INTELLIGENT FLIGHT CONTROL SYSTEMS

Robert F. Stengel*
Princeton University
Princeton, NJ USA

ABSTRACT

The capabilities of flight control systems can be enhanced by designing them to emulate functions of natural intelligence. Intelligent control functions fall in three categories. **Declarative** actions involve decision-making, providing models for system monitoring, goal planning, and system/scenario identification. **Procedural** actions concern skilled behavior and have parallels in guidance, navigation, and adaptation. **Reflexive** actions are spontaneous, inner-loop responses for control and estimation. Intelligent flight control systems learn knowledge of the aircraft and its mission and adapt to changes in the flight environment. Cognitive models form an efficient basis for integrating "outer-loop/inner-loop" control functions and for developing robust parallel-processing algorithms.

INTRODUCTION

Recounting personal experiences in confronting wind gusts, one of the Wright brothers wrote, "The problem of overcoming these disturbances by automatic means has engaged the attention of many ingenious minds, but to my brother and myself, it has seemed preferable to depend entirely on intelligent control" [1, 2]. The Wright brothers' piloting actions depended on proper interpretation of visual and inertial cues, demonstrating biological intelligent control. In the past, human pilots flew aircraft through manual dexterity, informed planning, and coordination of missions. As aircraft characteristics and technology have allowed, an increasing share of the aircraft's operation has come to rely on electro-mechanical sensors, computers, and actuators. Panel displays have enhanced decision-making, stability augmentation systems have improved flying qualities, and guidance logic has carried machine intelligence to the point of "hands-off" flying for much of a modern aircraft's mission.

In a contemporary context, intelligent flight control has come to represent even more ambitious plans to

- make aircraft less dependent on proper human actions for mission completion,
- enhance the mission capability of aircraft,
- improve performance by learning from experience,
- increase the reliability and safety of flight, and
- lower the cost and weight of aircraft systems.

This paper presents concepts for intelligent flight control through the aid of what were once called "artificial" devices for sensing, computation, and control. We distinguish between control functions according to a cognitive/biological hierarchy that is bounded on one end by declarative functions, which typically involve decision-making, and on the other by reflexive functions, which are spontaneous reactions to external or internal stimuli.

In a classical flight control context, declarative functions are performed by the control system's outer loops, and reflexive functions are performed by its inner loops. At an intermediate level, procedural functions -- like reflexive functions -- have well-defined input-output characteristics but of a more complicated structure. Traditional design principles suggest that the outer-loop functions should be dedicated to low-bandwidth, large-amplitude control commands, while the inner-loop functions should have high bandwidths and relatively lower-amplitude actions. There is a logical progression from the sweeping, flexible alternatives associated with satisfying mission goals to more local concerns for stability and regulation about a desired path or equilibrium condition.

FOUNDATIONS FOR INTELLIGENT FLIGHT CONTROL

Intelligent flight control design draws on two apparently unrelated bodies of knowledge. The first is rooted in classical analyses of aircraft stability, control, and flying qualities. The second derives from human psychology and physiology. The design goal is to find new control structures that are consistent with the reasons for flying aircraft, that bring flight control systems to a higher level of overall capability.

* Professor of Mechanical and Aerospace Engineering.

presented at the IMA Conference on Aerospace Vehicle Dynamics and Control, Cranfield Institute of Technology, Bedford, UK, September 7-10, 1992. (rev.: 9/14/92)

Supported by government grant. See Acknowledgments.
Aircraft Flying Qualities and Flight Control

An aircraft requires guidance, navigation, and control to perform its mission. As suggested by Fig. 1, a human pilot can interact with the aircraft at several levels, and his or her function may be supplanted by electro-mechanical equipment. The pilot performs three distinct functions: sensing, regulation, and decision-making. These tasks exercise different human characteristics: the ability to see and feel, the ability to identify and correct errors between desired and actual states, and the ability to decide what needs to be done next. The first depends on the body’s sensors and the neural networks that connect them to the brain. The second relies on motor functions enabled by the neuro-muscular system to execute learned associations between stimuli and desirable actions. The third requires more formal, introspective thought about the reasons for taking action, drawing on the brain’s deep memory to recall important procedures or data. Sensing and regulation are high-bandwidth tasks that allow little time for deep thinking. Decision-making is a low-bandwidth task that requires concentration. Each of these tasks exerts a workload toll on the pilot.

Pilot workload has become a critical issue as the complexity of systems has grown, and furnishing ideal flying qualities throughout the flight envelope has become an imperative. It is particularly desirable to reduce the need to perform high-bandwidth, automatic functions, giving the pilot time to cope with unanticipated or unlikely events. In the future, teleoperated or autonomous systems could find increasing use for missions that expose human pilots to danger.

Research on the flying (or handling) qualities of aircraft has identified ways to make the pilot’s job easier and more effective, and it provides models on which automatic systems might be based. The first flying qualities specification simply stated, "(the aircraft) must be steered in all directions without difficulty and all time (be) under perfect control and equilibrium" [3, 4]. Further evolution of flying qualities criteria based on dynamic modeling and control theory has led to the widely used U.S. military specification [5] and the succeeding military standard [6].

Flying qualities research has led to the development of control-theoretic models of piloting behavior. Most of these models have dealt with reflexive, compensatory tracking tasks using simple time-lag and transfer function models [7, 8] or linear-quadratic-Gaussian (LQG) optimal-control models [9, 10]. Some treatments go into considerable detail about neuro-muscular system dynamics [11, 12]. These models often show good correlation with experimental results, not only in compensatory tracking but in more procedural tasks: the progression of piloting actions from single- to multi-input strategies as the complexity of the task increases is predicted in [10], while test-pilot opinion ratings are predicted by a "Paper Pilot" in [13]. These results imply that computer-based control laws can perform procedural and reflexive tasks within the fit error of mathematical human-pilot models. Insight on the human pilot's declarative actions can be drawn from [14-16], which introduce the types of decisions that must be made in aerospace scenarios, as well as likely formats for pilot-vehicle interface.

Figure 1. Guidance, Navigation, and Control Structure.

Figure 1 also portrays a hierarchical structure for stability-augmentation, command-augmentation, autopilot, and flight-management-system functions that can be broken into reflexive and declarative parts. Stability augmentation is reflexive control provided by the innermost loop, typically implemented as a linear feedback control law that provides stability and improves transient response through an Estimation/Compensation block. Forward-loop control provides the shaping of inputs for satisfactory command response through a Control/Compensation block, again employing linear models. The combination of control and estimation can be used to change the flying qualities perceived by the pilot, or it can provide a decoupled system for simplified guidance commands [17-20]. A basic autopilot merely translates the human pilot's commands to guidance commands for constant heading angle, bank angle, or airspeed, while the Guidance block can be expanded to include declarative flight management functions, using inputs from Navigation sensors and algorithms.

Intelligent functions have been added to flight control systems in the past. Gain scheduling and switching improve performance in differing flight regimes and mission phases. Control theory, heuristics, and reduced-order optimization have been used to achieve near-optimal trajectory management in many flight phases (e.g., [21-23]). The Guidance, Navigation, and Control (GNC) Systems for Project Apollo's Command/Service and Lunar Modules provide an early example of intelligent aerospace control [24-26]. The state-of-the-art of aircraft flight control systems has progressed to comparable levels and beyond, as represented by systems installed in modern transport and fighter aircraft (e.g., [27, 28]).

Intelligent flight control1 can be justified only if it materially improves the functions of aircraft, if it saves

1 As used here "intelligent flight control" subsumes "intelligent guidance, navigation, and control."
the time and/or money required to complete a mission, or if it improves the safety and reliability of the system. Interesting philosophical problems can be posed. Must machine-intelligence be better than the human intelligence it replaces in order for it to be adopted? We are willing to accept the fact that humans make mistakes; if a machine has a similar likelihood of making a mistake, should it be used? Lacking firm knowledge of a situation, humans sometimes gamble; should intelligent machines be allowed to gamble? When is it acceptable for machine intelligence to be wrong (e.g., during learning)? Must the machine solution be "optimal," or is "feasible" good enough? Which decisions can the machine make without human supervision, and which require human intervention? In a related vein, how much information should be displayed to the human operator? Should intelligent flight control ever be fully autonomous? If the control system adapts, how quickly must it adapt? Must learning occur on-line, or can it be delayed until a mission is completed? All of these questions must be answered in every potential application of intelligent control.

Cognitive and Biological Paradigms for Intelligence

Intelligence is the "ability involved in calculating, reasoning, perceiving relationships and analogies, learning quickly, storing and retrieving information ..., classifying, generalizing, and adjusting to new situations" [29]. This definition does not deal with the mechanisms by which intelligence is realized, and it makes the tacit assumption that intelligence is a human trait. Intelligence relates not only to intellectuality and cognition but to personality and the environment [30].

The debate over whether-or-not computers ever will "think" may never be resolved, though this need not restrict our working models for computer-based intelligent control. One argument against the proposition is that computers deal with syntax (form), while minds deal with semantics (meaning), and syntax alone cannot produce semantics [31]. This does not limit the ability of a computer to mimic natural intelligence in a limited domain. Another contention is that thinking is "non-algorithmic," that the brain evokes consciousness through a process of natural selection and inheritance [32]. Consciousness is required for common sense, judgment of truth, understanding, and artistic appraisal, concepts that are not formal and cannot readily be programmed for a computer (i.e., they are not syntactic).

Conversely, functions that are automatic or "mindless" (i.e., that are unconscious), could be programmed, implying that computers have more in common with "unintelligent" functions. Gödel's Theorem [36] is offered in [33] as an example of an accepted proposition that may be considered non-algorithmic; the statement and proof of the theorem must themselves be non-algorithmic and, therefore, not computable. However, while the human curiosity, intuition, and creativity that led to Gödel's Theorem may not be replicable in a computer, the statement and proof are expressed in a formal way, so they might be considered algorithmic after all.

The notion that syntax alone cannot produce semantics is attacked as being an axiom that is perhaps true but not knowable in any practical sense [34]; therefore, the possibility that a computer can "think" is not ruled out. A further defense is offered in [35], which suggests that human inference may be based, in part, on inconsistent axioms. This could lead to rule-based decisions that are not logically consistent, that are affected by heuristic biases or sensitivities, that may reflect deeper wisdom, or that may be wrong or contradictory. For example, knowledge and belief may be indistinguishable in conscious thought; however, one implies truth and the other bias or uncertainty. One might also postulate the use of meta-rule bases that govern apparently non-algorithmic behavior. The process of searching a data base, though bound by explicit symbolic or numerical algorithms, may include randomized behavior (e.g., genetic algorithms) that are not immediately identifiable as algorithmic.

More to our point, it is likely that a computer capable of passing a flying-qualities/pilot-workload/control-theoretic equivalent of the Turing test [36] could be built even though that computer might not understand what it is doing. For intelligent flight control, the principal objective is improved control performance, that is, for improved input-output behavior. The computer can achieve the operative equivalent of consciousness on its own terms and in a limited domain, even if it does not possess emotions or other human traits.

Discussions of human consciousness naturally fall into using the terminology of computer science. It is convenient -- as well as consistent with empirical data -- to identify four types of thought: conscious, preconscious, subconscious, and unconscious [37]. Conscious thought is the thought that occupies our attention, that requires focus, awareness, reflection, and perhaps some rehearsal. Conscious thought performs declarative processing of the individual's knowledge or beliefs. It makes language, emotion, artistry, and philosophy possible. Unconscious thought "describes those products of the perceptual system that go unattended or unrehearsed, and those memories that are lost from primary memory through display or displacement" [37]. Within the unconscious, we may

---

2 As summarized in [32]: Any algorithm used to establish a mathematical truth cannot prove the propositions on which it is based. Or another [33]: Logical systems have to be fixed up "by calling the undecidable statements axioms and thereby declaring them to be true," causing new undecidable statements to crop up.

3 Turing suggested that a computer could be considered "intelligent" if it could "converse" with a human in a manner that is indistinguishable from a human conversing with a human.

4 Searle describes such a computer as a "Chinese Room" that translates Chinese characters correctly by following rules while not understanding the language in [31].
further identify two important components. *Subconscious thought* is procedural knowledge that is below our level of awareness but central to the implementation of intelligent behavior. It facilitates communication with the outside world and with other parts of the body, providing the principal home for the learned skills of art, athletics, control of objects, and craft. We are aware of perceptions if they are brought to consciousness, but they also may take a subliminal (subconscious) path to memory. *Preconscious thought* is pre-attentive declarative processing that helps choose the objects of our conscious thought, operating on larger chunks of information or at a more symbolic level. It forms a channel to long-term and implicit memory, and it may play a role in judgment and intuition.

Whether we adopt a single-processor model of consciousness such as Adaptive Control of Thought (ACT* as in [38]) or a connectionist model like Parallel Distributed Processing (PDP from [39]), we are led to believe that the central nervous system supports a hierarchy of intelligent and automatic functions with declarative actions at the top, procedural actions in the middle, and reflexive actions at the bottom. Declarative thinking occurs in the brain’s cerebral cortex, which accesses the interior limbic system for long-term memory [40]. Together, they provide the processing unit for conscious thought. Regions of the cerebral cortex are associated with different intellectual and physical functions; the distinction between conscious and preconscious function may depend on the activation level and duration in regions of the cerebral cortex.

The working memory of conscious thought has access to the spinal cord through other brain parts that are capable of taking procedural action (e.g., the brain stem for autonomic functions, the occipital lobes for vision, and the cerebellum for movement). Procedural action can be associated with subconscious thought, which supports voluntary automatic processes like movement and sensing. Voluntary signals are sent over the somatic nervous system, transmitting to muscles through the motor neural system and from receptors through the sensory neural system.

The spinal cord itself "closes the control loop" for reflexive actions long before signals could be processed by the brain. Nevertheless, these signals are available to the brain for procedural and declarative processing. We are all aware of performing some task (e.g., skating or riding a bicycle) without effort, only to waver when we focus on what we are doing. Involuntary regulation of the body’s organs is provided by the autonomic nervous system, which is subject to unconscious processing by the brain stem. "Bio-feedback" can be learned, allowing a modest degree of higher-level control over some autonomic functions.

Declarative, procedural, and reflexive functions can be built into a model of intelligent control behavior (Fig. 2). The *Conscious Thought* module governs the system by performing declarative functions, receiving information and transmitting commands through the *Subconscious Thought* module, which is itself capable of performing procedural actions. Conscious Thought is primed by *Preconscious Thought* [41], which can perform symbolic declarative functions and is alerted to pending tasks by Subconscious Thought. These three modules overlay a bed of deeper *Unconscious Thought* that contains long-term memory. They are capable of intellectual learning, and while their physical manifestation may be like the PDP model, they exhibit characteristics that are most readily expressed by the ACT* model.\(^5\)

The Subconscious Thought module receives information from the Sensory System and conveys commands to the Muscular System through peripheral networks. Voluntary Reflexive Actions provide low-level regulation in parallel with the high-level functions, responding to critical stimuli and coordinating control actions. High- and low-level commands may act in concert, or one may block the other. Voluntary Reflexive Actions can be trained by high-level directives from Subconscious Thought, while the learning capabilities of involuntary Reflexive Action are less clear. Control actions produce Body motion and can affect an external Controlled System, as in piloting an aircraft. In learned control functions, Body motion helps internalize the mental model of Controlled System behavior. The Body and the Controlled System are both directly or indirectly subjected to Disturbances; for example, turbulence would affect an aircraft directly and the pilot indirectly. The Sensory System observes External Events as well as Body and Controlled System motions, and it is subject to Measurement Errors.

There are many parallels and analogies to be drawn in comparing the functions of human and computer-based intelligence. It may be useful to ponder a few, especially those related to knowledge acquisition, natural behavior, aging, and control. Perhaps the most important observation is that learning requires error or incompleteness. There is nothing new to be gained by observing a process that is operating perfectly. In a control context, any operation should be started using the best available knowledge of the process and the most complete control resources. Consequently, learning is not always necessary or even desirable in a flight control system. *Biological adaptation is a slow process*, and proper changes in behavior can be made only if there is prior knowledge of alternatives. If adaptation occurs too quickly, there is the danger that misperceptions or disturbance effects will be misinterpreted as parametric effects. *Rest is an essential feature of intelligent biological systems*. It has been conjectured that REM (rapid-eye movement) *Sleep* is a time of learn-

---

5 ... although the actual processing mechanism is not clear. In a recent seminar at Princeton (March 9, 1992), Herbert Simon noted that if you open the cabinet containing a sequential-processing computer, the innards look very much like those of a parallel processor.
ing, consolidating, and pruning knowledge\(^6\) [42]. Systems can learn even when they are not functioning by reviewing past performance, perhaps in a repetitive or episodic way.

![Diagram of Cognitive/Biological Control Behavior](image)

**Figure 2. A Model of Cognitive/Biological Control Behavior.**

The cells of biological systems undergo a continuing birth-life-death process, with new cells replacing old; nature provides a means of transmitting genetic codes from cell to cell. Nevertheless, the central nervous system is incapable of functional regeneration. Once a portion of the system has been damaged, it cannot be replaced, although redundant neural circuitry can work around some injuries. Short-term memory often recedes into long-term memory, where it generally takes longer to be retrieved. With time, items in memory that are less important are forgotten, possibly replaced by more important information; hence, information has a half-life that depends upon its significance to our lives (and perhaps to its "refresh rate"). Humans develop the capability to form chords of actions that are orchestrated or coordinated to achieve a single goal. Response to an automotive emergency may include applying the brakes, disengaging the clutch, steering to avoid an obstacle, and bracing for impact all at once. We develop "knee-jerk" reactions that combine declarative, procedural, and reflexive functions, like clapping after the last movement of a symphony.

Nature also provides structural paradigms for control that are worth emulating in machines. First, there is a richness of sensory information that is hard to fathom, with millions of sensors providing information. This results in high signal-to-noise ratio in some cases, and it allows symbolic/image processing in others. Those signals requiring high-bandwidth, high-resolution channel capacity (vision, sound, and balance) have short, dedicated, parallel runs from the sensors to the brain. This enhances the security of the channels and protects the signals from noise contamination. Dissimilar but related sensory inputs facilitate interpretation of data. A single motion can be sensed by the eyes, by the inner ear, and by the "seat-of-the-pants" (i.e., by sensing forces on the body itself), corroborating each other and suggesting appropriate actions. When these signals are made to disagree in moving-cockpit simulation of flight, a pilot may experience a sense of confusion and disorientation.

There are hierarchical and redundant structures throughout the body. The nervous system is a prime example, bringing inputs from myriad sensors (both similar and dissimilar) to the brain, and performing low-level reasoning as an adjunct. Many sensing organs occur in pairs (e.g., eyes, ears, inner ears), and their internal structures are highly parallel. Pairing allows graceful degradation in the event that an organ is lost. Stereo vision vanishes with the loss of an eye, but the remaining eye can provide both foveal and peripheral vision, as well as a degree of depth perception through object size and stadiametric processing. Our control effectors (arms, hands, legs, feet) also occur in pairs, and there is an element of "Fail-Op/Fail-Op/Fail-Safe" design [43] in the number of fingers provided for manual dexterity.

**Structure for Intelligent Flight Control**

The preceding section leads to a control system structure that overlays the cognitive/biological model of Fig. 2 on the flight control block diagram of Fig. 1 and adds new functions. The suggested structure (Fig. 3) has super-blocks identifying declarative, procedural, and reflexive functions; these contain the classical GNC functions plus new functions related to decision-making, prediction, and learning. The black arrows denote information flow for the primary GNC functions, while the gray arrows illustrate the data flow that supports subsidiary adjustment of goals, rules, and laws.

Within the super-blocks, higher-level functions are identified as conscious, preconscious, and subconscious attributes, not with disregard for the philosophical objections raised earlier but as a working analog for establishing a computational hierarchy. The new functions relate to setting or revising goals for the aircraft's mission, monitoring and adjusting the aircraft's systems and subsystems, identifying changing characteristics of the aircraft and its environment, and applying this knowledge to modify the structures and parameters of GNC functions.

---

\(^6\) "In REM Sleep, the brain is barraged by signals from the brain stem. Impulses fired to the visual cortex produce images that may contain materials from the day's experiences, unsolved problems, and unfinished business." [42]
The suggested structure has implications for both hardware and software. Declarative functions are most readily identified with single-processor computers programmed in LISP or Prolog, as decision-making is associated with list processing and the statement of logical relationships. Procedural functions can be conceptualized as vector or "pipelined" processes programmed in FORTRAN, Pascal, or C, languages that have been developed for numerical computation with subroutines, arrays, differential equations, and recursions. Reflexive functions seem best modeled as highly parallel processes implemented by neural networks, which apply dense mappings to large masses of data almost instantaneously. Nevertheless, parallel processes can be implemented using sequential processors, and sequential algorithms can be "parallelized" for execution on parallel processors. The choice of hardware and software depends as much on the current state-of-the-art as on the closeness of computational requirements and GNC functions.

In the remainder of the paper, declarative, procedural, and reflexive control functions are discussed from an aerospace perspective. In practice, the boundaries between mission tasks may not be well defined, and there is overlap in the kinds of algorithms that might be applied within each group. A number of practical issues related to human factors, system management, certifiability, maintenance, and logistics are critical to the successful implementation of intelligent flight control, but they are not treated here.

DEclarative Systems

Goal planning, system monitoring, and control-mode switching are declarative functions that require reasoning. Alternatives must be evaluated, and decisions must be made through a process of deduction, that is, by inferring answers from general or domain-specific principles. The inverse process of learning principles from examples is induction, and not all declarative systems have this capability. Most declarative systems have fixed structure and parameters, with knowledge induced off-line and before application; declarative systems that learn on-line must possess a higher level of reasoning ability, perhaps through an internal declarative module that specializes in training.

Expert Systems

Expert Systems are computer programs that use physical or heuristic relationships and facts for interpretation, diagnosis, monitoring, prediction, planning, and design. In principal, an expert system replicates the decision-making process of one or more experts who understand the causal or structural nature of the problem [44]. While human experts may employ "nonmonotonic reasoning" and "common sense" to deduce facts that apparently defy simple logic, computational expert systems typically are formal and consistent, basing their conclusions on analogous cases or well-defined rules.

---

7 Expert systems can have tree or graph structures. In the former, there is a single root node, and all final (leaf) nodes are
A rule-based expert system consists of data, rules, and an inference engine [46]. It generates actions predicated on its data base, which contains measurements as well as stored data or operator inputs. An expert system performs deduction using knowledge and beliefs expressed as parameters and rules. Parameters have values that either are external to the expert system or are set by rules. An "IF-THEN" rule evaluates a premise by testing values of one or more parameters related by logical "ANDs" or "ORs," as appropriate, and it specifies an action that set values of one or more parameters.

The rule base contains all the cause-and-effect relationships of the expert system, and the inference engine performs its function by searching the rule base. Given a set of premises (evidence of the current state), the logical outcome of these premises is found by a data-driven search (forward chaining) through the rules. Given a desired or unknown parameter value, the premises needed to support the fixed or free value are identified by a goal-directed search (backward chaining) through the rules. Querying (or firing) a rule when searching in either direction may invoke procedures that produce parameter values through side effects [47].

Rules and parameters can be represented as objects or frames using ordered lists that identify names and attributes. Specific rules and parameters are represented by lists in which values are given to the names and attributes. The attribute lists contain not only values and logic but additional information for the inference engine. This information can be used to compile parameter-rule-association lists that speed execution [48]. Frames provide useful parameter structures for related productions, such as analyzing the origin of one or more failures in a complex, connected system [49]. Frames possess an inheritance property; thus a particular object lays claim to the properties of the object type.

Crew/Team Paradigms for Declarative Flight Control

Logical task-classification is a key factor in the development of rule-based systems. To this point, we have focused on the intelligence of an individual as a paradigm for control system design, but it is useful to consider the hypothetical actions of a multi-person aircraft crew as well. In the process, we develop an expert system of expert systems, a hierarchical structure that reasons and communicates like a team of cooperating, well-trained people might. This notion is expanded in [50-53]. The Pilot's Associate Program initially focused on a four-task structure and evolved in the direction of the multiple crew-member paradigm [54-56].

AUTOCREW is an ensemble of nine cooperating rule-based systems, each figuratively emulating a member of a World War II bomber crew: executive (pilot), copilot, navigator, flight engineer, communicator, spoofer (countermeasures), observer, attacker, and defender (Fig. 4) [53]. The executive coordinates mission-specific tasks and has knowledge of the mission plan. The aircraft's human pilot can monitor AUTOCREW functions, follow its suggestions, enter queries, and assume full control if confidence is lost in the automated solution. The overall goal is to reduce the pilot's need to regulate the system directly without removing discretionary options. AUTOCREW contains over 500 parameters and over 400 rules.

Figure 4. AUTOCREW Configuration with Pilot/Aircraft Interface (adapted from [52]).

AUTOCREW was developed by defining each member expert system as a knowledge base, according to the following principles:

- Divide each knowledge base into task groups: time-critical, routine, and mission-specific.
- Order task groups from most to least time-critical to quicken the inference engine's search.
- Break major tasks into sub-tasks according to need for communicating system functions.
- Identify areas of cooperation between knowledge bases.

The five main task groups for each crew member are:

- Time-critical
- Routine
- Mission-specific
- Other
Reasoning Under Uncertainty

Rule-based control systems must make decisions under uncertainty. Measurements are noisy, physical systems are subject to random disturbances, and the environment within which decisions must be made is ambiguous. For procedural systems, the formalism of optimal state estimation provides a rigorous and useful means of handling uncertainty [63]. For declarative systems, there are a number of methods of uncertainty management, including probability theory, Dempster-Shafer theory, possibility theory (fuzzy logic), certainty factors, and the theory of endorsements [64].

Bayesian belief networks [65], which propagate event probabilities up and down a causal tree using Bayes's rule, have particular appeal for intelligent control applications because they deal with probabilities, which form the basis for stochastic optimal control. We have applied Bayesian networks to aiding a pilot who may be flying in the vicinity of hazardous wind shear [66]. Figure 5 shows a network of the causal relationships among meteorological phenomena associated with microburst wind shear, as well as temporal and spatial information that could affect the likelihood of microburst activity. A probability of occurrence is associated with each node, and a conditional probability based on empirical data is assigned to each arrow. The probability of encountering microburst wind shear is the principal concern; however, each time new evidence of a particular phenomenon is obtained, probabilities are updated throughout the entire tree. In the process, the estimated likelihood of actually encountering the hazardous wind condition on the plane's flight path is refined.

The safety of aircraft operations near microburst wind shear will be materially improved by forward-looking Doppler radar, which can sense variations in the wind speed. Procedural functions that can improve the reliability of the wind shear expert system include extended Kalman filtering of the velocity measurements at incremental ranges ahead of the aircraft [67].

Probabilistic reasoning of a different sort has been applied to a problem in automotive guidance that may have application in future Intelligent Vehicle/Highway Systems [68-70]. Intelligent guidance for headway and lane control on a highway with surrounding traffic is based on worst-plausible-case decision-making. It is assumed that the intelligent automobile (IA) has imaging capability as well as on-board motion sensors; hence, it can deduce the speed and position of neighboring automobiles. Each automobile is modeled as a simple discrete-time dynamic system, and estimates of vehicle states are propagated using extended Kalman filters [63]. There are limits on the performance capabilities of all vehicles, and IA strategy is developed using time-to-collide, braking ratios, driver aggressiveness, and desired security factors. Plausible guidance commands are formulated by minimizing a cost function based on these factors [70]. Both normal and emergency expert systems govern the process, supported by procedural calculations for situation assessment, traffic prediction, estimation, and control (Fig. 6). Guidance commands are formulated by minimizing a cost function based on these factors [70].

Each of the expert systems discussed in this section performs deduction in a cyclical fashion based on prior logical structures, prior knowledge of parameters, and real-time measurements. Intelligent flight control systems must deal with unanticipated events, but it is difficult to
identify aeronautical applications where on-line declarative learning is desirable. Nevertheless, off-line induction is needed to formulate the initial declarative system and perhaps (in a manner reminiscent of REM Sleep) to upgrade declarative logic between missions.

Two approaches are considered in greater detail. The first is called rule recruitment [75], and it involves the manipulation of "dormant rules" (or rule templates). This method was applied in the development of an intelligent failure-tolerant control system for a helicopter. Each template possesses a fixed premise-action structure and refers to parameters through pointers. Rules are constructed and incorporated in the rule base by defining links and modifying parameter-rule-association lists. Learning is based on Monte Carlo simulations of the controlled system with alternate failure scenarios. Learned parameter values then can be defined as "fuzzy functions" [76] contained in rule premises.

The second approach is to construct classification or decision trees that relate attributes in the data to decision classes [52]. The problem is to develop an Expert Navigation-Sensor Management System (NSM) that selects the best navigation aids from available measurements. Several aircraft paths were simulated, and the corresponding measurements that would have been made by GPS, Loran, Tacan, VOR, DME, Doppler radar, air data, and inertial sensors were calculated, with representative noise added. The simulated measurements were processed by extended Kalman filters to obtain optimal state estimates in 200 simulations. Using the root-sum-square error as a decision metric, Analysis of Variance (ANOVA) identifies the factors that make statistically significant contributions to the decision metric, and the Iterative Dichotomizer #3 (ID3) algorithm [77-79] extracts rules from the training set by inductive inference. The ID3 algorithm quantifies the entropy content of each attribute, that is, the information gained by testing the attribute at a given decision node. It uses an information-theoretic measure to find a splitting strategy that minimizes the number of nodes required to characterize the tree. Over 900 examples were used to develop the NSM decision tree.

**PROCEDURAL SYSTEMS**

Most guidance, navigation, and control systems fielded to date are procedural systems using sequential algorithms and processors. Although optimality of a cost function is not always a necessary or even sufficient condition for a "good" system, linear-optimal stochastic controllers provide a good generic structure for discussion.

**Control and Estimation**

We assume that a nominal (desired) flight path is generated by higher-level intelligence, such as the human pilot or declarative machine logic. The procedural system must follow the path, \( \mathbf{x}^*(t) \) in \( t_0 < t < t_f \). Control is exercised by a digital computer at time intervals of \( \Delta t \). The \( n \)-dimensional state vector perturbation at time \( t_k \) is \( \mathbf{x}_k \), and the \( m \)-dimensional control vector perturbation is \( \mathbf{u}_k \). The discrete-time linear-quadratic-Gaussian (LQG) control law is formed as [63],

\[
\mathbf{u}_k = \mathbf{u}_k^* - C_B \hat{\mathbf{x}}_k - C_F \mathbf{y}_k = C_B \hat{\mathbf{x}}_k - C_F \mathbf{y}_k
\]  

(1)

\( \mathbf{y}_k^* \) is the desired value of an output vector (defined as \( \mathbf{H}_x \mathbf{x}_k + \mathbf{H}_u \mathbf{u}_k \)), and \( \hat{\mathbf{x}}_k \) is the Kalman filter estimate, expressed in two steps:

\[
\hat{\mathbf{x}}_k = \mathbf{H}_o \mathbf{B}_o \hat{\mathbf{x}}_k + \mathbf{K} \{ \mathbf{z}_k - \mathbf{H}_o \mathbf{B}_o \hat{\mathbf{x}}_k \}
\]  

(2)
The forward and feedback control gain matrices are \( C_F \) and \( C_B \), \( \Phi \) and \( \Gamma \) are state-transition and control-effect matrices that describe the aircraft's assumed dynamics, the estimator gain matrix is \( K \), and the measurement vector, \( z_k \), is a transformation of the state through \( H_{obs} \). The gains \( C_B \) and \( K \) result from solving two Riccati equations that introduce tradeoffs between control use and state perturbation and between the strengths of random disturbances and measurement error. \( C_F \), which provides proper steady-state command response, is an algebraic function of \( C_B \), \( \Phi \), \( \Gamma \), and \( H_{obs} \). All of the matrices may vary in time, and it may be necessary to compute \( K \) online. In the remainder, it is not essential that \( C_B \) and \( K \) be optimal (i.e., they may have been derived from eigenstructure assignment, loop shaping, etc.), although the LQR gains guarantee useful properties of the nominal closed-loop system [63].

The control structure provided by eq. 1 and 2 is quite flexible. It can represent a scalar feedback loop if \( z \) contains one measurement and \( u \) one control, or it can address measurement and control redundancy with \( z \) and \( u \) dimensions that exceed the dimension of the state \( x \). It also is possible to incorporate reduced-order modeling in the estimator. Assuming that \( \Phi \) and \( \Gamma \) have the same dimensions as the aircraft's dynamic model \((n \times n \text{ and } n \times m)\), the baseline estimator introduces nth-order compensation in the feedback control loop. The weights of the quadratic control cost function can be chosen not only to penalize state and control perturbations but to produce output weighting, state-rate weighting, and implicit model following, all without modifying the dynamic model [63]. Integral compensation, low-pass filter compensation, and explicit model following can be obtained by augmenting the system model during the design process, increasing the compensation order and producing the control structures shown in Fig. 7.

These cost weighting and compensation features can be used together, as in the proportional-integral/implicit-model-following controllers developed in [80]. Implicit model following is especially valuable when an ideal model can be specified (as identified in flying qualities specifications and standards [5, 6]), and integral compensation provides automatic "trimming" (control that synthesizes \( u^*k \) corresponding to \( x^*_k \) to achieve zero steady-state command error) and low-frequency robustness. Combining integral and filter compensation produces controllers with good command tracking performance and smooth control actions, as demonstrated in flight tests [81-83]. The LQG regulator naturally introduces an internal model of the controlled plant, a feature that facilitates control design [84]. It produces a stable approximation to the system inverse, which is at the heart of achieving desired command tracking.

The estimator in the feedback loop presents an efficient means of dealing with uncertainty in the measurements, in the disturbance inputs, and (to a degree) in the aircraft's dynamic model. If measurements are very noisy, the estimator gain matrix \( K \) is "small," so that the filter relies on extrapolation of the system model to estimate the state. If disturbances are large, the state itself is more uncertain, and \( K \) is "large," putting more emphasis on the measurements. Effects of uncertain parameters can be approximated as "process noise" that increases the importance of measurements, leading to a higher \( K \). If the system uncertainties are constant but unknown biases or scale factors, a better approach is to augment the filter state to estimate these terms directly. Parametric uncertainty introduces nonlinearity; hence, an extended Kalman filter must be used [63].

| a) Linear-Quadratic-Gaussian (LQG) Regulator. |
| b) Proportional-Filter LQG Regulator. |
| c) Proportional-Integral LQG Regulator. |
| d) Explicit-Model-Following LQG Regulator. |

**Figure 7. Structured Linear-Quadratic-Gaussian Regulators.**

**Stability and Performance Robustness**

Controlled system robustness is the ability to maintain satisfactory stability and performance in the presence of parametric or structural uncertainties in either the aircraft or its control system. All controlled systems must
possess some degree of robustness against operational parameter variations. The inherent stability margins of certain algebraic control laws (e.g., the linear-quadratic (LQ) regulator [63, 85-87]) may become vanishingly small when dynamic compensation (e.g., the estimator in a linear-quadratic-Gaussian (LQG) regulator) is added [88]. Restoring the robustness to that of the LQ regulator typically requires increasing estimator gains (within practical limits) using the loop-transfer-recovery method [89] or the stochastic robustness approach described below.

Subjective judgments must be made in assessing the need for robustness and in establishing corresponding control system design criteria, as there is an inevitable tradeoff between robustness and nominal system performance [90]. The designer must know the normal operating ranges and distributions of parameter variations, as well as the specifications for system operability with failed components, else the final design may afford too little robustness for possible parameter variations or too much robustness for satisfactory nominal performance. Robustness traditionally has been assessed deterministically [91, 92]; gain and phase margins are an inherent part of the classical design of single-input/single-output systems, and there are multi-input/multi-output equivalents based on singular-value analysis (e.g., [93]). A critical difficulty in applying these techniques is relating singular-value bounds on return-difference and inverse-return-difference matrices to real parameter variations in the system.

Statistical measures of robustness can use knowledge of potential variations in real parameters. The probability of instability was introduced in [94] and is further described in [95, 96]. The stochastic robustness of a linear, time-invariant system, is judged using Monte Carlo simulation to estimate the probability distributions of closed-loop eigenvalues, given the statistics of the variable parameters in the system's dynamic model. Because the system is either stable or not, the probability of instability has a binomial distribution; hence, the confidence intervals associated with estimating the metric from simulation are independent of the number or nature of the uncertain parameters [95].

Considerations of performance robustness are easily taken into account in Stochastic Robustness Analysis (SRA). Systems designed using a variety of robust control methods (loop transfer recovery, H∞ optimization, structured covariance, and game theory) are analyzed in [97], with attention directed to the probability of instability, probability of settling-time exceedence, probability of excess control usage, and tradeoffs between them. The analysis uncovers a wide range of system responses and graphically illustrates that gain and phase margins are not good indicators of the probability of instability. This also raises doubts about the utility of singular values, as they are multivariable equivalents of the classical robustness metrics. Incorporating SRA into the design of an LQG regulator with implicit model following and filter compensation leads to designs that have high levels of stability and performance robustness [98]. The reason for improvement is that SRA measures the actual effects of parameter variations on stability and performance rather than incremental changes in the nominal margins.

**Adaptation and Tolerance to Failures**

Adaptation always has been a critical element of stability augmentation. Most aircraft requiring improved stability undergo large variations in dynamic characteristics on a typical mission. Gain scheduling and control interconnects initially were implemented mechanically, pneumatically, and hydraulically; now the intelligent part is done within a computer, and there is increased freedom to use sophisticated scheduling techniques that approach full nonlinear control [81, 99].

One approach to improving failure tolerance is parallel redundancy: two or more control strings, each separately capable of satisfactory control, are implemented in parallel. A voting scheme is used for redundancy management. With two identical channels, a comparator can determine whether or not control signals are identical; hence, it can detect a failure but cannot identify which string has failed. Using three identical channels, the control signal with the middle value can be selected (or voted), assuring that a single failed channel never controls the plant. Parallel redundancy can protect against control-system component failures, but it does not address failures of plant components. Analytical redundancy provides a capability to improve tolerance to failures of both types. The principal functions of analytical redundancy are failure detection, failure identification, and control-system reconfiguration [47].

Procedural adaptation and failure-tolerance features will evolve outward, to become more declarative in their supervision and more reflexive in their implementation. Declarative functions are especially important for differentiating between normal and emergency control functions and sensitivities. They can work to reduce trim drag, to increase fatigue life, and to improve handling and ride qualities as functions of turbulence level, passenger loading, and so on. Gain-scheduling control can be viewed as fuzzy control, suggesting that the latter has a role to play in aircraft control systems [100-102]. Reflexive functions can be added by computational neural networks that approximate nonlinear multivariate functions or classify failures.

---

8 Real parameter variations affect not only the magnitude and relative phase angle of the system's Nyquist contour but its shape as well [63]. Therefore, the points along the contour that establish gain and phase margin (i.e., the corresponding Bode-plot frequencies) are subject to change.
Nonlinear Control

Aircraft dynamics are inherently nonlinear, but aerodynamic nonlinearities and inertial coupling effects are generally smooth enough in the principal operating regions to allow linear control design techniques to be used. Control actuators impose hard constraints on operation because their displacements and rates are strictly limited. Nonlinear control laws can improve control precision and widen stability boundaries when flight must be conducted at high angles or high angular rates and when the control-actuator limits must be challenged.

The general principles of nonlinear inverse control are straightforward [103]. Given a nonlinear system of the form,

$$\dot{x} = f(x) + G(x)u$$

where $G(x)$ is square ($m = n$) and non-singular, the control law

$$u = -G^{-1}f(x) + G^{-1} v$$

inverts the system, since

$$\dot{x} = f(x) + G(x)[-G^{-1}f(x) + G^{-1} v] = v$$

where $v$ is the command input to the system.

In general, $G(x)$ is not square ($m \neq n$); however, given an $m$-dimensional output vector,

$$\Delta y = Hx$$

it is possible to define a nonlinear feedback control law that produces output decoupling of the elements of $y$ or their derivatives such that $y(d) = v$. The vector $y(d)$ contains Lie derivatives of $y$,

$$y(d) = f^*(x) + G^*(x)u$$

where $d$ is the relative degree vector of differentiation required to identify a direct control effect on each element of $y$. $G^*(x)$ and $f^*(x)$ are components of the Lie derivatives, and $G^*(x)$ is guaranteed to be structurally invertible by the condition that defines relative degree [104]. The decoupling control law then takes the form

$$u = -(G^*(x))^{-1}f^*(x) + [G^*(x)]^{-1} v$$

The control law is completed by feeding $y$ back as appropriate to achieve desired transient response and by pre-filtering $v$ to produce the desired command response [105]. Because the full state is rarely measured and measurements can contain errors, it may be necessary to estimate $x$ with an extended Kalman filter, substituting $\hat{x}$ for $x$ in control computations.

Evaluating $G^*(x)$ and $f^*(x)$ requires that a full, $d$-differentiable model of aircraft dynamics be included in the control system; hence the statement of the control law is simple, but its implementation is complex (Fig. 8). Smooth interpolators of the aircraft model (e.g., cubic splines) are needed. Feedforward neural networks with sigmoidal activation functions are infinitely differentiable, providing a good means of representing this model on-line and allowing adaptation [106, 107]. Limitations to the inverse control approach are discussed in [108].

![Figure 8. Decoupling Nonlinear-Inverse Control Law.](image)

REFLEXIVE SYSTEMS

Inner-loop control is a reflexive (though not necessarily linear) function. To date, most inner loops have been designed as procedural control structures; computational neural networks may extend prior results to facilitate nonlinear control and adaptation. Neural networks can be viewed as nonlinear generalizations of sensitivity, transformation, and gain matrices. Consequently, compensation dynamics can be incorporated by following earlier models and control structures. Nonlinear proportional-integral and model following controllers, as well as nonlinear estimators, can be built using computational neural networks.

Computational Neural Networks

Computational neural networks are motivated by input-output and learning properties of biological neural systems, though in mathematical application the network becomes an abstraction that may bear little resemblance to its biological antecedent. Computational neural networks consist of nodes that simulate the neurons and weighting factors that simulate the synapses of a living nervous system. The nodes are nonlinear basis functions, and the weights contain knowledge of the system. Neural networks are good candidates for performing a variety of reflexive functions in intelligent control systems because they are potentially very fast (in parallel hardware implementation), they are intrinsically nonlinear, they can address problems of high dimension, and they can learn from experience. From the biological analogy, the neurons are modeled as switching functions that take just two discrete values; however, "switching" may be softened to "saturation," not only to facilitate learning of the synaptic...
weights but to admit the modeling of continuous, differentiable functions.

The neural networks receiving most current attention are static expressions that perform one of two functions. The first is to approximate multivariate functions of the form

\[ y = h(x) \]  

(9)

where \( x \) and \( y \) are input and output vectors and \( h(\cdot) \) is the (possibly unknown) relationship between them. Neural networks can be considered to be generalized spline functions that identify efficient input-output mappings from observations [109, 110]. The second application is to classify attributes, much like the decision trees mentioned earlier. (In fact, decision trees can be mapped to neural networks [111].) The following discussion emphasizes the first of these two applications.

An N-layer feedforward neural network (FNN) represents the function by a sequence of operations,

\[ r(k) = s(k)[w(k-1)r(k-1)] \rightarrow s(k)[r(k)] \]  

(10)

where \( y = r(N) \) and \( x = r(0) \), \( W(k-1) \) is a matrix of weighting factors determined by the learning process, and \( s(k)[\cdot] \) is an activation-function vector whose elements normally are identical, scalar, nonlinear functions \( \sigma_1(\eta_1) \) appearing at each node:

\[ s(k)[\eta(k)] = [\sigma_1(\eta_1(k)) ... \sigma_n(\eta_n(k))]^T \]  

(11)

One of the inputs to each layer may be a unity threshold element that adjusts the bias of the layer's output. Networks consisting solely of linear activation functions are of little interest, as they merely perform a linear transformation \( H \), thus limiting eq. 9 to the form, \( y = Hx \).

Figure 9 represents two simple feedforward neural networks. Each circle represents an arbitrary, scalar, nonlinear function \( \sigma_i(\cdot) \) operating on the sum of its inputs, and each arrow transmits a signal from the previous node, multiplied by a weighting factor. A scalar network with a single hidden layer of four nodes and a unit threshold element (Fig. 9a) is clearly parallel, yet its output can be written as the series

\[ y = a_0(\sigma_0(b_0x + c_0)) + a_1(\sigma_1(b_1x + c_1)) + a_2(\sigma_2(b_2x + c_2)) + a_3(\sigma_3(b_3x + c_3)) \]  

(12)

illustrating that parallel and serial processing may be equivalent.

Consider a simple example. Various nodal activation functions, \( \sigma_i \), have been used, and there is no need for each node to be identical. Choosing \( \sigma_0(\cdot) = (\cdot), \sigma_1 = (\cdot)^2, \sigma_2 = (\cdot)^3, \sigma_3 = (\cdot)^4 \), eq. 9 is represented by the truncated power series,

\[ y = a_0(b_0x + c_0) + a_1(b_1x + c_1)^2 + a_2(b_2x + c_2)^3 + a_3(b_3x + c_3)^4 \]  

(13)

It is clear that network weights are redundant (i.e., that the \( (a, b, c) \) weighting factors are not independent). Consequently, more than one set of weights could produce the same functional relationship between \( x \) and \( y \). Training sessions starting at different points could produce different sets of weights that yield identical outputs. This simple example also indicates that the unstructured feedforward network may not have compact support (i.e., its weights may have global effects) if its basis functions do not vanish for large magnitudes of their arguments.

\[ \sigma = (\cdot) \]

\[ \text{a) Single-Input/Single-Output Network.} \]

\[ \text{b) Double-Input/Single-Output Network.} \]

Figure 9. Two Feedforward Neural Networks.

The sigmoid is commonly used as the artificial neuron. It is a saturating function: \( \sigma(\eta) = 1/(1 + e^{-\eta}) \) for output in \((0,1)\) or \( \sigma(\eta) = (1 - e^{-2\eta})(1 + e^{-2\eta}) = \tanh \eta \) for output in \((-1,1)\). Recent results indicate that any continuous mapping can be approximated arbitrarily closely with sigmoidal networks containing a single hidden layer (\( N = 2 \)) [112, 113]. Symmetric functions like the Gaussian radial basis function \( (\sigma(\eta) = e^{-\eta^2}) \) have better convergence properties for many functions and have more compact support as a consequence of near-orthogonality [109, 114]. Classical B-splines [115] could be expressed in par-
allel form, and it has been suggested that they be used in multi-layered networks [116]. Hidden layers strengthen the analogy to biological models, though they are not necessary for approximating continuous functions, and they complicate the training process.

In control application, neural networks perform functions analogous to gain scheduling or nonlinear control. Consider the simple two-input network of Fig. 9b. The scalar output and derivative of a single sigmoid with unit weights are shown in Fig. 10. If u is a fast variable and v is a slow variable, choosing the proper weights on the inputs and threshold can produce a gain schedule that is approximately linear in one region and nonlinear (with an inflection point) in another. More complex surfaces can be generated by increasing the number of sigmoids. If u and v are both fast variables, then the sigmoid can represent a generalization of their nonlinear interaction.

For comparison, a typical radial basis function produces the output shown in Fig. 11. Whereas the sigmoid has a preferred input axis and simple curvature, the RBF admits more complex curvature of the output surface, and its effect is more localized. The most efficient nodal activation function depends on the general shape of the surface to be approximated. There may be cases best handled by a mix of sigmoids and RBF in the same network.

The cerebellar model articulation controller (CMAC) is an alternate network formulation with somewhat different properties but similar promise for application in control systems [117, 118]. The CMAC performs table look-up of a nonlinear function over a particular region of function space. CMAC operation can be split into two mappings. The first maps each input into an association space \( A \). The mapping generates a selector vector \( a \) of dimension \( n_A \), with \( c \) non-zero elements (usually ones) from overlapping receptive regions for the input. The second mapping, \( R \), goes from the selector vector \( a \) to the scalar output \( y \) through the weight vector \( w \), which is derived from training:

\[
y = w^T a
\]

Training is inherently local, as the extent of the receptive regions is fixed. The CMAC has quantized output, producing "staircased" rather than continuous output. A recent paper proposes to smooth the output using B-spline receptive regions [119].

The FNN and CMAC are both examples of static networks, that is, their outputs are essentially instantaneous: given an input, the speed of output depends only on the speed of the computer. Dynamic networks rely on stable resonance of the network about an equilibrium condition to relate a fixed set of initial conditions to a steady-state output. Bidirectional Associative Memory (BAM) networks [120] are nonlinear dynamical systems that subsume Hopfield networks [121], Adaptive-Resonance-Theory (ART) networks [122], and Kohonen networks [123]. Like FNN, they use binary or sigmoidal neurons and store knowledge in the weights that connect them; however, the "neural circuits" take time to stabilize on an output. While dynamic networks may operate more like biological neurons, which have a refractory period between differing outputs, computational delay degrades aircraft control functions.

Although neural networks performing function approximation may gain little from multiple hidden layers, networks used for classification typically require multiple layers, as follows from the ability to map decision trees to neural networks [111]. The principal values of performing such a mapping are that it identifies an efficient structure for parallel computation, and it may facilitate incremental learning and generalization.

![Figure 10. Example of Sigmoid Output with Two Inputs.](image-url)
The output is a scalar, and the network is trained (for example) with "1" corresponding to failure and "0" corresponding to no failure. The data patterns associated with each failure may require feature extraction, pre-processing that transforms the input time series into a feature vector [124]. As an alternative, the feature vector could be specified as a parity vector [127], and the neural network could be used for the decision-making logic in FDI.

<table>
<thead>
<tr>
<th>a) Radial Basis Function (RBF).</th>
</tr>
</thead>
<tbody>
<tr>
<td>b) x-Derivative of RBF.</td>
</tr>
</tbody>
</table>

Figure 11. Example of Radial Basis Function Output with Two Inputs.

**Reflexive Learning and Adaptation**

Training neural networks involves either supervised or unsupervised learning. In supervised learning, the network is furnished typical histories of inputs and outputs, and the training algorithm computes the weights that minimize fit error. FNN and CMAC require this type of training, as discussed below. In unsupervised learning, the internal dynamics are self-organizing, tuning the network to home on different cells of the output semantic map in response to differing input patterns [128].

**Backpropagation** learning algorithms for the elements of $W^{(k)}$ typically involve a gradient search (e.g., [129, 130]) that minimizes the mean-square output error

$$E = \left[ r_d - r(N) \right]^T \left[ r_d - r(N) \right] \quad (15)$$

where $r_d$ is the desired output. For each input-output example presented to the network, the gradient of the error with respect to the weight matrix is calculated, and the weights are updated by

$$W^{(k)}_{\text{new}} = W^{(k)}_{\text{old}} + \beta r(k-1) \left[ d(k) \right]^T \quad (16)$$

$\beta$ is the learning rate, and $d$ is a function of the error between desired and actual outputs. For the output layer, the error term is

$$d(N) = S\left[ W(N-1) r(N-1) \right] (r_d - r(N)) \quad (17)$$

where the prime indicates differentiation with respect to $r$. For interior layers, the error from the output layer is propagated from the output error using

$$d^{(k)} = S\left[ W(k-1) r(k-1) \right] \left[ W^{(k-1)} \right]^T d(k-1) \quad (18)$$

Search rate can be modified by adding momentum or conjugate-gradient terms to eq. 16.

The CMAC network learning algorithm is similar to backpropagation. The weights and output are connected by a simple linear operation, so a learning algorithm is easy to prescribe. Each weight contributing to a particular output value is adjusted by a fraction of the difference between the network output and the desired output. The fraction is determined by the desired learning speed and the number of receptive regions contributing to the output.

Learning speed and accuracy for FNN can be further improved using an extended Kalman filter [106, 107, 131]. The dynamic and observation models for the filter are

$$w_k = w_{k-1} + q_{k} \quad (19)$$

$$z_k = h(w_k, r_k) + n_k \quad (20)$$

where $w_k$ is a vector of the matrix $W_k$'s elements, $h(\cdot)$ is an observation function, and $q_k$ and $n_k$ are noise processes. If the network has a scalar output, then $z_k$ is scalar, and the extended Kalman filter minimizes the fit error between the training hypersurface and that produced by the network (eq. 15). It has been found that the fit error can be dramatically reduced by considering the gradients of the surfaces as well [106, 107]. The observation vector becomes
\[ z_k = \begin{bmatrix} h(w_k, r_k) \\ \frac{\partial h}{\partial r}(w_k, r_k) \end{bmatrix} + n_k \]  

with concomitant increase in the complexity of the filter. The relative significance given to function and derivative error during training can be adjusted through the measurement-error covariance matrix used in filter design.

Recursive estimation of the weights is useful when smooth relationships between fit errors and the weights are expected, when the weight-vector dimension is not high, and when local minima are global. When one of these is not true, it may speed the computation of weights to use a randomized search, at least until convergent regions are identified. Such methods as simulated annealing or genetic algorithms can be considered (and the latter has philosophic appeal for intelligent systems) [132-134].

The first of these is motivated by statistical mechanics and the effects that controlled cooling has on the ground states of atoms (which are analogous to the network weights). The second models the reproduction, crossover, and mutation of biological strings (e.g., chromosomes, again analogous to the weights), in which only the fittest combinations survive.

Statistical search methods can go hand-in-hand with SRA to train robust neural networks. Following [98], the randomized search could be combined with Monte Carlo variation of system parameters during training, numerically minimizing the expected value of fit error rather than a deterministic fit error.

We envision an aerodynamic model that spans the entire flight envelope of an aircraft, including post-stall and spinning regions. The model contains six neural networks with multiple inputs and scalar outputs, three for force coefficients and three for moment coefficients (for example, the pitch moment network takes the form \( C_m = g(x, u) \), where \( x \) represents the state and \( u \) the control). Input variables are not restricted to those having plausible aerodynamic effect, incorrect correlations may be created in the network; hence, attitude Euler angles and horizontal position should be neglected, while physically meaningful terms like elevator deflection, angle of attack, pitch rate, Mach number, and dynamic pressure should be included.

The aircraft spends most of its flying time within normal mission envelopes. Unless it is a trainer, the aircraft does not enter post-stall and spinning regions; consequently, on-line network training focuses on normal flight and neglects extreme conditions. This implies not only that networks must be pre-trained in the latter regions but that normal training must not destroy knowledge in extreme regions while improving knowledge in normal regions. Therefore, radial basis functions appear to be a better choice than sigmoid activation functions for adaptive networks.

Elements of the input vector may be strongly correlated with each other through the aircraft's equations of motion; hence, networks may not be able to distinguish between highly correlated variables (e.g., pitch rate and normal acceleration). This is problematical only when the aircraft is outside its normal flight envelope. Pre-training should provide inputs that are rich in frequency content, that span the state and control spaces, and that are as uncorrelated as possible. Generalization between training points may provide smoothness, but it does not guarantee accuracy.

Control Systems Based on Neural Networks

Neural networks can find application in logic for control, estimation, system identification, and physical modeling. In addition to work already referenced, additional examples can be found in [135-140].

Figure 12a illustrates an application in which the neural network forms the aircraft model for a nonlinear-in-inverse control law. The aircraft model of Fig. 9 is implemented with a neural network that is trained by a dedicated (weight) extended Kalman filter (the thick gray arrow indicating training). The extended Kalman filter for state estimation is expanded to estimate histories of forces and moments as well as the usual motion variables.

It is possible to conduct supervised learning on-line while not interfering with normal operation because the state Kalman filter produces both the necessary inputs and the desired outputs for the network training algorithm. There is no need to provide an ideal control response for training, as the form of the control law is fixed. Procedural and reflexive functions are combined in this control implementation, under the assumption that the direct expression of inversion is the most efficient approach.

Figure 12b shows a logical extension in which the inverse control law is implemented by neural networks. Inversion is an implicit goal of neural-network controllers [135, 136], and the formal existence of inversion networks has been explored [141]. Although Fig. 12b implies that the inversion networks are pre-trained and fixed, they, too, can be trained with the explicit help of the network that models the system [136].

If a desired control output is specified (Fig. 12c), then the formal model of the aircraft is no longer needed. The control networks learn implicit knowledge of the aircraft model. Referring to Fig. 10 and Eq. 1 and 2, control and estimation gains, state-transition and control-effect matrices, and measurement transformations can be implemented as static neural networks with either off-line or on-line learning.

Dividing control networks into separate feedback and forward parts may facilitate training to achieve design goals. A feedback neural network has strongest effect on homogeneous modes of motion, while a forward neural network is most effective for shaping command (forced) response. This structure is adopted in [139], where the
forward and feedback networks are identified as *reason* and *instinct networks*. In pre-training, it is plausible that the feedback network would be trained with initial condition responses first, to obtain satisfactory transient response. The forward network would be trained next to achieve desired steady states and input decoupling. A third training step could be the addition of command-error integrals while focusing on disturbance inputs and parameter variations in training sets.

Once baselines have been achieved, it could prove useful to admit limited coupling between forward and feedback networks to enable additional nonlinear compensation. In adaptive applications, networks would be pre-trained with the best available models and scenarios to establish satisfactory baselines; on-line training would slowly adjust individual systems to vehicle and mission characteristics.

Including the integral of command-vector error as a neural network input produces a *proportional-integral* structure (140), while placing the integrator beyond the network gives a *proportional-filter* structure (Fig. 10). The principal purpose of these structures is, as before, to assure good low- and high-frequency performance in a classical sense. Extension of neural networks to state and weight filters is a logical next step that is interesting in its own right as a means of more nearly optimal nonlinear estimation.

**CREW-STRUCTURED INTELLIGENT AIRCRAFT CONTROL**

The declarative AUTOCREW paradigm presented earlier can be expanded to include procedural and reflexive functions, recognizing that control of flight is just one of several control functions in the aircraft. An intelligent control system for a civil aircraft might take the form of Fig. 13; functions represented by crew-member equivalents are linked to each other by a communications network and to aircraft systems via a separate network. This concept remains to be explored.

**CONCLUSION**

Intelligent flight control systems can do much to improve the operating characteristics of aircraft. An examination of cognitive and biological models for human control of systems suggest that there is a declarative, procedural, and reflexive hierarchy of functions. Top-level aircraft control functions are analogous to conscious and preconscious thoughts that are transmitted to lower-level subsystems through subconscious, neural, and reflex-like activities. Human cognition and biology also suggest models for learning and adaptation, not only during operation but between periods of activity.

The computational analogs of the three cognitive/biological paradigms are expert systems, stochastic controllers, and neural networks. Expert systems organize decision-making efficiently, stochastic controllers optimize estimation and control, and neural networks provide rapid, nonlinear, input-output functions. It appears that many functions at all levels could be implemented as neural networks. While this may not always be necessary or even desirable using sequential processors, mapping declarative and procedural functions as neural networks may prove most useful as a route to new algorithms for the massively parallel processors of the future.
ACKNOWLEDGMENTS

This work has been supported by the Federal Aviation Administration and the National Aeronautics and Space Administration under Grant No. NGL 31-001-252 and by the Army Research Office under Contract No. DAAL03-89-K-0092.

REFERENCES


This report summarizes the research conducted during the academic year 1991–1992 under the FAA/NASA sponsored Joint University Program for Air Transportation Research. The year end review was held at Ohio University, Athens, Ohio, June 18–19, 1992. The Joint University Program is a coordinated set of three grants sponsored by the Federal Aviation Administration and NASA Langley Research Center, one each with the Massachusetts Institute of Technology (NGL-22-009-640), Ohio University (NGR-36-009-017), and Princeton University (NGL-31-001-252). Completed works, status reports, and annotated bibliographies are presented for research topics, which include navigation, guidance and control theory and practice, intelligent flight control, flight dynamics, human factors, and air traffic control processes. An overview of the year’s activities for each university is also presented.