Beliefs related to *defining characteristics* (DCs), that describe the general traits of places, objects and individuals. DCs are fundamental to understand them. The DCs value hardly changes over the time.
- Beliefs related to *transitory states* (TSs), characteristics whose values represent the current state of places, objects and individu-

als. TSs' values have a much more dynamic nature, compared to DCs.

- Beliefs related to *attitudes*, useful to determine the behaviour of an agent towards other environment components (places, objects and individuals). Attitudes' values are less variable than TSs', but more than DCs'.

Among the whole set of agent's beliefs, COGNITIVA distinguishes a small subset related to the agent itself, which is fundamental in the architecture. This subset constitutes what is called the agent's **personal model**, and includes DCs such as its *personality traits*, whose values determine the coherent and stable behaviour of the individual; TSs such as its *moods* and its *physical states*, identifying the state of the mind and the body of the agent, respectively; and also its *attitudes* towards others.

Many of these characteristics are intrinsically related, and exert some influence on other beliefs of the personal model. So, for instance, the personality traits influence the value of the emotions.

3.2 Management of the Past State: History

Agent behaviours that do not take into account events occurred in past moments are specially disappointing for human observers. The architecture proposed considers two mechanisms to maintain the agent's past history information:

- *Accumulative management of the past*: this is an implicit mechanism, related to the way in which beliefs are managed. External changes in the environment or internal modifications in the agent internal state may produce an update of the agent beliefs. However, this update is performed as a variation ---of higher or lower intensity--- of the previous beliefs, avoiding abrupt alterations in the individual behaviour.
- *Explicit management of the past state:* an accumulative effect of the past events may not be enough to manage efficiently the past state, because it does not consider information related to the events themselves or to the temporal instant in which they took place. Our proposal is to maintain that information in a structure, internal to the agent, containing propositions related to any significant event that happened, the temporal instant of its occurrence and its importance.

With all this information, an agent will be able to select appropriate behaviours based on something more than perceptions and beliefs.

3.3 Agent Perceptions Interpretation

The cognitive module receives from the sensors *perceptions* of the environment. As it was said above, the cognitive layer manages the incoming perceptual information through a module we have called *interpreter*, which performs a triple function:

- Acts as interface between the sensors and the rest of the cognitive module, making it independent from the perceptual module. The interpreter transforms the perceptions coming from the sensors into *percepts*[1], understandable by the rest of the components and processes of the cognitive module.
- Filters the received perceptions, discarding those not interesting, in that moment, for the agent.
- Directs the percepts towards the appropriate components and processes of the architecture. Besides, participates in the updating of the agent's information structures and provides the rest of the agent's components with significant, updated and on-time information.

To manage efficiently the updating of beliefs and past history, COGNITIVA incorporates *expectations* inspired in the proposal of [12], which has been adapted, in turn, from the OCC Model [13]. Expectations capture the agent predisposition towards the events that take or can take place. In COGNITIVA, expectations are valued depending on:

- *Expectancy:* Expressing the agent's expectation for an event occurrence.
- *Desire:* Indicating the agent's desirability of an event occurrence.

From the expectations towards a given event, the confirmation or disconfirmation of its occurrence will produce a rich set of emotions.

3.4 Management of the Desirable State: Concerns

The reactive layer provides the agent with a fast-response mechanism to face up to situations restricted by response time. However, reactive does not necessarily mean *automatic* or, even worse, *out-of-control*.

If a person is able to overcome his fear and cross a hanging bridge, just to obtain some benefit arriving to the other end, an agent should be able to "control" its reactions, to avoid undesirable behaviours and conflicts with other layers intentions.

---

[1] Name proposed by Pierce [11], in the context of visual perception, to design the initial interpretative hypothesis of what is being perceived.

The mechanism proposed, without switching off reactions, is to consider the agent *concerns*[2], elements managed by the deliberative and social processes and accessed by all the three layers, that represent the range of desirable values for the transitory states of an individual in a specific moment.

Every concern expresses the upper and lower acceptable threshold for a given transitory state, and has a priority associated.

Besides, concerns are influenced by the personality traits of the agent, to adapt them to the specific characteristics of the individual.

3.5 Reactive Processing of Percepts

The main purpose of the reactive layer is to effectively respond to time-demanding events produced by changes in the environment. Depending on the implicit wilfulness of the response, two kind of reactive processes have been distinguished:

- *Reflex processing:* Starting from changes in the environment, and taking into account the agent beliefs and concerns, it is able to produce appropriate responses with a minimum charge of wilfulness. This is the kind of *pre-attentive processes* that Allen considers enough for an agent to survive in environments in which these generically determined solutions do not usually fail [8].

- *Conscious reaction processing:* Not all reactive behaviours necessarily emerge as reflex response. There are some preconceived or learnt reactions, triggered by a certain event, which imply a slightly higher level of consciousness, and are executed as some kind of preconceived reactive *mini-plans*. These mini-plans do not suppose any deliberative

3.6 Objectives Maintenance: Goals

Beyond the pure reactive behaviours, the deliberative and social layers base their operation on two main concepts: *goals* and *plans*.

Goals represent the objectives the agent intends to direct its behaviour in the future to. Goals are effective as stable is the environment.

Goals and concerns are conceptually different, since the concerns do not manifest the agent aims, but the limits of the desirable value of the agent's transitory state. Thus, unlike goals, which once they have been reached are discarded, concerns are maintained as long as they keep on being desirable states for the agent.

---

[2] Not exactly the same idea of "concerns" of Wright's MINDER1 proposal [10].

Goals will be characterised by an objective situation pursued, the goal current state (a life-cycle for goals has been defined), the goal importance, a goal creation time stamp (to check anytime the goal validity) and the goal expiry time.

Goals may be produced from two different perspectives, according strictly to the personal capabilities of the agent, or considering interactions with other agents to make indirect use of their capabilities. Therefore, goals will be proposed from two different layers of the cognitive architecture: from the deliberative one and from the social one.

Also for both layers, the origin of the goals could be: (1) an external source, through the interpreted perceptions coming from the interpreter; (2) an internal source, through changes or absence of changes in the internal structures of the agent – beliefs, concerns or past history–; or (3) a mixture of external and internal sources. There will be a process in each one of the layers, very similar on their purpose, but different in their scope, to produce goals. Thus, in the deliberative layer there will be a *deliberative goals generator*, which will generate goals from the deliberative point of view, while in the social layer will be a peer process, called *social goals generator*, to generate goals from the social perspective.

Besides, both goals generators perform a second function: the management of those goals. Basically, throughout this function, the agent will cyclically check the validity of every goal stored. More precisely, goals reached will be deleted, goals having lost interest for the agent will be cancelled, and unreachable goals in a past moment, but still valid, will be left ready to be planned again.

3.6 Action Resolution: Plans

The other main concept in which is based the operation of the deliberative and social layer is that of *plans*. Plans are the "paths" outlined by the agent to reach its goals from the current situation. The kind of plans managed by the cognitive module will consist of an ordered set of *actions* to be executed, the *goal* to be reached with the plan, and some particular *parameters* to help the scheduler to properly organise the proposed plans.

Once a plan has been draft to reach a goal, it goes on to swell the available agent's plans. From that moment, it will be waiting to be incorporated to the scheduler agenda –although it could be eliminated if its associated goal is cancelled.

Plans are proposed simultaneously by the planning processes located in the deliberative and social layers. Both layers provide alternative or complementary solutions to reach every agent's goal and subgoal.

Thus, the resulting plan will be a mixture of "deliberative" and "social" actions.

Since the planning strategy and procedure will be very context dependent, it should be concreted in a later phase. On this stage, the generic planning functions will be carried out on each layer by a *deliberative planner* and a *social planer*, respectively.

Both planners also care for maintaining some coherence between the deliberative/social processes and the reactive ones, through the updating of the concerns. Every time a planner proposes any action that needs some particular value for the concerns, it includes a previous fictitious internal action to update properly the corresponding concern thresholds. Once the concerns have been modified and the action has been executed, the concerns will be returned to its original value.

3.7 Actions Organisation

To finish with the description of the elements of the COGNITIVA we will analyse the structure of the *scheduler*. This component takes the action proposed by the reactive processes and by the deliberative and social layers as plans, and schedules them properly in its *Agenda*. The Agenda is an action sequence ordered according to their time of execution. The scheduler will take in every cycle the next action to be executed and will send it to the effectors.

Any *action* consist on (1) some preconditions, to allow to the scheduler to parallelise the execution of some actions; (2) an operator, which will be executed when it arrive to the effectors; and (3) the action consequences, events –desirable or not– expected after the execution of the action, needed to update the value of the expectations.

From all the action received at a given moment, the scheduler will select one to be sent to the effectors for execution, keeping the rest ordered in the Agenda to be executed later. Then the perception expectations are updated. If the expectations are fulfilled after the action execution, the execution of the rest of the actions of the Agenda will continue; otherwise, the execution of the scheduled actions will be interrupted and the Agenda will be restructured.

3.8 Complete Execution Cycle of COGNITIVA

Fig. 2 summarises the operation cycle of COGNITIVA. All its described components and their interaction is depicted.

4 Remarkable Results

As it was comented in section 2, COGNITIVA bases its strenght on its ability to be adapted to different contexts and problems without increasing

substantially its cost. With this aim, it proposes an application process consisting on two specification phases: a functional one and a contextual one.

From the generic description of COGNITIVA, we have built a functional specification, based on fuzzy logic, which provides an operative implementation of the architecture, although it is still context independent.
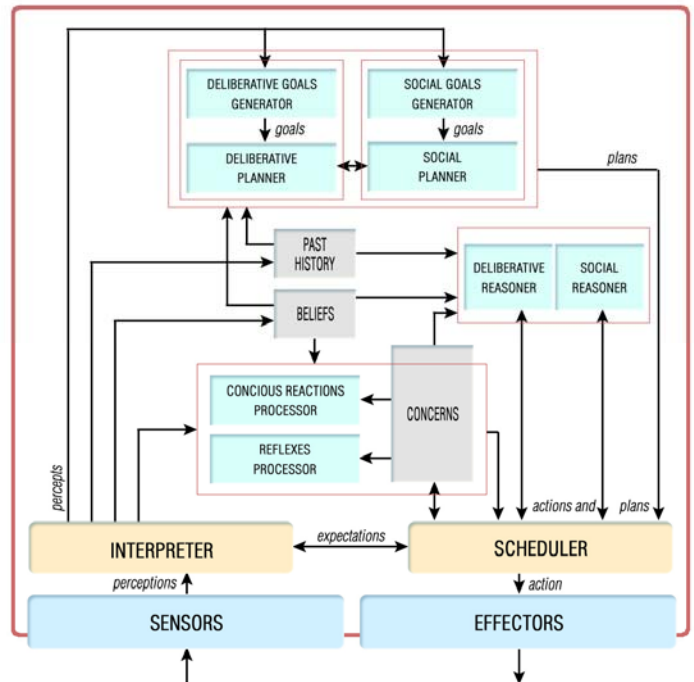

Fig. 2. Internal components and processes of COGNITIVA.

Then, to check its adaptivity, from that functional specification we have developed two contextual specifications, very different in their nature: in the first one, in a multiagent system simulating Vickrey's auctions, we have used COGNITIVA to model the behaviour of the auction bidders; in the second one, in a 3D Virtual Environment, we have used the architecture to guide the behaviours of its inhabitants.

The results obtained have been very satisfying, with effort measures of the development of both contextual specifications from the functional specification quite interesting. Furthermore, the productivity measure in both cases has shown a very promising learning curve in the application of the architecture.

5 Conclusions

Human behaviours rarely follow only rational patterns. Personality and emotions have much to do with them. For agents involved in task such as creativity or decision-making human simulation, the incorporation of that anthropomorphical features to

their architecture may be the only way to achieve believable behaviours.

Although some efforts have been done in that way, most of the architectures proposed hitherto are very context dependent, making it hard to reuse the same architecture to solve different problems.

We bet on a generic cognitive architecture, context-free, valid for many different problems. This generic architecture, COGNITIVA, is the base for building specific-context agents with behaviours influenced by personality and emotion, through the definition and particularisation of the generic functionalities and characteristics proposed. With this aim, with the architecture is provided a progressive specification process, which particularises it to the specific context, without an increasing of its complexity or computational cost.

References:

[1] Davis, D.N. and Lewis S.J., Computational Models of Emotion for Autonomy and Reasoning, *Informatica (Special Edition on Perception and Emotion Based Reasoning)*, Vol.27, No.2, 2003, pp. 159-165.

[2] Picard, R.W., Affective Computing, *Tech. Report 321*, MIT Media Laboratory, Perceptual Computing Section, 1995.

[3] LeDoux, J., *The Emotional Brain*, Simon and Schuster, New York, 1996.

[4] Adolphs, R., Tranel, D., Bechara, A., Damasio, H. and Damasio, A.R., Neuropsychological Approaches to Reasoning and Decision-Making, *Neurobiology of Decision-Making*, 1996, pp. 157-179.

[5] Gadanho, S.C., Learning Behavior-Selection by Emotions and Cognition in a Multi-Goal Robot Task, *Journal of Machine Learning Research*, Vol.4, 2003, pp. 385-412.

[6] Gratch, J. and Marsella, S., Evaluating the Modeling and Use of Emotion in Virtual Humans, *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS 2004)*, 2004, pp. 320-327, New York, ACM Press.

[7] de Sevin, E. and Thalmann, D., An Affective Model of Action Selection for Virtual Humans, *Proceedings of Agents that Want and Like: Motivational and Emotional Roots of Cognition and Action Symposium at the Artificial Intelligence and Social Behaviors 2005 Conference (AISB'05)*, 2005, Hatfield.

[8] Allen, S.R., *Concern Processing in Autonomous Agents*, Ph.D. Thesis, Faculty of Science of The University of Birmingham, School of Computer Science, 2001.

[9] Cañamero, D., Modeling Motivations and Emotions as a Basis for Intelligent Behavior, *Proceedings of the First International Symposium on Autonomous Agents (Agents'97)*, 1997, pp. 148-155, ACM Press, New York.

[10] Wright, I.P., *Emotional Agents*, Ph.D. Thesis, Faculty of Science of The University of Birmingham, School of Computer Science, 1997.

[11] Pierce, C.S., *Collected Papers*, The Belknap Press of Harvard University Press, Cambridge, 1965.

[12] Seif El-Nasr, M., Yen, J. y Ioerger, T.R., FLAME - A Fuzzy Logic Adaptive Model of Emotions, *Autonomous Agents and Multi-Agent Systems*, Vol.3, No.3, 2000, pp. 219-257.

[13] Ortony, A., Clore, G. y Collins, A., *The Cognitive Structure of Emotions*, Cambridge University Press, 1988.

[14] Velásquez, J.D., When Robots Weep: Emotional Memories and Decision-Making, *Proceedings of the Fifteenth National Conference on Artificial Intelligence and Tenth Innovative Applications of Artificial Intelligence Conference (AAAI'98/ AAAI'98)*, 1998.