ABSTRACT. Successful decisions largely depend on correct interpretation of data. Today, our ability to collect data outstrips our ability to interpret it, a situation called "information input overload." Information input overload is known to have a deleterious effect on decision makers. Full use of data, knowledge, and other information requires a system that can extract the critical decision factors and follow a decision tree to find related pieces of information. A knowledge-based system was built to aid the project management team responsible for identifying cotton fields at risk to pink bollworm and releasing sterile pink bollworm to help control the native pink bollworm population. The system uses object-oriented design, expert system techniques, a link to simulation models, and database management in an integrated system to optimize, improve, and ease the decision-making process. The system made significantly fewer mistakes than did human decision makers, while assigning treatments to high and low risk areas. In addition, the system thoroughly documents the decision-making process and the resulting recommendations, thus allowing use of adjuncts such as a GIS and simulation models of pest and crop populations.

Cotton Pest Management: A Knowledge-Based System to Handle Information Input Overload

Michelle Walters Geyer,[1] Ronaldo A. Sequeira,[2] David B. Paradice,[3] George C. Fowler,[3] Charlotte Miller,[4] and Robert T. Staten[1]

Most organizations use resources to collect facts and details (i.e., data) about their business's progress, performance, and efficiency. Decisions and subsequent actions are the products of this effort. To ignore or otherwise misuse data is to misuse potential resources (Courtney and Paradice 1988). In an ideal world, each datum would be correctly interpreted and evaluated in relationship to other related pieces of data. Thus, decisions would be based on complete and accurate interpretation of the situation. Decision makers may not take complete advantage of data due to several constraints. First, the volume of data may be too large to be assimilated. Second, the nature of the problem may be complex. Third, difficulties may be magnified and multiplied by the need to decide quickly. Finally, there are risks associated with missing a "trigger" point. Therefore, decision makers may use excessive amounts of

[1] USDA, APHIS, PPQ
Phoenix Methods Development Center
4125 E. Broadway Road
Phoenix, Arizona 85040

[2] USDA, ARS, CSR, Crop Simulation Laboratory
P.O. Box 5367
Mississippi State, Mississippi 39762

[3] Business Analysis Department
Texas A&M University
College Station, Texas 77843-4217

[4] USDA, APHIS, ISCD
Information Resource Management and Budget
6505 Belcrest Road, Room 717
Hyattsville, Maryland 20782

resources to compensate for incomplete understanding and to ensure success of the project. Investigators have studied how people use information (i.e., data in context or summarized to make it more usable) to make decisions, and have developed theories and applications to model and enhance the decision-making process (Gorry and Morton 1971, Keen 1980, Mason and Mitroff 1973, Sprague 1980). The components involved in a decision include the decision maker, the task involved, the reason for making a decision (or stimulus), the decision environment, and the decision strategy. Information input overload is a characteristic of the task. The decision process has been described by several models. The steps include intelligence, design, and choice (Simon 1960), and are closely tied to the decision components listed above. Gorry and Morton (1971) state:

> Most managers do not have great information needs. Rather, they have need of new methods to understand and process the information already available to them. (p. 65).

The development of database theory and practice was an early step toward managing data. Database management has since become an important area of object-oriented applications as well (Lindsjorn and Sjoberg 1988, Lindskov-Knudsen and Lehrmann-Madsen 1988, Tello 1989, Tsichritzis and Nierstasz 1988). Object-oriented programming refers to both a programming and design style that has been widely embraced by the scientific community as the best current approach to data management, software design, and software maintenance (Booch 1991, Meyer 1988). An advantage of object-oriented databases is the ability to deal with abstract data types. For example, you could select all the geographical data types in a geographical application if you employed abstract data typing. A goal of modern databases is to relieve users of the need to have a priori knowledge of data fields and files to issue retrieval commands.

DATA ANALYSIS AND INTERPRETATION CAN BE IMPEDED BY THE AVAILABILITY OF ABUNDANT DYNAMIC DATA AND THE URGENCY OF MAKING A DECISION.

This goal can be realized using object-oriented databases and abstract data types (Garvey and Johnson 1989). An abstract data type is defined as a set of data and its allowable operations. For example, the abstract data type "string" could be defined as a list of characters that can be appended, separated into individual characters, and counted.

Decision support system and expert system theory and application grew partly as an extension of management information systems (MIS) and partly in response to the continuing need to aid decision makers. Decision support and expert system technologies offer a catalyst for major organizational transformations in that they offer organizations the ability to reshape the nature and execution of fundamental business operations. Research at the intersection of application, design, and technology currently is quite limited. According to Henderson (1987, page 346), "The decision support system field must pursue this research opportunity aggressively." Stabell (1983, p. 233) states that the "diagnosis of decision making and the specification of changes in the decision processes are the activities that provide the key inputs to the design of decision support systems," and that this diagnosis will point out areas of potential improvement.

A knowledge-based system is a computer program that mimics human reasoning processes (heuristics) to solve complex problems (Stone et al. 1986). In this context, knowledge-based systems are prescriptive tools that analyze a situation and present a solution, whereas simulations are predictive (Reddy 1987). These tools are complementary. Beck and Jones (1989) state:

> We are currently in the middle of an evolutionary process in computer science where simulation, expert systems, databases, and natural language ideas are merging. Each of these methodologies has its own strengths and, when integrated into one framework, will offer an exciting, powerful decision support system.

The integration of knowledge-based systems and AI tools with existing databases and other systems is a critical area for future development. This development is projected to occur in the AI arena (Harmon and King 1985) and in organizational decision management (Paradice and Courtney 1989). Feigenbaum et al. (1989) predict the seamless merger of expert system activity and conventional data-processing. In agriculture, as in many other fields, current attention is focused on knowledge-based systems.

This study integrates research findings from MIS, decision theory, computer science, and behavioral decision making. It involves the study of decision support components and their integration, and draws from decision-making research and theory. This study is based on the interaction between information input overload and decision strategy as carried out by a knowledge-based system. Traditional components of decision making (databases, decision trees, post-analysis reports) have been integrated with new tools from AI and simulation modeling. The research results are synthesized and applied to the design of computer-based decision aids. The architecture was evaluated using an application from the domain of agricultural research.

Overcoming Information Input Overload

Data analysis and interpretation, and subsequent theory refinement, can be impeded by abundant, dynamic data and the urgency of making a decision. A person is thus faced with information input overload (Miller 1960). Information load is a decision task variable that can affect the "decision process" (Einhorn 1971). A psychiatrist, J.G. Miller, first reported "information input overload" in 1960. Later, it was described as "information presented at a rate too fast for a person to process" (Sheridan and Ferrell 1974). A person suffering from information input overload may react in one or a combination of ways including: 1) omission: temporary failure to respond to incoming information, 2) error: processing wrong information or incorrect response to correct information, 3) queuing: putting off some information, expecting to take care of it when things inevitably slow down, 4) filtering: processing certain types of information while ignoring other types of information; 5) responding in a general way: ignoring detail, responding with less accuracy than normal, and 6) escape: quitting (Miller 1960). Filtering (4) may take the form of summarizing or recoding the input, or giving up at some point, making a decision, and ignoring further input (Hiltz and Turoff 1985). Decision makers tend to shift from compensatory strategies to noncompensatory strategies as information load increases (Biggs et al. 1985, Cook 1987, Klayman 1985, Payne 1976). Klayman (1985) describes a compensatory strategy as one that uses all available and related information. The benefits of this strategy include the ability to trade the advantages of one dimension against the disadvantages of another dimension. Noncompensatory strategies use only some of the available and related information. A noncompensatory strategy lacks the benefits of the compensatory strategy and is the *best alternative given the effort expended*, as opposed to the optimal choice (Hogarth 1981, Simon 1955, Tversky 1969, 1972).

The noncompensatory strategy is used in response to information overload as a load-reduction strategy rather than as a direct decision-making strategy. Some researchers have suggested that this is an accuracy-versus-effort tradeoff (Cook 1987, Johnson and Payne 1985) since noncompensatory strategies require less effort than do their compensatory counterparts. Payne (1976) reported a sequential shift in decision strategies by students making decisions under increasing information load, and attributed the shift to the increase in cognitive effort required. As information load increased, the subjects shifted from high information processing strategies to low information processing strategies. Both the proportion and the variability of searched information decreased as the information load increased. Naturally, the noncompensatory/low information processing strategy may suffer

OBJECT-ORIENTED PROGRAMMING REFERS TO BOTH A PROGRAMMING AND A DESIGN STYLE THAT HAS BEEN WIDELY EMBRACED BY THE SCIENTIFIC COMMUNITY AS THE BEST CURRENT APPROACH TO DATA MANAGEMENT, SOFTWARE DESIGN, AND SOFTWARE MAINTENANCE.

Information Input Overload in a Research Environment

The United States Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS), Plant Protection and Quarantine (PPQ) Methods Development Center and Pink Bollworm Rearing Facility in Phoenix, Arizona, are working on a joint project with the California Department of Food and Agriculture (CDFA) in Bakersfield, California, to control the pink bollworm moth (PBW) in the San Joaquin Valley of central California. The PBW, *Pectinophora gossypiella* (Saunders), feeds almost exclusively on cotton (*Gossypium* spp.). PBW is often the key pest of cotton in Arizona, southern California, and northwestern Mexico. The larvae (immature stages) of PBW bore into the developing cotton fruit. The larvae feed on the cotton lint and seeds, causing significant damage and dramatically reducing the yield of cotton lint (Lee 1984, Pfadt 1978).

The San Joaquin Valley produces 80 to 90% of the cotton grown in California. This growing area includes as much as 1.2 million acres of cotton, covering 4000 to 5000 square miles and generating some 700 million dollars in revenue annually.

Federal and state agencies began PBW management in the San Joaquin Valley as a joint directive in 1968 with an integrated pest management approach (Metcalf 1980). Insect control efforts include heat, fumigation, and mechanical treatments at gin and oil mills, use of clean seed, quarantine of "raw" cotton products, cultural controls including early plow-down and host-free periods, occasional use of PBW pheromones, use of insecticides in emergencies (only once so far for PBW), monitoring of PBW populations, and the release of sterile PBW moths.

These protocols include timing and quantity of moth delivery to the San Joaquin Valley, location of releases, periodicity of releases, and geographical distribution of field monitoring crews. Key daily decisions involve number, timing, and location of sterile moth releases. Each year the season starts with releases of sterile moths on fields in areas where native moths were found in the previous year. Thus, the percent of area on which there are releases varies from year to year. Moth release typically begins before the first flower buds are present on cotton plants (pin square) and continue with daily shipments of moths throughout the season. As new native PBW are found, the releases are adjusted following a decision protocol and continue throughout the season.

Data Collection and Database Management

Field crews monitor PBW in the cotton fields using 12,000 to 33,000 surveillance stations (traps), following a pre-determined rotation. Field technicians check the traps, count native and sterile moths captured, and clean or replace the traps. They record the trap catches on a "trap slip" and give this information to the central office. Any suspected native PBW caught in a trap is sent to Visalia, where taxonomists verify if the moth is a native. Project management uses the field information to regulate activities such as the further release of sterile PBW moths, use of pheromones, cultural controls, etc.

Initially, project managers used hand-written records of field data as the basis for decision making. As the project progressed and the mass of data increased, management introduced computerized book-keeping procedures for field technician's records, inventory, trap catch entries, daily analysis, and weekly reports.

An MIS to Handle Information Input Overload

The system produced by this research is an alternative to the current system as described above. The knowledge-based system consists of several modules, including: (1) a data interface, (2) data structures encoded as objects, (3) decision-making protocol encoded as rules, (4) diagnostics in report form detailing the amount, frequency, and justification of treatment, and geographic area to be treated, (5) diagnostics in map form showing the treatment areas graphically, and (6) a link to a simulation model of problem development. Figure 2 depicts the current decision-making process and the new system; a description of each of the new system's components follows.

1. Relational database

The system takes the field data (Figures 2 and 3) and gives recommendations on where the next release of sterile moths should be. Field technicians make weekly counts of native and sterile PBW found in Delta traps (baited with PBW pheromone) placed in each cotton field in the San Joaquin Valley. The data are identified to the trap level, with complete information on the field and geographic location (township,

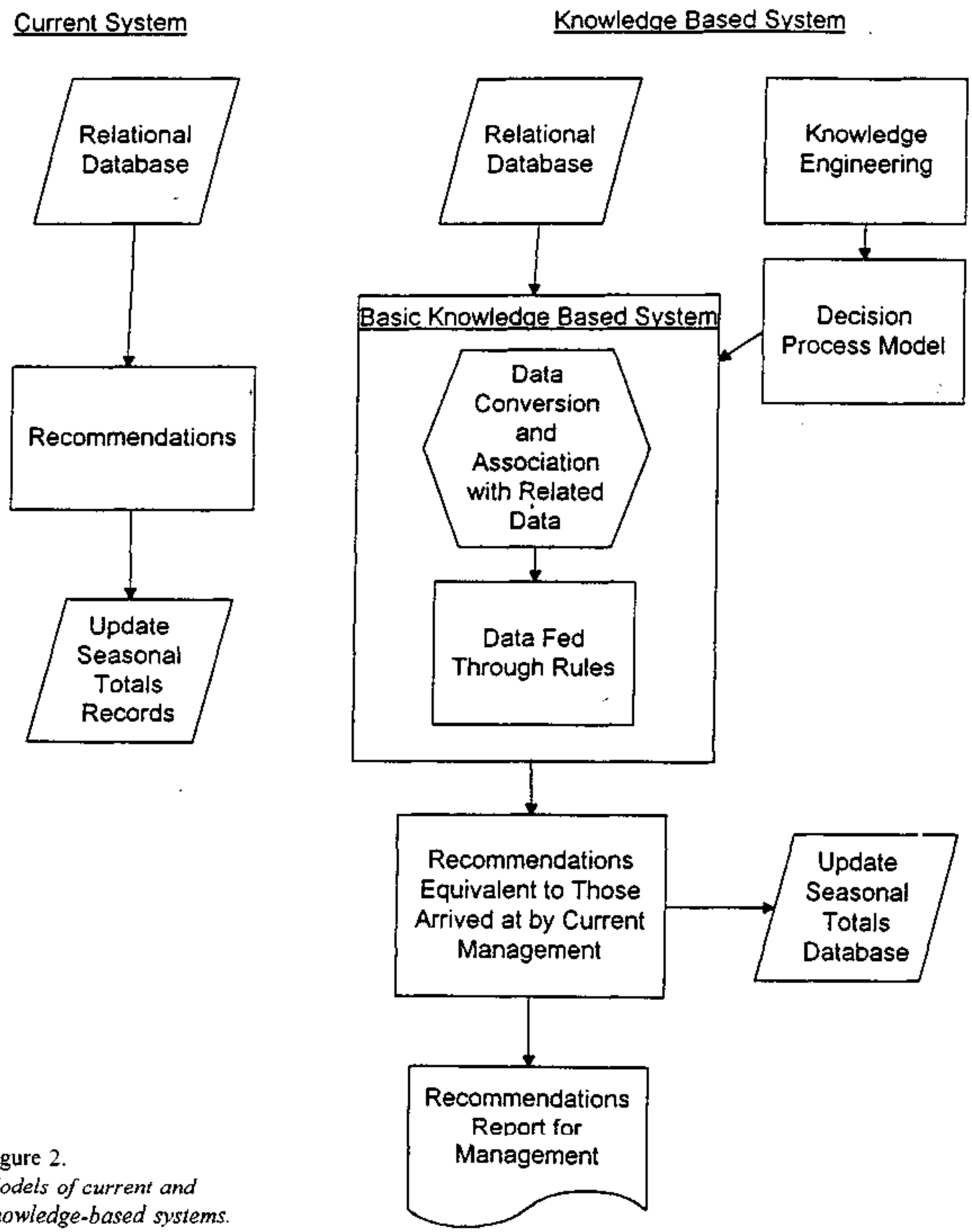

Figure 2.
Models of current and knowledge-based systems.

range, section), date, technician zone, and number and type of PBW moths captured.

2. Data interface

Because the knowledge-based system used an existing database, a C language preprocessor was constructed to render data to a form suitable for integration. The preprocessor (Figure 3) makes note of redundant database components, creates a unique key (identifier) for each data record, and transforms some data types (e.g., changes the date from traditional to Julian).

3. Data structures encoded as objects

The system takes the field data from the C language Preprocessor Data Interface, associates it with relevant geographic information, and gives recommendations on where the next release of sterile moths should be. A module converts the relational database to an object-oriented database (Figure 3). This module is within the knowledge-based system environment. Object-orientation represents not only the data structures (as in a relational database), but the relationship between the data structures, by using

```
Example from the Relational Data Base
County, date, twn, rng, sec, zone, trap no., acres, steriles, natives
"KERN","05/14/90",34,25,23,11,312,160,89,0
"KERN","05/14/90",34,25,23,11,313,160,91,1
"KERN","05/14/90",34,25,23,11,215,160,123,0
↘
  C Language Pre-Processor
County, Julian day, yy, (TRS key = township, range, section), (Trap key =
township, range, section, trap), zone, acres, sterile PBW, native PBW);
KERN 134 90 S342523 T34252312 15 152 89 1
KERN 134 90 S342523 T34252313 16 100 91 0
KERN 134 90 S353412 T35341215 1 64 123 60 1
↘
  File Reader/ Object Converter
↘
  Object Class
(defschema trap-class
(is-a trs-class)
 (trs S342523 )
 (sterilePBW 91)
 (nativePBW 1)
 (update-tot 0)
 (cumcatch 0)
 (assignment-date 134)
 (reason-for-release SINGLE-FIND-1X)
 (release-rate 3)
 (sample-date 134)
 (update-trap 1)
 (print-total 1))
```

Figure 3.
Illustration of data flow.

active objects. An active object is one that can be affected by another object during run-time. Thus it enhances the original data representation by including the hierarchical structure and critical domain-specific data. Definition of spatial relations is done by a submodule that uses geographical relationships. The inputs to this module are three relational files (Running Totals, Buffer Data, and Geographical Data).

4. Decision-making protocol

The organization has a complex decision-making protocol, based on 20 years of experience and scientific investigations (Table 1).

The ratio of sterile to native moths is a critical factor for decision makers. The ratio of ≥ 60 sterile moths to one native moth is considered to be a good ratio. Conversely, a ratio of ≤ 59 steriles to one native is considered to be a bad ratio. Experts have discovered that a ratio of at least 60 sterile moths captured for each native moth captured will repress reproduction by the native moths.

Knowledge engineering issues are an integral part of building a knowledge-based system. Consequently, an intensive literature review was conducted in the area of pink bollworm control, related integrated pest management (IPM) issues, knowledge-based and expert systems, database systems, and AI. Also, in-depth consultations were conducted with experts involved in monitoring, data analysis, and decision making for PBW management. A rule-based, knowledge-based system was built that captures human heuristics and research-based decision paths.

The system takes advantage of its hierarchical nature and deep coupling between the rule base and the knowledge base (including the object-oriented database) to enable "scale-dependent" decisions. For example, a cotton field is a member of a section of land. The section (described by its township, range, and section numbers) is a member of a county; it could also be considered a member of a township and a member of a range. The county is a member of a State of the Union. Decisions can be made at any of these scales of knowledge. Also, certain decisions (such as how many moths to release) are designated as belonging at one scale. The existence of these scales also allows for relatively easy changes to protocol, for instance, to make pheromone treatment decisions at the field level versus the section level.

The system documents why the recommendation was made and the date of decision. This leaves a trail through the decision process and

Table 1. *All possible conditions and events and the correct decision based on that set of conditions and events.*

Condition: Found One Native Insect							
and Treatment level is:	Zero	One		Three		Blitz	
and Ratio is:		Good	Bad	Good	Bad	Good	Bad
Then New Treatment	3	3	Blitz	3	Blitz		
Condition: Found More Than One Native Insect							
and Treatment level is:	Zero	One		Three		Blitz	
and Ratio is:		Good	Bad	Good	Bad	Good	Bad
Then New Treatment	Blitz	*3*	*Blitz*	*3*	Blitz		
Treatments in italics are those that were not specified by protocol. These levels of treatment were deduced to fill the blanks. The additions were cleared through the human experts.							
Note: blitz does not exist as a condition because it is downgraded to 3x after one day and is not repeated in a particular section in any one season.							

Table 2. *Representative methods for individual insect traps used in the object-oriented management of PBW sterile release.*

Method Name	Attached to Class	Objects Accessed	Effects	Conditions for Invocation
Multiple native find in a trap in a 1x rated section, with a bad ratio of sterile to native moths.	trap-class	Type Native PBW Sterile PBW Trap-class (self) assignment-date release-rate year-of-assignment sample-date blitz-data TRS	Modifies instance of trap-class by changing the value of its attributes for rate, reason for release, assignment-date, blitz-data, and classification-based-on. Also finds other traps within that TRS and modifies their rate to the highest rate and finds the geographical neighbors.	New trap catches are read into the system. If more than one new native PBW is found in a trap rated at 1x (lowest rate), and having a bad ratio of sterile to native PBW.

Table 3. *Representative method for individual township/range/section used in the object-oriented management of PBW sterile release.*

Method Name	Attached to Class	Objects Accessed	Effects	Conditions for Invocation
Find-Neighbors-of-Blitz	TRS-class	Type Change-date Type-based-on TRS-class (self) TRS-class (neighbor)	Identifies and modifies geographical neighbors of a Blitz TRS so that the surrounding sections have a rating of at least 1x (lowest) rating.	When new data is read in, a trap in a TRS has met criteria for Blitz (highest) rating.

is a part of the documentation process that was lacking in the previous decision-making process. This protocol is encoded as rules, and governs which actions are taken (see Tables 2 and 3 for examples of scale-based rules).

5. Presentation of diagnostics

After the now object-oriented data file is processed through the heuristic rules, the recommendations are presented in a written report form (Figure 4), also stored in a database format (Figure 5). Reasons are recorded for each recommendation.

6. Diagnostics in map form

The recommendations are also presented in geographical map form. This form is most suitable for the organization's tactical use of the recommendations.

7. Simulation Model

Links to a domain-relevant simulation model are included in the knowledge-based system environment. Simulation models have been proposed for event prediction, policy-making, as surrogates for real systems, and for management and decision support (Curry and Feldman 1987, Law and Kelton 1982, Mesterton-Gibbons 1988, Shannon 1976). Recently, simulation models have become widespread and their profitability demonstrated under field conditions in agricultural management (Ladewig and Taylor-Powell 1989, Ladewig and Thomas 1992). The simulation models used in agriculture consist of systems of difference equations and their overall output is the result of the interactions among the different processes represented by these equations.

The knowledge-based system described here used the plant model GOSSYM (Baker et al.

For Juldate 199 area S112104 based on SF1XGoodRatio is a new 3x area.
For Juldate 199 area T1121040066 based on SF1XGoodRatio is a new 3x area.
For Juldate 199 area S282532 based on BufferOfBlitz is a new 1x area.

Figure 4. *Example of written report format.*

199 T2924230357 S292423 5 90 198 199 28 ONEORMore3XBadRatio 1
199 T1220290015 S122029 5 90 198 199 2 MF1XNOSTERILE 1
199 T1120040067 S112004 1 90 198 145 0 BUFFER 0

Figure 5. *Example of database format.*

1983) to verify the presence of appropriate plant stages (those that will support populations of the infesting organism, PBW). GOSSYM is a dynamic, process-level simulation of cotton growth and yield. It is based on a materials balance algorithm that keeps track of carbon and nitrogen in the plant, and water and nitrogen in the soil root zone. At present, this model is being used for the management of upwards of 400,000 hectares of commercial cotton in Mississippi. GOSSYM also has served as the basis for other crop models and is used as a research tool by scientists in the United States and abroad (McKinion and Baker 1983).

Results and Analysis

Validation of the system was conducted during development of the prototype. Validation of a system refers to "substantiating that a system performs with an acceptable level of accuracy" and does the "right" job (O'Keefe 1987). Marcot (1987) described validation as "insuring that the meaning and content of the rules meet some carefully defined criteria of adequacy." Definition of those criteria is critical to successful validation (Bellman and O'Leary 1991, Marcot 1987). Modifications were made to make each of the system components react according to protocol and in proper sequence. Criteria for adequacy of recommendations were well defined, whether recommendations were made with or without the knowledge-based system. Conflict between rules was resolved by "trimming the decision tree," for instance, modifying an object so that it will not fire the same rule twice.

The goal of the empirical verification stage was to assess if the knowledge-based system significantly improved the accuracy and precision of decision makers suffering from information input overload. Verification refers to "doing the job right" so that the program does what the programmer planned for it to do.

Data Analysis

The system was tested using several hypotheses in order to judge the impact such a system could have on an organization and on its knowledge workers. The general hypothesis is followed by specific tests. The general or null hypothesis is:

> H_o: The integrated expert system, using acceptable protocol, is not valid and does not overcome the problem of information overload in the prescribed input domain.

This general hypothesis is adapted from Robert O'Keefe's "Validating Expert System Performance" (1987). Decisions over a 24-day period (July 16 through August 10, 1990) were used. July and August were used because this is typically an active period. The data were gathered from the project database for all of 1990. Once all the data were available on a daily basis, the recommendations from the knowledge-based system were matched to their corresponding reports from the human decision makers. Note that there was a two-day time lag in decisions made by the human decision makers, due to a bottleneck at the taxonomy lab. The data from the human decision makers (project management at the Bakersfield Pink Bollworm Laboratory of USDA/CDFA) were carefully examined and compared to project protocol. Then the recommendations from the knowledge-based system were compared to project protocol. As stated, the project protocol was developed by PBW sterile release experts and was designed to be used under light, as well as heavy, infestations. The tests were:

Test 1.

Test if the expert system incorrectly identifies areas (core zones) that need to be treated (too many or too few), more so than a human manager.

Treating too few acres is considered worse than treating too many because the biological effects of letting an infestation go untreated is a great risk, while treating too many is an economic expense but not an immediate threat to the project. Each system found all 155 core zones that needed treatment. Thus, the two systems performed equally well with the knowledge-based system finding each area and documenting its reasoning.

Test 2.

Test if the expert system assigns the wrong treatment to areas (core zones) that need treatment, more so than a human manager. There were four levels at which an area could be treated. These levels were: 0 (= no treatment); lowest (= 30 moths/acre); Mid-Rate (90 moths/acre); Blitz (500 moths/acre). The results of testing are shown in Table 4.

There were several types of mistakes that had obvious differences. Mistake type 4 was the assignment of low rate to a Blitz area. Project management made 22 mistakes of this type while the knowledge-based system made one. This type of mistake could be of high risk to the biological control project in that, by under-treating, the pest population could increase to a damaging and unmanageable level. While this is not the most serious type of mistake (mistake type 1 is the most serious) it may put the project at risk.

Mistake type 5 had the highest number by project management: 45 occurrences. This mistake involves assigning the low rate to a Mid-Rate area and is slightly less serious than the preceding mistake. It also puts the project at risk by under-treating proven risk areas. These areas (those properly assigned Blitz or Mid-Rate) are "proven" high risk because they have had at least one native moth captured.

Mistakes such as type 9, the assignment of Blitz to a low rate area, is a waste of resource. It does not directly threaten the project. The indirect threat is that resources are used in a non-optimal way and, worse, that those resources

Table 4. *Comparison of decisions made by project management and knowledge-based system.*

	Number of project management mistakes	Number of knowledge-based system mistakes
Mistake Type 1: assignment of zero rate to highest rate area	1	0
Mistake Type 2: assignment of zero rate to mid-rate area	0	0
Mistake Type 3: assignment of zero rate to lowest rate area	0	0
Mistake Type 4: assignment of lowest rate to highest rate area	22	1.
Mistake Type 5: assignment of lowest rate to mid-rate area	45	0
Mistake Type 6: assignment of mid-rate to highest rate area	9	0
Mistake Type 7: assignment of highest rate to zero rate area	1	0
Mistake Type 8: assignment of highest rate to mid-rate area	1	0
Mistake Type 9: assignment of highest rate to lowest rate area	10	0
Mistake Type 10: assignment of mid-rate to zero rate area	0	0
Mistake Type 11: assignment of mid-rate to lowest rate area	0	0
Mistake Type 12: assignment of lowest rate to zero rate area	0	0

may have been needed in a true high-risk area. Ten mistakes of this kind were made, the third most common mistake. Overall, the knowledge-based system performed much better than project management in performing "core" treatment assignment.

Test 3.

The expert system incorrectly identifies areas (buffer zones) indirectly affected by core zones, more so than a human manager.

As in Test 1, failure to identify is a more serious mistake than identifying too many buffers. The first part of the test, that is, properly identifying buffers, was performed equally well by project management and the knowledge-based system. The second part of the test was comparing mistakes of including too many areas, that is, areas that were not supposed to be under treatment. Project management assigned treatment to 52 of these areas during the course of the test, while the knowledge-based system did not make any mistakes of this nature. The results indicate that resources were used to treat extra buffer areas. This draws resources away from high risk areas.

The third part of the test was examining the mistake of omission of an area. Neither system made a mistake of this type.

Test 4.

Test if the expert system assigns the wrong treatment to areas (buffer zones) indirectly affected by core zones, more so than a human manager.

Buffers have only one level of treatment, the low rate. Assignment of a higher level is considered unnecessary and a waste of resources. Downgrading from a Mid-Rate or Blitz to the low rate is a very serious mistake. This mistake is not the same as assigning a rate of low to a core zone at the time that the core zone was identified. These mistakes involve buffer areas or the mistaken downgrading of a previous core area as a buffer area, and are not accounted for in previous tests.

Project management made 31 mistakes of downgrading an area previously designated as Mid-Rate to one designated as low rate. Protocol does not call for this action to be taken, even under conditions of moth shortage. The knowledge-based system made no mistakes of this type. Downgrading is a threat to the project and a serious mistake. The second type of mistake for

Table 5. *Explanation of knowledge-based system output.*

199	T2924230357	S292423	5	90	198	199	28	OneOrMore3X BadRatio	1
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1. This is last Julian date on which that record was examined.									
2. This is the unique key for a single trap. It includes the township, range and section numbers, as well as the trap number.									
3. This is the unique key for a TRS and includes the township, range and section numbers.									
4. This designates the current release rate for that section.									
5. This is the year of assignment.									
6. This is the Julian date on which the most recent native insect was caught.									
7. This is the action date, that is the date on which the data was examined and a recommendation made.									
8. This is the cumulative total of native moths captured in that trap.									
9. This is the reason for the current release rate.									
10. This designates that the section has been blitzed and insures that the section will not be re-blitzed, in accordance with protocol.									

this test is one of waste of resources. These are the same 52 mistakes by project management that were discussed in Test 3, where project management released at the low rate on areas where protocol dictates non-release. Therefore, the knowledge-based system followed protocol better than project management in assigning release rates to buffer areas.

In examining each set of records (human and knowledge-based system), it was often confusing as to what the human's decision had been and why the decision had been made. The knowledge-based system left a decision trail so that even later it was easy to track when and why the recommendation had been made, as well as what the recommendation had been (see Table 5 for an example and explanation of the documentation).

Many of the "mistakes" recorded in the above tests were made consciously by project management. As noted, 1990 was the worst year ever for PBW in the San Joaquin Valley. The problem of information input overload occurred in three levels within the organization. At the strategic level, project management needed to guard the project from losing control of the insect population and considered 1990 to be an anomaly. Therefore, strategic decision makers felt it necessary to consciously change protocol. At the management level, the project needed to use its limited supply of sterile insects in the most efficacious manner and, in doing so, was required to examine and analyze a large quantity of data each day. Tactical operations suffered information input overload directly in that the pilots of the release planes were unable to accurately learn different, more complex flight paths on a daily basis.

Because size of the PBW population was unprecedented, protocol changed significantly several times. Some, but not all, of these changes were documented. It was not possible to separate actual mistakes from changes in protocol. In addition, some of the changes were made in order to simplify operations. This switch to a non-compensatory strategy was a symptom of information input overload. This idea was discussed with project management and they requested inclusion of several different sets of protocol to match the PBW population level. Thus, had such a knowledge-based system been in place during the 1990 year, some, but not all, of these changes to protocol would have taken place. However, the changes would have been thoroughly documented and consistently carried out.

The results of the data analysis provided three major findings. First, they showed that a heuristics-based, object-oriented, knowledge-based system evaluated project data as well, or better than, its human decision-making counterparts. Second, they demonstrated that a knowledge-based system could be used to enforce the use of project protocol, even in the most trying of circumstances and under the full effects of information input overload. Third, they illustrated the completeness and clarity of documentation provided by the knowledge-based system as compared to a human decision maker. Lack of documentation is a common problem; conversely, good documentation practices are a great benefit (Singer 1992).

Questionnaire and Feedback from the Domain Experts

The benefit of the knowledge-based system to project management was further evaluated with a questionnaire. The system was first demonstrated and thoroughly explained to each expert. The expert was invited to use the system and ask questions about any facet of the functioning of the knowledge-based system and its output. Each expert then filled out the questionnaire. Statistical analysis was not appropriate due to the small number of participants. The results of the questionnaire are tabulated beneath each question below.

1. *Project management must examine a large amount of data each day.*

• Strongly agree	3
• Agree	1
• Neither agree nor disagree	0
• Disagree	0
• Strongly disagree	0

2. *I think this system facilitates the data assimilation required in managing PBW sterile release.*

- Strongly agree 1
- Agree 3
- Neither agree nor disagree 0
- Disagree 0
- Strongly disagree 0

3. *I think this system helps prevent my overlooking important data.*

- Strongly agree 0
- Agree 2
- Neither agree nor disagree 2
- Disagree 0
- Strongly disagree 0

One expert explained that he wants to compare the results of the knowledge-based system with his own analysis over the period of a cropping season. Therefore he could neither agree nor disagree. Further, he has only overlooked two important data points in his career, so that any improvement would be incidental.

4. *I think this system accurately represents the complexity of PBW decision making.*

- Strongly agree 2
- Agree 2
- Neither agree nor disagree 0
- Disagree 0
- Strongly disagree 0

5. *I think this system makes decisions in a timely manner.*

- Strongly agree 1
- Agree 2
- Neither agree nor disagree 1
- Disagree 0
- Strongly disagree 0

The expert who neither agreed nor disagreed was uncertain that the knowledge-based system would have all the data available in a timely manner. A bottleneck exists with the taxonomy laboratory in Visalia and not all data is entered into the computer soon enough. He was not concerned that the knowledge-based system ran too slowly.

6. *I think that this system makes consistent decisions under any data load.*

- Strongly agree 0
- Agree 2
- Neither agree nor disagree 2
- Disagree 0
- Strongly disagree 0

7. *I think that this system provides good documentation for its decisions.*

- Strongly agree 2
- Agree 2
- Neither agree nor disagree 0
- Disagree 0
- Strongly disagree 0

8. *I think that this system provides adequate explanation for its decisions.*

- Strongly agree 1
- Agree 2
- Neither agree nor disagree 1
- Disagree 0
- Strongly disagree 0

9. *Project management must make release decisions each day and within time constraints.*

- Strongly agree 4
- Agree 0
- Neither agree nor disagree 0
- Disagree 0
- Strongly disagree 0

10. *Project management has to apply a complex set of protocol to the data in order to make decisions.*

- Strongly agree 3
- Agree 1
- Neither agree nor disagree 0
- Disagree 0
- Strongly disagree 0

11. *Project management is or should be responsible for complete documentation of decisions and the circumstances under which each*

decision was made.

- Strongly agree 4
- Agree 0
- Neither agree nor disagree 0
- Disagree 0
- Strongly disagree 0

12. *Project management often faces "information input overload."*

- Strongly agree 2
- Agree 2
- Neither agree nor disagree 0
- Disagree 0
- Strongly disagree 0

13. *The system applies project protocol accurately and consistently.*

- Strongly agree 2
- Agree 2
- Neither agree nor disagree 0
- Disagree 0
- Strongly disagree 0

Responses were generally positive but conservative. The users wanted to see the knowledge-based system run over a period of time in order to evaluate it completely. They agreed that they often suffer from information input overload and that they are responsible for complete and accurate documentation of circumstances and actions. Each user was anxious to have his own copy to use the knowledge-based system in his office and felt that the development and use of the knowledge-based system would lead to further favorable advances in decision making.

An additional analysis was conducted to graphically demonstrate the effects of information overload on the decision makers. Figure 6 illustrates the relationship of the amount of information that should have been used in order to make decisions according to protocol versus the actual amount used by project management. These data were collected by determining how many bits of information were required to discover the most significant native moth found in a section (taking into account the number of native and sterile moths), the neighboring sec-

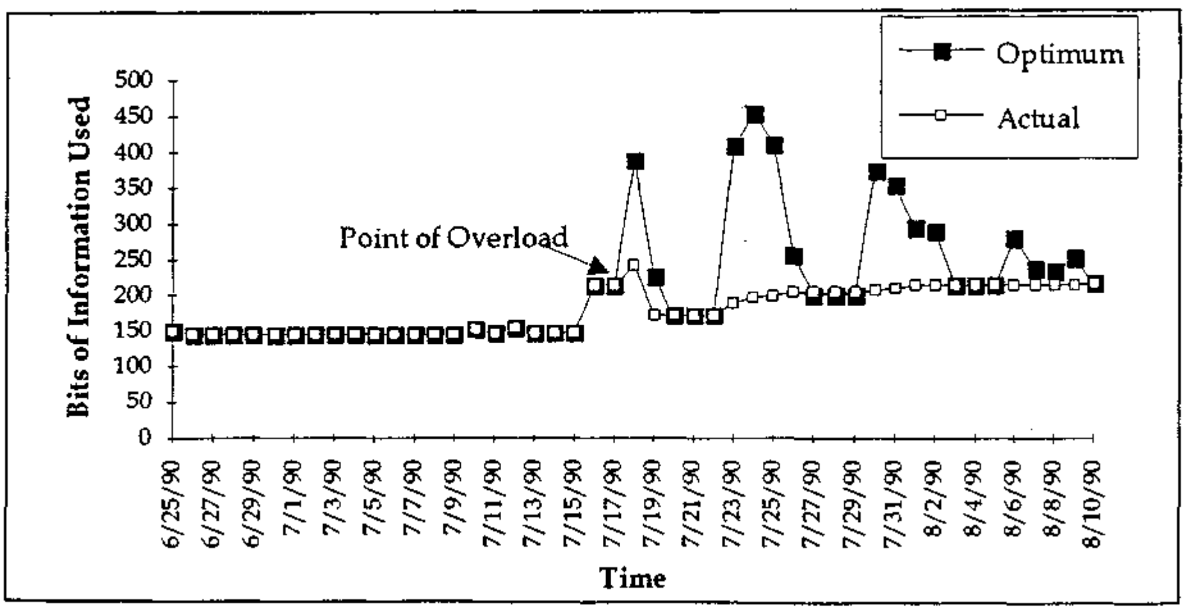

Figure 6. *Effects of information overload.*

tions and their status, if necessary, and the appropriate treatment given the previous treatments given to that section. The number of bits also includes keeping track of sections that have releases but did not have any native moths found on that day. On July 19, project management shifted decision strategies away from accepted protocol. Part of this shift was due to a shortage of moths for that day and partly due to the large amount of data that came in on that day. Moth shortages were not reported during the remainder of the study period, yet project management did not return to its former compensatory decision strategy. A shift in decision strategy from compensatory to non-compensatory is an indication of information input overload and further shows the benefits of building and using a knowledge-based system under these conditions.

Summary, Conclusions, and Recommendations

This study was undertaken because of the need of business and government for an efficient mechanism for better decisions under the all-too-common circumstance of information input overload. The knowledge-based system modeled the decision process well, used all relevant and available data, and enhanced the process by making significantly fewer mistakes and by thoroughly documenting its decision making. The knowledge-based system did not show any symptoms of information input overload. It was thus able to correctly evaluate large databases quickly (maximum run time was 20 minutes) and incorporate indirectly related pieces of data. Because the knowledge-based system used more data to support its recommendations and followed protocol, it reduced risk as well. The knowledge-based system produced a database that could go directly into a GIS to produce maps of the affected agricultural areas and the rates of treatment to those areas. The knowledge-based system also provided links to related insect and crop simulation models.

The extension of knowledge-based systems specifically to relieve information input overload for management is mandated by the ever increasing ability of machines to collect data (Weizer et al. 1991) and the relatively static ability of humans to process it (Miller 1960). The use of a knowledge-based system solves the information overload problem encountered by the strategic level in that users could quickly look at what current protocol recommends and why. In addition, another decision branch could easily be added to the knowledge-based system containing the new protocol used in situations similar to 1990. The knowledge-based system solves the management level problem in that it quickly applies protocol to the data and provides documentation.

The logical extension of this research to develop a geographical display of the treatment recommendations should go a long way to solving the tactical problems. The knowledge-based system produces a data file that is designed to go directly into a GIS. The GIS would produce a map showing local landmarks and the cotton fields. The fields would be color or pattern coded to indicate the treatment level required. This should reduce the information input load on the pilots and facilitate adherence to project protocol, even under the most trying of conditions.

In conclusion, a knowledge-based system is very effective at making consistent decisions, whether under the influences of information input overload or not. It was able to correctly evaluate large databases quickly, incorporate indirectly related pieces of data, and follow project protocol. Whereas many knowledge-based system applications have appeared in the literature, this research takes an analytical approach to their implementation and examines how specific MIS issues are addressed by not only expert systems but other important components of decision making such as object-oriented representations.

Acknowledgments

The authors are grateful to Reg Rosander, Bill Abel and Dan Keaveny for expert and technical advice; to Robert Jimersen, Eileen Kenny, Matthew Lee, and Nora McCaw for geographical data entry and error checking; and to Liesl Carland for data entry and for her able editing of the manuscript.

References

Ackoff, R.L. 1967. Management misinformation systems. Management Science 14(4): B147-B156.

Amarel, S. 1991. Themes and directions of artificial intelligence: Opportunities and challenges. Plenary speaker, AAAI-91 National Conference on Artificial Intelligence. Anaheim, California, July 14-19, 1991.

Baker, D.N., J.M. McKinion, and J.R. Lambert. 1983. GOSSYM: A simulator of cotton crop growth and yield. South Carolina Agricultural Experimental Station Technical Bulletin 1089, 134 pages.

Beck, H.W., and J.W. Jones. 1989. Simulation and artificial intelligence. Pages 141-160 in: Knowledge Engineering in Agriculture, J.R. Barrett and D.D. Jones, editors. American Society of Agricultural Engineers Monograph number 8, American Society of Agricultural Engineers, St. Joseph, Michigan.

Bellman, K.L., and D.E. O'Leary. 1991. Knowledge-based systems verification, validation, and testing. Pages MP5-1-MP5-193 in: Proceedings, 9th National Conference on Artificial Intelligence, Anaheim, California, July 1991.

Biggs, S.F., J.C. Bedard, B.G. Gaber, and T.J. Linsmeier. 1985. The effects of task size and similarity on the decision behavior of bank loan officers. Management Science 31: 1-18.

Booch, G. 1991. Object Oriented Design with Applications. The Benjamin/Cummings Publishing Company, Incorporated, Redwood, California.

Cook, G.J. 1987. An analysis of information search strategies for decision making. Ph.D. dissertation, Department of Decision Sciences, Arizona State University, Tempe, Arizona.

Courtney, J.F., and D.B. Paradice. 1988. Database Systems for Information Management. Times Mirror/ Mosby College Publishing, St. Louis, Missouri.

Curry, G.L., and R.M. Feldman. 1987. Mathematical foundations of population dynamics. Texas A&M University Press, College Station, Texas.

Einhorn, H.J. 1971. Use of nonlinear, non-compensatory models as a function of task and amount of information. Organizational Behavior and Human Performance 6: 1-27.

Feigenbaum, E., P. McCorduck, and H.P. Nii. 1989. The rise of the expert company: How visionary companies are using artificial intelligence to achieve higher productivity and profits. Vintage Books, division of Random House, New York.

Garvey, M.M., and M.S. Johnson. 1989. Introduction to object-oriented databases. Information and Software Technology 31(10): 521-528.

Gorry, G.A., and S. Morton. 1971. A framework for management information systems. Sloan Management Review 13(1): 55-70.

Grant, J. 1987. Logical Introduction to Databases. Harcourt Brace Jovanovich, Publishers, and Academic Press, New York.

Harmon, P., and D. King. 1985. Expert systems: Artificial intelligence in business. John Wiley, New York.

Henderson, J.C. 1987. Finding synergy between decision support systems and expert systems research. Decision Sciences 18: 333-349.

Hiltz, S.R., and M. Turoff. 1985. Structuring computer-mediated communication systems to avoid information overload. Communications of the Association of Computing Machinery 28: 680-689.

Hogarth, R.M. 1981. Beyond discrete biases: functional and dysfunctional aspects of judgment heuristics. Psychological Bulletin 90: 197-217.

Johnson, E.J., and J.W. Payne. 1985. Effort and accuracy in choice. Management Science 31(2): 395-414.

Keen, P.G.W. 1980. Adaptive design for decision support systems. Sloan Management Review 12(2): 15-25.

Klayman, J. 1985. Children's decision strategies and their adaptation to task characteristics. Organizational Behavior and Human Decision Making 35: 179-201.

Ladewig, H., and E. Taylor-Powell. 1989. An assessment of GOSSYM/COMAX as a decision support system in the U.S. cotton industry. Texas Agricultural Extension Service, College Station, Texas.

Ladewig, H., and J.K. Thomas. 1992. A follow-up evaluation of the GOSSYM/COMAX cotton program. Texas Agricultural Extension Service, College Station, Texas.

Law, A.M., and W.D. Kelton. 1982. Simulation modeling and analysis. McGraw-Hill, New York.

Lee, J.A. 1984. Cotton as a world crop. Pages 1-25 in: Cotton, R.J. Kohel and C.F. Lewis, editors. American Society of Agronomy, Incorporated, Crop Science Society of America, Incorporated, and Soil Science Society of America, Incorporated, Madison, Wisconsin.

Lindsjorn, Y., and D. Sjoberg. 1988. Database concepts discussed in an object oriented perspective. Pages 300-318 in: Proceedings, European Conference on Object-Oriented Programming, Oslo, Norway, August 15-17, 1988. Springer-Verlag, Berlin.

Lindskov-Knudsen, J., and O. Lehrmann-Madsen. 1988. Teaching object oriented programming languages. Pages 21-40 in: Proceedings, European Conference on Object-Oriented Programming, Oslo, Norway, August 15-17, 1988. Springer-Verlag, Berlin.

Marcot, B. 1987. Testing your knowledge base. AI Expert 2(8): 42-47.

Mason, R.O., and I.I. Mitroff. 1973. A program for research on management information systems. Management Science 19(5): 457-487.

McKinion, J.M., and D.N. Baker. 1983. Dynamic crop modeling: a synergism of computers, experimental research, and the scientific method. Pages 502-509 in: Analysis of ecological systems: State of the art in ecological modeling, W.K. Lauenroth, G.V. Sogerboe, and M. Flugg, editors.

Metcalf, R.L. 1980. Changing role of insecticides in crop protection. Annual Review of Entomology 25: 219-256.

Mesterton-Gibbons, M. 1988. A concrete approach to mathematical modeling. Addison-Wesley Publishing Company, Reading, Massachusetts.

Meyer, B. 1988. Object-oriented software construction. Prentice-Hall, New York.

Miller, J.G. 1960. Information input overload and psychopathology. American Journal of Psychiatry 116: 695-704.

O'Keefe, R. 1987. Validating expert system performance. IEEE Expert 2: 81-89.

Paradice, D.B., and J.F. Courtney Jr. 1989. Organizational knowledge management. Information Resources Management Journal 2(3): 1-13.

Payne, J.W. 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. Organizational Behavior and Human Performance 16: 366-387.

Pfadt, R.E. 1978. Fundamentals of applied entomology, 3rd edition. Macmillan Publishing Company, New York.

Reddy, V.R. 1987. Epistemology of knowledge based simulation. Simulation 48(4): 162-166.

Shannon, R. 1976. Simulation: The art and the science. Prentice Hall, New York.

Sheridan, T.B., and W.R. Ferrell. 1974. Man-machine systems: Information, control, and decision models of human performance. Massachusetts Institute of Technology Press, Cambridge, Massachusetts.

Simon, H.A. 1955. A behavioral model of rational choice. Quarterly Journal of Economics 69: 99-118.

Simon, H.A. 1960. The new science of management decisions. Prentice-Hall Publishing, Englewood Cliffs, New Jersey.

Singer, L.M. 1992. McGraw-Hill guide to effective communications for MIS professionals. McGraw-Hill, New York.

Sprague, R.H. 1980. A framework for the development of decision support systems. Management Information Systems Quarterly 4: 1-26.

Stabell, C.B. 1983. A decision oriented approach to building decision support system. Pages 222-233 in: Building Decision Support Systems, J.L. Bennett, editor. Addison -Wesley Publishing Company, Reading, Massachusetts.

Stone, N.D., R.N. Coulson, R.E. Frisbie, and D.K. Loh. 1986. Expert systems in entomology: Three approaches to problem solving. Bulletin of the Entomological Society of America 32: 161-166.

Tello, E.R. 1989. Object-oriented programming for artificial intelligence. Addison-Wesley Publishing Company, Reading, Massachusetts.

Tsichritzis, D.C., and O.M. Nierstrasz. 1988. Fitting round objects into square databases. Pages 283-298 in: Proceedings, European Conference on Object-Oriented Programming, Oslo, Norway, August 15-17, 1988. Springer-Verlag, Berlin.

Tversky, A. 1969. Intransitivity of preferences. Psychological Review 76: 31-48.

Tversky, A. 1972. Elimination by aspects: A theory of choice. Psychological Review 79: 281-299.

Volonino, L., C.C. Sekar, and R. Ramesh. 1987. Decision quality and information load. Pages 628-637 in: Proceedings, 20th Annual Hawaii International Conference on System Sciences, Honolulu, Hawaii, January 6-9, 1987. Western Periodicals, North Hollywood, California.

Weizer, N., G.O. Gardner III, S. Lipoff, M.F. Roetter, and F.G. Withington. 1991. The Arthur D. Little forecast of information technology and productivity. John Wiley, New York.

Michelle Walters Geyer is a computer specialist at the USDA/APHIS Methods Development Center in Phoenix, Arizona. She received a B.S. in agri-business at Arizona State University in 1976, an M.S. in entomology from Texas A&M University in 1986, and a Ph.D. from the latter in business analysis in 1993. She is currently working on information/expert systems for cotton pests, on development of information acquisition, analysis, documentation, and reporting for temperature treatments of fruit for export, and data sharing systems for two sterile insect rearing facilities.

Ronaldo A. Sequeira is a computer scientist at USDA/ARS Crop Simulation Lab in Mississippi State, Mississippi. He received his B.S. from the University of Florida, his M.S. from Texas A&M University, and his Ph.D. from the latter, in 1990, with the Industrial Engineering Department where he specialized in biosystems engineering. He is currently involved with the development of genetic algorithms, artificial life systems, and computer simulation of biological processes that are relevant for managed biosystems.

David B. Paradice is an associate professor in the Business Analysis Department at Texas A&M University. He received a B.S. in informaiton and computer sciences and an M.S. in industrial management from the Georgia Institute of Technology. He received his Ph.D. in business administration from Texas Tech University. His research interests include the application of emerging technologies, such as animation, sound, and video, to computer-based business processes and in support of managerial problem forumulation, as well as the influence of com-

puter-based systems on ethical decision-making processes.

George C. Fowler is a professor in the Business Analysis Department at Texas A&M University. He received his B.S. in 1969, M.B.A. in 1969, and Ph.D. in 1976 from Texas A&M University. His research interests include information systems education and curriculum development, small business systems and end user computing, and programming complexity. He also has strong interests in teaching and service and is his department's corporate liaison.

Charlotte Miller is Chief, Information Resources Management Branch, USDA, APHIS, ISCD in Hyattsville, Maryland. She received a B.A. in history from Carlton, and in 1976 earned a Ph.D. from the University of Florida with a certificate in Latin American Studies. Her recent interests include strategic informaiton resource planning for APHIS, including planning for business process reengineering efforts in that agency.

Robert T. Staten is the Center Director at the USDA/APHIS Methods Development Center in Phoenix, Arizona. He received his B.S. in entomology from New Mexico State University in 1960, an M.S. in 1965 and Ph.D. from University of California at Riverside in 1970. He is involved in the development and use of PBW pheromones for survey and mating disruption, sterile insect technology, and combining these two areas into a biorational management system for the eradication of the PBW from the western cotton growing regions. He is also involved in the rearing and release of beneficial insects to control the sweet potato whitefly.

SUBSCRIPTION INFORMATION

AI Applications is published three times a year in April, July, and October. Prepaid individual subscriptions $37; government agencies, university departments, libraries, private companies, and foreign subscriptions (including Canada) US$72 (US$16 extra optional for air mail to foreign addresses). Send subscription requests to: AI Applications, P.O. Box 3066, Moscow, Idaho 83843 USA (telephone 208-885-7033, fax 208-885-6226). Back issues are also available. Please write for information.