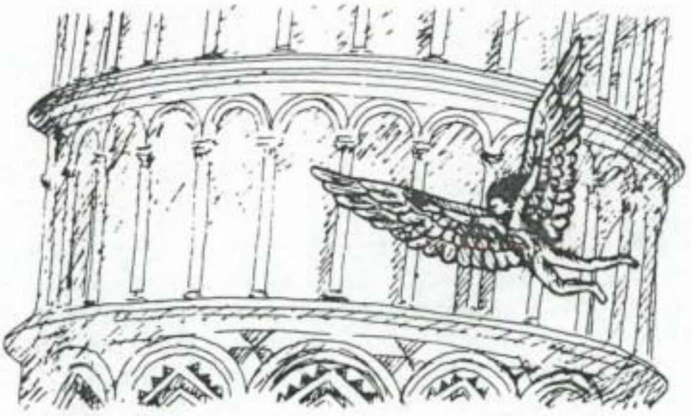

'91 I·C·A·R

Fifth International Conference on Advanced Robotics

Robots in Unstructured Environments

June 19-22, 1991
Pisa, Italy

Volume 2
91TH0376-4

1991 Fifth International Conference on Advanced Robotics

IEEE Catalog Number: 91TH0376-4

Library of Congress Number: 91-55202

ISBN Softbound: 0-7803-0078-5

Microfiche: 0-7803-0079-3

Additional copies of this publication are available from

IEEE Service Center
445 Hoes Lane
Piscataway, NJ 08854-4150

1-800-678-IEEE

Organisation of Robot Behaviour Through Genetic Learning Processes

Marco Dorigo
Politecnico di Milano
MP-AI Project
Dipartimento di Elettronica
Via Ponzio 34/5
20133 Milano
Italy
dorigo@ipmel1.polimi.it

Uwe Schnepf
Expert Systems Research Group
German National Research Center
for Computer Science (GMD)
P.O.Box 1240
5205 Sankt Augustin 1
Germany
usc@gmdzi.uucp

***Abstract*–Work in the field of artificial intelligence (AI) has led to the development of the so-called knowledge-based approach to build computational models of human intelligence. The main idea therein is that intelligence is based on symbol manipulation, or more explicitly, that thinking actually *is* symbol manipulation [1]. Behaviour-based robotics represents a different approach to modelling the interaction of an autonomous agent with its environment hence providing the basis for the development of cognitive capabilities in artificially intelligent systems. In this paper we present a machine learning approach based on genetic algorithms (GAs) and unsupervised reinforcement learning to the generation and organisation of robot behaviour. The implementation of an ethological model of behavioural organisation based on genetics-based machine learning will be outlined. Finally, the general implications of future work in behaviour-based robotics on the information processing paradigm currently dominating AI will be discussed.**

I. INTRODUCTION

Explaining the cognitive abilities of the brain purely in terms of symbol manipulation as in current AI implementations lacks the flexibility and expressiveness of natural cognitive systems. We believe that this is because concepts and the ability to verbally describe them are products of human cognition and not what cognition is actually based on. The notion of behaviour-based artificial intelligence has been put forward focusing on the assumption that cognitive capabilities actually require the ability to act and react robustly and flexibly in a dynamic environment. Hence building robots is a key issue in designing artificially intelligent systems [2]. Although we consider this approach to AI of great importance, we realise the deficiencies of current behaviour-based systems as far as the conceptual foundations of this approach are concerned [3]. In this paper, a framework is presented to build robot systems which are able to adapt themselves to the environment on the basis of reinforcement learning. We believe that an interesting approach to the realisation of this adaptive ability is the use of evolutionary algorithms. By means of these biologically inspired algorithms we let our robot develop at first very basic sensory-motor skills and later on a hierarchical structure in which some basic skills are grouped together to represent higher-level capabilities.

In the next section we overview genetic algorithms and the genetics-based machine learning paradigm (GBML). In Section 3 we present the general ideas about behaviour-based robotics and our own model of behavioural organisation. In Section 4 we propose a mapping of the GBML paradigm on the coordination learning problem and present our current work together with expected results. Finally, some directions for further research are sketched.

II. GENETIC ALGORITHMS, CLASSIFIER SYSTEMS AND GBML

Genetic algorithms are a computational device inspired to population genetics introduced by Holland[4]. They can be used either to optimize functions or as a rule-discovery device. We are here interested in their use as a rule-discovery system in the framework of genetics-based machine learning. A genetics-based machine learning system (also known as learning classifier system, LCS) is an adaptive system that learns by means of interaction with an environment. It belongs to the class of reinforcement learning systems in which learning is guided by rewards coming from the environment. The system is composed of three main modules (Fig.1): the performance subsystem, the apportionment of credit subsystem and the rule-discovery subsystem.

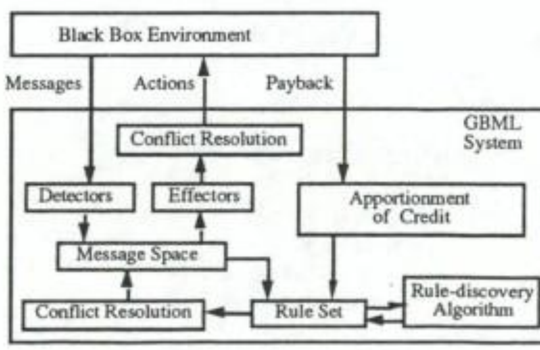

Fig. 1. Structure of a GBML system.

Manuscript received April 7, 1991. This work was partially supported by Italian National Research Council through grants PFR,2,Alpi and PFSI,2,PD.

7803-0078/91/0600-1456$01.00 ©1991 IEEE

The overall genetics-based machine learning system resulting from the interaction of these three subsystems is a parallel production system with the following peculiarities

- Rules are strings of bits over a three-valued alphabet (A={0,1,*}) with a condiction-action format.
- Rules fire in parallel.
- Rules are evaluated to find out their relative importance in the problem solving task to be learned.
- New rules are generated using existing rules which are considered to be the best by the evaluation mechanism.

The following subsections will shortly describe each of the three components.

A. The Performance Subsystem

The performance subsystem is composed of

- A set of rules, called classifiers.
- A list of messages, used to collect messages sent from classifiers and from the environment to other classifiers.
- An input and an output interface to the environment (detectors and effectors) to receive/send messages from/to the environment.
- A payback mechanism to reward/punish the system when it performs a useful/negative action.

The system works receiving messages from the environment and from other classifiers and uses them to activate internal rules (i.e. classifiers that send messages to other classifiers) and external rules (i.e. classifiers that send messages to effectors, causing actions and so modifying the environment). For a deeper discussion see [5].

B. The Apportionment of Credit Subsystem

The goal of the apportionment of credit subsystem is to give every classifier a value, called strength, proportional to its usefulness. In the context of genetics-based machine learning the most used algorithm is the bucket brigade (BB) in which a flow of payments (positive or negative, depending on the usefulness of the performed action) goes from the environment to external classifiers that caused the performed actions and from these external classifiers back to internal classifiers that caused the activation of external ones and so on until the beginning of the chain is reached. The expected result of the algorithm is the creation of sets of rules highly inter-related that perform useful actions and achieve high values of strength, increasing in this way the probability of being activated in presence of the right environmental and internal conditions (messages).

C. The Rule-Discovery Algorithm

The genetic algorithm is used in this context as a rule-discovery algorithm (see [5] or [6] for an introduction). Genetic algorithms are a class of stochastic algorithms that work modifying a population - or a set - of solutions (in GBML a solution is a classifier) to a given problem. Solutions are properly codified and a function, called fitness function, is defined to relate solutions to performance (i.e. to have some measure of the quality of the solution). In genetics-based machine learning the fitness of a classifier is given by its usefulness and so we can consider the apportionment of credit algorithm to be our fitness function. At every cycle a new population is created from the old one giving higher probability to reproduce (i.e. to be present again in the new population of solutions) to solutions with higher than average fitness. This new population is then modified by means of some genetic operators; the two most important genetic operators are crossover, which recombines individuals, and mutation that randomly changes some of the values of the genes constituting an individual. After these modifications are done the reproduction operator is applied again and the cycle repeated. There are results that show that this algorithm processes information contained in the structure of the solutions in a very efficient way (see for example [5]).

III. BEHAVIOUR-BASED ROBOTS

The classical approach to build a robot controller is to decompose the system to be developed into various subparts according to their functionality. Hence, there is a system for sensing, a system for situation interpretation, a system for choosing the appropriate actions according to the specified situations and a system to execute the actions chosen (see Fig. 2). Each of these subsystems is very complex too, using knowledge-based techniques to model the robot controller, plans etc.

Rodney Brooks's subsumption architecture [2] represents a successful attempt to build non-centralised, distributed robot controllers which are able to perform flexibly and robustly in an unstructured dynamic environment. The originality and radicalism of this approach lies in the decision to actually leave the main roads of traditional AI where explicit world modelling and reasoning techniques are considered to be a key to success. Although Brooks does not deny the value of knowledge representation and reasoning techniques for more complex tasks, he considers the use of these techniques as not appropriate and even damaging at the various levels of behavioural organisation tackled so far.

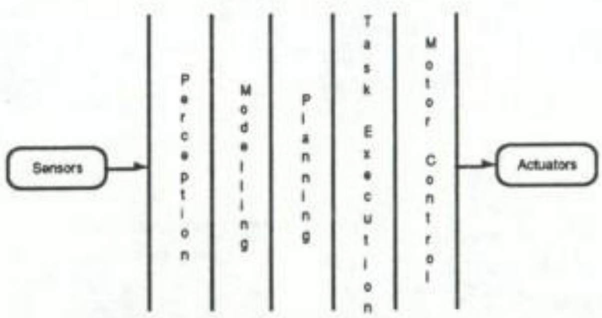

Fig.2. Vertical decomposition into functional modules.

Brooks suggested that, instead of building systems which intend to cover the complete range of human reasoning activities, one should follow an incremental path from very simple, but complete, systems to complex systems. The incrementality of this approach is constituted by a horizontal decomposition of tasks according to various "activities" in an autonomous agent (see Fig. 3) rather than a vertical decomposition based on functional entities. Each activity producing subsystem, or *behaviour*, is a complete system featuring sensing, reasoning and acting capabilities. Once such a behaviour has been implemented and tested, it remains unchanged and covers a particular aspect of the agent's complete behavioural repertoire.

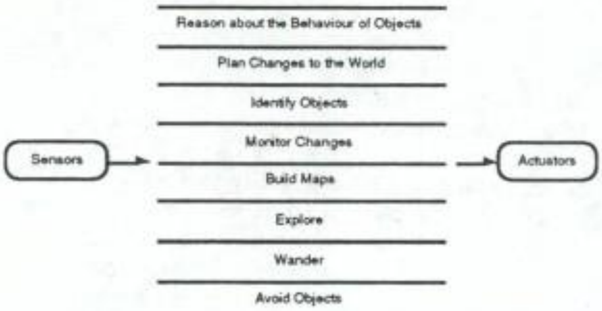

Fig. 3. Horizontal decomposition into competence modules.

By introducing several layers to enhance the agent's behavioural repertoire, it becomes necessary to define the appropriate organisation and ordering of the layers. To this end, Brooks used a hierarchically fixed organisation of the behaviours, building each new layer on top of already existing ones - the *subsumption architecture*. Each layer has been equipped with the ability to monitor and to influence the behaviour of the layer below it (see Fig. 4). The interaction of the competing layers determines the complexity of the overall behaviour of the autonomous system perceived by the observer. The control flow in the agent emerges from the changing activities of the distributed control flow resident in the various layers, but there is no central control involved.

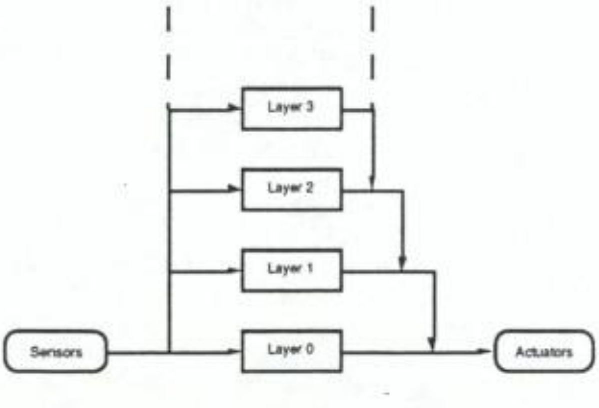

Fig. 4. The subsumption architecture.

Although there are several advantages in the behaviour-based approach to building intelligent autonomous systems, there currently exist some open questions in the design of behaviour-based robot controllers we would like to focus on:

- What is the appropriate design of a behaviour?
- What is the appropriate relation between various behaviours?
- Are there primitive or atomic behaviours?
- What is the conceptual relation between perception and action, i.e. what is the appropriate response to a particular sensor input?
- How can an agent use its existing behavioural repertoire to react to novel situations more flexibly?

As far as the behavioural organisation of a robot is concerned, these questions have to be answered in order to build flexible and robust robot systems.

A. Ethological Insights

Tinbergen [7] has developed a model of behavioural organisation in animals which we consider as one possible answer to the questions mentioned above. The Tinbergen model can be summarised as follows:

Each animal has a set of innate behaviours which enable it to react to the sensorial input from the environment. These behaviours are hierarchically organised in such a way that behaviours at higher levels are composed of behaviours at lower levels of the hierarchy. Behaviours at the same level of the behavioural organisation compete against each other in order to gain control over the animal. Each behaviour, or so-called *fixed action pattern*, is strictly related to a so-called *innate releasing mechanism*. Altogether, the fixed action pattern and the innate releasing mechanism form the so-called *instinct center*.

The mediation between the competing instinct centers at the same level of the hierarchy is performed by mutual inhibition. Each instinct center tries to prevent the other instinct centers at the same level from becoming active by sending inhibitory signals corresponding to its level of excitation.

Whether a fixed action pattern is released or not depends on the excitation and the threshold levels of the innate releasing mechanism. These levels depend on various parameters' values, such as the excitation from inner and outer sensors, hormones, excitation signals created within each instinct center, inhibitory signals from neighbouring centers, and additional excitation from instinct centers located at higher levels of the hierarchy. In this way, the releasing mechanisms act as a block which prevent the action patterns from gaining control over the animal randomly.

Once a fixed action pattern has been released it serves as a source of excitation for the various behaviours at the next lower level of the behavioural hierarchy. In the case that primitive action patterns are reached, the appropriate motor signals are released serving as input for nerve cells (e.g. in muscle groups, muscle strains, and motor neurons) and generate activities. This architecture is shown in Fig. 5.

But how is this behavioural organisation related to the emergence of intelligent behaviour or intelligence in general? As work in animal psychology [8] shows, the formation of concepts can already be observed in the behavioural organisation of simple animals such as fish. Although the concepts observable in fish behaviour are quite different from what we generally consider to represent a concept, evolutionary development has built on this abilities to create more intelligent systems and indicates that there is a direct link between the former and the latter.

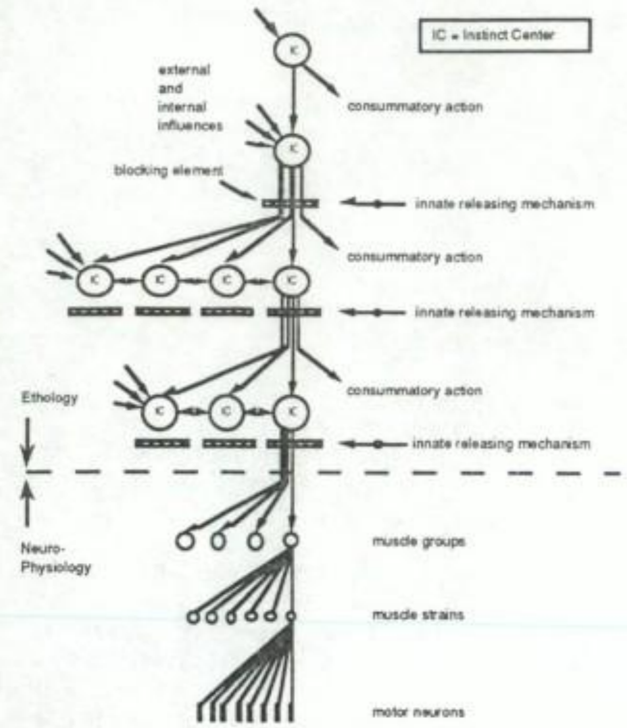

Fig.5. A hierarchy of instinct centers.

The general ability of an agent to adapt its behaviour to a dynamically changing environment and to "learn" is represented by self-organising processes based on the coincidence and correlation of motor and sensor activities. Both aspects, self-organisational processes to correlate useful and damaging patterns of behaviour to sensory input via environmental feedback, and evolutionary development of behavioural sequences, form the basic components of our computational approach to building autonomous systems. Hence the learning system to be developed has to fulfil two, very basic demands: First, it must have general learning capacities such as the ability to correlate arbitrary sensory input to useful behaviour, and the ability to coordinate behaviours as described by the Tinbergen model. Second, it has to incorporate evolutionary-based learning techniques such as the genetical learning of sensory-action-patterns, and the learning of competition or coordination patterns.

What has been described so far shows some great similarities to genetics-based machine learning as described in Section 2. The implementational model of the construction of the Tinbergen hierarchy using GBML techniques will be the subject of the following section.

IV. A HIERARCHICAL/PARALLEL APPROACH TO EVOLUTION-BASED REINFORCEMENT LEARNING (GBML*)

Our overall learning system will be composed of many classifier systems each one having as a learning goal the learning of a simple behaviour through interaction with its environment; the system as a whole has as a learning goal the coordination of activities.

A. The Programming Environment

Even at the lowest level of the described learning activities, i.e. the correlation between pieces of arbitrary sensory input and useful responses, the notions of disorder and dynamics play an important role, since only through environmental feedback the system assigns an internal *semantics* to the sensory input conceived. Hence the sensor invariants which serve to identify useful situations are constructed by the system in accordance with its necessities and change dynamically with those. For these reasons, we consider a real robot as absolutely necessary to study the emergence of adaptive behaviour.

We also need some simulation, but a simulation which is designed in close correlation to the real robot: since our learning approach is based on evolutionary techniques and starts from scratch as far as the robot's ability to react is concerned, we have to use simulation in order to generate the basic interaction abilities of the robot. Without simulation, the robot would possibly destroy itself while initially performing arbitrary actions the appropriateness of which has to be measured through the interaction. Additionally, simulations reduce the training cycle of classifier sets dramatically.

B. The Correlation of Sensor Stimuli and Useful Responses

The first task in the construction of useful robot behaviours consists in *designing* appropriate *action rules*, i.e. defining useful robot responses to given sensor inputs. This task will be performed by each individual classifier set.

The different classifier sets will be instantiated, i.e. will be set up and started, individually in accordance with new situations. In this way, each classifier set will learn a particular aspect of the robot interaction with the real world, e.g., passing through a door, getting out of a corner, following a corridor etc. Once a stable population of classifiers in a classifier set has been achieved during the learning procedure running in the simulation environment, the corresponding classifier set can be downloaded onto the robot. Assuming that future modifications of such a classifier set are minimal as the environment and the feedback model do not change dramatically, we expect to continue the adaptation process on the robot hardware, i.e. each classifier set will use the interaction with the real world to modify its classifiers permanently.

C. The Construction of the Behavioural Hierarchy

The next task consists in enabling the system to coordinate the behavioural modules in a more flexible way in order to achieve its demands or goals more efficiently. As outlined before, this coordination corresponds to the emergence of competing and coordinating activities between these behaviours. This can be achieved by the learning of messages which are to be exchanged between the various classifier sets. The learning procedure involved in these activities is similar to what has been explained before about the learning of appropriate system responses. But instead of learning robot activities, these classifier systems learn to send excitation and inhibition messages between each other. Additionally, a chaining of behavioural sequences can be performed using the same learning procedure. This corresponds to the most interesting aspect of using genetics-based machine learning: that we can achieve distributed problem solving by means of a mixture of the two most important ways of interaction of natural systems: competition and cooperation.

In this way, we expect the hierarchical behavioural organisation to have the following advantages over non-hierarchical ones:

- Emergent goal-directedness of the robot's overall behaviour. The learning of action chaining and excitation/inhibition signals contributes to the formation of useful behavioural sequences.
- Anticipation of future events or useful activities. As behaviours of a behavioural sequence support each other, initially active behaviours of such a chain give additional excitation to behaviours *located* at the end of this behavioural chain.
- Formation of concepts. As the behavioural modules at each level of the hierarchy higher than the basic ones correlate the various sensory invariants of their subordinate behaviours, they form more complex sensor invariants.

V. EXPECTED RESULTS

We expect this work to have some general implications on the information processing paradigm currently dominating AI (this applies to both, symbolic and connectionistic AI [8]). One could characterize this paradigm as the attempt to build cognitive architectures by mapping external information onto internal representation and performing appropriate reasoning steps on this representation.

Following the evolutionary-based approach to the development of behavioural organisation described so far, we hope to get a deeper understanding of cognition as a *construction* rather than a *matching process* due to the interaction of an autonomous agent with the real world. The clustering of information and the formation of concepts are not inherent to the sensor information available to an agent, but are constructed by the agent in accordance with its need to adapt in order to survive, and with the environment it is located in. These processes are dynamic and the feature detectors and concepts generated this way are subject to change if the environment or the agent's basic demands are changing.

VI. CONCLUSIONS

We have presented an approach to robotics that we consider an appropriate and innovative way to develop autonomous systems that perform tasks in a way that "looks as if they were intelligent". We have also introduced some very basic concept about genetic algorithms and their application to machine learning. We have claimed that our ultimate goal is to build robots which, starting from some designed but very simple skills, learn appropriate behaviours in order to survive in an unstructured environment. We have matched the behavioural organisation architecture with the learning paradigm in order to show how it could be possible to use evolutionary strategies to develop this kind of "intelligent" robot behaviour. First results obtained with a simulated robot are described in [9] and [10].

REFERENCES

[1] Harnad,St. "The symbol grounding problem", Physica D 42, p.335-346,1990.

[2] Brooks,R.A.,"Achieving artificial intelligence through building robots", Artificial Intelligence Memo 899, 1986, MIT, USA.

[3] Brooks,R.A.,"A robot that walks; emergent behaviours from a carefully evolved network", Artificial Intelligence Memo 1091, 1989, MIT, USA.

[4] Holland,J.H. "Adaptation in natural and artificial systems", 1975, Ann Arbor: The University of Michigan Press.

[5] Goldberg,D.E. "Genetic algorithms in search, optimization & machine learning", 1989, Addison-Wesley.

[6] Dorigo,M. "Genetic algorithms: The state of the art and some research proposal", Technical Report No. 89-058, Politecnico di Milano, Italy.

[7] Tinbergen,N., "The study of instincts", Oxford University Press, 1966.

[8] Kampis,G. "Perception, cognition, and models: Steps towards a non-symbolic paradigm", Evolutionary Systems Group, Department of Ethology, L.Eötvös University, 1990, Budapest, Hungary.

[9] Dorigo,M., Schnepf,U. "A bootstrapping approach to robot intelligence: First results", Technical Report No. 90-068, Politecnico di Milano, Italy, December 1990.

[10] Dorigo,M., Sirtori,E., "ALECSYS: A parallel laboratory for learning classifier systems", Proceedings of Fourth International Conference on Genetic Algorithms - July 13-16, 1991 - San Diego - California.