Adding Active Mental Entities to Autonomous Mobile Robot Control Architectures

Pietro Baroni*, Daniela Fogli*, Giovanni Guida*, Silvano Mussi^*

*Università di Brescia, Dipartimento di Elettronica per l'Automazione,
Via Branze 38, 25123 Brescia, Italy, e-mail: {baroni, guida, fogli}@bsing.ing.unibs.it
^Consorzio Interuniversitario Lombardo per l'Elaborazione Automatica,
Via Sanzio 4, 20090 Segrate (MI), Italy, e-mail: mussi@icil64.cilea.it

Abstract: This paper aims to lay down the foundations of an approach to autonomous mobile robot (AMR) navigation based on the explicit representation of mental entities (desire, intention, obligation, etc.) underlying robot behavior, considered as autonomous active entities. The approach is intended to integrate concepts from the area of distributed architectures and of mental entity representation and it is claimed to allow the achievement of both robust and sophisticated robot behavior. Starting from an analysis of the evolution of AMR control architectures, the introduction of mental entities in the context of distributed control architectures is motivated as a further step in the direction of making explicit the deep reasons that underlie the behavior of intelligent systems. A simple application example concerning an autonomous mail delivery robot is finally provided: the example demonstrates the main features and the potential of the proposed approach in facing unusual and complex tasks.

1 INTRODUCTION

In the design of control architectures for autonomous mobile robots (AMRs) there has been a progressive evolution from hierarchical paradigms, based on the sense-plan-act loop [14][16][5], to behavior-based architectures featuring the explicit representation of behaviors [3][1][6][9] and distributed control architectures [2][7], where robot control is distributed among autonomous interacting agents. Moreover, other researchers [4][12][13] have suggested to endow robots with so-called “mental states”, in order to improve their ability to operate autonomously in unknown environments.

This paper aims at laying down the foundations of an approach to AMR control based on the explicit representation of mental entities (desire, intention, obligation, etc.) underlying the robot behavior, considered as autonomous agents. The approach is grounded on two assumptions:

- the adoption of a distributed paradigm for the control of AMRs is crucial for obtaining effective behavior in complex situations;
- the explicit representation of mental entities as autonomous agents is crucial for supporting advanced AMR performance.

The proposed approach is based on the concept of active mental entity and can account for the mechanisms by which external AMR behavior can be generated as a result of internal mental processes.

2 AMR CONTROL PARADIGMS

2.1 Evolution and state-of-the-art

Several researchers are involved in studying how to make a robot an *intelligent autonomous agent*. For example, Pfeifer [11] and Steels [15] analyze cognition mechanisms to understand how to design robots that can operate in a variety of unforeseeable environmental situations and are capable of autonomously updating their world models. A useful way to analyze AMR control paradigms is to study how their behavior is obtained. Two main stages of development may be identified.

Coded behavior

This paradigm, often called *hierarchical paradigm*, is based on the principle of explicitly coding how the robot should behave in any specific situation. Robot control is achieved through a sense-plan-act loop: a suitable sensory system identifies the current situation and then a hierarchical plan determines the actions to perform.

Coded behavior does not provide indeed any explicit representation of robot behavior, but only

focuses on actions, fully determined, through a hierarchical plan, by the sensory data.
This approach is presently considered obsolete and has been generally replaced by the more advanced explicit behavior paradigms.

Explicit behavior
This paradigm, often called *behavior-based paradigm*, is based on the assumption that robots are intelligent problem solvers and their problem-solving capabilities are determined by their behavior. Behavior is understood as a complex process, governed by a suitable control regime. Clearly, the control regime adopted for managing robot behavior greatly influences its performance. Considering proposals, based on this paradigm, developed and experimented in recent years, we distinguish control regimes into two main categories: rigid control and flexible control. These are illustrated below.

Explicit behavior with rigid control
The first proposal based on the explicit behavior paradigm is the *subsumption architecture* developed by Brooks [3], who advocates a problem decomposition strategy based on task-achieving behaviors, which replaces the old way of decomposing the control system of a robot into sensing, planning and acting, adopted by coded behavior approaches. To this end, a set of competence levels are defined; each level provides an informal specification of a class of behaviors such as "avoid contact with objects", "wander aimlessly around without hitting things", "explore the world", etc. Layers of the robot control system are then built in correspondence to competence levels. Each layer is implemented by simple modules which exchange messages and operate in parallel. In this way, the AMR architecture can be developed incrementally, starting from lower-level layers and proceeding step by step towards higher-level ones. Unfortunately, this tends to result in a hierarchical organization, with a great amount of interdependence between behaviors, lacking modularity and flexibility. Moreover the need of implementing high level operation strategies can hardly be achieved from the simple combination of purely reactive behaviors.
To overcome this problem, a control mechanism is needed; thus, more recent proposals develop behavior-based architectures with centralized control.
Balch et al. describe a trash-collecting team of robots, that can implement several reactive behaviors, each one devoted to accomplish a particular step of the office cleanup task [1]. A centralized structure determines the correct sequence of behaviors needed to solve a specific cleanup problem (sequenced coordination).
Also Konolige proposes a behavior-based architecture for his robot Erratic [6]. Here, the behaviors are composed by means of fuzzy rules and are coordinated by a decision system based on an intention schema. At any moment, several schemata (10 to 15) are able to operate in parallel, monitoring conditions and coordinating behaviors. The system can react in real time to environment changes.
Finally, Montgomery et al. [9] built a flying robot which uses a behavior-based control architecture and which is capable of locating and manipulating objects and transporting them from one location to another. Behaviors are organized hierarchically: at the highest level of control, a *sequencer* determines the behaviors to activate and the parameters to instantiate to achieve a desired subgoal. The authors underline the difficulty of increasing the complexity of a behavior-based system due to the possible coupling of behaviors. To overcome this limitation they propose, as a future work, the integration of a fuzzy-rule system with the behavior-based controller.

Even if centralized control is intended to improve purely behavior-based architectures, it may, however, become a bottleneck in complex problems. It is quite rigid and generally does not allow dealing with multiple goals simultaneously. Moreover, organizations with centralized control do not support incremental system development and makes maintenance and extension hard and costly.

Explicit behavior with flexible control
Distributed robot control architectures aim at overcoming the limitations discussed above. The issue of distributed control has been explored by Liscano et al. [7], who propose a mobile robot with a blackboard architecture for coordinating and

integrating several real-time activities. According to the classical blackboard concept [10], this architecture includes a collection of highly independent modules, called knowledge sources, with their own inference mechanisms and a local knowledge base to perform specific tasks. Blackboard structure does not support, however, reactive behavior capabilities. Thus, the work of Liscano et al. presents a modification of the approach to manage real-time issues by avoiding that all communications are forced to go through the blackboard.

In [2], Baroni et al. propose a more advanced distributed architecture, called *specialist net*, that includes many specialists, each able to carry out a specific activity.

Specialists cooperate according to a task-sharing approach. In this approach, when an agent is faced with a task too large or too complex for its capabilities, it may request assistance to other agents available in the architecture. It first decomposes the task at hand into more manageable sub-tasks and then it attempts to find other specialists with the appropriate competence to handle them. Specialists include a set of dedicated modules, in charge of decomposing complex problems, allocating the sub-problems resulting from decomposition, solving problems falling within their competence domain, and communicating with other agents.

2.2 Limitations and open issues

All the approaches to the problem of controlling AMRs presented in the previous section can not support sophisticated behaviors in front of unusual and complex situations. Modules (or behaviors) which compose the described robot control systems have limited capabilities and allow only simple tasks to be carried out.

Referring to Brooks' robots, Pfeifer [11] says that "if they are to perform more intelligent tasks they will need to be equipped by cognition", where cognition means the ability of dealing with desires, inhibitions, intentions, etc. In more general terms, it seems reasonable to envisage a new AMR control paradigm that, pushing further the evolution surveyed in the previous section, goes beyond explicit behavior. Behavior can be understood as the external visible result of a hidden cognitive activity occurring inside the robot control system. Behavior does not need to be explicitly defined and managed, as in the explicit behavior paradigm, but it is determined, in an indirect way, by the cognitive activity.

We call the new paradigm proposed above *indirect behavior*, and we claim that it can support more sophisticated and effective solutions of the AMR control problem than explicit behavior.

3 ENDOWING ROBOTS WITH ACTIVE MENTAL ENTITIES

3.1 The intuitive concept of mental entity

Human behavior is often explained by using terms such as desire, inhibition, intention, hope, obligation, prohibition, etc. These terms denote *mental entities*, i.e. entities that are inside the mind of a person and which are responsible of his/her external behavior.

The complex processes that occur in human mind involving mental entities are called *mental processes*.

The totality of our actions is guided, consciously or not, by our mental processes.

Behavior and mental activity are complex interconnected processes. In a simple framework, behavior can be understood as the combined result of mental activity and perception of the world. No behavior can be explained without taking into account (at least) these two components.

3.2 Why modeling mental entities for AMR control ?

In order to explore the motivations underlying the use of mental entity models for AMR control, let us examine some different descriptions of the operation of a trash collecting robot. From an external point of view, robot operation can be described in very simple terms such as: the robot is able to move around, recognize trash, approach it, and collect it.

The overall robot operation can then be described, according to the coded behavior paradigm, as a move-recognize-approach-collect loop. In this description only the desired operation of the robot is considered and can be coded in algorithmic terms. However, this solution is weak and fragile when confronted with real-world situations. For example, the goal to avoid moving obstacles can hardly be dealt with in such a simple approach. In fact, this goal may arise asynchronously and has to be achieved in real-time, without waiting for the completion of the current step of the AMR control algorithm. Moreover, it can not be treated simply as an interrupt, since after the obstacle has been avoided the state of the robot and of the world has changed and the previous process can not be simply resumed.

Behavioral approaches have been designed to overcome this kind of limitations. According to the explicit behavior paradigm, robot capabilities are modeled as independent behaviors, such as "in presence of trash, collect it" or "in presence of an obstacle, avoid it". This approach to behavior modeling corresponds in some way to a modular rule-based representation. In fact, each behavior corresponds to a separate part of the problem, just like a single rule corresponds to an isolated chunk of domain knowledge cased in a condition-action frame. There is no explicit representation of the reasons why obstacles should be avoided, as well as there is no justification for the relation between the IF-part and the THEN-part of a rule. The desired robot operation arises in this case from a not predetermined combination of single behaviors, just like useful reasoning results from a not predetermined combination of rules. Attention is therefore paid to specify behaviors, whereas global operation results as a side-effect of correct behavior specification.

In behavior-based systems with rigid control the connections between sensory data and behaviors and between different behaviors are strictly programmed and there is a hierarchy of control levels where higher-level behaviors subsume the lower-level ones. Therefore, no dynamic interaction between behaviors is allowed and the overall behavior organization turns out to be complex, rigid, and difficult to design, test, extend and modify.

To overcome these difficulties, dynamic interactions between autonomous agents, each one playing a specific role in the overall system organization, are explicitly represented in behavior-based systems with flexible distributed control.

For instance, the agent in charge of obstacle avoidance may cooperate with the agent in charge of approaching trash in order to produce a motion strategy which satisfies both needs.

Distributed control approaches are based on the fundamental assumption that the global behavior of the robot is the result of the interaction between elementary agents, each one in charge of a specific activity. No attention is paid, however, to the reasons underlying agents' activity and to the motivations which push agents to interact. In a sense, agents are understood as behavioral black boxes; they can provide specified behaviors but the internal mechanisms that determine such behaviors are not modeled.

This implies that agent behavior is constrained to remain the same in all situations and does not take into account conditions that may affect the reasons that makes such behavior appropriate (or not). Consider, for example, obstacle avoidance. Such behavior is generally appropriate, since the robot may damage itself through a collision with an obstacle or since an obstacle may obstruct the shortest path towards a target. However, in the case an obstacle can not damage the robot and can be pushed (for example, it is an unlocked door or an empty cardboard box obstructing a corridor) the robot might prefer to hit the obstacle, rather than searching (possibly in vain) a way to avoid it. In order to choose between these alternatives, the robot should have an explicit representation of the motivations that lie behind its behavior. It should be able to reason about the intentions from which its behavior derives, in order to tailor the behavior to meet such intentions in the most appropriate way, compatibly with the current situation. This approach implies a significant focus shift with respect to the behavior-based paradigm; behaviors are no more considered as the primitive entities from which the overall robot operation derives, but are themselves produced by the mental activity occurring inside the robot control architecture. This corresponds to a deeper modeling activity, which aims to bring to light the mental roots of behavior, according to the assumption that, like better operation can be achieved by modeling the behaviors underlying it, better behaviors can be achieved by modeling the mental activity on which behaviors are based. This way, it is possible to improve robot intelligence in facing new and critical situations.

Modeling mental activity encompasses modeling mental entities and mental entity interactions. Mental entities are not considered as passive entities (mere information structures, mere static objects, operating according to stated and fixed procedural rules) but as active entities, i.e. entities endowed with autonomy. In fact, since mental entities have to dynamically interact to produce a globally intelligent behavior, it is essential that they are provided with individual and independent operation capabilities. Therefore, we call them *active mental entities*, stressing that they are autonomous and can operate and cooperate according to their own goals and strategies. So, recalling that "intelligence emerges from the interaction of the components of the systems" [8], in our proposal, system intelligence emerges from interaction of active mental entities.

4. AN EXAMPLE

In this section we present a simple example that shows, in intuitive terms, how the proposed ideas could be implemented in an actual robot, allowing it to deal successfully with complex problems and unusual situations. It is assumed that the robot control is managed by a distributed architecture composed by autonomous agents, each capable to carry out specific activities. Each agent is endowed with active mental entities: the interactions of active mental entities determine agent's (and consequently robot's) behavior. For the sake of clarity and simplicity we consider here only two types of active mental entities, namely intentions and persuasions.

We say that an agent has an *intention* when it commits itself to pursue an achievement. Intentions are modeled as active mental entities, since we assume that an intention is an autonomous entity definitely committed to reach its achievement, possibly cooperating or conflicting with other intentions.

We say that an agent has a *persuasion* when it has some belief about the truth value of a given proposition. Persuasions are modeled as autonomous active entities since we assume that a persuasion is not just the passive result of some perceptual or reasoning activity, but it is definitely committed to reinforce itself, i.e. to find new elements to believe or disbelieve in the related proposition and to verify the already available ones. In such activity a persuasion may cooperate or conflict with other persuasions.

The example concerns a department mail delivery robot, to which the user consigns an envelope to be delivered to Mr. X. Interaction with the user is managed by a specialised agent UI (User Interaction), which is characterized by a primitive intention whose subject is "obey-the-user". According to this intention, a new intention has to be generated whose subject is "deliver-mail-to Mr. X". However, since UI has no specific competence on mail delivering, it has to address the request of creating this new intention to another competent agent. By resorting to its interface, UI identifies the MD (Mail Delivery) agent to which the request is addressed: the new intention whose subject is "deliver-mail-to Mr. X" is therefore created within MD. MD has both general knowledge about the mail delivery task and specific knowledge about the department personnel, as far as mail delivery is concerned. So MD knows that Mr. X is actually a department employee, that his office is office number 9, and that normally, his office hours are from 9.00 a.m. to 5.00 p.m.. Since current time is 3.00 p.m. (this information is provided on request by a specialised CLOCK agent), MD generates the persuasion that Mr. X is now present in the department.

This persuasion supports the persuasion that the intention "deliver-mail-to Mr. X" can be actually achieved. On these grounds, the intention "deliver-mail-to Mr. X" can then elaborate a strategy to deliver mail to Mr. X. To this purpose, however, some more detailed persuasion about where Mr. X is has to be instantiated. In absence of more specific information, using default knowledge that normally an employee is in his office, the following persuasion is generated: "Mr. X present-in-office". On the basis of this persuasion, the following simple strategy is eventually generated:

task 1: go to office 9;
task 2: deliver the envelope to Mr. X.

Task 1 is considered first: it still concerns a quite generic and high-level task and must therefore be associated to a new intention. A request of generating such intention is therefore addressed by MD to the MM (Movement Management) agent.

MM, generates the intention "go-to-office 9" and, exploiting his knowledge about building topology, it generates then a strategy to reach office 9 and begins to realize it.

It has to be stressed that, meanwhile, all the mentioned mental components do not remain simply waiting for the accomplishment of the selected strategy. On the contrary, they remain active: intentions are continuously looking for better strategies and persuasions are continuously looking for new evidences supporting them. Moreover, it can be assumed that some agents in charge of sensory acquisition are always active, because their output is necessary to other agents in charge of very essential primitive intentions, such as preserving robot integrity. Sensory data can therefore be used by intentions and persuasions for the activity mentioned before.

While the robot moves towards office 9, the intention "deliver-mail-to Mr. X" may elaborate the following alternative strategy:

task 1: find Mr. X around in the department
task 2: go near Mr. X
task 3: deliver the envelope to Mr. X.

This strategy relies on the persuasion that Mr. X is not in his office but somewhere else in the department: "Mr. X present-but-not-in-office". However, before the strategy becomes operative, the persuasion has to find some support. To this purpose, "Mr. X present-but-not-in-office" may generate intentions like "recognize-voice-of Mr. X" and "recognize-face-of Mr. X" to be addressed to agents specialized in processing audio and video input coming from sensory devices.

If sufficient resources are available, these activities can be carried out in parallel while the robot moves toward office 9, according to the first generated strategy.

Let us suppose now that, while moving, the robot gets near a glass wall behind which there is Mr. X. Then the vision system recognizes Mr. X in front of the robot and, therefore, persuasion "Mr. X present-but-not-in-office" gets strong support, whereas "Mr. X present-in-office" is dismissed, since direct evidences prevail always over default knowledge.

As a consequence, also the first strategy generated on the ground of "Mr. X present-in-office" and the relevant intentions are abandoned (note however that they will revive if, for instance, it will be realized that Mr. X recognition was erroneous).

The second proposed strategy becomes then active and, since task 1 has been achieved (find Mr. X around in the department), task 2 is pursued (go near Mr. X).

A request to create the intention "go-near-Mr. X" is therefore addressed to the agent MM (which, in the meanwhile, has dismissed all the activities implied by the first strategy). Suppose now that Mr. X is standing just in front of the robot. The goal of navigating towards such a fixed target is reduced to the goal go-forward. While the robot is moving forward, the VC (Video Camera) and SRS (Sonar Range Sensors) agents acquire and process data about the external world. Doing this, they continuously generate or update persuasions about the environment, whose subject is communicated to the agent CA (Collision Avoidance).

Suppose now that, while the robot is approaching the target, VC and SRS communicate to CA two contradicting persuasions: SRS has the persuasion that "there-is-an-obstacle-on-the-path", whilst VC has the persuasion "no-obstacle-on-the-path". CA recognizes that it is impossible to take a decision, given these contradicting persuasions, and therefore decides that the conflict should be solved. It puts the two persuasions face to face by notifying each of them of the existence of the opponent persuasion.

The persuasions "there-is-an-obstacle-on-the-path" and "no-obstacle-on-the-path" enter therefore a debate in order to solve the conflict. First of all, an analysis of the motivations supporting them is carried out: "there-is-an-obstacle-on-the-path" is supported by the fact that sonar received reflected echoes, "no-obstacle-on-the-path" is supported by the fact that in the image collected by video camera nothing but Mr. X is seen.

The persuasions "there-is-an-obstacle-on-the-path" and "no-obstacle-on-the-path" are then in charge of searching evidence or other persuasions corroborating their supports or undermining the opponent's ones. For instance, "no-obstacle-on-the-path" can resort to general knowledge (this knowledge may be provided by the SRS agent itself) that sonar readings are often erroneous and notify it to "there-is-an-obstacle-on-the-path". In turn "there-is-an-obstacle-on-the-path" may reply that sonar readings are erroneous in specific conditions (near wall corners, in presence of noise sources, etc.) that are not met in the present case. Moreover "there-is-an-obstacle-on-the-path" may attack directly "no-obstacle-on-the-path" by resorting to general knowledge provided by BT (Building Topology) agent, that, in the building, there are invisible obstacles (such as transparent glass walls).

Since "no-obstacle-on-the-path" is not able to reply to these arguments, "there-is-an-obstacle-on-the-path" prevails: the presence of an obstacle is accepted and CA intervenes to modify the motion plan, going around the glass wall and eventually reaching Mr. X.

5. CONCLUSION

The application of the proposed approach to AMR navigation has been described in order to show its potential in practical contexts.

Due to space limitations and for the sake of clarity, only the basic ideas of our approach and an intuitive and informal description of its application have been given. However, a (preliminary) complete formal description of all the presented concepts has been developed by the authors and a software implementation in C++ on a Sun workstation is in progress.

References

[1] T. Balch, G. Boone, T. Collins, H. Forbes, D. Mackenzie, J. C. Santama'na, Io, Ganymede, and Callisto A Multiagent Robot Trash-Collecting Team. *AI Magazine*, Summer 1995, 39-51.

[2] P. Baroni, G. Guida, S. Mussi, A. Vetturi, A Distributed Architecture for Control of Autonomous Mobile Robots. *Proc. of ICAR'95, 7th Int. Conf. on Advanced Robotics*, Sant Feliu de Guixols, SP, 1995, 869-877.

[3] R. A. Brooks, A Robust Layered Control System For A Mobile Robot. *IEEE Journal of Robotics and Automation*, RA-2. April 1986, 14-23.

[4] P. R. Cohen, H. J. Levesque, Intention is choice with commitment. *Artificial Intelligence* 42(3) (1990), 213-261.

[5] C. Isik, A. M. Meystel, Pilot Level of a Hierarchical Controller for an Unmanned Mobile Robot. *IEEE J. Robotics and Automation*, Vol. 4, no. 3, June 1988, 241-255.

[6] K. Konolige, ERRATIC Competes with the Big Boys. *AI Magazine,* Summer 1995, 61-67.

[7] R. Liscano, A. Manz, E. R. Stuck, R. E. Fayek, J. Tigli, Using a Blackboard to Integrate Multiple Activities and Achieve Strategic Reasoning for Mobile-Robot Navigation. *IEEE Expert*, April 1995, 24-36.

[8] M. Minsky, Steps Toward Artificial Intelligence. *Proc. IRE 49*, Jan. 1961, 8-30.

[9] J. F. Montgomery, A. H. Fagg, G. A. Bekey, The USC AFV-I. A Behavior-Based Entry in the 1994 International Aerial Robotics Competition. *IEEE Expert*, April 1995, 16-22.

[10] H. P. Nii, Blackboard Systems: The Blackboard Model of Problem Solving and the Evolution of Blackboard Architectures. *The AI Magazine,* Summer 1986, 38-53.

[11] R. Pfeifer, Cognition - Perspectives from autonomous agents. *Robotics and Autonomous Systems,* 15 (1995) 47-70.

[12] A. S. Rao, M. P. Georgeff, Modeling Rational Agents within a BDI-Architecture, *Proc. of KR&R-91 Int. Conf. on Knowledge Representation and Reasoning,* Cambridge, MA, 1991.

[13] Y. Shoham, Agent Oriented Programming. *Artificial Intelligence*, 60, 51-92, 1993.

[14] M.G. Slack, Planning paths through a spacial hierarchy: Eliminating stair-stepping effects. *Proc. SPIE Conference on Sensor Fusion,* 1988.

[15] L. Steels, When are robots intelligent autonomous agents?. *Robotics and Autonomous Systems,* 15 (1995) 3-9.

[16] C. E. Thorpe, Path relaxation: Path planning for a mobile robot. *Proc. 4th American National Conference on Artificial Intelligence*, Austin, TX, 1984, 318-321.