UMI Number: 9528625

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

A dissertation entitled

A Knowledge Based Decision Support System using Distributed Artificial Intelligence

submitted to the Graduate School of the
University of Wisconsin-Madison
in partial fulfillment of the requirements for the
degree of Doctor of Philosophy

by

Rahul Bhaskar

Degree to be awarded: December 19___ May 19___ August 19 95

Approved by Dissertation Readers:

Major Professor

May 22, 1995
Date of Examination

Dean, Graduate School

A Knowledge-Based Decision Support System using Distributed Artificial Intelligence

by

Rahul Bhaskar

A dissertation submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

(BUSINESS)

at the

UNIVERSITY OF WISCONSIN - MADISON

1995

ACKNOWLEDGMENTS

I would like to take this opportunity to thank the people who made this thesis possible. Without them, I would have not completed this project. They were my guides, inspiration and helpers during many vicissitudes that my life presented.

First and foremost, I would like to thank Professors Amit Gupta and S. C. Park for giving the guidance and their help throughout the process. I could not have done this without them. They were not only my academic advisors but friends who cared for my professional and personal growth.

My deepest gratitude to my other committee members, Professor Jim Morris, Professor Scott Webster, and Professor Jeffrey Russell for their time, advice, and motivation.

I would also like to thank Professor Robert Miller who helped and encouraged me with his guidance, motivation, wisdom, and insights into the world of statistics. Special thanks to Professor Jim Morris. His gentleness, knowledge and experience were always reassuring through the thick and thin of the years as a graduate student.

To the most important people in my life, those who always inspired me and who I love the most, I dedicate this work to my parents, sisters and their families. Their strong belief in me and profuse love keep me warm during many cold storms.

ACKNOWLEDGMENTS

I would like to take this opportunity to thank the people who made this thesis possible. Without them, I would have not completed this project. They were my guides, inspiration and helpers during many vicissitudes that my life presented.

First and foremost, I would like to thank Professors Amit Gupta and S. C. Park for giving the guidance and their help throughout the process. I could not have done this without them. They were not only my academic advisors but friends who cared for my professional and personal growth.

My deepest gratitude to my other committee members, Professor Jim Morris, Professor Scott Webster, and Professor Jeffrey Russell for their time, advice, and motivation.

I would also like to thank Professor Robert Miller who helped and encouraged me with his guidance, motivation, wisdom, and insights into the world of statistics. Special thanks to Professor Jim Morris. His gentleness, knowledge and experience were always reassuring through the thick and thin of the years as a graduate student.

To the most important people in my life, those who always inspired me and who I love the most, I dedicate this work to my parents, sisters and their families. Their strong belief in me and profuse love keep me warm during many cold storms.

To my uncles, Mr. J. C. Bhasker, Dr. S. K. Saroj, and their families who have helped me in every aspect and have kept me company by calling and visiting frequently, I say Thank You. Your love has been an unending source of joy in my life.

Thanks to my landlubber friends in Madison and elsewhere (Dan and Ruth Bachhuber, Robin Matthews, and Seema Kapani). I appreciate the comments made by Dan, Ruth and Greg in helping me finish this thesis.

I salute all my sailor friends (including those in heaven) for keeping me company on many long and lonely voyages and helping me survive many sea storms and pirates. I will always be a sailor at heart because of them! I pray that we will all find the haven of our final destination.

My special thanks to Mr. David Matthews, special-agent-in-charge at the Wisconsin Division of Narcotics Enforcement, Madison office, for being my domain expert in the field of drug enforcement and for his enriching friendship and brotherhood.

There are many more people than I can mention, who made this research and thesis possible. Thank You All!

Madison, Wisconsin

Rahul Bhaskar

May 1995

To my uncles, Mr. J. C. Bhasker, Dr. S. K. Saroj, and their families who have helped me in every aspect and have kept me company by calling and visiting frequently, I say Thank You. Your love has been an unending source of joy in my life.

Thanks to my landlubber friends in Madison and elsewhere (Dan and Ruth Bachhuber, Robin Matthews, and Seema Kapani). I appreciate the comments made by Dan, Ruth and Greg in helping me finish this thesis.

I salute all my sailor friends (including those in heaven) for keeping me company on many long and lonely voyages and helping me survive many sea storms and pirates. I will always be a sailor at heart because of them! I pray that we will all find the haven of our final destination.

My special thanks to Mr. David Matthews, special-agent-in-charge at the Wisconsin Division of Narcotics Enforcement, Madison office, for being my domain expert in the field of drug enforcement and for his enriching friendship and brotherhood.

There are many more people than I can mention, who made this research and thesis possible. Thank You All!

Madison, Wisconsin

Rahul Bhaskar

May 1995

ABSTRACT

A KNOWLEDGE-BASED DECISION SUPPORT SYSTEM USING

DISTRIBUTED ARTIFICIAL INTELLIGENCE

Rahul Bhaskar

Under the supervision of Assistant Professor Amit Gupta

At the University of Wisconsin - Madison

The main purpose of this research is to use selected tools and methodologies from the distributed artificial intelligence (DAI) field to create a functioning knowledge-based decision support system (KBDSS). This KBDSS performs telephone analysis and financial analysis. Furthermore, an observational study is conducted to compare the prototype of the KBDSS, which we call Sherpa, to the system currently in use.

The architecture of the prototype comprises a set of problem solvers, all operating under the guidance of meta-level knowledge. Each problem solver works as a classifier system. The problem solving knowledge of each solver is developed using machine learning and/or knowledge engineering.

As part of the observational study, variables are identified and hypotheses formulated. Statistical analysis is carried out to test the hypotheses. It is found that Sherpa outperformed the currently used system in all aspects tested.

TABLE OF CONTENTS

LIST OF FIGURES

To my parents Tilak and Sudarshan, for the gift of this life, and once again, to the great monkey whose blanket continually protects me.

"There are not enough jails, not enough policemen, not enough courts to enforce a law not supported by the people" (Hubert Humphery, 1/1/65, Speech at Williamsburg, VA)

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

The main thrust of this research is to use selected tools and methodologies from the distributed artificial intelligence (DAI) field to create a functioning knowledge-based decision support system (KBDSS). This KBDSS is designed to perform telephone record and financial analyses[1]. Furthermore, an observational study was conducted to compare the prototype of the KBDSS, which we call 'Sherpa', to the system in current use at the Wisconsin Division of Narcotics Enforcement (DNE).

Any drug investigation has many steps. It includes target selection, data collection, data evaluation, analysis (including telephone record and financial analyses) and dissemination of information and data. The specific area of drug investigation automated by Sherpa in this research is using the telephone and financial analyses to classify the suspect into one of the five categories: 1) Source, 2) Customer, 3) Broker 4) Pursue (for more information), and 5) Non-Suspect.

[1] Please refer to Appendix A for detailed description of a drug investigation and financial and telephone analysis.

1.2 Distributed Artificial Intelligence

DAI is a subfield of artificial intelligence (AI) concerned with the concurrence in AI computations. The DAI architecture has proven to be very beneficial in the development of advanced and complex reasoning systems (Laasri and Maitre 1989). DAI is concerned with the cooperative solution of problems by a decentralized group of agents. The agents may range from a simple processing element to complex entities. The problem solving strategy of these systems is cooperative. This means that mutual sharing of information is necessary to allow the group as a whole to generate a solution. The group of agents is decentralized because both control and data are logically and often geographically distributed (Taha et al., 1994). DAI offers increased speed, reliability and ability to handle applications with a natural spatial distribution, as well as the ability to tolerate uncertain data and knowledge (Smith and Davis, 1981). Because such systems are highly modular, they also offer conceptual clarity as well as simplicity of design and maintenance.

1.3 Research Procedures

There are two parts to this research. The first part is concerned with the construction of the system and the second part with comparing the current system with Sherpa. The research aspect of constructing Sherpa is briefly described in the next section followed by an outline for comparing the current system to Sherpa in section 1.2.2. These aspects will be described in complete detail in Chapters 3 and 4, respectively.

1.3.1 Construction of Sherpa

Drug investigation involves breaking down the overall problem into a subset of problems that are simpler to solve. Then, the problems in this subset are solved individually. The solutions obtained are combined together to form the overall answer. Designed as a KBDSS that can solve the problems in a similar manner, Sherpa consists of individual problem solving modules. These modules were designed using the following methodologies:

1. Knowledge Engineering: Structured interviews, observations and other methods of knowledge engineering were used to elicit information about the drug investigation process as well as the work dynamics of the DNE.

2. Compiling Rules by Induction: The specific examples collected in the interviews were generalized into rules by using an induction algorithm.

3. Developing Problem Solvers: In this step, the system's other specialized problem solvers were developed.

These steps to construct Sherpa are described in complete detail in Chapter 3.

1.3.2 Study to Compare Present System with the Prototype of Sherpa

Statistical analysis was done to compare the current system with Sherpa. The four variables that were measured to perform this analysis were:

Time: Time is the duration it takes to obtain telephone and financial analysis after the receipt of the data by the DNE.

Leads: A "*Lead*" is defined as the information suggesting whether to pursue further investigative activity and the nature of that activity. This variable was measured by noting the number of leads to classify the suspected drug dealers using the current system and Sherpa.

Identification Frequency: This is the number of drug dealers belonging to the middle and upper levels.

Evidence: This variable measures the amount of evidentiary information to determine the classification of the suspect. This is the information that an agent can use to confirm to other personnel his convictions about the suspect's classification.

The following hypotheses based on the four variables listed above were developed.

<u>Hypothesis 1</u>: An agent will be able to obtain telephone and financial records in less time with Sherpa than with the current system.

<u>Hypothesis 2</u>: Sherpa will help the narcotics agent identify a greater number of leads for further investigation than the current system.

Hypothesis 3: The information provided by Sherpa will help identify more middle and upper level drug dealers.

Hypothesis 4: The amount of evidentiary information provided by Sherpa will be higher than the current system.

To test these hypotheses, an experimental study was conducted by introducing a functional prototype of the Sherpa system at the Appleton office of the Wisconsin DNE. The Madison office of the DNE was treated as the control group. This choice was justified because the personnel at the Madison DNE office helped in developing Sherpa. There are a total of 15 agents at the Madison (8) and the Appleton (7) DNE offices. The demographics of regions covered by these offices are also comparable. The description of the components of the two offices and the state regions under their jurisdiction are highlighted in Chapter 4.

Actual cases were solved by the agents using the Sherpa system and the current system. Comparisons were made and the values of the dependent variables were determined. Detailed results of this comparative study are described in Chapter 4. In Chapter 3, a description of the construction of Sherpa using the steps described in section 1.3.1 is given. Chapter 2 discusses the literature relevant to this research.

CHAPTER 2

LITERATURE SURVEY

This chapter provides the literature survey on various methodologies used in this research. In section 2.1, the literature survey of the distributed problem solving methodologies is provided. This is followed by section 2.2, which provides a survey of literature in the methodologies of inductive learning and knowledge engineering. Then, Section 2.3 provides the literature survey in the area of constructing classification models.

2.1 Distributed Problem Solving System

This research is based on the concepts of distributed problem solving (DPS) which is a subsystem of distributed artificial intelligence (DAI). One of the defining characteristics of DPS is that it is a cooperative activity of a group of decentralized and loosely coupled knowledge sources (KSs) or problem solvers. The KSs cooperate in the sense that none of them has sufficient information to solve the entire problem by itself. Therefore, they have to work as a team and share knowledge about the tasks, results, goals, and constraints to solve the problem. "Decentralized" in this context means that both control and data are logically and often geographically distributed (i.e., there is no global control or global data storage). "Loosely coupled" in this context means that individual KSs spend most of their time in computation rather than communication (Davis and Smith, 1983).

In a pure distributed problem-solving process, the problem is divided into tasks, and special task performers (agents) are designed to solve these tasks for this specific problem. All interaction (such as cooperation, coordination if any) strategies are incorporated as an integral part in the design of the system.

Most problem-solving processes in a DPS system consist of four phases. These are as follows (Shaw and Whinston, 1989):

1. The decomposition of the problem into sub-problem tasks;
2. The allocation of the sub-problem tasks among the agents;
3. The solving of the sub-problem tasks among the agents;
4. The integration of the solutions, obtained in phase 3, to obtain the global solution.

A DPS system has significant advantages over a single, monolithic, centralized problem solver. These advantages include: 1) faster problem-solving by exploiting parallelism, 2) reducing necessary communication by transmitting only high-level partial solutions rather than raw data to a central site, 3) increased flexibility by creating problem solvers with different abilities dynamically combined, to solve the current problem, and 4) increased reliability by allowing other problem solvers to replace failed ones (Durfee et al., 1989).

To date, there has been insufficient research on distributed problem solving systems. Corkhill (1982) has been developing self organizing problem-solvers, but his work along these lines is still incomplete (Pattison et al., 1987). On the other hand, Shaw and Whinston (1989) have provided interesting new techniques by introducing a market mechanism through which agents can compete.

Werner (1989) has developed a unified theory of communication, cooperation and social structure that fulfills a necessary condition for the design of cooperating complex agents. Star (1989) has suggested basing DAI on a social metaphor, using boundary objects as the data structure for intuition. Gasser et al., (1989) has developed a new framework for representing and re-using organizational knowledge.

New theories, both pragmatic and abstract, are emerging from recent work. Sathi and Fox (1989) apply constraint-direct reasoning to negotiations about resource allocation in scheduling problems. By examining probabilistic interactions among agents, they are extending the earlier work of Rosenchein et al., (1986) on models of rationality and cooperation without communication. In addition, several researchers are beginning to urge that DAI research use social theories as a methodological and epistemological formulation. Again, new theories of large scale dynamic systems, genetic systems and human organizations promise global insights for systems that can be modeled in realistic simple terms.

A wide range of perspectives exists on planning in multi-agent worlds. Agents can be organized along a spectrum from those dealing with simple, mathematically characterizable agents working in abstract and constrained worlds to approximately useful but theoretically uncharacterized multi-agent planners.

Decker et al. (1989) provide a collection of 73 dimensions with which to describe any particular DAI system or theory, as well as a metric for each dimension. Their schema allows different systems to be described and compared in a large feature space. For example, they compare their method of partial global planning with the contract net system of Davis and Smith (1983).

Finally, an array of methodological approaches is appearing and being debated in DAI research. Nii et al., (1990) have reported numerous experimental studies to evaluate their multilevel concurrence control mechanism in CAGE and POLIGON. Star (1989) suggests the "Durkheim test" for "social utility" to replace the Turing test for individual intelligence. Her earlier investigation of the construction of robust knowledge in scientific practice suggests ways of refining the units of analysis and objects of study for DAI. To paraphrase Star (1989), researchers in DAI must have "open theories" that recognize the inherently multi-perspective nature of knowledge and account for the processes of theory revision. Such open theories could provide the boundary objects through which the systems would interact.

2.2 Inductive Learning and Knowledge Engineering:

Two methods used to create knowledge sources in Sherpa are inductive learning and knowledge engineering. The survey of the literature related to these methodologies is presented below.

2.2.1 Inductive Learning:

Most applications of artificial intelligence are based on constructing a model of knowledge used by a human expert. This approach has had a major impact mainly since the early 1980s and is illustrated by numerous case studies reported in "The rise of Expert Company" (Fiegenbaum et al., 1988). In some cases the task that an expert performs can be thought of as "classification"- assigning things to categories or classes determined by their properties. For instance, Feigenbaum et al., (1989) cite a system developed by American Express to assist credit authorizers. The primary properties for this application are details of a proposed transaction and the particular customer's credit history. The classes correspond to a recommendation to approve or to decline the transaction. Expert systems at Texas Instruments (TI) that help in the preparation of capital expenditure proposals can also be viewed as classification-based; each item on the proposal must be judged to be consistent or inconsistent with TI's policies.

The book by Quinlan (1992) describes a set of computer programs that construct a classification model by discovering and analyzing patterns found in a set of records. The key requirements for tasks that lend themselves to this inductive approach are as follows:

1. **Attribute Value Descriptions:** The data to be analyzed must be a flat file - all information about one object or a case must be expressible in terms of a fixed collection of properties or attributes. Each attribute may have either discrete or numeric values, but the attributes used to describe a case must not vary from one case to another.

2. **Predefined Classes**: The categories to which cases are to be assigned must have been established a priori. In machine learning, this is deemed *supervised* learning, as

compared to *unsupervised* learning in which appropriate groups of cases are found by analysis. Fisher et al., (1991) provide a thorough treatment of unsupervised learning.

3. Discrete Classes: This requirement has to do with the categories to which cases are assigned. The classes must be sharply delineated. A case either does or does not belong to a particular class and the number of cases must be much greater than the number of classes. One group of tasks that does not have discrete classes is concerned with the prediction of continuous values such as the price of gold or the temperature at which an alloy will melt. Similar tasks in which a continuous-valued class is broken into vague categories such as hard, quite hard, flexible, quite soft, soft should be approached with caution. Friedman (1988) and Brieman et al., (1984) describe analogous methods for handling continuous classes.

4. Sufficient Classes: Inductive generalization proceeds by identifying patterns in data. As this differentiation usually depends on statistical tests of one kind or another, there must be sufficient cases to allow these tests to be effective. The amount of data required is affected by factors such as the number of properties and classes and the complexity of the classification model. As these factors increase, more data is needed to construct a reliable model.

5. "Logical" classification models: Decision trees or production rules essentially restrict the description of a class to a logical expression whose primitives are statements about the values of particular attributes. One common form of classification model that does not satisfy this requirement is the *linear discriminant* (Nilsson, 1965) in which weighted contributions of attributes are combined arithmetically and compared to a threshold. The corresponding descriptions of classes are thus arithmetic rather than logical.

2.2.2. Knowledge Engineering

Mockler and Dologite, (1992) describe knowledge engineering as "a term that is used to describe the knowledge-based systems (KBS), developmental job of situation analysis and representation, and computer system design and implementation."

Knowledge engineering varies according to the need of the project. When experts are developing their own system using an expert system shell, one person can do all the jobs involved in the knowledge engineering: acquiring and defining knowledge (situation analysis), formulating and representing knowledge (situation representation and system design), and engineering and programming the software (system design and development).

In many situations, a knowledge engineer works with experts and system users, and performs additional planning and managing tasks. In these situations, the knowledge engineering can involve interpersonal tasks, such as interviewing, as well as many other tasks involved in coordinating the experts, the users and various technical people who may be required for developing the actual system. There are many knowledge acquisition techniques available to a knowledge engineer (Boose 1988, Boose 1989). Some of the more commonly used techniques are:

1. Techniques for observing the experts.
2. Techniques for representing a domain expert's overall decision making approach. (Grabowski, 1988).
3. Techniques used in gathering data.
4. Techniques for acquiring and formulating knowledge (manual and automated).
5. Techniques used in prototyping.

These are explained in detail in Chapter 3.

2.3. Classification Models

Another methodology used in this research concerns classification models. The purpose of a classification model is to develop accurate predictors for future cases. Many times the decision tree obtained by an inductive learning program is hard to understand. Several ways of surmounting this comprehension barrier have been explored. Shapiro (1987) for instance breaks down a single large tree into a hierarchy of smaller trees. His *structured induction* approach can lead to trees that, individually and collectively, are easier to understand.

C4.5, (Quinlan, 1992) produces a classification model as *production rules*, a format that appears to be more easily understandable than trees. This program uses a simplified form of production rule L -> R in which the left-hand side L is a conjunction of attribute-based tests and the right-hand side R is a class. A default class (one for which no rule's left-hand side is satisfied) is also designated. To classify a case using a production rule model, the ordered list of rules is examined to find the first case in which the left-hand side is satisfied by the case. The predicted class is the one nominated by the right-hand side of this rule. If no rule's left-hand side is satisfied, the case is predicted to belong to the default class. In C4.5, a module of the program examines the (original) decision tree produced by C4.5 and derives from it a set of production rules.

There are many kinds of classifiers in addition to an inductive learning classifier such as C4.5. An overview of these is provided next.

A. Instance-based classifiers

The underlying philosophy in the instance-based systems is to classify a case by recalling a similar case whose class is known and to predict that the new case will have the same class. Following are the central issues of instance-based systems: (Quinlan, 1992)

- What training cases should be remembered? Too many cases retained will cause the system to be unwieldy and slow. The ideal situation would be to retain prototypical cases that together summarize all the important information. This is an approach that can be observed in medical and legal textbooks.

- How can the similarity of cases be measured? If all the attributes are continuous, we can compute the distance between two cases as the square root of the sum of squares of attribute differences, perhaps first scaling attributes so that their contributions are comparable. When some attributes are not continuous, however, the interpretation of this distance is difficult. Moreover, if there are many irrelevant attributes, two similar cases may appear to be quite dissimilar because they have different values of unimportant attributes.

- How should the new case be related to the remembered cases? Two alternatives are to use the single most similar remembered case, or to use several similar cases with predictions weighted by their differing degrees of similarity.

B. Neural Networks

In its modern form, a neural network consists of units connected by links. Three kinds of units are distinguished: input units, which introduce information from the environment to the network; output units, which give results; and all other units, which are "hidden" from the environment. Each link has an associated weight and some units have a bias that appears below the unit. To process a case, the input units are first assigned numbers between 0 and 1 representing the attribute values. Each unit's input, *I,* is determined as the sum of the weighted output of the units connected into it, plus the unit's bias, and the unit's output determined as:

$$1/(e^{-I}+1) \qquad (2.1)$$

which ranges from 0 to 1 as I ranges from minus infinity to infinity.

The values of network weights and biases are learned through repeated examination of the training cases. The deviation of each output unit's output from its correct value for each case is *back-propagated* through the network; all relevant connection weights and unit biases are adjusted, using gradient descent, to make the actual output closer to the target. The training continues until the weights and biases stabilize.

Normal networks can be used to predict real values, not just classes. In a classification context, however, the predicted class is encoded in some way. For instance, if the number lies between 2^{b-1} and 2^{b} so that b bits are required to identify a class, each of b output units can be made to represent one bit of the class code. Alternatively, one output unit can be associated with every class. (Quinlan, 1992)

C. Statistics

Statistical methods provide a rich body of knowledge for classification purposes. One example is maximum likelihood classifiers that often use normally distributed attribute values to determine the distribution means, variance, and covariance. As a general rule, however, statistical techniques tend to focus on tasks in which all the attributes have continuous or ordinal values. Many of them are also *parametric*, assuming some form for the model and then finding appropriate values for the model's parameters from the data. For instance, a linear classifier assumes values, then finds the particular linear combination that gives the best fit throughout the training cases. Maximum likelihood classifiers often assume that attribute values are normally distributed, then use the training data to determine the distribution means, variances and covariance (Sherphered et al., 1988).

CHAPTER 3

METHODOLOGIES TO CONSTRUCT SHERPA

In this chapter, we describe the methodologies used in this research to construct the Sherpa system. These methodologies are: 1) distributed problem solving (DPS) methodology that is a sub-field of distributed artificial intelligence (DAI), and 2) the rule formulation methodologies.

In the section 3.1, a description of DAI and DPS is presented. Then, section 3.2 describes the rule formulation methodologies. This is followed by section 3.3 and 3.4 in which the use of the rule formulation and DPS methodologies are described, respectively.

3.1 Distributed Artificial Intelligence

DAI, a sub-field of artificial intelligence, is concerned with concurrence in artificial intelligence computation. A DAI system has significant advantages over a centralized problem solver (Durfee et al., 1989). These advantages include: 1) faster problem solving by exploiting parallelism, 2) reducing necessary communication by transmitting only high level partial solutions, rather than raw data to a central site, 3) increased flexibility by creating problem solvers with different abilities, dynamically combined, to solve the current problem, and 4) increased reliability by allowing other problem solvers to replace the failed ones.

Huhns (1987), Bond and Gasser, (1988), and Durfee et al., (1989) identified the following reasons for increased utilization of DAI systems:

1. The use of concurrent and distributed computing is wide spread. This is due to the availability of more powerful computers which are also becoming cheaper. Additionally, new technologies such as 'client-server' are facilitating the proliferation of such computing.

2. Many artificial intelligence applications are inherently distributed. The application may be spatially distributed, such as interpreting and integrating from distributed sensors or controlling a set of robots that work together on a factory floor. On the other hand, the application may be functionally distributed, such as developing a sophisticated architectural expert system composed of individual experts in specialtized areas such as structural engineering, electrical wiring and room layout.

3. A DAI system supports the principles of modular design and implementation. The ability to structure a complex problem into relatively self-contained processing modules leads to systems which are easier to build, debug and maintain.

4. There are problems that are too large for centralized systems. These problems can only be solved by the cooperation of several independent systems. For instance, multiple expert systems with different but possibly overlapping expertise can cooperate to solve the problems that are outside the scope of a single system.

5. Distributed artificial intelligence can help to improve the available techniques for representing and using the knowledge of a certain domain.

3.1.1 Distributed Problem Solving

The methodology within DAI, used to develop Sherpa is known as Distributed Problem Solving (DPS). This section provides a detailed description of DPS[2].

---

[2] The use of DPS methodology to construct Sherpa is given in section 3.3.

One of the defining characteristics of DPS is that it is a cooperative activity of a group of decentralized[3] and loosely-coupled[4] problem solvers. The problem solvers cooperate in the sense that none of them has sufficient information to solve the entire problem by itself; therefore, they have to work as a team and share knowledge about the tasks, results, goals, and constraints to solve the problem. (Davis and Smith, 1983).

In applying DPS, the problem is divided into tasks, and special performers (agents) are designed to solve each task. All interaction (including cooperation and coordination) strategies are incorporated as an integral part in the design of the system. To stress the aspect of cooperation in distributed problem solving, the term cooperative distributed problem solving is used. Durfee et al., (1989) defined cooperative DPS as "the study of how a loosely coupled network of problem solvers (agents) can work together to solve problems beyond their individual capabilities." Each problem-solving agent in the network is capable of sophisticated problem solving. Agents can work independently but the problem faced by them cannot be completed without cooperation. Cooperation is necessary because no single agent has the sufficient expertise, resources, or information to solve a problem. Most

---

[3] "Decentralized" means that both control and data are logically and often geographically distributed (i.e., there is no global control or global data storage).

[4] "Loosely Coupled" means that individual problem solvers spend most of their time in computation rather than communication.

problem-solving processes handled in a DPS system consist of four phases. They are as follows:

1. The decomposition of the problem into sub-problems;
2. The allocation of the sub-problem tasks among the agents;
3. The solving of the sub-problem tasks by the assigned agents; and
4. The integration of the solutions, obtained in phase 3, to obtain the global solution.

These four phases can be further defined as follows:

(I) Problem Decomposition

In a typical decomposition process, a single "supertask" is decomposed into smaller subtasks, each of which requires less knowledge or fewer resources. Most decomposition processes flow directly from the descriptions of available operators for problem-solving. This indicates that they depend on the designers' forethought in operator construction and description. The key element in problem decomposition is to minimize the costs of knowledge and resource distribution (Bond and Gasser, 1988).

The decomposition choices are primarily dependent on how a problem is described. In a conventional DPS system, many of these tasks are performed by the designer. Decomposition may be achieved along the following lines:

1. Select tasks that are inherently decomposable: In this approach, the given representation of the task contains its decomposition. The description of the states, of the state-space, and of the operators, leads to a decomposition.

2. Decomposition by programmer: This approach is useful in systems that have few known principles or methods for automatically decomposing tasks.

3. Hierarchical planning: A hierarchical planner performs genuine task decomposition. It generates tasks which become goals to work on. However, this provides only a partial solution because the decomposition depends on the prior descriptions of the available execution operators which are usually static.

(II) Task Allocation

A task can be allocated by sending to an agent the entire problem description, a solution method, and a control trigger, that the agent simply enacts. Alternatively, an agent can be provided with the data to which it can apply methods the agent already possesses, or it can receive methods to apply to the available data. Finally, an individual agent may receive only a control or responsibility "trigger", while gathering for itself its problems, its data and knowledge, and its solution methods. This range of choices indicates a need for some basis to decide what level of responsibility and choice should be assigned.

(III) Resource Allocation

Resources are the products consumed to accomplish problem-solving. Resource allocation is a sub-problem of task allocation in a DPS system. Allocating resources or tasks is a way of prioritizing the tasks, because tasks without resources cannot be solved. Among the approaches used for resource allocation are the following:

1. Resource allocation using specialist "sponsor" agents (De Jong, 1988) that allocate fixed portions of resources in a manner analogous to research sponsors.

2. Resource allocation based on criticality of the tasks (Lesser et al., 1988).

3. Resource allocation via resource pricing.

(IV) Solving of the Sub-Problems by the Assigned Agents and Integration of the Solution

The sub-problems are solved by the agents according to their specialty. These answers are complete solutions to the sub-problem. These solutions may have to be integrated into a solution to the overall problem.

3.2 Methodologies to Form Rules for a Rule-Based System

To implement the DPS methodology, different sub-problem solving agents are developed. The methods to develop these agents are collectively called 'rule formulation'. There are two methodologies employed in this research to develop rules for Sherpa[5]. These are 1) Inductive Learning and 2) Knowledge Engineering. A description of these methodologies follows...

---

[5] Description of these modules and other issues in designing and implementing the architecture of Sherpa is presented in section 3.4.

3.2.1 Inductive Learning

Inductive learning methodology can be used to assign objects to classes determined by their properties. As previously mentioned in Chapter 2, the key requirements for the tasks that lend themselves to the inductive approach are as follows[6]:

1. Attribute Value Description;
2. Predefined Classes;
3. Discrete Classes;
4. "Sufficient" Classes; and
5. "Logical" Classification Models.

To implement the inductive learning approach, a computer program, C4.5 (Quinlan, 1992) is used. This program generates a classifier in the form of a *decision tree*, with a structure that is either

- a *leaf*, indicating a class, or
- a *decision node*, that specifies some test to be carried out on a single attribute value, with one branch and subtree for each possible outcome of the test.

[6] For complete detail of these properties refer to Chapter 2.

A decision tree can be used to classify a case by starting at the root of the tree and moving through it until a leaf is encountered. At each non-leaf decision node, the case's outcome for the test at the node is determined and attention shifts to the root of the subtree corresponding to this outcome. When this process leads to a leaf, the class of the case is predicted to be that which is recorded at the leaf.

When classification tasks become more intricate, however, even simplified decision trees can grow to unwieldy proportions. As an example, the simplified decision tree for a particular four-piece chess endgame has 158 nodes and would require more than three pages to be reproduce in detail. This tree, although extremely accurate, is too complex to be understood by anyone! (Quinlan, 1992). Michie, (1986) writes of a similar chess classification model:

> "Recent results have shown that programs constructed by systems such as Quinlan's ID3 can be, in one sense, 'super-programs' and at the same time quite incomprehensible to people....The ID3-synthesized program clearly qualifies as a super-program."

Ways of overcoming this barrier have been explored. Shapiro, (1987) breaks down a single large tree into a hierarchy of smaller trees; his *structured induction* approach can lead to trees that, individually and collectively, are easier to understand.

C4.5 employs a different way to achieve the same goal by re-expressing a classification model as *production-rules*, a format that appears to be more comprehensible than trees. The program uses a simplified form of production rule L-->R in which the left had side, L, is a conjunction of attribute-based tests and the right hand side, R, is a class. One of the classes is also designated as a default. To classify a case using a production rule model, the ordered list of rules is examined to find the first case in which the left-hand side is satisfied. The predicted class is the one nominated by the right-hand side of this rule. If no rule's left-hand side is satisfied, the case is predicted as belonging to the default class.

An example of the C4.5 output is given in the Figure 1. The program generates four rules (Rules 2-5) in this example from a decision tree. Let us examine Rule 5. The rule number (in this case 5) is arbitrary, being derived from the order of leaves in the original tree, and serves to identify the rule. The left-hand side contains two tests, and a case that satisfies both of them is classified as <u>good</u>, the class identified by the right-hand side of the rule. The program also predicts that this classification will be correct for 93% of unseen cases which satisfy this rule's left-hand side. After listing three similar rules, the program selects <u>good</u> as the default class for this domain.

3.2.2 Knowledge Engineering

The other methodology used to formulate rules for Sherpa is known as "Knowledge Engineering". Mockler and Dologite, (1992) describe knowledge engineering as "a term that is used to describe the KBS developmental job of situation analysis and representation, and computer system design and implementation."

As outlined in Chapter 2, there are many commonly used techniques to accomplish knowledge engineering. A brief description of these follows:

1) **Process Tracing:** These techniques are used to follow a domain expert's train of thought or reasoning as he or she completes a task. These techniques are shown in Figure 2 and are discussed in a greater detail below.

Direct observation of familiar tasks: This technique helps the knowledge engineer observe what information is important to the domain expert in performing a job, as well as to identify the critical situation factors affecting the expert's job performance. Wherever possible, several tasks familiar to the expert can be observed. An analysis of these observations supplemented by an analysis of the available texts and technical manuals describing such tasks, can be used to generate a preliminary knowledge base (Grabowski, 1988, Hoffman, 1987).

Recollection of past experience: This technique involves asking experts how they arrived at a solution to a problem in the past and how the various aspects of the situation affected their decision. Using such a technique often produces detailed models of the expert's reasoning process (Grabowski, 1988).

Simulated familiar tasks: This technique uses information accumulated from past observations, interviews, and the written material available. It allows an expert to simulate the performance of a familiar task using a case study (Grabowski, 1988). The availability of actual data is a prerequisite for using this technique (Hoffman, 1987). Having been given an actual case, the expert would talk aloud about how to solve the problem simulated in the case study. By using such simulated scenarios the expert is forced to concentrate on the specific aspects of the case or problem that helps to produce a more detailed operational picture of how the expert goes about the task. Where a variety of familiar tasks are involved and time is limited, simulating familiar tasks is a very useful knowledge acquisition technique.

Constrained process tracing: This technique involves having an expert perform a familiar task or discuss a simulated or recollected task, but under a time constraint or information constraint, or both. Where the information-constraint technique is used, contextual information is withheld, forcing the expert to rely heavily upon his or her knowledge and reasoning skills (Hoffman, 1987).

Difficult cases: This technique involves observing experts dealing with a case that has unusual, unfamiliar or challenging features. It is used in the later stages of knowledge acquisition to refine the reasoning behind a decision and to expand the applicability of a system to a wider range of situations. This technique is often used to get evidence about the subtle, or more refined aspects of the expert's reasoning. (Hoffman, 1987)

2) **Protocol Analysis:** During observations of the expert, protocol analysis is used to analyze, organize and represent a domain expert's overall decision-

making approaches and the specific methods used to move from one knowledge state to another (Grabowski, 1988). In using this technique, a knowledge engineer is able to formulate a sequenced map of the expert's decision-making thoughts.

An example of protocol analysis is verbal protocol. It is the verbalization of an expert thinking aloud, which can be tape-recorded, transcribed or otherwise recorded.

There are two types of verbal protocols: concurrent and retrospective. A description follows:

1) Concurrent protocols: These are obtained by recording an expert's verbalization of thoughts and reasoning processes while solving a problem. One advantage of this type of protocol analysis is that it requires little equipment. In addition, the knowledge engineer may not have to be present when the expert encounters a tough case in a real-world situation. Using this technique, the expert is equipped with a small tape recorder and instructed on how to make a verbal protocol of personal observations (Hoffman, 1987).

2) Retrospective Protocols: These are protocols obtained by asking a domain expert to review audio, video or transcription records of the expert's previous verbalization after the task is completed The expert is then questioned about decision-making criteria he or she employed at certain points in the recorded session. Retrospective protocols are often used when concurrent protocols are suspected of interfering with the expert's task performance, or when the task performance is suspected of interfering with the expert's ability to offer a coherent protocol (Wright and Ayton, 1987).

3.3 Sherpa Architecture

Sherpa system was constructed based on the aforementioned methodologies. The outline of the architecture is shown in Figure 3. The various modules of Sherpa are the language system, the problem processing system, meta-level knowledge, knowledge sources, and databases. The Following are the descriptions of each module.

Language System: The language system represents the human-computer interaction module of Sherpa. The Language system is a graphical-user interface module that is used to input data and gain output information[7] from Sherpa.

An example of the language system module for the Sherpa prototype (telephone analysis) is shown in Figure 4. Minimal text entry is required. Instead he/she is required to point and click on his/her selections. This makes the interface user-friendly.

Problem Processing System: The problem processing system is the module by which the past classifying information on any data is compared to the classifying information being analyzed in the present problem. Any information that pre-exists in the system is combined with the new information and made available to the agent. For example, in the telephone record analysis, the telephone numbers of the suspect are matched with the pre-existing criminal records and made available to the investigator. This provides the user with more valuable information.

The common telephone record available to the investigator is shown in Figure 5. As one might expect, it is useful only if the agent knows the specific information he/she wants from the system. On the other hand, Sherpa presents information as shown in Figure 6. This telephone record analysis,

---

[7] Information is defined as data of direct use to the agent in his work.

available through the problem processing module of Sherpa, is more specific and provides the investigator with more useful information. In the example provided, the modified record provides the investigator with the criminal record information on the incoming and outgoing phone calls of the suspect.

Meta-Level Knowledge: This is the module that gives "knowledge about knowledge". It determines the priority of different knowledge sources in solving the problem. Sherpa lets the agent choose the desired analysis. Using the meta-level knowledge rules, correct information is automatically made available to the agent for decision-making.

Knowledge Sources: These are independent sources to extract information from the data available to Sherpa. These sources are independent in that they do not invoke one another and ordinarily have no knowledge of each other's expertise or behavior. They are also cooperative in that they contribute solution elements to a shared problem. Two knowledge sources were developed for this research. These are the telephone and financial analyses. These knowledge sources are described in complete detail in the section 3.4.

Databases: Various databases are used in this module. The data from these databases is provided to other modules in Sherpa. This data is integrated with the current data in the modules. In this manner, logically and geographically distributed data is available to the agents for decision-making.

3.4 Application of DPS Methodologies to Construct Sherpa

The four phases of DPS methodology described in section 3.3 were used to construct Sherpa. The description of these phases is provided in the following sections.

3.4.1 Problem Decomposition

The drug investigation (explained in Appendix A) consists of decomposing the overall problem into simpler sub-problems. Each sub-problem solution requires a different data source. As can be seen in the dependency diagram in Figure 7 these sources of data are: 1) Financial data; 2) Toll Record; 3) Surveillance data; and 4) Other sources (e.g., trash check and mail check). Each source contributes to the correct classification of the suspect. These sources can be used independently to classify a suspect. They may also serve to confirm each other's results.

As an example of analysis working independently, a telephone record analysis may be used to determine the relationship of the suspect with a known criminal. This is made possible by observing the calls made by the suspect. If a call is made to a person who has a criminal record, then suspicion about the caller increases. An example of sources working together would be a suspect with a high phone bill which elicits the need to determine sources of income (financial analysis).

3.4.2 Task Allocation in Sherpa

Any task in Sherpa is allocated to a particular knowledge source based on the nature of the data and the capability of the appropriate knowledge source to solve it. For example, if the source of information is telephone records, then the toll analysis knowledge source is allocated the task of analyzing these records. By applying the developed rules, the toll record analysis module analyzes the data.

3.4.3 Resource Allocation

The resource allocation in Sherpa is done based on the rules formulated as a result of knowledge engineering. For example, the financial records are more difficult to analyze due to the large number of records and variety of sources. Therefore, resources should be allocated to telephone records first. Telephone records also allow an agent to determine relationships between people. Therefore analysis of these records may suggest the need for further investigation and the nature of this investigation.

3.5 Application of Rule Formulation Methodologies to Construct Sherpa

There are two knowledge sources that were developed to create Sherpa and perform research for this thesis. These knowledge sources are described in the following sections:

3.5.1 Toll Analysis

Telephone Analysis is used to determine the following relationships: 1) between individuals, 2) between individuals and organizations, and 3) among organizations. Telephone analysis is undertaken only when a 'reasonable suspicion' about the individual or the organization exists and is corroborated by other intelligence indicators.

In the manual system, the telephone analysis involves drawing link charts manually. An example of a link chart is presented in Figure 8. The second step in the analysis is to interpret the chart and determine the relationships in the activity being analyzed. In the link chart the connections are made between two parties by drawing lines between them. If a group of people belong to an organization they are shown by a rectangle. This type of manual analysis becomes very cumbersome to draw and analyze when numerous subjects are involved.

The module for toll analysis was developed using inductive learning. Some attributes of the toll records were available in most cases. These attributes included: 1) Frequency: How many times the call was made to or from a particular number, 2) Time of day: What time during the day the call was made?, and 3) Day of the week: What day of the week the call was made? An example of the rules formulated as a result of this inductive generalization is given in Figure 9. These rules automate the toll records analysis which was previously done manually.

3.5.2 Financial Analysis

The second knowledge source developed for Sherpa was the financial analysis module. It was developed using the knowledge engineering techniques described in section 3.3. Financial gain is one of the primary motives in criminal activity. Therefore, skills in concealed income analysis are important to the development of inferences related to most criminal activities (ANACAPA Training Manual, 1992).

The purpose of the financial analysis is to aid the investigator in determining the extent to which a person, group or organization is receiving and benefiting from money obtained from non-legitimate sources. It can help in strengthening suspicions about a criminal's classification.

In the manual system, a person's 'net worth' is calculated. It is defined as the 'difference between assets and liabilities'. An example of this analysis is given in the Figure 10. As shown in the figure, liabilities are calculated by counting the subject's total liabilities from all sources. Some of the sources are illustrated in the Figure 11. These liabilities are subtracted from the assets and net worth is calculated. (The figure also shows the calculation of adjusted gross income.)

The system constructed after integrating these modules was tested at Wisconsin Division of Narcotics Enforcement. The observational study and its results are discussed in Chapter 4.

CHAPTER 4

PLANNING AND ANALYSIS OF OBSERVATIONAL STUDY

One of the aims of this thesis is to compare the currently used telephone and financial analysis with the analysis done by Sherpa. This comparison can be used to develop Sherpa into a completely functional problem solving system. This chapter describes the hypotheses, the experimental design and the statistical analysis that was done to compare the two systems. The organization is as follows: Section 4.1 states the purpose of this study and the hypotheses that were tested; section 4.2 describes the experimental design and related issues, and section 4.3 is a report on the results that were obtained from this study.

4.1 The Research Setting

The two systems which are compared for the purposes of this study are: 1) system in current use which consists of a manual system as well as geographically and logically segregated computer systems and 2) Sherpa, which is an integrated KBDSS system which gives drug investigators access to all the necessary data and information. Additionally, Sherpa provides agents with financial analysis, telephone analysis, and other information to make a decision on the classification of a suspect.

The move toward technology usage (such as the introduction of more sophisticated and powerful computers) at the DNE inspired top management to support the project. Top management was also interested in this project, because it would indicate to them the future direction that DNE should take technologically.

The current system in the DNE is very cumbersome. This is due to the combined effects of various factors. These include the fact that only two analysts serve thirty agents, they lack analysis tools (hardware and software), and have minimal telecommunication facilities between analysts (situated in Madison) and the DNE agents (situated in Madison, Appleton, Eau Claire, and Milwaukee). After the observational study is accomplished a fully functional system will be developed at DNE. This system will help the DNE use geographically and logically distributed data and information.

To conduct the study and compare the present system with Sherpa, hypotheses were formulated. These hypotheses are discussed in the next section.

4.2 Hypotheses

Four performance measures are examined in the current study: time, identification frequency, leads, and evidentiary information. Based upon these measures, hypotheses were formulated as follows:

Hypothesis 1: *An agent will be able to obtain telephone and financial records in less time with Sherpa than with the current system.*

When the telephone records are requested from the telephone company by an agent (with legal approval), they are sent to the analysts in Madison since they are the only authorized people able to receive these documents. They enter the data into a specialized database manually. After this entry, the records are analyzed. The analysis is either faxed or mailed to the agents. The agent may request a follow-up analysis which may be delayed due to the workload of the analysts.

The time between receiving the data and obtaining the analysis will be reduced significantly when using Sherpa. This is hypothesized to be true because data are scanned into the system by an agent and imported directly into a specialized database.

<u>Hypothesis 2</u>: *Sherpa will help the narcotics agent identify a greater number of leads for further investigation than the current system.*

A *Lead* is defined as 'the information suggesting whether to pursue further investigative activity and the nature of the activity.' In the present system, during the initial stages of the analysis only a particular item can be inspected. For example, the format of data and information available to the investigator may be a phone bill as shown in the Figure 5. In this figure, if the investigator can identify (because of a hunch) some name, phone number, or other information, he may decide to pursue an investigation.

Sherpa provides the same information in a more usable manner. An example of the phone bill information available is given in Figure 6. The data provided to the investigator in this case contains more useful information. For example, the previous criminal record of a called party may provide a lead to an agent. In addition, Sherpa provides the conclusions based on the rules. Therefore, it is hypothesized that the agent will be able to obtain more leads from information provided by Sherpa.

<u>Hypothesis 3</u>: *The information provided by Sherpa will help identify more middle and upper level drug dealers.*

Middle and upper level drug dealers can be defined in lay person terms as the drug dealers who are not 'street pushers'. Their identification is one of the main purposes of Sherpa. It is hypothesized that since the information available to the agents will provide them with identifiable patterns in the data, it will be possible to recognize the information that distinguishes a middle and upper level drug dealer with higher frequency. For example, Sherpa provides an agent information confirming a repeat offender by providing him with information from distributed databases.

<u>Hypothesis 4</u>: *The amount of evidentiary information provided by Sherpa will be higher than the current system.*

Evidentiary information is information that an agent can use to prove to others his/her convictions about the suspect's classification. In the current system, the evidentiary information available to the agent is restricted, because the only way he/she can get the right evidentiary information is by knowing the nature of the information from a previous experience. Sherpa provides information in a more meaningful form. It is integrated and shows readily identifiable patterns as shown in Figure 6.

4.2.1 Variables

To perform statistical analysis and test above mentioned hypotheses, four variables are measured. These variables are:

Time: It is the duration it takes to obtain telephone and financial analysis after the receipt of the data by the DNE.

Leads: A "*Lead*" is defined as the information suggesting whether to pursue further investigative activity and the nature of that activity.

Identification Frequency: This is the number of drug dealers belonging to middle and upper level.

Evidence: This variable measures the amount of evidentiary information to determine the classification of a suspect. This is the information that an agent can use to prove to other personnel of his/her convictions about the suspect's classification.

4.3 Administration of the Observational Study

Two DNE offices were chosen for the purposes of this study. They are located in Madison and Appleton. They were selected on the basis of comparable number of agents and their experience as shown in Figure 11. Figure 12 lists the counties which are served by these regional offices. The demographics of these counties are comparable as is indicated by the total adult population and the total number of drug violations.

The Sherpa system was developed using the methodologies described in the Chapter 3. The Madison special agent-in-charge (SAC) was involved in the initial development of Sherpa. Therefore, to avoid any bias due to his personal interest in the success of Sherpa, Madison was selected as the control office. The control was set to measure any extraneous conditions, such as changes in the organizational policy, that could influence the observed results. The Appleton office was selected as the treatment site. This study was conducted for three months (January 7 - April 7, 1995). The chronology of this study can be found in the Appendix D.

After the pre-test at the Appleton and Madison offices, training was provided to all the 8 agents at the Appleton office including the SAC. The training was performed in a group for 1.5 hours. They were told that the purpose of this research was to test the prototype of a system which had the potential of being developed into a fully functional system.

To compare the current system with Sherpa, closed cases from the last three years were screened for their telephone and financial analysis content. This was done with the help of the SAC in Appleton. Three cases from the SAC's closed cases and three cases from a senior agent's closed cases were selected for analysis using the current system and Sherpa. First, these cases were solved using the current system and the variables mentioned in section 4.1 were measured. These cases consisted of multiple financial and telephone analysis. The value of the variables were measured for each analysis. Sherpa was then used by the agents to analyze the same cases (agents did not solve the same cases using both methods). Each agent was asked to make an assessment using Sherpa and determine the value of the variables. The agents whose cases were being used were not involved in this part of the study.

The data collected in the observational study was statistically analyzed. The results are presented in the section 4.4.

4.4 Study Results

Cochran, (1983) mentions that in the estimation of the difference between the values of variables after the treatment and before treatment, two major statistical aids are *tests of significance* and *confidence interval.* The *test of significance* relates to the question: "Is there convincing evidence that exposure to the possible causal force has any effect at all?" It specifies whether the difference in the mean values of the same variable, obtained by two methods is non zero. The Confidence interval relates to the question, "How large is the difference between the means." (Cochran, 1983)

There was an additional concern that measurements three months after the introduction of a computer system would reflect the effects of other factors influencing the organization, beyond the effects of technology. These factors could consist of changes in the organizational structure, changes in the procedures, changes in accessibility. To measure the effects of the introduction of Sherpa rather than these other factors, a pre-test and a post-test were conducted at the control group office in Madison.

In a systematic analysis, the change in various factors over time was observed. These are reported in Figure 13. No changes that would affect the measurement of the effectiveness of Sherpa at Appleton were found. What happened in the organization long after the introduction of the technology is beyond the scope of this study. The analysis suggests that the three month data collection period used was adequate.

4.4.1 Results

To perform the statistical analysis, t-tests for paired samples were performed. Pairs in this case were the mean values of the same variable using the current system and using Sherpa. Standard deviations were calculated. Due to the large difference in the standard deviation of two samples, log transformation was performed. This reduced the difference between the standard deviations. T-statistics for paired samples thus obtained are shown in Figure 14.

The results in Figure 14 support each of the hypotheses.

Hypothesis 1: The 95% confidence interval of the difference in the mean value is (-2.142, -1.239) and the t-value is 7.98. This supports the hypothesis that an agent will be able to obtain telephone and financial records in less time using Sherpa than using the current system.

Hypothesis 2: The 95% confidence interval of the difference in the mean value of the leads obtained by the two methods is (-1.974, -1.456) and the t-value is 14.13. This supports the hypothesis that Sherpa will help the narcotics agent identify more leads for further investigation compared to the current system.

Hypothesis 3: The 95% confidence interval of the difference in the mean value of the identification frequency is (-0.723, -0.417) and the t-value is 7.94. This supports the hypothesis that the information provided by Sherpa will help identify more middle and upper level drug dealers.

Hypothesis 4: The 95% confidence interval of the difference in the mean value of the evidentiary information is (-1.883, -1.209) and the t-value is 9.77. This supports the hypothesis that the amount of evidentiary information provided by Sherpa will help identify a greater number of middle and upper level drug dealers.

4.5 Discussion

The introduction of Sherpa improved the classification of suspects when measured in terms of all the variables observed. In addition, the method used allowed a conclusion to from that the changes observed were the results of Sherpa, and not the result of methodological artifacts, preexisting differences between the groups, or other extraneous factors. No policy changes in the DNE were observed during the time of the study which could have affected the results.

Overall, Sherpa performed significantly better than the current system. This indicates that, based upon the observational study, a fully functioning system should be developed to help drug agents at the DNE. Future research should also be undertaken to further expand the role of Sherpa. This and other research issue are discussed in Chapter 5.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATION

5.1 Conclusions

The two main objectives of this thesis, as stated in Chapter 1, were: 1) to introduce a methodology that uses the distributed artificial intelligence capabilities to develop a KBDSS and 2) to compare the analysis provided by the KBDSS with the analysis provided by the current system.

These research objectives were achieved through the development of a KBDSS to perform the financial and telephone record analysis. The strategy of the KBDSS was to distribute problem solving among three modules of the Sherpa system to solve the problem. The development of these modules was explained in Chapter 3. An observational study was conducted to compare Sherpa with the current analysis system. The results were discussed in Chapter 4. The research conclusions can be summarized as follows:

- Distributed artificial intelligence has many potential benefits where information, resources or expertise are naturally distributed or where they can intentionally distributed to improve the speed, modularity, and/or reliability of the system.

- The distributed artificial intelligence architecture suits the area of drug intelligence more than the single agent system. This is due to the fact that drug investigation requires several geographically and logically distributed databases and DAI is a suitable solution.

- Computerization is an important means of improving drug investigation. Drug intelligence (the most important part of a drug investigation) can be strengthened by using the computerized KBDSS developed in this dissertation.

- The analysis step in drug intelligence is the most crucial step. The goal of performing more efficient analysis can be accomplished using a system like Sherpa!

- The distributed problem solving concept utilized in this thesis helps integrates the different sources of information.

5.2 Recommendations for Future Research

This research may be applied to different domains in law enforcement and, indeed, in many other domains. As a result of this research, the author recommends the following items to be addressed in the future.

1. Enhance the capabilities of the developed KBDSS by incorporating different models for analyzing each investigation task. This addition to the system will enable the drug investigator to analyze the outcome of each module in a "what-if" scenario.

2. Introduce learning capabilities into the system. This can be achieved by attaching modules with machine learning capabilities as shown in the Figure 3.

3. Measure the effects of the system over an extended time period. For example, a study could be performed to measure learning by new agents at the DNE using Sherpa.

4. Measure the effects of such a system in giving feedback on performance to the agents.

5. Conduct a study to determine the changes in the organizational structure due to the introduction of Sherpa.

6. Introduce improved telecommunication capabilities with a multimedia input and output from Sherpa over geographically distant computer terminals.

7. Increase the use of image technology to enable the agents to access the mug shots of criminals.

REFERENCES

ANACAPA Training Manual (1992). Manual of ANACAPA Law Enforcement Training. San Diego, CA: ANACAPA Press.

Bond, A. H, & Gasser, L. (1988). Readings in Distributed Artificial Intelligence. Palo Alto, CA: Morgan Kauffman.

Boose, J. H. (1989). A Survey of Knowledge Acquisition Techniques and Tools. Knowledge Acquisition, March 1989, 3-36.

Boose, John H., & Gaines B. R. (1988). Knowledge Acquisition for Knowledge-Based Systems. San Diego, CA: Academic Press.

Brieman L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and Regression Trees. Belmont, CA: Wadsworth.

Davis, R. (1980). Report on the workshop on Distributed AI. SIGART Newsletter, 73, 42-52.

Davis, R., & Smith, R. G. (1983). Negotiation as a metaphor for Distributed Problem Solving. Artificial Intelligence, 20 , 63-109.

Druffee, E. H., Lesser, V. R., & Corkhill, D. D. (1989). Cooperative Distributed Problem Solving. In Barr, Cohen, and Fiegenbaum (Eds.). The Handbook of Artificial Intelligence, IV, 85-147.

Fiegenbaum, E. A., Mccorduck, K. B. & Nii, H. P.(1988). The Rise and Fall of the Expert Company. New York: Times Books.

Friedman, J. H. (1988). Multivariate Regression Splines. Technical Report 102, Laboratory for Computational Statistics. Stanford, CA: Stanford University.

Gallant, S. (1988). Connectionist Expert Systems. Communications of the ACM, 31:2, 152-169.

Grabowski, M.(1988). Knowledge Acquisition Methodologies: Survey and Empirical Assessment. Proceedings of the Ninth International Conference on Information Systems. SIM, IMS, and ACM

Hecht-Nielsen, R. (1988). Application of Counterpropagation Networks. Neural Networks, 1:2, 131-139.

Hoffman, R. R. (1987). The problem of extracting knowledge of expert from the perspective of experimental psychology. AI Magazine, Summer 1987, 53-67.

Huhns, M. N. (1987). Distributed Artificial Intelligence. San Mateo, CA: Morgan Kaufmann.

Laasri, H. & Maitre, B. (1989). Flexibility and Efficiency in Blackboard Systems: Studies and Achievements in ATOME. In Jagannathan, Dodhiwala V. R., & Baum L. S. (Eds.), Blackboard Architecture and Applications (pp. 309-322). San Diego, CA: Academic Press.

Lesser, V. R., Pavlin, J. & Durfee, E. H. (1988). Approximate Processing in Real Time Problem Solving. AI Magazine, 9:1, 49-61.

McClelland, J. & Rumelhart, D. (1988). Explorations in parallel Distributed Processing. Cambridge, MA: MIT Press.

Michie, D. (1986). On Machine Intelligence. Chichester, U.K.: Ellis Horwood.

Nisbett, R. & Wilson, T. (1977). Telling more than we can know: Verbal Reports on Mental Process. Psychological Review, 84:3, 231-259.

Mockler, R. J. & Dologite, D. G. (1992). An Introduction to Expert Systems. New York, NY: McMillan.

Nilsson, N. J. (1955) Learning Machines. New York, NY: McMillan.

Nii, H. P., Aiello, N., & Rice, J. (1990). Experiment on Cage and Poligon: Measuring the performance of parallel blackboard system. Distributed Artificial Intelligence. San Mateo, CA: Morgan Kauffman.

Pattison, H. E., Corkhill, D. D., & Lesser, V. R., (1987). Instantiating Descriptions of Organization Structures. In Michael N. Huhns (Eds.), Distributed Artificial Intelligence (pp. 59-96). San Mateo, CA: Morgan Kauffman.

Piramuthu, S., Park, S. C., Raman, N., & Shaw, M. J. (1993). Integration of Simulation Modeling and Inductive Learning in Adaptive Decision Support Systems Decision Support Systems, 9, 127-142.

Quinlan, J. R. (1992). C4.5 programs for Machine Learning. San Mateo, CA: Morgan Kauffman.

Rosenchein, J. S. & Breese, J. S. (1986). Communication-Free Interaction among Rational Agents. In Michael N. Huhns (Eds.), Distributed Artificial Intelligence. San Mateo, CA: Morgan Kauffman.

Sathi, A. & Fox, M. S. (1989). Constraint-Directed Negotiation of Resource Reallocations. Distributed Artificial Intelligence. San Mateo, CA: Morgan Kauffman.

Shaw, M. J. & Whinston, A. (1989). Learning and Adaptation in Distributed Artificial Intelligence Systems. In Gasser, L. & Huhns, M. N. (Eds.) Distributed Artificial Intelligence (pp. 413 - 429). San Diego, CA: Morgan Kauffman.

Smith, R. G. & Davis, R. (1981). Framework for cooperation in Distributed Problem Solving. IEEE Transactions on Systems, Man, Cybernetics, SMC-11:1, 61-70.

Star, S. L. (1989). The structure of ill-structured solutions: Boundary objects and Heterogeneous Distributed Problem Solving In Gasser, L. & Huhns, M. N. (Eds.) Distributed Artificial Intelligence (pp. 413 - 429). San Diego, CA: Morgan Kauffman.

Taha, M. A., Park, S. C. and Russell, J. S. (1994). An Intelligent Decision Support System for Construction Contractors Prescreening, Under review by European Journal of Operation Research.

Werner (1989). Cooperating Agents: A Unified Theory of Communication and Social Structure. Distributed Artificial Intelligence. San Mateo, CA: Morgan Kauffman.

APPENDIX A

CRIMINAL AND DRUG INVESTIGATION PROCESS

"Knowledge of past, present, and future criminal activity which results from the planned collection of information which, when evaluated, combined with other relevant information and analyzed, provides the user with a rational decision making." (Hamilton, 1994) can be defined as the criminal intelligence. This intelligence collection is the primary motive of any drug or criminal investigation.

The Investigation Process

A drug investigation is the product of a process consisting of a series of highly interrelated components. A failure or weakness in any one of these components will seriously impair the entire process and reduce the quality of product. These components, as shown in Figure A.1, are as follows:

- Target Selection - The systematic selection of targets to insure that intelligence efforts are directed toward targets which are in acceptable balance of utility (worth), probability of a successful result, and resources expended.

- Collection - Collection consists of both data collection planning and actual collection of information regarding the target(s) of intelligence operations. The collection must first be carefully planned before any information gathering occurs. Information is collected from both overt (open) and covert (closed) sources.

- Data Evaluation - All collected information cannot be validated as factual. Intelligence may be facts, opinions, rumors, and/or inferences. Many times one piece of intelligence may contradict another piece of intelligence. It is imperative that each intelligence report include both an evaluation of the source's reliability and the reported information's validity.

- Storage and Retrieval System - Collected intelligence must be promptly recorded on the appropriate intelligence form and placed in a storage and retrieval system. The system must include:

    - rapid user access;
    - selectivity of retrieval;
    - documentation of each dissemination;
    - periodic system audit and, when appropriate, information purging;
    - physical security to protect the files; and
    - system security to protect the information flow.

● Data Collection - Webster's dictionary defines collation as "assembling in proper order to clarify or give meaning to information." Taking a deck of playing cards and arranging them into four piles, one of each suit, in numerical order is one example of collation as is arranging a stack of surveillance reports in chronological order so that the oldest report is on top and the newest is on the bottom before reading them.

Many techniques used by the intelligence analysts such as link charts, various flow charts, visual investigative analysis, etc. are in fact data collation techniques used in what is commonly referred to as the data description and integration phase of the intelligence process.

● Analysis - Analysis is the heart of an intelligence system. Without it no interpretation or additional meaning can be given to raw, fragmented information which has been collected. Analysis is a mini - process within the overall intelligence process. The analytical process involves:

- data integration and clarification;
- inference development;
- inference testing; and
- finalizing inferences which are relevant and meaningful to the user (an intelligence product).

The link between each of the steps described above is inductive or inferential logic, which allows meaningful interpretation of the data to be made based upon the analyst's own knowledge, experience and expertise. Logic allows bits and pieces of information to be assembled in a meaningful manner.

• Dissemination - Dissemination is where the entire intelligence process most frequently breaks down. A saying in the law enforcement is that "intelligence is similar to gold nuggets because everyone wants as much as they can get, but no one wants to share what they have.

Analysis is the most important part of the intelligence process. Sherpa is developed to be support analysis phase in a drug investigation. Without analysis, fragmented information remains just that. No additional meaning or interpretation can be derived from it. Analysis is a frequently neglected but absolutely essential component of the intelligence system.

All intelligence must be promptly recorded on the appropriate agency form and placed in an intelligence file (in the traditional system) where it can be accessed by personnel both within the agency and from other agencies who "need" the information.

Data Evaluation - Not all the information in an intelligence system can be validated as factual. Intelligence may take the form of fact, opinion, rumor, or inference. The reliability of information sources varies widely. Many times individual pieces of information contained in an intelligence system may contradict one another.

Reliability is determined by the source's past performance and is an index of the Consistency with which or how the information is reported. Validity is an index of accuracy or truth of the information. A "model" data evaluation scale includes the following data evaluation rating on each intelligence report:

A. Reliable

B. Usually Reliable

C. Fairly Reliable

D. Not usually reliable

E. Unreliable

F. Reliability cannot be judged

Reasonable Suspicion

All information contained in an intelligence file must conform to the agency's data collection criteria and there must be a "reasonable suspicion" that the individual, group, business, organization, is involved in the criminal activity. Reasonable suspicion is defined as simply more likely to be true than not.

It is that degree of proof which would indicate to a reasonable person that there is a suspicion that the individual, group, business, or organization is involved in criminal activity. Raw, fragmented information may not fully conform to the reasonable suspicion criterion when it is initially collected. This type of intelligence must be processed to a point where it can be disseminated. In other words, when it has been placed in the master intelligence "file."

Working files are not generally disseminated and access is limited on strict "need-to-know" basis. When additional information has been collected and the reasonable suspicion criterion has been satisfied, the information may be processes into the master intelligence file and will be available for dissemination within the agency intelligence dissemination policy.

Dissemination Policy - Intelligence is disseminated based on two factors; those who have "right to know" and "need-to-know". Right-to-know is determined by an individual or agency's official capacity or authority and relationship to the information being requested. The agency dissemination policy should include a "right-to-know" criteria based upon both position within the agency and assignment. "Need-to-know" is based on the requesting agency or agency's representatives purpose in requesting the information. A direct official involvement in an investigation and a reason for requesting the information must exist.

When the intelligence is disseminated, the dissemination must be recorded on a dissemination control log recording the following information:

- date of the request;
- name of the individual requesting the information;
- agency Name (If disseminated to an outside law enforcement agency);
- reason for request (need - to - know);
- information requested; and
- Intelligence unit member handling the request.

Auditing and Purging - Information stored in criminal intelligence files must be periodically audited and, where appropriate, purged to insure the file is current, accurate and relevant to both the collection criteria and the needs and objectives of the agency. A primary purpose is to safeguard the right of privacy of the individuals contained in the file as guaranteed under both federal and state law. Law enforcement agency must keep stored intelligence data on individuals current and accurate. To ensure that the review and purging of intelligence files is done systematically, agencies must develop purging criteria based on time schedules.

Operational procedures to purge, as well as the manner of destruction for purge materials, must be established. When an intelligence file is initially processed it must be done so in a manner that insures that at the appropriate time, which could be six months to five years depending on the state laws, the type of intelligence data being processed, e.g., the file will automatically be audited to determine if it should be retained or purged.

General considerations for auditing intelligence information to determine if it should be retained or purged include:

- the age of the intelligence data - when was the data first placed in the file?

- frequency of use - how often the data has been used since it was first entered?

- the reason for initially collecting the data - does the purpose for which the data was collected still exist?

- nature of information - how important, significant or sensitive is the information?

- reliability and validity - is the information a confirmed fact, hypothesis, opinion or unsubstantiated rumor?

- completeness and accuracy - when was the information last updated, what effort is required to validate or update the information?

- usefulness - what is the information's present and potential utility?

- impact on the intelligence unit's operation - what effect will the removal of this information from the intelligence file have on the intelligence unit's peration?

The objective of an intelligence purging policy is to totally eliminate ntelligence data from the system at that point in time when one or more of the ollowing conditions apply:

- it is no longer useful because the purpose for which it as collected has been satisfied or no longer exists.
- it is obsolete because its age makes it unreliable for presentt purposes and it is not worth the effort it would take to update it.
- there is no longer any legitimate reason for maintaining the information in an active intelligence file.
- the information in the file has been found to be invalid, untrue or misleading.
- all information on the subject of the file has been ordered "sealed" by competent judicial authority.

Appendix B

General Crime and Arrest Facts at a glance (1993)
(Data from Wisconsin Department of Justice)

- The 1993 **crime rate** for Wisconsin was 4,091 serious crimes for each 100,000 residents.

- A total of 404,920 **arrests** were made in 1993. While total **adult arrests** increased by 3 %, **juvenile arrests** increased by 8.4 %.

- Over $154,000,000 in **property** was stolen during 1993, a 3.7 % decrease from the nearly $164,000,000 in property stolen in 1992.

- Nearly sixty % of all **murder** offenders were between 15 and 24 years of age.

- The number of **sexual assault** increased by 4.7 %, from 7,151 in 1992 to 7,847 in 1993.

- The proportion of **robberies involving a firearm** have increased from 41 % of the total in 1989 to 53 % of the total robberies in 1993.

- The number of reported **burglaries** has decreased by 20.4 % from 1983 to 1993.

- There were major increases in certain **arrests** including weapon law violations (10.5 %), fraud (6 %), drug law violations (11 %), runaways (13 %),and curfew law violations (10.6 %).

- Persons between 18 and 20 years of age accounted for nearly 53 % of the 33,015 arrests for liquor law violations.

- Four Wisconsin **law enforcement officers** were killed while in the performance their duty.

- There were 762 reported **assaults on law enforcement officers**; an 13.6 % decrease from 882 in 1992.

Appendix C

ORGANIZATIONAL BACKGROUND

This appendix explains how the Wisconsin Department of Justice is organized in order to highlight some of the Department's key responsibilities that are closely related to this thesis.

1. Wisconsin Department of Justice

The Department of Justice is the leading law enforcement agency in the state. It supports and initiates various law enforcement activities in the state. State criminal investigators are in the Division of Narcotics Enforcement and the Division of Criminal Investigation. The latter has bureaus specializing in organized crime and general investigations, arson and white collar crime. The Division of Legal Services, Criminal Litigation Unit, prosecutes important criminal cases and provides assistance to local prosecutors. The Division of Law Enforcement Services contains the state crime information bureau and the state crime laboratories. Figure C.1 is an organizational chart of the department.

2. Division of Narcotics Enforcement

The Division of Narcotics Enforcement (DNE) leads and coordinates Wisconsin's state and local drug enforcement efforts. The DNE emphasizes investigations of individuals, groups and organizations involved in high-level drug trafficking. Another major goal of the division is to support local drug enforcement efforts by providing advanced investigative and technical services.

The DNE has three bureaus, dealing with investigative operations, internal operations and special operations, as shown in the organizational chart below. The investigative operations are divided among the four sub-state regions, with headquarters in Madison, Milwaukee, Appleton, and Eau Claire.

The Special Operations Bureau houses the intelligence Unit, as well as the unit that provides technical investigative services and the marijuana eradication program. The technical intelligence section houses the two analysts who provide services to the full division. The direct access to the intelligence data is strictly limited to the intelligence staff.

Appendix D - Chronology of the Observational Study

ACTIVITY	DATE
A. Pre test conducted at Appleton and Madison	7 January (Madison) and 9 January (Appleton) 1995
B. Installed functioning system	Tuesday January 31, 1995
C. Trained the agents	Tuesday February 7, 1995
D. Evaluated the variables in six cases. Three of these cases were closed cases of special agent in-charge and other three were of a senior agent that were closed also. The variables were measured with the help of agents. One agent per case was used.	Tuesday February 7 - 20th February 1995
E. Announcement was made that all the agents could use the Sherpa system for solving the cases they were working on but only after they had worked on the old cases using Sherpa in point D.	February 7, 1995
F. Agents analyzed these cases using	February 7 - March 31 1995

Sherpa, after they had evaluated the variables in the cases solved manually. The agents who solved a particular case manually was assigned a different case to solve using Sherpa.	
G. Post Test	April 13(Appleton) April 17 (Madison)

* In a case, there may be many instances where toll records and financial analysis were performed.

APPENDIX E

Total Arrests Made by Wisconsin Law Enforcement Agencies[8]

(1992 and 1993)

Offense	1992	1993	% Change
Murder	449	521	+ 16.0
Forcible Rape	699	686	- 1.9
Robbery	2,440	2,300	- 5.7
Aggravated Assault	4,532	4,469	- 1.4
Total Arrests for violent offenses	**8,120**	**7,976**	**- 1.8**
Burglary	7,108	6,956	- 2.1
Theft	46,634	47,793	+ 2.5
Motor Vehicle Theft	4,922	4,963	+ 0.8
Arson	524	550	+ 5.0
Total Arrests for Index Offences	**59,188**	**60,262**	**+ 1.8**
Arrests for all other offenses	319,866	336,682	+ 5.2
Total Arrests	**387,174**	**404,920**	**+ 4.6**

[8] Data from Wisconsin Department of Justice

APPENDIX F

Number of Drug Arrests by County under the Jurisdiction of Madison and Appleton DNE offices, 1993

County	Total Drug Violations
Adams County	24
Brown Co	345
Calumet	21
Columbia	79
Crawford	20
Dane	1049
Dodge	116
Door	41
Florence	34
Fond Du Lac	131
Forest	16
Grant	51
Green	26

Green Lake	25
Iowa	17
Jefferson	204
Juneau	34
Kewaunee	36
La Crosse	132
Lafayette	13
Langlade	40
Lincoln	40
Manitowoc	428
Marinette	68
Marquette	13
Menominee	28
Monroe	137
Oconto	30
Oneida	99
Outagamie	185
Richland	30
Rock	347
Sauk	97

Shawano	48
Sheboygan	126
Vernon	48
Vilas	41
Walworth	320
Waupaca	60
Waushara	15
Winnebago	149

```
Read 40 cases (16 attributes) from example-dss

Processing tree 0
Final rules from tree 0:
Rule 5: wage increase first year > 2.5
        statutory holidays > 10
        - > class good[93.0%]
Rule 4:wage increase first year > 4
        - > class good[90.6%]
Rule 3: wage increase first year <= 4
        statutory holidays <= 10
        - > class bad[87.1%]
Rule 2: wage increase first year <= 2.5
        working hours > 36
        - > class bad[85.7%]
Default class: good
```

Figure 1 C4.5: Example of an Output

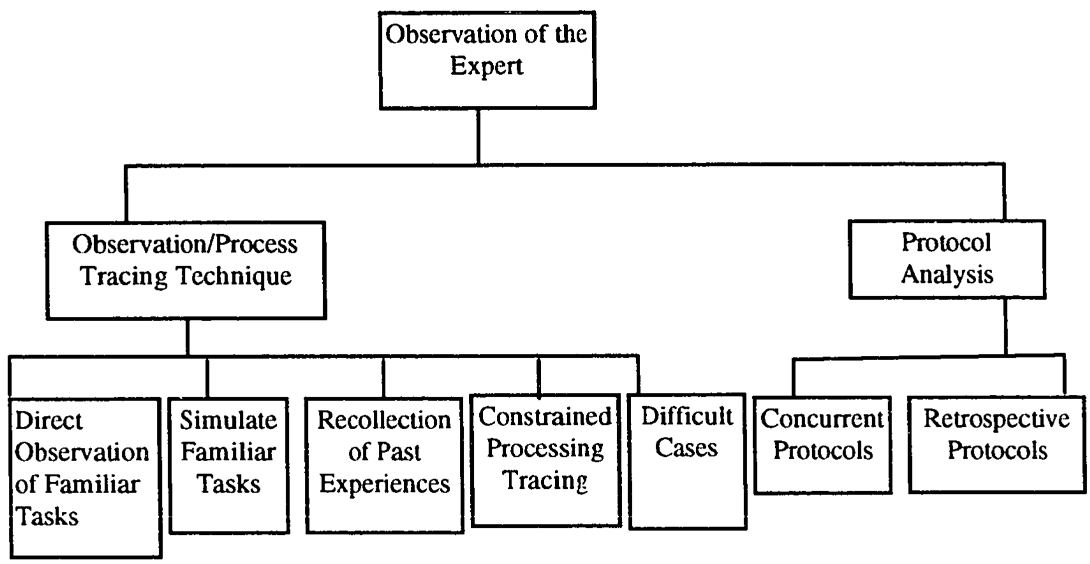

Figure 2 Knowledge Engineering

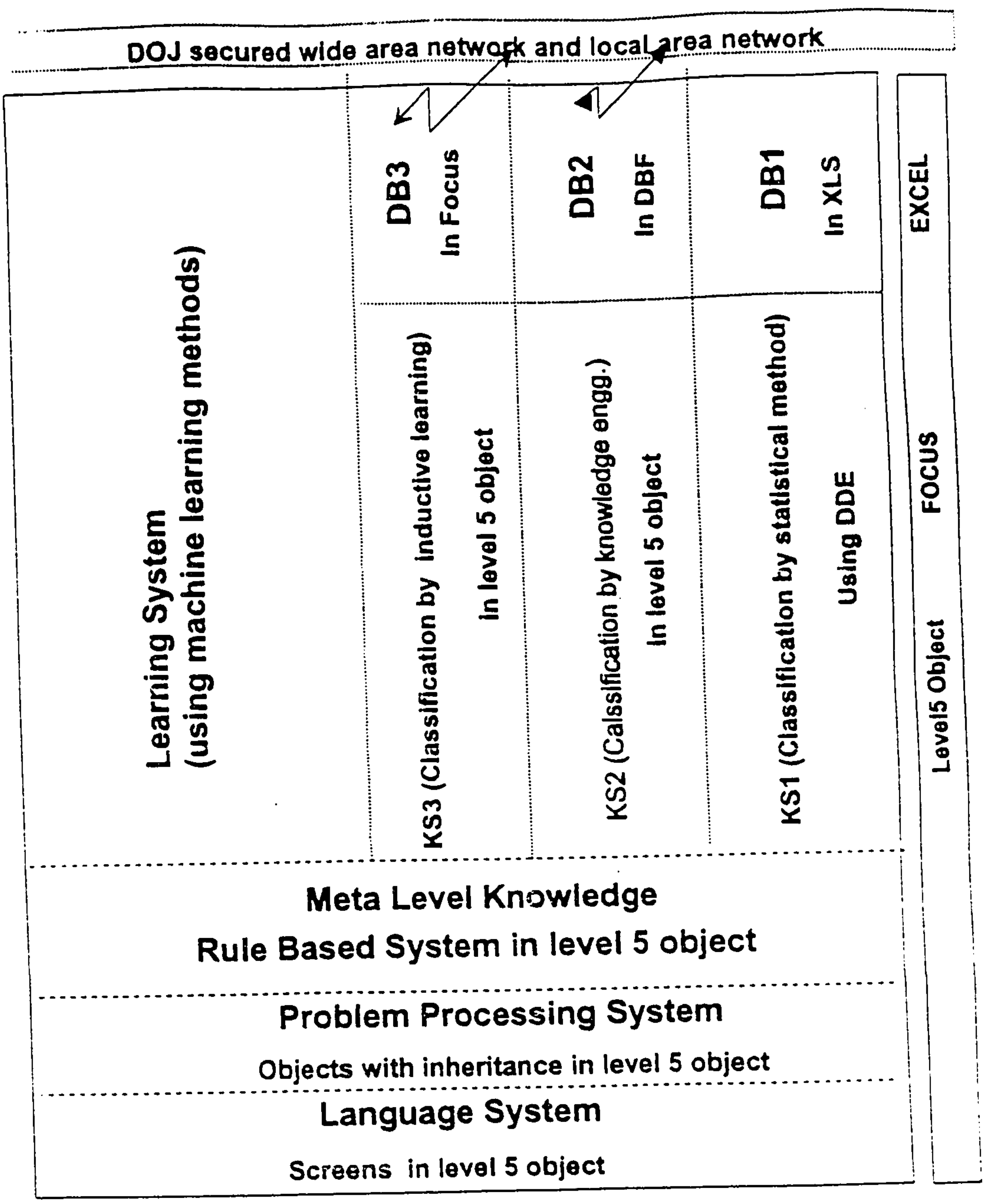

Figure 3 Sherpa Architecture

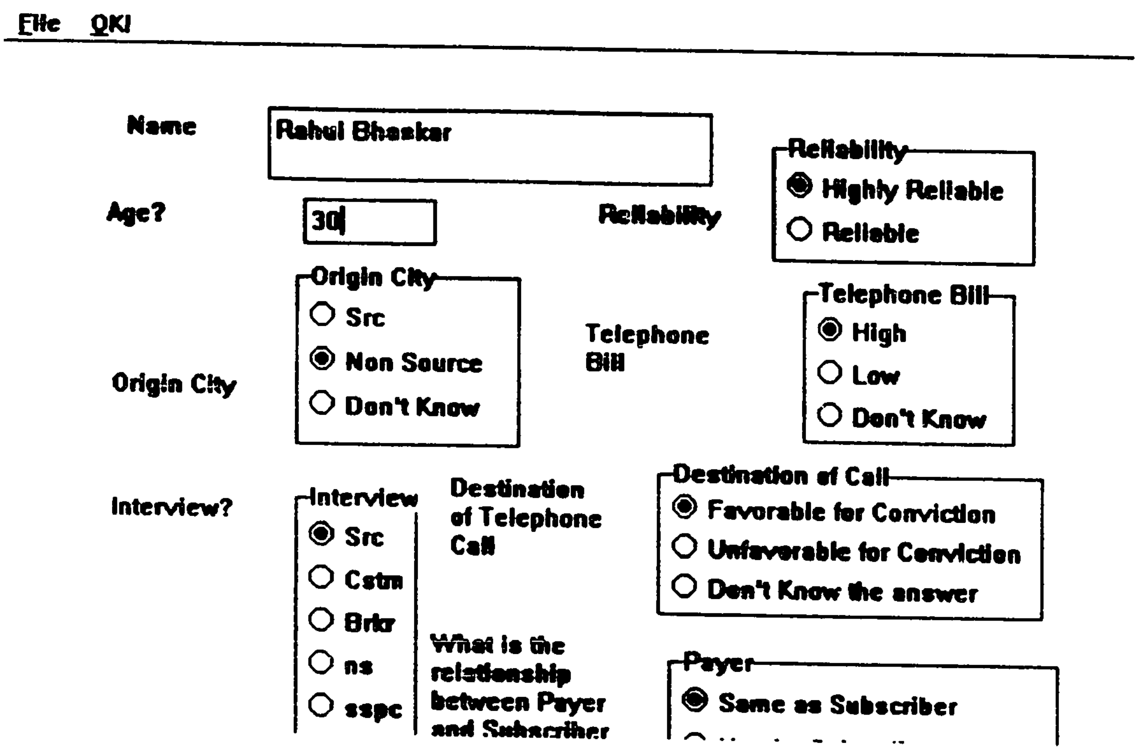

Figure 4 Language Module for Telephone Analysis

STATE RES MAY 1, 1990 04-01 20-00 99-00 0 -001 414

WISCONSIN BELL DETAILED CHARGES MAY 1, 1990 Page 4

To order new service or move service, call 1 800 924-5678.
To discuss bill payment, call 1 800 924-1500.
For all other reasons, call 1 800 924-1000.
Office hours are 7:30am - 6:30pm Monday thru Friday.
FOR FASTER SERVICE CALL TUESDAY THRU FRIDAY.

Local Usage Services
You placed 452 call(s) with your Flat Rate Call Plan.

Information Charges
You placed 36 call(s) to 1-411
35 are billed at $.25 each 8.75

Long Distance

NO	DATE	TIME	PLACE CALLED			NUMBER	CODE	MIN		RSC	TCT	L
1	4-06	937P	OSHKOSH	WI	414			1	.14	18	8	1
2	4-07	857A	OSHKOSH	WI	414			9	.60	18	8	1
3	4-07	1152A	OSHKOSH	WI	414			1	[illegible]	18	8	1
4	4-07	1159A	OSHKOSH	WI	414			1	.08	18	8	1
5	4-07	1159A	OSHKOSH	WI	414			8	.53	18	8	1
6	4-09	106P	GREEN BAY	WI	414			2	.45	18	8	1
7	4-09	131P	GREEN BAY	WI	414			3	.65	18	8	1
8	4-10	745P	OSHKOSH	WI	414			1	.14	18	8	1
9	4-11	1117A	GREEN BAY	WI	414			1	.25	18	8	1
10	4-12	503P	OSHKOSH	WI	414			6	.66	18	8	1
11	4-18	1208P	BONDUEL	WI	715			1	.26	18	8	1
12	4-18	225P	BONDUEL	WI	715			1	.26	18	8	1
13	4-18	253P	BONDUEL	WI	715			1	.26	18	8	1
14*	4-18	530P	LOCAL CALL	TO	414			2	1.65	06	8	1
			FROM APPLETON	WI	414							
15*	4-18	541P	LOCAL CALL	TO	414			1	1.65	06	8	1
			FROM APPLETON	WI	414							
16	4-18	1009P	OSHKOSH	WI	414			1	.14	18	8	1
17	4-19	540P	ARGONNE	WI	715			1	.18	18	8	1
18	4-19	542P	ARGONNE	WI	715			2	.36	18	8	1
19	4-19	550P	OSHKOSH	WI	414			3	.35	18	8	1
20	4-19	733P	BONDUEL	WI	715			1	.16	18	8	1
21	4-19	807P	OSHKOSH	WI	414			1	.14	18	8	1
22*	4-20	1224A	LOCAL CALL	TO	414			2	1.65	06	8	1
			FROM APPLETON	WI	414							
23	4-21	203P	OSHKOSH	WI	414			1	.08	18	8	1
24	4-21	731P	OSHKOSH	WI	414			2	.15	18	8	1
25	4-23	840P	OSHKOSH	WI	414			7	.76	18	8	1
26	4-24	610P	OSHKOSH	WI	414			1	.14	18	8	1
27	4-24	847P	OSHKOSH	WI	414			2	.24	18	8	1
28	4-28	240P	OSHKOSH	WI	414			10	.66	18	8	1

Figure 5 Common Telephone Record Available in the Present System

TIMBUKTU DEPT OF JUSTICE
Telephone Combination Subscriber Report - Was Printed on: Wednesday April 20, 1994 15:19:31
SUBJECT(S): AB0000

Telephone Number	# of Calls	Subscriber Name & ID#	Subscriber Address
(414) 2	1	OLD BANK 666 N ROAD APT 666	
(414) 2	5	FRANCIS E BADGUY AB0000 666 S JACKSON APT 6	
(414) 2	1	MOBIL PHONE	
(414) 3	7	MOBILE TELEPHONE	
(414) 3	40	FRANCIS E BADGUY AB0000 666 S JACKSON APT 6	

Many calls to a person with a crime record (shown by a docket number) is a clear indication that the person calling has suspicious connections.

Figure 6 Telephone Record Information Provided by Sherpa

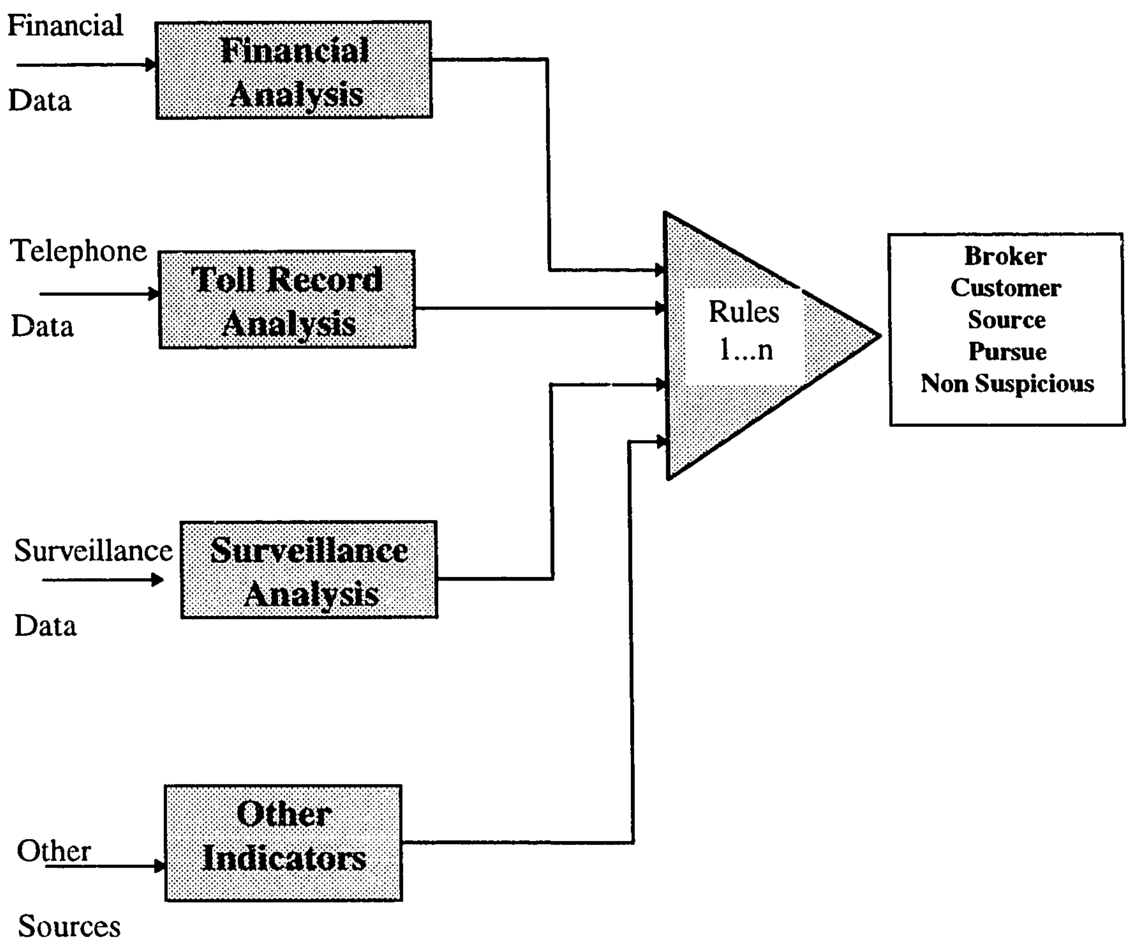

Figure 7 Dependency Diagram

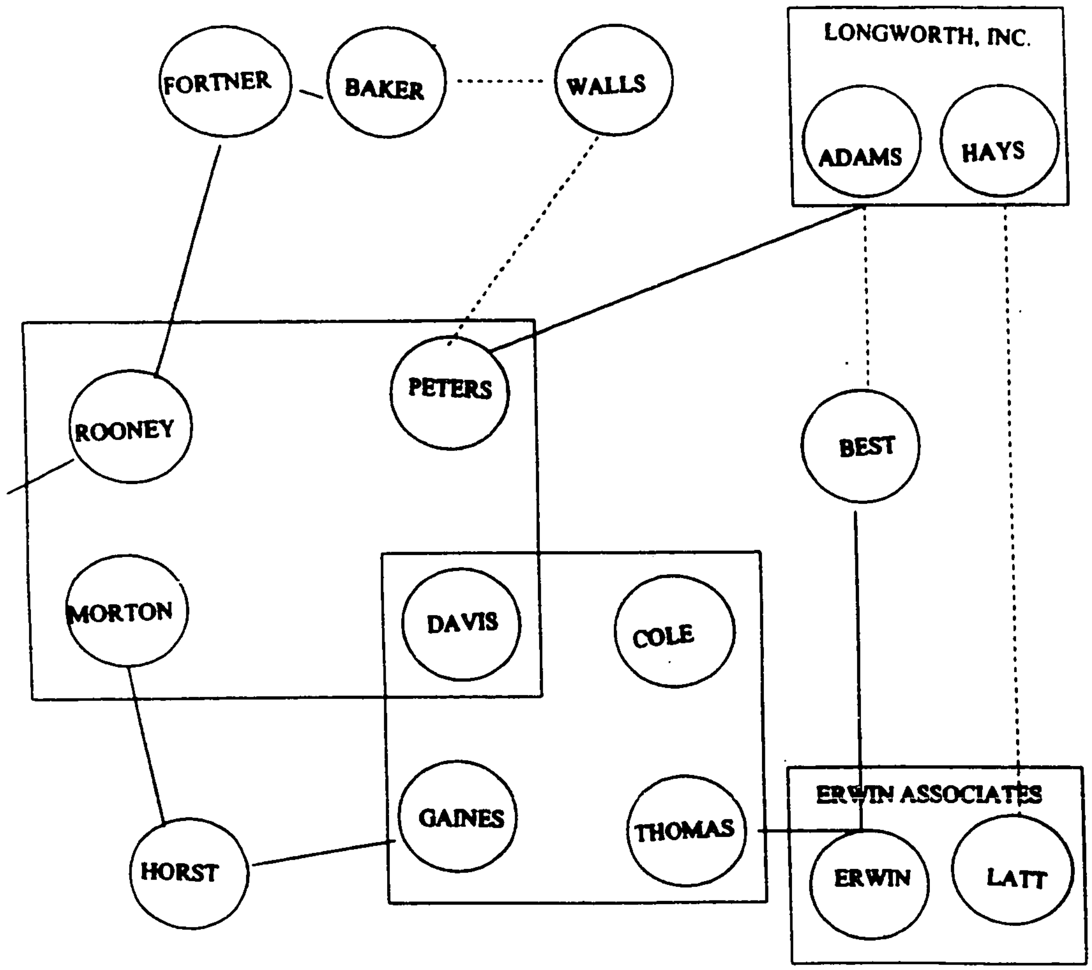

Figure 8 Link chart for Telephone Analysis in the Current system

Decision Tree:

If interview = broker then clasify the subject as broker.
If interview = customer then classify the subject as Customer.

If interview = source and citystate = source then classify subject as Source/Broker.

If interview = source and citystate =nonsource then classify subject as Source.

If interview = nonsource and citystate = source then classify subject as Suspect.

Figure 9 Example of Telephone Analysis Rules

Bank and Loan Companies
Master file of customers - depositors, debtors, safe deposit box holders
Deposit - withdrawal Records
Checking Account Application with handwritten examplars and some information on customer
Joint and business accounts show associates

Credit Card Agencies

Credit Report Agencies

Securities Brokers

Transfer Agents
Maintained by Companies to record traffic in stocks of that company, listing the owners and how much they hold

Educational Institutions, Public and Private
Records of student scholarship applications, vehicles owned by the students and registered with the University

Bonding Companies

Insurance Companies

Abstract and Title Companies

Drug Stores

Accountants and Business Services

Hospitals

Newpapers

Figure 10 Sources of Information for Financial Analysis

Variable	Madison Average (7 Agents)	Appleton Average (8 Agents)
Age	36.63	31.63
Years in Law Enforcement	7.96	8.13
Years in Narcotics Enforcement	6.82	5.39
Formal Education (in years)	14.00	15.69

Figure 11 Comparative Study of Madison and Appleton DNE Agents[9]

[9] Total Number of Agents in Wisconsin = 31

Appleton DNE office	Madison DNE office
Total Adult Population (1990 Census) =	Total Adult Population (1990 census)=
805,509	899,915
Total Drug Violations(1993) = 2019	Total Drug Violation (1993)= 2744
Counties Served by Appleton office[10]	**Counties Served by Madison office**
Brown, Calumet, Door, Florence, Fond du Lac, Forest, Green Lake, Kewaunee, Langlade, Lincoln, Manitowoc, Marinette, Marquette, Menominee, Oconto, Oneida, Outgamie, Shawno, Shaboygan, Vilas, Waupaca, Waushara, Winnebago	Adams, Columbia, Crawford, Dane, Dodge, Grant, Green, Iowa, Jefferson, Juneau, La Crosse, Lafayette, Monroe, Richland, Rock, Sauk, Vernon, Walworth

Figure 12 Demographics of the Areas Under the Jurisdiction of Madison and Appleton DNE Offices

[10] Total Number of Counties in Wisconsin = 72

Pre test and Post test Item (conducted at Madison)	Difference in the mean value before and after the treatment at Appleton
Quickness of the receipt of the analysis after the toll and financial analysis data is submitted.	No change
Number of cases for which toll and financial analysis data is obtained.	No change
Accessibility of the data for financial and toll analysis	No change
Correlation of data and agent's decision to send data for analysis	No change
Training in the use of toll and financial records	No change
Self reported ability to use toll records and financial records for investigation	No change
Criticality of toll records in solving a case	No change
Value of toll and financial records in planning further investigative technique	No change
Value of toll and financial records when considering the use of resources (manpower, financial resources, subpoena or grand jury)	No change
Value of toll and financial records to estimate the time required to conduct an investigation	No change
Value of toll and financial analysis in assessing the classification of a suspect	No change
Value of toll and financial analysis in assessing the target's vulnerability	No change

Figure 13 Results of Pre-test and Post-test study

Evidentiary Information

Variable		Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
LEVIDENC	Log of Evidence				1.0963	.548	.137
		16	.557	.025			
LSEVIDEN	Log of Sherpa Evidence				2.6423	.745	.186

Paired Differences Mean	SD	SE of Mean	t-value	df	2-tail Sig
-1.5460	.633	.158	-9.77	15	.000
95% CI (-1.883, -1.209)					

Identification

Variable		Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
LID	Log of identification				1.4130	.370	.092
		16	.729	.001			
LSID	Log of Sherpa Identification				1.9827	.405	.101

Paired Differences Mean	SD	SE of Mean	t-value	df	2-tail Sig
-.5698	.287	.072	-7.94	15	.000
95% CI (-.723, -.417)					

Figure 14 Results of the Statistical Analysis-Part-I (T-Test for Paired Samples)

Leads

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
LLEADS Log of Leads				1.0871	.459	.115
	16	.526	.036			
LSLEADS Log of Sherpa Leads				2.8024	.530	.133

Paired Differences

Mean	SD	SE of Mean	t-value	df	2-tail Sig
-1.7153	.485	.121	-14.13	15	.000
95% CI (-1.974, -1.456)					

Time

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
LSTIME Log of Sherpa time				.7413	1.417	.354
	16	.811	.000			
LTIME				2.4320	.967	.242

Paired Differences

Mean	SD	SE of Mean	t-value	df	2-tail Sig
-1.6907	.848	.212	-7.98	15	.000
95% CI (-2.142, -1.239)					

Figure 14 Results of the Statistical Analysis-Part-II (T-Test for Paired Samples)

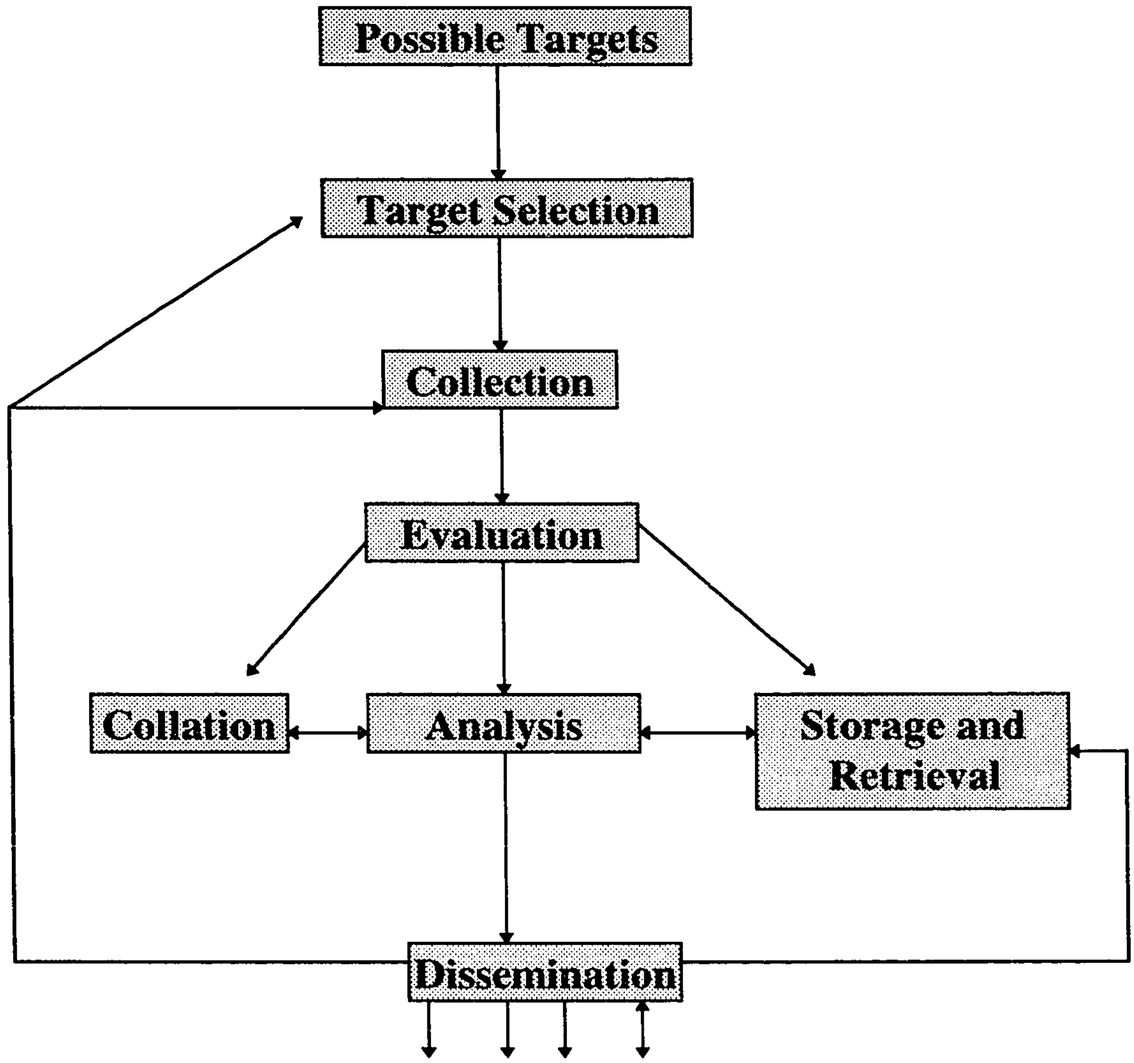

Figure A.1 Drug Crime Intelligence-Gathering Process

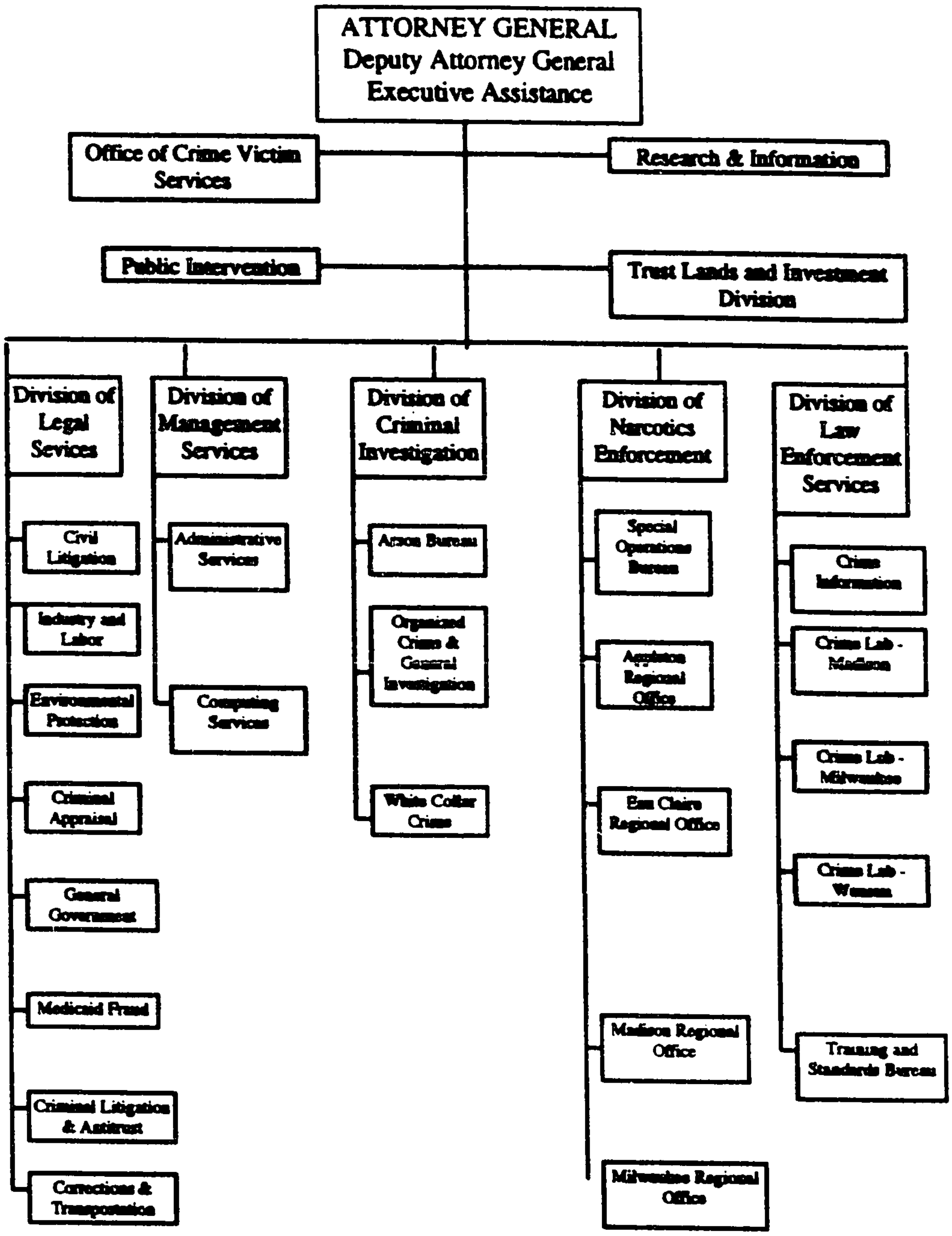

Figure C.1 Wisconsin Department of Justice Organizational Chart

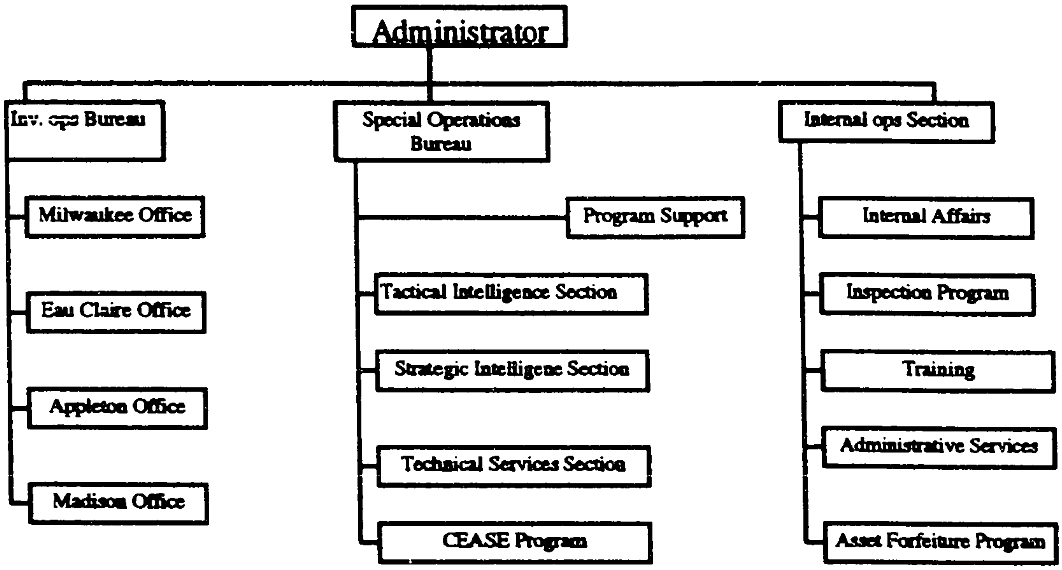

Figure C.2 Wisconsin Division of Narcotics Enforcement Organizational Chart