Constructing the Correct Diagnosis When Symptoms Disappear

Nancy E. Reed
Department of Computer and Information Science
Linköpings Universitet, S-581 83 Linköping, SWEDEN
nanre@ida.liu.se

Abstract

When multiple defects (also called diseases or faults) are present, there is a possibility of *interactions* between the defects. When defects interact, the *cues* (data obtainable) for a combination of defects is not a simple sum of the cues observable for the component defects. Expected cues may be missing, altered, or new cues may appear. Each of these alterations of cues makes diagnosis more difficult, as the correct defect combination may not even be considered (triggered) by a diagnostic system. We present an algorithm for heuristic solution construction that integrates multiple types of information about the case. Solutions are evaluated based on how many of the abnormal cues are accounted for, with a method that combines cues that may be altered due to interactions between defects. The method can account for cues that combine with one another in three basic ways, set union, additively and ordered dominance (some values mask other values) or with a combination of those basic ways.

For the solution space of one task, diagnosing congenital heart defects, we considered seven major defects and found the solution space (exhaustive) was reduced by approximately 50% because some of the defects could not physically occur together. Experimental results on cases from hospital files demonstrate the effectiveness of the heuristic solution construction algorithm to generate the correct solution early which reduced the number of solutions explored (compared to an exhaustive search) even further on most cases. With the computational power of current workstations, even cases requiring exploration of this entire solution space required less than 4 minutes of CPU time per case.

Introduction

Cue refers to a piece of data available about the case (observed) or one expected from a defect. Cues may be either normal (expected of a normal patient) or abnormal (also called *symptoms*). Cues include test results, patient interviews, physical exams, and the patient's history. *Single defect* refers to a single physical abnormality. *Disease* or *fault* are terms also used frequently. Each defect has a name that uniquely identifies it. *Multiple defect* refers to the coexistence of two or more physical abnormalities (defects), independent of any causal relationship(s). A multiple defect with a unique name will be called a *complex* defect.

Diagnosing multiple defects continues to be a difficult problem in many domains, especially medical domains. When multiple defects might be present, the number of potential solutions to each problem is greatly increased. Multiple defects are *interacting* when the cues from the multiple defect case are not set additive (Patil 1988) when compared to the cues for the component defects. Diagnosis is even more difficult if the defects interact. In particular, when defects interact, expected abnormal cues may be combined, missing, or altered, and new abnormal cues may appear.

Bylander, et al. (1991) have shown that abduction problems are in general intractable. One exception is finding one best explanation for an ordered, independent, monotonic abduction problem. Interacting defects are clearly not in the tractable category, since cues may cancel. As a result, solutions to multiple defect problems will continue to require a great deal of computational power. Within these constraints, a combination of efficient heuristic solution construction algorithms and increasingly powerful computers allows us to tackle interesting diagnostic problems, one of which is described in this paper.

Diagnostic Control Algorithm

This is a decision-support approach to diagnosis, in other words, the goal is not "a diagnosis", but rather to produce evidence that compares alternative solutions. This approach uses a ranking of solutions based on how many of the abnormal cues in the case are accounted for and identifying which one(s) are not. Any solutions accounting for all or almost all of the abnormal cues can be considered potential diagnoses.

This approach applies to multiple interacting defects and synthesizes ideas from a number of diagnostic approaches including set covering (Peng and Reggia 1990), recognition-based reasoning (Thompson *et al.* 1983; Johnson *et al.* 1988) and abduction and hypothesis assembly (Bylander *et al.* 1991; Fischer 1991;

Josephson and Josephson 1994).

This computational model uses two primary modes of reasoning. First, a forward chaining style including recognition-based reasoning is performed until all cues have been accepted. Then, an abductive style consisting of alternating solution construction and evaluation is performed until an adequate solution is found or all alternatives have been considered (Reed 1995; Reed *et al.* 1997). Heuristics are used in the construction of alternative solutions to focus on the most promising solutions first.

The modules applicable when new cues are available include two *identify features* modules and two *recognize defect* modules for recognition-based reasoning (RBR) described next. The *identify solution type* module searches for cues that can focus problem solving on a subset of the solution space. The solution type may be identified as a single defect, a named defect, a complex defect, a multiple defect, or some combination of those types. The *identify essential defects* module searches for cues that are only produced by one specific defect. These are also called pathognomonic cues. If a cue can only be caused by one defect and that cue appears, then the corresponding defect must be a component of the solution. These defects are called *essential* (Fischer 1991).

The two recognition-based reasoning (RBR) modules applicable when new cues are available propose and evaluate hypotheses (Thompson *et al.* 1983; Johnson *et al.* 1988). The *propose hypotheses* module activates physiological and defect hypotheses based on observed cues in the case. The *review hypotheses* module evaluates all active hypotheses with new information as it becomes available. Hypotheses can be in exactly one of four states, *dormant* (inactive), *proposed* (believed relevant), *accepted* (believed true), and *rejected* (believed false). All hypotheses start in a dormant state, meaning they are not currently considered relevant to the case. The other three states all describe active hypotheses. Evidence is gathered to support or oppose active hypotheses. If enough positive evidence accumulates, a proposed hypothesis will be accepted. If enough negative evidence accumulates, a proposed or accepted hypothesis will be rejected. Rejecting a hypothesis is final. Once rejected, a hypothesis cannot change state.

The second step of the control algorithm is to evaluate the current solution. If the RBR modules accept one defect after all observed cues have been processed, that solution is considered the "current solution" and is evaluated using a metric described in the next section. If the current solution explains all abnormal cues (or at least some specified cutoff value), then that is the only solution evaluated and the problem is considered solved.

In all other cases, when recognition-based reasoning (RBR) accepts no defects or more than one defect, or if the accepted defect does not explain all the abnormal cues, then solutions are constructed and evaluated by the *construct solutions* and *evaluate solutions* modules, respectively.

Evaluating Solutions

The metric used to compare solutions is called *evidence points* and is defined below (Reed *et al.* 1997). Solutions are evaluated based on the ratio of explained to total abnormal cues. For a case C and a solution S, the abnormal cues observed in the case ($Obs._C$) and expected for the defects in the solution ($Exp._S$). The best solutions are those that have the highest evidence point ratios (meaning the fewest unexplained abnormal cues). The evidence point (EP) formula is shown next.

$$EvidencePoints(C, S) = \frac{\sum Explained\ Abnormal(Obs._C\ or\ Exp._S)}{\sum Explained\ Abnormal + \sum Unexplained\ Abnormal}$$

Cues are categorized as either important or ignored. Important abnormal expected cues for each defect are classified in one of two categories - *required* or *optional*. Required expected cues are "always" present in a case when the defect is present (unless they are missing or altered due to interactions between defects). Optional expected cues are often present in cases when the defect is present, but their absence does not need a reason. Ignored cues are those that are either not important for determining a diagnosis or are not useful for discriminating between defects. Ignored cues are not included in the evaluation of solutions using the evidence points metric.

EP calculations generalize to solutions containing more than one defect and account for interactions between defects as follows. The EP calculations *cluster* cues. Cues of the same type from the case and all defects in the solution are considered together. Each type of cue has a specific combination method. The methods currently available include three basic ways - set union, additively, and ordered dominance (where "stronger" values mask other values), or a combination of the three basic ways based on characteristics of the cue, case or domain. These combination methods allow the correct interpretation of altered cues due to interacting defects.

When a case is presented, it is assumed that all abnormal cues of the important types are included, as is usually done by physicians documenting a case. If the observation of a specific cue is not possible, that cue is given a value of *unknown* for that case. Unknown cues and all expected cues of the same type are not included in the EP formula when solutions are evaluated.

Heuristic Solution Construction

Candidate solutions are constructed using the heuristic algorithm summarized in Table 1. First, solutions containing one defect (single or complex) are explored, then those with two defects, etc, up to the *maximum number of defects per solution*, which is domain dependent. Heuristic solution construction makes use of the defects proposed and accepted by the recognition-based reasoning (RBR) modules and any essential defects or solution type identified (all modules active on new cues).

In both heuristic and exhaustive search modes, solution construction and evaluation will stop when either the *first* sufficient solution or *all* solutions up to the same number of defects as the first have been constructed and evaluated. If no solutions explaining enough abnormal cues are found, processing continues until all heuristic or exhaustive solutions have been constructed and evaluated (up to the maximum number of defects per solution).

Heuristically generated solutions are constructed using the following modules: *include essential defects, cover cues, add associated defects, match solution type, eliminate incompatible defects,* and *eliminate duplicate solutions*. The *add essential defects* module makes sure that any essential defects identified are included in the solutions constructed. The *cover cues* module includes defects that explain significant abnormal observed cues in the case as identified by the RBR modules (accepted, proposed, or rejected defect hypotheses).

1. Construct 1 defect solutions in this order:
   Essential defects.
   RBR *accepted* defects.
   RBR *proposed* (including rejected) defects.
2. **For** NUMDEF = 2 **to** MAXDEFPERSOLN **do**
   Start with solutions of (NUMDEF -1) defects,
   form solutions containing NUMDEF defects by
     adding defects in the following order:
       essential defects.
       RBR accepted defects
       RBR proposed defects.
       associated defects (Common,Occasional,Rare):
   Remove duplicate solutions.
   Remove solutions of incompatible type(s).
   Remove solutions with incompatible defects.
 End For

Table 1: Heuristic solution construction algorithm.

The *add associated defects* module finds and adds defects that co-occur with a minimum of some specified frequency with some defect already under consideration to a solution. The frequencies that defects occur with other defects are classified into four categories, common, occasional, rare, and never. For each pair of defects, D_i and D_j, the rate that D_j occurs when defect D_i is present is contained in a database. Defects that never occur together can be due to the physical properties of the defects. Often the frequency that D_k appears when D_l is present is the same as that of D_l appearing when D_k is present. These appear as symmetrical entries in the matrix. However, it is possible that the frequencies will be different due to the fact that D_k and D_l can occur with greatly different frequencies.

The *eliminate duplicates*, *match solution type*, and *eliminate incompatibles* modules prune unproductive solutions that may be generated and should be self explanatory.

Exhaustive Solution Construction

Exhaustive solution construction mode, when selected, first constructs and evaluates solutions using the heuristic module. Then all solutions not constructed in the heuristic mode are constructed and evaluated. The exhaustive mode may also be automatically invoked if the heuristic mode did not find any solutions capable of explaining all or most of the important abnormal cues.

Example Domain

This section describes characteristics of the domain of pediatric cardiology. There is a standard set of data collected in this domain including history, physical exam, blood tests, cardiac auscultation, X-ray, and EKG data. Based on consultation with an expert, we chose between 2 and 5 important types of cues in each of 4 critical test areas (cardiac auscultation, EKG, physical exam, and X-ray). Cues in other areas were ignored.

Defects

In this domain, four kinds of physical defects can occur – communication defects (holes), obstructions near valves (or insufficiencies), absent or mis-connected vessels, and electromechanical defects. Electromechanical defects (other than secondary manifestations of other defects) were excluded from this investigation. They are covered in the work of others including Bratko et al. (1989) and Downing and Widman (1991).

The 7 common defects selected for this study are described next. Aortic Stenosis (AS) and Pulmonary Stenosis (PS) are valvular defects (obstructions) that restrict the flow of blood through the aortic or pulmonary valves, respectively. Atrial Septal Defect (ASD) and Ventricular Septal Defect (VSD) are communication defects, where blood flows between two normally unconnected chambers of the heart (the upper two and lower two respectively).

Tetralogy of Fallot (TF) is a complex defect with four components: VSD, PS, the aorta usually overrides the VSD, and right ventricular hypertrophy (thickening of the chamber wall) is present. Total Anomalous Pulmonary Venous Connection (TAPVC) is another complex defect. It occurs when all the vessels from the lungs connect to the right atrium instead of the left atrium. A hole in the atrial septum (ASD) is necessary with this defect, otherwise oxygenated blood could not flow to the body.

Partial Anomalous Pulmonary Venous Connection (PAPVC) is when some, but not all, of the vessels are mis-connected as in TAPVC. There need not be a hole in the atrial septum with this defect, therefore it is considered a single defect. The maximum number of (named) defects per case is considered to be three.

Defect associations and incompatibilities

Figure 1 shows the association relationships among the 7 cardiac defects examined in this study. The diagonal of the figure is crossed out since each defect can occur only once in a patient. In the top row, ASD and VSD occasionally occur with AS, while PAPVC, PS, and TAPVC rarely occur with AS. TF never occurs with AS. PS and VSD never occur with TF because they are part of the definition of TF. Similarly with ASD and TAPVC.

DEFECT \ ASSOCIATED DEFECT	Aortic Stenosis	Atrial Septal Defect	Partial APVC	Pulmonary Stenosis	Total APVC	Tetralogy of Fallot	Ventricular Septal Defect
Aortic Stenosis	X	O occasional	R rare	R rare	R rare	N never	O occasional
Atrial Septal Defect	O occasional	X	C common	C common	N never	C common	C common
Partial APVC	R rare	C common	X	R rare	N never	R rare	R rare
Pulmonary Stenosis	R rare	C common	R rare	X	R rare	N never	C common
Total APVC	R rare	N never	N never	R rare	X	R rare	R rare
Tetralogy of Fallot	N never	C common	R rare	N never	R rare	X	N never
Ventricular Septal Defect	O occasional	C common	R rare	C common	R rare	N never	X

Figure 1: Defect association relationships identified.

Some of the 7 defects are incompatible, they do not or cannot (physically) occur together, and are identified by an N in a cell of Figure 1. Ignoring the order of defects in a solution and using the maximum of 3 defects per case, a maximum of 37 possible combinations of defects can occur, 7 containing single (including complex) defects, 16 with 2 defects, and 14 with 3 defects.

In a domain with no incompatible defects (and 7 single defects where the order of defects is again ignored), there would be 63 possible combinations to search (7 single defects, 21 combinations of 2 defects, and 35 combinations of 3 defects). The search space is reduced by almost half (41% fewer) due to characteristics of the domain.

If the maximum of 3 defects per case is lifted, the reduction in the number of possible solutions is even greater - there are only 44 possible solutions containing the above defects consistent with the incompatibilities - an additional 6 possibilities with 4 defects and 1 possible combination of 5 defects. If there were no incompatible defects, there would be 127 possible solutions. The number of solutions to explore is reduced by over 65% due to the defect incompatibilities of this domain.

Tests on Hospital Cases

The diagnostic algorithm has been implemented and a knowledge base constructed for the task of diagnosing congenital heart defects. The knowledge base contains the 7 common defects described above (5 single and 2 complex). A total of 78 cases with single, complex, and multiple defects were available from hospital files for knowledge base construction and testing. Each case contained 1, 2, or 3 of the 7 defects as determined by surgery or cardiac catheterization (performed after the initial expert diagnosis). Approximately one-third of the cases were used for knowledge base construction and the other two-thirds (53 cases) were used in a "blind" test. On these cases, the original expert diagnosis is compared to the results obtained using recognition-based reasoning alone (RBR) and to the results obtained using the computational model with the heuristic solution construction (HSC) algorithm (which includes the RBR modules). It should be noted that the experts saw the patients in person while both RBR and HSC used selected information from the written records. Therefore, the experts had access to more information about the cases.

	Result	Expert	RBR	HSC
S	Correct or ranks high	32	18	30
S	UTD or partial	1	10	0
S	Incorrect or ranks low	0	5	3
C	Correct or ranks high	4	2	4
C	UTD or similar defect	2	2	1
C	Incorrect or ranks low	0	2	1
M	Correct or ranks high	8	0	7
M	UTD, or near the top	5	7	3
M	Incorrect or ranks low	1	7	4

S - single, **C** - complex, **M** - multiple, UTD - Unable to diagnose.

Table 2: Results on all 53 test cases.

Table 2 summarize the results on all the test cases. In the 33 single defect cases (S), the experts correctly identified all but 1 case where the expert gave a partially correct diagnosis. RBR alone correctly diagnosed approximately half (18) of the cases, incorrectly diagnosed 5 cases and was unable to come to a conclusion in the rest. HSC was almost as good as the experts, ranking the correct diagnosis as the best single defect solution (29/33) or very close (1/33) in approximately 90% of the cases. The center of the table shows the results on the 6 complex defect test cases (C). The experts gave the correct diagnosis in 2/3 of the cases and diagnosed a clinically similar defect to the actual defect in the other cases. RBR correctly diagnosed 1/3 of the cases, diagnosed incorrect defects in another 1/3 of the cases and was unable to diagnose the remaining 1/3. HSC ranked the correct solution at or close to the top in 2/3 of the cases and ranked a clinically similar defect higher in 1 case, again approaching the level of the experts.

On the 14 multiple defect cases (M) shown at the bottom of the table, the experts gave the correct diagnosis in over half (8) of the cases, gave alternative or partial diagnoses in approximately 1/3 (5) of the cases and incorrectly diagnosed the remaining case (with clinically very similar defects). RBR did not correctly diagnose any cases. In half of the cases, RBR diagnosed incorrect defects, in the remaining half of the cases no conclusion was reached. HSC ranked the correct solution at or near the top in half of the cases, in approximately 1/4 (3) of the cases, the correct solution ranked reasonably high, while in the remaining cases, the correct solution was not near the top. On the case misdiagnosed by the experts, HSC ranked the same incorrect, but clinically similar solution well above the correct solution (which was ranked low).

The RBR knowledge base was updated during the construction phase, but was not originally designed for multiple defects. This clearly shows as none of the multiple defect cases were correctly diagnosed by RBR. The same RBR knowledge base was used in HSC, however, and contributed to the correct diagnosis of 7 multiple defect cases, which is detailed in the next section.

In the cases where the correct solutions did not rank at the top, we analyzed all unexplained abnormal cues (missed points in the EP formula) and determined that they were due to either atypical cues present in the case, or inaccurate expected cues in the knowledge base. More knowledge base development effort can reduce or eliminate the second category. Atypical cues will always be a hazard when working with real data. The EP metric's evaluation of cases is used to highlight these "unexplained" cues to bring them to the attention of the user. Even including these unexplained abnormal cues, the HSC algorithm demonstrates a large improvement from a previous method, RBR alone, applied to the cases and approached the level of the original expert diagnoses.

Generating the correct solution

To improve the percentage of correct solutions generated by HSC in the future, we next analyze the test cases above to determine which reasoning methods generated the correct solutions.

Results on the 53 test cases are shown in Table 3. RBR activated (accepted, proposed, or rejected) the correct solution in the majority of the cases (29/33 or 88%), although only reached the correct diagnosis in about half (18). On one case, both the correct defect and an incorrect defect were accepted.

In general, this means that RBR is very good at activating the correct defect on single defect cases. It is not infallible, however. The evaluation of solutions with the EP metric gives something like a "second opinion" on potential solutions and resulted in a much higher number of correct diagnoses being rated at or near the top compared to other solutions (Table 2). However, potential solutions must be constructed before they can be evaluated. In cases where no defects were proposed, or the proposed defects did not explain very much of the data, all possible single defects were evaluated (exhaustive search).

Solution Generation	S	C	M
RBR accepted (all)	18	2	0
RBR accepted (1 of 2 or 3)	N/A	N/A	3
RBR accepted (2 of 1, 2 or 3)	1	N/A	0
RBR accepted incorrect	2	2	2
RBR accepted correct+incorrect	1	0	1
RBR rejected (1 of 1, 2 or 3)	2	0	1
RBR proposed all	6	0	2
RBR proposed (1 of 1, 2 or 3)	2	0	1
RBR proposed (2 of 2 or 3)	N/A	0	1
RBR proposed only incorrect	0	2	3
No defects proposed	2	0	0
Associated defects	N/A	N/A	2

S - single, C - complex, M - multiple

Table 3: Correct solutions generated on all 53 test cases.

In the complex and multiple defect cases, RBR activated the correct defect in 2/3 of the complex cases and at least one of the component defects in all but 3 (11/14) of the multiple defect cases. In two multiple defect cases, associated defects were necessary to construct the correct solutions since only one component defect was activated by RBR. In one case, one correct defect was proposed, and in the other, one correct defect was accepted by RBR. The other component of the correct solution was not proposed by RBR in either case. Both second defects were commonly associated with a defect under consideration. Unfortunately, in one of the cases, a single defect explained all the case cues, so this single defect solution would be preferred over a two defect solution (which was correct). The other case contained one atypical cue that was not explained by the correct two-defect solution, although the correct solution was the best two-defect solution. Two three-defect solutions (supersets of the correct solution) were able to explain all the observed cues in that case.

Discussion

Generating and testing large numbers of potential solutions to one problem has been computationally prohibitive until recently. We use a combination of heuristics to focus on the most promising solutions first, combined with reasonably fast computers. The experiments reported were performed using sun 4 (40 Mhz Sparc) workstations running Unix BSD 4.3 and Lucid Common Lisp. The fastest of the 53 problems mentioned above were solved in 3 seconds of CPU time. Even when all possible (37) solutions were explored and verbose printout was requested (generating 30-50 pages of text per case), the longest cases took less than 4 minutes of CPU time.

HSC performed very well and explored only one or a very few solutions on most single-defect cases, where there is no possibility of interactions between defects. More time and computation was focused on cases that were more difficult - ones containing multiple defects or presenting atypical cues. Effort is reduced in the evaluation of all solutions because only "important" cues are processed (14 types in the domain investigated). In addition, only cues of the same type are matched in clusters to calculate the EP, so there is no exponential growth there. The amount of time spent was well within the limits of current systems, and increasingly larger problems will be feasible as computer power doubles every few years.

Other successful approaches to the diagnosis of multiple interacting defects include model-based reasoning (de Kleer and Williams 1987; Reiter 1987), qualitative reasoning (Bratko *et al.* 1989; Downing and Widman 1991), complete simulation models (Wu 1991; Jang 1993) and probabilistic reasoning based on variations of Bayes theory (Kleiter 1992; Szolovits and Pauker 1993; Heckerman *et al.* 1995).

The significant differences and advantages of this computational model center on correctly explaining cues modified due to interactions between defects, especially in domains where complete simulation models are not available or cannot feasibly be constructed. All defects in a solution are evaluated in a cluster, when necessary, to explain abnormal cues in this method. Other methods, like symptom clustering, group cues together and explain them with one disease, but do not use a combinations of defects to explain one cue (with the exception of some additive combinations). For probabilistic methods, the interaction between defects means that the probability of an abnormal cue is altered compared to the probability calculated from the component defects. Thus the collection and use of additional statistics is necessary for each combination of defects and each type of cue where interactions are present.

Future work is planned to experiment on larger sets of defects and to explore the usefulness of different weights associated with each type of cue, giving more (or less) importance to selected cues.

Summary

Multiple interacting defects occur in many domains. On real problems, the entire search space of solutions may not need to be explored except on the most difficult cases. In the domain examined, we found that the number of possible solutions was greatly reduced, by almost half, due to incompatible defects. There were also a relatively small number (3) of maximum defects per solution. Current computational resources easily processed even an exhaustive generation and evaluation of solutions. Other domains may produce similar results.

References

Ivan Bratko, I. Mozetič, and N. Lavrač. *KARDIO: A Study in Deep and Qualitative Knowledge for Expert Systems.* MIT Press, Cambridge, MA, 1989.

Tom Bylander, D. Allemang, M. C. Tanner, and J. R. Josephson. The computational complexity of abduction. *Artificial Intelligence*, 49:25–60, 1991.

Johan de Kleer and B. C. Williams. Diagnosing multiple faults. *Artificial Intelligence*, 32(1):97–130, 1987.

K. L. Downing and L. E. Widman. Extending model-based diagnosis to medicine. In *5th Workshop on Qualitative Reasoning*, pp. 263–273, Austin, TX, 1991.

Olivier Fischer. *Cognitively Plausible Heuristics to Tackle the Computational Complexity of Abductive Reasoning.* PhD thesis, Ohio State University, 1991.

David Heckerman, J. S. Breese, and K. Rommelse. Decision-theoretic troubleshooting. *Communicaitons of the ACM*, 38(3):49–57, 1995.

Y. Jang. *HYDI: A hybrid System with Feedback for diagnosing Multiple Disorders.* PhD thesis, MIT, 1993.

Paul E. Johnson, J. B. Moen, and W. B. Thompson. Garden path errors in diagnostic reasoning. In L. Bolc and M.J. Coombs, editors, *Expert System Applications*, pp. 395–427. Springer-Verlag, 1988.

John R. Josephson and S. G. Josephson, editors. *Abductive Inference: Computation, Philosophy, Technology.* Cambridge University Press, 1994.

Gernot D. Kleiter. Bayesian diagnosis in expert systems. *Artificial Intelligence*, 54(1):1–32, 1992.

Ramesh S. Patil. Artifical intelligence techniques for diagnostic reasoning in medicine. In Shrobe and the AAAI, editors, *Exploring Artificial Intelligence*, pp. 347–379, San Mateo, CA, 1988. Morgan Kaufmann.

Yun Peng and James A. Reggia. *Abductive Inference Models for Diagnostic Problem-Solving.* Springer-Verlag, New York, N.Y., 1990.

N. E. Reed, M. Gini, P. E. Johnson, and J. H. Moller. Diagnosing congenital heart defects using the fallot computational model. *Artificial Intelligence in Medicine*, 10:25–40, 1997.

N. E. Reed. *Diagnosing Multiple Interacting Defects with Cue Combination Descriptions.* PhD thesis, University of Minnesota, June 1995.

Raymond Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32:57–95, 1987.

Peter Szolovits and S. G. Pauker. Categorical and probabilistic reasoning in medicine revisited. *Artificial Intelligence*, 42:167–180, 1993.

W. B. Thompson, P. E. Johnson, and J. B. Moen. Recognition-based diagnostic reasoning. In *Proc. 8th IJCAI*, pp. 236–238, Karlsruhe, W. Germany, 1983.

Thomas D. Wu. A problem decomposition method for efficient diagnosis and interpretation of multiple disorders. *Comp. Meth. and Prog. in Biomed.*, 35:239–250, 1991.