More on Goal-Directed Diagnosis

Ron Rymon
University of Pennsylvania

Follow this and additional works at: http://repository.upenn.edu/cis_reports

Recommended Citation


This paper is posted at ScholarlyCommons. http://repository.upenn.edu/cis_reports/296
For more information, please contact repository@pobox.upenn.edu.
More on Goal-Directed Diagnosis

Abstract
In many diagnosis-and-repair domains, diagnostic reasoning cannot be abstracted from repair actions, nor from actions necessary to obtain diagnostic information. We call these exploratory-corrective domains. In TraumAID 2.0, a consultation system for multiple trauma management, we have developed and implemented a framework for reasoning in such domains which integrates diagnostic reasoning with planning and action. In this paper, we present Goal-Directed Diagnosis (GDD), the diagnostic reasoning component of this framework. Taking the view that a diagnosis is only worthwhile to the extent that it can affect subsequent decisions, GDD focuses on the formation of appropriate goals for its complementary planner.

Comments
More On Goal-Directed Diagnosis

MS-CIS-92-81
LINC LAB 239

Ron Rymon

University of Pennsylvania
School of Engineering and Applied Science
Computer and Information Science Department
Philadelphia, PA 19104-6389

November 1992
More On Goal-Directed Diagnosis

Ron Rymon
Department of Computer and Information Science
University of Pennsylvania
Room 419C, 3401 Walnut Street, Philadelphia, PA 19104
Tel: (215) 898-0326, Fax: (215) 898-0587
E-mail: rymon@linc.cis.upenn.edu

In Proceedings DX-92, October 1992

Abstract

In many diagnosis-and-repair domains, diagnostic reasoning cannot be abstracted from repair actions, nor from actions necessary to obtain diagnostic information. We call these exploratory-corrective domains. In TraumAID 2.0, a consultation system for multiple trauma management, we have developed and implemented a framework for reasoning in such domains which integrates diagnostic reasoning with planning and action. In this paper, we present Goal-Directed Diagnosis (GDD), the diagnostic reasoning component of this framework. Taking the view that a diagnosis is only worthwhile to the extent that it can affect subsequent decisions, GDD focuses on the formation of appropriate goals for its complementary planner.

1 Prologue

In many domains, it is common to distinguish reasoning and activity concerned with what problems need be addressed from that reasoning concerned with how to address those problems. As such, Artificial Intelligence (AI) subsumes as separate sub-disciplines diagnosis research, seeking the source (or sources) of a system’s faulty behavior, and planning research, concerned with the construction of action plans to achieve certain goals. Based on that dichotomy, most diagnostic programs take a diagnosis as their objective.

In last year’s workshop, in a paper titled “Towards Goal-Directed Diagnosis”, we argued this may be inadequate. In trauma management, for example, therapy is the ultimate objective and diagnosis is the “price” that one has to pay in order to achieve that objective. In such domains, diagnosis should only persist so long as it can affect those decisions for which it was carried out in the first place. We called that the Goal-Directed Diagnosis (GDD) principle.

![Figure 1: Basic Cycle of Reasoning, Planning and Action](image)

In addition, we showed that in systems where activity is necessary for both diagnostic and therapeutic purposes, it is advantageous to integrate diagnostic reasoning and planning capabilities. An architecture called Exploratory-Corrective Management (ECM) was proposed, in which a repeated basic cycle of diagnostic reasoning, planning and action is used (Figure 1). The ECM architecture was shown to satisfy the following desiderata for exploratory-corrective domains:

1. It allows interleaving diagnosis and repair.
2. It positions the diagnostic reasoner to
   (a) set diagnostic and therapeutic goals;
   (b) use incoming evidence to monitor the actions and/or other events;
   (c) reason about changes in knowledge or state to adapt current goals.
3. It positions the planner to mediate between concurrent diagnostic and therapeutic needs.

TraumAID 2.0 [Webber et al 92] is a consultation system for the diagnosis and treatment of multiple trauma which implements the ECM architecture. A formalization of its goal-directed diagnostic reasoning is the focus of this paper, whereas its complementary planning paradigm, called Progressive Horizon Planning (PHP), is described elsewhere [Rymon et al 92]. Each consultation session in TraumAID 2.0 consists of several cycles of reasoning, planning and acting.
The paper is structured as follows: Section 2 reviews related work. Section 3 begins the introduction of GDD with a formal rule-based language and a simple inference from observations to conclusions and goals. Diagnostic problems are defined in this language and their solutions correspond to the closure of that inference. Taking a higher level perspective on the role of GDD reasoning within a complete ECM session, Section 4 begins with a meta-level algorithmic description of the ECM algorithm, continues with a short discussion of issues pertaining to the actual implementation of diagnostic and therapeutic strategies in such systems, and concludes with an illustrative example of a complete session. Section 5 extends the paradigm to effectively deal with contradictory information.

2 Related Work

2.1 Formal Diagnosis

Recent years have seen significant advances in formal approaches to diagnosis. A large number of approaches and frameworks have been suggested: probabilistic classifiers and discriminators, logical consistency-based and abductive paradigms, graph-based formulations in which causal and coincidental relations are modeled, etc. Some of the frameworks combine and/or unify two or more paradigms. However, all these these formalizations take diagnosis — broadly defined as a characterization of the current state of affairs — as their solution and goal. Our Goal-Directed Diagnostic paradigm takes a different view of diagnosis and its objectives.

To understand this view, consider that general theories of diagnosis give rise to a (sometimes large) number of hypothetical characterizations that are consistent with, or that explain the observed failure. In such instances, it is necessary to choose among such possibilities. An important observation made by [Poole & Provan 90] is that the optimality of a diagnosis must depend on post-diagnosis goals. To that end, [Provan & Poole 91] advocate the use of utilities in order to choose among different potential diagnoses. [Poole & Provan 91] take that approach one step further, realizing that there is often no need for a complete explanation and that the granularity of a solution depends on its uses, and also on available tests.

Realistically, diagnosis is rarely an independent process: more often it serves the purpose of another process, e.g. repair. The GDD paradigm, presented here, is part of a total approach for reasoning in exploratory-corrective domains which combines diagnostic reasoning, planning and action [Rymon 93]. In that framework, the main purpose of diagnostic reasoning is generating proper goals; diagnosis purely for characterization’s sake becomes a secondary concern.

[Friedrich et al 91] share much of that insight. Considering diagnosis as part of an overall diagnosis-and-repair process, they correctly observe that repair does not always require a complete diagnostic explanation. Unlike other formalizations, their theory has no explicit notion of a diagnosis. Instead, a sequence of tests and repair actions is sought, that if applied to the current state, will imply (as in a logical proof) a restoration of the diagnosed system to a proper working condition. Presented not as a theory of diagnosis but as a theory of repair planning, their work applies a possible-models planning approach [Winslett 88] to a diagnostic domain. The link between diagnosis and repair planning is also emphasized by [Pepper & Kahn 87]. [Rushby & Crow 91] formalize reconfiguration, a limited form of repair, using an extension of Reiter’s theory of diagnosis [Reiter 87].

In domains such as trauma management, planning and action are necessary for both diagnosis and repair. At the same time, diagnostic reasoning and activity are necessary to (a) set goals for, (b) monitor the execution of, and (c) verify the actual results of actions.

Current work on sequential diagnosis has often taken a simplistic view of information acquisition. While the potential, or expected, discriminatory power of a given piece of information for the problem at hand is considered, the activity necessary to obtain it is not (i.e. either only simple questions are considered, or a simple cost function is attached to every piece of information). Of course, researchers have realized that this approach is inadequate:

“One limitation of DART in its present form is that it doesn’t take into account the cost of tests.... Unfortunately, the evaluation of test cost can be quite difficult because of dependencies on sequencing and grouping.” [Genesereth 84]

Planning researchers, as early as STRIPS, have studied these issues extensively. The ECM architecture thus uses its planning capabilities not only for repair, but also for information acquisition. It also uses planning to mediate between diagnostic and therapeutic activities when multiple problems concurrently require both [Rymon 93]. That strong tie between diagnosis and therapy is common in medical practice and is evident in many AI medical management systems. What we add here is a formalization of the relationship between the two.

2.2 AMORD

AMORD [deKlerk et al 77] is a general purpose problem solver which has accompanying truth maintenance and planning facilities. The main thesis behind AMORD is that combinatorially explosive reasoning can often be avoided if a problem solver reasons explicitly about its own reasoning. Explicitly encoded reasoning control is shown effective in avoiding irrelevant, or useless inferences. TraumAID’s GDD reasoner shares this intuition.

\footnote{In an article in the Encyclopedia of Artificial Intelligence, p. 589, referring to MYCIN, Shortliffe and Rennels write: “Although MYCIN is often described as a diagnostic program, its principal motivation was therapy planning”}
The key difference between the two paradigms is their distinct objectives: while AMORD’s objective is to avoid irrelevant or useless inferences, the objective of GDD is to avoid unimportant, and possibly dangerous, actions. Although GDD can also be viewed as a general purpose reasoning scheme, the result of the distinct objectives is that even though both systems compute beliefs and goals, AMORD focuses on the former, whereas GDD focuses on the latter. A more detailed technical comparison between AMORD and the MVL-based inference system used to implement GDD can be found in [Rymon 93].

3 A Formulation of Goal-Directed Diagnosis

Goal-Directed Diagnosis (GDD) begins with the point of view that diagnosis is only worthwhile if it has the potential to affect future decisions. Thus, while we accept the common definition of a diagnosis as a case characterization, we believe that different purposes can lead to different characterizations of the same situation. For example, different purposes may lead to different refinement efforts. GDD allows explicit encoding of purposes, which it uses to guide its problem solving. More specifically, throughout a problem solving session, the GDD reasoner will maintain both a belief – a description of the current characterization, and an attitude – encoding a sense of purpose by pointing to goals worth pursuing.

In a recurrent cycle, the GDD reasoner takes as input a diagnostic problem, characterized by (1) observations; and (2) mappings (rules) from from observations to conclusions (belief), and from observations and conclusions to goals (attitude). A solution to such problem is a new attitude-belief assignment. Goals, propositions regarded as relevant by the current attitude, are the addressed by the accompanying planner and served by the actor. New observations result in a modified diagnostic problem, and a new cycle is initiated.

In this section, we describe a rule-based language for specifying diagnostic problems in GDD a corresponding inference scheme.

3.1 Underlying Framework: Multi-Valued Logics

Multi-Valued Logics (MVL) [Ginsberg 88] is a formal framework for inference in which each proposition is assigned not only a truth value, corresponding to the strength of belief in that proposition being true or false, but also a knowledge assessment, measuring roughly the amount of knowledge used to derive such belief. Bilattices, in which one partial order corresponds to the truthfulness measure and the other to the knowledge one, are used by MVL as domains for truth-value assignment.

Our goal-directed diagnostic paradigm (GDD) was first formalized using a three-valued logic: true, false and unknown [Rymon et al 91]. This paper presents an MVL-based reformulation of GDD. The immediate result of this reformulation is the ability to extend the inference paradigm to more expressive domains in a consistent fashion. Importantly, it allows explicit representation of contradictions (cf. section 5).

In this new formalization of GDD, each proposition is assigned a value drawn from the cross product of two bilattices: one representing belief, the other attitude. The notion of belief is interpreted regularly, whereas the attitude component represents problem-solving control information and measures the relevance of acquiring information about, or achieving the condition described by, the particular proposition. Importantly, each of these bilattices still maintains both truth and knowledge sub-components.

Within Ginsberg’s paradigm, domain knowledge is expressed by general first-order formulae, and thus requires an underlying theorem prover. In contrast, TraumAID uses a rule-based representation which is simpler to handle, but incomplete in general. Instead of reconstructing TraumAID’s knowledge, we chose to specialize Ginsberg’s theory to the rule-based case. While the material presented next is self-contained, the reader is referred to [Ginsberg 88] for a more complete coverage of MVL.

3.2 Attitude and Belief

During the diagnostic process, the GDD reasoner will maintain and update an attitude and a belief for propositional statements. To remain general, propositions may be any fact about the patient or the world that the reasoner may know to hold, may know not to hold, may assume, may want to know whether hold, may want to achieve, may be confused about, etc. The reasoner’s attitude towards and/or belief in a given proposition will change over time as a result of new information becoming available, new inferences drawn, activity carried out, etc.

Definition 3.1 A Lattice

A lattice is a triple (L, ∧, ∨), where L is its domain, ∧ ("meet") and ∨ ("join") are binary operations on L onto itself that are (1) idempotent, i.e. a ∧ a = a, a ∨ a = a; (2) commutative, i.e. a ∧ b = b ∧ a, a ∨ b = b ∨ a; (3) associative, i.e. (a ∧ b) ∧ c = a ∧ (b ∧ c), (a ∨ b) ∨ c = a ∨ (b ∨ c), and (4) obey the absorption laws, i.e. a ∧ (a ∨ b) = a, a ∨ (a ∧ b) = a.

Alternatively, a lattice can be defined as a partially ordered set, any two elements of which have a greatest lower bound (glb) and a least upper bound (lub). The two definitions coincide by taking lub(a, b) = a ∨ b, and glb(a, b) = a ∧ b. A lattice is said to be complete if lub and glb can be defined for any subset of the lattice’s elements. Thus, any lattice with a finite domain is complete.

Definition 3.2 A Bilattice

A bilattice is a sextuple (B, ∧, ∨, ⊕, ⊗, △) such that:

1. L₁ def = (B, ∧, ∨) and L₂ def = (B, ⊕, ⊗) are both complete lattices; and

2. →B → B is a mapping such that:
propositions.

Extensions to more complex bilattices are discussed in weakly grounded

if belief is restricted to \{T, F\}. We shall say that it is

relevant) and \(I\) (for irrelevant) in the attitude bilattice.

Technically, the extreme points in the belief bilattice are
called \(T\), and \(F\), whereas same points are called \(R\) (for

\[\text{Figure 2: Basic Truth-Knowledge Bilattice}\]

Ginsberg discusses bilattices that are based on two partial

orders: truth-wise \((\leq_t)\) and knowledge-wise \((\leq_k)\). That

is, each proposition is described by how strongly we believe it is true, and by how much knowledge was involved in inferring that belief. The smallest nontrivial bilattice (Figure 2) has four points: \(T\) (absolute truth), \(F\) (false), \(U\) (unknown), and \(\perp\) (contradictory). In that bilattice, the \(\leq_t\) partial order defines one lattice \((B, A, V)\), whereas the \(\leq_k\) defines another lattice \((B, \cdot, +)\). In any truth-knowledge bilattice, negation reverses the truth capacity of a proposition, leaving its knowledge capacity unchanged. In particular, within the basic bilattice, \(\neg\) maps \(T\) to \(F\) and vice versa, leaving \(U\) and \(\perp\) untouched.

Two, possibly distinct, bilattices are used in GDD as domains for attitude and belief respectively:

**Definition 3.3 Attitude and Belief**

Given a set of primitive propositions \(H\) \(\text{def} \{h_i\}_{i=1}^n\),

- an **attitude** maps \(H\) to an attitude bilattice \(B_A\)
- a **belief** maps \(H\) to a belief bilattice \(B_B\)
- an **attitude-belief** combines the two and maps \(H\) to

the cross product \(B_A \times B_B\). Conversely, it can also be viewed as a pair \((\phi_A, \phi_B)\) of attitude and belief.

For our purposes here, we shall take both \(B_A\) and \(B_B\) to be represented by the 4-point bilattice. The belief bilattice, following Ginsberg’s suggestion, is defined by the truth-knowledge partial orders. In the attitude bilattice, a proposition is described by its relevance \((\leq_r)\) and the knowledge used to derive that relevance \((\leq_k)\). Notably, one’s knowledge with respect to the truthfulness of a proposition need not equal, in general, one’s knowledge with respect to the relevance of that same proposition. Technically, the extreme points in the belief bilattice are called \(T\), and \(F\), whereas same points are called \(R\) (for relevant) and \(I\) (for irrelevant) in the attitude bilattice. Extensions to more complex bilattices are discussed in 

[Ginsberg 88, Rymon 93]

We shall say that an attitude-belief is **strongly grounded** if belief is restricted to \(\{T, F\}\). We shall say that it is **weakly grounded** if such restriction holds for all relevant propositions.

### 3.3 Goals

Goals are a **semantic** interpretation of an attitude-belief assignment to propositions. Generally speaking, a proposition \(p\) is a goal if its attitude assignment is high on the relevance partial order \((\leq_r)\). Of course, not every such proposition need be attempted; it may already be believed to hold (belief=\(T\)). In general, one has to define which combinations of relevance and achievement levels need be addressed, in what order, etc. Such considerations, however, are beyond the scope of, and do not affect the operation of our GDD reasoner. Within the ECM architecture, such determination is left to the planner.

A goal’s nature, in particular its characterization as diagnostic or therapeutic, plays an important role in deciding whether or not to address it. Taking a somewhat simplified view, a diagnostic goal, aimed at determination, may be regarded satisfied whenever the proposition is proved either true or false. A therapeutic goal, on the other hand, will only be satisfied with the actual achievement of the condition it describes (i.e. belief=\(T\)).

Rules used by GDD to express knowledge have their antecedents expressed solely in belief terms. However, in encoding diagnostic and therapeutic strategies in GDD, we found it useful to also use the relevance (or irrelevance) of a goal. Thus, a mapping from the attitude- to the belief-bilattice was added:

\[
\text{attitude}(a) \text{ def} \begin{cases} 
T & a = R \\
F & a = I \\
a & \text{otherwise}
\end{cases}
\]

That mapping models, roughly, the belief in relevance of the given proposition.

### 3.4 A Language for Representing Diagnostic Knowledge

We use Prolog-like rules to represent knowledge. Two types of rules are be used: one for inferring belief, the other for inferring goals.

**Definition 3.4 An Antecedent**

An **antecedent** is recursively defined as either:

1. a primitive proposition \(h\); or
2. \(\text{attitude}(h)\), where \(h\) is a proposition; or
3. \(\neg a\), where \(a\) is an antecedent; or
4. \(\text{true}(a)\), \(\text{false}(a)\), \(\text{unknown}(a)\), or \(\text{contradictory}(a)\), where \(a\) is an antecedent; or
5. \(\text{known}(a)\), \(\text{unless}(a)\), or \(\text{compatible-with}(a)\), where \(a\) is an antecedent.

**Definition 3.5 Rules**

A rule \(R\) has two parts: a **body**, or a premise, which is a set of antecedents; and a **header**, or consequent, which is a single proposition. GDD has two types of rules:
1. **Evidential rules affect belief.** They map evidence and lower-level conclusions to conclusions:

\[ \text{Ant}_1 \land \text{Ant}_2 \land \ldots \land \text{Ant}_r \Rightarrow d \]

For example, the following evidential rule concludes whether a patient’s shock is due to abdominal bleeding:

\[ \text{Shock} \land \neg \text{Single\_wound\_to\_upper\_chest} \land \text{unless(Pericardial\_Tamponade)} \land \text{unless(Massive\_Hemothorax)} \land \text{unless(Tension\_Pneumothorax)} \Rightarrow \text{Shock\_of\_possible\_abdominal\_origin} \]

2. **Goal Setting rules** affect attitude. They map evidence and conclusions to goals, or more precisely to attitude:

\[ \text{Ant}_1 \land \text{Ant}_2 \land \ldots \land \text{Ant}_r \Rightarrow g \]

For example, the following goal-setting rule concludes whether it is relevant to know whether a patient has hematuria.

\[ \text{Gunshot\_wound\_to\_abdomen} \land \text{Bullet\_in\_abdomen} \Rightarrow \text{Hematuria} \]

A fact may often be provable in several alternative ways. Similarly, a goal may need to be set in a variety of contexts. We therefore allow a conclusion to be made via a number of evidential rules, and a goal to be set via a number of goal-setting rules.

Conversely, we allow the same proposition to serve as the header of both goal-setting and evidential rules. In particular, as header to a goal-setting rule, a diagnostic or therapeutic goal means that the rule is used to conclude that the goal is worth adopting. An evidential rule whose header is that same goal is used to conclude whether or not it has been satisfied. Similarly, a goal-setting rule whose header is a clinical condition is used to conclude that it is relevant to investigate that condition. A similarly headed evidential rule is used to conclude whether or not the condition holds. Consider, for example, the diagnosis and repair of a **pericardial tamponade**. During that process, corresponding diagnostic and therapeutic goals are instantiated, addressed, and satisfied:

1. **Setting a Diagnostic (knowledge) goal:**
   \[ \ldots \Rightarrow \text{Pericardial\_Tamponade} \]
   "It is necessary to know if the problem exists".

2. **Satisfying a Diagnostic (knowledge) goal:**
   \[ \ldots \Rightarrow \text{Pericardial\_Tamponade} \]
   "Conclude that the problem exists".

3. Setting a Therapeutic goal:
   \[ \ldots \Rightarrow \text{Relieve\_pressure\_pericardial\_sac} \]
   "It is necessary to address the problem".

4. Satisfying a Therapeutic Goal:
   \[ \ldots \Rightarrow \text{Relieve\_pressure\_pericardial\_sac} \]
   "The problem has been successfully addressed".

3.5 **A Diagnostic Problem Instance**

**Definition 3.6 A Diagnostic Problem Instance**

An instance of a diagnostic problem is a quadruple \( P \triangleq (H, RB, M_0, OBS) \) such that:

- \( H = \{ h_1, h_2, \ldots, h_n \} \) is a set of propositions;
- \( RB \) is a set of evidential and goal-setting rules;
- \( M_0 \subseteq H \) is a set of observed manifestations (i.e., propositions for which we can, initially, assert that they are either true or false, with some degree of confidence);
- \( OBS : M_0 \rightarrow B_B \) is a belief (restricted to \( M_0 \)).

3.6 **Inference**

In the basic form of inference, to be introduced next, a closure of the observations given the evidential and goal-setting rules is computed. The basic inference is simpler and more intuitive than that in [Ginsberg 88]. In Section 5, we will have to extend it to a form that is closer to Ginsberg’s in order to support reasoning about contradictory information. In order to define the basic inferential closure, let us first define the semantic interpretation of rules.

**Definition 3.7 Belief Assignment for an Antecedent**

Let \( (\phi_A, \phi_B) \) be the current attitude-belief (defined on primitive propositions). For any antecedent \( (\text{Ant}) \), the belief in that antecedent, \( \phi_B(\text{Ant}) \), is defined as follows:

1. If \( \text{Ant} \) is a proposition \( h \), \( \phi_B(\text{Ant}) \triangleq \phi_B(h) \);
2. If \( \text{Ant} \) has the form \( \text{attitude} \), \( \phi_B(\text{Ant}) \triangleq \text{attitude}\(\phi_A(h)\)\);
3. If \( \text{Ant} \) has the form \( \neg \text{Ant}' \), where \( \text{Ant}' \) is an antecedent, \( \phi_B(\text{Ant}) \triangleq \neg \phi_B(\text{Ant}') \);
4. If \( \text{Ant} \) has the form \( f(\text{Ant}') \), where \( \text{Ant}' \) is an antecedent, then
   - If \( f \) is one of four belief predicates (true, false, unknown, contradictory), then
     \[ \phi_B(f(\text{Ant}')) \triangleq \begin{cases} T & \phi_B(\text{Ant}') = T \\ F & \text{otherwise} \end{cases} \]
     Similarly for known, unless, and compatible-with:
   - \( \phi_B(\text{known}(\text{Ant}')) \triangleq \begin{cases} F & \phi_B(\text{Ant}') = U \\ \bot & \phi_B(\text{Ant}') = \bot \\ T & \text{otherwise} \end{cases} \)

---

2 A pericardial tamponade is a condition where blood fills the pericardial sac, pressuring the heart and interfering with its operation.
• \( \phi_B^*(\text{unless}(\text{Ant}')) \) \( \overset{\text{def}}{=} \neg \phi_B^*(\text{Ant}') \lor \neg \phi_B^*(\text{known}(\text{Ant}')) \)

• \( \phi_B(\text{compatible} - \text{with}(\text{Ant}')) \) \( \overset{\text{def}}{=} \phi_B(\text{Ant}') \lor \neg \phi_B^*(\text{known}(\text{Ant}')) \)

Belief predicates will be useful where reasoning in absolute terms is necessary, since they map terms to \{T,F\}. Thus, \( h \) and \( \text{true}(h) \), are interpreted differently. The same is true of \( \neg h \) and \( \text{false}(h) \). Similarly, \( \neg \text{known}(h) \) refers to any situation in which a concrete belief cannot be reached whereas \( \text{unknown}(h) \) describes the particular such situation \( (U) \) resulting from lack of information. In the basic bilattice, another member of the former class is \( \perp \) which corresponds to contradictory information.

**Definition 3.8** Belief Assignment for a Rule’s Body

Let \( \langle \phi_A, \phi_B \rangle \) be an attitude-belief, \( R \) a rule with antecedents \( \{\text{Ant}_i\}_{i=1}^k \), then define

\[
\phi_B^*(\text{body}(R)) \overset{\text{def}}{=} \land_{i=1}^k \phi_B^t(\text{Ant}_i)
\]

**Definition 3.9** Consistent Inference (closure)

An attitude-belief \( \langle \phi_A, \phi_B \rangle \) is a consistent inference for a problem instance \( P \) iff

1. It coincides with OBS, i.e. \( \forall h \in M_0, \phi_B(h) = \text{OBS}(h) \);
2. For any proposition \( d \in H \) up to \( M_0 \), let \( \{R_i\}_{i=1}^l \) be all the evidential rules with \( d \) in their header, then

\[
\phi_B(d) = \lor_{i=1}^l \phi_B^t(\text{body}(R_i));
\]
3. Similarly, for any proposition \( d \in H \), let \( \{R_i\}_{i=1}^l \) be all the goal-setting rules with \( d \) in their header, then

\[
\phi_A(d) = \text{attitude}^{-1}(\lor_{i=1}^l \phi_B^t(\text{body}(R_i))); \]

Solving a diagnostic problem requires computing a consistent inference. While in general, there is no guarantee that such an inference is unique, computable, or even exists, the following algorithm is a straightforward, greedy attempt at computing it. It begins with the observations, assuming all other propositions to be unknown \( (U) \). Then, in a forward-chaining fashion, it repeatedly enforces Definition 3.9 until a fixed point (inferential closure) is reached.

**Algorithm 3.10** Computing an Inferential Closure

1. Start off with the observations, by setting

\[
\phi_A(h) \overset{\text{def}}{=} \text{U}, \text{for all } h \in H
\]

\[
\phi_B(h) \overset{\text{def}}{=} \begin{cases} \text{OBS}(h) & h \in M_0 \\ \text{U} & \text{otherwise} \end{cases}
\]

2. Mark all the rules (indicating they need to be fired).
3. Until none of the rules is marked, let \( R \) be such rule and let \( d \) be its header.
   - If \( R \) is an evidential rule, let \( \{R_i\}_{i=1}^l \) be the set of all evidential rules for \( d \), and let \( v \overset{\text{def}}{=} \lor_{i=1}^l \phi_B^t(\text{body}(R_i)) \). If \( \phi_B(d) \neq v \) then set \( \phi_B(d) = v \) and mark all rules in which \( d \) appears in the antecedents as needing to be fired.
   - Similarly, if \( R \) is a goal-setting rule, let \( \{R_i\}_{i=1}^l \) be the set of goal-setting rules for \( d \), and let \( v \overset{\text{def}}{=} \lor_{i=1}^l \phi_B^t(\text{body}(R_i)) \). If \( \text{attitude}(\phi_A(d)) \neq v \) then set \( \phi_A(d) = \text{attitude}^{-1}(v) \) and mark all rules in which \( d \) appears in the antecedents as needing to be fired.
   - Unmark \( R \).

4. Return \( \langle \phi_A, \phi_B \rangle \).

**3.7 Solving a Diagnostic Problem**

The formal definition of a solution in the GDD framework emphasizes its distinction from other diagnostic frameworks.

**Definition 3.11** A Diagnostic Explanation (Diagnosis)

Let \( \langle \phi_A, \phi_B \rangle \) be the inferential closure for a problem instance \( P \), then \( \phi_A \) is the diagnostic explanation (diagnosis) for \( P \). It is weakly complete if \( \phi_A, \phi_B \) is weakly grounded (i.e. no relevant propositions are unknown), and strongly complete if \( \phi_A, \phi_B \) is strongly grounded (i.e. all propositions are known).

Most formal diagnostic frameworks are content with a diagnosis as a solution. In GDD, we emphasize the importance of the goals, and consequently of the actions, adopted while arriving at a diagnosis.

A complete diagnostic session consists of several cycles through the ECM architecture. In a given diagnosis-and-repair scenario, if the GDD principle is adopted, it is unlikely that a strongly complete diagnosis be sought. In the case of a weakly complete diagnosis, while all diagnostic goals have been fulfilled, some of the therapeutic goals may still need attention (recall that a therapeutic goal is satisfied only when it is assigned a T value). The important point, however, is that the degree to which a diagnosis is refined is to a great extent a function of potential actions, and of their expected outcomes. In the ECM architecture, the termination criterion is thus part of the planner. Therefore the importance of the attitude part of the inferential closure: it defines the goals which will be pursued next. Goals are the “architectural duty” of the GDD reasoner.

**Definition 3.12** A Solution to a Diagnostic Problem

A solution to a diagnostic problem is the inferential closure, i.e. the pair \( \langle \phi_A, \phi_B \rangle \).

**4 The Diagnostic-Therapeutic Process**

Having defined a solution to a single diagnostic problem, it is important to understand the overall process, i.e. a complete diagnostic-therapeutic session, as captured by repeated cycles of the ECM architecture. We begin with a meta-level algorithmic description of that process. We
then discuss shortly issues pertaining to the implementation of diagnostic and therapeutic strategies within that architecture, using the GDD paradigm. An illustrative example of a complete diagnostic-therapeutic session concludes this section.

4.1 Integrating Diagnosis, Planning, and Action: The ECM Algorithm

Algorithm 4.1 calls the diagnostic reasoner whenever new information appears. The solution it provides is used to guide the complementary planner in the choice of activity which, in turn, may provide new information to start a new cycle.

Algorithm 4.1 ECM Diagnosis-and-Repair Algorithm

1. Initialize \( \langle \phi_A, \phi_B \rangle \) to coincide with OBS;
2. Compute an inferential closure for \( \langle \phi_A, \phi_B \rangle \);
3. Construct a plan \( P \) for the combination of goals indicated by \( \langle \phi_A, \phi_B \rangle \);
4. Unless \( P \) is empty do
   - Execute \( P \) until the first new piece of information \( (h,v) \) comes in;
   - Update \( \langle \phi_A, \phi_B \rangle \) to reflect \( (h,v) \);
   - Go to step 2;

Note that the criterion to terminate diagnosis is not necessarily related to the concreteness (groundness) of the current characterization (the diagnostic explanation, or diagnosis). Diagnosis terminates when the planner returns an empty plan, e.g. when all goals have been addressed or when no means are available for addressing remaining goals. Also note that any newly provided information, whether it has been called for or not, whether it is acquired via diagnostic or via therapeutic activity, will be used by the algorithm to refine its current diagnosis and possibly trigger new goals.

4.2 Goal Inhibition: using GDD to encode Strategies

To demonstrate its usefulness, we have used GDD to implement a number of diagnostic and therapeutic strategies drawn from the literature and from our experience in TraumAID. Two mechanisms that were useful in that task have already been presented: goal-setting rules, as a mean to drive goals, and planning, as a mean to resolve conflicts between competing goals. Goal inhibition, a third mechanism which we have found useful in implementing such strategies, is presented next. For lack of space, we will not present the strategies themselves; the interested reader is referred to [Rymon 93].

Trauma management presents an on-going interplay between diagnosis and therapy in which goals often have to be delayed or even ignored. Goal interaction is generally addressed by TraumAID’s planner, but since planning is extremely demanding computationally, much of the planning effort can be saved if unnecessary, non-contributing, goals are inhibited before being passed on to the planner.

The basic approach to goal inhibition is to qualify goal-setting rules by the negation of any inhibition condition. Consider, for example a patient with a right lumbar injury. Normally, such injury would suggest a possible duodenal injury:

\[
\text{Right lumbar wound} \Rightarrow \text{Duodenal injury}
\]

A standard test for duodenal injury is a CT scan. However, this lengthy and costly procedure should not be pursued independently if there is already a perceived need for a laparotomy. The surgical procedure will expose any duodenal injury, if present. Thus, one could qualify the above rule as follows:

\[
\text{Right lumbar wound} \land \neg \text{Laparotomy required} \Rightarrow \text{Duodenal injury}
\]

While correct, the problem with such inhibition scheme is that if the inhibition condition is complex, rules become even more complex and hard to maintain. In addition, overloading rules makes them less interpretable: the separate function of each of a rule’s antecedents becomes unclear (cf. [Clancey 83]). Instead, inhibiting relationships between goals and between goals and conclusions can be specified separately. For each goal \( g \), we can define an inhibition clause which will then be compiled (as a macro expansion) into all of that goal’s goal-setting rules.

Definition 4.2 Goal Inhibition Clause

Given a goal \( g \), \( \text{inhibit}(g) \) is a clause (or clauses) specifying the conditions under which \( g \) should not be pursued.

Technically, the macro expansion procedure for a goal inhibition clause creates a new internal proposition called \( \text{inhibit}(g) \) and a set of rules headed by that proposition that correspond to the condition itself. In the above example, the following new rule is added:

\[
\text{Laparotomy required} \Rightarrow \text{inhibit(Duodenal injury)}
\]

Then, we pad the body of every goal-setting rule for \( g \) as follows: \( \text{body}(R) \overset{\text{def}}{=} \text{body}(R) \land \neg \text{true(inhibit}(g)) \)

An alternative approach to goal inhibition is based on an extension of GDD to allow negative conclusions (goals) in rule headers (cf. Section 5). Note that so far, rule headers have simply been proposition names and have been interpreted positively, i.e. a rule has been viewed as an argument for its header-proposition. The formal interpretation of rules headed by a negated proposition is somewhat complex, and requires work. Informally, we can think of a rule headed by \( \neg p \) as an argument against \( p \). Such extension to our representation allows a more natural (and convenient) form of goal-inhibition: it can simply be viewed as a “reason” not to pursue that goal. Taking that view, it is natural to consider that particular
reason together with all other reasons for and against that goal’s pursuit. Goal-inhibition rules are used for that purpose. Let \( g \) be the said goal, we add a rule of the form: inhibit(\( g \)) \( \rightarrow \neg g \).

Under the framework discussed in Section 5, a contradictory attitude (\( \perp \)) will be assigned to a goal for which both a goal-setting and a goal-inhibition rule have succeeded. In our previous example, we will write:

\[
\text{Laparotomy_required} \rightarrow \neg \text{Duodenal_injury}
\]

Another way in which goals can be effectively inhibited, and which is sometimes preferable in implementing diagnostic strategies is to “force” a concrete value (i.e. \( T \) or \( F \)) on the goal’s underlying proposition. Such an approach will often be preferable when the real value of the inhibited goal can be inferred at the presence of the inhibiting condition; scaled diagnosis, discussed next, is a case in point. Unfortunately, that is not always possible: for example, nothing can be concluded about whether or not a patient has a duodenal injury in the above example. The net effect of that approach is that the goal remains relevant, but is regarded achieved by the planner and thus not addressed.

4.3 Example

To illustrate the effect of multiple diagnostic and therapeutic cycles in TraumAID 2.0’s ECM architecture, we will follow a case from the initial observations and the suspected diagnosis, to its validation, its treatment and the verification of its success.

Consider a patient presenting in the emergency room in stable condition, suffering a gun-shot wound to the left chest. A new diagnostic problem is instantiated with these observations. Let \( \langle \phi_A, \phi_B \rangle \) denote the system’s current attitude-belief. Initially \( \phi_A(h)=\phi_B(h)=U \), for all propositions \( h \in H \). As soon as the observations are reported, \( \phi_B \) is set accordingly.

In the next stage, the closure of \( \langle \phi_A, \phi_B \rangle \) is computed. In particular, we use the following goal-setting rule to set the diagnostic goal \( \text{Simple_Hemothorax(Left)} \), aimed at knowing whether or not the patient suffers a hemothorax\(^3\).

\[ (1) \quad \text{Chest_Wound(Left)} \vdash \text{Simple_Hemothorax(Left)} \]

At this point, if no other issues arise, control is transferred to the planner. A planner such as TraumAID’s current planner would recommend a Survey Chest X-Ray as a means of obtaining the desired information.

Suppose the physician orders an X-ray, and from the X-ray reports signs of hemothorax and a compound fracture to the left ribs. While the latter information had not been solicited, the system’s belief toward both propositions will be updated to \( T \). (If X-ray reports are assumed to be complete, beliefs about all other features of the X-ray will be updated to \( F \).) While each of these updated beliefs may trigger further investigation, for this example we will ignore all but the hemothorax finding. That finding triggers the evidential rule:

\[ (2) \quad \text{X-ray_shows.Simple_Hemothorax(Left)} \quad \Rightarrow \quad \text{Simple_Hemothorax(Left)} \]

The system thus changes its belief in the presence of hemothorax from \( U \) to \( T \). That change may be interpreted as a satisfaction of the diagnostic goal set by (1). Note too that we must distinguish a hemothorax finding from the condition of having a hemothorax, since the condition can be diagnosed in other ways, such as through the presence of decreased breath sounds, as in the following rule:

\[ (3) \quad \neg \text{Radiography_available} \land \\
\text{Decreased_breath_sounds(Left)} \quad \Rightarrow \quad \text{Simple_Hemothorax(Left)} \]

The presence of a hemothorax triggers the following goal-setting rule:

\[ (4) \quad \text{Simple_Hemothorax(Left)} \quad \vdash \quad \text{Rx.Simple_Hemothorax(Left)} \]

The attitude toward this therapeutic goal is thus updated from \( U \) to \( R \), and the goal is referred to the planner. A planner such as TraumAID’s would recommend addressing it through the insertion of a chest tube. Evidence that a chest tube has been inserted leads to a goal becoming relevant of ensuring its proper placement, and of checking that it is functioning correctly. In addition to these two cycles of reasoning and activity (a subsequent X-ray is required to check proper placement), the following rule is evaluated to check that the treatment goal for the simple hemothorax is actually satisfied:

\[ (5) \quad \neg \text{Chest_tube_is_draining_blood(Left)} \land \\
\text{Chest_tube_is_draining_blood(Left)} \quad \Rightarrow \quad \text{Rx.Simple_Hemothorax(Left)} \]

In summary, we have tracked the hemothorax from the initial wound report, through its suspicion as more investigation is recommended, continuing with the acquisition of more evidence that allows for concluding its presence and the need to address it, and finally, making sure that the treatment actually works.

5 Reasoning with Contradictory Information

5.1 Explicit Representation of Negation

Even though \( \perp \) is part of both our belief and attitude bilattices, it will never appear in practice given the basic inference paradigm. One may argue this to be a good feature of our inference system, since \( \perp \) will often lead to a host of problems (see [Ginsberg 88]). On the other
hand, \( \bot = T + F \), and can thus be used to represent the co-

presence of contradictory evidence. Explicit representa-

tion of contradictions provides a system with important

flexibility. Consider, for example, the diagnosis of a Tension

Pneumothorax\(^4\) (TP). Two alternative methods can

be used for that purpose:

1. \( \text{X-ray\_shows\_Tension\_Pneumothorax} \Rightarrow \text{Tension\_Pneumothorax} \)
2. \( \text{Needle\_aspiration\_chest(Positive)} \Rightarrow \text{Tension\_Pneumothorax} \)

Medically, positive indication in one of the tests is enough

for the diagnosis to be justifiably made. However, a physi-

cian will also assume lack of tension pneumothorax if any

of the two tests has come up negative. It is medically un-

justified to take the other test as well. Unfortunately, by

our semantic interpretation of the above rules, TP will be

tagged unknown (\( F \lor U \)) in such case. It was not proved,

but cannot be dismissed since not all the rules have failed.

To account for that description, we may choose to aug-

ment the above rules as follows:

1'. \( \text{X-ray\_shows\_Tension\_Pneumothorax} \land \)
\( \text{compatible\_with( Needle\_aspiration\_chest(Positive))} \Rightarrow \text{Tension\_Pneumothorax} \)

2'. \( \text{Needle\_aspiration\_chest(Positive)} \land \)
\( \text{compatible\_with(X\_ray\_shows\_Tension\_Pneumothorax)} \Rightarrow \text{Tension\_Pneumothorax} \)

The added antecedents will cause the failure of both rules

if any of the tests comes up negative. Otherwise, if any or

both tests come up positive, and there is no contradictory

evidence, TP is positively concluded.

We now consider the case in which tests provide con-

tradictory evidence. As so far presented, both rules will

fail and a negative TP will be assessed. While for some

propositions that may be a plausible solution, it is not the

medical interpretation in the case of a TP. In that case,

medical practice will default to positive treatment.

So far, in our semantics, a proposition was true or false if

it was so asserted or concluded. It was unknown if it was

not reported or could not be concluded. In other words,

unknown arose from lack of knowledge. Contradictions

belong to a different class of unknowns. To reason about

contradictions, we first extend the original rule representa-

tion to allow a header to be a negated proposition. Then,

we can simply add two more rules to the original ones:

3. \( \neg\text{X-ray\_shows\_Tension\_Pneumothorax} \Rightarrow \neg\text{Tension\_Pneumothorax} \)
4. \( \neg\text{Needle\_aspiration\_chest(Positive)} \Rightarrow \neg\text{Tension\_Pneumothorax} \)

\(\text{A tension pneumothorax is a condition in which air leaks into the chest cavity between the chest wall and the lung to the point beyond which the return of blood to the heart is at risk.}\)

The added rules model explicitly the consequences of neg-

ative test results. In addition, we redefine the inference

to a form which is closer to Ginsberg's and which allows

combining negative and positive contributions:

**Definition 5.1 Consistent Inference (redefined)**

An attitude-belief \( \langle \phi_A, \phi_B \rangle \) is a consistent inference for a

problem instance \( P \) iff

1. It coincides with OBS, i.e \( \forall h \in M_0 \phi_B(h) = \text{OBS}(h) \).
2. For any proposition \( d \in H - M_0 \), let \( \{ R_i \}_{i=1}^{k} \) be all the
evidential rules with \( d \) in their header, \( \{ R_i \}_{i=1}^{k} \) all the

evidential rules with \( \neg d \) in their header, then

\[
\phi_B(d) = \sum_{i=1}^{k} (\phi_B(body(R_i)) \lor U) + \\
+ \sum_{i=1}^{k} (-\phi_B(body(R_i)) \land U);
\]

3. Similarly, for any proposition \( d \in H \), let \( \{ R_i \}_{i=1}^{k} \) be all the
goal-setting rules with \( d \) in their header, \( \{ R_i \}_{i=1}^{k} \) all the
goal-setting rules with \( \neg d \) in their header, then

\[
\phi_A(d) = \text{attitude}^{-1}(\sum_{i=1}^{k} (\phi_B(body(R_i)) \lor U) + \\
+ \sum_{i=1}^{k} (-\phi_B(body(R_i)) \land U));
\]

Using the new inference procedure, given the above rules,

TP will be assigned a \( \bot \) belief, to indicate the contradic-

tory information with respect to its presence.

**5.2 Dealing with Contradictions**

Apart from their actual representation, two major concerns

need be addressed when one represents and reasons about

contradictory information. The first problem is a technical

one: under the above inference scheme, \( \bot \) propagates

itself. We call that problem pollution, and a few hints

as to how to deal with it can be found in [Rymon 93].

The second problem is the semantic interpretation of \( \bot \),
i.e. what to do with a proposition for which a \( \bot \) belief

(\(\text{attitude}^{-1}\)) is assigned?

Unfortunately, the answer here varies. In some cases, a

positive default is appropriate whereas in other circum-

stances a negative one is more plausible. In some circum-

stances, none of the conclusions can be safely made and

further investigation is warranted. In other cases a default
to a particular treatment (e.g. the TP case) is appropriate

whereas in yet other circumstances, a conclusive decision

is irrelevant to solving the problem.

The GDD principle provides us with a major escape by

stating that \( \bot \) is irrelevant to solving the problem.

We use the \( \bot \) to indicate that we are dealing with a

contradictory proposition. Then, \( \bot \) can be used to

encode the conclusion of a proposition which is not

safely made: 

\[ \text{contradictory(Tension\_Pneumothorax)} \Rightarrow R_x\_\text{Tension\_Pneumothorax} \]
6 Summary and Early Evaluation

We assume that diagnosis is only worthwhile to the extent that it can affect decisions, and so have presented a formalization of a goal-directed diagnostic paradigm (GDD). In contrast to other diagnostic paradigms, GDD defines a solution to a diagnostic problem not only in explanation terms, but also and more importantly in terms of the recommendations it implies.

GDD works within the exploratory-corrective management (ECM) architecture in which control cycles between a diagnostic reasoner, planner, and an actor. In this architecture, diagnosis is used to initiate both diagnostic and therapeutic goals, to adapt those goals to reflect new information, and to verify their actual achievement. Planning is used to address and mediate between concurrent diagnostic and therapeutic needs.

The new MVL-based formalization of GDD extends the original paradigm and allows reasoning with contradictory information and with defaults. A number of diagnostic and therapeutic strategies were shown to be naturally implementable in this framework [Rymon 93].

TraumAID [Webber et al 92] is a consultation system for multiple trauma management which has been developed over the past eight years as a collaboration between the Department of Computer and Information Science in the University of Pennsylvania and the Department of Surgery in the Medical College of Pennsylvania. Its new version, TraumAID 2.0, implements an ECM architecture in which diagnostic reasoning is complemented with a planning capability. TraumAID 2.0 has been successfully validated on 266 theoretical cases. Under a grant from AHCPR and with the help of a panel of national trauma management experts, we began evaluating its recommendations with respect to actual care provided by physicians in MCP, as well as to that of an earlier version.

The MVL-based formalization of GDD is new and extends the diagnostic reasoning in TraumAID 2.0. It has been implemented in prototype form and has been functionally tested on parts of TraumAID's knowledge-base.

Acknowledgement

Many of the ideas in this paper were shaped while implementing TraumAID 2.0 and as a result of discussions within the TraumAID research group. Particular thanks go to Bonnie Webber, John Clarke, Greg Provan and Charlie Ortiz. I would also like to thank Peter Szolovits for important comments and suggestions, Matt Ginsberg for helping me understand MVL, and anonymous reviewers.

References


