Sociopathic Knowledge Bases: Correct Knowledge Can Be Harmful Even Given Unlimited Computation

by

David C. Wilkins and Yong Ma

August 1989
This paper studies a situation in which correct knowledge is harmful to a problem solver even given unlimited computational resources. A knowledge base is defined to be sociopathic if all the tuples in the knowledge base are individually judged to be correct and a subset of the knowledge base gives better performance than the original knowledge base independent of the amount of computational resources that are available. Almost all knowledge bases that contain probabilistic rules are shown to be sociopathic and so this problem is very widespread.

Sociopathicity has important consequences for rule induction methods and rule set debugging methods. Sociopathic knowledge bases cannot be properly debugged using the widespread practice of incremental modification and deletion of rules responsible for wrong conclusions a la Trievesias; this approach fails to converge to an optimal solution. The problem of optimally debugging sociopathic knowledge bases is modeled as a bipartite graph minimization problem and shown to be NP-hard. Our heuristic solution approach is called the Sociopathic Reduction Algorithm and experimental results verify its efficacy.
Sociopathic Knowledge Bases:
Correct Knowledge Can be Harmful
Even Given Unlimited Computation

David C. Wilkins and Yong Ma

Knowledge-Based Systems Group
Department of Computer Science
University of Illinois
405 North Mathews Ave
Urbana, IL 61801

August 1989

Submitted for Publication:
Artificial Intelligence Journal
Sociopathic Knowledge Bases:
Correct Knowledge Can be Harmful
Even Given Unlimited Computation

David C. Wilkins and Yong Ma
Department of Computer Science
University of Illinois
405 North Mathews Avenue
Urbana, IL 61801

Abstract

This paper studies a situation in which correct knowledge is harmful to a problem solver even given unlimited computational resources. A knowledge base is defined to be sociopathic if all the tuples in the knowledge base are individually judged to be correct and a subset of the knowledge base gives better performance than the original knowledge base independent of the amount of computational resources that are available. Almost all knowledge bases that contain probabilistic rules are shown to be sociopathic and so this problem is very widespread.

Sociopathicity has important consequences for rule induction methods and rule set debugging methods. Sociopathic knowledge bases cannot be properly debugged using the widespread practice of incremental modification and deletion of rules responsible for wrong conclusions a la Teiresias; this approach fails to converge to an optimal solution. The problem of optimally debugging sociopathic knowledge bases is modeled as a bipartite graph minimization problem and shown to be NP-hard. Our heuristic solution approach is called the Sociopathic Reduction Algorithm and experimental results verify its efficacy.
Contents

1 Introduction 3

2 Inexact Reasoning and Rule Interactions 4

3 Debugging Rule Sets and Rule Interactions 6
   3.1 Types of rule interactions ................................ 6
   3.2 Traditional methods of debugging a rule set .............. 7

4 Minimizing Sociopathic Interactions 8
   4.1 Bipartite graph minimization formulation .................. 9

5 Sociopathic Reduction Algorithm 14
   5.1 The Sociopathic Reduction Algorithm ...................... 14
   5.2 Example of sociopathic reduction ............................ 16
   5.3 Experience with the Sociopathic Reduction Algorithm ....... 18

6 Related Work 20

7 Summary and Conclusion 21

8 Acknowledgements 21

Appendix 1: Calculating $G$. 22

References 73
1 Introduction

Reasoning under uncertainty has been widely investigated in artificial intelligence. Probabilistic approaches are of particular relevance to rule-based expert systems, where one is interested in modeling the heuristic and evidential reasoning of experts. Methods developed to represent and draw inferences under uncertainty include the certainty factors used in Mycin (Buchanan and Shortliffe, 1984), fuzzy set theory (Zadeh, 1979), and the belief functions of Dempster-Shafer theory (Shafer, 1976) (Gordon and Shortliffe, 1985). In many expert system frameworks, such as Emycin, Expert, MRS, S.1, and Kee, the rule structure permits a conclusion to be drawn with varying degrees of certainty or belief. This paper addresses a concern common to all these methods and systems.

In refining and debugging a probabilistic rule set, there are three major causes of errors: missing rules, wrong rules, and deleterious interactions between good rules. The purpose of this paper is to explicate a type of deleterious interaction and to show that it (a) is indigenous to rule sets for reasoning under uncertainty, (b) is of a fundamentally different nature from missing and wrong rules, (c) cannot be handled by traditional methods for correcting wrong and missing rules, and (d) can be handled by the method described in this paper.

In section 2, we describe the type of deleterious rule interactions that we have encountered in connection with automatic induction of rule sets, and explain why the use of most rule modification methods fails to grasp the nature of the problem. In section 3, we discuss approaches to debugging and refining rule sets and explain why traditional rule set debugging methods are inadequate for handling global interactions. In section 4, we formulate the problem of reducing deleterious interactions as a bipartite graph minimization problem and show that it is NP-hard. In section 5, we present a heuristic method called the Sociopathic Reduction Algorithm. Finally, our experiences in using the Sociopathic Reduction Algorithm are described.

A brief description of terminology will be helpful to the reader. Assume there exists a collection of training instances, each represented as a set of feature-value pairs of evidence and a set of hypotheses.
Rules are in Horn clause form: conclude(H, CF) :- E, where E is a conjunction of evidence, H is a hypothesis, and CF is a certainty factor or its equivalent.

A rule that correctly confirms a hypothesis generates true positive evidence; one that correctly disconfirms a hypothesis generates true negative evidence. A rule that incorrectly confirms a hypothesis generates false positive evidence; one that incorrectly disconfirms a hypothesis generates false negative evidence. False positive and false negative evidence can lead to misdiagnoses of training instances.

2 Inexact Reasoning and Rule Interactions

When operating as an evidence-gathering system (Buchanan and Shortliffe, 1984), an expert system accumulates evidence for and against competing hypotheses. Each rule whose preconditions match the gathered data contributes either positively or negatively toward one or more hypotheses. Unavoidably, the preconditions of probabilistic rules succeed on instances where the rule will be contributing false positive or false negative evidence for conclusions. For example, consider the following rule:

```
conclude(klebsiella, 0.77) :-
    finding(surgery, yes),
    finding(gram_neg_infection, yes)
```

The frequency with which R1 generates false positive evidence has a major influence on its CF of 0.77, where $-1 \leq \text{CF} \leq 1$. Indeed, given a representative set of training instances, such as a library of medical cases, the certainty factor of a rule can be given a probabilistic interpretation\(^1\) as a function $G(z_1, z_2, z_3)$, where $z_1$ is the fraction of the positive instances of a hypothesis where the rule premise succeeds, thus contributing true positive or false negative evidence; $z_2$ is the fraction of the negative instances of a hypothesis where the rule premise succeeds, thus contributing false positive or true negative evidence;

\(^1\)See Appendix 1 for a description of the function $G$. The calculations of $G$ give a purely statistical interpretation to CFs, and hence do not incorporate orthogonal utility measures as was done in MYCIN (Buchanan and Shortliffe, 1984).
and $z_3$ is the ratio of positive instances of a hypothesis to all instances in the training set. For R1 in our domain, $G(.43,.10,.22) = 0.77$ by the formulas in Appendix A, because statistics on 104 training instances yield the following values:

\[
\begin{align*}
    z_1 & : \text{E true among positive instances} = 10/23 \\
    z_2 & : \text{E true among negative instances} = 8/81 \\
    z_3 & : \text{H true among all instances} = 23/104
\end{align*}
\]

Hence, R1 generates false positive evidence on eight instances, some of which may lead to false negative diagnoses. But whether they do or not depends on the other rules in the system; hence our emphasis on taking a global perspective. The usual method of dealing with situations such as this is to make the rule fail less often by specializing its premise (Michalski et al., 1983). For example, surgery could be specialized to neurosurgery, and we could replace R1 with:

\[
\text{conclude(klebsiella, 0.92) :-} \\
\text{finding(neurosurgery, yes),} \\
\text{finding(gram_neg_infection, yes)}
\]

(R2)

On our case library of training instances for the R2 rule, $G(.26,.02,.22) = 0.92$, so R2 makes erroneous inferences in two instances instead of eight. Nevertheless, modifying R1 to be R2 on the grounds that R1 contributes to a misdiagnosis is not always appropriate; we offer three objections to this frequent practice. First, both rules are inexact rules that offer advice in the face of limited information, and their relative accuracy and correctness is explicitly represented by their respective CFs. We expect them to fail, hence failure should not necessarily lead to their modification. Second, all probabilistic rules reflect a trade-off between generality and specificity. An overly general rule provides too little discriminatory power, and a overly specific rule contributes too infrequently to problem solving. A policy on proper grain size is explicitly or implicitly built into rule induction programs; this policy should be followed as much as possible. Specialization produces a rule that usually violates such a policy. Third, if the underlying problem for an incorrect diagnosis is rule interactions, a more specialized rule, such as the specialization of R1 to R2, can be viewed as creating a potentially more dangerous rule. Although it only makes an incorrect inference in two
instead of eight instances, these two instances will be now harder to counteract when they contribute to misdiagnoses because R2 is stronger. Note that a rule with a large CF is more likely to have its erroneous conclusions lead to misdiagnoses. This perspective motivates the prevention of misdiagnoses in ways other than the use of rule specialization or generalization.

Besides rule modification, another common method of nullifying the incorrect inference of a rule in an evidence-gathering system is to introduce counteracting rules. In our example, these would be rules with a negative CF that concludes Klebsiella on the false positive training instances that lead to misdiagnoses. But since these new rules are probabilistic, they will introduce false negatives on some other training instances, and these may lead to misdiagnoses. We could add yet more counteracting rules with a positive CF to nullify any problems caused by the original counteracting rules, but these additional rules introduce false positives on yet other training instances, and these may lead to other misdiagnoses. Also, a counteracting rule is often of less quality in comparison to rules in the original rule set; if it were otherwise the induction program would have included the counteracting rule in the original rule set. Clearly, adding counteracting rules may not be necessarily the best way of dealing with misdiagnoses made by probabilistic rules.

3 Debugging Rule Sets and Rule Interactions

Assume we are given a set of probabilistic rules that were either automatically induced from a set of training cases or created manually by an expert and knowledge engineer. In refining and debugging this probabilistic rule set, there are three major causes of errors: missing rules, wrong rules, and unexpected interactions among good rules. We first describe types of rule interactions, and then show how the traditional approach to debugging is inadequate.

3.1 Types of rule interactions

In a rule-based system, there are many types of rule interactions. Rules interact by chaining together, by using the same evidence for different conclusions, and by drawing the same conclusions from different collections of evidence. Thus one of the lessons learned from research on MYCIN was that complete modularity of rules is not possible to achieve when
rules are written manually (Buchanan and Shortliffe, 1984). An expert uses other rules in a set of closely interacting rules in order to define a new rule, in particular to set a CF value relative to the CFs of interacting rules.

Automatic rule induction systems encounter the same problems. Moreover, automatic systems lack an understanding of the strong semantic relationships among concepts to allow judgments about the relative strengths of evidential support. Instead, induction systems use biases to guide the rule search (Michalski et al., 1983). The rule sets that are later analyzed for sociopathicity in this paper were generated by the induction subsystem of ODYSSEUS. The inductive biases used in this system are rule generality, whereby a rule must cover a certain percentage of instances; rule specificity, whereby a rule must be above a minimum discrimination threshold; rule colinearity, whereby rules must not be too similar in classification of the instances in the training set; and rule simplicity, whereby a maximum bound is placed on the number of conjunctions and disjunctions (Wilkins, 1987).

3.2 Traditional methods of debugging a rule set

The standard approach to debugging a rule set consists of iteratively performing the following steps:

- **Step 1.** Run the system on cases until a false diagnosis is made.
- **Step 2.** Track down the error and correct it, using one of five methods pioneered by Teiresias (Davis, 1982) and used by knowledge engineers generally:
  - Method 1: Make the preconditions of the offending rules more specific or sometimes more general.²
  - Method 2: Make the conclusions of offending rules more general or sometimes more specific.
  - Method 3: Delete offending rules.
  - Method 4: Add new rules that counteract the effects of offending rules.

²Ways of generalizing and specializing rules are nicely described in (Michalski et al., 1983). They include dropping conditions, changing constants to variables, generalizing by internal disjunction, tree climbing, interval closing, exception introduction, etc.
Method 5: Modify the strengths or CFs of offending rules.

This approach may be sufficient for correcting wrong and missing rules. However, it is flawed from a theoretical point of view, with respect to its sufficiency for correcting problems resulting from the global behavior of rules over a set of cases. It possesses two serious methodological problems. First, using all five of these methods is not necessarily appropriate for dealing with global deleterious interactions. In section 2 we explained why in some situations modifying the offending rule or adding counteracting rules leads to problems, and misses the point of having probabilistic rules, and this eliminates methods 1, 2 and 4. If rules are being induced from a representative set of training cases, modifying the strength of the rule is illegal, since the strength of the rule has a probabilistic interpretation, being derived from frequency information derived from the training instances, and this eliminates method 5. Only method 3 is left to cope with deleterious interactions. The second methodological problem is that the traditional method picks an arbitrary case to run in its search for misdiagnoses. Such a procedure will often not converge to a good rule set, even if modifications are restricted to rule deletion. The example in section 5.2 illustrates this situation.

Our perspective on this topic evolved in the course of experiments in induction and refinement of knowledge bases. Using “better” induction biases did not always produce rule sets with better performance, and this prompted investigating the possibility of global probabilistic interactions. Our original approach to debugging was similar to the Teiresias approach. Often, correcting a problem led to other cases being misdiagnosed, and in fact this type of automated incremental debugging seldom converged to an acceptable set of rules. It might have if we we engaged in the common practice of “tweaking” the CF strengths of rules. However this was not permissible, since our CF values were derived from a representative set of training cases, and have a precise probabilistic interpretation.

4 Minimizing Sociopathic Interactions

Assume there exists a large set of training instances, and a rule set for solving these instances has been induced that is fairly complete and contains rules that are individually judged to be good. By good, we mean that they individually meet some predefined quality standards such
as the biases described in section 3.1. Further, assume that the rule set misdiagnoses some of the instance in the training set. Given such an initial rule set, the problem is to find a rule set that meets some optimality criteria, such as to minimize the number of misdiagnoses without violating the goodness constraints on individual rules. Now modifications to rules, except for rule deletion, generally break the predefined goodness constraints. And adding other rules is not desirable, for if they satisfied the goodness constraints they would have been in the original rule set produced by the induction program. Hence, if we are to find a solution that meets the described constraints, the solution must be a subset of the original rule set. More formally:

Definition 1 (Sociopathic Knowledge Base) A knowledge base is sociopathic if and only if (1) all the tuples in the knowledge base are individually judged to be good; and (2) a subset of the knowledge base gives better performance than the original knowledge base independent of the amount of available computational resources.

By the definition of a sociopathic knowledge base, the best rule set is viewed as the element of the power set of rules in the initial rule set that yields a global minimum weighted error. A straightforward approach is to examine and compare all subsets of the rule set. However, the power set is almost always too large to work with, especially when the initial set has deliberately been generously generated. The selection process can be modeled as a bipartite graph minimization problem as follows.

4.1 Bipartite graph minimization formulation

A bipartite graph \( G = (V, E) \) is a graph whose vertices \( V \) can be partitioned into two sets \( V_1 \) and \( V_2 \) so that every edge in \( E \) joins a vertex in \( V_1 \) to a vertex in \( V_2 \). For each hypothesis in the set of training instances, define a directed bipartite graph \( G = (V, E) \), with its vertices \( V \) partitioned into two sets \( I \) and \( R \), as shown in Figure 1. Elements of \( R \) represent rules, and the evidential strength of \( R_j \) is denoted by \( CF_j \). Each vertex in \( I \) represents a training instance; for positive instances \( M_i \) is 1, and for negative instances \( M_i \) is -1. Arcs \( [R_j, I_i] \) connect a rule in \( R \) with the training instances in \( I \) for which its preconditions are

\[ \text{If we discover that this solution is inadequate for our needs, then introducing rules that violate the induction biases is justifiable.} \]
satisfied; the weight of arc \([R_j, I_i]\) is \(CF_j\). The weighted arcs terminating in a vertex in \(I\) are combined using an evidence combination function \(F\), which is defined by the user. The combined evidence classifies an instance as a positive instance if the combined evidence is above a user specified threshold \(CF_t\). In the example in section 5.2, \(CF_t\) is 0, while for Mycin, \(CF_t\) is 0.2.

\[
\text{Instance Set} \quad \text{Rule Set}
\]

\[
\begin{align*}
I_1 (M_1) & \quad \bullet R_1 (CF_1) \\
I_2 (M_2) & \quad \bullet R_2 (CF_2) \\
& \quad \vdots \\
I_m (M_m) & \quad \bullet R_n (CF_n)
\end{align*}
\]

Figure 1: Bipartite Graph Formulation. The left hand nodes, \(I_1, \ldots, I_m\) represent a case library of \(m\) training instances, where \(M_i\) indicates whether an instance is a positive or negative example of a hypothesis. The right hand nodes, \(R_1, \ldots, R_n\) represent a knowledge base of probabilistic rules, where \(CF_j\) is the strength of the rule. The links show which training instances \(I_1, \ldots, I_m\) satisfy the preconditions of rule \(R_j\).

More formally, assume that \(I_1, \ldots, I_m = \text{training set of instances}, \text{and } R_1, \ldots, R_n = \text{rules of an initial rule set}. Then we want to minimize:

\[
z = \sum_{i=1}^{m} b_i e_i
\]

subject to the constraints

\[
e_i = \begin{cases} 
0 & \text{if } F(a_{i1}r_1, \ldots, a_{in}r_n) > CF_t \quad \text{for } M_i = 1 \\
0 & \text{if } F(a_{i1}r_1, \ldots, a_{in}r_n) \leq CF_t \quad \text{for } M_i = -1 \\
1 & \text{otherwise}
\end{cases}
\]

10
\[ \sum_{j=1}^{n} r_j \geq R_{\text{min}} \]  \hspace{2cm} (4)

where

\begin{align*}
    b_i &= \text{bias constant to preferentially favor instances;} \\
    r_j &= \text{if } R_j \text{ is in solution rule set then 1 else 0;} \\
    a_{ij} &= \text{if arc } [R_j, I_i] \text{ exists then } CF_j \text{ else 0;} \\
    CF_t &= \text{the } CF \text{ threshold for positive classification;} \\
    F &= \text{n-ary function for combining } CFs, \text{ where } \\
    &\quad \text{the time to evaluate is polynomial in } n; \\
    R_{\text{min}} &= \text{minimum number of rules in solution set;}
\end{align*}

The problem is to find a subset of \( R \) such that the global weighted error \( z \) is minimum. That is, the solution formulation solves for \( r_j \); if \( r_j = 1 \) then rule \( R_j \) is in the final rule set. The main tasks of the user are to specify the evidence combination function \( F \) and to set up the \( a_{ij} \) matrix, which associates rules and instances and indicates the strength of the the associations. Note that the value of \( a_{ij} \) is zero if the preconditions of \( R_j \) are not satisfied in instance \( I_i \). Preference can be given to particular instances via the bias \( b_i \) in the objective function \( z \). For instance, the user may wish to favor the selection of rules that will not misdiagnose certain instances by setting the corresponding \( b_i \) to a very high value.

The \( R_{\text{min}} \) constraint forces the solution rule set to be above a minimum size. This prevents finding a solution that is too specialized for the training set, giving good accuracy on the training set but having a high variance on other sets, which would lead to poor performance.

\textbf{Theorem 1} \textit{The bipartite graph minimization problem for heuristic rule set optimization is NP-hard.}

\textbf{Proof:} To show that the bipartite graph minimization problem (BGMP) is NP-hard, we shall reduce Satisfiability problem (SAT) to it. The major difficulty is that we have to use
numerical combination functions to determine logical truth values of clauses. Assume there are \( l \) boolean variables \( A_1, \ldots, A_l \) and \( k \) clauses \( C_1, C_2, \ldots, C_k \), where \( C_i \) is a disjunction of some literals. For example, \( C_4 = A_1 \lor \neg A_3 \lor A_4 \).

1. **Input transformation:** SAT clauses are mapped into graph instance nodes and the atoms of the clauses are mapped into rule nodes. Arcs connect rule nodes to instance nodes when the respective literals appear in the respective clauses. Let \( m = k \) and \( n = l \). Let each clause represent a positive instance, then set \( M_i = 1 \) for \( 1 \leq i \leq m \). Let \( CF_j \) to be 1 for \( j = 1, 2, \ldots, n \). For each instance node \( I_i \) (corresponding to \( C_i \)), define the combination function as follows:

\[
F(a_{i1}r_1, \ldots, a_{in}r_n) = 1 - \prod_{j=1}^{n} (1 - g(a_{ij}r_j))
\]

(5)

where

\[
g(a_{ij}r_j) = \begin{cases} 
  a_{ij}r_j & \text{if } A_j \text{ appears in } C_i \\
  1 - a_{ij}r_j & \text{if } \neg A_j \text{ appears in } C_i \\
  0 & \text{otherwise}
\end{cases}
\]

(6)

Note that \( a_{ij} = CF_j = 1 \) if either \( A_j \) or \( \neg A_j \) appears in \( C_i \). Thus the \( g(a_{ij}r_j) \) function can be simplified to:

\[
g(a_{ij}r_j) = \begin{cases} 
  r_j & \text{if } A_j \text{ appears in } C_i \\
  1 - r_j & \text{if } \neg A_j \text{ appears in } C_i \\
  0 & \text{otherwise}
\end{cases}
\]

(7)

Since every clause is of the same importance, let \( b_i = 1 \) for all \( i \), for the objective function \( z \). Let \( R_{\min} = 0 \) to make its associate constraint trivially true. Finally, choose \( CF_i \) to be 0.

2. **Output transformation:** The output transformation is that (1) if \( R_j \) remains in the final rule set, \( A_j \) is assigned to be true; otherwise, it is assigned to be false; (2) SAT is satisfied only if \( z = 0 \), i.e., all the instances are correctly classified.

3. **Justification:** First, it is clear that the input and output transformations can be performed in polynomial time. Second, we will show that \( C_i \) is satisfied iff the corresponding \( I_i \) is correctly classified in the final rule set, i.e., \( e_i = 0 \). To help understand the functionality of \( g(a_{ij}r_j) \), let us rewrite it as follows:
\[ g(a_{ij}r_j) = \begin{cases} 
1 & \text{if } A_j \text{ appears in } C_i \text{ and } r_j = 1, \text{ or} \\
0 & \text{if } \bar{A}_j \text{ appears in } C_i \text{ and } r_j = 0 \\
0 & \text{otherwise} 
\end{cases} \] (8)

If part: Assume that \( e_i = 0 \), i.e., \( F(a_{i1}r_1, \ldots, a_{in}r_n) > 0 \) (\( F \) must be 1), then at least one \( g(a_{ij}r_j) \) is 1. By the definition of \( g(a_{ij}r_j) \) above, either \( A_j \) appears in \( C_i \) and \( r_j = 1 \) or \( \bar{A}_j \) appears in \( C_i \) and \( r_j = 0 \). In either case, according to the output transformation, the corresponding clause \( C_i \) is satisfied (true).

Only if part: Assume that \( C_i \) is satisfied by the truth assignment in the final rule set. Then there must exist some atom \( A_j \) such that either \( A_j \) is in \( C_i \) and it is assigned to be true or \( \bar{A}_j \) is in \( C_i \) and assigned to be false. In either case, \( g(a_{ij}r_j) = 1 \), by the output transformation and the definition of the function. Therefore, \( F(a_{i1}r_1, \ldots, a_{in}r_n) = 1 \) and \( e_i = 0 \).

To summarize, \( g(a_{ij}r_j) \) being 1 corresponds intuitively to the positive contribution made by \( A_j \) to \( C_i \).

Finally, it’s shown that SAT is satisfiable iff BGMP so constructed has a minimum objective value 0. If BGMP has a solution with \( z = 0 \), then \( e_i = 0 \) for all \( i \), because \( b_i = 1 \). Therefore each \( C_i \) is satisfied and thus SAT is satisfiable. Conversely, if the SAT is satisfiable then each \( C_i \) can be satisfied by some truth assignment of atoms. Clearly, the final rule set of the BGMP formulation (of SAT) can be easily constructed with \( z = 0 \), according to that assignment. \( \square \)

**Corollary 1** Given a positive real number \( B \), the problem of determining if there exists a rule set whose global weighted error \( z \) is less than or equal to \( B \) in the bipartite graph formulation for heuristic rule set optimization is NP-complete.

**Proof:** To show that this decision problem is in NP, we notice that it is easy to construct a polynomial algorithm for checking whether or not the (weighted) number of misdiagnosis by any given subset of \( R \) is less than or equal to \( B \). It is NP-hard by an argument similar to that in the proof of the above theorem. \( \square \)
5 Sociopathic Reduction Algorithm

In this section, a heuristic method called the Sociopathic Reduction Algorithm is described, and an example is provided based on the graph shown in Table 1.

5.1 The Sociopathic Reduction Algorithm

The following heuristic hill-climbing search method, the Sociopathic Reduction Algorithm, is one that we have developed and used in our experiments:

- Step 1. Assign values to penalty constants. Let $p_1$ be the penalty assigned to a poison rule. A poison rule is a strong rule giving erroneous evidence for a case that cannot be counteracted by the combined weight of all the rules in the rule base that give correct evidence. Let $p_2$ be the penalty for contributing false positive evidence to a misdiagnosed case, $p_3$ be the penalty for contributing false negative evidence to a misdiagnosed case, $p_4$ be the penalty for contributing false positive evidence to a correctly diagnosed case, $p_5$ be the penalty for contributing false negative evidence to a correctly diagnosed case, and $p_6$ be the penalty for using weak rules. Let $h$ be the maximum number of rules that are removed at each iteration. Let $R_{\text{min}}$ be the minimum size of the solution rule set.

- Step 2. Optional step for very large rule sets: given an initial rule set, create a new rule set containing the $n$ strongest rules for each case.

- Step 3. Find all misdiagnosed cases for the rule set. If none exists, stop. Otherwise, collect and rank the rules that contribute evidence toward these erroneous diagnoses. The rank of rule $R_j$ is $\sum_{i=1}^{6} p_i n_{ij}$, where:
  - $n_{1j} = 1$ if $R_j$ is a poison rule or its deletion leads to the creation of another poison rule and 0 otherwise.
  - $n_{2j}$ = the number of misdiagnoses for which $R_j$ gives false positive evidence;
  - $n_{3j}$ = the number of misdiagnoses for which $R_j$ gives false negative evidence;
  - $n_{4j}$ = the number of correct diagnoses for which $R_j$ gives false positive evidence;
- \( n_{sf} \) = the number of correct diagnoses for which \( R_j \) gives false negative evidence;
- \( n_{sf} \) = the absolute value of the CF of \( R_j \);

- Step 4. Eliminate the \( h \) highest ranking rules.
- Step 5. If the number of misdiagnoses is decreased, go to step 3.
- Step 6. Else, if the number of misdiagnoses begins to increase and \( h \neq 1 \), then
  - Undo the last deletion, i.e., take back the most recently removed \( h \) rules.\(^4\)
  - \( h \leftarrow h - 1 \).\(^5\)
  - Goto step 3.

- Step 7. Otherwise, i.e., if the number of misdiagnoses is increased and \( h = 1 \), then undo the last rule deletion; output the final rule set and stop.

Each iteration of the algorithm produces a new rule set, and each rule set must be rerun on all training instances to locate the new set of misdiagnosed instances. If this is particularly difficult to do, the \( h \) parameter in step 4 can be increased, but there is the potential risk of converging to a suboptimal solution. For each misdiagnosed instance, the automated reasoning system that uses the rule set must be able to explain which rules contributed to a misdiagnosis. Hence, we require a system with good explanation capabilities.

The nature of an optimal rule set differs between domains. Penalty constants, \( p_i \), are the means by which the user can define an optimal policy. For instance, via \( p_2 \) and \( p_3 \), the user can favor false positive over false negative misdiagnoses, or visa versa. For medical expert systems, a false negative is often more damaging than a false positive, as false positives generated by a medical program can often be caught by a physician upon further testing. False negatives, however, may be sent home, never to be seen again.

In our experiments, the value of the six penalty constants was \( p_i = 10^6 - i \). The \( h \) constant determines how many rules are removed on each iteration, and its value is about 5. \( R_{min} \) is the minimum size of the solution rule set, usually about 90% of the original set; its usefulness was described in section 4.1.

\(^4\)It is this step that makes it a hill-climbing algorithm.
\(^5\)Since the \( h \) is usually small, say about 5, the next incremental step of 1 is the simplest, although the more complicated schema of step decrements can be implemented for a relatively big number of \( h \).
Table 1: An example for Sociopathic Reduction algorithm. There are ten training instances that are classified as positive (+) or negative (-) instances of the hypothesis. There are six rules shown with their CF strength. The marks indicate the instances to which the rules apply, i.e., when an instance satisfies the premises clauses of a rule.

5.2 Example of sociopathic reduction

In this example, which is illustrated in Table 5.1, there are ten training instances $I_0, \ldots, I_9$, classified as positive or negative instances of the hypothesis. There are six rules $R_1, \ldots, R_6$ shown with their CF strength. The marks ($\times$) indicate the instances to which the rules apply, i.e., when an instance satisfies the premises clauses of a rule. To simplify the example, define the combined evidence for an instance as the sum of the evidence contributed by all applicable rules, and let $CF_i = 0$. Rules with a CF of one sign that are connected to an instance of the other sign contribute erroneous evidence. Two cases in the example are misdiagnosed: $I_4$ and $I_5$. The objective is to find a subset of the rule set that minimizes the number of misdiagnoses. Before the details are examined, the following points concerning examples should be made.

First, it can be shown that it is impossible to have an example using rules with out degree less than 5 that has all the points to be made from this example, if there are the equal number of positive and negative training instances. The argument is trivial for the rules with out degree of 1 and 2. For a rule with out degree of 3, assume that it has a positive CF value and is to be deleted. Then, it must misdiagnose some negative instance to become a
rule to be blamed. And, in order to have a positive CF, it must provide (positive) evidence for two positive instances, provided that the number of positive instances is equal to that of negative instances. Therefore, the number of correct diagnoses for which it gives false positive evidence must be zero, since the only negative instance that it connects to is the misdiagnosed one. Then, its ranking vector is \( (n_{1j}, n_{2j}, n_{3j}, n_{4j}, n_{5j}, n_{6j}) = (0, 1, 0, 0, 0, CF) \) which results in the smallest ranking quantity that a blamed rule with positive CF can have. Thus, the algorithm will not guarantee to chose it for deletion. The argument for rules with out degree of 4 is similar to the above, or the CF values are zeroes if the rules connect to two positive instances and two negative ones. It may be possible to devise a heuristic algorithm which gives a better computational performance from this observation.

The second point to make is that the CF values attached to the rules are the real values that are calculated based on the formula given in the appendix. Take \( R_1(+.33) \) for example.

\[
\begin{align*}
  z_1 : & \text{ E true among positive instances } = \frac{3}{5} \\
  z_2 : & \text{ E true among negative instances } = \frac{2}{5} \\
  z_3 : & \text{ H true among all instances } = \frac{5}{10}
\end{align*}
\]

Then,

\[
   z_4 = \frac{z_1z_3}{z_1z_3 + z_2(1 - z_3)} = 0.60
\]

Since \( z_4 > z_3 \),

\[
   CF = \frac{z_4 - z_3}{z_4(1 - z_3)} = \frac{1}{3} = 0.33
\]

Now the examination of the example is to be preceded. Assume that the final rule set must have at least four rules, hence \( R_{min} = 4 \). Let \( p_i = 10^{6-i} \), for \( 0 \leq i \leq 5 \), thus choosing rules in the highest category, and using lower categories to break ties.

On the first iteration, two misdiagnosed instances are found, \( I_4 \) and \( I_5 \), and four rules contribute erroneous evidence toward these misdiagnoses, \( R_1, R_2, R_4, \) and \( R_5 \). Their ranking vectors are shown in Table 2. Clearly, \( R_1 \) has the highest ranking quantity \( \sum_{i=1}^{6} p_in_{ij} \), thus
Table 2: The ranking vectors of blamed rules

|   | \( n_{1j} \) | \( n_{2j} \) | \( n_{3j} \) | \( n_{4j} \) | \( n_{5j} \) | \( n_{6j} \) |
|---|---|---|---|---|---|---|
| \( R_1 \) | 0 | 1 | 0 | 1 | 0 | 0.33 |
| \( R_2 \) | 0 | 1 | 0 | 0 | 0 | 0.75 |
| \( R_4 \) | 0 | 0 | 1 | 0 | 1 | 0.33 |
| \( R_5 \) | 0 | 0 | 1 | 0 | 0 | 0.75 |

it is chosen for deletion. On the second iteration, one misdiagnosis is found, \( I_4 \), and two erroneous rules contribute erroneous evidence, \( R_4 \) and \( R_5 \). Rules are ranked and \( R_4 \) is deleted. This reduces the number of misdiagnoses to zero and the algorithm successfully terminates.

The same example can be used to illustrate the problem of the traditional method of rule set debugging, where the order in which cases are checked for misdiagnoses influences which rules are deleted. Consider a Teiresias style program that looks at training instances and discovers \( I_4 \) is misdiagnosed. There are two rules that contribute erroneous evidence to this misdiagnosis, rules \( R_4 \) and \( R_5 \). It wisely notices that deleting \( R_4 \) causes \( I_6 \) to become misdiagnosed, hence increasing the number of misdiagnoses; so it chooses to delete \( R_5 \). However, no matter which rule it now deletes, there will always be at least one misdiagnosed case. To its credit, it reduced the number of misdiagnoses from two to one; however, it fails to converge to an rule set that minimizes the number of misdiagnoses.

5.3 Experience with the Sociopathic Reduction Algorithm

Some preliminary experiment with the Sociopathic Reduction Algorithm has been completed, using the Mycin case library which is a collection of 112 solved cases that were obtained from records at the Stanford Medical Hospital. The rule set of about 370 rules was the one after (1) correcting an incorrect domain theory, and (2) using apprenticeship learning to extend an incomplete domain theory (Wilkins and Tan, 1989). The Sociopathic Reduction Algorithm removed 21 rules from the knowledge base after 8 iterations. In Table 3, it is shown that about 10% improvement over the knowledge base tested is obtained.

Although our work is pretty much theoretical research oriented one example of experiments is not sufficient by any means. Thus, our ongoing experiments involve two kinds
| Disease                | Number Cases | Before Reduction | After Reduction |
|------------------------|--------------|------------------|-----------------|
|                        | TP | FN | FP | TP | FN | FP |
| Bacterial Meningitis   | 16 | 14 | 2 | 13 | 12 | 4 | 4 |
| Brain Abscess          | 7  | 1  | 6 | 0  | 1  | 6 | 0 |
| Cluster Headache       | 10 | 8  | 2 | 0  | 8  | 2 | 0 |
| Fungal Meningitis      | 8  | 3  | 5 | 0  | 4  | 4 | 0 |
| Migraine               | 10 | 6  | 4 | 0  | 7  | 3 | 0 |
| Myco-TB Meningitis     | 4  | 4  | 0 | 1  | 4  | 0 | 3 |
| Primary Brain Tumor    | 16 | 3  | 13| 0  | 10 | 6 | 1 |
| Subarach Hemorrhage    | 21 | 16 | 5 | 3  | 16 | 5 | 4 |
| Tension Headache       | 9  | 8  | 1 | 3  | 8  | 1 | 1 |
| Viral Meningitis       | 11 | 10 | 1 | 12 | 10 | 1 | 6 |
| None                   | 0  | 0  | 0 | 7  | 0  | 0 | 12 |
| Totals                 | 112 | 73 | 39| 39 | 80 | 32| 32 |

Table 3: The Sociopathic Reduction Algorithm, when applied to this knowledge base, improves the performance by about 10%.

of tests. First, we divide the cases into a training set and a validation set with 70% vs. 30% each, so that it can be shown that the performance improvement is carried over to the validation set. To be more accurate, we would like to randomly split the cases five times and then average the improvements. Second, we like to apply the method just described to various knowledge bases available, for example, a knowledge base after correction of wrong rules only, a knowledge base after case-based learning application, and so on.
6 Related Work

The original contribution of this paper is to show that correct knowledge can be harmful independent of problem-solving efficiency and that this problem is widespread. Another contribution is to show that the problem of harmful knowledge can be minimized and problem-solving performance improved by a particular form of knowledge base reduction, and that the optimal reduction is NP-hard.

The theme of correct knowledge being harmful has been studied by a number of other investigators. Minton has investigated how the learning of correct search control knowledge can slow down a problem solver; his solution approach is to quantify the potential utility of a new piece of control knowledge and only add those with a high utility (Minton and Carbonell, 1987). Markovitch and Scott have shown that any deductively learned knowledge affects the cost of searching a problem space; their solution approach is to use filter functions that can determine whether a piece of past knowledge that has been deductively learned should be used on a current problem (Markovitch and Scott, 1989). Still another approach is to modify learned search control knowledge to increase problem-solving speed (Prieditis and Mostov, 1987).

The theme of improving problem-solving accuracy via knowledge base reduction has been studied in conjunction with eliminating or reducing wrong knowledge. For example, the genetic algorithm used in conjunction with a classifier system eliminates as much as half of a knowledge base; it eliminates rules that has not contributed to past problem-solving successes (Holland, 1986). Another approach is to perform a global analysis of a knowledge base and eliminate those rules that are redundant or inconsistent (Ginsberg et al., 1988).

Learning systems that perform induction from noisy training instances have also addressed the problem of wrong knowledge. The RULEMOD program of META-DENDRAL selects a subset of rules that have wide applicability, thereby reducing the number of wrong rules (Buchanan and Mitchell, 1978). RULEMOD also selects rules that jointly form a good global cover and hence shares our concern for finding rules that work well together. The TRUNC program of AQ15 deletes those disjunctions of non-probabilistic induced rules that cover the fewest cases (Michalski et al., 1986a; Michalski et al., 1986b). The reduced knowledge bases produced by RULEMOD and TRUNC give equal or superior performance.
7 Summary and Conclusion

Traditional methods of debugging a probabilistic rule set are suited to handling missing or wrong rules, but not to handling deleterious interactions between good rules. This paper describes the underlying reason for this phenomenon. We formulated the problem of minimizing deleterious rule interactions as a bipartite graph minimization problem and proved that it is NP-hard. A heuristic method was described for solving the graph problem, called the Sociopathic Reduction Algorithm. In our experiments, the Sociopathic Reduction Algorithm gave good results.

We believe that the rule set refinement method described in this paper, or its equivalent, is an important component of any learning system for automatic creation of probabilistic rule sets for automated reasoning systems. All such learning systems will confront the problem of deleterious interactions among good rules, and the problem will require a global solution method, such as we have described here.

Our future research in this area is to create a theory of sociopathicity that subsumes all AI techniques for uncertainty reasoning, including certainty factors, Bayesian methods, probability methods, Dempster-Shafer theory, fuzzy reasoning, belief networks, and non-monotonic reasoning. For our progress to date, see (Ma and Wilkins, 1990a; Ma and Wilkins, 1990b; Ma and Wilkins, 1990c).

8 Acknowledgements

We thank Marianne Winslett for suggesting the bipartite graph formulation and for detailed comments, and thank Bruce Buchanan for earlier major collaboration on this work (Wilkins and Buchanan, 1986). We also express our gratitude for the helpful discussions and critiques provided by Bill Clancey, Ramsey Haddad, David Heckerman, Eric Horovitz, Curt Langlotz, Peter Rathmann and Devika Subramanian.

This work was supported in part by NSF grant MCS-83-12148, ONR grant N00014-88K-0124, and an Arnold O. Beckman research award to the first author. We are grateful for the computer time provided by the Intelligent Systems Lab of Xerox PARC and SUMEX-AIM at Stanford University.
Appendix 1: Calculating $G$.

Consider rules of the form $\text{conclude}(H, \text{CF}) :- \text{E}$. Then $\text{CF} = G = G(x_1, x_2, x_3) = \text{empirical predictive power of rule } R$, where:

- $x_1 = P(E^+|H^+) = \text{fraction of the positive instances in which } R \text{ correctly succeeds}$
  (true positives or false negatives)
- $x_2 = P(E^+|H^-) = \text{fraction of the negative instances in which } R \text{ incorrectly succeeds}$
  (false positives or true negatives)
- $x_3 = P(H^+) = \text{fraction of all instances that are positive instances}$

Given $x_1, x_2, x_3$, let

- $x_4 = P(H^+|E^+) = \frac{x_1 x_3}{x_1 x_3 + x_2 (1 - x_3)}$.

If $x_4 > x_3$ then $G = \frac{x_4 - x_1}{x_3 (1 - x_3)}$ else $G = \frac{x_4 - x_1}{x_3 (1 - x_3)}$.

This probabilistic interpretation reflects to the modifications to the certainly factor model proposed by (Heckerman, 1986).
References

Buchanan, B. G. and Mitchell, T. M. (1978). Model-directed learning of production rules. In Waterman, D. A. and Hayes-Roth, F., editors, *Pattern-Directed Inference Systems*, pages 297–312. New York: Academic Press.

Buchanan, B. G. and Shortliffe, E. H. (1984). *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, Mass.: Addison-Wesley.

Davis, R. (1982). Application of meta level knowledge in the construction, maintenance and use of large knowledge bases. In Davis, R. and Lenat, D. B., editors, *Knowledge-Based Systems in Artificial Intelligence*, pages 229–490. New York: McGraw-Hill.

Ginsberg, A., Weiss, S. M., and Politakis, P. (1988). Automatic knowledge base refinement for classification systems *Artificial Intelligence, 35*(2):197–226.

Gordon, J. and Shortliffe, E. H. (1985). A method for managing evidential reasoning in a hierarchical hypothesis space. *Artificial Intelligence, 26*(3):323–358.

Heckerman, D. (1986). Probabilistic interpretations for Mycin's certainty factors. In Kanal, L. and Lemmer, J., editors, *Uncertainty in Artificial Intelligence*, pages 167–196. New York: North Holland.

Holland, J. H. (1986). Escaping brittleness: the possibilities of general-purpose learning algorithms applied to parallel rule-based systems. In Michalski, R. S., Carbonell, J. G., and Mitchell, T. M., editors, *Machine Learning, Volume II*, volume 2, chapter 20, pages 593–624. Los Altos: Morgan Kaufmann.

Ma, Y. and Wilkins, D. C. (1990a). An analysis of Bayesian evidential reasoning. Working Paper KBS-90-001, Department of Computer Science, University of Illinois.

Ma, Y. and Wilkins, D. C. (1990b). Computation of rule probability assignments for Dempster-Shafer theory and the sociopathicity of the theory. Working Paper KBS-90-002, Department of Computer Science, University of Illinois.

Ma, Y. and Wilkins, D. C. (1990c). Sociopathicity properties of evidential reasoning systems. Working Paper KBS-90-016, Department of Computer Science, University of Illinois.

Markovitch, S. and Scott, P. D. (1989). Utilization filtering: a method for reducing the inherent harmfulness of deductively learned knowledge. In *Proceedings of the 1989 IJCAI*, pages 738–743, Detroit, MI.

Michalski, R. S., Carbonell, J. G., and Mitchell, T. M., editors (1983). *Machine Learning: An Artificial Intelligence Approach*. Palo Alto: Tioga Press.
Michalski, R. S., Mozetic, I., and Hong, I. (1986a). The AQ15 inductive learning system: An overview and experiments. Technical Report ISG 86-20, UIUCDCS-R-86-1260, Department of Computer Science, University of Illinois.

Michalski, R. S., Mozetic, I., Hong, J., and Lavrac, N. (1986b). The multi-purpose incremental learning system AQ15 and its testing application to three medical domains. In Proceedings of the 1986 National Conference on Artificial Intelligence, pages 1041–1045, Philadelphia, PA.

Minton, S. and Carbonell, J. G. (1987). Strategies for learning search control rules: An explanation-based approach. In McDermott, J., editor, Proceedings of the 1987 IJCAI, pages 228–235, Milan.

Prieditis, A. E. and Mostov, J. (1987). PROLEARN: towards a prolog interpreter that learns. In Proceedings of the 1987 National Conference on Artificial Intelligence, pages 494–498.

Shafer, G. A. (1976). Mathematical Theory of Evidence. Princeton: Princeton University Press.

Wilkins, D. C. (1987). Apprenticeship Learning Techniques For Knowledge Based Systems. PhD thesis, University of Michigan. Also, Knowledge Systems Lab Report KSL-88-14, Dept. of Computer Science, Stanford University, 1988, 153pp.

Wilkins, D. C. and Buchanan, B. G. (1986). On debugging rule sets when reasoning under uncertainty. In Proceedings of the 1986 National Conference on Artificial Intelligence, pages 448–454, Philadelphia, PA.

Wilkins, D. C. and Tan, K. (1989). Knowledge base refinement as improving an incorrect, inconsistent, and incomplete domain theory. In Proceedings of the Sixth International Conference on Machine Learning, pages 332–337, Ithaca, NY.

Zadeh, L. A. (1979). Approximate reasoning based on fuzzy logic. In Proceedings of the 1979 IJCAI, pages 1004–1010, Tokyo, Japan.
| Name                     | Address                                         |
|--------------------------|-------------------------------------------------|
| Dr. Susan Chipman        | Personnel and Training Research                 |
|                          | Office of Naval Research                        |
|                          | Code 11459                                      |
|                          | Arlington, VA 22211-2000                        |
| Dr. William J. Glassy    | NRL                                             |
|                          | 2448 Harper Street                             |
|                          | Palo Alto, CA 94304                            |
| Dr. Norman Cliff         | Department of Psychology                        |
|                          | Univ. of St. California                         |
|                          | Los Angeles, CA 90089-1001                     |
| Dr. Paul Cohen           | Computer Science Department                     |
|                          | Southern Illinois University                    |
|                          | Center                                          |
|                          | Amherst, MA 01002                              |
| Dr. Alisa Colling        | BBN                                             |
|                          | 22 Mainland Street                             |
|                          | Cambridge, MA 02138                            |
| Dr. Stanley Colyer       | Office of Naval Technology                      |
|                          | Code 222                                        |
|                          | 400 N. Quincy Street                           |
|                          | Arlington, VA 22217-1000                       |
| Dr. Greg Cooper          | Stanford University Knowledge Systems Lab      |
|                          | P. O. Box 807                                  |
|                          | Stanford, CA 94305                             |
| Dr. Richard L. Coulson   | Dept. of Physiology                             |
|                          | School of Medicine                              |
|                          | Southern Illinois University                    |
|                          | Carbondale, IL 62901                           |
| Dr. Meredith P. Crawford | University of Illinois                         |
|                          | Computer Science                               |
|                          | Champaign, IL 61801                            |
| Dr. Hans P. Crambag      | Faculty of Law                                 |
|                          | University of Limburg                          |
|                          | P.O. Box 68                                   |
|                          | Middeltricht                                    |
|                          | The NETHERLANDS 8000 MD                       |
| Dr. Kenneth B. Cross     | Anapara Sciences, Inc.                          |
|                          | P.O. Drawer Q                                  |
|                          | Santa Barbara, CA 92102                        |
| Dr. Cary Cibec           | Intelligent Instructional Systems               |
|                          | Texas Instruments AI Lab                      |
|                          | P.O. Box 86046                                 |
|                          | Dallas, TX 75280                                |
| Prof. Veronica Dahl      | Department of Computer Science                  |
|                          | Simon Fraser University                         |
|                          | Burnaby, British Columbia                      |
| Dr. John P. Dalphin     | Chair, Computer Science Dept.                  |
|                          | Towson State University                        |
| Prof. Michael G. Dyer    | Computer Science Department                     |
|                          | UCLA                                            |
| Dr. John Elle            | Navy Personnel R&D Center                       |
| Dr. Gerald F. Dejong    | Dept. of Computer Science                      |
|                          | University of Illinois                         |
|                          | Urbana, IL 61801                               |
| Dr. Thomas E. DeZarn    | Project Engineer, AT                           |
|                          | General Dynamics                               |
| Dr. Thomas G. Dietzler   | Dept. of Computer Science                      |
|                          | Oregon State University                        |
|                          | Corvallis, OR 97230                            |
| Dr. Ronna Dillon        | Department of Guidance and Educational Psychology |
|                          | Southern Illinois University                    |
|                          | Carbondale, IL 62901                           |
| Dr. J. Stuart Dunn      | Faculty of Education                            |
|                          | University of British Columbia                 |
|                          | Vancouver, BC 3470                             |
| Dr. Krizhan Donze       | George Mason University                         |
|                          | Dept. of Computer Science                      |
|                          | Fairfax, VA 22030                              |
| Dr. Pierre Dugort       | Organization for Economic                      |
|                          | Cooperation and Development                    |
|                          | 2, rue Andre-Pascal                             |
| Dr. Ralph Drey           | V-P Human Factors                              |
|                          | JPL Systems                                    |
|                          | Suite 1206                                     |
| Dr. Jean-Ellis           | Navy Personnel R&D Center                       |
| Dr. Susan Epstein       | 144 S. Mountain Avenue                         |
|                          | Montclair, NJ 07042                            |
| Dr. E. L. Fink         | Telenor, Inc.                                   |
| Dr. D. E. Fink          | 525 University Avenue                          |
|                          | Palo Alto, CA 94301                            |
| Dr. J. E. Fink          | Lockheed Austin Division                       |
|                          | 4400 Dowling Road                              |
| Dr. Glenn E. Fink       | Office of Personnel Management                  |
|                          | Office of Examination Development              |
|                          | 1400 E St., NW                                 |
| Dr. Lorraine E. Fink    | Naval Air Development Center                    |
|                          | Wernertown, PA 19744                           |
| Dr. Lawrence M. Fink    | Stanford University Medical Center              |
|                          | 1203 Medical Computer Science                   |
| Dr. P. A. Fink          | University of New England                      |
|                          | Department of Psychology                       |
|                          | Ambridge, New South Wales 22511                 |
| Mr. Nicholas S. Fink    | Dept. of Computer Science                      |
|                          | Oregon State University                        |
|                          | Corvallis, OR 97230                            |
| Dr. Kenneth D. Fink     | Division of Computer Science                    |
|                          | University of Illinois                         |
|                          | 800 N. Mathews Avenue                          |
| Dr. Kenneth Fink        | Division of Computer Science                    |
|                          | The University of Florida                      |
|                          | 11000 University Parkway                       |
|                          | Pensacola, FL 32514                            |
| Name               | Institution                                    | Address                                      |
|--------------------|------------------------------------------------|----------------------------------------------|
| Dr. Charles Pery   | Department of Computer Science                 | Carnegie-Mellon University, Pittsburgh, PA 15213 |
| Dr. Barbara A. Fox | University of Colorado                          | Department of Linguistics, Boulder, CO 80305 |
| Dr. Mark Fox       | Carnegie Mellon University                      | Robotics Institute, Pittsburgh, PA 15213     |
| Dr. Carl H. Frederiksen | Dept. of Educational Psychology            | McGill University, 2600 Tatterson Street, Montreal, Quebec CANADA H3A 1Z1 |
| Dr. John R. Frederiksen | BBN Laboratories                  | 10 Matheras Street, Cambridge, MA 02238     |
| Dr. Norman Frederiksen | Educational Testing Service                  | Princeton, NJ 08541                            |
| Dr. Alfred R. Frey     | AFOSR/AFNL, Bldg. 410                  | Bolling AFB, DC 20232-6448                   |
| Dr. Peter Friedland | Chief, AI Research Branch                      | Mail stop 144-17 NASA Ames Research Center, Moffett Field, CA 94305 |
| Dr. Michael Friendly | Psychology Department                        | York University, Toronto, ON, CANADA M4J 1P2 |
| Col. Dr. Ernst Pries | Herpetologisch Dienst                          | Maria Theresien-Kaserne, 1536 Wien, AUSTRIA  |
| Dr. Robert M. Gagne  | Knowledge Science Institute                    | University of Calgary, Calgary, Alberta CANADA T2N 1N4 |
| Dr. Brian R. Gagne   | Knowledge Science Institute                    | University of Calgary, Calgary, Alberta CANADA T2N 1N4 |
| Dr. Dorothy Genter  | Department of Psychology                      | University of Illinois, 402 E. Daniel, Champaign, IL 61820 |
| Dr. Donald Genter   | Philips Laboratories                          | 244 Statesborough Road, Briarcliff Manor, NY 10410 |
| Dr. Helen Gigney    | National Science Foundation                   | 1600 G Street N.W., Room 904, Washington, DC 20550 |
| Dr. Art Greenough   | Dept. of Psychology                           | Memphis State University, Memphis, TN 38119 |
| Dr. Philip Gillies  | Army Research Institute                       | PERI-II, 1601 Eisenhower Avenue, Alexandria, VA 22333-4400 |
| Dr. Allan Glase    | AT&T Bell Laboratories                        | Holmdel, NJ 07733                             |
| Dr. Robert Glasser  | Learning Research & Development Center        | University of Pittsburgh, 3600 O'Hara Street, Pittsburgh, PA 15206 |
| Dr. Marvin D. Glick | Homestead Terrace                           | The University of N.C., 1104 11th Street, NC 27606 |
| Dr. Dwight J. Glassing | ARI Field Unit                      | P.O. Box 6787, Presidio of Monterey, CA 92944-5211 |
| Dr. Joseph Gagnon  | Computer Science Laboratory                   | SRI International, 330 Ravenswood Avenue, Menlo Park, CA 94025 |
| Dr. Richard Golden | Psychology Department                         | Stanford University, Stanford, CA 94305     |
| Michael Haben       | DORNBIER GMBH                                 | P.O. Box 1403, D-7000 Friedrichshafen 1 WEST GERMANY |
| Dr. Richard G. Gagnon | Knowledge Science Institute                  | University of Illinois, Department of Computer Science, Urbana, IL 61801 |
| Dr. Sherry Goal    | AFRRI/OML                                    | Brooks AFB, TX 71226-5901                    |
| Dr. T. Gondrey     | Georgia Institute of Technology               | School of Industrial and Systems Engineering, Atlanta, GA 30332-0250 |
| Dr. Wayne Gray     | Artificial Intelligence Laboratory             | NYNEX, 100 Westchester Avenue, White Plains, NY 10604 |
| Dr. Al Geuse      | Dept. of Psychology                           | Memphis State University, Memphis, TN 38119 |
| Dr. Dik Gregory   | Admistratively Research Establishment/ABX    | Queens Road, Teddington, Middlesex, ENGLAND TW100LN |
| Dr. Glenn Griffin | Naval Education and Training Program Support Activity/Instructional Technology Impl. Div. | Panama City, FL 32408-1000                   |
| Dr. Benjamin N. Greenfield | IBM T.J. Watson Lab   | P.O. Box 704, Yorktown Heights, NY 10568 |
| Dr. Stephen Greaves | Center for Adaptive Systems                  | Room 246, 111 Cummington Street, Boston, MA 02115 |
| Dr. David Hauser   | Naval Education and Training Program Support Activity/Instructional Technology Impl. Div. | Panama City, FL 32408-1000                   |
| Dr. Barbara Hayes-Roth | Knowledge Systems Laboratory                | 701 Welch Road, Palo Alto, CA 94304             |
| Dr. James Hendler  | Center for Computer Science                   | University of Maryland, College Park, MD 20742 |
| Dr. Michael Henken | DORNBIER GMBH                                 | P.O. Box 1403, D-7000 Friedrichshafen 1 WEST GERMANY |
| Dr. Chris Hammond  | Dept. of Computer Science                     | University of Chicago, 1105 E. 16th Street, Chicago, IL 60617 |
| Dr. Marjory Hendler | Center for Computer Science                   | University of Maryland, College Park, MD 20742 |
| Dr. James Hendler  | Department of Educational Development         | University of Delaware, Newark, DE 07118        |
| Dr. Geoffrey Hunt  | Computer Science Department                   | University of Toronto, Sandford Firming Building, 10 King's College Road, Toronto, Ontario CANADA M4S 1A4 |
| Name                      | Address                                      | Phone   |
|---------------------------|----------------------------------------------|---------|
| Dr. Haym Hirsh            | Dept. of Computer Science                    |         |
|                           | Rutgers University                           |         |
|                           | New Brunswick, NJ 08903                      |         |
| Dr. James E. Hoffman      | Department of Psychology                     |         |
|                           | University of Delaware                        |         |
|                           | Newark, DE 19711                             |         |
| Dr. John H. Holland       | Dept. of EE and CS                            |         |
|                           | Room 3111                                    |         |
|                           | University of Michigan                        |         |
|                           | Ann Arbor, MI 48109                          |         |
| Ms. Julia S. Hough        |                                              |         |
|                           | Philadelphia, PA 19144                        |         |
| Dr. Jack Hunter           |                                              |         |
|                           | 2322 Cowdridge Street                        |         |
|                           | Lansing, MI 49040                            |         |
| Dr. Ed Hutchins           | Intelligent Systems Group Institute for      |         |
|                           | Cognitive Science [C-011]                     |         |
|                           | UCED                                         |         |
|                           | La Jolla, CA 92032-1004                      |         |
| Dr. Wayne Ike             | Dept. of Information and CS                  |         |
|                           | University of California, Irvine             |         |
|                           | Irvine, CA 92717                             |         |
| Dr. Robin Jeffries        | Hewlett-Packard Laboratories, 3L             |         |
|                           | P.O. Box 10495                               |         |
|                           | Palo Alto, CA 94308-0071                     |         |
| Dr. Lewis Johnson         | USC                                          |         |
|                           | Information Sciences Institute                |         |
|                           | 4578 Admiralty Way, Suite 1001               |         |
|                           | Marina Del Rey, CA 92092                     |         |
| Dr. Daniel B. Jones       | U.S. Nuclear Regulatory Commission           |         |
|                           | NBR/SEEB                                      |         |
|                           | Washington, DC 20555                         |         |
| Mr. Paul L. Jones         | Research Division                            |         |
|                           | Chief of Naval Technical Training            |         |
|                           | Building East                                |         |
|                           | Naval Air Station Memphis                    |         |
|                           | Millington, TN 38054-5540                    |         |
| Dr. Randolph Jones        | Information and Computer Science             |         |
|                           | University of California                     |         |
|                           | Irvine, CA 92717                             |         |
| Mr. Roland Janse          | Mike Mars, R-108                             |         |
|                           | Bedford, MA 01730                            |         |
| Dr. Wendy Kellogg         | IBM T. J. Watson Research Ctr.               |         |
|                           | P.O. Box 704                                 |         |
|                           | Yorktown Heights, NY 10956                   |         |
| Dr. Gary Kreese           |                                              |         |
|                           | 618 Spenser Avenue                           |         |
|                           | Pacific Grove, CA 98088                      |         |
| Prof. Anwaad K. Ashi      | Department of Computer Science               |         |
|                           | University of Pennsylvania                  |         |
|                           | R-260 Meade School                           |         |
|                           | Philadelphia, PA 19104                       |         |
| Dr. Douglas Kiley         | Department of Psychology                     |         |
|                           | University of North Carolina                 |         |
|                           | Department of Statistics                     |         |
|                           | Chapel Hill, NC 77514                        |         |
| Dr. J.A.S. Kelso          | Center for Complex Systems                   |         |
|                           | Building MT 3                                |         |
|                           | Florida Atlantic University                  |         |
|                           | Boca Raton, FL 23801                        |         |
| Dr. Pat Langley           |                                              |         |
|                           | NASA Ames Research Center                    |         |
|                           | Mail Stop 44-17                             |         |
|                           | Moffett Field, CA 40035                      |         |
| Dr. Michael Kaplan        | Office of Basic Research                     |         |
|                           | U.S. Army Research Institute                 |         |
|                           | 1201 Eisenhower Avenue                      |         |
|                           | Alexandria, VA 22332-1600                   |         |
| Ms. Shyam Kapur           | Dept. of Computer Science                    |         |
|                           | Cornell University                           |         |
|                           | 4128 Upton Hall                              |         |
|                           | Ithaca, NY 14853                            |         |
| Dr. Dennis Kibler         | Department of Information & Computer Science |         |
|                           | University of California                    |         |
|                           | Irvine, CA 92717                             |         |
| Dr. Thomas Kijian         | AFRL/UT                                      |         |
|                           | Williams AFB, AZ 52400-6457                  |         |
| Dr. J. Peter Korstad      | Army Research Institute                      |         |
|                           | Orlando Field Unit                           |         |
|                           | r/o PM TRADE-E                               |         |
|                           | Orlando, FL 32812                            |         |
| Dr. Walter Kistach        | Department of Psychology                     |         |
|                           | University of Colorado                       |         |
|                           | Boulder, CO 80300-0345                      |         |
| Dr. Yves Kedraff          | Geophysical Survey and Mapping               |         |
|                           | Geophysical Survey and Mapping               |         |
|                           | Boulder, CO 80300-0345                      |         |
| Dr. Stephen Kershaw       | AFRL/UT                                      |         |
|                           | Williams AFB, AZ 52400-6457                  |         |
| Dr. J. Peter Korstad      | Army Research Institute                      |         |
|                           | Orlando Field Unit                           |         |
|                           | r/o PM TRADE-E                               |         |
|                           | Orlando, FL 32812                            |         |
| Dr. Alan M. Langford      | Learning R and D Center                     |         |
|                           | 2025 O'Shea Street                           |         |
|                           | University of Pittsburgh                     |         |
|                           | Pittsburgh, PA 15200                        |         |
| Dr. Keith R. Leiwi        | Honeywell S and RC                           |         |
|                           | 2450 Technology Drive                        |         |
|                           | Minneapolis, MN 15418                       |         |
| Dr. John Levine           | Learning R and D Center                     |         |
|                           | University of Pittsburgh                     |         |
|                           | Pittsburgh, PA 15200                        |         |
| Dr. Les S. Levy           | 1D Darade Drive                              |         |
|                           | Morristown, NJ 07460-1900                   |         |
| Name                        | Address                                      |
|-----------------------------|----------------------------------------------|
| Dr. Jedidiah L. Olsen       | Graduate School of Business                  |
| University of Michigan      | Ann Arbor, MI 48109-2846                     |
| Office of Naval Research    | Code 11403                                |
| 800 N. Quincy Street        | Arlington, VA 22211-4900                     |
| (6 Capsules)                |                                             |
| Dr. James A. Beggs         | University of Maryland                      |
| School of Medicine          | Department of Neurology                      |
| 32 South Greene Street      | Baltimore, MD 21201                         |
| Prof. Tom O'Shea            | Institute for Defense Analyses              |
| The Open University         | 1833 Beechard St. Alexandria, VA 22311        |
| Dr. Bruce Parker            | Computer Science Department                  |
| University of Texas         | Taylor Hall 1-764                          |
| Austin, TX 78712-1700       |                                             |
| Dr. Joseph Parkin           | Department of Computer and Information Science |
| University of California    | New Brunswick, NJ 08902                      |
| Dr. J. Ross Quillan         | School of Computing Sciences                |
| N.W. Institute of Technology |
| University of California    | Amherst, MA 01003                           |
| Dr. Paul S. Rehboort        | University of Southern California           |
| Information Sciences Institute |
| 8876 Adams Way             |
| Marias Del Rey, CA 95292    |
| Dr. Peter Pinaki            | Graduate School of Education                |
| EMST Division               | 6651 Teleman Hall                          |
| University of California    | Berkeley, CA 94702                          |
| Prof. Christoph K. Riehbeck | Department of Computer Science               |
| Yale University             | P.O. Box 2164                              |
| New Haven, CT 06510-2164    |                                             |
| Dr. Jeffrey L. Sennett      | Department of Computer Science               |
| Carnegie Mellon University  | Pittsburgh, PA 15213                        |
| Dr. Jeffrey C. Schimmer     | School of Computer Science                  |
| Carnegie Mellon University  | Pittsburgh, PA 15213                        |
| Lawrence School             | Psychological & Quantitative Foundations    |
| College of Education        | University of Iowa                         |
| Iowa City, IA 52242         |                                             |
| Dr. Ramesh Patil            | MIT                                          |
| Laboratory for Computer      |
| Science Room 418            | 445 Technology Square                       |
| Cambridge, MA 02139         |                                             |
| Dr. Stephen Reder           | NRREL                                       |
| 181 SW Main, Suite 670      | Portland, OR 97204                          |
| Dr. Joseph R. Quillan       | School of Computing Sciences                |
| N.W. Institute of Technology |
| University of California    | Amherst, MA 01003                           |
| Dr. Roy A. Reeman          | Computer Science Department                  |
| University of Southern      |
| California                  | Los Angeles, CA 90030                        |
| Dr. David C. Ries           | Department of Computer & Information Science |
| George Mason University     | 4600 University Drive                       |
| Fairfax, VA 22030           |                                             |
| Dr. Jeffrey C. Schimmer     | School of Computer Science                  |
| Carnegie Mellon University  | Pittsburgh, PA 15213                        |
| Dr. Jeffrey L. Sennett      | Department of Computer Science               |
| Carnegie Mellon University  | Pittsburgh, PA 15213                        |
| Lawrence School             | Psychological & Quantitative Foundations    |
| College of Education        | University of Iowa                         |
| Iowa City, IA 52242         |                                             |
| Dr. Paul S. Rehboort        | University of Southern California           |
| Information Sciences Institute |
| 8876 Adams Way             |
| Marias Del Rey, CA 95292    |                                             |
| Dr. Peter Pinaki            | Graduate School of Education                |
| EMST Division               | 6651 Teleman Hall                          |
| University of California    | Berkeley, CA 94702                          |
| Dr. Stephen Reder           | NRREL                                       |
| 181 SW Main, Suite 670      | Portland, OR 97204                          |
| Dr. Joseph R. Quillan       | School of Computing Sciences                |
| N.W. Institute of Technology |
| University of California    | Amherst, MA 01003                           |
| Dr. Roy A. Reeman          | Computer Science Department                  |
| University of Southern      |
| California                  | Los Angeles, CA 90030                        |
| Dr. David C. Ries           | Department of Computer & Information Science |
| George Mason University     | 4600 University Drive                       |
| Fairfax, VA 22030           |                                             |
| Dr. Jeffrey C. Schimmer     | School of Computer Science                  |
| Carnegie Mellon University  | Pittsburgh, PA 15213                        |
| Dr. Jeffrey L. Sennett      | Department of Computer Science               |
| Carnegie Mellon University  | Pittsburgh, PA 15213                        |
| Lawrence School             | Psychological & Quantitative Foundations    |
| College of Education        | University of Iowa                         |
| Iowa City, IA 52242         |                                             |
| Name                      | Address                                                                 | City, State, Zip |
|---------------------------|--------------------------------------------------------------------------|-----------------|
| Dr. Michael G. Shanks     | NASA Ames Research Ctr. Mail Stop 389-1                                  | MERRITT FLD, CA 50250 |
| Dr. Derek S. Siman        | Computing Science Department The University                              | Aberdeen ABF JFX S Scotland, UNITED KINGDOM |
| Dr. Patrick Suppes        | Stanford University Institute for Mathematical Studies in the Social Sciences | Stanford, CA 94305-4115 |
| Dr. Kurt Van Luehn        | Department of Psychology Carnegie-Mellon University                      | Robotics Park, Pittsburgh, PA 15210 |
| Dr. Walter L. Sallin      | Honeywell SARC 2600 Technology Drive                                    | Minneapolis, MN 55418 |
| Dr. Richard Setton        | GTE Labs                                                                 | Waltham, MA 02114 |
| Dr. Robert L. Simpson, Jr. | Dr. Howard Caingie-Mlo Uiest Crawford Hall                              | College Park, MD 20742 |
| Dr. Edward E. Smith       | Department of Psychology University of Michigan                          | 200 Packard Road Ana. Ariz. 85233 |
| Dr. William Swartout      | USC                                                                      | Marina Del Rey, CA 90292 |
| Mr. Rajiv Tadepalli       | Microsoft Research Institute Department of Computer Science              | New Brunswick, NJ 08900 |
| Dr. Rand S. Siman         | Schlumberger Technologies Lab. Schlumberger Palo Alto Research          | Palo Alto, CA 94304 |
| Dr. Eliot Schwary         | EE/CS Department University of Michigan                                  | Ann Arbor, MI 48109-2127 |
| Dr. Terry W. Thordyke      | PMC Corporation                                                          | Central Engineering Labs 1554 Coleman Ave., Box 180 Santa Clara, CA 95050 |
| Dr. Howard Shebe          | Symbolec, Inc.                                                           | Eleven Cambridge Center, Cambridge, MA 02142 |
| Dr. Chris Teng            | Department of Computer Science University of Illinois                    | New Brunswick, NJ 08900 |
| Dr. David B. Wilkin        | Department of Computer Science University of Illinios                    | 405 N. Matthews Avenue Urbana, IL 61801 |
| Dr. Randall Shoemaker     | Naval Research Laboratory Cens 5210                                      | 455 Overlook Avenue, S.W. Washington, DC 20376-5000 |
| Dr. Frederick Steinhauser  | CIA-ORD                                                                  | Amos Building Washington, DC 20588 |
| Dr. Douglas Towsen        | Behavioral Technology Lab University of Southern California              | 1844 S. Eleon Ave. Redondo Beach, CA 90277 |
| Dr. Keith B. Wendt        |PMC Corporation                                                          | Central Engineering Labs 1505 Coleman Ave., Box 180 Santa Clara, CA 95050 |
| Dr. Bernhard Siler        | Information Sciences                                                   | Fundamental Research Laboratory GTE Laboratories, Inc. 45 System Road Waltham, MA 02254 |
| Dr. Ted Steinkra          | Dept. of Geography                                                       | University of South Carolina Columbia, SC 29628 |
| Dr. Michael S. Simon      | Department of Computer Science and Psychology                            | Carnegie Mellon University Pittsburgh, PA 15213 |
| Dr. Leon Storl             | Dept. of Computer Engineering and Science                                | Case Western Reserve University Cleveland, Ohio 44106 |
| Lt. Col. Edward Trautman  | Naval Training Systems Center                                            | 12060 Research Parkway Orlando, FL 32808 |
| Dr. Paul B. Williams      | Institute for Simulation and Training The University of Central Florida | 12456 Research Parkway, Suite 300 Orlando, FL 32826 |
| Dr. Robert L. Simmons, Jr. | DARPA/ISTD                                                              | 1600 Wilson Blvd Arlington, VA 22209-2208 |
| Dr. Michael J. Strait      | UMD Computer Graduate School                                            | College Park, MD 20745 |
| Dr. Paul E. Uzofor        | Department of Computer and Information Science                           | University of Massachusetts Amherst, MA 01003 |
| Dr. Zita M. Simons        | Chief, Technologies for Skill Acquisition and Retention                  | 1051 Eisenhower Avenue Alexandria, VA 22332 |
| Dr. Harold P. Van Cott    | Committee on Human Factors National Academy of Sciences                 | 101 Constitution Ave. Washington, DC 20548 |
| Dr. Edward Wernick         | Honeywell S and RC                                                      | 1660 Technology Drive Minneapolis, MN 55418 |
ONR DISTRIBUTION LIST [ILLINOIS/WILKINS]

Dr. Paul T. Wegig
Army Research Institute
8001 Eisenhower Ave.
ATTN: PERI-R.
Alexandria, VA 2233-1600

Dr. Joseph Wehl
Alphatech, Inc.
3 Burlington Executive Center
111 Middlesex Turnpike
Burlington, MA 01803

Dr. Beverly P. Wexel
Dept. of Computer and Information Sciences
University of Massachusetts
Amherst, MA 01003

Dr. Ronald R. Yager
Machine Intelligence Institute
Iona College
New Rochelle, NY 10801

Dr. Wassam Yagdami
Dept. of Computer Science
University of Easter
Princes of Wales Road
Easter EX4 4PT
ENGLAND

Dr. Joseph L. Young
National Science Foundation
Room 320
1800 G Street, N.W.
Washington, DC 20550

Dr. Maria Zemankova
National Science Foundation
1800 G Street N.W.
Washington, DC 20550

Dr. Uri Zernik
GE - CRD
P. O. Box 8
Schenectady, NY 12301