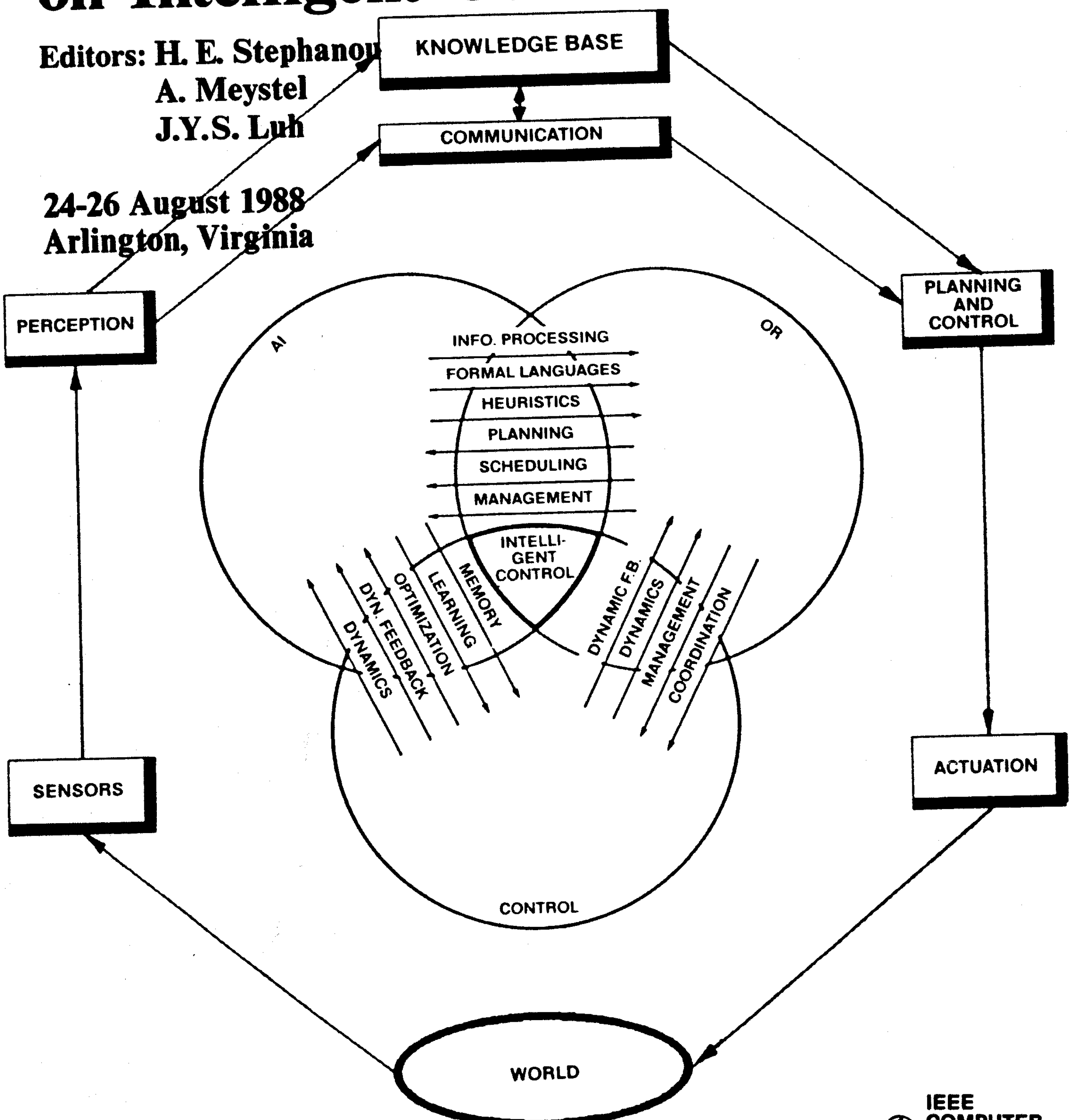
Proceedings

IEEE International Symposium on Intelligent Control 1988

Editors: H. E. Stephanou
A. Meystel
J.Y.S. Luh

24-26 August 1988
Arlington, Virginia



The Institute of Electrical and Electronics, Engineers, Inc.

IEEE COMPUTER SOCIETY PRESS

MACHINE LEARNING IN A DYNAMIC WORLD

Panel Discussion

Mieczyslaw M. Kokar
Panel Organizer and Editor
Northeastern University
Boston, Massachusetts

Panelists:

P. J. Atnsaklis, University of Notre Dame
K. A. DeJong, George Mason University
A. L. Meyrowitz, Office of Naval Research
A. Meystel, Drexel University
R. S. Michalski, George Mason University
R. S. Sutton, GTE Laboratories

Machine Learning was established as a research discipline in the 1970's and experienced a growth expansion in the 1980's. One of the roots of machine learning research was in Cybernetic Systems and Adaptive Control. Machine Learning has been significantly influenced by Artificial Intelligence, Cognitive Science, Computer Science, and other disciplines; Machine Learning has developed its own research paradigms, methodologies, and a set of research objectives different from those of control systems. In the meantime, a new field of Intelligent Control has emerged. Even though Intelligent Control adheres more closely to the traditional control systems theory paradigms – mainly quantitative descriptions, differential equations models, goal-oriented system design, rigid mathematical formulation of goals and models – it has also deviated from the traditional systems theory approach.

The two fields have moved forward without much interaction between them – different conferences, different journals, different researchers. Machine Learning has been concerned primarily with general learning mechanisms and methodologies for software implementation of learning systems. Intelligent Control has concentrated more on the dynamics of real physical systems and practical control problem solving. Because the two disciplines have at least one goal in common – automatic acquisition of knowledge about the world – they should have more interaction. The lack of interdisciplinary communication may lead to some undesirable results: establishing different terminologies for the same phenomena, repetitive work (discovering the same things independently), and lower quality research (ignoring the results established by the other discipline).

The goal of this panel was to analyze the interactions between Machine Learning and Intelligent Control. The panel consisted of several researchers both from the area of Intelligent Control and from Machine Learning.

The panelists were asked to concentrate on such general issues as:

- the need for the interaction,
- common research topics,
- common results,
- common methods.

The more specific topic of this panel was **machine learning and world dynamics.**

- Machine Learning devotes very little attention to the dynamics of real world. How does the lack of dynamics in the machine learning models affect the practical value of the machine learning results?

- Intelligent Control concentrates on formally specified problems while ignoring common-sense knowledge. How much of progress in controlling real-world systems can be made without taking into account all kinds of imprecise information?

In the following we present the views of the panelists. This presentation is based on the written material submitted to the panel organizers by the panelists. The presentation has been divided into four sections. The first section is devoted to the trends in the development of the two disciplines, Machine Learning and Intelligent Control. In the second section we present the panelists' views on the main paradigms within the two disciplines. The following section contains the common challenges to the both communities which occur as a result of the need to control complex dynamic systems. Finally, in the last section some similarities between the research issues and methods in the two disciplines are identified.

MAIN TRENDS IN MACHINE LEARNING and INTELLIGENT CONTROL

Antsaklis *Notes from the past.* In the 60's, adaptive control and learning received a lot of attention in the control literature. It was not always clear, however, what was meant by those terms. The comment by Y.Tsypkin (1971) describes quite clearly the atmosphere of the period, which, I should say, also has some striking similarities with the today's atmosphere:

"It is difficult to find more fashionable and attractive terms in the modern theory of automatic control than the terms of adaptation and learning. At the same time, it is not simple to find any other concepts which are less complex and more vague."

Adaptation, learning, self-organizing systems and control were competing terms for similar research areas, and K. S. Fu says (1970):

"The use of the word 'adaptive' has been intentionally avoided here ... adaptive and learning are behavior-descriptive terms, but feedback and self-organizing are structure, or system configuration, descriptive terms. Nevertheless the terminology war is still going on .... It is certainly not the purpose of this paper to get involved with such a war."

The term pattern recognition was also appearing together with adaptive, learning and self-organizing systems in the control literature of that era. It is obvious that there was no agreement as to the meaning of these terms and their relation.

Today, twenty or more years later, we have made some progress, at least in agreeing about the meaning of certain terms and we have come full cycle in the popularity of certain research areas. Certainly *pattern recognition* is today a research discipline in its own right, developing and using an array of methods ranging from conventional algorithms to artificial intelligence methods implemented via symbolic processing. The term *selforganizing* systems has almost disappeared from use in the control literature. *Adaptive control* has gained renewed popularity in the past decade mainly emphasizing studies in the convergence of adaptive algorithms and in the stability of adaptive systems; the systems considered are primarily systems described by differential (or difference) equations where the coefficients are (partially) unknown. In an attempt to enhance the applicability of adaptive control methods, *learning control* has been recently reintroduced in the control literature.

Evolution of control systems. I consider the introduction of learning in control as part of the continuing evolution of the control methods to address more complicated and demanding control problems.

Typically, control systems are dynamic systems and they involve feedback mechanisms. The system to be controlled, usually called the plant, and the decision making controller are distinct entities and they are both described by differential or difference equations. Conventional control systems are designed using mathematical models of physical systems. A mathematical model which captures the dynamical behavior of interest is chosen and then control design techniques are applied, aided by CAD packages, to design the mathematical model of an appropriate controller. The controller is then realized via hardware of software and it is used to control the physical system. The procedure may take several iterations. The mathematical model of the system must be "simple enough" so that it can be analyzed with available mathematical techniques, and "accurate enough" to describe the important aspects of the relevant dynamical behavior. It approximates the behavior of a plant in the neighborhood of an operating point. The first mathematical model to describe plant behavior for control purposes is attributed to J. C. Maxwell who in 1868 used differential equations to explain instability problems encountered with James Watt's flyball governor; the governor was introduced in 1769 to regulate the speed of steam engine vehicles. Control theory made significant strides in the past 120 years, with the use of frequency domain methods and Laplace transform in the 30's and 40's and the introduction of the state space methods in the 60's. Optimal control in the 50's and 60's, stochastic, robust and adaptive methods in the 60's to today, have made it possible to control more accurately significantly more complex dynamical systems than the original flyball governor. The need to achieve the demanding control specifications on increasingly complex dynamical systems has been addressed in the past by using more complex mathematical models, such as nonlinear and stochastic, and by developing more sophisticated design algorithms for, say, optimal control. Complex mathematical models, however, can seriously inhibit our ability to develop control algorithms, because of their mathematical complexity and the inability of existing mathematical methods to meet our growing needs. Fortunately, simpler plant models can be used in the control design, for example linear models, where well developed algorithms do exist. This is possible because of the feedback used in control and *fixed controllers* are designed so to guarantee stability, robustness, and performance. *Adaptive control* is used to attain the control objectives when the plant parameter variations are too large for the control objectives to be achieved via fixed controllers.

It is clear that the development of control theory has been driven by the need to attain increasingly demanding control objectives on more complex dynamical systems under increasing uncertainty in the plant and environment.

Meystel Learning Control Systems are expected of improving their operational behavior in real time as well as from operation to operation. These improvements have to be performed by the control system with no human involvement, i.e., autonomously. Learning autonomously - this is the ultimate capability of the control system, and the ultimate challenge for the researcher in the control area. But **learning what?** Different researchers in different time were offering different answers to this question.

Initially it was clear that the only thing we do not know is the external world. Thus, for K.S.Fu (1971) the term *learning controller* meant equipping the controller with a set of devices with human-like capabilities, i.e., pattern recognition and decision making. These devices were to learn the reality. For Y.Tsypkin (1971) the system of learning consisted of a preassigned set of operations, and the control inputs were to be learned: "Under the term learning in a system, we shall consider a process of forcing the system to have a particular response to a specific input signal (action) by repeating the input signals and then correcting the system externally." The subsequent two decades can be considered a period of clarification and establishing a new scientific paradigm which is suggested by *intelligent control.*

Meyrowitz Research in machine learning, control, and their integration is sponsored by the Office of Naval Research within the ONR Intelligent Systems Program. That Program has two primary components: Artificial Intelligence, where the objective is to understand automation and extension of human intellectual skills, and Robotics, where the concern is with understanding the design of intelligent sensor-based mechanical systems. The common element of intelligence creates a broad overlap of research interest across the two areas. As progress is made in artificial intelligence, we expect to see the discovery of automated techniques crucial to advanced aids to human decision making; at the same time, those techniques are likely to play an important role in the controlling software for intelligent robots.

The issues of automated inductive learning, reasoning by analogy, and scientific discovery are receiving special emphasis in a Knowledge Acquisition Accelerated Research initiative. The interest is not just in extending these areas individually, but also in better understanding their integration, and in deriving theoretical results on the limitations of automated capabilities for learning.

Some progress has been made in understanding how learning can contribute to robotic control. An important example is found in the research of Alberto Segre (previously at the University of Illinois, now at Cornell). In an experimental system called ARMS, Segre demonstrated the use of explanation-based learning in having a robotic system learn as an apprentice, i.e., observing humans solve assembly tasks and extracting for future reference solutions to such problems. One drawback to this work, noted by Segre himself, was the fact that ARMS solved problems only in a simulated environment. This essentially insulated the system from real-world concerns related to complexity and uncertainty.

Uncertainty can, in fact, arise in a variety of ways in the real world. The dynamic character of world events, the variable duration of actions and changes of world states, the lack of completely accurate world models, the lack of knowledge about consequences of actions, and the imperfection of sensors can all contribute to the uncertainty with which robotic systems must cope. Segre, as well as Gerald DeJong (University of Illinois), are among a number of researchers now attempting to automate learning capabilities for robots which must be controlled under uncertain conditions, and there is consequently a growing preference for experimentation with actual devices rather than with simulation.

Research in machine learning which will impact control must take into account not only the uncertainty of information available to a robot, but also the novel character of autonomous robotic systems in having to function of a continuing basis. This will require the ability to learn diverse kinds of knowledge. It will, moreover, require learning through observation, analysis, exploration, and experimentation as robots seek to take advantage of sensors used in an active way. There are deep research challenges in this regard, in terms of having robots exhibit generality in being able to cope with the broad spectrum of real world events, and having robots exhibit efficiency in learning and processing the great quantity of information to which they will be exposed.

PARADIGMS AND DEFINITIONS

Antsaklis Recently a number of papers dealing with learning control methods have appeared in the literature. It is quite unfortunate that there is no effort made in many of these works to relate to previous work on learning, either to the results which have appeared in the learning control literature of the 60's and early 70's or the machine learning literature; and consequently, they are bound to duplicate work done in the past and fall into similar traps. Another related unfortunate fact is that the term "learning" is being used by certain authors quite loosely to describe what was known before as a converging algorithm, or an adaptive algorithm. As a result, the real contributions to the theory of learning control in the recent literature appear to be very few indeed; the contributions are mainly task oriented with no real attempt made to address the applicability of the method and its limitations, or to identify the method in terms of the Machine Learning classifications. It should be said, however, that learning is achieved, in a certain sense, when an adaptive control algorithm is used to adapt the controller parameters so that stability is maintained. In this case the system learns and the knowledge acquired is the new values for the parameters. Note, however, that if later the same environmental changes occur again and the system is described by exactly the same parameters identified earlier, the adaptive control algorithm still needs to recalculate the controller, and perhaps the plant parameters, since nothing was kept in memory. So, in that sense the system has not learned. It has certainly learned what to do when a certain type of changes takes place. In particular, it has been told exactly what to do, that is it was given the adaptive algorithm, and this is knowledge acquisition by rote learning. The knowledge represented by the new values of the controller and the plant parameters, and the circumstances under which these values are appropriate, are not retained.

Lessons to be learned. Learning is becoming an integral part of the theory and methods to control a system. This is due to the increased requirements imposed on the controller and the autonomous capability expected from future control systems. It has been said that if we ignore the past, we are in danger of reinventing the wheel and then renaming it! It is important therefore to be aware of the developments which have taken place in the past and learn from them. Learning in control is an exciting area and it has a lot to contribute. It is up to us to identify the promising methods, to relate and compare them to existing methods in machine learning and to the earlier learning control literature. Only in this way we can make significant progress and establish an identity for the area of Learning Control.

Michalski Machine learning has been closely associated with the problems of automated control since the time adaptive systems were introduced. The interpretation of the term 'learning' has changed over time. We can distinguish at least two different interpretations; they are closely related to two different paradigms in the evolution of control theory. The two paradigms are:

- Conventional Control Theory, and
- Intelligent Control Theory.

In the conventional control theory paradigm everything (plant's models, environment, and the controller) needs to be defined in terms of parameters, variables, and equations. Learning within this paradigm was understood as modifying parameters of the plant's model and of the controller (adaptation).

Intelligent control theory is a blend of traditional control theory, artificial intelligence (AI), and operations research (OR). In this paradigm the control system attempts to represent and use knowledge. Learning means here creating knowledge structures, not just parameter adjustments.

In the same way as intelligent control is an extension of the traditional control theory, machine learning in the AI sense is an extension of the adaptive control approach. In other words, the AI approach to machine learning can be viewed as a superset of the adaptive systems approach.

Learning here is equated with building, modifying, or improving descriptions (see Michalski, 1986). The descriptions can be in the form of declarative statements, procedures, control algorithms, simulation models, or theories. The process of learning involves a source of information (either a teacher of the environment), and a learner (knowledge recorder).

Within the AI approach to machine learning we can distinguish three main learning paradigms:

- empirical learning,
- analytic learning, and
- constructive learning.

Empirical learning (see Figure 1) utilizes little of background knowledge. It is based on knowledge-poor inductive inference. The three directions within this paradigm are:

- symbolic learning,
- genetic algorithms, and
- connectionist systems.

Empirical learning creates new, hypothetical knowledge.

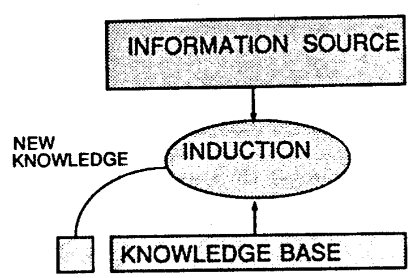


Figure 1. Empirical Learning

Analytic learning (Figure 2) requires large amounts of background knowledge. The basic inference type it uses is deduction. The learned knowledge is a new representation for the input information. It is not considered new as it is in the deductive closure of the knowledge base before the process of learning took place.

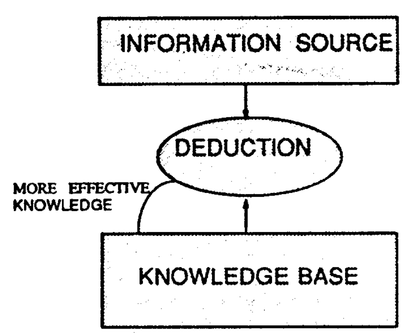


Figure 2. Analytical Learning

Constructive learning (Figure 3) is a goal-oriented method. The learning strategy depends on the relevance of the background knowledge to the task at hand. The inference can be either inductive or deductive. The created knowledge can be either new

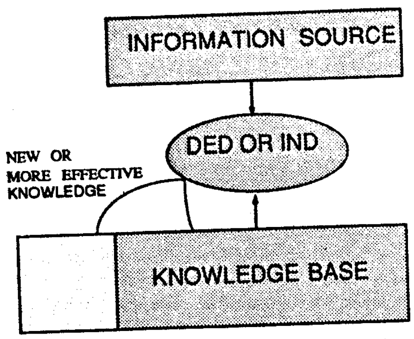


Figure 3. Constructive Learning

(if learned through induction) or more effective (if learned through deduction).

A control system with constructive learning capabilities is shown in Figure 4. The central unit of this system – inference engine – is connected to four other modules. The system interacts with the external world through sensors and actuators, where the process of perception takes place. The results of all kinds of inference are stored in the knowledge base. The knowledge base has also information about control goals. The other two functional units perform evaluation and selection of the most relevant information to the goal of learning.

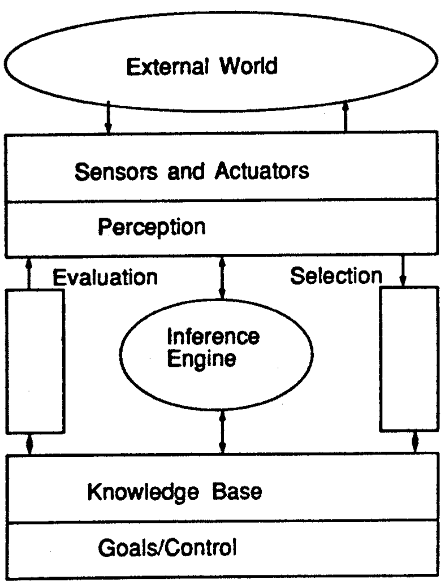


Figure 4. A Control System with Learning Capabilities

Sutton Ultimately, the problem of Artificial Intelligence comes down to making a sequence of decisions over time so as to achieve certain goals. AI is thus a control problem, at least in a trivial sense, but also in a deeper sense. This view is to be contrasted with AI's traditional view of itself, in which the central paradigm is not that of control, but of *problem solving* in the sense of solving a puzzle, playing a board game, or solving a word problem. Areas where the problem solving paradigm does not naturally apply, such as robotics and vision, have been viewed as outside mainstream AI. I think that the control viewpoint is now much more profitable than the problem solving one, and that control should be the centerpiece of AI and machine learning research.

If both AI and more traditional areas of engineering are viewed as approaches to the general problem of control, then why do they seem so different? In the 1950's and early 1960's these fields were not clearly distinguished. Pattern recognition, for example, was once a central concern of AI and only gradually shifted to become a separate specialized subfield. This happened also with various approaches to learning and adaptive control. I would characterize the split as having to do with the familiar dilemma of choosing between obtaining clear, rigorous results on the one hand, and exploring the most interesting, powerful systems one can think of on the other. AI clearly took the latter "more adventurous" approach, utilizing fully the experimental methodology made possible by digital computers, while the "more rigorous" approach became a natural extension of existing engineering theory, based the pencil-and-paper of theorem and proof (see Figure 1). This is not in any way to judge these fields.

The most striking thing indicated in the figure is not that some work was more rigorous and some more adventurous, but the depth of the gulf between work of these two kinds. Most AI work makes absolutely no contact with traditional engineering algorithms, and vice versa. Perhaps this was necessary for each field to establish its own identity, but now it is counterproductive. The hottest spot in both fields the one between them. The current enormous popularity of neural networks is due at least in part to its seeming to span these

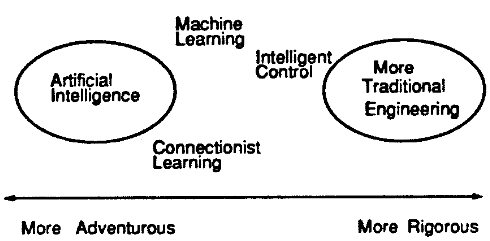


Figure 5. AI vs. More Traditional Engineering

two – the applications potential of rigorous engineering approaches and the enhanced capability of AI. Intelligent control is also in this position.

My conclusion then is that there is indeed a very fruitful area that lies more or less between Intelligent Control and Machine Learning (including connectionist or neural net learning), and which therefore presents an excellent opportunity for interdisciplinary research.

While I am advocating more interdisciplinary research, let me also warn of its pitfalls, in the hope that we can avoid them better than our predecessors. One meaning of "interdisciplinary" is "between disciplines". The problem of working *between* disciplines is that if you are not actually *in either* discipline you may end up with *no* discipline at all. That is, with no constraints and methodologies for measuring and ensuring progress. The best way to prevent this is for interdisciplinary work to fully accept some of the evaluative standards of its neighbor disciplines, to truly span them rather than fall between them. One of the best defenses is to maintain *early falsifiability*-some way of proving your interdisciplinary idea mistaken before you've had to learn all about a whole new discipline.

Meystel *Definition of Learning.* We can introduce the definition of *learning* which incorporates most of the views and approaches in this area (see Meystel, 1988). It blends the ideas of knowledge and skill acquisition and/or modification, change of behavior via change in the program, efficiency and quality of operation as a goal of learning, and achieving these changes in the structure of the controller as well as in the program of operation as a result of discovery. This definition is formulated as to be applied for the couple: Control System - Controlled System which is adequate to a multiplicity of technological problems, and allows for powerful interpretation in a number of practical cases (see Saridis, 1977).

Definition 1. Learning is a process based upon experience of operation, of development (modifying and/or restructuring) of representations and/or of algorithms of Control Systems, which provides their operation with better value(s) of the cost-functional(s) considered to be a subset of the (externally given) assignment for the Controlled System.

Mechanisms of Learning. We gave answers to the question of "learning what". Another question that arises is: **learning how?** Let us discuss some of the terms from Definition 1. *Development* consists of *modification* and *restructuring*. Development can be a part of design prior to system operation, and can be a part of normal system operation. Thus, it can be said that learning is a perpetual redesigning of a system. Modifying can be understood as parametric adjustment of these algorithms with no changes in their structure, (e.g., making corrections in the range of the rules without changing the rules). Cost-functional is considered to be assigned by the user who formulated the specifications. (In fact, a more complicated case can be considered when the user assigns only general policies, and lower level cost-functionals are formulated and then modified, and restructured by the system itself. We won't talk about this type of systems in this paper, they seem to be out of reach now, our immediate goal is to discuss learning control systems which can work under preassigned cost-functional. If these learning control systems can be realized then our effort will evolve).

Restructuring presumes not only modification of existing rules but also creation of the new rules, or even the new meta-rules (rules for the rules) which ends up with some rule hierarchy. This a more complex, more sophisticated case of learning. The highest level of learning takes place when the goals of the functioning could be reconsidered based on the results of operation of the subsystem of learning. The following analogy can be of importance: the system with learning is permanently undergoing the on-line process of design. The system with no learning has been designed in the past once and forever.

Both *learning via modifying* and *learning via restructuring* can be based upon two types of knowledge: knowledge obtained from an external source ("learning by being told"), and knowledge created within the learning system ("learning by discovery"). The external source of knowledge is evaluated by the degree of belief, and therefore it just substitutes for the internal subsystem which creates the knowledge within the learning system (and whose results of operation are also evaluated by the degree of belief). Thus, the systems with learning by being told can be considered a subset of systems with learning by discovery, in which the process of discovery is externalized.

Typology of Learning. The following different types of learning processes and schemes are visualized by specialists in psychology (Cagne, 1965) (they will be interpreted in a fashion meaningful within the framework of control systems theory).

Type 1. Signal Learning, their association by similarity, their discrimination by the lack of resemblance. Labeling the associations. Generalization upon the set of labels. Actually, development of the phenomenological part of *world representation* is presumed for this type of learning.

Type 2. Stimulus-response Learning. Causal relationships are being recorded at this stage and organized into a system, or a rule base.

Type 3. Associating units of "stimulus-response", labeling the associations. Generalization upon the set of labels.

Type 4. Associating labels (words): unifying the results of learning of Type 1 through 3 into a set of consistent models (e.g., mathematical models, analogical models, computational algorithms, and others). These models form a hierarchical system.

Type 5. Discriminating objects with the same labels (i.e., belonging to the same class). It is expected that the rules of discrimination can be formulated as generalized rules and can be applicable at each level of the hierarchy of labels.

Type 6. Concept learning: recognition of the general units existing within the joint system of world representation, and the causal relationships discovered within the system of world representation.

Type 7. Rule learning: judgments concerned with control are being stated. According to the system of representation and the concepts associated with this system, the system of rules can also be organized as ?a hierarchy.

Type 8. Problem solving using the rules collected within the system (see Type 7).

It would be instrumental to consider also the following definition of containing a condensed representation of the processes characteristic for Learning Control Systems.

Definition 2. Operation of Learning Control System is defined as a system of nested hierarchical generalizations performed over the redundant stored information about the current and/or prior experience of operation. This system of generalizations changes the world representation as well as the algorithms of control available for selection during the problem solving.

Problems with the conventional control theory paradigm. Conventional control was dealing with a broad variety of problems in a spectrum of devices starting with a speed regulator in the early steam engines, and ending with stabilizing a goal oriented group of the spacecrafts. However, the following problems are unequivocally considered to be difficult for solving in the paradigm of conventional control theory:

a). optimum control of nonlinear systems, because the models of the nonlinear systems are traditionally inconvenient for using the well established paradigm of linear control theory, nonlinear control theory is perpetually in its embryonic state, optimum control solutions cannot typically be found for important nonlinear systems;

b). optimum control of stochastic systems, because the models of stochastic systems presume knowledge of probabilistic parameters and characteristics of the systems which cannot be provided in practice of design;

c). control of multilink manipulators (either 6-DOF, or redundant ones), because the models of *plants* turn out to be so huge, so unencompassible that even the off-line solutions can make a predicament to a control engineer, not to talk about the on-line control which is really required;

d). control of redundant systems, or any other type of systems that lead to so called "ill-posed" problems, because in order to solve them one has to introduce a regularizing functional which is usually done based upon wishful assumptions rather than on information known about the system;

e). control of autonomous robots, because most of the information to be taken into account is not known at the beginning, cannot be supported by off-line solutions, require the on-line interpretation;

f). control of systems with multi-sensor feedback information, because the multi-sensor information must be integrated, which means that some conceptualizing activities are presumed;

g). on-line control of systems with incomplete initial knowledge of the model, and/or of the environment, because we do not know how to incorporate new knowledge of the world within the model of the plant, and/or the model of the controller.

In all these problems, neither the knowledge assumed at the stage of design could be considered complete and satisfactory, nor the process of design could be completed unless new knowledge would be additionally supplied. An opportunity to view all these systems in a different, unconventional way as systems with never ending design stage, has appeared as a result of the broad application of computers, and in particular, industrial computer systems equipped with a variety of transducers-sensors (see Meystel, 1985).

COMMON CHALLENGES

Antsaklis *Learning and the autonomous control system.* There are needs today that cannot be successfully addressed with the existing conventional control theory. They mainly pertain to the area of uncertainty. Control systems must perform well under significant uncertainties in the plant and the environment and thy must be able to compensate for system failures. To achieve this, the capacity for certain degree of autonomy is necessary. Such autonomous, intelligent behavior is a very desirable characteristic of the future, advanced control systems. To accomplish this, decision making abilities should be added to conventional control systems to meet the increased control requirements. The controller's capacity to learn from past experience clearly is an integral part of such autonomous intelligent controllers, as the ability to learn is one of the fundamental attributes of intelligent behavior. The study and computer modeling of learning processes in their multiple manifestations constitutes the subject matter of Machine Learning. The research area called *Machine Learning* developed quite independently from the control area in the past twenty years.

Where can learning be used in the control of systems? As it was already mentioned, learning plays an essential role in the autonomous control of systems (see Antsaklis, 1989). It appears that several types of learning, from rote learning to learning from observation and discovery, can be utilized there. My research group is reporting at this conference some interesting initial results in (Gao et. al.), where learning is introduced to enhance the adaptive control of a space antenna. There are of course many areas in control where learning can be used to advantage and these needs can be briefly classified as follows:

1. **Learning about the plant.** That is learning how to incorporate changes and then how to derive new plant models.

2. **Learning about the environment.** This can be done using methods ranging from passive observation to active experimentation.

3. **Learning about the controller.** For example, learning how to adjust certain controller parameters to enhance performance.

4. **Learning new design goals and constraints.**

Meystel This list can be enhanced by adding one synthetic domain which can include all of the above and more:

5. **Learning the control algorithms.** We always presume that *conceptually* we are capable of prescribing what to do and how to control. In practice, however, it turns out that our concepts evolve, and later we can see better the set of control rules to be applied. We can imagine a controller which *learns autonomously how to control*, is capable of arriving with new concepts of operation, and develops better algorithms of control which can better address the problems it encounters in the reality.

Fundamental Problems in the Area of Learning Control Systems. The area is somewhat different from the area of *adaptive control systems.* The usual goal of adaptive control systems is to achieve good performance when changes emerge either in the control system, or in the environment. The underlying premise of the adaptive control can be formulated as follows: any system is designed based upon a number of assumptions, and these assumptions do not take into account all realities of operation. The deviations from our assumptions are possible. So, the system should be capable of trimming itself properly to a realistic course of events, it should adjust to the deviations that emerge.

Finally, a desire appears to gradually improve the performance based upon the whole experience of operation: from task to task, taking into account the actual experience of operation. This is the **first problem** of the area of learning control systems evolving from the usual circle of problems in adaptive control: **learning control system parameters from the repetitive operation.** Certainly, from a simple memorization of the changes expected we quickly arrive to the set of techniques for the **world modeling**, which means that generalization and concept formation are the possible tools for solving the first fundamental problem of the area.

The Internal Model Principle (Wonham, 1976) states that control system must incorporate the model of the plant as well as the model of the external world, in order to provide a consistent operation of the system. This principle yields for a priori representation of the plant within the controller memory (or it implies the need for learning this representation). This representation should be control oriented: it should describe the plant only to a degree which is required by a particular control operation. Different operations require different models. This is why for Albus (1975) learning control meant collecting in the memory of the controller all available plant's experiences relevant to the set of procedures performed by the controller and eventually serving as a world representation.

This means that the learning controller can reflect only our current view of the plant's model (and the *exosystem* model). Either we know this model, or we assume it to be known. But what if at the beginning of the operation we did not know the model of the plant at all? At first, it sounds as an idle question. Why should we *not know* the model of the plant? Firstly, we never can know the real model of the plant, we are dealing always with approximations commensurate with our knowledge of the plant and with the apparatus of analysis at hand. Secondly, some plants have different models at different time. Thirdly, some models are so complicated that on-line communication is either impossible, or very expensive.

From the engineering point of view, it would be much better to have a universal modular controller which does not require any prior knowledge of the plant and allows for using it with any plant. It will incorporate the model of the plant after some initial working together with the plant: it will learn it. **Learning the model of the plant and the model of the exosystem** is the second major problem of the learning control systems area.

Learning control algorithms also can be improved from memorizing the results of repetitive operations and inferring upon these results. Indeed, the decision making scheme can be improved by observing how it operates in a multiplicity of cases (which could have been forgotten or otherwise overlooked at the stage of control algorithm design). Unlike in the previous case, not the external factors are to be analyzed and modeled, but the concepts of decision making in dealing with these external factors. Here we face the **third problem** of the area of learning control systems: **learning control concepts (and algorithms) from the repetitive operation.** We can see that the tools of generalization and concept formation are expected to be valid in dealing with this fundamental problem as well.

The fourth problem is looming in a multiplicity of applications. This problem is concerned with learning the goals of operation which are unknown at the beginning of the operation. It would be prudent, however, to address this problem in the framework of a separate analysis concerned with the situation in the area of control systems for unmanned intelligent machines.

Sutton We now turn to those areas of AI and Control where the common problems are more striking than the common results, and where an interdisciplinary approach might help.

The Symbol Grounding Problem. It is commonplace in AI systems to have symbols such as CLYDE, ELEPHANT, and GREY, and to represent knowledge as propositions over these symbols, e.g., (IS-A CLYDE ELEPHANT), (COLOR CLYDE GREY). Like the definitions in a dictionary, the knowledge in these systems is typically entirely self-referential; there is no sensor that can tell the system that an object is GREY. GREY in fact has no meaning to the system except insofar as it is a COLOR property of certain objects. Such a symbol is said to be ungrounded. Often, much higher level symbols are left ungrounded in AI systems, such as shape, functional descriptions, e.g., THING-THAT-CAN-BE-SAT-UPON, or structures, e.g., SUPPORTED-BY in the blocks world. Ungrounded symbols are perfectly adequate for some kinds of reasoning processes. However, for other purposes, for relating to the real world and real world problems, and particularly for connecting with introspections which are still a critical part of the AI methodology, the presence of so many ungrounded symbols has come to be recognized as a critical problem. Intelligent control approaches are complimentary here as they begin completely connected to the real world, with sensors and actuators, and only build up symbols from there. In AI systems there is often reasoning with symbolic labels for sensory properties, but no way of sensing then; in intelligent control systems it is often exactly the reverse.

The Perfect Model Disease. Ron Rivest has coined this term to describe an "illness" that AI (and, to a lesser extent, control theory) has had for many years and is only now beginning to recover from. The illness is the assumption and reliance upon having a perfect model of the world. In toy domains such as the blocks world, puzzle solving, and game playing this may be adequate, but in general of course it is not. Without a perfect model, everything becomes much harder–or at least much different–and so we have been reluctant to abandon the perfect model assumption. The alternative is to accept that our models of the world will always be incomplete, inaccurate, inconsistent, and changing. We will need to maintain multiple models, at multiple levels of abstraction and granularity, and at multiple time scales. It is no longer adequate to view imperfections and inconsistencies in our models as transients and to perform steady-state analysis; we must learn to work with models in which these imperfections will *always* be present. This means certainty equivalence approaches are not enough and dual control approaches are needed.

Control without Reference Signals. The dogma in control is to assume that some outside agency specifies a desired trajectory for the plant outputs in such a way that controls or control adjustments can be determined. For many problems, however, this is simply not appropriate. Consider a chess game. The goal is clearly defined, but in no sense does one ever have a desired trajectory for the game or the moves to be made. Suppose I want a robot to learn to walk bipedally. Producing target trajectories for the joint angles and velocities is a large part of the problem, a part which needs to be addressed by learning, not just by analysis and a priori specification. In my opinion, most real control problems are of this sort – in most cases it is natural to provide a specification of the desired result that falls far short of the desired trajectories usually assumed in conventional and adaptive control. This problem will become more and more common as we begin to consider imperfect and weak models, and particularly for systems with long-delayed effects of controls on goals. *Reinforcement learning* represents one approach to this problem (Mendel & McLaren, 1970; Sutton, 1984).

Knowledge Representation. Finally, both AI and control are stumbling on the difficult problem of how to create higher-level representations of their environments and actions. In pattern recognition, a form of this is called *feature extraction*, and in AI is called the *new terms* problem. Broadly speaking, the standard approach in both fields has been for the knowledge representation problem to be solved by the researcher, not the machine. That is a substantial barrier for intelligent control.

COMMON RESULTS

Sutton There are a number of existing results in AI and control theory with immediate relevance for the other field.

Dynamic Programming and A Search.* These two techniques have long been known to be closely related, if not identical. Nevertheless, the complete relationship remains obscure. More importantly, many results have been obtained independently for each technique. How many of these results carry over to the other field? Amazingly, such inter-relations remain almost completely unexplored, at least in the open literature. There are almost certainly important additional insights to be obtained for either of both fields by simply organizing their results from a unified perspective.

Back-propagation. Back-propagation is a connectionist (neural net) learning technique for learning real-valued nonlinear mappings from examples, that is for nonlinear regression (see Rumelhart, Hinton & Williams, 1986). Such a function has many possible uses in control – for learning nonlinear control laws, plant dynamics and inverse dynamics. The important thing is not back-propagation as a particular algorithm – it's clearly limited and will probably be replaced in the next few years – but the idea of a general structure for learning nonlinear mappings. This will remain of relevance to intelligent control.

Temporal-Difference Learning. This is a kind of learning specialized for predicting the long-term behavior of time series. It was first used in a famous early AI program, Samuel's checker player (Samuel, 1959), and since has been used in Genetic Algorithms (Holland, 1986) and in adaptive control in the role of a learned "critic" (Barto, Sutton & Anderson, 1983; Werbos, 1987). The basic idea is to use the change of temporal difference in prediction in place of the error in standard learning processes. Consider a sequence of predictions ending in a final outcome, perhaps a sequence of predictions about the outcome of a chess game, one made after each move, followed by the actual outcome. A normal learning process would adjust each prediction to look more like the final outcome, whereas a temporal-difference learning process would adjust each prediction to look more like the prediction that follows it (the actual outcome is taken as a final prediction for this purpose). If the classic LMS algorithm is extended in this manner to yield a temporal-difference algorithm, then, surprisingly, the new algorithm both converges to better predictions and is significantly simpler to implement (Sutton, 1988).

REFERENCES

Albus, J. (1975). A new approach to manipulator control: The cerebellar model articulation controller (CMAC). *Transactions ASME, Journal on Dynamic Systems, Measurements, and Control, vol. 97*, 220-227.

Antsaklis, P. J. (1989). Towards Intelligent Autonomous Control Systems: Architecture and Fundamental Issues. *Journal of Intelligent and Robotic Systems.*

Barto, A. G., Sutton, R. S., & Anderson, C. W. (1983). Neuronlike elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, and Cybernetics, 13*, 834-846.

Cagne, R. M. (1965). The Conditions of Learning. New York, NY: Holt, Reinehart and Winston.

Fu, K. S. (1970). Learning Control Systems – Review and Outlook. *IEEE Transactions on Automatic Control.*

Fu, K. S. (1971). Learning Control Systems and Intelligent Control Systems: An Intersection of Artificial Intelligence and Automatic Control. *IEEE Transactions on Automatic Control, vol. AC-16*, 70-72.

Gao, Z., et al. (1988). Learning for the Adaptive Control of Large Flexible Structures. In the Proceedings of this conference.

Holland, J. H. (1986). Escaping brittleness: The possibilities of general-purpose learning algorithms applied to parallel rule-based systems. In R. S. Michalski, J. G. Carbonell, and T. M. Mitchell (Eds.), *Machine Learning: An artificial intelligence approach, Vol. II*, Los Altos, CA: Morgan Kaufmann.

Mendel, J. M. & McLaren, R. W. (1970). Reinforcement-learning control and pattern recognition systems. In J. M. Mendel and K. S. Fu (Eds.), *Adaptive Learning and Pattern Recognition Systems*, New York, Academic Press.

Meystel, A. (1985). Baby-robot: On the Analysis of Cognitive Controllers for Robotics. *Proceedings of the IEEE Conference of Man and Cybernetics*, Tuscon, AZ, 327-332.

Meystel, A. (1988). Learning Control System: A Framework for the Theory. *Technical Report, Department of Electrical and Computer Engineering, Drexel University*, Philadelphia, PA.

Michalski, R.S. (1986). Understanding the Nature of Learning: Issues and Research Directions. In R.S.Michalski, J.G.Carbonell, and T.M.Mitchell (Eds.), *Machine Learning: An Artificial Intelligence Approach, vol. 2*, Morgan Kaufmann Publishers, Inc.

Michalski, R. S., and Ko, H. (1988). On the Nature of Explanation or Why Did the Wine Bottle Shutter. In *Proceedings of the Workshop on Explanation-Based Learning*, Stanford University.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. 1: *Foundations*. Cambridge, MA: MIT Press.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal on Research and Development, 3, 210-299.* Reprinted in E. A. Feigenbaum, & J. Feldman (Eds.), *Computers and thought.* New York: McGraw-Hill.

Saridis, G. N. (1977). Self-Organizing Control of Stochastic Systems, Marcel Dekker, Inc.

Sutton, R. S. (1984). *Temporal credit assignment in reinforcement learning.* Doctoral Dissertation, Department of Computer and Information Science, University of Massachusetts, Amherst.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning, 3*, 9-44.

Tsypkin, Y. (1971). Adaptation and Learning in Automatic Systems. New York, NY: Academic Press.

Werbos, P. J. (1987). Building and understanding adaptive systems: A statistical/numerical approach to factory automation and brain research. *IEEE Transactions on Systems, Man, and Cybernetics, 17*, 7-20.

Wonham, W. M. (1976). Towards an Abstract Internal Model Principle. *IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-6, No. 11*, 735-740.