M. D. Lemmon and P. J. Antsaklis, "Towards a Working Characterization of "Intelligent" Supervisory Control," Technical Report of the ISIS (Interdisciplinary Studies of Intelligent Systems) Group, No. ISIS-93-007, Univ of Notre Dame, November 1993.

Towards a Working Characterization of "Intelligent" Supervisory Control

Technical Report of the ISIS Group
at the University of Notre Dame
ISIS-93-007
November, 1993

Michael Lemmon and Panos J. Antsaklis Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556

Interdisciplinary Studies of Intelligent Systems

M. D. Lemmon and P. J. Antsaklis, "Towards a Working Characterization of "Intelligent" Supervisory Control," Technical Report of the ISIS (Interdisciplinary Studies of Intelligent Systems) Group, No. ISIS-93-007, Univ of Notre Dame, November 1993.

Preface

In May 1993, a task force was created to look into the area of Intelligent Control and define what is meant by the term. The task force was created at the invitation of the Technical Committee on Intelligent Control of the IEEE Control Systems Society, and its findings were aimed mainly towards serving the needs of that society.

After the task force was formed in May, a position paper representing a particular point of view was aired to "get the ball rolling". It certainly achieved that ! Views were exchanged over email and animated discussions were conducted off and on during the whole summer. A first outline of this report was sent to all members in late July. At the end of August a meeting took place at the 1993 International Symposium on Intelligent Control in Chicago and several task force members and non-members exchanged views on the subject. It became apparent that a consensus was emering. Participants of that meeting sent their comments in writing to all the task force members in September and a final report was drafted in October of 1993.

The document found within this technical report of the ISIS group is the position paper which was used to start the deliberations of the Intelligent Controls task force.

Panos J. Antsaklis and Michael Lemmon
November 10, 1993

M. D. Lemmon and P. J. Antsaklis, "Towards a Working Characterization of "Intelligent" Supervisory Control," Technical Report of the ISIS (Interdisciplinary Studies of Intelligent Systems) Group, No. ISIS-93-007, Univ of Notre Dame, November 1993.

Towards a Working Characterization of "Intelligent" Supervisory Control

M.D. Lemmon *and P.J. Antsaklis
Dept. of Electrical Eng
University of Notre Dame
Notre Dame, IN 46556, USA
(219)-631-8309
(Internet) lemmon@maddog.ece.nd.edu

May 5, 1993

Abstract

Supervisory control systems are systems in which the plant is controlled by a discrete event system called the supervisor. Such control systems are often referred to as *intelligent* since the supervisor's behaviour emulates the actions of human operators. This paper, however, argues that such emulation is insufficient for an intelligent system. It is argued that such systems are intelligent only if the system is able to interpret the meaning of a supervisor's actions without any external assistance. In other words, the interpretation of the supervisor's actions is supplied by the system internally. This viewpoint of intelligent systems is consistent with viewpoints on human cognition held by certain factions of the AI and cognitive psychology communities. The significance of this particular perspective on intelligent systems is that it provides a working characterization of intelligence which emphasizes the importance of adaptation and control in machine intelligence.

1 Introduction

Supervisory control refers to the use of discrete event systems (DES) in control. The DES supervisor issues logical directives which direct or supervise the plant's behaviour. Figure 1 provides a block diagram for such a supervisory control system. This figure shows the plant, the supervisor, and an interface connecting the plant and supervisor. The supervisor is, of course, a discrete event system. The plant can be either another DES or a continuous-state system. If the plant is a continuous-state system, then the combined system shown in figure 1 is called a hybrid dynamical system [Stiver 1992].

---

*The partial financial support of the National Science Foundation (IRI91-09298 and MSS92-16559) is gratefully acknowledged

M. D. Lemmon and P. J. Antsaklis, "Towards a Working Characterization of "Intelligent" Supervisory Control," Technical Report of the ISIS (Interdisciplinary Studies of Intelligent Systems) Group, No. ISIS-93-007, Univ of Notre Dame, November 1993.

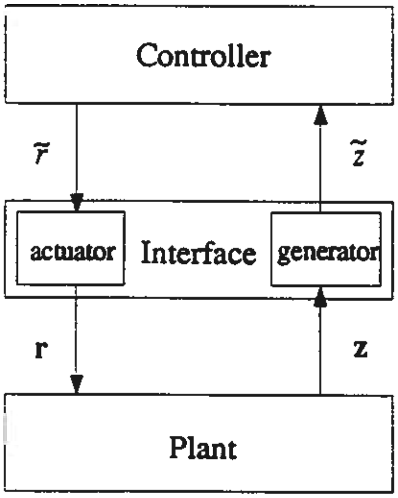

Figure 1: Block Diagram of Supervisory Control System

Supervisory control systems are often referred to as *intelligent* control systems because the actions of the supervisor are formulated in a way which mimics the actions of human-operated systems. In other words, control is based on high level decisions which are reminiscent of the way in which people control complex plants. The potential advantages of such an approach to control are a high degree of flexibility in the face of plant and environmental uncertainty. This means that the automated system exhibits a degree of autonomy rivaling that of human-operated systems [Antsaklis 1989].

2 Intelligence and Meaning

This notion of machine *intelligence*, however, is not entirely satisfying. At issue is the notion that mimicry of human decision making constitute intelligence. The traditional formulation of such supervisor controllers involve the assignment of meanings or interpretations to the symbols manipulated by the supervisor. Such interpretations allow us to form "explanations" [Pylyshyn 1984] for what the supervisor is attempting to do. In a temperature control system, for instance, a certain range of temperatures might be designated as "TOO HOT", thereby necessitating a control action to cool the entire system. The "intelligence" of the system is buried in its intepretation of the symbol "TOO HOT". But where does this association originate. In general, it is the system designer who determines what is meant by "TOO HOT", consequently it is not the system, but rather the system designer who is intelligent. If this is the case, then on what basis can we assert that straightforward symbol manipulation is indicative of an intelligent system?

This same fundamental argument has been leveled against production based inference as a model for human cognition. This argument was originally formulated as John Searle's famous Chinese room argument [Searle 1984]. In this thought experiment we are given a machine (room) into which Chinese ideograms are passed and out of which are passed responses (once again in the form of ideograms) to these inputs. Inside the machine is a native English speaker who has no

M. D. Lemmon and P. J. Antsaklis, "Towards a Working Characterization of "Intelligent" Supervisory Control," Technical Report of the ISIS (Interdisciplinary Studies of Intelligent Systems) Group, No. ISIS-93-007, Univ of Notre Dame, November 1993.

understanding of Chinese, but outputs the responses (in Chinese) with regard to a collection of rules. If the observer outside the room applies a Turing test and the "room" passes, does this mean the system is intelligent? Clearly the English speaker has no idea what he is doing and the natural impulse is to say no. In other words, "blind" manipulation of symbols is insufficient for characterizing an intelligent system. To accomplish such a characterization we need to know what is going on inside the room.

At the heart of Searle's complaint is the notion of a symbol's meaning. This problem is also referred to as the symbol grounding [Harnad 1990] or symbol binding problem. Symbol grounding refers to methods by which symbols of a formal system acquire semantic content or meaning. Such meaning can be categorized into two groups. Symbols can be either extrinsically or intrinsically grounded. The labels of extrinsic and intrinsic refer to the way in which symbol groundings are acquired. Meaning is acquired when the symbol is associated with some nonsymbolic item. For example, consider the symbol "TOO HOT" which was introduced above. The meaning of "TOO HOT" is a range of temperatures. If that range is determined by the system designer, then the acquisition of "meaning" is external to the system and hence we say that the symbol is extrinsically grounded.

Intrinsic grounding arises when the system assigns its own meaning to the symbol. For example, symbols can acquire content through internal associations with other symbols [Chalmers 1992]. To be more concrete, consider the preceding discussion of the symbol "TOO HOT". If the system has a way of assigning this label to a range of temperatures which is independent of an external supervisor, then we can say that the symbol is intrinsically derived or grounded. A good example of intrinsic grounding can be found in biological organisms. The label "TOO HOT" is derived from a pain avoidance mechanism which is built into the organism. Since the derivation of the symbol's meaning is internal to the system itself, it can be argued that the system (organism) understands what the symbol "TOO HOT" means. The advantage of this understanding, as will be discussed below, is that the system is able to adapt to environmental changes in a way which preserves the semantic content of the symbols its uses, thereby maintaining a high degree of independence (or autonomy) from external supervisors.

The relation of symbol meaning to intelligence is crucial to Searle's characterization of intelligent systems. In particular, the reason why the Chinese room was not "intelligent" was because the English speaking agent had no internal mechanism for assigning content to the ideograms which he was manipulating. The ideograms were not intrinsically grounded or rather the system has no understanding of the meaning of these symbols. Searle asserts that without such understanding there is no intelligence.

The preceding arguments therefore suggest the following characterization of intelligent systems.

> A necessary prerequisite for intelligence is the system's ability to dynamically assign and use symbol grounding without any external supervision.

The application of this characterization to the supervisory control system shown in figure 1 allows us to determine when such systems are truly intelligent. For example, consider a traditional supervisory control system in which the symbols used by the supervisor (i.e. the label TOO HOT) are fixed beforehand by the system designer. Such a supervisory control system is not intelligent because the symbols are externally grounded. If, however, the system determines its own meaning for the label TOO HOT, then we might have an intelligent system. System which can dynamically assign

M. D. Lemmon and P. J. Antsaklis, "Towards a Working Characterization of "Intelligent" Supervisory Control," Technical Report of the ISIS (Interdisciplinary Studies of Intelligent Systems) Group, No. ISIS-93-007, Univ of Notre Dame, November 1993.

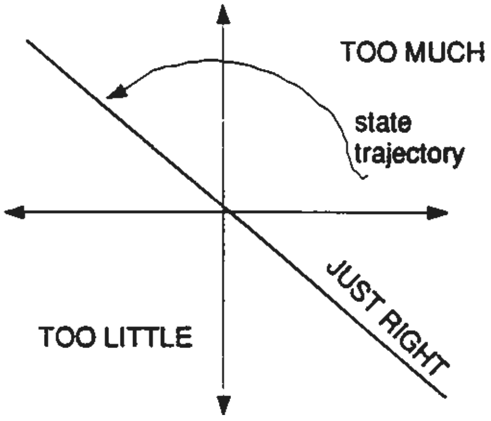

Figure 2: Example: symbols bindings for variable structure system

symbol bindings in this way are *adaptive* in nature. Therefore a necessary condition for intelligence is adaptation.

From this perspective, then, traditional supervisory control and perhaps most "intelligent" control schemes are not intelligent at all. Can we find a simple example of an "intelligent" supervisory control system. The following example, originally discussed in [Lemmon 1993] provides such an example which passes the characterization of intelligence given above.

3 Example

Consider a dynamical system whose plant dynamics are represented by the following differential equation

$$\dot{\mathbf{x}} = \begin{cases} f_1(\mathbf{x}) & \text{if } \mathbf{s}'\mathbf{x} > 0 \\ f_2(\mathbf{x}) & \text{if } \mathbf{s}'\mathbf{x} < 0 \end{cases} \tag{1}$$

Where $\mathbf{x} \in \Re^n$ and $\mathbf{s} \in \Re^n$. This system is a variable structure control [Utkin 1977] system. The hyperplane defined by the vector $\mathbf{s}$ is called a *switching surface*. The switching surface partitions the state space into three regions consisting of two halfspaces and the boundary between these halfspaces.

The two halfspaces generated by the switching surface, $\mathbf{s}$, are assigned the symbolic labels "TOO MUCH" and "TOO LITTLE" as is shown in figure 3. The closed set separating these regions is given the label "JUST RIGHT". This surface is determined by the switching surface associated with vector $\mathbf{s}$. The control objective is to keep the system state out of the regions TOO MUCH and TOO LITTLE and constrain it to the region JUST RIGHT. This can be done provided the surface characterized by vector $\mathbf{s}$ is an attracting invariant set of the system.

A switching surface which is invariant with respect to control strategies f_1 and f_2 is said to be a *sliding mode*. Because we require this surface to be invariant, we can use Lyapunov inequalities to test whether or not a given surface is "consistent" with the hypothesis that it is also a sliding mode. This test is give below

$$\mathbf{s}'\mathbf{x}\mathbf{s}'\dot{\mathbf{x}} < 0 \tag{2}$$

for all $\mathbf{x} \in \Re^n$ where $\dot{\mathbf{x}}$ is as given by equation (1). This test serves as a sufficient condition for $\mathbf{s}$ to be a sliding mode and this test can be incorporated into an algorithm which iteratively searches for an $\mathbf{s}$ which is a sliding mode. The specific form of this algorithm is found in [Lemmon 1993] and uses the central-cut ellipsoid method to search for the $\mathbf{s}$ which satisfies the inequality system. The details of the algorithm and its convergence properties are not discussed here. It has been shown that the procedure converges after a finite number of updates to a sliding mode, provided one exists. Of more interest to us here is the flow chart for the algorithm which has a direct bearing on our notion of adaptive symbol binding.

The algorithm of interest is an inductive inference protocol [Angluin 1983] consisting of the following steps. In [Lemmon 1993] it was called the invariant subspace identification or ISID algorithm. The flowchart of the algorithm is shown in figure 3. The steps of this algorithm are itemized below.

1. Form an *initial hypothesis* which asserts that $\mathbf{s}$ is a sliding mode for the system.
2. Perform an *experiment* which measures the system state $\mathbf{x}$ and system state velocity $\dot{\mathbf{x}}$.
3. Use inequality (2) as an *oracle* to test whether or not the experimental data is consistent with the hypothesis that the switching surface, $\mathbf{s}$, is a sliding mode. In other words, use the data in inequality (2) and see whether it is satisfied or not.
4. If the data is declared to be consistent with the hypothesis then do nothing and go to step 2.
5. If the data is declared inconsistent, then use the central cut ellipsoid algorithm [Bland 1981] to find a new $\mathbf{s}$ which is consistent with the current data and all past data.

We therefore see that the algorithm works as follows. It forms an initial hypothesis and then checks its consistency with the plant's current behaviour. If an inconsistency is determined, the vector $\mathbf{s}$ (i.e. the hypothesis) is changed. Note, however, that this hypothesis also determines the associations between the symbols TOO MUCH, TOO LITTLE and JUST RIGHT with specific subsets of the plant's state space. In other words, our system is adjusting symbol bindings with regard to the internal principle embodied in the inequality system given by equation (2). The Boolean functional represented by this inequality test is sometimes called an *oracle*. This basic cycle of experiment, oracle query, and update is then repeated until a solution is found. In [Lemmon 1993] it was shown that this procedure must terminate after a finite number of updates.

Given the above example system, we now apply our preceding characterization of "intelligence". Clearly the plant symbols TOO MUCH, TOO LITTLE, and JUST RIGHT are grounded with respect to specific subsets of the state space. Is this grounding extrinsic or intrinsic? This is determined by seeing what happens if the system or plant changes. Consider, for instance, what happens if the plant changes so that the initial s is not a sliding mode. The oracle will then detect this inconsistency and begin modifying the s. But as was discussed above, the s determines the symbol bindings and therefore our system is readjusting bindings with regard to the internal principle embodied by the oracle. Under this viewpoint, the above system passes the working characterization given above

M. D. Lemmon and P. J. Antsaklis, "Towards a Working Characterization of "Intelligent" Supervisory Control," Technical Report of the ISIS (Interdisciplinary Studies of Intelligent Systems) Group, No. ISIS-93-007, Univ of Notre Dame, November 1993.

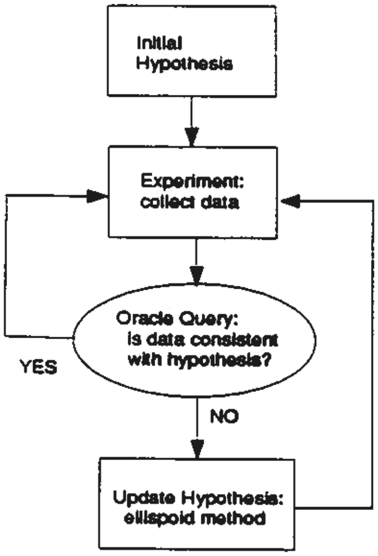

Figure 3: Flowchart for ISID Algorithm

and we can say that the symbols possess intrinsic semantic content. In other words this system is intelligent.

The intrinsic content embodied by the oracle is, of course, hardwired into the system. The choice of the oracle (i.e the linear inequalities) represents a choice by the system designer. There can be other oracle structures used, in which different internal event principles are used. Therefore in some sense it is still through the ingenuity of the designer that this system appears to possess intelligent processing. However, the fact remains that this system is assigning meaning without explicit external help. In other words, meaning is assigned internal to the system. This fact is true regardless of the origin of that internally realized principle (i.e. the oracle) guiding the binding process. Whether the internal principle comes from a designer, or emerges out of evolutionary forces (i.e. biological organisms), the fact remains that the system, once in possession of this organizational principle, is assigning meaning by itself. The organizational principle provides the basis for the system's understanding of a symbol and since understanding is a necessary condition for intelligence, such systems are intelligent.

4 Discussion

The preceding example of a so-called "intelligent" supervisory control systems prompts the following observations which are enumerated below.

- As noted above, a necessary condition for intelligent systems is that they be *adaptive*. Without adaptive symbol binding, the system cannot be using any internal mechanisms for interpreting the symbols.

M. D. Lemmon and P. J. Antsaklis, "Towards a Working Characterization of "Intelligent" Supervisory Control," Technical Report of the ISIS (Interdisciplinary Studies of Intelligent Systems) Group, No. ISIS-93-007, Univ of Notre Dame, November 1993.

- As argued above, ***intelligence is an internal property of the system not a behaviour.*** The immediate consequence of this observation is that intelligent systems cannot be determined by passive observation of behaviour. The Chinese room experiment is an example of a system which behaves intelligently, but is clearly not intelligent. Intelligence must be ***actively tested*** for. It is only in this way that an external observer can determine if the system is adaptively assign symbol bindings.

- The property of intelligence, while intellectually interesting, will only be important from an engineering perspective if its adds some functionality to the system not already present. Whether or not a symbolic system is intelligent can and is argued endlessly. The pragmatic reason for focusing on "intelligent" supervisory cntrol systems is that they endow systems with enhanced autonomy. Examining the anticipated applications of supervisory control, it is apparent that they are meant for complex and unpredictable systems. This means that "events" and their bound symbols may change unexpectedly. If this occurs, then it is well within the realm of possibility for our not-so intelligent supervisor to happily chunk away and produce a stream of nonsensical control directives. The reason this occurs is because the supervisor has no understanding of the significance or meaning of the symbols it is manipulating. Whether we choose to call this intelligent or not is somewhat irrelevant from an engineering perspective. The end result is the same, a system whose autonomy is circumscribed by an a priori and possibly ad hoc set of symbol bindings. The pragmatic solution to the problem is to allow the sytem to determine bindings itself with regard to a hardwired organizational principle it can check its performance against. In other words, for high autonomy we need a system which determines its own symbol meanings and as we noted above such systems are intelligent in the sense proposed by J. Searle.

- It has been argued that one aspect of intelligence is the ability to efficiently organize information. This perspective is not within the realm of our definition. While efficient knowledge representation is important, the necessity of such efficiency for "intelligence" is debatable. In this paper's view, intelligence is associated with meaning and understanding. This does not necessarily mean that efficiently organized systems are intelligent. Let's examine source coding theory. An important problem here is the efficient (i.e. smallest number of bits) representation of a message in a way which minimizes signal distortion. Does a coding scheme which achieves the Shannon bound also imply intelligence? No. There is little correlation in this case between the efficient representation of the message and the message's semantic content.

- Another issue concerns the relationship between control and intelligence. In other words how necessary is control for intelligence and vice versa. This paper takes the rather narrow view that intelligence is tied to intrinsic symbol binding. In the preceding example, we saw that such symbol binding can be done by associating symbols with dynamical ***invariants*** of the system. It might then be suggested that intelligent systems are systems which actively search for "invariants" in their sensory stream, where invariance is with respect to the system's behaviour. Another way of viewing this is that intelligent systems define symbol meanings in terms of their behavioural choices or actions. There is, in other words, a very strong link between intelligence and control which suggests that the latter is necessary for the former. If this is accepted then the term "intelligent control" has a significance which goes well beyond its being an umbrella term for a number of ad hoc decision based control schemes. Control represents a missing element in traditional AI's view of computational intelligence and explains why many facets of the robotics community are pushing the notion of active perception and subsummation or colony style robotic architectures.

The preceding discussion has introduced a perspective on machine intelligence in supervisory control which focuses on the way in which systems ground the meaning of symbolic objects. It was argued that a necessary condition for a system to be intelligent is that it ground its symbols using an internal mechanism which does not require assistance from a teacher outside of the system. An example of such a system was provided. It was noted that this notion of intelligence is consistent with viewpoints held by certain factions of the AI community. It was also noted that this viewpoint of intelligence leads to some interesting observations which shed considerable light on the nature of machine intelligence and supervisory control. These insights suggest that control and adaptation are necessary attributes of intelligent systems and that the role of intelligence in control is the enhancement of system autonomy. The objective of this short position paper was to raise these ideas and issues for consideration by the intelligent controls community.

References

[Angluin 1983] D. Angluin, C.H. Smith, "Inductive Inference: Theory and Methods." *Computing Surveys*, 15(3):237-269, September 1983.

[Antsaklis 1989] P. J. Antsaklis, K. M. Passino S. J. and Wang, "Towards Intelligent Autonomous Control Systems: Architecture and Fundamental Issues", *Journal of Intelligent and Robotic Systems*, Vol.1, pp. 315-342, 1989.

[Bland 1981] R.G. Bland, D. Goldfarb, M.J. Todd, "The Ellipsoid Method: a Survey", *Operations Research*, 29:1039-1091, 1981.

[Chalmers 1992] D. J. Chalmers, "Subsymbolic Computation and the Chinese Room", in *The Symbolic and Connectionist Paradigms: closing the gap*, ed: John Dinsmore, Lawrence Erlbaum Associates, pp. 25-48, 1992.

[Harnad 1990] S. Harnad, "The Symbol Grounding Problem", *Physica D*, vol. **42**, pp. 335-446, 1990.

[Lemmon 1993] M. D. Lemmon, "Inductive Inference of Invariant Subspaces", *Proceedings of the American Control Conference*, San Francisco, California, June 2-4, 1993.

[Pylyshyn 1984] Z. Pylyshyn, *Computation and cognition*, Cambridge, MA, Bradford/MIT Press, 1984.

[Searle 1984] J. R. Searle, *Minds, brains, and science*, Cambridge, MA: Harvard University Press, 1984.

[Stiver 1992] J. A. Stiver, P. J. Antsaklis, "Modeling and Analysis of Hybrid Control Systems", *Proceedings of the 31st Conference on Decision and Control*, pp. 3748-3751, Tucson AZ, December 1992.

[Utkin 1977] V.I. Utkin (1977), "Variable Structure Systems with Sliding Modes", *IEEE Transactions on Automatic Control*, Vol. AC-22:212-222.