Assumption Management in Qualitative Medical Diagnosis

Keith Downing  
Jeff Shrager

University of Oregon  
Eugene, Oregon  
97403

Xerox PARC  
3333 Coyote Hill Rd.  
Palo Alto, CA  
94304

June 27, 1988

Abstract

The Assumption-Based Diagnostician (ABD) project demonstrates the applicability of formal techniques of assumption-based reasoning to qualitative diagnosis, a task complicated by many types of assumptions. Our primary goal is to connect the medical-knowledge level to the logical level by automating the conversion of high-level diagnostic concepts such as "symptoms" and "faults" into the constructs of the Assumption-Based Truth Maintenance System (ATMS). Due to the simplicity of the ATMS logic, this compilation requires the formalization of many of the implicit assumptions of diagnostic reasoning. When the proper assumptions are included, the results of interpretation construction correspond to all diagnostic explanations consistent with the fault and symptom information. To combat the plethora of possible but implausible diagnoses, meta-assumptions put additional constraints upon explanation generation and provide an operational characterization of hierarchical diagnosis.
1 Introduction

This paper presents a method of qualitative medical diagnosis based on assumption management. The Assumption-Based Diagnostician (ABD) links the knowledge level to the logical level by compiling medical concepts into propositional logic, where ATMS (de Kleer, 1986a,b) interpretation construction is used to generate diagnostic explanations.

The motivation for an assumption-based approach arises from three important processes of medical diagnosis:

1. Causal Abduction - If the domain knowledge is encoded in rules having causes as antecedents and effects as consequents, then the process of constructing an explanation of symptoms in terms of causes involves backward reasoning (abduction) over those causal rules.

2. Theory Management - A typical diagnostic cycle consists of observation followed by the formation of explanatory theories, which are then tested and accepted, rejected, or revised.

3. Qualitative Reasoning - Medical experts explain causality mainly in terms of qualitative derivatives and ordinal relationships between quantities (Campbell, 1983; Kuipers and Kassirer, 1984). For example, a common description of the causal pathway through which sodium intake affects blood pressure is:

   Increased sodium intake results in a stronger concentration of plasma sodium, which raises the oncotic pressure between the interstitial (between cell) spaces and the plasma — causing the osmosis of water into the bloodstream. The ensuing elevation of blood volume then initiates a rise in cardiac blood output, which raises the blood pressure (Abstracted from Guyton (1986)).

All three of these processes involve inexact reasoning, which often incurs non-monotonic reasoning in that assumptions are made and later retracted. In abduction and theory formation, causal explanations represent defeasible assumptions, because in the absence of closed-world assumptions (e.g., “High salt intake and nervous tension are the only possible causes of high blood pressure.”) such explanations cannot be logically deduced from the observed effects. Furthermore, the local nature of constraint propagation combines with the non-determinism of qualitative arithmetic to create considerable ambiguity in qualitative reasoning about physical systems — thereby requiring auxiliary assumptions, such as the confluence heuristics of de Kleer and Brown (1985). Qualitative reasoning therefore adds yet another degree of imprecision to the inherently non-monotonic process of diagnosis - making assumption management a vital aspect of qualitative diagnosis.

In the ABD methodology, some of the assumptions of qualitative diagnostic reasoning (along with the knowledge of faults, symptoms, and their causal interactions) are expressed in the propositional calculus of the ATMS. The ABD system converts high-level causal relationships between faults and symptoms into an ATMS causal network. Then, when provided with a set of symptom premises, the ATMS performs interpretation construction, which returns all possible collections of fault and auxiliary assumptions capable of deriving those symptoms via the causal justification structure. These possible diagnoses could then be evaluated relative to criteria such as the likelihood of the fault assumptions alone or in combination with others, or the feasibility of the auxiliary assumptions.

2 Formalizing Diagnostic Problems in Biological Domains

In many electrical and mechanical domains, diagnosis can be aptly described as “reasoning from unusual behaviors (symptoms) to structural abnormalities (faults)”. The pairing of faults with structural defects seems accurate in these domains, because many faults involve major structural aberrations such as shorted resistors, disconnected wires, or dead motors. Fortunately for living organisms, most ailments do not involve major structural damage; totally severed arteries, non-functioning organs and similar traumas occur less frequently than shifts of equilibrium conditions to “illness” states. Hence, medical diagnosis attempts to explain observed changes in certain symptomatic parameters by changes in other parameters that could cause those symptoms. The “fault” is often some primary causal parameter whose proper modification (via therapy) will relieve the problem.

Still, the tenuous dichotomy between symptoms and faults has no absolute basis; it only exists relative to factors such as the available therapies, the region of the body under diagnosis, the granularity of diagnostic reasoning, the observability of changes (i.e. symptoms are visible; faults are not), and the relevant forms of causal interaction. For example, a cardiologist, who deals primarily with hydraulic and neuro-electrical forms of causation, would view a clogged artery as a fault that contributes to the symptom of high blood pressure; on the other hand, a dietician,
whose main concern is the chemical composition of the blood, would describe the same blocked artery as a symptom of an excessive cholesterol concentration.

Any formalization of diagnostic reasoning requires some commitment to a specific symptom-fault division. In previous work (Downing, 1987), this was achieved through the implicit use of "Causally-Bounded Diagnosis" (CBD). CBD relativizes symptoms and faults to a "causal border" defined by some mechanism of causal interaction. For instance, relative to a hydraulic causal mechanism, clogged arteries are faults, while elevated blood pressure is a symptom, since only the latter results from hydraulic activity and is a true effect from the hydraulic standpoint (Figure 1).

CBD involves two steps:

1. Causal Circumscription - Given a network consisting of different types of causal relationships between system parameters, choose one mechanism of causal interaction (e.g., hydraulic, electrical, or kinematic). Use this "relevant" mechanism to circumscribe some portion of the total causal network so that all parameters within the circumscription are involved in at least one causal relationship of the selected type. Diagnosis will take place only within the circumscription; hence diagnosis will involve only the circumscribed parameters and the relevant causal relationships between them.

2. Symptom-Fault Relativization - Define fault parameters as those parameters not causally influenced by other circumscribed parameters. Faults are thus independent and static relative to the circumscription since their only influences come from uncircumscribed parameters or "irrelevant" causal mechanisms. The dependent or symptomatic parameters are those causally affected by at least one circumscribed parameter.

As an example of CBD, consider Figure 1, which includes the causal pathway from sodium intake to blood-pressure change. The arc labels "+" and "−" designate direct and inverse causal influences, respectively, and are further annotated by the type of causal interaction. Choosing the hydraulic activities of the circulatory loop as the relevant causal mechanism yields the indicated circumscription (dashed box). Relativization designates total blood volume (BV), venous compliance (VC), and arterial resistance (AR) as fault parameters, while cardiac output (CO) and arterial pressure (AP) represent symptomatic parameters. Hence the three single-fault hypotheses that explain ∂AP+ are ∂BV+, ∂VC−, and ∂AR+.

3 The ATMS as an Explanation Generator

To generate explanations for symptoms within a CBD model, the Assumption-Based Diagnostician employs ATMS interpretation construction. As a brief review, the ATMS stores data in nodes, which are connected by justifications. The justification records how a node was derived from other nodes. ABD uses three types of nodes: premises are true in all contexts - they require no justifications. Assumptions are self-justifying, and are therefore defeasible premises. Conclusions are justified by other premises, assumptions, or conclusions.

A collection of non-contradictory assumptions constitutes an environment. A context denotes the deductive closure (via justifications) of an environment. A nogood is a minimal set of contradictory assumptions - that is, one with no contradictory subsets. An interpretation is a maximal environment - that is, one with no consistent supersets. Each interpretation spans a single context, but one context may be derived from many different interpretations. The crucial observation is that nogoods constrain interpretation generation, since no environment can subsume a nogood.

The following simple example illustrates these concepts:

\[
\begin{align*}
A \Rightarrow e \\
e, B \Rightarrow f & \quad \text{nogoods: (A,C), (C,D)} \\
C \Rightarrow e & \quad \text{interpretations: (A,B,D), (B,C)} \\
D \Rightarrow d & \quad \text{contexts: (A,B,D,e,f,d), (B,C,e)} \\
c, d \Rightarrow \bot \\
e, c \Rightarrow \bot
\end{align*}
\]

1Assumptions are capitalized, conclusions appear in lower-case, and justifications are implications with conjunctions as antecedents. ⊥ denotes "false", a contradiction.
An ABD "explanation" (similar to Poole's (1986) default-logic explanations, which are based on extension membership) for a collection of nodes is an interpretation whose context contains those nodes. Hence,

\[ \text{explanation}(\{f\}) := \{(A, B, D)\} \]

(1)

As illustrated in the next section, a group of one or more nodes may have many ABD explanations.

4 Causally-Bounded Diagnosis with the ATMS

A CBD problem can utilize the ATMS explanation capabilities once the CBD network has been translated into the proper collection of ATMS constructs. Since simple causality seems easily represented by justifications, the translation of CBD models into ATMS structures appears to be trivial. However, the modelling of multiple interacting faults requires a complex network of assumptions, conclusions, nogoods, and justifications. Otherwise, interpretation construction would return an explanation set that is incomplete and/or unsound. Medical diagnosis requires these complex ATMS structures, because fault interactions occur quite frequently in self-regulating systems like the human body as compensatory mechanisms respond to one fault by creating an opposing one. For example, in response to the increased blood pressure incurred by raised blood volume, the sympathetic nervous system increases venous compliance and decreases heart rate; both changes then lower blood pressure.

At one level, the mapping between CBD and ATMS is straightforward. Since independent/fault parameters have no causes within the circumscription, they parallel ATMS assumptions, which represent "deductively independent" datum. Similarly, since symptoms are causally affected by other circumscribed parameters, they are appropriately characterized by ATMS conclusions, which are justified by other nodes. For both fault and symptom parameters, the salient property is the sign of its derivative. The mutually-contradictory modes \( \delta X^+, \delta X^-, \) and \( \delta X_0 \) sufficiently represent all qualitative first-order states of the parameter \( X \): increasing, decreasing and steady, respectively. Justifications between the modes of different parameters reflect causal influences, for instance, if \( X \) directly influences \( Y \), then the justifications:

\[ \delta X^+ \Rightarrow \delta Y^+, \delta X^- \Rightarrow \delta Y^- \]

capture that relationship. Similarly, if \( X \) inversely influences \( Y \), then these justifications are added:

\[ \delta X^+ \Rightarrow \delta Y^-, \delta X^- \Rightarrow \delta Y^+ \]

(3)

The complexity of the CBD-ATMS translation arises in the effort to account for interacting faults. Although fault parameters require only the three mode assumptions mentioned above, each symptomatic parameter requires a three-tiered ATMS representation (Figure 2). In this model, faults cannot directly affect the qualitative derivatives of symptomatic parameters. Instead, they exert influences upon those parameters. Two or more opposing influences are then resolved to derive a single observed symptom. This three-layered representation enhances the modularity of the ATMS model. Consequently, new fault parameters, symptom parameters, and causal relationships can be incrementally added without modifying any of the pre-existing ATMS justifications. 3

In the ATMS network of Figure 2, the fault parameters, BV and VC, affect CO through the "Influence \( \delta CO \)" nodes. If they exert opposing influences (e.g., when \( \delta BV^+ \) and \( \delta VC^- \) both hold), then the two influence nodes must combine with one of the three resolvent assumptions to derive the appropriate observation. Otherwise, if CO has a single influence (e.g., "Influence \( \delta CO^+ \)"), then that influence must combine with the assumption that no opposing influences exist ("Not Influence \( \delta CO^- \)") to justify the appropriate observation ("Observe \( \delta CO^+ \)"). When passed to the interpretation constructor, the ATMS network above yields five explanations for elevated cardiac output:

\[ \text{explanation}(\{\delta CO^+\}) = \begin{cases} 
(\delta BV^+, \delta VC^-, \neg INF \delta CO^-) \\
(\delta BV^-, \delta VC^+, \neg INF \delta CO^-) \\
(\delta BV^+, \delta VC^-, \neg INF \delta CO^-) \\
(\delta BV^+-, \delta VC^-, RES \delta CO^+) \\
(\delta BV^-, \delta VC^+, RES \delta CO^+) 
\end{cases} \]

(4)

Notice that the "maximal" property of interpretations forces the inclusion of the irrelevant assumption \( D \) in the explanation of \( f \). The ATMS defaults as many assumptions as possible during interpretation construction – yielding many overspecific explanations. We are exploring an approach to this problem in current research.

3For a more detailed analysis and justification of this representation, see Downing and Shrager (1988).
The physiology behind these explanations is quite simple: if total blood volume increases, then the heart will naturally pump more blood per minute (i.e. cardiac output rises). If venous compliance increases, then the veins (acting as capacitors) hold more blood, thus delaying its return to the heart and decreasing cardiac output. Hence, increases in both BV and VC exert opposing influences upon CO. In the model of Figure 2, any explanation of high cardiac output must include $\partial BV^+$ and/or $\partial VC^-$, along with the appropriate auxiliary assumptions (i.e. the "not influence" and "resolvent" assumptions). The ATMS network thus combines physiological knowledge with the logic of causal interaction to render a complete set of possible explanations for $\partial CO^+$.

5 Meta-Assumptions

The ambiguities of qualitative arithmetic and the simplicity of propositional logic lead to many implausible (but logically possible) ABD diagnostic explanations (e.g., the hypothesis that all independent parameters are faulted.). Additional physiological or diagnostic knowledge could provide a bias to prune these implausible explanations; but still, the generate-and-test scenario of interpretation construction followed by knowledge-directed pruning results in an "intermediate interpretation bulge" (Forbus, 1986) (i.e., the intermediate set of interpretations is much larger than the filtered set), since all possible interpretations consistent with the basic causal circumscription are initially generated. For instance, in a CBD environment in which n fault parameters affect a symptomatic parameter Y, there are at least $3^n - 2^n$ possible explanations for any change in Y. Consequently, when $n = 10$, interpretation construction yields over 50,000 possible explanations!

One promising remedy for the intermediate interpretation bulge is the compilation of global or "meta" constraints into the basic causal network. By providing additional constraint on interpretation construction, this compilation effectively moves tests into the generator and reduces the number of implausible (yet possible) diagnoses.

In The Assumption-Based Diagnostician (ABD), global constraints are encoded as meta-assumptions such as the single-fault assumption and assumptions that rank independent parameters according to the strength of their influence upon a given dependent parameter (e.g., blood volume influences cardiac output more strongly than does venous compliance). Meta-assumptions are compiled into ATMS nogoods, which contain existing assumptions plus a single new assumption designating the global constraint. These nogoods restrict interpretation construction without burdening it with extra assumptions, and they lend non-monotonicity to meta-assumptions⁴. Thus, nogood compilation moves the external-knowledge filter into the explanation generator.

Meta-assumption compilation begins with a restricted predicate-calculus description of the constraint, which is negated and simplified to create a nogood schema. This serves as a template for the generation of all "violators" (i.e., minimal assumption combinations that violate the meta-assumption). The addition of the meta-constraint's newly-created ATMS assumption to each violator produces a set of nogoods that will enforce the meta-constraint during interpretation construction.

The following example illustrates the operationalization of the "influence-domination" meta-assumption:

Cardiac Output (CO) is more strongly affected by changes in blood volume (BV) than by changes in venous compliance (VC).

To rephrase the constraint in terms of relationships indigenous to the causal network of Figure 2:

When BV and VC exert opposing influences upon CO, and no other faults influence CO in the same direction that VC does; then the resolvent qualitative influence upon CO will equal BV's qualitative influence.

This has the following logical description:

\[
\text{influence}(BV, CO) \neq \text{influence}(VC, CO) \land \text{influence}(BV, CO) \neq 0 \land \text{influence}(VC, CO) \neq 0 \land \not\exists P : \text{influence}(P, CO) = \text{influence}(VC, CO) \\
\Rightarrow \text{resolvent}(CO) = \text{influence}(BV, CO)
\]

⁴Declared "true", a new meta-assumption becomes a part of every valid interpretation; hence, the interpretation constructor need never consider it. When not in effect, the meta-assumption is marked "false" and ignored. Non-monotonicity comes from the fact that when the meta-assumption is false, it will not appear in any interpretations. Thus, none of its containing nogoods will be subsets of any interpretations and therefore will not help prune them.
This is negated and simplified to form a nogood schema:

\[
\begin{align*}
\text{influence}(B, C) \neq \text{influence}(V, C) \land \text{influence}(B, C) \neq 0 \land \text{influence}(V, C) \neq 0 \\
\exists P \supseteq \text{influence}(P, C) = \text{influence}(V, C) \\
\wedge \text{resolvent}(C) \neq \text{influence}(B, C)
\end{align*}
\] (6)

Grounding this schema in the assumptions of the ATMS causal network results in a set of four nogoods, each of which contains the influence-dominance assumption, \(\text{dominates}(B, V, C, C)\):

\[
\begin{align*}
\{ \text{Dominates}(B, V, C, C), \text{d}B+, \text{d}V+, \text{RES } \text{d}C0- \} \\
\{ \text{Dominates}(B, V, C, C), \text{d}B+, \text{d}V+, \text{RES } \text{d}C00 \} \\
\{ \text{Dominates}(B, V, C, C), \text{d}B-, \text{d}V-, \text{RES } \text{d}C0+ \} \\
\{ \text{Dominates}(B, V, C, C), \text{d}B-, \text{d}V-, \text{RES } \text{d}C00 \}
\end{align*}
\] (7)

By declaring \(\text{Dominates}(B, V, C, C)\) true (and thereby forcing every interpretation to contain it), the interpretation constructor is restrained by the new nogoods and generates a smaller set of explanations for \(\theta C0+\):

\[
\begin{align*}
\text{explanation}(\theta C0+) = \{ \text{Dominates}(B, V, C, C), \text{d}B+, \text{d}V0, \text{RES } \text{d}C0- \} \\
\{ \text{Dominates}(B, V, C, C), \text{d}B0, \text{d}V-, \text{RES } \text{d}C0- \} \\
\{ \text{Dominates}(B, V, C, C), \text{d}B+, \text{d}V-, \text{RES } \text{d}C0+ \} \\
\{ \text{Dominates}(B, V, C, C), \text{d}B-, \text{d}V-, \text{RES } \text{d}C00 \}
\end{align*}
\] (8)

The addition of the Single-Fault Assumption (SFA) results in four more nogoods:

\[
\begin{align*}
\{ \text{SFA}, \text{d}B+, \text{d}V+ \} \\
\{ \text{SFA}, \text{d}B+, \text{d}V- \} \\
\{ \text{SFA}, \text{d}B-, \text{d}V+ \} \\
\{ \text{SFA}, \text{d}B-, \text{d}V- \}
\end{align*}
\] (9)

These further restrict interpretation construction to two explanations:

\[
\begin{align*}
\text{explanation}(\theta C0+) = \{ \text{SFA}, \text{Dominates}(B, V, C, C), \text{d}B+, \text{d}V0, \text{RES } \text{d}C0- \} \\
\{ \text{SFA}, \text{Dominates}(B, V, C, C), \text{d}B0, \text{d}V-, \text{RES } \text{d}C0- \}
\end{align*}
\] (10)

Meta-assumptions characterize a hierarchical approach to diagnosis wherein higher levels of abstraction contain more meta-assumptions (and hence are more highly constrained by nogoods) than lower regions. Thus, ABD provides an operational formalism for a hierarchy of diagnostic assumptions similar to Davis's (1985) "ordered categories of failure". However, Davis's hierarchy organizes multiple representations as defined by alternate mechanisms of causal interaction. In contrast, Causally-Bounded Diagnosis defines the basic diagnostic environment relative to the causal circumscription carved out by a single causal mechanism. Meta-assumptions then constrain behaviors (and hence explanations) within that circumscription. So while Davis's assumptions enable the expansion of qualitative diagnosis from one to many representations, ABD portrays the assumptions that underlie reasoning within a single qualitative representation.

In summary, by restricting interpretation construction, meta-assumptions compensate for the extreme generality of the basic ATMS causal network, which otherwise leads to an overabundance of possible diagnoses. These defeasible constraints define a hierarchical approach to diagnosis in which the higher abstraction levels embody stricter notions of "plausibility" than do the lower, less-constrained levels. Hence, meta-assumptions provide a structured means of moving explanation testing into the generator - thereby alleviating the intermediate interpretation bulge and making assumption-based diagnosis more manageable.

6 The Assumption-Based Diagnostician

The Assumption-Based Diagnostician (ABD)\(^5\) works as follows:

\(^5\)ABD was developed in Xerox Commonlisp and runs in conjunction with de Kleer's ATMS.
1. The user enters the symptomatic and fault parameters along with the qualitative influences that relate them. This defines the diagnostic environment.

2. ABD compiles the parameters and influences into an ATMS causal network by introducing the ATMS assumptions, nodes, justifications, and contradictions that define the logical semantics of the high-level concepts.

3. At any time after step 2, the user can enter meta-assumptions, which ABD compiles into nogoods. Meta-assumptions can be deactivated and reactivated when desired.

4. The user enters specific symptomatic observations to define a particular diagnostic problem.

5. ABD biases the causal network by converting to premises those nodes corresponding to the symptomatic observations.

6. Finally, ABD runs ATMS interpretation construction over the biased causal network, generating all possible explanations in terms of fault, auxiliary and meta assumptions.

To fully explore the explanation space, the user may repeat steps 3 through 6 as many times as desired with different observations and/or meta-assumptions.

7 Discussion

De Kleer and Brown (1983) detail the role of implicit assumptions in modeling physical systems. They indicate that both novices and experts employ assumptions to direct reasoning. However, experts have explicated their assumptions and thus recognize the limitations of their models while also understanding how to transform those models by adjusting certain crucial assumptions. Assumption management is thus an integral part of expert reasoning. In addition, the non-monotonicity of diagnosis and the ambiguity of qualitative reasoning suggest that assumption management should play a major part in automated qualitative diagnosis.

The Assumption-Based Diagnostician uses the ATMS to formalize some of the tacit assumptions of medical diagnosis. De Kleer and Williams (1987) have demonstrated the ATMS’s utility for handling the non-monotonicity of hypothesis generation and testing during diagnosis; but, they only discuss ATMS assumptions representing beliefs that specific components are functioning properly. Many other types of assumptions participate in the diagnostic process but previously lacked appropriate logical semantics. ABD formalizes some of these assumptions and illustrates their participation in diagnostic explanation. These new assumptions fill many of the gaps between the diagnostic-knowledge level and the logical level, but they put huge demands on the interpretation constructor, which then generates a plethora of possible but implausible diagnoses. ABD meta-assumptions reduce this explanation explosion while providing a logic-based operational model for hierarchical diagnosis.

In conclusion, the exploitation of ATMS interpretation construction to generate a complete and sound set of diagnostic explanations necessitates the formalization of certain implicit assumptions of qualitative reasoning. The ensuing interpretation bulge motivates the need for meta-constraints and the concomitant mechanism for hierarchical diagnosis. Thus, the Assumption-Based Diagnostician unites three important lines of research by grounding qualitative reasoning and hierarchical diagnosis in the ATMS. In short, ABD links “causes to clauses” to illustrate the role of assumption management in medical diagnosis.

Acknowledgments.

We owe thanks to Johan de Kleer for clearing up some of our ambiguities and assumptions concerning the ATMS, qualitative physics and this paper. Also, I (Keith) am grateful to Nils Peterson for his lively and enlightening discussions of human physiology. Finally, Sally Douglas has provided advice, support and general guidance throughout my perpetual graduate career.

References.


13. Poole, David L., *Default Reasoning and Diagnosis as Theory Formation*, University of Waterloo Tech. Rept. CS-86-08, March 1986.