

# Principles of Diagnosis: Current Trends and a Report on the First International Workshop

Walter Hamscher

■ *Automated diagnosis is an important AI problem not only for its potential practical applications but also because it exposes issues common to all automated reasoning efforts and presents real challenges to existing paradigms. Current research in this area addresses many problems, including managing and structuring probabilistic information, modeling physical systems, reasoning with defeasible assumptions, and interleaving deliberation and action. Furthermore, diagnosis programs must face these problems in contexts where scaling up to deal with cases of realistic size results in daunting combinatorics. This article presents these and other issues as discussed at the First International Workshop on Principles of Diagnosis.*

Diagnosis has historically provided an obliging rock for each succeeding generation of AI researchers to blunt their axes on. Occasionally, someone chips a golden nugget from this rock: Diagnosis has motivated research efforts and new techniques, resulting in discoveries for the field that researchers still explore. The rule-based techniques in MYCIN (Shortliffe 1976) and the model-based reasoning techniques in SOPHIE III (Brown, Burton, and de Kleer 1982), to name two examples, initiated the development of technologies that are integral parts of general AI practice today. Silver nuggets, individually less valuable but far more numerous, appear as well: AI applications in diagnosis appear to have been—and probably remain—the single largest category of expert systems in use (Harmon 1988). The diagnosis problem in general might never be solved, but this fact hardly matters—just look at all the nuggets the work has produced already. Researchers continue to pound away at the diagnosis problem, each with his/her own axe. Further nuggets for AI, both gold and silver, will surely be loosened by this relentless pounding under different paradigms.

Currently, two dominant paradigms exist within the diagnosis community:

the probabilistic paradigm and the logic-based paradigm. Not surprisingly, some of the most interesting research transcends these categories; nevertheless, these categories are convenient shorthand for certain recognized communities with different approaches to diagnosis.

In the probabilistic paradigm, diagnostic knowledge is typically represented as a set of associations between disorders and their symptoms, with the task of the program being to find the set of disorders that is most likely given the symptoms. Although the probabilistic paradigm is dominated by Bayesian approaches (Szolovits and Pauker 1978), relying on mechanisms such as belief networks (Pearl 1986), other approaches such as Dempster-

---

*...two dominant paradigms exist within the diagnosis community...probabilistic ...and...logic-based...*

---

Shafer belief measures and connectionism lead to similar architectures (Shortliffe 1976; Gordon and Shortliffe 1985; Peng and Reggia 1989).

In the logic-based paradigm, diagnostic knowledge is typically represented as a first-order theory of operation for the system being diagnosed, and the diagnosis task is to find the set of diagnoses that are logically possible given the symptoms. There are *consistency-based approaches*, in which the diagnosis need only be consistent with the symptoms (Davis 1984; Genesereth 1984; Reiter 1987), and *abductive approaches*, in which the diagnosis must entail the symptoms (Patil, Szolovits, and Schwartz 1981;

Poole 1989). The term *model-based diagnosis* has historically been coextensive with logic-based diagnosis, although the real distinction is one of emphasis; research in model-based diagnosis emphasizes the nature of the model of the patient and its role in defining the space of potential diagnoses (Davis and Hamscher 1988).

The First International Workshop on Principles of Diagnosis (Dx-90), sponsored by Price Waterhouse Technology Centre and the American Association for Artificial Intelligence (AAAI) in cooperation with the Association for Computing Machinery (ACM), was held in July 1990 at Stanford University. Like any workshop, Dx-90 served as a snapshot of an area of research in progress, making explicit the current issues and themes. At Dx-90, seven themes recurred in a variety of forms, suggesting the following questions to which researchers are currently devoting their attention:

**Characterizing of diagnoses:** What is a diagnosis? A foundational step of automated diagnosis is to define what constitutes an acceptable and computable answer within a particular representation.

**Nonmonotonic reasoning:** How can an automated diagnostician emulate the ability of a human diagnostician to manipulate and revise assumptions about the nature of the system being diagnosed, the nature of the fault, and so on?

**Abduction:** Can technology developed for diagnosis be generalized to perform other abductive reasoning tasks, and conversely, does the general notion of *abduction* as “inference to the best explanation” provide any computational leverage on the more specific problem of diagnosis?

**Synthesis of logic and probability:** Can a single diagnosis framework exploit complementary advantages of probabilistic and logic-based diagnosis?

**Scaling:** What computational architectures can deliver acceptable performance and accuracy for diagnosis problems of realistic size?

**Device modeling:** What principles of constructing and using device models in the model-based paradigm will lead to useful diagnoses along with computational efficiency?

**Broadening of automated diagnosis:** How can diagnosis programs, which often focus on the narrow problem of computing the most probable diagnosis, exploit the richer context and capabilities available to human diagnosticians?

## The Characterizing of Diagnoses

In the consistency-based paradigm, a system being diagnosed is characterized by a set of components whose status is either normal or abnormal. A *diagnosis* is a set of abnormal components (along with the implicit complementary set of normal components) such that the resulting state of the system is consistent with the symptoms observed (Reiter 1987). This definition is similar to that used by Reggia, Nau, and Wang (1983), in which a diagnosis is a set of disorders, the union of whose symptoms include all observed symptoms. A *minimal diagnosis* as defined by de Kleer and Williams (1987) is one in which abnormality is minimized; that is, changing the status of any abnormal component to normal would make the diagnosis inconsistent with the symptoms. In principle, the minimal diagnoses are a compact characterization of the set of all diagnoses because every superset of a minimal diagnosis could be a diagnosis.

This characterization is adequate if the abnormality of a component is consistent with all possible behaviors. However, if components are constrained in their possible behaviors when they fail, then the notion of a minimal diagnosis is inadequate: Some supersets of a minimal diagnosis might not be consistent with the observed symptoms and, hence, not be diagnoses. As a result, researchers working within the consistency-based diagnosis paradigm have begun to question the adequacy of the notion of minimal diagnoses.

One alternative characterization of the set of diagnoses is as a set of *kernel diagnoses*. The intuition is as follows: Reasoning about the symptoms presented by a system produces *conflicts*, a sentence describing the diagnoses in conjunctive normal form (CNF). However, the description is more useful if the sentence is converted to disjunctive normal form (DNF) because in this form, each alternative diagnosis is explicitly stated as a separate clause. Previous research has focused on conflicts such as "either component *A* or *B* is abnormal, and either component *B* or *C* is abnormal," but if abnormality is not consistent with every behavior, it is possible to derive sentences of a more general form, such as "either *A* is abnormal or *B* is normal, and either *B* is abnormal or *C* is normal."

---

## *A minimal diagnosis...is one in which abnormality is minimized...*

---

Converting these more general sentences to DNF yields kernel diagnoses that provide a more constraining characterization of the space of diagnoses (de Kleer, Mackworth and Reiter 1990). Similar intuitions are exploited in the GDE+ system of Struss and Dressler (1989). Saraswat, de Kleer, and Raiman (1990) demonstrated that these syntactic characterizations of diagnoses can usefully be supplanted by model-theoretic characterizations.

An alternative characterization of diagnoses exploits the intuition that diagnoses can be clustered according to the ways in which they account for symptoms (Pople 1977; Wu 1990). In medical diagnosis, thousands of individual disorders are possible, and it is not unusual for diagnoses to involve more than three underlying disorders. Under these circumstances, the length of the sentences makes full conversion from CNF to DNF impractical. In the clustering approach, the symptoms (effectively, conflicts) are instead clustered into an intermediate form. Intuitively, symptoms are clustered when they can all be explained by a single disorder, and the clusters are mutually exclusive. To illustrate, let  $d_1$ ,  $d_2$ , and  $d_3$  be disorders, and  $s_1$  and  $s_2$  be symptoms, with  $s_1 \rightarrow d_1 \vee d_3$ , and  $s_2 \rightarrow d_2 \vee d_3$ . Suppose that we observe both symptoms  $s_1$  and  $s_2$ . In CNF, the diagnosis is  $(d_1 \vee d_3) \wedge (d_2 \vee d_3)$  and, after transforming into DNF, is  $d_3 \vee (d_1 \wedge d_2)$ . In the DNF form, each clause refers to a separate combination of disorders that could have caused the observed symptoms. The alternative is to represent the diagnoses as  $(s_1 s_2)$ , meaning that both symptoms are caused by the same disorder, and  $(s_1) (s_2)$ , meaning that there are two different disorders, one causing  $s_1$  and the other causing  $s_2$ . In cases where there are many symptoms and disorders, this intermediate form is useful because it allows attention to be focused on *parsimonious diagnoses* (those with fewer clusters), yet it is much easier to compute than a full conversion to DNF. Although the results are still in-

conclusive, clustering might be able to computationally exploit the fact that certain sets of  $d_i$  tend to account for the same  $s_j$  (Wu 1991).

## Nonmonotonic Reasoning

The notion of abnormality drawn on by work in logic-based diagnosis is nonmonotonic (Ginsberg 1987), and several researchers are actively seeking to exploit nonmonotonic reasoning in diagnostic settings. Human diagnosticians always reason under assumptions that can require revision—there are fewer than  $n$  faulty components, the system description is correct, and so on—so nonmonotonic reasoning is, in some sense, an inevitable element of diagnosis. Reiter (1987) explicitly drew a connection between diagnoses and extensions of normal default theories that laid a foundation for this intuition.

One approach is to view a set of assumptions as coherent if it has a valid extension in a nonmonotonic framework (Dressler 1990). One kind of assumption important in logic-based diagnosis is a *correctness assumption*, that is, an assumption that a particular component  $C$  in a system is working normally. Formally, it is the assumption that  $ab(C)$  is false, where  $ab(.)$  is the abnormality predicate. For this kind of assumption, a coherent assumption set is a diagnosis because it is consistent yet minimizes abnormality. This approach addresses, in a different way, the same problem addressed by de Kleer, Mackworth, and Reiter (1990): In general, not all supersets of minimal diagnoses are valid diagnoses.

A different approach is to make a connection between minimal diagnoses and the logic programming notion of a stable model (Gelfond and Lifschitz 1988). A *stable model* is essentially an extension in an autoepistemic theory. Among the advantages of this approach is that it allows the use of existing algorithms and results on computing stable models of logic programs in a logic-based diagnostic setting; further, it might allow the system description itself to be written as a nonmonotonic theory and, thus, be more compact than an ordinary first-order theory (Eshghi 1990).

Intuitively speaking, components that have exhibited no symptoms of failure can (defeasibly) be exonerated. It is reasonable to expect that nonmonotonic reasoning in a diagnostic setting should exploit this intuition.

A promising approach is to compute alibis, in the terminology of Raiman (1989). An *alibi* is a sentence that concludes that a particular component is normal if certain other components are. Alibis can be constructed by circumscribing  $ab(.)$  in the theory consisting of the union of the description of the correct behavior of the device with the observations of its actual behavior (Raiman 1990). For example, suppose that there is a trivial system description consisting of one component  $C$ , with the identity behavior  $\neg ab(C) \wedge (i = x) \rightarrow (o = x)$ . If we observe  $i = 0$  and  $o = 0$ , then circumscribing  $ab(.)$  results in the sentence  $(i = 0) \wedge (o = 0) \rightarrow \neg ab(C)$ ; that is,  $C$  is defeasibly exonerated because it has not exhibited any symptoms. A key advantage of this approach is that it can reduce the number of observations needed to conclude a diagnosis when used within a system such as GDE (de Kleer and Williams 1987) that selects observations.

### Abduction

Diagnosis can be viewed as an instance of abduction, where a proper diagnosis is essentially an explanation of how a particular set of disorders caused the observed symptoms. Abduction contrasts with *deduction*, which preserves truth, and *induction*, which preserves consistency. Abduction preserves explanations, and because abduction is inference to the best explanation, varying the definitions of best and explanation yields various known approaches to logic-based diagnosis (Console and Torasso 1990; Zadrozny 1990). The abductive approach to diagnosis, in which diagnoses must entail the symptoms, has at least one advantage over approaches in which diagnoses need only be consistent with the symptoms: It is more restrictive (Console, Dupré, and Torasso 1989; Poole 1989). Because the set of possible diagnoses is virtually always combinatoric no matter what the framework, any criterion that restricts this space without sacrificing validity is desirable.

There are several ways to extend the consistency-based approach to achieve this restriction. One method is to use a completion semantics (Clark 1978) for models of abnormal behavior. Suppose that some symptom  $s$  can be caused by just two known disorders  $d_1$  and  $d_2$ , so that  $d_1 \rightarrow s$  and  $d_2 \rightarrow s$ . By adding the axiom  $s \rightarrow (d_1 \vee d_2)$ —the completion of  $s$ —the

only consistent diagnoses will be those in which the disorders entail every observed symptom (Struss and Dressler 1989). Another method of extending the consistency-based approach is to ensure that the observations are derived in all extensions of a nonmonotonic theory (Dressler 1990). Still another method arises from extending the consistency-based paradigm with probabilistic information. In a Bayesian framework, the posterior probabilities of diagnoses that fail to entail the symptoms will naturally be lower than the posteriors of diagnoses that do not (de Kleer and Williams 1987, 1989). A more direct approach is to treat abduction as an inference procedure separate from deduction (Poole 1989).

---

*An alibi is a sentence that concludes that a particular component is normal if certain other components are.*

---

Abductive frameworks must be used with caution. It is rarely practical, even in engineered and well-understood systems such as digital circuits, to enumerate all the fault modes of every component type. Completion axioms are safe only when defeasible (Struss and Dressler 1989). In less well-understood domains, such as physiological systems, the domain experts might be willing to enumerate many disorders, but the incompleteness is always acknowledged either by a willingness to admit diagnoses that do not guarantee entailment of every observed symptom (Pople 1982; Reggia, Nau, and Wang 1983) or by the admission of unknown causes (Console, Dupré, and Torasso 1989) to account for symptoms. Another cautionary note arises from the fact that the most reliable criterion for the best explanation (hence for the best diagnosis) presumes a strong notion of causality. That is, the preferred explanations are those that explain how the underlying disorders caused the observed symptoms to appear. Although much progress has been made in the representation of causality in general (Shoham 1987), in practice, it remains difficult to construct static knowledge bases that

can adequately represent the multitude of possible causal relationships among events.

Because diagnosis can be described as a special case of abduction, it is also worthwhile to try exploiting computational mechanisms developed for model-based diagnosis to address more general abduction problems. The system CROSBY (Hamscher 1991a), based on SHERLOCK (de Kleer and Williams 1989), finds the most probable explanations of financial data. These explanations can—but do not necessarily—include hypotheses that the financial data contain errors.

### Synthesis of Logic and Probability

The two main approaches to diagnosis—probabilistic and logic based—have not yet successfully been reconciled. The attempts to synthesize these approaches is an active research area because of the strong motivations for such a synthesis.

For example, probabilistic information can be incorporated into the inner loop of the computations in a logic-based approach, specifically by ensuring that inferences are drawn only for the most probable sets of assumptions (de Kleer 1991). This modification dramatically improves the performance of the program by exploiting the observation that in a diagnostic setting, only the most probable combinations of disorders are of interest.

Also, probabilistic approaches to diagnosis, particularly when formulated as belief networks, have been shown to be practical, and important progress has been made in solving large-scale networks. However, the disorder-symptom formulation of diagnosis, used so widely in the probabilistic paradigm, is less flexible than a logic-based formulation. Difficulties include the necessity of enumerating all disorders, the disjointness of disorders from symptoms, the incorporation of contextual information, and difficulties in representing temporal relations among events. Probabilistic information can also be difficult to acquire, and as a result, it would be desirable to derive the probabilistic information from domain knowledge about causation and association, that is, from the sort of knowledge on which the logic-based approach focuses. Deriving a probabilistic formulation of a diagno-

sis problem from a partial causal model is an important and challenging problem.

Finally, most real diagnosis problems occur in situations where the available problem-solving resources must be traded against the utility of finding exact diagnoses. Diagnosis should minimize overall cost, where the cost of diagnosis includes the cost of computation, the acquisition of additional patient observations, the reaching of possibly incorrect diagnoses, appropriate therapies, and so on. Logic-based approaches offer little guidance or power with respect to this issue, but probabilistic approaches offer decision theory as an initial framework. Research that exploits the strengths of both approaches would be valuable for this reason.

## Scaling and Performance

As a practical matter, diagnosis programs must run acceptably fast, even when the set of possible diagnoses is composed of multiple disorders from a space of many hundreds of possible disorders. Researchers in the probabilistic paradigm have been concerned with this scaling issue for some time. Three common simplifying assumptions aid this effort: (1) the disorders are mutually exclusive and exhaustive, (2) the symptoms are conditionally independent, and (3) there is no masking or interaction of symptoms. One way to embody the latter assumption in a belief net formulation is by *noisy or gate* combination of symptoms; that is, each symptom could be caused by any supporting disorder, with the link events independent.

These assumptions are used in the QMR-BN project (Shwe et al. 1990), the goal of which is to demonstrate real-time diagnosis (that is, compute the most likely diagnoses given a fixed set of symptoms) using Bayesian methods on a large knowledge base. QMR (quick medical reference) is a successor to the Internist system (Pople 1977, 1982), and QMR-BN is a belief net reformulation of its knowledge base, containing some 600 disorders, 4,000 symptoms, and 40,000 associations between disorders and symptoms. In QMR-BN, the effects of multiple disorders on a common symptom are combined as a *leaky noisy or gate*, meaning that for each symptom there is an additional leak event that can cause the symptom to

occur even when none of the known disorders is present (members of the QMR-BN project are considering alternative formulations because the diagnoses found are sensitive to the leak probabilities and, in practice, are difficult to establish). Leak events play a role exactly analogous to unknown causes in logic-based abductive approaches. The class of belief nets conforming to these assumptions is termed BN20.

Because inference in general belief nets is NP-hard (Cooper 1990), attention has focused on approximate solutions. For example, the TOP N procedure (Henrion 1991) generates only the most likely  $n$  diagnoses. The performance is impressive: A Mac II implementation of TOP N was able to analyze a case with 14 positive and 10 negative symptoms in only 17 seconds, achieving an error bound of .02 on the probabilities of 7000 hypotheses. TOP N examines only the most likely combinations of disorders (which account for the majority of the probability mass anyway) and computes only the relative probabilities of different disorders instead of their absolute probabilities. TOP N begins by examining the null hypothesis (that is, no disorders and all positive symptoms caused by leak events), using a branch-and-bound strategy to compute the relative probability of each incremental hypothesis. TOP N's admissibility criterion rests on the notion of *marginal explanatory power*, which is the increase or decrease in the relative probability of a hypothesis  $H$  after including an additional disorder  $d$ . The BN20 assumptions let any search path be pruned when its marginal explanatory power is less than one and allow for relatively simple computations of upper and lower bounds on relative probabilities.

An alternative to the sequential search approach exemplified by TOP N is a connectionist approach retaining the formulation of the diagnosis problem as that of finding the most plausible set of independent disorders  $D$  accounting for a set of symptoms  $S$ . Peng and Reggia (1989) demonstrate this approach using an existing knowledge base in a medical domain that consists of causal strengths  $c_{ij}$  representing how frequently disorder  $d_i$  causes symptom  $s_j$ ; in effect,  $p(d_i \text{ causes } s_j | d_i)$ . A novel activation rule is used in which all  $c_{ij}$  are excitatory, and all the possible causes of each observed symptom (that is,  $\{d_i | c_{ij} >$

$0, \text{ and } s_j \text{ is observed}\})$  are mutually inhibitory. Intuitively speaking, given an activated set of symptoms, those disorders most strongly associated with these symptoms will be activated, but solutions will be preferred in which the fewest disorders account for any observed symptoms. Simulation results illustrate a potential for fast convergence to a solution owing to the massively parallel nature of connectionist computations, assuming computation parameters have been selected properly to avoid oscillation between solutions and other undesirable behaviors (Reggia, Peng, and Tuhim 1990). However, on some data sets, the technique finds the correct answer (in effect, a minimal diagnosis) 76 percent of the time and gets one of the three best answers 92 percent of the time. These results suggest that further research needs to be done to investigate this apparent trade-off between computational speed and diagnostic accuracy.

## Device Modeling

Constructing a formal description of the system to be diagnosed is a key step in the logic-based diagnosis paradigm. For this reason, research in this paradigm usually focuses on engineering domains in which computational models already exist, hence the phrase model-based diagnosis. New research in this area is broadening the scope of device models to which model-based diagnosis can be applied.

### Dynamic Devices

Until recently, most work in model-based diagnosis focused on devices with *static behavior*, that is, devices whose output at time  $t$  depend only on the input at  $t$ . Devices with *dynamic behavior*—whose output at  $t$  can be a function of input arbitrarily far in the past—are substantially more important to actual engineering practice, and techniques for diagnosing them are an important research subject.

For digital circuit diagnosis, an important demonstration domain, progress has been made by extending conventional constraint-based behavior representations (Sussman and Steele 1980; Davis and Shrobe 1983) to manipulate propositions representing signal values over time intervals. For example, in XDE (Hamscher

1991b), signal values are represented using propositions limited to fixed end points, but SIDIA (Guckenbiehl and Schäfer-Richter 1990) and GMODS (Holtzblatt, Neiberg, and Piazza 1991) use episode-based temporal constraint propagators similar to TCP (Williams 1986) extended in different ways to maintain assumption-based truth maintenance system (ATMS) labels (de Kleer 1986) for their inferences. The DIANA system (Dague, Jehl, and Taillibert 1990) for diagnosing analog circuits takes the approach of propagating arrays of real intervals indexed by discrete time points.

A practical difficulty encountered in extending the standard model-based approach with dynamic behavior is that not only is the space of generated diagnoses potentially combinatoric (as with static devices), but the cost of predicting behavior also increases dramatically, thus increasing the cost of testing each diagnosis for consistency (Hamscher 1991b). As a result, using structural hierarchy to coordinate the use of multiple levels of behavioral abstraction becomes a dominant theme in such systems.

The need to exploit the existence of nonintermittent failure modes is another theme in diagnosing dynamic digital circuits. A nonintermittent fault in a component results in the component always producing the same output given the same input. Although relevant to static devices, the ability to filter out diagnoses involving intermittent faults turns out to be an even more important constraint for dynamic devices (Hamscher and Davis 1984). A number of existing logic-based diagnosis systems do not exploit the diagnostic power available if nonintermittency is assumed, but they also fail to properly diagnose intermittent faults. Raiman et al. (1991) demonstrate how to incorporate a nonintermittency assumption into SHERLOCK (de Kleer and Williams 1989). The intuitions behind this work are (1) the indexing of all observations and inferences with a discrete time stamp, (2) a general nonintermittency axiom (the same input result in the same output at all times), and (3) nonintermittent misbehavior modes for components (for example, each component has both an abnormal mode and a non-intermittently abnormal mode). The ATMS inference machinery underlying SHERLOCK can represent such axioms without difficulty, enabling SHERLOCK to derive diagnoses that involve any

---

*...research in this paradigm usually focuses on engineering domains in which computational models already exist...*

---

combination of component faults, intermittent or otherwise. Empirical results indicate that in large combinatorial digital circuits, the nonintermittency assumption filters out an average of 20 percent of the diagnoses.

Another complication raised by model-based diagnosis of dynamic digital devices concerns the selection of informative observations. The key problem is that probing a signal yields a complete history, but many diagnoses predict only partial histories, so that computing the probability of an observation given a candidate (hence the expected information gain from making the observation) is not well defined. Guckenbiehl and Schäfer-Richter (1990) extended the approach used in GDE (de Kleer and Williams 1987) and XDE (Hamscher 1991b) by considering observations of entire signal histories and filtering the resulting probe suggestions on the basis of circuit structure.

### Continuous, Dynamic Devices

Although digital circuits are a popular domain for investigating model-based diagnosis, continuous systems are also an important research area. One approach to the diagnosis of continuous dynamic systems is represented by TEXSYS (Glass, Erickson, and Swanson 1991), a monitoring and diagnosis system for BATBS, a prototype thermal bus for space station Freedom. The thermal bus is a fluid system with 30 major components, most exhibiting nonlinear behavior, along with 90 sensors for flow, pressure, and temperature. There are several good reasons for selecting a model-based approach for such a problem: The target device is an engineered system consisting of identifiable components with limited interactions, its architecture is expected to undergo continuous modification, reasoning could be simplified by the expectation that no multiple faults would occur, and real-

time operation seems feasible because of the relatively slow operation of the target system. However, in TEXSYS, the lack of computational models for the individual components and the overall system behavior is a key issue. The resulting hybrid architecture uses a structural model of BATBS and qualitative behavior models for the short-term behavior of simple components (pipes, for example), but most of the diagnostic reasoning is performed by symptom-fault association rules for each primitive component and certain aggregate components.

TEXSYS illustrates that better theories of modeling continuous systems are a prerequisite for progress in this area. Progress in qualitative modeling of physical systems (Weld and de Kleer 1990) suggests that it might be an appropriate approach. For example, MIMIC (Dvorak and Kuipers 1989) demonstrates how a continuous system represented by qualitative differential equations in QSIM (Kuipers 1986) could be used to simulate the effects of hypothesized faults. Similarly, Ng (1990) demonstrates how QSIM can be used to detect inconsistent diagnoses in the logic-based approach of Reiter (1987).

The MIDAS system for continuous online monitoring of chemical plants (Oyeleye, Finch, and Kramer 1990; Rose and Kramer 1991) represents an alternative to qualitative simulation. QUAFF, the qualitative reasoning component of MIDAS, begins with a representation of the target system as a set of linear first-order differential equations. It is assumed that there is a nominal steady state—an assumption that usually holds in this domain—and the equations are rewritten to represent deviations from this steady state. This is used to construct a graph in which the nodes represent qualitative variables ranging over  $\{-,0,+\}$ , and the arcs represent the signs of the coefficients in the linear equations. The QUAFF algorithm then performs an analysis of feedback effects in this graph, finding both the initial and final response of each variable to an initial disturbance. In an empirical test on a complex system of 147 variables and 185 arcs, QUAFF was able to find unambiguous final disturbances for 85 percent of 107 initial disturbances corresponding to system component malfunctions. The results of QUAFF's feedback analysis are incorporated into an augmented graphic representation.

MIDAS uses this augmented graph representation to continuously monitor the target system and update its diagnoses. In the chemical plant domain, the time scale of disturbance propagation can range from minutes to hours, so MIDAS records each deviation of a sensor from normal as a separate event and updates its candidate set in real time after each such event. In the same empirical test mentioned earlier, with 76 randomly selected malfunctions, MIDAS ranked the actual malfunction equal to or above all other candidates in 82 percent of the cases and included the actual malfunction among the candidates after 99 percent of all events.

### Structural Failures

In general, logic-based diagnosis approaches correctly diagnose faults that result in modifications of one or more individual component behaviors but cannot propose faults that result in structural modifications. For example, *bridge faults*—solder splashes on digital circuit boards—result in new connections between structural components. One approach is to separately check for structural modifications with known behavioral effects (Davis 1984). The number of potential bridge faults can be controlled by checking only where solder splashes are most likely to occur. A different approach is to simply include potential bridge faults as insulator components in the initial description of the target device (Preist and Welham 1990). An extension of the digital circuit model to distinguish between values that are transmitted and received enables an abductive diagnosis engine to diagnose bridging faults just as it diagnoses ordinary component failures. However, one of the difficulties with this approach is that in the digital domain, the implementation of a strictly hierarchic design can involve sharing components and introducing physical adjacencies that violate this strict hierarchy (Davis 1984; Hamscher 1988); this approach appears to require the device description to be augmented with the behavioral effects of these physical adjacencies.

### Hierarchy and Abstraction

The importance of using multiple levels of abstraction has been a recurrent theme throughout the history of automated diagnosis and is particu-

---

## *The emphasis on finding therapeutic actions is... relevant when working under real-time constraints.*

---

larly relevant in the diagnosis of complex devices. Mozetic (1990) empirically demonstrates the familiar claim that hierarchic diagnosis can result in exponential speedup (that is, diagnosis takes time logarithmic in the time that would be required if only the most detailed level of description were available). In the generate-and-test framework of KARDIO (Bratko, Mozetic, and Lavrac 1989), a simulation model of a heart is used to both propose multiple-disorder diagnoses from symptoms and verify that the sets of disorders account for all symptoms. Models ordered with respect to abstraction level are used in the traditional fashion: Diagnoses at the most abstract level are found first and limit the proposal and verification of diagnoses at the next, more concrete level. Mozetic (1990) further formalizes general conditions on adjacent levels of abstraction and the use of partial evaluation to derive the abstract models from the more concrete. For example, in a digital circuit setting, one would begin with a description of the behavior of the components of a Boolean gate using real-valued voltages and currents and use definitions of concepts such as “high” and “low” to derive an abstract behavioral description. Exponential speedup was demonstrated in empirical tests on a semiautomatically constructed four-level heart model with 943 leaf diagnoses.

### Broader Formulations of Diagnosis

Much AI diagnosis research has focused on computing the plausible diagnoses for a fixed set of symptoms. However, human diagnosticians work within a richer context. In particular, they often perform diagnosis in contexts where therapeutic actions can and should be taken even before deciding on a final diagnosis.

*Therapy* can be defined as an interleaved process of using both diagnosis and repair to suppress undesired

symptoms. In a logic-based framework, it is possible to distinguish between components that are necessary in any diagnosis as opposed to those that are possible or irrelevant (Friedrich, Gottlob, and Nejd1 990). Diagnosis (defined as abduction on theories consisting of horn clauses describing the behavior of connected components) is known to be NP-complete (Bylander et al. 1991). However, finding the components that are necessarily part of a diagnosis is computable in polynomial time, even though determining whether a hypothesis is included in a minimal diagnosis remains NP-complete. Thus, a therapeutic procedure can run in polynomial time, although it might perform unnecessary repairs.

The emphasis on finding therapeutic actions is particularly relevant when working under real-time constraints. The REACT system (Ash et al. 1990) assumes a single underlying disorder in the setting of an intensive care unit, and a solution is characterized by having an appropriate action ready at a deadline, in this case, the point at which intervention must be performed to sustain the life of the patient. REACT represents the set of alternative diagnoses as a subset-superset lattice of diagnoses and moves downward through this lattice in response to accumulating evidence. In REACT, each therapeutic action has a fixed utility, and the mapping from every competing set of alternative diagnoses to actions is made computationally trivial. As a result, REACT always has some action that it can recommend, albeit possibly suboptimal, given any stage of the diagnosis process.

### Summary

Dx-90 was a successor to the Workshop on Model-Based Diagnosis held in Paris in July 1989. The broadening of the title and emphasis was motivated by an explicit desire to foster interaction among different communities interested in diagnosis. The 50 participants at Dx-90 were virtually unanimous in their support for another workshop to be held in 1991. As a result, CISE Technologie Innovative hosted Dx-91 on 14–16 October 1991 in Milan. Future workshops will be held annually, alternating between North America and Europe. The diversity of topics addressed at Dx-90 (and within the AI diagnosis community generally) illustrates the

continued vigor of the area and its potential for producing further nuggets for AI in general. At the moment, three of the themes discussed previously seem to present the greatest potential for such progress.

First, diagnosis applications have demonstrated the value of synthesizing probabilistic and categorical knowledge. Incorporating probabilistic information into a logically formulated problem yields computational benefits; exploiting knowledge about localized causal relationships organizes and constrains representations of uncertainty. As this synergy develops further, it might form the basis of a new knowledge system architecture that is likely to supplant other architectures that are grounded strictly in one paradigm or the other.

Second, diagnosis demonstrates the centrality of modeling decisions and the importance of ontological commitment in programs that aspire to solve realistic problems. Diagnosing engineered artifacts with dynamic and continuous behavior provides a vivid illustration: Some issues of (abstract) importance within the logic-based and probabilistic paradigms virtually fade to insignificance in comparison to the computational consequences of modeling decisions such as making a steady-state assumption for a continuous system, representing the behavior of a dynamic system with temporal abstractions, or considering structural failures that violate the design hierarchy of the target device. Diagnosis is a good task with which to study the interaction between such representational commitments and program performance and motivates improvements in techniques for representing all types of complex physical systems.

Third, diagnostic problem solving presents a realistic challenge and demonstration vehicle for computational theories of nonmonotonic reasoning. The diagnosis of engineered artifacts is of particular relevance. Not only are assumptions about the failures in the device defeasible, but even the correct behavior of any engineered system of realistic complexity can (and often must) be reasoned about under many simplifying assumptions, all defeasible in the face of malfunctions. A flexible architecture that exercises meaningful control over modeling assumptions while it performs diagnosis would be a significant achievement not only for diagnosis but also for AI in general.

## Acknowledgments

For the preparation of this report, Max Henrion, Eric Horvitz, and Olivier Raiman shared their impressions of the workshop and clarified details of their presentations; Peter Patel-Schneider and Ramesh Patil suggested helpful revisions; and Renée Wolfe assisted with manuscript preparation. Price Waterhouse, AAAI, ACM SIGART, Tom Gruber, and Bob Englemore provided administrative support for Dx-90.

## References

- Ash, D.; Vina, A.; Seiver, A.; and Hayes-Roth, B. 1990. Action-Oriented Diagnosis under Real-Time Constraints, Technical Report, KSL-90-39, Knowledge Systems Laboratory, Stanford Univ.
- Bratko, I.; Mozetic, I.; and Lavrac, N. 1989. *KARDIO: A Study in Deep and Qualitative Knowledge for Expert Systems*. Cambridge, Mass.: MIT Press.
- Brown, J. S.; Burton, R.; and de Kleer, J. 1982. Pedagogical, Natural Language, and Knowledge Engineering Issues in SOPHIE I, II, and III. In *Intelligent Tutoring Systems*, eds. D. Sleeman and J. S. Brown, 227–282. New York: Academic.
- Bylander, T.; Allemang, D.; Tanner, M. C.; and Josephson, J. R. 1991. Some Results Concerning the Computational Complexity of Abduction. *Artificial Intelligence* 49(1–3): 25–60.
- Clark, K. 1978. Negation as Failure. In *Logic and Data Bases*, eds. H. Gallaire and J. Minker, 293–322. New York: Plenum. Also in 1987. *Readings in Nonmonotonic Reasoning*, ed. M. L. Ginsberg, 311–325. San Mateo, Calif.: Morgan Kaufmann.
- Console, L., and Torasso, P. 1990. Integrating Models of the Correct Behavior into Abductive Diagnosis. In *Proceedings of the Ninth European Conference on Artificial Intelligence*, ed. L. C. Aiello, 160–166. London: Pitman.
- Console, L.; Dupré, D. T.; and Torasso, P. 1989. A Theory of Diagnosis for Incomplete Causal Models. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, 1311–1317. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.
- Cooper, G. F. 1990. The Computational Complexity of Probabilistic Inference on Bayesian Belief Networks. *Artificial Intelligence* 42(2–3): 353–405.
- Dague, P.; Jehl, O.; and Taillibert, P. 1990. An Interval Propagation and Conflict Recognition Engine for Diagnosing Continuous Dynamic Systems. In *Expert Systems in Engineering: Lecture Notes in Artificial Intelligence 462*, eds. G. Gottlob and W. Nejdl, 16–31. New York: Springer-Verlag.
- Davis, R. 1984. Diagnostic Reasoning Based on Structure and Behavior. *Artificial*

*Intelligence* 24(1): 347–410. Also in 1985. *Qualitative Reasoning about Physical Systems*, ed. D. Bobrow, 347–410. Cambridge, Mass.: MIT Press.

Davis, R., and Hamscher, W. C. 1988. Model-Based Reasoning: Troubleshooting. In *Exploring Artificial Intelligence: Survey Talks from the National Conferences on Artificial Intelligence*, ed. H. E. Shrobe, 297–346. San Mateo, Calif.: Morgan Kaufmann. Also in 1990. *Artificial Intelligence and MIT: Expanding Frontiers*, eds. P. H. Winston and S. A. Shellard, 380–429. Cambridge, Mass.: MIT Press.

Davis, R., and Shrobe, H. 1983. Representing the Structure and Behavior of Digital Hardware. *IEEE Computer* 16(10): 75–82.

de Kleer, J. 1991. Focusing on Probable Diagnoses. In *Proceedings of the Ninth National Conference on Artificial Intelligence*, 842–848. Menlo Park, Calif.: American Association for Artificial Intelligence.

de Kleer, J. 1986. An Assumption-Based TMS. *Artificial Intelligence* 28(2): 127–162. Also in 1987. *Readings in Nonmonotonic Reasoning*, ed. M. L. Ginsberg, 280–297. San Mateo, Calif.: Morgan Kaufmann.

de Kleer, J., and Williams, B. C. 1989. Diagnosis with Behavioral Modes. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, 1324–1330. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

de Kleer, J., and Williams, B. C. 1987. Diagnosing Multiple Faults. *Artificial Intelligence* 32(1): 97–130. Also in 1987. *Readings in Nonmonotonic Reasoning*, ed. M. L. Ginsberg, 372–388. San Mateo, Calif.: Morgan Kaufmann.

de Kleer, J.; Mackworth, A.; and Reiter, R. 1990. Characterizing Diagnoses. In *Proceedings of the Eighth National Conference on Artificial Intelligence*, 324–330. Menlo Park, Calif.: American Association for Artificial Intelligence.

Dressler, O. 1990. Diagnoses as Coherent Assumption Sets. In *Expert Systems in Engineering: Lecture Notes in Artificial Intelligence 462*, 47–52. eds. G. Gottlob and W. Nejdl. New York: Springer-Verlag.

Dvorak, D., and Kuipers, B. J. 1989. Model-Based Monitoring of Dynamic Systems. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, 1238–1243. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Eshghi, K. 1990. Diagnoses as Stable Models. In *Working Notes of the First International Workshop on Principles of Diagnosis*, ed. W. C. Hamscher, 39–48, Technical Report 14, Price Waterhouse Technology Centre, Menlo Park, California.

Friedrich, G.; Gottlob, G.; and Nejdl, W. 1990. Hypothesis Classification, Abductive Diagnosis, and Therapy. In *Expert Systems in Engineering: Lecture Notes in Artificial Intelligence 462*, eds. G. Gottlob

- and W. Nejdil, 69–78. New York: Springer-Verlag.
- Gelfond, M., and Lifschitz, V. 1988. The Stable Model Semantics for Logic Programming. In *Proceedings of the Fifth International Conference on Logic Programming*, 1070–1080. Cambridge, Mass.: MIT Press.
- Genesereth, M. 1984. The Use of Design Descriptions in Automated Diagnosis. *Artificial Intelligence* 24(1): 411–436. Also in 1985. *Qualitative Reasoning about Physical Systems*, ed. D. Bobrow, 411–436. Cambridge, Mass.: MIT Press.
- Ginsberg, M. L. 1987. *Readings in Nonmonotonic Reasoning*. San Mateo, Calif.: Morgan Kaufmann.
- Glass, B. J.; Erickson, W. K.; and Swanson, K. J. 1991. TEXSYS: A Large-Scale Demonstration of Model-Based Real-Time Control of a Space Station Subsystem. In *Proceedings of the Seventh Conference on Artificial Intelligence Applications*. Washington, D.C.: IEEE Computer Society.
- Gordon, J., and Shortliffe, E. H. 1985. A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space. *Artificial Intelligence* 26(3): 323–357.
- Guckenbiehl, T., and Schäfer-Richter, G. 1990. SIDIA: Extending Prediction-Based Diagnosis to Dynamic Models. In *Working Notes of the First International Workshop on Principles of Diagnosis*, ed. W. C. Hamscher, 74–82, Technical Report 14, Price Waterhouse Technology Centre, Menlo Park, California.
- Hamscher, W. C. 1991a. Model-Based Financial Data Interpretation. In *Proceedings of the First International Conference on Artificial Intelligence Applications on Wall Street*. Washington, D.C.: IEEE Computer Society.
- Hamscher, W. C. 1991b. Modeling Digital Circuits for Troubleshooting. *Artificial Intelligence* (Special Issue on Qualitative Physics) 51(1–3). Forthcoming.
- Hamscher, W. C. 1988. Model-Based Troubleshooting of Digital Systems, Technical Report 1074, Artificial Intelligence Lab, Massachusetts Institute of Technology.
- Hamscher, W. C., and Davis, R. 1984. Diagnosing Circuits with State: An Inherently Underconstrained Problem. In *Proceedings of the Fourth National Conference on Artificial Intelligence*, 142–147. Menlo Park, Calif.: American Association for Artificial Intelligence.
- Harmon, P. 1988. Expert Systems in Use. In *The Rise of the Expert Company*, E. A. Feigenbaum, P. McCorduck, and H. P. Nii, 273–316. New York: Times.
- Henrion, M. 1991. Search-Based Methods to Bound Diagnostic Probabilities in Very Large Belief Nets. In *Uncertainty and Artificial Intelligence: Proceedings of the Seventh Conference*, eds. B. D. D'Ambrosio, P. Smets, and P. P. Bonissone, 142–150. San Mateo, Calif.: Morgan Kaufmann.
- Holtzblatt, L. J.; Neiberg, M. J.; and Piazza, R. L. 1991. Temporal Reasoning in an Assumption-Based Reason Maintenance System, Technical Report M91-22, The MITRE Corporation, Bedford, Massachusetts.
- Kuipers, B. J. 1986. Qualitative Simulation. *Artificial Intelligence* 29(3): 289–338. Also in 1990. *Readings in Qualitative Reasoning about Physical Systems*, eds. D. Weld and J. de Kleer, 236–260. San Mateo, Calif.: Morgan Kaufmann.
- Mozetic, I. 1990. Reduction of Diagnostic Complexity through Model Abstractions, Technical Report TR-90-10, Austrian Research Institute for Artificial Intelligence, Vienna, Austria.
- Ng, H. T. 1990. Model-Based, Multiple-Fault Diagnosis of Time-Varying, Continuous Physical Devices. In *Proceedings of the Sixth Conference on Artificial Intelligence Applications*, 9–15. Washington, D.C.: IEEE Computer Society.
- Oyeleye, O. O.; Finch, F. E.; and Kramer, M. A. 1990. Qualitative Modeling and Fault Diagnosis of Dynamic Processes by MIDAS. *Chemical Engineering Communications* 96:205–228.
- Patil, R. S.; Szolovitz, P. S.; and Schwartz, W. 1981. Causal Understanding of Patient Illness in Medical Diagnosis. In *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, 893–899. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.
- Pearl, J. 1986. Fusion, Propagation, and Structuring in Belief Networks. *Artificial Intelligence* 29(3): 241–288.
- Peng, Y., and Reggia, J. 1989. A Connectionist Model for Diagnostic Problem Solving. *IEEE Transactions on Systems, Man, and Cybernetics* 19:285–298.
- Poole, D. 1989. Normality and Faults in Logic-Based Diagnosis. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, 1304–1310. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.
- Pople, H. E. 1982. Heuristic Methods for Imposing Structure on Ill-Structured Problems: The Structuring of Medical Diagnostics. In *Artificial Intelligence in Medicine*, ed. P. Szolovits, 119–190. Boulder, Colo.: Westview.
- Pople, H. E. 1977. The Formation of Composite Hypotheses in Diagnostic Problem Solving: An Exercise in Synthetic Reasoning. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, 1030–1037. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.
- Preist, C., and Welham, R. 1990. Modelling Bridge Faults for Diagnosis in Electronic Circuits. In *Working Notes of the First International Workshop on Principles of Diagnosis*, ed. W. C. Hamscher, 69–73, Technical Report 14, Price Waterhouse Technology Centre, Menlo Park, California.
- Raiman, O. 1990. A Circumscribed Diagnosis Engine. In *Expert Systems in Engineering: Lecture Notes in Artificial Intelligence 462*, eds. G. Gottlob and W. Nejdil, 90–101. New York: Springer-Verlag.
- Raiman, O. 1989. Elements of a Theory of Diagnosis. Ph.D. diss., Dept. of Computer Science, University of Paris 6.
- Raiman, O.; de Kleer, J.; Saraswat, V.; and Shirley, M. H. 1991. Characterizing Non-intermittent Faults. In *Proceedings of the Ninth National Conference on Artificial Intelligence*, 849–854. Menlo Park, Calif.: American Association for Artificial Intelligence.
- Reggia, J. A.; Nau, D. S.; and Wang, P. 1983. Diagnostic Expert Systems Based on a Set Covering Model. *International Journal of Man-Machine Studies* 19(5): 437–460.
- Reggia, J. A.; Peng, Y.; and Tuhim, S. 1990. The Role of Connectionist Methods in Abductive Diagnostic Problem Solving. In *Working Notes of the First International Workshop on Principles of Diagnosis*, ed. W. C. Hamscher, 83–91, Technical Report 14, Price Waterhouse Technology Centre, Menlo Park, California.
- Reiter, R. 1987. A Theory of Diagnosis from First Principles. *Artificial Intelligence* 32(1): 57–96. Also in 1987. *Readings in Nonmonotonic Reasoning*, ed. M. L. Ginsberg, 352–371. San Mateo, Calif.: Morgan Kaufmann.
- Rose, P., and Kramer, M. 1991. Qualitative Analysis of Causal Feedback. In *Proceedings of the Ninth National Conference on Artificial Intelligence*, 817–823. Menlo Park, Calif.: American Association for Artificial Intelligence.
- Saraswat, V.; de Kleer, J.; and Raiman, O. 1990. Contributions to a Theory of Diagnosis. In *Working Notes of the First International Workshop on Principles of Diagnosis*, ed. W. C. Hamscher, 33–38, Technical Report 14, Price Waterhouse Technology Centre, Menlo Park, California.
- Shoham, Y. 1987. *Reasoning about Change*. Cambridge, Mass.: MIT Press.
- Shortliffe, E. H. 1976. *MYCIN: Computer-Based Consultations in Medical Therapeutics*. New York: American Elsevier.
- Shwe, M.; Blackford, M.; Heckerman, D. E.; Henrion, M.; Horvitz, E. J.; Lehman, H.; and Cooper, G. F. 1990. Probabilistic Diagnosis Using a Reformulation of the INTERNIST-1/QMR Knowledge Base: II. Evaluation of Diagnostic Performance, Technical Report, KSL-90-68, Knowledge Systems Laboratory, Stanford Univ.
- Struss, P., and Dressler, O. 1989. Physical Negation: Integrating Fault Models into the General Diagnostic Engine. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, 1318–1323. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.
- Sussman, G. J., and Steele, G. L. 1980.

(continued on page 37)

choose to become. They might, in fact, learn to communicate in several natural languages. These remarkable phenomena of language learning have amazed most of us at one time or another, and it is only natural that we have tried to use computers to study or even duplicate them—with only partial success to report at this date.

The AAAI Spring Symposium on Machine Learning of Natural Language and Ontology (MLNLO) provided an opportunity to get together and discuss the partial successes and the research challenges that lie ahead. It was a rare opportunity because the work has tended to be reported in fragments, a thesis here or there, a paper at an AI or computational linguistics conference, another at a psychology or linguistics or child language conference or in a philosophy journal. The field is naturally highly multidisciplinary, and the interested researchers all speak their own languages—not just natural languages but specialized disciplinary dialects laden with the theoretical constructs and assumptions of each discipline. Thus, this symposium provided a forum for useful interchange of ideas.

“Learning of natural language” is a simple-sounding phrase that covers a number of phenomena. On the one hand, there are various aspects of language to be learned, such as the sounds that are significant to a particular language (phonology), words (lexicon) and their variations (morphology), the structure of meaningful utterances (syntax), and meaning and its relation to the lexicon and syntactic structure (semantics). On the other hand, there are the different components of learning: inducing the data to be learned from raw linguistic and nonlinguistic data; somehow codifying these data into an internalized, structured system that can be used in an automatic manner; and generalizing to be able to deal with new input never heard before and produce new output never uttered before. The learning of *ontology*, the understanding of what exists in the world, is closely linked with the learning of language.

At the symposium, 50 participants discussed contributions in all these areas, with 20 full-length presentations and a similar number of “advertising spots” that allowed virtually all groups some air time. It should also be mentioned that a parallel symposium focused on connectionist natural language processing (CNLP) and

that nontraditional computing has clearly exerted its influence on the field of language learning. Not only were a number of applications of connectionist and genetic techniques presented, but a joint final session was held with CLNP.

However, the main efforts are still closely linked to contemporary AI and linguistic theory. The field is beginning to attack various practical applications in areas where the knowledge is rich enough to allow modest learning, and it is providing increasing challenge and support to psycholinguistics and linguistics research. In this respect, participants spent time focused on a number of special topics, such as the extent to which language mechanisms are language specific and linguistic properties are innate, the conditions under which it is formally possible to learn a language, the recognition of ungrammatical sentences, the development of the ability to use metaphors, the modeling of second-language learning, and the question of how lexical symbols become grounded in reality. The treatments presented complemented cognitive theory with computational implementation.

At the end of the symposium, we took time before the joint CNLP panel to review the value of the symposium and look to the future. It was resolved that we instigate a regular program of MLNLO events; a newsletter; resource sharing (software, texts, and so on); and further symposia, workshops, and conferences. The first such event was a one-day workshop on natural language learning to be held at the 1991 International Joint Conferences on Artificial Intelligence in Sydney, Australia, on August 25. As befitted its shorter length, this workshop had a tighter focus, with a major goal being an analysis of proposed language-learning models to allow comparing and contrasting of the theoretical perspectives and the hypotheses embodied; the implementation techniques and learning algorithms; and the implications of the virtues, failings, and results of particular implementations and modeling experiments.

The symposium participants also felt that the working notes of the MLNLO symposium were a landmark volume worthy of further distribution. Thus, the working notes will immediately be made available (through the German AI Institute in

Kaiserslautern [DFKI D-91-09]) to a wider audience in the form of a technical report and an edited book. These publications will allow the expanded presentation of selected papers and, perhaps, additional invited papers from some who could not attend. Information can be obtained from powers@informatik.uni-kl.de or from reeker@ida.org.

David Powers

University of Kaiserslautern

Larry Reeker

Institute of Defense Analyses

---

## Principles of Diagnosis

(continued from page 23)

Constraints: A Language for Expressing Almost-Hierarchical Descriptions. *Artificial Intelligence* 14(1): 1–40.

Szolovits, P., and Pauker, S. G. 1978. Categorical and Probabilistic Reasoning in Medical Diagnosis. *Artificial Intelligence* 11:115–144.

Weld, D. S., and de Kleer, J. 1990. *Readings in Qualitative Reasoning about Physical Systems*. San Mateo, Calif.: Morgan Kaufmann.

Williams, B. C. 1986. Doing Time: Putting Qualitative Reasoning on Firmer Ground. In *Proceedings of the Fifth National Conference on Artificial Intelligence*, 105–112. Menlo Park, Calif.: American Association for Artificial Intelligence. Also in 1990. *Readings in Qualitative Reasoning about Physical Systems*, eds. D. Weld and J. de Kleer, 353–360. San Mateo, Calif.: Morgan Kaufmann.

Wu, T. D. 1991. Domain Structure and the Complexity of Diagnostic Problem Solving. In *Proceedings of the Ninth National Conference on Artificial Intelligence*, 855–861. Menlo Park, Calif.: American Association for Artificial Intelligence.

Wu, T. D. 1990. Efficient Diagnosis of Multiple Disorders Based on a Symptom Clustering Approach. In *Proceedings of the Eighth National Conference on Artificial Intelligence*, 357–364. Menlo Park, Calif.: American Association for Artificial Intelligence.

Zadrozny, W. 1990. The Logic of Abduction. In *Working Notes of the First International Workshop on Principles of Diagnosis*, ed. W. C. Hamscher, 8–17, Technical Report 14, Price Waterhouse Technology Centre, Menlo Park, California.



**Walter Hamscher**, who organized and chaired Dx-90, earned his Ph.D. in artificial intelligence from the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology in August 1988. He is currently a research manager in the Systems Research Group at Price Waterhouse Technology Centre, where he and his colleagues are developing a model-based approach to a variety of tasks in auditing and financial analysis.