Logic Engineering in Medicine*

Peter Lucas
Department of Computer Science, Utrecht University
Padualaan 14
3584 CH Utrecht, The Netherlands
E-mail: lucas@cs.uu.nl

Abstract

The safety-critical nature of the application of knowledge-based systems to the field of medicine, demands the adoption of reliable engineering principles with a solid foundation for their construction. Logical languages with their inherent, precise notions of consistency, soundness and completeness offer such a foundation, thus promoting scrutinous engineering of medical knowledge. Moreover, logic techniques provide a powerful means for getting insight into the structure and meaning of medical knowledge used in medical problem solving. Unfortunately, logic is currently only used on a small scale for building practical medical knowledge-based systems. In this paper, the various approaches proposed in the literature are reviewed, and related to different types of knowledge and problem solving employed in the medical field. The appropriateness of logic for building medical knowledge-based expert systems is further motivated.

Keywords & Phrases: logic programming in medicine, logic engineering, knowledge-based systems, expert systems, medical knowledge representation.

1 Introduction

In the words of Sir Arthur Conan Doyle, Dr. Watson is portrayed as an archetypical medical doctor, displaying deductive qualities far inferior to those of Sherlock Holmes, so that Holmes' sarcastic comments on Watson's theories often turn out to be justified. Given that the work of a medical doctor has much in common with that of a detective, Doyle's opinion of the medical profession may be too negative. Yet, there is some truth in Doyle's picture: doctors have never displayed much interest in and mastery of formal techniques. Instead, they tend to emphasize the important role played by clinical intuition in the diagnosis and treatment of disease, and often say that formal techniques are inappropriate for capturing the subtleties of the medical decision-making process. An important reason for this belief might be the strongly patient-centred nature of clinical medicine, which always left little time and interest to bring medicine on a firm footing, by framing it according to some formal theory of medical decision processes. The beneficial effects of such a theory for the individual patient were not immediately obvious. Of course, these medical doctors are right, but not entirely. It is increasingly recognized that clinical intuition, however practically important it may be, constitutes a basis too weak for modern medicine [Macartney, 1988]. Although many doctors are still reluctant to adopt

*This paper was published in: *The Knowledge Engineering Review*, **10**(2), 153–179, 1995.

formal techniques in medical decision making, formal techniques are slowly taking their place. For example, there are an increasing number of groups working on formal consensus among medical specialists concerning diagnostic and treatment management, for example in the field of oncology. Any joint effort to express precisely which diagnostic and therapeutic actions are most appropriate for patients with a given disease, may improve health care. Modern society requires that doctors are capable of justifying their professional decisions and actions to the patient, a trend which, undoubtedly, will only increase in the near future. Obviously, this will be possible only if the doctor is provided with detailed knowledge about the diagnostic and treatment procedures which are generally accepted, and is supported in applying them. Methods and techniques developed within the areas of decision-support systems in general, and medical expert systems in particular, may contribute considerably to this process.

Generally speaking, there are two frameworks that support current trends in formal medicine, referred to in this paper as:

- decision theory, and
- symbolic reasoning technology.

Decision theory, including decision analysis, probabilistic and decision methods, has been the framework most widely adopted in the medical field. The decision-theoretic approaches are attractive from a medical perspective, because they focus on the handling of *uncertainty*, a subject of much concern in medical decision making. *Symbolic reasoning technology* focuses on the *structure* of medical knowledge and medical problem solving, with additional emphasis on capturing the *meaning* of medical knowledge used in problem solving. These two frameworks are not necessarily incompatible as is demonstrated by the existence of probabilistic belief networks and influence diagrams, where uncertainty and structure is combined in a unifying framework [Kim & Pearl, 1983; Pearl, 1988; Shachter, 1986]. Nevertheless, much of the structure of medical knowledge cannot be captured in terms of decision theory – simply because probability theory is not sufficiently expressive for representing various kinds of semantic relationships – where there are ways of handling uncertainty within a symbolic reasoning framework (although this work does not have the canonical status of probability theory) [Duda et al., 1976; Shortliffe & Buchanan, 1975; Cohen, 1985; Krause et al., 1995].

In this paper, research concerning the application of symbolic reasoning technology to the medical field, using logic as the principal tool, is reviewed. This particular approach to building systems is sometimes referred to as *logic engineering*, a name coined by the Computational Logic Group at Imperial College in London to the application of logic programming techniques, and popularised by the Advanced Computation Laboratory at the Imperial Cancer Research Fund in London [Fox et al., 1990b]. The meaning of the term 'logic engineering' in this paper is taken to be wider; of interest here is the use of techniques from the fields of logic programming, mathematical logic, logical knowledge representation and theorem proving for building medical applications. Although much work has been done in the application of symbolic reasoning technology to the medical field, this paper focuses almost exclusively on the logical formalization of medical knowledge and general medical principles, such as diagnosis and treatment management. The use of specific techniques from the field of logic programming to the medical field is also covered. Only limited attention is given to related topics, such as qualitative reasoning, if the applied techniques are not logical in nature. The application of Prolog in developing particular medical systems, without exploiting some logical characterization of medical knowledge, falls outside the scope of the paper as well.

Although logical techniques are now considered fundamental to both computer science and artificial intelligence, the situation in medical informatics is quite different. As has been discussed above, in medicine there is a large gap between theory and practice. Bridging this gap using logical techniques is a major challenge.

The organization of the paper is as follows. In Section 2, the grounding of logic engineering into the field of theorem proving and logic programming is reviewed. In Section 3, we focus on the knowledge-representation aspects of using logic for representing medical knowledge. Section 4 is concerned with the specification of medical problem solving using logical techniques. The paper is rounded off by summarizing what has been achieved by the field, and a number of future research directions are suggested.

2 Grounding of the field

Most of the logical and theorem proving techniques have emerged from research with mathematical and computer science (including artificial intelligence) applications in mind [Chang & Lee, 1973; Kowalski, 1979]. This research has revealed some, more or less severe, limitations of using logic for solving mathematical or computer science problems [Kowalski, 1990; Wos, 1988; Wos et al., 1992]. Much research effort has been invested in the last two decades to overcome some of these limitations, which has yielded deep insight into the subject, (cf. [Lloyd, 1987; Stickel, 1986]), and several new and interesting logical languages and systems have emerged. Many of these languages are based on the language Prolog [Sterling & Shapiro, 1986]. Examples of such languages are Gödel – a language which provides, among others, special language facilities for better control over reasoning [Hill & Lloyd, 1994] – and Login and Life – which are Prolog-like languages with built-in inheritance relationships among terms [Aït-Kaci & Nasser, 1986; Aït-Kaci & Podelski, 1993]. New programming paradigms have also been introduced in the realm of logic programming, such as logic object-oriented programming in the language *L & O*, [McCabe, 1992], – which allows programs to be organized into classes, and to reason over expressed relationships among those classes –, constraint logic programming in languages such as CLP($\mathcal{R}$), [Jaffar & Lassez, 1987], and Chip, [Dincbas et al., 1988], – which allow the declarative expression of equality among the real and rational numbers, respectively –, and inductive logic programming which concerns learning new relationships from logical data and background knowledge [Muggleton, 1992]. Furthermore, general purpose theorem provers, such as Otter, [McCune, 1990; Wos et al., 1992], and PTTP, [Stickel, 1988], provide rich environments for experimenting with reasoning strategies.

Although many application fields are likely to benefit from the achievements of this research, it is important to realize that many of these, including the limitations, will not automatically carry over to all fields. Medicine, for example, is characterized by a huge body of factual knowledge concerning specific instances, laid down in thick medical textbooks and scientific journals. This contrasts, for example, with mathematics and computer science, where it is customary to strive for generality. Where with mathematics many of the most elementary expressions cannot be specified in first-order logic, and require at least second-order logic, in medicine much of the knowledge is almost propositional in nature. Hence, considering the use of logic for representing medical knowledge requires rethinking the limitations encountered in typical (mathematical) theorem-proving and logic programming applications. As shall become clear, the language of first-order predicate logic offers sufficient expressiveness for the representation of medical knowledge, although there is room for additional machinery, such as meta-level reasoning, [Van Harmelen, 1991; Maes & Nardi, 1988; Weyhrauch, 1980], and non-standard logic, [Łukaszewicz, 1990; Smets et al., 1988; Turner, 1984].

From a medical point of view, there are several advantages linked with the logic engineering approach:

- Logic has a well-defined syntax and semantics, resulting in clearly defined meanings of the represented medical knowledge in terms of the relationships among pieces of knowledge.
- Soundness and completeness results for logical deduction are well-established. Hence, beforehand it is known that any derived medical conclusion can be trusted as following from the represented medical knowledge, and that none of the relevant conclusions will be missed.
- Well-founded methods for verification are available. For example, conflicting medical data and knowledge can be identified.

Medical applications are often safety-critical. Consequently, the well-developed formal underpinning of logic engineering may contribute to the safe application of medical expert systems in health care [Fox, 1993]. Of course, no guarentee of the medically sound content of such systems can be derived from this. Methods for the validation of the resulting systems remain indispensable [O'Keefe, 1987; Wyatt & Spiegelhalter, 1990].

Logic is not the only class of formalisms having the nice properties mentioned above. They are shared, for example, by the algebraic specification languages [Ehrig & Marh, 1985; Ehrig & Marh, 1990; Spivey, 1989]. Algebraic techniques have only been applied on a small scale in designing medical systems (cf. [Todd, 1994; Hammond & Davenport, 1995]).

3 Logic representation of medical knowledge

In this section, the declarative aspects of medical knowledge are considered from a logical point of view; we shall comment on the application of particular logical specifications of medical knowledge, when appropriate. One word on notation: in this paper we use a mixture of standard logic and logic programming notation, because the research reviewed is drawn from both fields.

3.1 Semantic data representation

A good starting point for thinking about the logical representation of medical knowledge are the data required in solving medical problems, part of which are patient findings. In practical implementations, medical data are often represented in a record-like fashion, without giving proper attention to the medical meaning of individual elements [Lucas, 1993]. The logical basis of data has been studied extensively in the field of database systems. We shall briefly review some of the consequences of database theory for the representation of medical data.

Consider, for example, the following term representation of an instance of the entity *patient*:

$$\begin{aligned}&\mathit{patient}(\mathit{name} \Rightarrow \mathit{patient1},\\&\qquad \mathit{sex} \Rightarrow \mathit{male},\\&\qquad \mathit{age} \Rightarrow 66\end{aligned}$$

$$complaint \Rightarrow [oppressive\_pain, nausea])$$

containing data of a patient with cardiac disease, where elements such as *sex* represent attributes, and elements such as *male*, represent values of attributes. In predicate logic, attributes such as *age* are single-valued, i.e. they may take at most one value at a time. A single-valued attribute a is interpreted as a function $a : E \rightarrow V$ from the set of distinguished entities E to the set of values V of the attribute; the representation of the information about *age* in the term above is therefore

$$age(patient1) = 66$$

where the equality predicate enforces the single-valuedness [Lucas & Van der Gaag, 1991]. Attributes such as *complaint* are multi-valued, i.e. they may take more than one value at a time. A multi-valued attribute A is interpreted in predicate logic as a binary relation $A \subseteq E \times V$. Hence, a multi-valued attribute is represented in predicate logic as a predicate symbol. For example,

$$\begin{array}{l} Complaint(patient1, oppressive\_pain) \wedge \\ Complaint(patient1, nausea) \end{array}$$

is the predicate logic representation of information concerning the complaints of patient 1 given above.

Note that the predicate logic representation of the data reveals the precise meaning of the elements, as well as the various roles they play in the example above. For example, *age* maps a specific patient to a single number, representing the patient's age; the role of the value *nausea* is that of a complaint of the patient. Patient data can be grouped further by including a reference to the test or procedure from which the data are derived. For example,

$$Complaint(patient1, medical\_interview, oppressive\_pain)$$

indicates that the finding has been obtained from the medical interview of the patient.

Another important logical aspect of medical data is the handling of unknown data. Due to the overwhelming number of facts in medicine, it is not possible to explicitly represent all those facts, even when they concern a single patient. In particular, the logical interpretation of the medical data involves the tests or procedures by which the data have been obtained. One possible, and often adopted, logical interpretation of medical data is as follows [Lucas, 1993]:

- if a test has a single possible outcome, but has not been carried out, or the test result is unknown, then the corresponding attribute value is assumed to be unknown; otherwise, only one value is represented as indicated above.
- if a test may have more than one simultaneous outcome, and the test has not been carried out, or all test results are unknown, then the attribute value is assumed to be unknown; otherwise, all test results obtained are represented, augmented with the remaining possible results as negative literals.

For the patient above, if oppressive pain and nausea are the only complaints entered, and fever is another, unmentioned possibility, it is assumed that

$$\neg Complaint(patient1, fever)$$

holds. If nothing is known about the patient's complaints, no unit clauses concerning '*Complaint*' are represented. A weak form of the closed world assumption (CWA) is therefore taken to hold in many medical applications, i.e. in contrast with the original CWA, [Reiter, 1978], absent information is not automatically assumed to be negative. Little is known about the precise conditions under which this special CWA is acceptable from a medical point of view. It may be necessary to distinguish explicitly between negative information obtained from the weak CWA, and negative information explicitly entered by the doctor, with respect to the conclusions drawn using negative information.

3.2 Medical representation models

Medical problem solving involves a plethora of types of medical knowledge, as was already evident in the early days of medical expert systems [Pople, 1982]. In this section, the logical formalization of some of these types of knowledge, and the current state of the art in research, is briefly reviewed. In Table 1, an overview of logic engineering research is listed.

Logic model	Application	References
causal	simulation	[Bratko et al., 1989; Lucas, 1993]
	diagnosis	[Console et al., 1989]
		[Torasso & Console, 1989]
		[Fox et al., 1990a; Fox et al., 1990b]
anatomic	design	[Hammond et al., 1993]
	diagnosis	[Lucas, 1993]
taxonomic	diagnosis	[Huang et al., 1993]
heuristic	diagnosis	[Fox et al., 1990b]
		[Lucas, 1993]
		[Moser & Adlassnig, 1992]
functional	diagnosis	[Bratko et al., 1989; Coiera, 1990]
	monitoring	[De Geus et al., 1991]
	treatment	[De Geus et al., 1991]
safety	treatment	[Hammond et al, 1994; Hammond & Sergot, 1995]

Table 1: Application of logic representation models in medicine.

In applying logic to the construction of medical knowledge-based systems it is becoming increasingly common to make a distinction between several layers of knowledge. Apart from being concerned with the representation of basic objects and relationships among objects in medicine, medical representation models also deal with the representation of medical decision structures. Both facets can be represented in logic, although, to achieve a perspicuous specification, often at different layers. In this paper, it is sufficient to distinguish only two such layers; the first layer, called the *object-layer*, is concerned with medical data and relationships among the data, and the second layer, called the *meta-layer*, comprises a specification of reasoning methods to perform particular medical tasks, such as diagnosis or making treatment decisions, using information from the object-layer [Fox et al., 1990a]. The meta-layer may also incorporate descriptions of properties of the object-layer relations.

3.2.1 Causal medical models

Causality is an important structural concept in medicine because the pathogenesis of disease as well as the effects of treatment are described in medical textbooks and journals in terms of cause-effect relationships. Causality was the basic modelling concept in several early medical expert systems such as CASNET/GLAUCOMA [Kulikowski & Weiss, 1982] and ABEL [Patil et al., 1982]; it has also been investigated by H.E. Pople [Pople, 1973; Pople, 1977] prior to the development of the INTERNIST-I system [Miller et al., 1982] and was again incorporated in the CADUCEUS system [Pople, 1982]. With the advent of heuristic expert systems, [Buchanan & Shortliffe, 1984], interest in causality diminished somewhat, although it might be argued that causality remained an important concept in qualitative reasoning, where the behaviour of a system is described in terms of structure and function of components. However, no explicit representation of the notion of causality can be found in most qualitative reasoning methods, such as QSIM [Kuipers, 1986; Coiera, 1992]. By and large, there is now renewed interest in the use of the notion of causality in designing and building medical expert systems.

One of the problems with modelling cause-effect relations for medical problem solving is that medical knowledge is usually incomplete and uncertain. The incompleteness results from the inadequate elucidation of the causal chains contributing to (patho)physiological mechanisms. The uncertainty may be viewed as a consequence of this incompleteness, in the sense that if the contributing factors, including their interaction in producing an effect, are incompletely known, all one can say is that the particular effects have been observed repeatly under certain conditions, and may be attributed to certain causes. A consequence of this limitation is that the application of logic for modelling cause-effect relationships approximates medical reality only roughly. However, the notion of causality remains applicable, even when the knowledge of cause-effect relationships is incomplete [Coiera, 1992; Console et al., 1989]. Incompleteness of knowledge renders causal relationship only less detailed. This contrasts with qualitative reasoning models, which are somewhat more demanding with respect to knowledge concerning the (possibly abnormal) structure and function of parts of the human body. In Section 3.2.5, this subject is discussed in more detail.

Whether or not a logical model of causal knowledge is sufficiently accurate depends, of course, also on the use that is made of it. The application of causal knowledge for diagnostic problem solving is a very active area of research; it will be dealt with in Section 4.3. Here we shall mainly deal with the declarative aspects of causal knowledge, and only discuss how causal knowledge is employed as far as this promotes understanding.

There are two common ways in which the causal relation is formalized in logic. Firstly, causality can be formalized by means of a binary predicate symbol $\mathit{Causes}(x, y)$, denoting that the *cause* x has y as an *effect*. Axiomatization of the meaning of the causal relation always includes transitivity

$$\forall x \forall y \forall z ((\mathit{Causes}(x, z) \wedge \mathit{Causes}(z, y)) \rightarrow \mathit{Causes}(x, y))$$

Sometimes, axiomatization includes antisymmetry:

$$\forall x \forall y (\mathit{Causes}(x, y) \rightarrow \neg \mathit{Causes}(y, x))$$

Yet, in other formalizations, the predicate may be reflexive

$$\forall x \, \mathit{Causes}(x, x)$$

and, hence, not antisymmetric. From a medical point of view, few causal relations are reflexive – e.g. diseases or findings are not caused by themselves – so that in a medical setting reflexivity is not acceptable in general. This approach to the modelling of the causal relation has been followed, for example, in the Oxford System of Medicine (OSM), [Fox et al., 1990b], a large medical expert system for general practitioners, and the more recent DILEMMA system, [Huang et al., 1993], where part of the specifications concerns the causal relation. In contrast with the OSM, in the DILEMMA system, causal knowledge is represented using the meta-predicate *Domain*:

$$\mathit{Domain}(x, \mathit{causes}, y)$$

meaning 'x causes y', where particular instances for the variables x and y, represent object-constants and *causes* represents an object-predicate, essentially with a meaning identical to that of the predicate symbol *Causes*. The *Domain* predicate obtains meaning at the meta-layer. Properties of relations, such as transitivity for the causal relation, and the semantic relationship among relations are represented in the system at the meta-layer (cf. Section 4.1). For example, the meta-relation

$$\mathit{Meta\text{-}relation}(\mathit{causes}, \mathit{inverse}, \mathit{caused\ by})$$

expresses that the '*caused by*' relation is the inverse of the '*causes*' relation. Although the same information could also be expressed at the object-layer by

$$\forall x \forall y (\mathit{Causes}(x, y) \leftrightarrow \mathit{Caused\_by}(y, x))$$

the meta-information that this relation is an instance of the *inverse* relation would be lost.

In the second formalization, causality is modelled using logical implication, i.e. the implication

$$\mathit{cause} \rightarrow \mathit{effect}$$

represents a cause-effect relationship [Console et al., 1989]. Due to the special meaning attached to logical implication, reflexivity and transitivity are satisfied. However, this logical formulation of the causal relationship also poses some problems, not encountered in the first formalization of causality, being a consequence of the special meaning of logical implication.

Consider, for example, the following logical representation of a medical cause-effect relationship

$$\mathit{Disorder}(\mathit{infection}) \rightarrow \mathit{Complaint}(\mathit{fever})$$

In this case, we do not want the effect that by adding

$$\neg \mathit{Complaint}(\mathit{fever})$$

the logical consequence

$$\neg \mathit{Disorder}(\mathit{infection})$$

is obtained. Deriving that a patient does not have an infection from the fact that there are no fever complaints, seems too strong from a medical point of view. On close inspection it turns out that the formalization of this notion of causality corresponds to the if part of the logical formulation of the notion of *correlation* [Konolige, 1992]. Positive correlation between two entities C and E, can be formulated by the bi-implication:

$$C \leftrightarrow E$$

Here,

$$\{C \leftrightarrow E, \neg E\} \vdash \neg C$$

rightfully does hold in general; if C is positively correlated to E, then $\neg E$ is positively correlated to $\neg C$. However, the primary use of the formalization of causality using logical implication is in reasoning methods (to be discussed in Section 4.3) where logical deduction is applied in a forward-driven fashion, i.e. from the antecedent to the consequent of an implication. Then, the problem does not occur, but note that this solution is obtained at the (deliberate) cost of incompleteness. When adopting a logic programming approach, the theory, as well as the formulas representing data, may be restricted to positive (definite) clauses. Then, the problem does not occur either, simply because negative data are not allowed, or only applied with respect to part of the logical theory. However, as the example above indicates, it is necessary to have the ability to represent negative data in medicine; hence, the restriction to definite clauses may be too strong.

The formalization of causality using implication has the advantage that several extensions can be accommodated in logic in a straightforward way, where for the first approach to modelling causality significant additional axiomatization would be required. Thus, the expressiveness of the causal relation is increased, making it more suitable for representing the subtleties of medical knowledge. For example, the notion of *conditional causality*, meaning that the satisfaction of a cause-effect relationship is dependent on the truth of some condition, can be formulated in logic as follows [Console & Torasso, 1990a]:

$$\mathit{cause} \wedge \mathit{condition} \rightarrow \mathit{effect}$$

Conditional causality is useful in medicine in relating two (physiological) states to each other, where one state passes into another state when some condition is satisfied. One possible condition is duration in time. For example,

$$\begin{array}{l}(\mathit{state}(\mathit{heart}) = \mathit{ischaemia} \ \wedge \\ \ \mathit{duration}(\mathit{state}(\mathit{heart})) > 30) \\ \rightarrow \mathit{state}(\mathit{heart\text{-}muscle}) = \mathit{necrosis}\end{array}$$

i.e., if the blood supply of the heart is diminished (ischaemia) and the duration of that state of low supply surpasses 30 minutes, the heart muscle cells will die (necrosis). In the logical framework of L. Console and P. Torasso ([Console & Torasso, 1990a]), a condition is logically distinguished from a cause by the fact that a condition must have been deduced to be satisfied, where causes can only be hypothesized, or derived from other hypothesized causes.

Representing medical knowledge using one of the formalizations of causality above may yield a large number of formulas. Therefore, it may be difficult to maintain the overview of the knowledge base. Using a more abstract representation of the causal knowledge may help in preserving the overall picture of the contents of a logical knowledge base. In their CHECK system, Console and Torasso employ the following abstract notation for expressing conditional causality between two states S and S'

$$\forall x(S(x) \wedge C \rightarrow S'(f(x)))$$

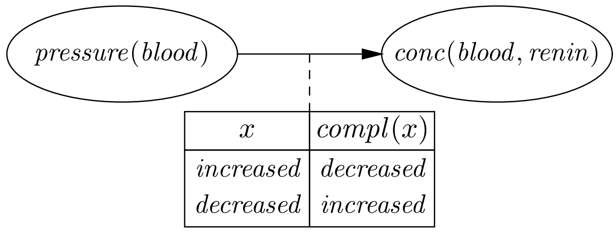

Figure 1: Abstract representation of causal knowledge.

where the function term $f(x)$ is the abstract representation of the functional dependence of the state S' on the state S, specified separately by means of the function definition f [Torasso & Console, 1989]. For example, consider the following abstract causal relation:

$$\forall x(\mathit{pressure}(\mathit{blood}) = x \rightarrow \mathit{conc}(\mathit{blood}, \mathit{renin}) = \mathit{compl}(x))$$

(renin is a substance released from the kidneys if the blood pressure drops), where the condition C is assumed to be universally true, and can therefore be left out. If

$$\mathit{pressure}(\mathit{blood}) = \mathit{decreased}$$

is the state given, then

$$\mathit{conc}(\mathit{blood}, \mathit{renin}) = \mathit{increased}$$

is concluded, where $\mathit{increased} = \mathit{compl}(\mathit{decreased})$. This concise notation makes it possible to represent causal knowledge as a graph, where each vertex stands for a set of ground atoms, and the causal relation is given in the form of a table of function values associated with the arcs of the graph, as is done for the implication above in Figure 1. This representation has much in common with a belief-network representation (where arcs have a probabilistic interpretation), [Pearl, 1988], and *qualitative function constraints* in qualitative reasoning, as proposed by B.J. Kuipers (cf. Section 3.2.5), where, in contrast with the causal relation discussed above, the notion of time has been incorporated [Kuipers, 1986]. In [Console & Torasso, 1991b], Console and Torasso propose to extend their causal model by adding the explicit representation of temporal relations.

As we have remarked above, the modelling of medical knowledge using the logical formalization of causality may be difficult. With the aim of obtaining a more flexible notion of causality, Console and Torasso distinguish between weakly ('may cause') and strongly ('must cause') causal relations (actually, these relations are called 'possible' and 'necessary', respectively, in their paper) [Console & Torasso, 1990a]. The formalization discussed above concerns only the strongly causal relation; the following formula

$$\mathit{cause} \wedge \alpha \rightarrow \mathit{effect}$$

expresses the weakly (or possibly) causal relation; the literal α is called an assumption literal. As distinct from the condition C in conditional causality, an assumption literal cannot be derived; it can only be assumed. Even if the cause holds, the effect will not be derived when the assumption literal is assumed to be false. This is precisely the effect required. The various forms of causal relations using implication were investigated within the medical field using the system CHECK, and applied, among others, to the diagnosis of disorders of the liver.

As can be seen in Table 1, formalized causal knowledge is usually applied for diagnosis.

3.2.2 Anatomical models

Although knowledge of the anatomical structures in the human body is important for problem solving in some medical specialties, anatomical reasoning models have not been applied very frequently. The potential complicated nature of such models may explain this lack of research. Neurology is a medical specialty in which knowledge of anatomical relations is of crucial importance in the process of diagnosis [Jaspers, 1990; Reggia, 1978; Tuhrim, 1991]. The first medical expert system that used anatomical knowledge as part of a problem-solving model, but which was not based on logic, is LOCALIZE [First et al., 1982]. Work on the logical representation of anatomical knowledge is relatively scarce, although the domain seems eminently suitable for applying logical techniques.

In [Lucas, 1993], the axiomatization and application of basic anatomical relations for the diagnosis of lesions of the facial nerve is described. Here, knowledge about the way in which parts of the nerve segments are connected to each other sufficed for diagnostic application of that knowledge. Such simple anatomical relations can be axiomatized in a straightforward manner by means of a binary *Connected* predicate, defined to be transitive and irreflexive; $\mathit{Connected}(x, y)$ means that a nerve runs from level x up to level y. The (extensional) anatomical knowledge was provided for by identifying levels with anatomical structures through which the nerve runs. For example,

$$\mathit{Connected}(\mathit{stapedius\_nerve}, \mathit{geniculate\_ganglion})$$

describes a segment of the facial nerve from the point where the stapedius nerve splits off to the point where the geniculate ganglion develops. Anatomical structures and signs were related to each other by the axiom:

$$\forall x \forall y (\mathit{Lesion}(x) \wedge \mathit{Connected}(y, x) \rightarrow \mathit{Lesion}(y))$$

meaning that a lesion at level x includes all the signs of a lesion at the lower level y. This is a characteristic feature of neurological problems. Next it is straightforward to specify the relationship between a lesion at a certain level and the specific anatomical structures that will be affected by this lesion. For example, a lesion of the facial nerve at the stapedius nerve will affect the stapedius muscle, causing the patient to perceive all sound as too loud (hyperacusis):

$$\mathit{Lesion}(\mathit{stapedius\_nerve}) \leftrightarrow \mathit{Affected}(\mathit{stapedius\_muscle})$$
$$\mathit{Affected}(\mathit{stapedius\_muscle}) \leftrightarrow \mathit{Complaint}(\mathit{hyperacusis})$$

[Hammond et al., 1993] describes the application of anatomical reasoning in the design of removable partial dental prostheses, called removable partial dentures, within the RaPiD system. The design process is carried out under the guidance of general design rules which the designer is not allowed to break. The design rules are represented as *logical integrity constraints*, which are syntactically indistinguishable from non-Horn clauses of the form [Kowalski, 1979]:

$$A_1, \ldots, A_m \leftarrow B_1, \ldots, B_n$$

(A_i and B_j are atomic formulas) which is equivalent to

$$\leftarrow B_1 \wedge \cdots \wedge B_n \wedge \neg\exists A_1 \wedge \cdots \wedge \neg\exists A_m$$

where $\exists A_i$ denotes the existential closure of the variables occurring in an atom A_i. Let KB denote the knowledge base comprising the prosthesis design knowledge, I the collection of integrity constraints and U the input of the user. In the RaPiD system, the integrity constraints are simply represented as (generalized) Prolog clauses, yielding:

$$\texttt{:-}\ B_1, \ldots, B_n, \texttt{not}\ A_1, \ldots, \texttt{not}\ A_m.$$

i.e. logical negation is taken as negation by finite failure. For example, the following integrity constraint

$$\begin{aligned} \text{IC12} : \leftarrow & \mathit{tooth}(\mathit{Tooth}, \mathit{present}), \\ & \mathit{in\_arch}(\mathit{Tooth}, \mathit{Arch}), \\ & \mathit{centre\_position}(\mathit{Tooth}, \mathit{Point}), \\ & \mathit{not\ on\_arch}(\mathit{Point}, \mathit{Arch}). \end{aligned}$$

expresses that a tooth can only drift along the arch of the jaw. The set of clauses $\text{KB} \cup I$ is consistent. When using the RaPiD system the design made by the dentist, U, is checked agains the logical integrity constraint I by determining whether or not $\text{KB} \cup I \cup U$ is inconsistent, which can be simply done by the Prolog interpreter. For example, integrity constraint IC12 is applied automatically in the design process to ensure that the artificial tooth that is manipulated by the designer does remain on the arch of the jaw. This work may, in fact, be viewed as an special instance of *consistency-based diagnosis* [Reiter, 1987; de Kleer et al., 1992]. In consistency-based diagnosis, data are compared to the logical specification of a system; inconsistency is interpreted as an indication that something is wrong with the system. In contrast with consistency-based diagnosis, here the conclusion is that something is wrong with the data.

It is instructive to compare the logic work discussed above with the algebraic specification of anatomical knowledge. [Todd, 1994] describes a formal model of a program that assists in the localization of nerve lesions in the upper limb. The anatomical model underlying the program was developed using the specification language Z, [Spivey, 1989], a language based on set theory. In comparison with the logical specification of anatomical knowledge, as proposed in [Lucas, 1993], the applied algebraic techniques in the work of Todd render it rather easy to express how the anatomical model can be applied as part of a reasoning method, such as diagnosis. The specification of reasoning behaviour using logic is less straightforward, because it requires the development of a logical meta-theory (cf. Section 4). However, the logical specification of the meaning of objects in terms of relationships among objects, often yields semantically appealing results, which may be taken as an advantage of logic over algebraic techniques. Furthermore, meta-theory design is an inherent aspect of knowledge-based system design using logical specification language, such as $(\text{ML})^2$ [Van Harmelen & Balder, 1992].

3.2.3 Taxonomic knowledge

Having as a concept its origin in biology, taxonomic knowledge playes an important role in medical problem solving. This kind of knowledge is used in the description of diseases and disease categories, and also applied to describe human anatomy. The application of the hierarchical relationships among diseases in logical formulations of diagnostic problem solving will be discussed below. The logical meaning of taxonomic (also referred to as hierarchical) knowledge has been extensively studied, and has (at least for uncomplicated cases) been settled for some time [Hayes, 1979; Lucas & Van der Gaag, 1991]. One of the problems of using the standard logical representation of taxonomic knowledge, as introduced in [Hayes, 1979], is that in medicine every general description of a medical concept knows several exceptions.

The representation of exceptions in a logical framework requires non-standard logics, such as default logic [Besnard, 1989]. Solutions to this problem have been developed for the related subjects of semantic nets ([Touretzky, 1986]) and frames ([Lucas & Van der Gaag, 1991]). The precise extent of the problem in medicine has not been explored as yet.

It is worth noting that the logical representation of taxonomic knowledge yields implications of the form:

$$\forall x(C_1(x) \rightarrow C_2(x))$$

expressing that C_1 is a specialisation of C_2 (or C_2 is a generalisation of C_1). Such formulas overload the meaning of logical implication as a causal relation discussed in Section 3.2.1. Hence, although representing causality by means of logical implication is attractive, there are undesirable effects resulting from this representation, in addition to those mentioned in Section 3.2.1.

3.2.4 Heuristic knowledge

Heuristic medical knowledge concerns the imprecisely known relationships between patient and laboratory findings and general or specific diagnostic or treatment-decision categories, based on practical experience [Buchanan & Shortliffe, 1984; Clancey, 1985]. The typical heuristic, rule-based expert system, although usually described in terms of a production system, may in fact also be viewed as a logical theory with some extralogical features. Translation of a rule-based system to a logical theory may be useful to get insight into the precise, logical meaning of the medical concepts represented in the original knowledge base. Some research has been carried out, using logical techniques, concerning the analysis of existing medical expert systems. An important consequence of undertaking a logical analysis of an existing, heuristic medical knowledge base is that it forces the developers of a medical system, including the medical professionals involved, to think carefully about the medical knowledge that is represented. In [Bezem, 1988; Lucas & Van der Gaag, 1991; Lucas, 1993] it is discussed how a knowledge base of a rule-based system can be translated into first-order logic. [Moser & Adlassnig, 1992] discuss the translation to first-order logic of a medical expert system based on binary associations between diseases and findings.

The logical analysis of a medical knowledge base can also be advantageous in that the result of the logical reformulation can be checked with respect to its consistency. This approach may help in discovering inaccuracies in the knowledge base, and thus support the process of building a reliable medical expert system. Such consistency tests were carried out for the logical reformulation of the HEPAR system, a rule-based expert system for the diagnosis of disorders of the liver and biliary tract [Lucas & Janssens, 1991], and the CADIAG-1/BIN system, a medical expert system in the broad domain of internal medicine [Adlassnig et al., 1985].

In heuristic, medical expert systems, a knowledge base represents knowledge that holds for a class of patients. This observation gives rise to the logical formulation of heuristic rules using a universally quantified variable over the domain of patients, expressing that a rule applies to an entire class of patients. An example of such a heuristic rule concerning the disorder of the liver called Wilson's disease, taken from the logical reformulation of the HEPAR system, is the following:

$$\forall x(\mathit{Duration}(x, \mathit{complaint\_lab}, \mathit{chronic}) \wedge$$

$(disorder(x) = hepatocellular) \wedge$
$(age(x) < 25) \wedge$
$(caeruloplasmin(x, labresult, biochemistry) > 20) \wedge$
$(urinary\_copper(x, labresult, biochemistry) > 1)$
$\rightarrow Diagnosis(x, Wilson\text{'}s\_disease))$

Translation of the HEPAR system into predicate logic yields a collection of logical implications, which, after automatic conversion to clausal form, could be fed into a resolution-based theorem prover to check for the consistency of the knowledge base after input of data of a particular patient. This way a number of inconsistencies in the original knowledge base could be discovered [Bezem, 1988].

The work of Moser and Adlassnig in the logic translation of CADIAG-1/BIN reveals another, elegant, use of the application of logic in medical expert systems [Moser & Adlassnig, 1992]. In the CADIAG-1/BIN system, medical knowledge is described in terms of a fixed collection of binary relations. For example, the binary relation *oc* (*o*bligatory occurrence and *c*onfirmation) expresses that a particular finding f is necessary and sufficient for accepting the occurrence of a disease d in the patient. In medicine, such findings are usually referred to as pathognomonic. This is denoted as

$$f \text{ oc } d$$

placing the relation symbol 'oc' in infix position. For example,

'*Kayser-Fleischer rings*' oc '*primary biliary cirrhosis*'

Kayser-Fleischer rings are brown rings at the periphery of the cornea, caused by deposits of copper. These are typical for the liver disease primary biliary cirrhosis (PBC); observing Kayser-Fleischer rings is necessary and sufficient for diagnosing PBC.

The translation of Moser and Adlassnig involved an in-depth analysis of the meaning of the binary relations in CADIAG-1/BIN. For example, the logical meaning of the 'oc' relation was defined as:

$$\forall x \forall y (x \sqsubseteq y \wedge y \sqsubseteq x)$$

where the binary predicate symbol $\sqsubseteq$ stands for non-empty set inclusion, i.e. $c_1 \sqsubseteq c_2$ fails to hold if c_1 stands for the empty set. Hence, in the expression f oc d stands f for the set of patients with finding f, and d for the set of patients with disease d. For example, the meaning with respect to the patients with primary biliary cirrhosis is: all patients with Kayser-Fleischer rings are patients suffering from PBC, and vice versa; there is at least one such patient.

The other binary relations were analysed in a similar fashion, yielding similar, though more complicated, logical expressions as for the 'oc' relation. A generic, logical specification of the meaning of these binary relations yielded a short specification of a consistency checker in Prolog, which was subsequently applied for the analysis of the knowledge base. Some very subtle medical inconsistencies were detected by this remarkably simple method [Moser & Adlassnig, 1992].

3.2.5 Physiological knowledge

(Patho)physiological models can be applied in anaesthesiology in monitoring and controling the physiological state of patients undergoing surgery. These kinds of model have a long standing tradition in medicine, having their roots in control theory of biological systems [Grodins, 1963].

(Patho)physiological models usually consist of algebraic equations, where some of the variables in the equations can be measured in the patient, and others cannot, but instead must be computed from the equations and the known variables. Because the values of the variables, such as the blood pressure in a patient, evolve over time, i.e. the variable is a function of time, the notion of time is often an essential ingredient of such models. Qualitative reasoning is the field traditionally concerned with the study of formal techniques for the specification of, and reasoning with, such models, where it is assumed that the (numerical) values of the variables are only roughly known [Coiera, 1992]. For example, in the QSIM approach proposed by B.J. Kuipers, the time-dependent variation in the value of a variable is approximated by the notion of *qualitative state* [Kuipers, 1986; Kuipers, 1994]. In a qualitative state, a discrete qualitative value (called a landmark value) is assigned to a variable for a particular time-point, including the direction towards which the variables has changed, such as 'increased', 'steady' and 'decreased'. The modelling of state change in time is a central feature of qualitative reasoning. Because the QSIM theory is based on the qualitative interpretation of differential equations, the theory is essentially algebraic and not logical in nature. Most other work in qualitative reasoning in medicine based on differential equations is beyond the scope of the present paper. In [Coiera, 1990], the application of QSIM for the generation of disease histories as part of the diagnostic process of acid-base disturbances, is described. In this work, QSIM has been incorporated in a logic programming framework. For a review of the subject of qualitative reasoning, including other approaches than QSIM, the reader is referred to [Coiera, 1992].

A QSIM-like model consists of a qualitative description of a system of differential equations. Hence, little detail with respect to the values of the variables is required. Nevertheless, in medicine even a qualitative description of functional behaviour in terms of QSIM can be too demanding. Even if such information is available, the technique may not be at the right level of abstraction for representing a problem domain. Furthermore, the generality of the qualitative algebra on which QSIM is based, may lead to inefficient or spurious problem-solving behaviour. These were the main reasons for adopting an approach different from QSIM in the design of the KARDIO system, an expert system for the diagnosis of disturbances of the rhythm of the heart, known as cardiac arrhythmias [Bratko et al., 1989]. KARDIO is the best-known example of a successful logic engineering approach to qualitative reasoning in medicine, based on a functional domain model.

Basically, KARDIO's knowledge base consists of a logical formalization of a qualitative simulation model of the (normal and abnormal) electrical activity of the heart. The simulation model can be triggered by the assumption of the presence of a particular (combination of) cardiac arrhythmias in the patient. This sets up a chain of events, finally leading to a collection of findings which represents an electrocardiogram (ECG), corresponding to the cardiac arrhythmias assumed to be present. The various activation steps are represented by means of rules, such as, for example

```
heart(atrial_focus:permanent(Rhythm, Rate))
⇒
```

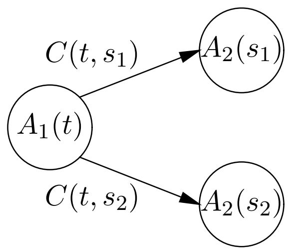

Figure 2: Non-deterministic state transition.

```
permanent(atrial_focus:form(Origin, Rhythm, Rate)) when
atrial_focus(Origin, Rhythm, Rate)
```

which expresses that if there is an assumed permanent centre of (abnormal) autonomous activity in the atrium (an atrial focus), which reveals a particular rhythm (e.g. regular or irregular) and rate, then the atrial focus with particular origin is accepted to be permanent, if a medical description of such an atrial focus declares it to be a valid possibility. (As in Prolog, strings starting with an upper-case letter denote variables.) Note that the adopted representation is record-like, because '`heart`' represents an object and '`atrial_focus`' represents an attribute (cf. Section 3.1). The expression

```
atrial_focus(Origin, Rhythm, Rate)
```

following the `when` keyword, states a *restriction* that must be fulfilled before the conclusion,

```
permanent(atrial_focus:form(Origin, Rhythm, Rate))
```

can be assumed. In general, rules have the following logical meaning

$$\forall x \exists y (A_1(x) \rightarrow (A_2(y) \wedge C(x, y))) \tag{1}$$

where $A_i(t)$, $i = 1, 2$ represent states of the heart that can be assumed, and $C(x, y)$ represents a restriction that relates two (possibly more in general) states to each other. The logical implication expresses that if some state $A_1(t)$ is assumed, and, for example, $C(t, s_1)$ and $C(t, s_2)$ hold, the disjunction

$$A_2(s_1) \vee A_2(s_2)$$

can be derived. This expresses two alternative states that can be activated by the state $A_1(t)$, where at least one of the two, $A_2(s_1)$ or $A_2(s_2)$, must actually occur (this is like a non-deterministic state transition, see Figure 2). The atomic formula $C(x, y)$ in formula (1) restricts the disjunction with respect to the arguments of the predicate A_2. If $C(x, y)$ had been left out, the more general conclusion $\exists y A_2(y)$ would have been reached in the present case [Bratko et al., 1989]. If no restrictions $C(t, s)$ are satisfied, then adopting the CWA, $\forall y \neg C(t, y)$ is assumed, and an inconsistency will arise. By removing $A_1(t)$ from the assumptions, it is possible to restore consistency. The derivation of a disjunction with respect to A_2 using formula (1) above will then be blocked.

Based on the collection of rules, denoted by KB, the satisfied restrictions $\mathcal{C}$ and state assumptions $\mathcal{A}$, that stand for cardiac arrhythmias that are assumed, KARDIO is capable of predicting alternative ECG descriptions (expressed as a disjunction of individual descriptions),

$$\mathrm{KB} \cup \mathcal{C} \cup \mathcal{A} \vdash \mathrm{ECG}$$

```
module cardiac_output_model(number : SV, HR, CO, KL, PLA).

table intermittent(CO).
table continuous(HR, PLA).
table unknown(SV).
table constant(KL).

clause constraints(CO, SV, HR, KL, PA) when
       CO = SV * HR and
       SV = KL * ln(PLA).
close.
```

Figure 3: Logic specification of constraints.

where $\mathrm{KB} \cup \mathcal{C} \cup \mathcal{A}$ is assumed to be consistent. This may be viewed as a qualitative simulation process. In Section 4.3, we shall briefly deal with the diagnostic interpretation of the rules in KARDIO.

Where the approaches above use a qualitative model of a medical domain, there are certain situations where sufficient numerical patient data are available. The equipment used in anaesthesiology collects a large amount of numerical information during surgery, that is vital for monitoring the patient's condition. These numerical data can also be used to drive the algebraic equations of a (patho)physiological model. Modern constraint logic programming languages make it possible to represent and manipulate such algebraic equations in a declarative way [Van Hentenryck, 1991; Jaffar & Maher, 1994]. RL, [Van Denneheuvel et al., 1990], is a constraint logic programming and database language that has been applied for this purpose. In [De Geus et al., 1991], the application of constraint logic programming to a physiological model of human circulation and gas exchange, represented and applied by means of RL, is described. We shall briefly discuss this work.

Consider the, highly simplified, logic specification of the function of the heart as a pump, in the syntax of the RL language (this specification follows [De Geus et al., 1991]) shown in Figure 3. Here, `CO` stands for the cardiac output (blood volume expelled by the heart per unit of time), `SV` for the stroke volume (blood volume expelled by the heart by a single contraction of the heart chambers), `HR` for the heart rate, `PLA` for the pressure in the left atrium, and `KL` is a parameter called the 'left heart parameter'. These variables are distributed over four database tables, or relations. The '`intermittent`' table stands for the variables measured only a limited number of times during surgery, whereas the '`continuous`' variables `HR` and `PLA` are measured continuously in the patient. The variable `SV` is not measurable, but can be computed. Finally, there are also variables that can be considered constants with respect to the patient, for example `KL`. A constraint clause, called `constraint`, represents equality constraints among the variables; the clause can be used to compute values for unknown variables, based on values of known variables. It is also possible to reduce (simplify) the model equations, based on known values for variables. The specification above can be used for:

- *interpretation* of measurements, i.e. computation of the unknown variables based on measured values of known variables;

- *prediction* of effects, i.e. based on particular additional assumptions about the variables, values for certain variables can be predicted.

In the following query, values for known variables are used for computing the values for the requested variables `CO` and `SV`, which are arguments to a call to the constraint solver `syminfer`, indicating that these are the variables to be determined:

```
syminfer(CO,SV) when constraint(CO, SV, HR, KL, PA) and
                HR = 70 and
                PLA = 16 and
                KL = 28.

solutions = {CO = 5432, SV = 77.6}
```

The constraint solver `syminfer` yields a symbolic solution, i.e. the requested variables are symbolically expressed as a function of the known variables. In this case, the set of equations is completely reduced to a set of variables with constant values.

In practical situations, the known variables are subject to measurement errors; this may give rise to logical inconsistency of the domain model when those variables are entered. Such errors can be simply accommodated for in a constraint model by introducing error variables, e.g.

```
PLAerror = 16  - PLA.
```

replaces the constraint '`PLA = 16`' above. In this way, inconsistency can be resolved; values for the error variables will be computed automatically. Note that a limitation of this approach is that only a snapshot of the physiological state of a patient is interpreted; no explicit temporal labeling of the information has been incorporated in the model.

Although there is currently no approach for representing and reasoning about time generally agreed upon, some researchers have attempted to actually exploit the notion of time in a medical application. Specific applications of temporal reasoning in medicine are discussed in Section 4.2 with respect to treatment management.

4 Medical problem solving

In the previous section, we have reviewed the declarative representation of medical knowledge in logic, and some remarks were made concerning the application of the logical representations. Such representations can be used for a variety of medical tasks, such as diagnosis, treatment selection and management, prognostic assessment, patient monitoring, etc. In early medical expert systems, the application of domain knowledge was carried out by a, more or less, general purpose inference engine. For some of the aforementioned representations and domains, domain-independent reasoning methods are appropriate. However, in many medical problem solving situations, medical knowledge is applied in a quite specific way. The specification of such problem solving methods requires a precise, unambiguous formalism. A disadvantage of algorithmic specification techniques is that it is impossible to describe the reasoning in terms of the logical specification only. Instead, the reasoning is described in terms of operations on data elements which are the result of a translation of the logical representation into algorithmic terms. Logical techniques are better suited for this purpose, because reasoning can be described in a declarative way, i.e. in terms of relations among logic formulas [Genesereth & Nilsson, 1987]. The logic specification of medical problem solving methods is an important subject of research, and is complementary to the declarative specification of object knowledge, discussed in the previous section.

4.1 Decision knowledge and task models

The (formal) specification of the specific way in which a medical task, such as diagnosis, is carried out is often referred to as a *task model*. A task model can be viewed as part of the meta-layer of a knowledge base. Work on the declarative specification of task models using logic has been carried out by J. Fox et al. in the context of the Oxford System of Medicine (OSM) [Clark et al.,1990; Fox et al., 1990b], and the DILEMMA system [Huang et al., 1993].

As has been mentioned in Section 3.2.1, the DILEMMA object-knowledge is represented using the meta-relation

$$Domain(O, R, O')$$

where O, O' represent concepts (objects) in the domain, and R is the name of a binary relation in which the objects participate. Specific patient data are represented by means of the relation *Patient_record*, containing all information at a particular time point for one patient. The tasks that can be used to solve medical problems, as well as a declaration of the specific manner in which this task is used, called the *role* of the task, are also represented at the object-layer. A task is applied to a relation represented at the object-layer. For example

$$Task(refining, diagnosis, may\ be\ caused\ by)$$

indicates that in the process of refining a candidate diagnosis, use is made of the 'may be caused by' object-relation. The predicate symbol *Domain* is, like *Task*, a meta-predicate, i.e. its meaning is defined at the meta-layer. Another example of a task, besides 'diagnosis' is 'treatment'; another example of a role of a task besides 'refining' is 'proposing' (a candidate for a diagnosis), etc.

The actual knowledge of how the tasks must be combined to solve a medical problem, is part of the meta-layer, not of the object-layer. For example, in the process of diagnosis, first a candidate diagnosis is proposed by means of the 'proposing' role of the diagnosis task; after that the diagnosis candidate may be refined.

Although the meaning of predicate symbols, such as *Domain* and *Task*, could be defined at the meta-layer, again using predicate logic, the developers of the DILEMMA decision support system have adopted a more pragmatic approach by expressing the meaning of the object-layer knowledge by means of first-order schemas, which can be viewed as as special-purpose (meta-layer) inference rules. The syntax and meaning of the object-layer has been described in Section 3.2.1.

For example, the schema depicted in Figure 4 provides a declarative specification of the reasoning method that is used for proposing a candidate solution, e.g. a candidate diagnosis as part of diagnostic problem solving. The schema consists of a name, *propose_candidate*, an antecedent consisting of three conditions, and a consequent. The condition

$$Context(Taskname, Focus)$$

represents the problem-solving context, which consists of a particular task, such as *diagnosis* (which is substituted for the meta-variable *Taskname*), that is carried out, and the particular propose_candidate

Context(Taskname, Focus),
Task(proposing, Taskname, Relation),
Domain(Focus, Relation, Candidate)

Candidate_proposed(Candidate, context(Taskname, Focus))

Figure 4: Proposing inference rule.

finding or disease that is being considered, substituted for the meta-variable *Focus*. The condition

$$Task(proposing, Taskname, Relation)$$

expresses that the particular task is carried out using medical knowledge specified by means of *Relation*. Finally, all knowledge that is used to solve the problem, concerning the relevant *Relation* is part of the object knowledge, represented using the *Domain* predicate. The consequent expresses that the particular candidate solution, e.g. a diagnosis, has been found by means of the given *Taskname* using a relation with respect to *Focus*. For example, suppose that the system has just confirmed that the patient suffers from *cardiac_ischaemia*, which is expressed by:

$$Context(diagnosis, cardiac\_ischaemia)$$

Then, based on this information the diagnostic process proceeds using causal knowledge:

$$Task(proposing, diagnosis, causes)$$

Since the object-layer contains the fact

$$Domain(cardiac\_ischaemia, causes, MI)$$

myocardial infarction (MI) is proposed as a candidate diagnosis:

$$Candidate\_proposed(MI, context(diagnosis, cardiac\_ischaemia))$$

Other medical tasks, such as treatment, follow-up and prognostic assessment, are represented using the same kind of schemas. For example, taxonomic knowledge concerning disorders is used as part of diagnostic problem solving in order to refine a solution, such as the refinement of a diagnosis from a disease category to a more specific disease. The specific inference schema employed is depicted in Figure 5.

4.2 Treatment management

Decisions concerning the choice of the appropriate treatment for a patient, involve various sorts of medical knowledge. One of the first systems by which treatment planning was studied in depth, is ONCOCIN, a well-known treatment management system for medical care of oncology patients [Hickam et al., 1985]. No attempts were made in the ONCOCIN project to refine_candidate

Context(Taskname, Focus), Candidate_proposed(Candidate, context(Taskname, Focus)), Task(refining, Taskname, Relation), Domain(Candidate, Relation, Sub_candidate)
Candidate_proposed(Sub_candidate, context(Taskname, Focus)), Candidate_refined((Candidate, Sub_candidate), context(Taskname, Focus))

Figure 5: Refining inference rule.

analyse the treatment knowledge encoded in the system using logical techniques. Research with respect to the logical aspects of treatment management has commenced only recently. Treatment selection has mainly been studied in the field of medical decision analysis. We shall not go into details here, but only mention that several of the problems that have been addressed in medical decision analysis, have as yet not been studied in depth with respect to their logical meaning.

The OaSiS system, an expert system for decision support in oncology much like the ONCOCIN system, has been developed using techniques from logic programming [Hammond et al, 1994; Hammond & Sergot, 1995]. One of the central aspects of treatment planning for cancer patients receiving chemotherapy is that the treatment may need adjustment, based on therapy results and side effects of administered drugs. For example, the number of white blood cells may decrease considerably under a regime of chemotherapy, requiring lowering of the dosage of certain cytotoxic drugs. Medical knowledge concerning the automatic adjustment of the dosage can be expressed in logic form as a number of equations, and applied using retrieved laboratory findings of the patient.

A more difficult application of techniques from logic programming to treatment management concerns checks for patient conditions or drug interactions that may diminish the efficacy of the treatment. Logical integrity constraints of the form introduced in Section 3.2.2 have been shown to provide a suitable representation for medical knowledge concerning forbidden conditions and actions [Hammond et al, 1994; Hammond & Sergot, 1995]. For example, the following integrity constraint:

$$\begin{array}{l}\leftarrow \mathit{user\_suggestion}(\mathit{perform}(\mathit{Action1}, \mathit{Plan})),\\ \quad \mathit{part\_of}(\mathit{Action2}, \mathit{Plan}),\\ \quad \mathit{produces\_effect}(\mathit{Action2}, \mathit{Effect}),\\ \quad \mathit{hazardous}(\mathit{Effect}),\\ \quad \mathit{aggravates}(\mathit{Action1}, \mathit{Effect}),\\ \quad \mathit{is\_avoidable}(\mathit{Action1}, \mathit{Plan}).\end{array}$$

expresses that if treatment action *Action1* is part of the designed treatment plan (*Plan*), and a second treatment action (*Action2*) will have side effect *Effect*, and this side effect is made worse by *Action1*, then *Action1* should be avoided as part of the treatment plan.

Treatment planning always involves some reasoning about time, which raises the issue of appropriate logical formalization of time. Some researchers have chosen for the representation of time within a standard logic framework. For example, in [Soper et al., 1991], the application of temporal knowledge for resource management in patients undergoing vascular surgery is discussed. Their approach is based on event calculus, which remains within standard logic [Kowalski & Sergot, 1986]. A disadvantage of this approach is that it does not really do justice to the special status of time. Although part of temporal reasoning can be incorporated in standard logical reasoning, there is much to say for separating out the time-dependent information, and using temporal logic for reasoning about treatment adjustment in time explicitly. Temporal logic goes beyond standard logic in that the same logical formula may have a different interpretation at different times [Turner, 1984]. Theoretical work on temporal and deontic logic (the modal logic of permission and obligation), and modal logic more generally, may provide a formal underpinning for the explicit manipulation of knowledge concerning the temporal and obligatory aspects of treatment planning [Das et al., 1993]. A technique to extend Prolog for temporal reasoning, and which would be interesting to investigate in medical applications, is described in [Hrycej, 1992].

4.3 Diagnostic reasoning

There are two different, frequently applied logical models of diagnostic reasoning in medicine; diagnosis is either described as a deductive process, or as an abductive reasoning process [Charniak & McDermott, 1985].

In the *deductive model* of diagnostic reasoning, a medical knowledge base KB and a collection of patient data F, are used to derive a diagnosis D

$$\mathrm{KB} \cup F \vdash D$$

where, for example, D is a conjunction of the form

$$\mathit{Diagnosis}(\mathit{patient1}, \mathit{Wilson's\_disease}) \wedge \cdots \wedge \mathit{Diagnosis}(\mathit{patient1}, \mathrm{PBC})$$

and KB might consist of heuristic rules similar to those discussed in Section 3.2.4. Diagnostic problem solving is often described using the hypothesize-and-test paradigm; this is expressed by the following formulation, equivalent to the one above:

$$\forall d \in D: \ \mathrm{KB} \cup F \cup \{\neg d\} \vdash \bot$$

where the negation of the disorder d is the logic representation of a hypothesis; the refutation indicates that the hypothesis has been rejected, i.e. the disorder d has been confirmed. If resolution is the principal inference rule applied, then a hypothesis-driven reasoning strategy can be designed in terms of a resolution strategy. The set-of-support strategy with negative hyperresolution for non-Horn theories, [Wos et al., 1992], or SLDNF resolution for generalized Horn-clause theories, [Lloyd, 1987; Lucas & Van der Gaag, 1991], are examples of strategies displaying such a reasoning behaviour. The former techniques have been applied in the logical reformulation of the HEPAR system [Lucas, 1993]; the latter have been applied in KARDIO [Bratko et al., 1989]. In the KARDIO system, for the purpose of diagnosis, the simulation model was converted to a diagnostic knowledge base by interchanging the conditions and conclusions in rules (cf. formula (1) in Section 3.2.5). A diagnosis D for a given ECG in a patient is a collection of cardiac arrhythmias, deductively obtained as follows

$$\mathrm{KB} \cup \mathcal{C} \cup \mathrm{ECG} \vdash D$$

where ECG represents the logical description of the patient's electrocardiogram, and $\mathcal{C}$ are conditions that must be fulfilled.

In the *abductive model* of diagnosis, such as the model proposed by L. Console and P. Torasso, the pathological behaviour of a biological system is specified in terms of cause-effect relationships as discussed in Section 3.2.1. More restrictive models of diagnosis, based on set theory instead of logic, are described in [Peng & Reggia, 1990], [Josephson & Josephson, 1994] and [Wu, 1991]. Diagnostic problem solving is described as the problem of accounting for a given set of observed patient findings F by supplying the knowledge base KB with a set (conjunction) of hypotheses D which, after computation of the deductive closure, accounts for each of the given observed findings. Formally, abductive diagnosis is defined as follows. Let

- KB be a logic specification of causal, medical knowledge (usually it is assumed that KB is a Horn clause theory);
- $F = F_p \cup F_n$ be a set of observed patient findings, with F_p denoting the present, abnormal findings – described by ground, positive unit formulas – and F_n the findings assumed or entered to be absent –described by ground, negative unit formulas ($F_p \cap F_n = \varnothing$).

A *diagnosis* D is then a set of (ground) disorder literals, i.e. literals expressing medical disorders, such as $\mathit{Disorder}(\mathit{patient1}, \mathit{myocardial\_infarction})$, and assumption literals α, such that [Console & Torasso, 1990a; Konolige, 1992]:

1. $\forall f \in F_p : \mathrm{KB} \cup D \vdash f$;
2. $\forall f \in F_n : \mathrm{KB} \cup D \cup \{f\} \nvdash \bot$ (or, equivalently, $\mathrm{KB} \cup D \cup F_n$ is consistent).

Hence, a diagnosis D *covers* (condition 1 above) all observed (positive) findings in F_p, and is *consistent* (condition 2 above) with all (negative) findings included in the set F_n. Often, it is required that D is subset minimal, i.e. D is a minimal conjunction of disorder and assumption literals for which the covering and consistency conditions above are satisfied. This means that none of the disorders can be left out from the set D without sacrificing the property that every finding can be accounted for. Subset minimality, and similar criteria, are domain-independent measures, that are applied to reduce the number of alternative diagnoses that can be produced [Peng & Reggia, 1990; Tuhrim, 1991]. They fall short in providing a medically relevant reason for selecting or disregarding certain diagnoses. Disorders that display considerable overlap with other disorders may not appear in any subset-minimal diagnosis. This is hardly acceptable from a medical point of view. In the following, we shall not pay further attention to such domain-independent selection criteria.

The set F_n is defined as follows:

$$F_n = \{\neg R(c) \mid R(d) \in F, c \neq d\} \cup \{\neg R'(e) \mid \neg R'(e) \text{ has been entered by the user}\}$$

where R, R' stand for predicates representing particular data groups, for example a finding corresponding to the result of a test or of the medical interview. Only observable findings for which at least one result has been obtained by performing a particular diagnostic test, are included in F_n. The other findings are assumed to be unknown. Note that a weak form of the closed world assumption with respect to findings is taken to hold in this case (cf. Section 3.1). It is interesting to note that the deductive method of diagnosis employed in KARDIO can be viewed as a form of abduction due to the interchange of state conditions

$$\forall x\,(Pain(x, cardiac\_ischaemic) \rightarrow Complaint(x, oppressive\_pain))$$
$$\forall x\,(Pain(x, cardiac\_ischaemic) \rightarrow Complaint(x, pain\_radiating\_to\_left\_arm))$$
$$\forall x\,(Disorder(x, myocardial\_infarction) \rightarrow Pain(x, cardiac\_ischaemic))$$
$$\forall x\,(Disorder(x, myocardial\_infarction) \rightarrow Sign(x, vomiting))$$
$$\forall x\,(Disorder(x, heart\_failure) \rightarrow Complaint(x, dyspnoea))$$
$$\forall x\,(Disorder(x, heart\_failure) \rightarrow Sign(x, cyanosis))$$
$$\forall x\,(Disorder(x, heart\_failure) \rightarrow Sign(x, pitting\_edema))$$

Figure 6: Diagnostic causal model.

and a state conclusion. The available literature (e.g. [Bratko et al., 1989]) does not provide sufficient detail to establish whether or not a diagnosis in KARDIO satisfies the covering and consistency conditions. It seems likely that it does not, because not all rules in KARDIO have a causal or state-transition reading.

A disadvantage of the abductive model of Console and Torasso is that all observed findings F_p must be explained by the covering condition. The underlying assumption is that all possible disorders that may cause the particular findings in the medical domain are included in the knowledge base, i.e. a form of domain closure is assumed to hold. Hence, every observed finding must be the effect of the presence of at least one of the described disorders. However, in medicine there are many situations in which this condition fails to hold. The huge number of interactions among concepts in medicine makes it feasible to represent only a small fraction of the entire body of domain knowledge. Furthermore, it is not very realistic to expect that a knowledge base will cover all fields of medicine. In practice, the assumptions of complete accuracy and exhaustiveness of a knowledge base fail to hold.

We illustrate the problem by the following example. Consider the logical specification KB presented in Figure 6, in which only strongly causal rules are included, and the following set of observed findings concerning cardiac disorders:

$$\begin{aligned} F_p = \{ & Complaint(patient1, oppressive\_pain), \\ & Complaint(patient1, pain\_radiating\_to\_left\_arm), \\ & Complaint(patient1, dyspnoea), \\ & Sign(patient1, vomiting), \\ & Sign(patient1, cyanosis)\} \end{aligned}$$

The set of absent findings is equal to:

$$F_n = \{\neg Sign(patient, pitting\_edema)\}$$

Obviously, if it is assumed that

$$D = \{Disorder(patient1, myocardial\_infarction)\}$$

then

$$\mathrm{KB} \cup D \vdash Complaint(patient1, oppressive\_pain)$$

and

$$\mathrm{KB} \cup D \cup F_n \nvdash \bot$$

but

$$\mathrm{KB} \cup D \nvdash \mathit{Complaint}(\mathit{patient1}, \mathit{dyspnoea})$$

Hence, D is not a diagnosis, because not all findings in F_p are covered, although D is consistent with KB and F_n. On the other hand, if $\mathit{Disorder}(\mathit{patient1}, \mathit{heart\_failure}) \in D'$, then

$$\mathrm{KB} \cup D' \cup F_n \vdash \bot$$

although the covering condition might be fulfilled. Hence, in the present case no diagnosis exists. Intuitively, this result may be taken as warning that if the knowledge base is correct, the given combination of findings F_p cannot have been observed. Unfortunately, if the knowledge base is not completely accurate, then the abductive model gives up too early.

From a medical point of view, a more realistic approach to abductive diagnosis in the medical domain would be to cover as many as possible of the findings in the set of observed abnormal findings F_p, i.e. $F_m \subseteq F_p$, such that

$$\mathrm{KB} \cup D \vdash F_m \tag{2}$$

and

$$\mathrm{KB} \cup D \cup F_n \nvdash \bot \tag{3}$$

with the original set of negative findings F_n. Now, if the set F_m in the covering condition (2) is equal to:

$$\begin{aligned} F_m = \{ & \mathit{Complaint}(\mathit{patient1}, \mathit{oppressive\_pain}), \\ & \mathit{Complaint}(\mathit{patient1}, \mathit{pain\_radiating\_to\_left\_arm}), \\ & \mathit{Sign}(\mathit{patient1}, \mathit{vomiting})\} \end{aligned}$$

where $F_m \subseteq F_p$, it follows that the diagnosis is equal to

$$D = \{\mathit{Disorder}(\mathit{patient1}, \mathit{myocardial\_infarction})\}$$

and the covering and the consistency conditions are both satisfied. In contrast with the original model of Console and Torasso, it is possible to have some remaining observed findings that cannot be accounted for. This is fairly common in medical practice. Note that only the definition of the set F_p in the covering condition has been adapted. Part of the attraction of this abductive model derives from the happy circumstance that a diagnosis can be computed efficiently. A single, for example, subset-maximal diagnosis D that disregards no relevant disorder, can be computed in polynomial time, where the original abductive problem, in which alternative diagnoses are computed, is NP hard [Bylander et al., 1992].

Still, a disadvantage of this adapted abductive model of diagnosis is that a disorder may be ruled out by the assumption of the absence of certain findings, included in the set F_n, which for certain medical domains might place too much emphasis on *assumed* negative findings. Doing away with the CWA for observed findings, thus accepting the open-world assumption for the knowledge base, yields yet another abductive diagnostic model. The set F_n only comprises negative findings entered by the user; the definition of the consistency condition, however, is left unchanged. Given this definition of the set F_n, there are good reasons to modify the definition of the covering condition, as follows

1′. $\forall d \in D \exists f \in F_p$: $\mathrm{KB} \cup \{d\} \vdash f$, where D is maximal.

This version of the covering condition expresses that only disorders having at least one associated finding in common with the set of observed findings, may be included in a diagnosis.

Above we have discussed deductive and abductive models of diagnosis. Abductive diagnosis has been considered in some detail, because many researchers believe that medical diagnosis is essentially abductive in nature [Josephson & Josephson, 1994; Peng & Reggia, 1990; Pople, 1973]. As we have argued, abductive diagnosis should not be taken as a fixed concept. Several different notions of abductive diagnosis, based on the nature of the medical domain, can be designed. Both the characteristics of the patient data and the medical knowledge base are important in this respect. For example, adoption of the open-world assumption for observed findings might be a good design option in medical emergency situations, where there is insufficient time to collect all data; the closed-world assumption might be taken to hold in the clinic. In [Console & Torasso, 1990b; Console & Torasso, 1991a], some other variations in the definition of abductive diagnosis are discussed, but not from a medical perspective. It is also worthwhile to design a special abductive reasoning scheme, based on a detailed analysis of a medical problem, for computational reasons. It is not unlikely that the resulting scheme will be less troublesome with respect to the computational complexity than abduction in general.

5 Conclusions

In this article, we have reviewed the application of logical techniques for building knowledge-based systems in medicine, referred to as 'logic engineering'. In particular, the tailoring of logical techniques to the representation of medical declarative and problem-solving knowledge, has been reviewed. The application of logic engineering to the field of medicine brings into the picture the question of for which problems in medicine the approach would be especially suitable. It has appeared to be difficult to get the products of this, and similar, research accepted by the medical community, in spite of accumulating evidence of the value of these systems [De Dombal, 1991]. Various reasons have been suggested as an explanation for this situation, such as the inadequate computational infrastructure in health care [Shortliffe, 1991]. Medical doctors are still not accustomed to the idea of a workstation as a general problem solving tool, contrasting practitioners in many other disciplines. This situation is likely to change as soon as information technology in general has gained greater acceptance in the clinic.

Logic engineering is still mainly a research area; many of the techniques and applications discussed in the paper have been developed fairly recently. It seems, therefore, unjust to judge the merits of logic engineering by the number of applications that have left the laboratory into the clinic. Nevertheless, several of the systems briefly discussed above indicate that logic engineering has something to offer in terms of clarity, accuracy, reliability, safety, and elegance. As a language for modelling medical knowledge, logic has several advantages. Firstly, by using logic for building or analysing a medical expert system, the developer is enforced to take a closer look at the meaning of the medical concepts involved. Although logic does not shield the developer completely from semantic inaccuracies, a logic engineering approach eases the identification of such problems. Secondly, even if the semantic analysis of a medical problem is cumbersome because of the vague, empirical nature of the medical knowledge involved, logic provides sufficient freedom to leave particular aspects of a specification open.

There are several other formal approaches, that share the aforementioned features of logic engineering, in particular decision theory, which has been discussed in the introduction, and algebraic specification languages. Although, algebraic specification language, such as OBJ, ([Futatsugi et al., 1985]), and Z, ([Spivey, 1989]), offer similar advantages as full predicate logic, they generally provide little language support for the structuring of information. This turns out to be one of the nicests feature of predicate logic, as has been illustrated by several of the examples above. Such a structure is usually exploited by the reasoning process. In this respect, algebraic specification languages have little to offer. Moreover, many algebraic specification languages, such as Z, do not fully support the operationalization of a specification, because no compiler or interpreter is available. This contrasts the logic engineering approach, in which the execution of a logical specification is of central concern. Nevertheless, for some medical problems which are algebraic in nature, algebraic specification might be more suitable than other techniques [Todd, 1994; Hammond & Davenport, 1995].

Learning new medical knowledge using logical techniques, i.e. inductive logic programming, is an interesting new research direction. Inductive logic programming has been applied to the medical field using the MOBAL system, ([Sommer et al., 1994]), an extensive machine-learning environment that incorporates a variety of techniques. For example, the system has been used for learning knowledge about the treatment of boys with maldescensus testis [Morik et al., 1993; Wrobel, 1990]. Only limited experience in applying inductive logic programming to medicine exists as yet.

One of the limitations of the logic-based approach to medical problem solving is that a methodology of problem solving in the medical field is still lacking. Medical problem solving is usually only described in very global terms. For example, although in many diagnostic systems, such as the DILEMMA system, diagnostic reasoning is carried out in stages from proposing candidate solutions to refining, and finally rejecting or accepting candidates, such a theory of diagnosis is as yet not part of the core of the medical sciences. We expect that such medical expert systems will be more readily accepted by the medical community if supported by fundamental research into the nature of the medical problem solving process.

Another limitation of the current logic engineering research is that no generally accepted formal methods for constructing logic-based medical expert systems is available. This situation may improve if formal specification languages for building medical expert systems become available. In contrast with the specification languages applied in software engineering, knowledge-based specification languages are usually based on logic. Hence, they share the advantages of logic mentioned above. Presently, there is still little experience in the application of such languages in the building process, which is a subject of on-going research. Recently, Fox et al., ([Fox, 1993; Krause & Glowinski, 1993]), have investigated the use of the logic-based specification language $(\mathrm{ML})^2$ for the development of knowledge-based systems, a language that is embedded in the KADS methodology for building knowledge-based systems [Van Harmelen & Balder, 1992]. Besides formal specification, clarifying the logical meaning of medical knowledge will remain a subject of study for at least one more decade to come.

Acknowledgements. I would like to thank Peter Hammond, Ernest Rotterdam, Ameen Abu-Hanna and an anonymous reviewer for their useful comments on earlier drafts of this article.

References

[Adlassnig et al., 1985] Adlassnig, K-P, Kolarz, G, Scheithauser, G, Effenberger, W and Grabner, G, 1985. "CADIAG: approaches to computer-assisted medical diagnosis", *Comp. Biol. Med.* **15** 315–335.

[Aït-Kaci & Podelski, 1993] Aït-Kaci, H and Podelski, A, 1993. "Towards a meaning of life", *J. Logic Programming* 16 195–234.

[Aït-Kaci & Nasser, 1986] Aït-Kaci, H and Nasser, R, 1986. "LOGIN: A logic programming language with built-in inheritance", *J. Logic Programming* **3** 185–215.

[Besnard, 1989] Besnard, P, 1989. *An Introduction to Default Logic*, Springer-Verlag, Berlin.

[Bezem, 1988] Bezem, M, 1985. "Consistency of rule-based expert systems", In: E Lusk and R Overbeek (eds.), *Proceedings of the 9th International Conference on Automated Deduction*, pp 151–161, Springer-Verlag, Berlin.

[Bratko et al., 1989] Bratko, I, Mozetič, I and Lavrač, N, 1989. *KARDIO: A Study in Deep and Qualitative Knowledge for Expert Systems*, The MIT Press, Cambridge, Massachusetts.

[Buchanan & Shortliffe, 1984] Buchanan, BG and Shortliffe, EH, 1984. *Rule-based Expert Systems: the MYCIN Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley, Reading, Massachusetts.

[Bylander et al., 1992] Bylander, T, Allemang, D, Tanner, MC and Josephson, JR, 1992. "The computational complexity of abduction", In: RJ Brachman, HJ Levesque and R Reiter (eds.), *Knowledge Representation*, pp 25–60. The MIT Press, Cambridge, Massachusetts.

[Chang & Lee, 1973] Chang, CL and Lee, RCT, 1973. *Symbolic Logic and Mechanical Theorem Proving*, Academic Press, New York.

[Charniak & McDermott, 1985] Charniak, E and McDermott, D, 1985. *Introduction to Artificial Intelligence*, Addison-Wesley, Reading, Massachusetts.

[Clancey, 1985] Clancey, WJ, 1985. "Heuristic classification", *Artificial Intelligence* **27** 289–350.

[Clark et al.,1990] Clark, DA, Fox, J, Glowinski, AJ, O'Neil, MJ, 1990. "Symbolic reasoning for decision making", In: K Borcherding, OI Larichev and DM Messick, *Contemporary Issues in Decision Making*, pp 57–75, Elsevier (North-Holland), Amsterdam.

[Cohen, 1985] Cohen, PR, 1985. *Heuristic Reasoning about Uncertainty: An Artificial Intelligence Approach*, Pitman, London.

[Coiera, 1990] Coiera, EW, 1990. "Monitoring diseases with empirical and model-generated histories", *Artificial Intelligence in Medicine* **2** 135–147.

[Coiera, 1992] Coiera, EW, 1992. "The qualitative representation of physical systems", *The Knowledge Engineering Review* **17**(1) 55–77.

[Console et al., 1989] Console, L, Dupré, DT and Torasso, P, 1989. "A theory of diagnosis for incomplete causal models", In: *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pp 1311–1317.

[Console & Torasso, 1990a] Console, L and Torasso, P, 1990. "Hypothetical reasoning in causal models", *International Journal of Intelligent Systems* **5** 83–124.

[Console & Torasso, 1990b] Console, L and Torasso, P, 1990. "Integrating models of correct behaviour into abductive diagnosis", In: *Proceedings of ECAI'90* pp 160–166.

[Console & Torasso, 1991a] Console, L and Torasso, P, 1991. "A spectrum of logical definitions of model-based diagnosis", *Computational Intelligence* **7**(3) 133–141.

[Console & Torasso, 1991b] Console, L and Torasso, P, 1991. "On the co-operation between abductive and temporal reasoning in medical diagnosis", *Artificial Intelligence in Medicine* **3**(3) 291–311.

[Das et al., 1993] Das, SK, Clarke, M and Fox, J, 1993. "A logic for reasoning about safety in decision support systems", *2nd European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, Granada.

[Van Denneheuvel et al., 1990] Van Denneheuvel, S, Van Emde Boas, P, De Geus, F and Rotterdam, E, 1990. "RL, a language for constraint solving", In: *Logic Programming and Database Processing* Computer Science in the Netherlands.

[Dincbas et al., 1988] Dicbas, M, Van Hentenryck, P, Simonis, H, Aggoun, A, Graft, T and Berthier, F, 1988. "The constraint logic programming language CHIP", In: *Proceedings of the International Conference in Fifth Generation Computer Systems*, Tokyo, Japan.

[De Dombal, 1991] De Dombal, FT, Dallos, V and McAdam, WA, 1991. "Can computer-aided teaching packages improve clinical care in patients with acute abdominal pain?", *British Medical Journal* **302** 1495–1497.

[Duda et al., 1976] Duda, RO, Hart, PE and Nilsson, NJ, 1976. "Subjective Bayesian methods for rule-based inference systems", In: *AFIPS Conference Proceedings of the 1976 National Computer Conference* **45** 1075–1082.

[Ehrig & Marh, 1985] Ehrig, H and Mahr, B, 1985. *Fundamentals of Algebraic Specification I: Equations and Initial Semantics*, Monographs on Theoretical Computer Science **6**, Springer-Verlag, Berlin.

[Ehrig & Marh, 1990] Ehrig, H and Mahr, B, 1990. *Fundamentals of Algebraic Specification II: Module Specifications and Constraints*, Monographs on Theoretical Computer Science **21**, Springer-Verlag, Berlin.

[First et al., 1982] First, MB, Weimer, BJ, McLinden, S and Miller, RA, 1982. "LOCALIZE: computer-assisted localization of peripheral nervous system lesions", *Computers and Biomedical Research* **15** 525–543.

[Fox, 1993] Fox, J, 1993. "On the soundness and safety of expert systems", *Artificial Intelligence in Medicine* **5** 159–179.

[Fox et al., 1990a] Fox, J, Clark, DA, Glowinski, A and O'Neil, MJ, 1990. "Using predicate logic to integrate qualitative reasoning and classical decision theory", *IEEE Transactions on Systems, man, and Cybernetics* **20** 347–357.

[Fox et al., 1990b] Fox, J, Gordon, C, Glowinski AJ and O'Neil, M, 1990. "Logic engineering for knowledge engineering: the Oxford System of Medicine", *Artificial Intelligence in Medicine* **2** 323–339.

[Futatsugi et al., 1985] Futatsugi, K, Goguen, JA, Jouannaud, JP and Meseguer, J, 1985. "Principles of OBJ2", In: *Proceedings of the Symposium on Principles of Programming Languages*, pp 52–66, ACM, New-York.

[Genesereth & Nilsson, 1987] Genesereth, MR and Nilsson, NJ, 1987. *Logical Foundations of Artificial Intelligence*, Morgan Kaufmann, Palo Alto, California.

[De Geus et al., 1991] De Geus, F, Rotterdam, E, Van Denneheuvel, S and Van Emde Boas, P, 1991. "Physiological modelling using RL", In: M Stefanelli, A Hasman, M Fieschi and J Talmon (eds.), *AIME91*, Lecture Notes in Medical Informatics, pp 198–210, Springer-Verlag, Berlin.

[Grodins, 1963] Grodins, FS, 1963. *Control Theory and Biological Systems*, Columbia University Press, New York.

[Hammond et al., 1993] Hammond, P, Davenport, JC and Fitzpatrick, FJ, 1993. "Logic-based integrity constraints and the design of dental prostheses", *Artificial Intelligence in Medicine* **5**(5) 431–446.

[Hammond et al, 1994] Hammond, P, Harris, AL, Das, KS and Wyatt, JC, 1994. "Safety and decision support in oncology", *Meth. Inf. Med.* **33**(4) 371–381.

[Hammond & Davenport, 1995] Hammond, P and Davenport, JC, 1995. "Eliciting and modelling the design knowledge of multiple experts", submitted for publication.

[Hammond & Sergot, 1995] Hammond, P and Sergot, M, 1994. "Computer support for protocol-based treatment of cancer", *J. Logic Programming*, to appear.

[Van Harmelen, 1991] Van Harmelen, F, 1991. *Meta-level Inference Systems*, Pitman, London.

[Van Harmelen & Balder, 1992] Van Harmelen, F and Balder, J, 1992. "$(\mathrm{ML})^2$: a formal language for KADS models of expertise", *Knowledge Acquisition* **4** 127–161.

[Hill & Lloyd, 1994] Hill, PM and Lloyd, JW, 1994. *The Gödel Programming Language*, The MIT Press, Cambridge, Massachusetts.

[Hayes, 1979] Hayes, PJ, 1979. "The logic of frames", In: D Metzing (ed.), *Frame Conception and Text Understanding*, pp 46–61, Walter de Gruyter and Co., Berlin.

[Van Hentenryck, 1991] Van Hentenryck, P, 1991. "Constraint logic programming", *The Knowledge Engineering Review* **6**(3) 151–194.

[Hickam et al., 1985] Hickam, DH, Shortliffe, EH and Bishoff, MB et al, 1985. "The treatment advice of a computer-based cancer chemotherapy protocol advisor", *Annals of Internal Medicine* **103** 928–936.

[Hrycej, 1992] Hrycej, T, 1993. "A temporal extension of Prolog", *J. Logic Programming* **16** 114–145.

[Huang et al., 1993] Huang, J, Fox, J, Gordon, C and Jackson-Smale, A, 1993. "Symbolic decision support in medical care", *Artificial Intelligence in Medicine* **5**(5) 415–430.

[Jaffar & Lassez, 1987] Jaffar, J and Lassez, J-L, 1987. "Constraint Logic Programming", *Proceedings of the 14th POPL*, pp 111–119.

[Jaffar & Maher, 1994] Jaffar, J and Maher, MJ, 1994. "Constraint logic programming: a survey", *J. Logic Programming* **19** 503–581.

[Jaspers, 1990] Jaspers, RBM, 1990. *Medical Decision Support: an Approach in the Domain of Brachial Plexus Injuries*, PhD thesis, Delft University of Technology.

[Josephson & Josephson, 1994] Josephson, JR, and Josephson, SG, 1994. *Abductive Inference: Computation,Philosophy, Technology*, Cambridge University Press, Cambridge.

[O'Keefe, 1987] O'Keefe, RM, Balci, O, Smith, EP, 1987. "Validating expert system performance", *IEEE Expert* **4** 81–89.

[Kim & Pearl, 1983] Kim, JH and Pearl, J, 1983. "A computational model for causal and diagnostic reasoning in inference systems", *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, pp 190–193, Karlsruhe.

[de Kleer et al., 1992] De Kleer, J, Mackworth, AK, Reiter, R, 1992. "Characterizing diagnosis and systems", *Artificial Intelligence* **56** 197–222.

[Konolige, 1992] Konolige, K, 1992. "Using default and causal reasoning in diagnosis", *Proceedings of the Workshop on Principles of Knowledge Representation and Reasoning*, Boston.

[Kowalski, 1979] Kowalski, RA, 1979. *Logic for Problem Solving*, North-Holland, Amsterdam.

[Kowalski & Sergot, 1986] Kowalski, RA and Sergot, MJ, 1986. "A logic-based calculus of events", *New Generation Computing*, **4**(1) 67–95.

[Kowalski, 1990] Kowalski, RA, 1990. "Problems and promises of computational logic", In: JW Lloyd (ed.), *Computational Logic*, pp 1–35, Springer-Verlag, Berlin.

[Krause & Glowinski, 1993] Krause, P and Glowinski, A (1993). "Formal specification and medical decision support", *Applied Artificial Intelligence* **7** 237–256.

[Krause et al., 1995] Krause, P, Ambler, S, Elvang-Coransson, M and Fox, J, 1995. "A logic of argumentation for reasoning under uncertainty", *Computational Intelligence*, to appear.

[Kuipers, 1986] Kuipers, BJ, 1986. "Qualitative simulation", *Artificial Intelligence* **29** 289–388.

[Kuipers, 1994] Kuipers, BJ, 1994. *Qualitative Reasoning*, The MIT Press, Cambridge, Massachusetts.

[Kulikowski & Weiss, 1982] Kulikowski, CA and Weiss, SM, 1982. "Representation of expert knowledge for consultation: the CASNET and EXPERT projects", In: P Szolovits (ed.), *Artificial Intelligence in Medicine*, pp 21–55, Westview Press, Boulder Colorado.

[Lloyd, 1987] Lloyd, JW, 1987. *Foundations of Logic Programming*, 2nd ed., Springer-Verlag, Berlin.

[Łukaszewicz, 1990] Łukaszewicz, W, 1990. *Non-monotonic Reasoning: Formalization of Commonsense Reasoning*, Ellis Horwood, Chichester.

[Lucas & Van der Gaag, 1991] Lucas, PJF and Van der Gaag, LC, 1991. *Principles of Expert Systems*, Addison-Wesley, Wokingham.

[Lucas & Janssens, 1991] Lucas, PJF and Janssens, AR, 1991. "Development and validation of HEPAR, an expert system for the diagnosis of disorders of the liver and biliary tract", *Medical Informatics* **16** 259-270.

[Lucas, 1993] Lucas, PJF, 1993. "The representation of medical reasoning models in resolution-based theorem provers", *Artificial Intelligence in Medicine* **5**(5) 395–414.

[Maes & Nardi, 1988] Maes, P and Nardi, D, 1988. *Meta-level Architectures and Reflection*, North-Holland, Amsterdam.

[Macartney, 1988] Macartney, FJ, 1988. "Diagnostic logic", In: C Philips,(ed.), *Logic in Medicine*, British Medical Journal, London.

[McCabe, 1992] McCabe, FG, 1992. *Logic and Objects*, Prentice-Hall, New York.

[Miller et al., 1982] Miller, AM, Pople, HE, Myers, JD, 1982. "INTERNIST-I, an experimental computer-based diagnostic consultant for general internal medicine", *New England Journal of Medicine* **307** 468–476.

[McCune, 1990] McCune, WW, 1990. *OTTER 2.0 Users' Guide*, Report ANN-88/44, Mathematics and Computer Science Division, Argonnne National Laboratory, Argonne, Illinois.

[Morik et al., 1993] Morik, K, Potamias, G, Moustakis, V and Charissis, G, 1993. "Model-based learning support to knowledge acquisition: a clinical case study", In: S Andreassen (ed.), *Artificial Intelligence in Medicine*, pp 434–439, IOS Press, Amsterdam, 1993

[Moser & Adlassnig, 1992] Moser, W and Adlassnig, K-P, 1992. "Consistency checking of binary categorical relationships in a medical knowledge base", *Artificial Intelligence in Medicine* **4** 389–407.

[Muggleton, 1992] Muggleton, S, 1992. "Inductive logic programming", In: S Muggleton (ed.), *Inductive Logic Programming*, pp 3–27, Academic Press, London.

[Patil et al., 1982] Patil, RS, Szolovits, P and Schwartz, WB, 1992. "Modelling knowledge of the patient in acid-base and electrolyte disorders", In: P Szolovits (ed.), *Artificial Intelligence in Medicine*, 191–226, Westview Press, Boulder Colorado, [Pearl, 1988] Pearl, J, 1988. *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufman, San Mateo, California.

[Peng & Reggia, 1990] Peng, Y and Reggia, JA, 1990. *Abductive Inference Models for Diagnostic Problem Solving,* Springer-Verlag, New York.

[Pople, 1973] Pople, HE, 1973. "On the mechanization of abductive logic", *Proceedings of the 3rd International Joint Conference on Artificial Intelligence.*

[Pople, 1977] Pople, HE, 1977. "The formation of composite hypotheses in diagnostic problem solving: an exercise in synthetic reasoning", *Proceedings of the 5th International Joint Conference on Artificial Intelligence.*

[Pople, 1982] Pople, HE, 1982. "Heuristic methods for imposing structure on ill-structured problems: the structuring of medical diagnosis", In: P Szolovits (ed.), *Artificial Intelligence in Medicine*, pp 119–190, Westview Press, Boulder Colorado.

[Reggia, 1978] Reggia, J, 1978. "A production system for neurological localization", *Proceedings of the 2nd Annual Symposium on Computer Applications in Medical Care*, pp 254–260.

[Reiter, 1978] Reiter, R, 1978. "On closed world databases", In: H Gallaire and J Minker (eds.), *Logic and Databases*, Springer-Verlag, Berlin.

[Reiter, 1987] Reiter, R, 1987. "A theory of diagnosis from first principle", *Artificial Intelligence* **32** 57–95.

[Shachter, 1986] Shachter, RD, 1986. "Evaluating influence diagrams", *Operation Research*, **34**(6) 871–882.

[Shortliffe, 1991] Shortliffe, EH, 1991. "Knowledge-based systems in medicine", In: K-P Adlassnig, G Grabner, S Bengtsson and R Hansen (eds.), *Medical Informatics Europe 1991*, pp 5–9, Springer-Verlag, Berlin.

[Shortliffe & Buchanan, 1975] Shortliffe, EH and Buchanan, BG, 1975. "A model of inexact reasoning in medicine", *Mathematical Biosciences* **23** 351–379.

[Smets et al., 1988] Smets, P, Mamdani, A, Dubois, D and Prade, H (eds.), 1988. *Non-Standard Logics for Automated Reasoning*, Academic Press, London.

[Sommer et al., 1994] Sommer, A, Emde, W, Kietz, J-U and Wrobel, S, 1994. *MOBAL 3.0 User Guide*, German National Research Centre for Computer Science (GMD), AI Research Division I3.KI, St. Augustin.

[Soper et al., 1991] Soper, P, Ranaboldo, C and Abeysinghe, G, 1991. "A temporal model for clinical and resource management in vascular surgery", In: D Karagiannis (ed.), *Database and Expert Systems Applications*, pp 549–552, Springer-Verlag, Berlin.

[Spivey, 1989] Spivey, JM, 1989. *The Z Notation: A Reference Manual*, Prentice Hall, New York.

[Sterling & Shapiro, 1986] Sterling, L and Shapiro, E, 1986. *The Art of Prolog*, The MIT Press, Cambridge, Massachusetts.

[Stickel, 1986] Stickel, M, 1986. “An introduction to automated deduction”, In: W Bibel and Ph Jorrand (eds.), *Fundamentals of Artificial Intelligence*, pp 75–132, Springer-Verlag, Berlin.

[Stickel, 1988] Stickel, M, 1988. “A Prolog technology theorem prover: implementation by an extended Prolog compiler”, *J. Automated Reasoning* **4**(4) 353–380.

[Todd, 1994] Todd, BS, 1994. “A probablistic simulation model to assist the localization of nerve lesions”, *Medical Informatics* **19**(3) 209–227.

[Torasso & Console, 1989] Torasso, P and Console, L, 1989. *Diagnostic Problem Solving: Combining Heuristic, Approximate and Causal Reasoning*, North Oxford Academic Publishers Ltd, London.

[Touretzky, 1986] Touretzky, DS, 1986. *The Mathematics of Inheritance Systems*, Pitman, London.

[Tuhrim, 1991] Tuhrim, S, Reggia, J and Goodall, S, 1991. “An experimental study of criteria for hypothesis plausibility”, *Journal of Experimental and Theoretical Artificial Intelligence* **3** 129–144.

[Turner, 1984] Turner, R, 1984. *Logics for Artificial Intelligence*, Ellis Horwood, Chichester.

[Weyhrauch, 1980] Weyhrauch, RW, 1980. “Prolegomena to a theory of mechanized formal reasoning”, *Artificial Intelligence***13** 133–170.

[Wos et al., 1992] Wos, L, Overbeek, R, Lusk, E and Boyle, J, 1992. *Automated Reasoning: Introduction and Applications*, 2nd ed., Prentice-Hall Inc, Englewood Cliffs, New Jersey.

[Wos, 1988] Wos, L, 1988. *Automated Reasoning: 33 Basic Research Problems*, Prentice-Hall Inc, Englewood Cliffs, New Jersey.

[Wrobel, 1990] Wrobel, S, 1990. *Application of MOBAL to the Medical Domains of ICS/FORTH*, Technical Note, GMD, P2154/31/1.

[Wu, 1991] Wu, TD, 1991. “A problem decomposition method for efficient diagnosis and interpretation of multiple disorders”, *Computer Methods and Programs in Biomedicine* **35** 239–250.

[Wyatt & Spiegelhalter, 1990] Wyatt, J and Spiegelhalter, DJ, 1990. “Evaluating medical expert systems: what to test for and how?”, *Medical Informatics* **15**(3) 205–217.