Psychiatric Diagnosis from the Viewpoint of Computational Logic

Joseph Gartner* Terrance Swift† Allen Tien‡ Carlos Viegas Damásio§
Luís Moniz Pereira¶

Abstract

While medical information systems have become common in the United States, present systems have mostly addressed clerical aspects of medicine such as billing, record management and similar tasks. Deeper problems, such as aiding the process of diagnosis, have largely remained unexplored for commercial systems. This is not surprising, since automating diagnosis requires considerable sophistication both in the understanding of psychiatric epidemeology and in knowledge representation techniques. This paper is an interdisciplinary study of how recent results in logic programming, non-monotonic reasoning, and knowledge representation can aid in psychiatric diagnosis. We argue that to logically represent psychiatric diagnosis as codified in the *Diagnostic and Statistical Manual of Mental Disorders, 4th edition* requires abduction over programs that include both explicit and non-stratified default negation, as well as dynamic preference rules. We show how such programs can be translated into abductive frameworks over normal logic programs and implemented using recently introduced logic programming techniques. Finally, we describe how such programs are used in a commercial product *Diagnostica*.

*Medicine Rules Inc, 25 East Loop Rd, Stony Broon, NY 11794 agartner@earthlink.net

†Department of Computer Science, SUNY at Stony Brook, Stony Brook, NY tswift@cs.sunysb.edu

‡Medical Decision Logic Inc 7921 Ruxway Rd, Baltimore MD 21204-3515 atien@jhmi.edu

§A.I. Centre, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2825-114 Caparica, Portugal. cd@di.fct.unl.pt

¶A.I. Centre, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2825-114 Caparica, Portugal. lmp@di.fct.unl.pt

1 Introduction

Medical information systems have become an active area of software development in the United States, with a market of over 10 billion dollars per year. Typically, these systems have as their goals either to cut the costs of medical treatment or to ensure that treatments are performed in a standard, well-documented manner. Traditional medical information systems address well-known problems such billing or shift-scheduling; or problems related to workflow management such as simple monitoring of treatment plans. However, important areas of medical practice, such as patient assessment (diagnosis) have not often been addressed by medical information systems, partly because the process of medical reasoning is difficult to automate. The purpose of the *Diagnostica* system developed by Medicine Rules, Inc (Stony Brook, NY) is twofold. As a research system, it explores how the process of psychiatric assessment can be represented by extensions of classical logic and forms a focus of an interdisciplinary collaboration between computer scientists and research psychiatrists. Just as importantly, as a commercially available product Diagnostica seeks to aid psychiatrists, psychologists and psychiatric social workers in diagnosing patients in an efficient and systematic manner.

Accurate diagnoses can be difficult to make, even for a trained psychiatrist. For instance a confused, elderly patient could suffer either from Alzheimer's Dementia or a Major Depressive Disorder (sometimes colloquially called *pseudo-dementia*). In the latter case, the patient may be treatable with medication; in the former case the patient may not be. Similarly, it may be difficult to determine whether a child has Attention Deficit Disorder (treated by medication) or an Adjustment Disorder (treated by therapy or by changing the child's environment). Diagnostic procedures concerning such disorders have been codified by the American Psychiatric Association in the fourth edition of its reference book *Diagnostic and Statistical Manual of Mental Disorders*, or *DSM-IV* [8], which is widely used in the United States. These procedures specify various *criteria* that a patient must satisfy in order to meet a diagnosis for a mental disorder. As an example, criteria for Asperger's Disorder, a Childhood Pervasive Development Disorder, is shown in Figure 1. As terminology we use the term *criterion* to specify both the conditions comprising a rule (e.g. criteria 1–5 in Figure 1) and the "symptoms" that the patient exhibits, e.g. criteria 1.a–1.d in Figure 1 which are sometimes called *base criteria*.

Criterion 1 reflects the *polythetic* nature of psychiatric diagnoses, in which there need be no essential characteristic or criterion of a diagnosis. Instead, multiple prototypes with varying features are used to group together a wide range of disparate phenomena into a diagnosis. At the same time there may be a significant amount of symptom overlap between different diagnoses. For instance, the failure to develop peer relationships can, under different circumstances, indicate schizophrenia, autism, and many other disorders. The issues of multiple prototypes and symptom overlaps leads to occasional difficulty and even ambiguity in distinguishing between the 618 DSM-IV diagnoses, as in the cases mentioned above. Because of these complications, while most American psychiatrists use DSM-IV, few use it to its full advantage. Studies have shown that clinical psychiatrists err in using DSM-IV by not considering all possible diagnoses, while research psychiatrists err by not excluding diagnoses quickly enough.

As indicated by Figure 1, DSM-IV diagnostic rules have a clear formulation that lends itself to formulation as a logic program: thus a patient meets criteria for a diagnosis if the body of the diagnosis, expressed as a logical rule, is satisfied. However, DSM-IV makes certain demands on knowledge representation, including the need to exclude certain diagnoses to prove other diagnoses, the need to represent incomplete knowledge, and the need for hypothetical reasoning during di-

1. Qualitative impairment in social interaction, as manifested by at least two of the following
   (a) marked impairment in the use of multiple nonverbal behaviors such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction;
   (b) failure to develop peer relationships appropriate to developmental level
   (c) a lack of spontaneous seeking to share enjoyment, interest,m or achievements with other people (e.g. by a lack of showing, bringing, or pointing out objects of interest to other people).
   (d) lack of social or emotional reciprocity
2. Restricted repetitive and stereotyped patterns of behavior, interests, and activities, as manifested by at least one of the following:
   (a) encompassing preoccupations with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus
   (b) apparently inflexible adherence to specific nonfunctional routines or rituals.
   (c) stereotyped and repetitive motor mannerisms (e.g., hand or finger flapping or twisting, or complex whole-body movements)
   (d) persistent preoccupation with parts of objects
3. The disturbance causes clinically significant impairment in social, occupational, or other important areas of functioning
4. There is no significant clinical delay in cognitive development or in the development of age-appropriate self-help skills, adaptive behavior (other than in social interaction) and curiosity about the environment in childhood.
5. Criteria are not met for another Pervasive Development Disorder or Schizophrenia

Figure 1: A Diagnostic Criterion for Asperger's Disorder

agnosis that require sophistication in their logical formulation and implementation. This paper explores how recently introduced techniques in logic programming and non-monotonic reasoning can be used to represent aspects of diagnosis as codified in DSM-IV. Specifically:

- We show that modeling DSM-IV requires non-stratified negation in order to handle ambiguities in diagnoses; we argue that both default and explicit negation are required to codify DSM-IV as is hypothetical reasoning.
- We show how practical clinical assessment using DSM-IV can be based on interpreting non-stratified negation in DSM-IV through the well-founded semantics [16] augmented by a novel form of preference logic whose semantics we define.
- We describe how the Diagnostica system is based on a partial implementation of these techniques, and discuss an important use for abduction to construct differentials for diagnoses.

Section 2 discusses the knowledge representation problems of DSM-IV in detail. Section 3 shows how these problems can be addressed in an abductive framework that includes logical preferences; while Section 4 provides a 3-valued semantics for these logical preferences and compares it to other semantics in the literature. Finally, we discuss the current version of Diagnostica which partially implements abduction frameworks for preference logic programs. For readability by non-specialists, nearly all discussion of the semantics of our Preference Logic Programs is confined to Section 4. However, we employ standard logic programming terminology throughout.

2 The Nature of Knowledge in DSM-IV

From the perspective of knowledge representation, several factors distinguish the process of psychiatric assessment.

Exclusion Criteria In making a diagnosis, a psychiatrist may need to ensure that certain criteria are fulfilled, while others are excluded. One example of an *exclusion criterion* is criterion 5 for Asperger's disorder (Figure 1) which specifies that criteria must not be met for Schizophrenia or for any other Pervasive Development Disorder (a class that includes Autism, Retts, Childhood Disintegrative Disorder, Asperger's, and Pervasive Development Disorder Not Otherwise Specified). Exclusion criteria occur frequently in DSM-IV diagnosis, with some variability in the phrasing of the negative conditions. Other exclusion criteria may state that "criteria are not better accounted for" by another diagnosis or class of diagnoses, (e.g. in Major Depressive Disorder a criterion requires that symptoms be not better accounted for by Schizophrenia) or that a patient has "not ever" experienced a syndrome (e.g. in Major Depressive Disorder a criterion requires verification that a patient has not ever had a manic episode).

Usually the use of exclusion criteria indicates a priority for how diagnoses are to be made and so the DSM-IV rules are generally stratified through exclusion criteria. For instance, most diagnoses in the class of Mood Disorders require the exclusion of Substance Abuse or Bereavement. In other cases, diagnoses may be non-stratified through exclusion criteria. In the case of dissimilar diagnoses, the non-stratification may be considered an error in DSM-IV; however there are several cases in which the lack of stratification reflects a lack of consensus about how to differentiate the diagnoses. We consider each of the non-stratified classes in turn.

Two diagnoses, Adjustment Disorder and Alzheimer's Dementia illustrate the first class. Both may be considered to be "default" diagnoses, that are to be made only if no other diagnoses are reasonable; exclusion rules for these diagnoses are very broad and can be cyclic. For instance, within the criteria of Cognitive Disorders, a diagnosis of Alzheimer's Dementia should be made only if no other cognitive disorder is more likely for the patient; accordingly, the exclusion rule for Alzheimer's Dementia states

- The disturbance is not better accounted for by another Axis I disorder (e.g. Major Depression, or Schizophrenia).[1]

Adjustment Disorder, which can also be considered as a default diagnosis, has a similarly broad exclusion. Interpreting DSM-IV rules strictly logically, it is possible to have a set of positive criteria that are met such that that a patient has Adjustment Disorder if his symptoms are not better met by Alzheimer's Dementia and that a patient has Alzheimer's Dementia if his symptoms are not

[1] An Axis I disorder is any mental disorder that is not a personality disorder or mental retardation.

better met by Adjustment Disorder. It should be noted, however, that this sort of loop through exclusion criteria is not likely to occur in practice, as it is not likely that a given patient would meet positive criteria for both diagnoses at the same time.

To understand the second class of mutually exclusive diagnoses, consider again the exclusion criterion (5) of Asperger's Disorder. Other Pervasive Development Disorders, such as Autism or Childhood Development Disorder contain similar exclusion rules, so that choosing among the three disorders may be indeterminate according to a logical interpretation of the DSM-IV rules. In the case of the Pervasive Development Disorders, the lack of stratification reflects not only the practical clinical problem of distinguishing Asperger's Disorder from, say, Autism, but also the fact that researchers continue to debate the validity of Asperger's Disorder as a distinct diagnosis altogether (see e.g. [13, 17]). The diagnoses of Asperger's Disorder and Autism is not a unique example of this type of stratification. The diagnoses Adjustment Disorder with Disturbance of Emotions and Conduct, Adjustment Disorder, and Attention Deficit/Hyperactivity Disorder are also linked through exclusion criteria and can be difficult to differentiate [9, 10, 7], as can several other sets of diagnoses.

Thus, while most diagnoses are stratified via exclusion rules, many are not. In many cases, the lack of stratification is accountable by the informality of the DSM-IV rules as with Alzheimer's Dementia and Adjustment Disorder. In these cases the DSM-IV rules should arguably be tightened to avoid inadvertent mistakes caused by exclusion rules that are too broad. However in other cases, such as Asperger's Disorder and Autism the lack of stratification has a deeper nature and reflects the similarity of the disorders themselves.

Incomplete Knowledge: If there are no indications that a patient has an uncommon symptom or case history, certain criteria may be ruled out by default. For instance, the diagnosis of Dissociative Fugue disorder depends on determining that the patient has no medical condition that could also account for the observed symptoms, a determination that may be difficult, if not impossible, to make with absolute certainty. Similarly, many diagnoses depend on a history of the patient that may be impossible to obtain, or may be unreliable from patients or their significant others (e.g. criterion 4 for Asperger's Disorder). For instance,

- A 5-year old child in foster care speaks normally. The physician has no way of obtaining a case history, so that the physician concludes by default that there is no evidence of a significant delay in language acquisition.
- A case history is taken from the child's parents and it is explicitly determined that there was no significant delay in language acquisition.

In the first case, the diagnosis may need to be made on less than perfect information, and there is a need to distinguish information that is assumed false because there is no evidence to support it from information that is explicitly known to be false.

Temporal Reasoning: DSM-IV often requires sophisticated temporal reasoning to represent the duration and occurrence of various symptoms. Indeed, certain closely related diagnoses be distinguished primarily through the duration of the symptoms. An example are the diagnoses Brief Psychotic Disorders, in which delusional symptoms last less than one month, Schizophreniform Disorder (symptoms last at least one month but less than six), and Schizophrenia (in which symptoms have lasted more than six months). Furthermore, temporal reasoning also may be used to

determine whether a patient is diagnosed with single or multiple disorders. For instance, if a patient is both depressed and anxious, he will be treated for anxiety only if the symptoms of an anxiety disorder preceded those of the depression — otherwise the anxiety is taken to be a symptom of the depression itself.

Hypothetical Reasoning: Diagnoses sometimes rely on hypothetical reasoning by the physician, particularly with regard to time. An instance of this is Adjustment Disorder, which has the criterion

- Once the stressor (or its consequences) has terminated, the symptoms do not persist for more than an additional 6 months.

Taken literally, this criterion implies that a physician cannot diagnose a patient as undergoing Adjustment Disorder, while the patient is undergoing it. Similarly, hypothetical reasoning about the expected duration of symptoms may be used to differentiate between the diagnoses of Schizophrenia or Schizophreniform Disorder.

3 Representing DSM-IV as a Logic Program

From the discussion above, it is apparent that modeling DSM-IV as a logic program requires the use of non-traditional techniques. The first three of the factors mentioned above: DSM-IV Exclusion Criteria, Incomplete Information, and Hypothetical Reasoning have been formalized and partially implemented. However, an adequate logic programming representation for the fourth factor, time, in DSM-IV is still to be determined.

3.1 Exclusion Criteria

In order to explain our approach to handling exclusion criteria, we first discuss the actions that should be taken when diagnoses are linked through mutual exclusion rules. First, there are certain diagnoses that are not considered to be similar, but that logically may have loops through exclusion criteria: for instance Alzheimer's Dementia and Adjustment Disorder. Positive criteria should not be satisfied for both of these disorders for any patient at a given time; if this happens, it should be considered an error condition. Second, certain diagnoses are known to be similar but mutually exclusive. In the case of Asperger's and Autism, only one of the diagnoses should be made true: that is, the epidemiological theory underlying DSM-IV states that a patient cannot have both Asperger's and Autism. At the same time if positive criteria are met for both Diagnoses, the action to take is ambiguous. Some clinicians would prefer Asperger's under the principle that if the diagnosis isn't clearly Autism the lesser diagnosis of Asperger's should be made. Other clinicians who don't believe that there is a separate Asperger's disorder separate from Autism would prefer the diagnosis of Autism. Third, in cases such as Pervasive Development Disorders and Schizophrenia which are also linked through exclusion rules, the relationship as specified in DSM-IV is complicated. If a patient has a Pervasive Development Disorder, the additional diagnosis of Schizophrenia is also made if the patient has had prominent delusions or hallucinations for over a month. In other words, Schizophrenia and Pervasive Development Disorders are usually mutually exclusive, but both diagnoses are warranted in certain cases.

Our approach to representing these different kinds of exclusions is based on modeling the exclusions using default negation augmented by abnormality conditions and preference rules. The

resulting program is then evaluated under the well-founded semantics. The portion of the diagnostic rule for Asperger's disorder relevant to exclusion criteria is

```
aspergers:-
   exclude(aspergers,retts),
   exclude(aspergers,autism),
   exclude(aspergers,childhood_disintegrative_disorder),
   exclude(aspergers,pervasive_development_disorder_nos),
   exclude(aspergers,schizophrenia).
```

where `exclude/2` is defined as

```
exclude(Diag1,Diag2):-
   abnormal_situation(Diag1,Diag2).
exclude(Diag1,Diag2):-
   not(Diag2).
```

In the case of Schizophrenia and Pervasive Development disorders, definition of an abnormal situation allows both diagnoses to be true by allowing the exclusion criterion to be satisfied by a means other than negation. At the same time, a set of mutually exclusive diagnoses will be undefined under the well-founded semantics if the positive criteria are met for each diagnosis in the set and if no abnormality conditions are defined. Such a situation is useful for representing cycles through the second clause of *exclusion/2* such as occurs with Alzheimer's Dementia and Adjustment Disorder, as the truth-value undefined can explicitly represent an error that is taken to occur when positive criteria are simultaneously met for both diagnoses.

Both exclusion criteria and abnormality rules model conditions that occur explicitly in DSM IV. However, as discussed previously, there may be similar, mutually exclusive diagnoses, such as Asperger's and Autism for which it should not be an absolute error if positive criteria are simultaneously satisfied for both. In these cases, other criteria, unspecified in DSM-IV may be brought to bear, and it is useful to allow the clinician to state the conditions under which she prefers one diagnosis to another. She would do so by a *preference rule* of the form:

```
prefer(Diagnosis1,Diagnosis2):- Body.
```

A semantics of such preference rules, based on a transformation into normal programs that can be evaluated under the well-founded semantics, is discussed in Section 4. Here, we note that our framework for preference logic is quite general, in that it allows the truth value of preferences (i.e. atoms formed over the predicate *prefer/2*) to depend on the truth value of literals that depend on other preferences, allows preferences to be defined about other preferences, and assigns cyclic preferences the truth value of undefined.

Example 3.1 The following programs illustrate, at a highly abstract level, the actions of Preference Logic Programs on some of the psychiatric diagnoses discussed so far. Let P_1 contain the rules.

```
aspergers :- not autism.
autism :- not aspergers.
```

```
major_depression_disorder.
alzheimers.
```

P_1 abstracts DSM IV diagnosis rules, discussed previously, for Asperger's Disorder and Autism, which are related through exclusion rules, and for Major Depression Disorder and Alzheimer's which are not related through exclusion rules. Suppose a psychiatric practice did not believe in the validity of the Asperger's diagnosis and preferred to diagnose patients with Autism. Suppose further that they believed that DSM-IV diagnostic criteria for Major Depressive Disorder and Alzheimer's were too coarse, and wanted to flag an error in the case when a diagnosis might be ambiguous [2]. In this case the practice could add the following preference rules:

```
prefer(autism,aspergers).

prefer(major_depression_disorder, alzheimers).
prefer(alzheimers,major_depression_disorder).
```

In this case, P_1 together with the preference rules has `autism` true, `aspergers` false, and both `major_depression_disorder` and `alzheimers` undefined.

Next suppose that a particular psychiatrist in a practice wishes to diagnose patients to have Major Depression Disorder rather than Alzheimer's in all cases (perhaps because he is part of a study about the efficacy of an experimental medication for depression). The psychiatrist would add the preference rule.

```
prefer(prefer(major_depression_disorder, alzheimers),
       prefer(alzheimers,major_depression_disorder)).
```

We summarize our treatment of exclusion rules in DSM-IV. Representation of DSM-IV knowledge is kept in the diagnosis rules themselves, including the *exclude/2* and *abnormal_situation/2* predicates. Preference rules allow the user to adjust how exclusion rules are interpreted using knowledge not contained in DSM-IV and, as mentioned above, both cyclic preferences and preferences about preferences may make sense in certain situations. Indeed preference rules could be used in place of the predicate *abnormal_situation/2*; the predicate *abnormal_situation/2* was introduced in order to maintain a distinction between DSM-IV knowledge and that represented by the user.

3.2 Incomplete Information

It is well-known from knowledge representation literature that information that is assumed false because there is no evidence to support it can be represented by default negation; while information that is explicitly known to be false can be represented by explicit negation. Explicit negation can be added to the well-founded semantics without increasing its computational complexity under the well-founded semantics with explicit negation [1] and directly embedded into the well-founded semantics [6]

3.3 Speculative Information

More speculative information, such as that needed to conclude an Adjustment Disorder can be represented using abduction, which allows hypothetical reasoning. Since preference rules can be

[2] The psychiatric literature, in fact, offers support for this view. See [14] for a survey of recent literature.

transformed into normal program rules (Section 4) and evaluated with the well founded semantics (with explicit negation) no special semantics for abduction is needed beyond what is present in the literature: e.g. the three-valued abductive frameworks for extended logic programs of [6]. Because preference logic programs are translated into normal programs, preferences are treated no differently than any other predicate in a program. As a result, the truth value of preferences may depend on particular abductive scenarios, and abductive integrity rules may call preferences just as they may call goals with any other predicates. Furthermore, Definition 4.1 ensures that any abductive dependency of a preference is propagated to literals whose truth depends on these preferences.

It is important to note that abduction plays a larger role in psychiatric diagnosis beyond what is needed to model hypothetical reasoning in DSM-IV, a topic to which we now turn.

Abduction and Differential Diagnoses As has been discussed above, clinicians often need to distinguish between closely related diagnoses. Often this is done through exclusion rules as has been discussed, but other times there are wording differences between positive criteria for similar diagnoses that can be used to as a *differential* between the diagnoses. Indeed, understanding differentials for related diagnoses is a fundamental element of clinical training; and applying these differentials is a fundamental element of clinical practice. Providing dynamic differentials for diagnoses can easily be done through abduction. The idea is that, if the differential is required between diagnosis D_1 and diagnosis D_2 then D_1 should be abduced in the presence of the integrity constraint $\bot$:- D_2 which, using the method of [2], produces the conditions for failing D_2. The abductive context will then provide the differential for the diagnoses.

In order for abduction to be practical for constructing differentials several conditions must hold. First the differential should be as specific as possible, which requires abducibles to be specific and to have an easily understood relationship with one another. At the same time, both the abductive contexts themselves should be minimal as should the number of abductive contexts returned. These goals are to some extent in conflict. For specificity, abducibles should be drawn from atomic propositions that represent the *symptom state* of a patient, and restricted to those atoms of the symptom state that are not known to be true or explicitly false. The most obvious representation of a symptom state makes use of DSM-IV base criteria. Alternatively, the symptom state may consist of elements of other assessment methodologies, such as the World Health Organization's Schedules for Clinical Assessment in Neuropsychiatry [15], which are mapped into DSM-IV base criteria. Adding structure to representation of symptom states benefits the abduction routines: for example if two elements are known to be inconsistent, perhaps because they are antonyms, the inconsistency constraints can be used to restrict abductive solutions.

At the same time, the number of abductive contexts generated should not overwhelm the user. For instance, if criteria 1.a—1.d and 2.a—2.d of Asperger's Disorder in Figure 1 were set as abducibles, then there may be as many as 24 different minimal abductive solutions to the goal `?- aspergers`. To reduce the number of solutions the abduction routines make use of special presentation routines. For instance, when abducing through a criterion in which at least n of a list of base criteria must be true, and for which k of the base criteria are true and l are explicitly false in the symptom state, the abductive solutions are grouped so that the user is presented with a statement of the form at least $(n - k)$ of a revised list (i.e. excluding the explicitly false base criteria) must be present. When abducing base criteria through exclusion rules, a large number of abductive solutions may also be derived. Thus, abduction is not allowed within exclusion rules: rather the exclusion rule itself is returned to the user, after ensuring that the excluded rule is not enforced by the symptom state and presently abduced abducibles.

4 Three-valued Preference Logic Programs

We now define the Preference Logic Programs upon which the representation of DSM-IV is based.

Definition 4.1 A *Preference Logic Program (PLP)* $[P, Pref]$ is a set of normal rules P along with a set of $Pref$ of *preference rules* (or *preferences*) of the form

- $prefer(Term_1, Term_2)$:- $Body$.

Arguments of $prefer/2$ are restricted to be atoms of $[P, Pref]$ are called *preference atoms.*

Assume that $[P, Pref]$ does not contain the predicate symbols $overridden/2$ or $preferred/2$. The *normal embedding* of $[P, Pref]$, $[P, Pref]_{norm}$, is the smallest program containing

1. The rules r' defined as follows. An atom A is *potentially preferred* if there is a preference rule $prefer(A_1, A_2)$:- $Body$ and A unifies with A_2 with mgu θ. Let r be a rule H :- B in P. If H is potentially preferred, then

   $$r' = H \text{ :- } B, not\ overridden(H).$$

   Otherwise $r' = r$.

2. The rules

   $$overridden(A_1) \text{ :- } preferred(A_2, A_1), A_2.$$

   $$preferred(A_1, A_2) \text{ :- } prefer(A_1, A_2).$$
   $$preferred(A_1, A_2) \text{ :- } preferred(A_1, A_3), prefer(A_3, A_2).$$

□

Clearly if the set of preference rules in a PLP $[P, Pref]$ is empty, the normal embedding will have no effect on P beyond adding the rules for $overridden/2$ and $preferred/2$. Definition 4.1 allows preferences to be dynamic in the sense that their truth-value may depend on the truth value of other parts of the program, including other preferences. In addition, preferences can be declared on preferences themselves.

Since $[P, Pref]_{norm}$ is a normal program, it can be evaluated under any semantics for normal programs. For the purposes of this paper, we restrict our attention to the well-founded semantics, under which preferences may have the value true, false, or undefined. It is immediate from Definition 4.1 that cyclic preferences (i.e. atoms of the form $prefer(A, A)$ for some atom A) are either false or undefined in $WFM([P, Pref])$.

4.1 Relation to other Preference Formalisms

The transformational semantics above extends the possible worlds semantics for PLPs as described in [12] which is concerned with what may be termed *static* PLPs [3].

[3] In [12], both P and $Pref$ may be locally stratified: for simplicity of presentation we restrict P and $Pref$ to definite programs in this section.

Definition 4.2 Let $[P, Pref]$ be a PLP in which P and $Pref$ are definite programs. We say that a ground atom A_1 *depends* on a ground atom A_2 if there is a path from A_1 to A_2 in the dependency graph of P. A *derived* atom in $[P, Pref]$ is one that depends on a preference atom. A *base* atom is an atom that is neither a preference atom nor a derived atom. Preferences in $[P, Pref]$ are *static* if all atoms in the bodies of rules in $Pref$ are base atoms. □

For a static PLP $[P, Pref]$, the semantics of $Pref$ is taken as its minimal model, together with that of the base atoms of P. Based on these observations, we can compute preferences, as it were, apart from P and define a relation $<_{pref}$ between atoms such that $A_1 <_{pref} A_2$ if A_2 is transitively preferred to A_1 (using the relation $prefer/2$ of $Pref$).

The possible worlds semantics of preference logic programs is based on *strongly optimal worlds.*

Definition 4.3 Let $[P, Pref]$ be a PLP whose preferences are static and such that P is definite. A set W of atoms over P is *reduced* if there is no $A_1, A_2 \in W$ such that $A_1 <_{pref} A_2$. If W is also a subset of the minimal model of P, then it is called a *world.* A world W_1 is *strongly preferred* to a world W_2 if for each $A_2 \in W_2$ there is an $A_1 \in W_1$ such that $A_2 = A_1$ or $A_2 <_{pref} A_1$. A world W is *strongly optimal* if for all other worlds, $W <_{pref} W_1 \Rightarrow W = W_1$.

The operator T_P denotes the standard inference operator for definite programs. A world W is *supported* if $W \subseteq T_P(W)$. A program $[P, Pref]$ has the optimal subproblem property if every strongly optimal world for $[P, Pref]$ is supported. □

Example 4.1 Consider the PLP P_2:

```
prefer(p(a),p(d)).
prefer(p(b),p(d)).

p(a):- p(d).
p(b).
p(d).
```

There are five worlds for P_2: $\{p(a), p(b)\}$, $\{p(a)\}$, $\{p(b)\}$, $\{p(d)\}$ and $\emptyset$. The world $\{p(a), p(b)\}$ is strongly optimal. However $T_P(\{p(a), p(b)\}) = \{p(b), p(d)\}$ so that P_2 does not have the optimal subproblem property.

Theorem 4.1 *Let $[P, Pref]$ be a static PLP with the optimal subproblem property, such that $<_{Pref}$ is acyclic. Let $[P, Pref]_{norm}$ be the normal embedding of P. Then*

1. *There is a unique strongly optimal world, $\mathcal{W}$ for $[P, Pref]$.*
2. *$WFM([P, Pref]_{norm})$ is two-valued and A is true in $WFM([P, Pref]_{norm})$ iff $A \in \mathcal{W}$.*

Proof:

1. Consider the world $\mathcal{W}$ constructed by computing the two-valued minimal model M_P for P, and then removing all and only those atoms $A \in M_P$ such that $\exists A' \in M_P$ *s.t.* $A <_{pref} A'$. Clearly $\mathcal{W}$ is a world. $\mathcal{W}$ must also be strongly optimal, for suppose there were some other world W' such that $\mathcal{W} <_{pref} W'$. Then $\exists A \in W'$ such that there is no element in $\mathcal{W}$ that is identical to A or preferred to it. By construction of $\mathcal{W}$, A would then not be in M, contradicting the assumption that W' is a world.

2. Since the preferences for $[P, Pref]$ are static and non-cyclic and P is definite, it can be seen from Definition 4.1 that $[P, Pref]_{norm}$ is locally stratified, and thus two-valued. The remainder of the proof is a straightforward induction on the strata of $[P, Pref]_{norm}$ to show that $WFM([P, Pref]_{norm}) = \mathcal{W}$.

$\diamondsuit$

In terms of other related work, a recent paper [5] provides a two-valued fixed-point semantics for static PLPs with the optimal subproblem property (but allowing cyclic preferences). Unlike the semantics presented here, neither the semantics of [5] nor its implementation requires the use of negation. Since [5] contains a theorem similar to Theorem 4.1 the three-valued semantics here also extends the two-valued semantics of [5] for cases in which preferences are non-cyclic. Finally, it should be noted that the atom-based approach to preferences presented above is distinct from those of [3, 4, 11] all of which define preferences on rules rather than on solutions.

4.2 Abductive Frameworks for Preference Logic Programs

Definition 4.1 indicates how a preference logic program can be translated into a normal program. This transformation can be modified to translate extended preference logic programs to extended programs simply by treating objective literals (atoms or their explicit negation) as atoms. The resulting abductive framework for the translated program $[P, Pref]_{norm}$ has the form

$$\langle [P, Pref]_{norm}, A, I \rangle$$

in which A is a set of abducibles and I a set of integrity rules. This framework, in which P and I may include non-stratified negation, can then be directly evaluated by the Abdual method (See [2] for details of Abdual and of the frameworks it evaluates). if A is empty, Abdual reduces to an evaluation of a query under the well-founded semantics and has polynomial data complexity in this case. Using the terminology of [2], this result can easily be extended to preference logic programs:

Proposition 4.1 *Let $[P, Pref]$ be a PLP whose ground instantiation is finite, and $[P, Pref]_{norm}$ be its normal embedding (Definition 4.1). Then Abdual evaluation of a query to the abductive framework $\langle [P, Pref]_{norm}, \emptyset, I \rangle$ has a complexity that is polynomial in the size of those rules in $P \cup Pref \cup I$ whose body is empty.*

Proof: Straightforward from Theorem 3.3 of [2] and Definition 4.1 which ensures that the size of $[P, Pref]_{norm}$ is polynomial in the size of $P \cup Pref$. $\diamondsuit$

5 The Diagnostica System

Investigation into the logical representation of DSM-IV was sparked by the desire to help automate DSM-IV in a commercial system, Diagnostica, a beta version of which is available `http://www.medicinerules.com`. We note that full implementation of Diagnostica, using the techniques of Section 3 is not yet complete, so that in this section we briefly describe its current, partial implementation whose reasoning facilities are based on the XSB system (cf. `http://www.cs.sunysb.edu/~sbprolog`).

Figure 2 shows part of the user interface of Diagnostica in which a clinician is entering symptoms for Attention-Deficit Hyperactivity Disorder (ADHD). The window on the left represents the graph

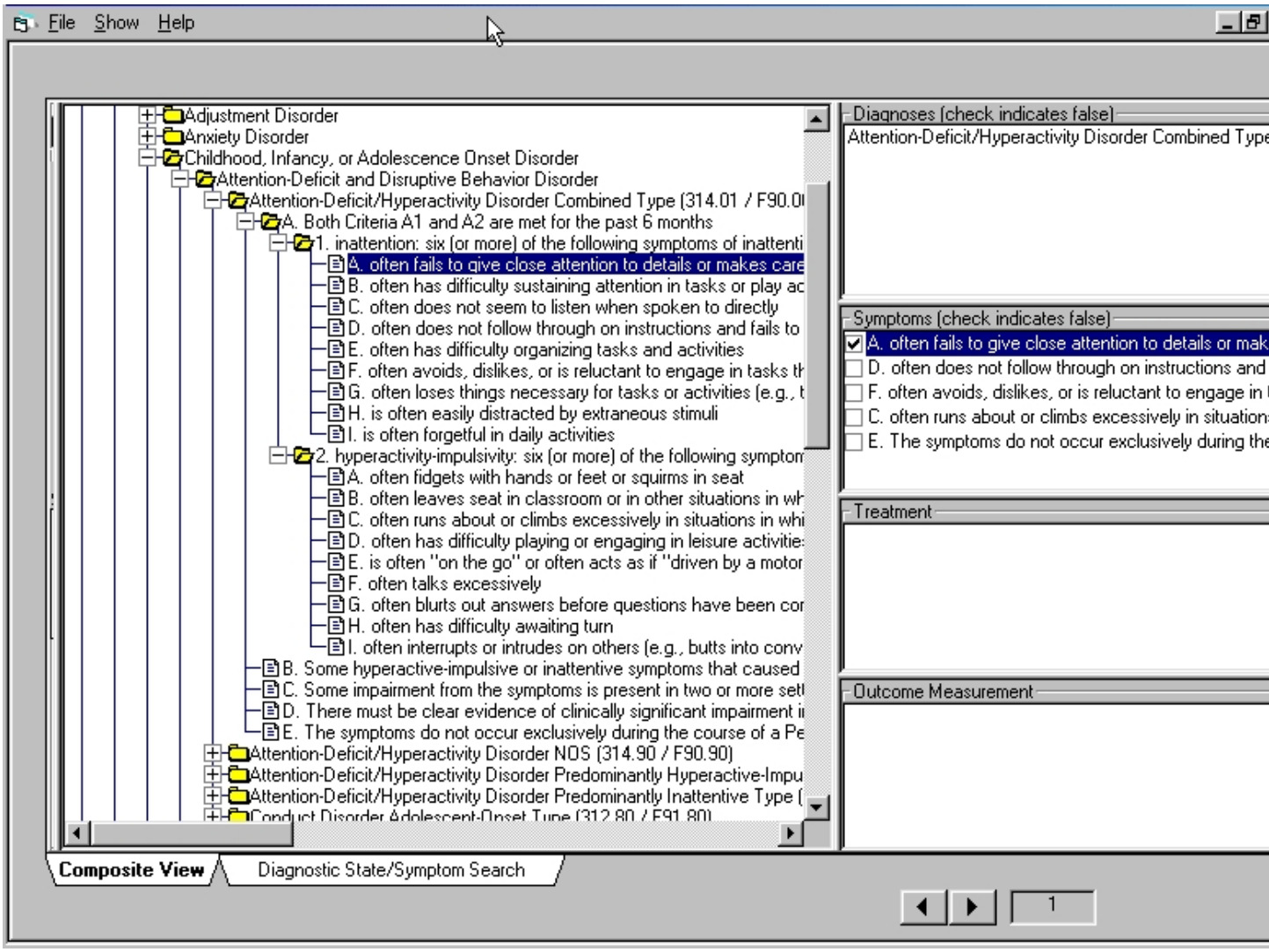

Figure 2: Using Diagnostica for ADHD

of DSM-IV rules; the top left window represents the diagnosis being viewed; the window below it represents base criteria that are true or false for a given diagnosis. As a user verifies base criteria various diagnoses, other than ADHD, may become partially satisfied and the user may wish to view these partially satisfied diagnoses along with the criteria necessary to derive the diagnoses. The user may view these diagnoses by pressing the tab "Diagnosis State Symptom Search".

The current user interface for Diagnostica thus uses abduction in a simple, but clinically relevant way. True differentials for diagnoses are not yet available to the user, nor are screens for adding or manipulating preferences. While the inclusion of these features is planned for future versions, it is worthwhile noting that designing a visual user interface for these features is not always straightforward. Specifying the body of a preference, for instance, must be carefully structured so that users obtain the effects that they intend, and creating a cogent user interface to allow the creation of such rules is an open problem.

6 Summary

Non-stratified programs are sometimes considered to be of little use for practical problems. However, translation of DSM-IV diagnostic rules into logical rules shows that sets of closely related diagnoses form non-stratified recursive components, so that non-stratified negation is semantically meaningful in this translation. Indeed, it is difficult to see how DSM-IV could be adequately coded without non-stratified negation. The well-founded semantics is used to represent the semantics of DSM-IV rather than, say, stable models for several reasons. It is convenient to use the *undefined* truth value to represent error conditions for diagnoses such as Alzheimer's Dementia and Adjustment Disorder. Second, the addition of preference rules under the well-founded semantics allows resolution of non-stratified loops while retaining a polynomial complexity of evaluation when abduction is not required. The addition of abduction to well-founded preference logic programs allows representation of hypotheses used in diagnoses as well as a means of constructing differentials for diagnoses.

The need to implement these aspects of DSM-IV in Diagnostica has helped spur the development of the Abdual evaluation method [2] as well as the Preference Logic presented here. At the same time, development of these formalisms has been necessary in order to understand how to implement abduction and preferences in Diagnostica. Further work for the representation of psychiatric diagnosis includes the representation of temporal reasoning in Diagnostica as well as studying how to guide the diagnostic process using epidemiological information (expressed as probabilities of various symptoms or syndromes that map into DSM-IV criteria). This work should lead to the commercial validation of further knowledge representation techniques, and perhaps to better understanding of the knowledge representation techniques themselves.

Acknowledgements

The authors would like to thank C.R. Ramakrishnan and David S. Warren for their work on a preliminary version of Diagnostica; Michael Gelfond for a discussion on knowledge representation in psychiatry; and Bharat Jayaraman for a discussion on the possible-worlds semantics of preference logic programs. A U.S. Patent is pending for the use of several of these techniques within the framework of medical reasoning.

References

[1] J. Alferes, C. Damásio, and L. M. Pereira. A logic programming system for non-monotonic reasoning. *Journal of Automated Reasoning*, 14(1):93–147, 1995.

[2] J. Alferes, L. M. Pereira, and T. Swift. Well-founded abduction via tabling dual programs. In *International Conference on Logic Programming*, pages 426–440. MIT Press, 1999.

[3] G. Brewka. Well-founded semantics for extended logic programs with dynamic preferences. *Journal of Artificial Intelligence Research*, 4:19–36, 1996.

[4] G. Brewka and T. Eiter. Preferred answer sets. In *Proceedings of the 6th Conference on Principles of Knowledge Representation and Reasoning*, pages 86–97. Morgan Kaufmann, 1998.

[5] B. Cui, T. Swift, and D. S. Warren. A case study in using preference logic grammars for knowledge representation. In *International Conference on Logic Programming and Non-Monotonic Reasoning*, pages 206–220. Springer-Verlag, 1999. LNAI 1730.

[6] C. Damásio and L. M. Pereira. Abduction over 3-valued extended logic programs. In *International Conference on Logic Programming and Non-Monotonic Reasoning*, pages 29–42. Springer-Verlag, 1995. LNAI 1265.

[7] P. B. de Mesquita; W. S. Gilliam. Differential diagnosis of childhood depression: using comorbidity and symptom overlap to generate multiple hypotheses. *Child Psychiatry Hum Dev*, 24:157–172, 1994.

[8] *Diagnostic and Statistical Manual of Mental Disorders.* American Psychiatric Association, Washington,DC, 4th edition, 1994. Prepared by the Task Force on DSM-IV and other committees and work groups of the American Psychiatric Association.

[9] R. Famularo, R. Kinscherff, and T. Fenton. Psychiatric diagnoses of maltreated children: preliminary findings. *J Am Acad Child Adolesc Psychiatry*, 31:863–867, 1996.

[10] J. Ford, R. Racusin, W. Daviss, C. Ellis, and J. Thomas. Trauma exposure among children with oppositional defiant disorder and attention deficit-hyperactivity disorder. *J Consult Clin Psychol*, 67:786–789, 1999.

[11] M. Gelfond and T. C. Son. Reasoning with prioritized defaults. In *Logic Programming and Knowledge Representation*, pages 164–223. Springer-Verlag, 1997. LNAI 1471.

[12] K. Govindarajan, B. Jayaraman, and S. Mantha. Preference logic programming. In *International Conference on Logic Programming*, pages 731–746. MIT Press, 1995.

[13] M. Prior, R. Eisenmajer, S. Leekam, L. Wing, J. Gould, B. Ong, and D. Dowe. Are there subgroups within the autistic spectrum? a cluster analysis of a group of children with autistic spectrum disorders. *Can J Psychiatry*, 43(6):589–595, August 1998.

[14] L. Rosenstein. Differential diagnosis of the major progressive dementias and = depression in middle and late adulthood: a summary of the literature of the early 1990s. *Neuropsychol Rev*, 8:109–167, 1998.

[15] *Schedules for Clinical Assesment in Neuropsychiatry.* World Health Organization, 1996. Version 2.1.

[16] A. van Gelder, K. Ross, and J. Schlipf. Unfounded sets and well-founded semantics for general logic programs. *JACM*, 38(3):620–650, 1991.

[17] F. Volkmar, A. Klin, and D. Pauls. Nosological and genetic aspects of asperger syndrome. *J Child Psychol Psychiatry*, 39(6):893–902, September 1998.