Robots as Responsible Agents

Eugénio Oliveira

NIA&R-Artificial Intelligence & Robotics Group Faculty of Engineering, University of Porto Rua dos Bragas, 4099 Porto Portugal eco@garfield.fe.up.pt

ABSTRACT

The quest for real autonomous robots leads us to discuss the problem about the best possible control architecture enabling that important characteristic. It has been broadly accepted that an hybrid architecture, i.e. putting together both reactive and deliberative paradigmes is needed to efficiently execute tasks in realistic dynamic environments. Our proposal, which is being implemented to control a RobuterII (from robotsoft) mobile platform, involves the use of a two-layers architecture, using symbolic representation for knowledge and goals at the deliberative level and sub-symbolic neural networks for implementing the behaviors at the reactive level. One of the main problems we are now addressing is how to make these two levels to communicate, to interat without being completly depending from each other. The Multi-agent system framework gives a flexible strategy for single agents cooperation and enables a set of behaviours to have a certain degree of autonomy. This reactive layer works together with the cognitive control agent where goals and commitments are logically represented through a simple modal logic.

1. INTRODUCTION

In the quest for the intelligent, autonomous mobile robot, we are faced with several fundamental questions:

Which Information and what kind of data should be processed, how much reasoning capabilities have to be present, how reflexive an entity (an Agent) should be, in order to display, up to a certain degree, autonomy, flexibility, responsiveness and competence. Extreme positions that have been hold for the past ten years on mobile robots control, both by behaviorists as well as cognitivists, tend to become closer, recognizing that both approaches have to include some part of other ones strategy. This fact lead to the emergence of new different hybrid architectures which still have to be characterized following different axes like:

- Which classes of problems are the most suitable for which kind of architecture?
- How much of the problem should be dealt with, in a more reactive way, by simple behaviors, and which part needs to besolved through a more complex reasoning process, mediating the sensing and action phases?
- Which control architecture to use: Centralized? Hierarchical Decentralized? Distributed?
- Do the world as well as other external entities play a role on this architecture? And, in consequence should Agents be taken just as Co-operative? Competitive? Benevolent? Selfish?
- Which implementation paradigms are the most appropriate for representing either behaviors or planning and reasoning capabilities? Sub-symbolic? Symbolic? Procedural?
- How to interface the different "levels" of control, using different types of information and representation paradigm (data, rules, behaviors, neural networks)?

These are some of the most important questions that have been recently addressed and did not get yet satisfactory answers. In our opinion, only an architecture based on the modular, decentralized and co-operative model of MultiAgents Systems [1] [2] can be rich enough to accommodate all the flexibility required for performant mobile robots. This choice immediately implies that, first, the robots' functionalities have to be distributed across different Agents and second, that it is not enough to have reactive Agents but, on the contrary, a certain amount of knowledge has to be present at least to enable both self-learning and Inter-Agent Co-operation. Including some "introspection" as well as reasoning capabilities upon other ones external behavior and potentialities may be a key feature to make it to converge the overall meaning of the different agents actions belonging to the community to specific useful goals.

As preliminary conclusions, let us stress that first, action should take place as a result of multiple interactions between Agent's potentialities and the environment and, second, we have to find out some harmony between deliberation and reaction, preventing the robot from a possible schizoid character. Notice that we are here exactly following the principle that is observed in the cortical visual system of primates where reflective and instinctive behaviors are clearly distinct (Goodale and Milner, Trends in Neuroscience, 1992).

2. NATURAL EVOLUTIONISM?

It is a known fact that existing hardware is responsible for important limitations on robots implementations and we will not overcome this situation in the next years. Therefore we are explicitly facing the following dilemma:

- Shall we go on developing sophisticated software, including the possibility to use knowledge and some kind of reasoning - may be inspired by human capabilities- despite the present huge hardware limitations? Or

- Shall we stick to the current hardware possibilities and, in accordance to that, try to implement simpler features, at the "instinct" level, and tackle only problems that imply animal-like behaviors and not those ones where "intellectual" activities have to be involved?

Clearly, the first approach is the one taken by Artificial Intelligence research community, since works on classical planning to more recent excitement on Distributed AI in general and Multi-Agent Systems in particular. On the other hand, "behaviorists", guided by Brooks (MIT) [3], they believe that we should start by building up flexible but simple robots (preferable lots of them) which, reacting to the real environment would make consistent and meaningful patterns of behavior to emerge out of their mutual interaction.

This debate is far from being over but, nevertheless, there is a big number of recent contributions supporting hybrid approaches where the reflective capabilities, although being present, do not prevent higher levels of abstraction involving both knowledge and decision making guided by inference. Waiting for a "natural" evolution of robots following, step by step, each new development of the hardware seems not to be reasonable once history proves the non-linearity of scientific evolution, giving the possibility of building up models and hypothesis, long before they could be completely experimentally tested.

3. GENERAL ARCHITECTURE

Taking as granted that this latter philosophy appears to be the most promising one, we have now to focus on several important problems that are the direct consequences of the previous choice.

First of all, the fundamental problem we are faced with, is about which one could be the most suitable architecture to accommodate this double-face entity capable of reasoning and reacting in accordance to the current situation. Shall it be a two-level control entity as it seems to be the most natural, due to this evidence which opposes reasoning and behavior? Or would it be better to consider a more completely hierarchical structure, centered on a master module? And the two levels, shall they be independent, working in parallel, in sequence, or alternatively? Shall they communicate? And if the answer is yes, how can they understand each other, being of a different nature (reasoning and reacting)?

Our proposal points out a two-level architecture, including a lower layer implementing a sub-symbolic system which encodes a number of distinct and elementary behaviors and, in parallel, an higher level layer where available knowledge and learning capabilities make it possible justified decision-making guiding possibly less immediate, but still situated, actions (see fig.1).

The different nature of the two proposed layers arises a problem of interfacing which is not trivial. We may look at this two layer interfacing mechanism in two rather different ways:

- First of all as an active pathway between both sub-symbolic(implicit) and symbolic (explicit) representations, conveying in both ways data produced in one of them but still having the possibility to be meaningful, or at least useful, by triggering other ones functionalities. Here we are putting an emphasis in the heterogeneity of the nature of the exchanging information.

- Secondly, this interface may take a different aspect, like a kind of translator which is able to pick up knowledge and definitively convert it into a more efficient (although implicit) sub-symbolic representation or, vice-versa, by abstracting upon behaviors' results leading to more general knowledge, explicitly described. Here we are emphasising the process of when and how to make each level to influence the other one.

Once a choice has been made on the use of both paradigms (reactive plus cognitive) to embody an Agent, a number of consequences follows at the level of the Agent capabilities when interacting with the environment (which, of course, includes other Agents). What we cannot say, is that it is indifferent the use of Neural Networks or Symbolic Programming to implement an Artificial System functionality as it seems to be believed by M. Arbib and J. Liaw [4] in opposition to what happens with the brain where Neural

Networks is argued to be the only medium where the structure meets function.

4. MOBILE ROBOT - THE AGENT

Until now we were talking of Agents in general. They could be Expert Systems co-operating in a distributed environment or just softbots, software programs "snoozing" around a network of available local users description pages. From now on, our Agents will have a body to sense and, may be, to modify a dynamic external world. We want one (or several) mobile robot(s) to evolve intelligently in the world, having goals and using its autonomy to be externally recognized as useful.

Three main questions have to be answered from the beginning:

1- Which basic behaviors should be given to the robot and how to implement them following the sub-symbolic paradigm?

2- Which high level reasoning capabilities should be available to make it possible to the robot to preview, to be committed and to learn? And, associated with this, how to represent the minimum apriori knowledge to make the robot "socially acceptable"?

3- How to enable a dynamic upgrade of the robot specific reactions ? Which situations do necessarily impliy reasoning ? In different words, what this exactly means is "how to implement, for a concrete robot, the above mentioned parallel two-layered architecture including their mutual interface ?".

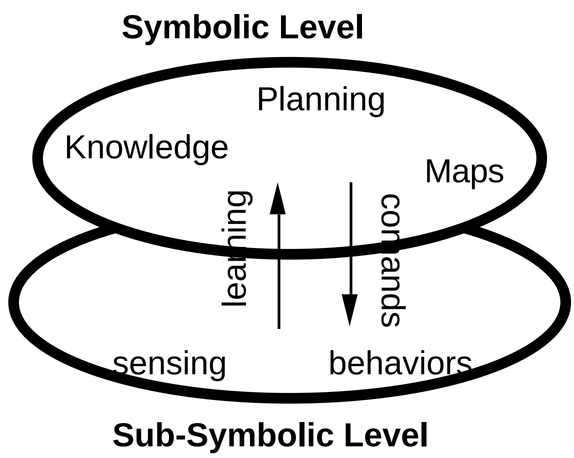

Fig.1 A two-layers control architecture for the Robot

To the first of these last questions, our is a classical and simple answer got from the observance of any leaving mechanism. The most basic behaviors of our mobile robot are: Sensing; Signaling; Moving around; Feeding; Fleeing; Avoiding obstacles. In order to be seen as "instinctive-like" encoded, they have to be implemented as independent, modular and simple neural networks.

The second question is somewhat more difficult and implies choosing among many possibilities, sometimes without a clear justification for the final choice. The first and most significant high level capability of the Agent (Robot) is the possibility to learn, not just meaning that it can adapt itself to new circumstances (which can be done also at the sub-symbolic level) but mainly that it can improve its "consciousness" of itself, the others and the rest of the world. What has been said has an immediate consequence that a certain accumulation of knowledge takes place as well as an apriori knowledge has to be inseminated at the Agents' symbolic knowledge level. It is also expected that at least some of this knowledge will take the form of topological and geographical information.

We are not implementing just intelligent systems but, mainly, interactive agents and, therefore, we reserved a special place for knowledge on how to interact and co-operate, giving, from the beginning, a social dimension to our robot. Finally, in order to give the robot the possibility to be useful for multiple purposes and in a goal-driven way, we gave the possibility for goals representation in the robots "mind". Several modal logics have been put forward as candidates to represent Agents Intentions. We need a Logic of Knowledge but also able to accommodate Believes, Goals, Commitments and Intentions.

We follow Cohen and Levesque [5] in the proposal for the formalization of the notion of intention through a logic of the rational agenthood where the primitive modalities are:

Believe: (BEL x ϕ) Agent x believes in ϕ

Goal: (GOAL x ϕ) Agent x has objective ϕ

Happens: (Happens α) Action α follows

Done: (DONE α) Action α just happened

BEL accessibility relation (in the possible worlds semantic) is Euclidean, transitive and serial leading to a KD45 logic of BEL. GOAL relation is serial leading to a KD logic. HAPPENS and DONE are now augmented with operators ´; ´
and ´? ´.
The operator LATER, (always) and ◊ (sometimes) are now defined as:

$$\Diamond \alpha \overset{def}{=} \exists\ x\ (\text{HAPPENS } x; \alpha?)$$

$$\alpha \overset{def}{=} \sim \Diamond \sim \alpha$$

$$(\text{LATER } p) \overset{def}{=} \sim p \wedge \Diamond p$$

Also (BEFORE p q) p happens before q can be defined:

◊ ~ (GOAL x (LATER p))

Now we may say that an Agent has a "persistent goal" ϕ if:

1- His goal is to make ϕ to be true and he does not believe ϕ is true at the moment.

2- Before giving up, one of the two following conditions hold:

i) Agent believes ϕ is already true,

ii) Agent believes that ϕ will never be true.

Therefore we can define the operator "persistent goal", which is of great importance in the context of autonomous robots, as follows:

$$(\text{p\_GOAL } x\ p) \overset{def}{=} (\text{GOAL } x\ (\text{LATER } p)) \wedge (\text{BEL } x \sim p) \wedge [\text{BEFORE} \; ((\text{BEL } x\ p) \vee (\text{BEL } x \sim p)) \; \sim (\text{GOAL } x\ (\text{LATER } p))]$$

The Agent intention to do an action is now defined as:

$$(\text{INTEND } x\ \alpha) \overset{def}{=}$$

(P_GOAL x [DONE x (BEL x (HAPPENS α))?; α])

In our opinion these modalities are powerful enough to represent inside our Robot's (agent) "mind", appropriate knowledge, believes, intentions and commitments to and about action.

The result of the robots "mind" activity may be either the enhancement of its inner, through the acquisition of new knowledge, or believes, (information, rules, procedures) or just a decision making process about what commands to select to immediately direct the robots actions.

What have to be stressed here, once again, is that this more deliberative activity is not to be always preceding every robots action but, instead, may take place either in parallel with the robot behaviors actions or even a posteriori, as a result of some robot activity.

5. MULTI-AGENT INTERACTION

Our robot was seen, in the framework of the last sections of this paper, as an agent that can interact with the world, possibly including also other robots. Now it is time to precise how to control one robot's action by means of these two main capabilities, we know are present at the robot control level: Deliberation and Reactivity [6].

Being strongly against giving control of the robot's activity either to a planning module or to a reactive mechanism, gives us the responsibility to find out some other way of defining what happens mediating sensing and action. And the answer seems to be that robots actions should be the result of a co-operative interaction among different distributed agents inside the robot, which includes primitive behaviors and decision-making output commands in a flexible, non-previewed order and varying from one extreme - immediate reaction, to the other extreme - deliberative planning activity. It is, of course expected that most of the actions selection will be the consequence of these two intermingled capabilities achieved through co-operative interaction of multiple agents.

The "structural coupling" of evolved species and external world referred to by Maturana and Varela, puts the accent in the interaction and not either on planning capabilities or on world representations inside the robot. This means that what has to be of major importance to the robot is the association between external world perceived patterns and internal schemata which may include behaviors as well as reasoning steps to make the robot evolve in accordance to the perceived world.

How do these schemata look like? What kind of patterns are going to activate these schemata and/or appropriate behaviors?

Having abandoned the so-called "correspondence method" (between the static world and internal representations), we must here focus upon the potential interactions, or situations characterized by the dynamic coupling of the robot and its environment.

Clearly, there are simple patterns (signals captured from the world), triggering simple and immediate behaviors - stop, fleeing, waving around, sensorial acquisition, signaling - or even not so simple but still at the "instinctive" level (carefully approaching, feeding, ...). These are situations that do not present special problems, once after recognition of a set of signals (a pattern), some procedures (or just commands) ought to be executed.

Nevertheless, there are other situations where reasoning has to be involved either to interpret a more complex dynamic pattern or in order to pursue the execution of a previously given task. Local goals are then generated either through the incoming information (from the outside) or/and by robots intentions. But in both cases the following action, independently on how it was decided, has to take into account the selected interaction schemata which captures the current situation. Therefore, the robot action is referring to the current interaction, possibly together with long term intentions (desires) and not just to a previously fabricated plan or some internal image of the world.

We have written above that the two different control levels of the robot - symbolic and sub-symbolic - could work in parallel, and it is now time to clarify how is this possible, to control robot's action through two almost independent instances of decision-making capabilities. That is for sure that deliberation cannot, most of the time, be so precise that planned actions can be directly executed. Plans are here seen as composed of schemata activation which, by themselves have certain preconditions to be fulfilled through the local examination of the interacting environment.

A schemata is then activated by a pattern and may contain different kinds of actions. What does a pattern look like?

A set of signal values coming from the available active sensors. This means that the pattern format depends much on the type as well as the number of sensors.

The action part of the schemata includes several slots:

1- which behaviors are to be activate;

2- which information is to be conveyed to the higher level;

3- which further sensorial acquisition actions will take place.

Each one of these three slots may be empty, but also all of them may be simultaneously instantiated (see fig2). Several primitive behaviors may be simultaneously active like reactive agents. Moreover, through the "informing planner" channel, they can convey information to the cognitive agent, working at the symbolic level. It is also true that the latter is able to send some some useful advices to the lower level, in order to influence the way behaviors may be combined (see [6] for details). In conclusion we may say that, emerging from the agents interaction we can recognize an a apropriate, flexible and evolutive global robot behavior.

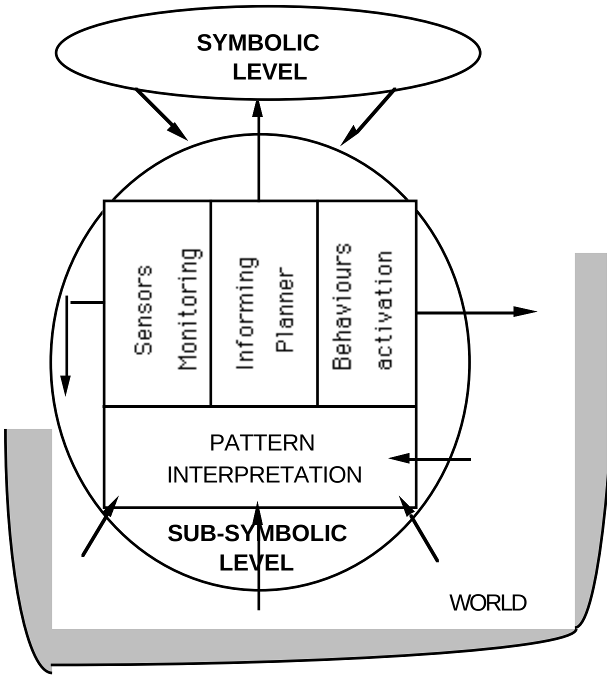

Fig.2 Robot sub-symbolic activities

6. THE IMPLEMENTATION

We will not discuss here details about our work on a particular robot to which has been given an hybrid control architecture inspired by some of the concepts here explained. Some details can be found in [6].

We are using a Robuter II mobile platform (from the French company Robotsoft) equipped with a set of ultrasonic sensors plus proximity sensors and using two different computing capabilities: on-board computer with a "real-time" operating system and a remote Unix sparc workstation communicating via modem with the robot platform.

At the sub-symbolic level a set of behaviors, implemented through very simple neural networks are organized in a subsumpsion-like architecture which dynamicaly changes according to advices coming from the planning agent [6] and in the upper, deliberative layer, planning capabilities are now starting to be enhanced with modal logics in accordance with what has been described so far in this paper.

Although some of the intended features have not yet been experimented, and therefore we have no real outcome the preliminary results, mainly regarding some learning capabilities about behaviors, are indeed promising, despite the poor performance of the existing sensors.

It is our intention to investigate further on this problem of how to keep in parallel two main capabilities for guiding the robot's actions in a real scenario.

7. REFERENCES

[1] G. O'Hare and N.Jennings (eds), Foundations of Distributed Artificial Intelligence, John Wiley & Sons, 1996

[2] A.Bond and L.Gasser (eds), Readings in distributed Artificial Intelligence, Morgan Kaufmann, S. Mateo, CA, 1988

[3] Rodney Brooks, "A Robusted Layered Control System for a Mobile Robot", in IEEE Journal of Robotics and Automation, Vol RA2, Nº1, 1986.

[4] M. Arbib and J. Liaw, Artificial Intelligence Journal, V.72, N.1-2, pgs 53-79, Jan 1995

[5] Cohen,P. and Levesque, H. "Intention = Choice + Commitment", in Proceedings of AAAI, Seattle, pgs. 410-415, 1987.

[6] M. Neves, Eugénio Oliveira, " A control architecture for an autonomous mobile robot", in Proceedings of the

First International Conference on Autonomous Agents, Marina del Rey, CA, USA, ACM Press, 1997.