Chapter 2
An Introduction to Knowledge Representation and Reasoning in Healthcare

Arjen Hommersom and Peter J.F. Lucas

2.1 Development of the Field

Healthcare and medicine are, and have always been, very knowledge-intensive fields. Healthcare professionals use knowledge of the structure (molecular biology, cell biology, histology, gross anatomy) and functioning of the human body as well as knowledge of methods and means, some of them described by clinical guidelines, to diagnose and manage disorders. In addition, knowledge of how healthcare is organised is essential for the management of a patient's disease.

Already in the early days of research in artificial intelligence, researchers realised that healthcare and medicine would be suitably challenging fields to drive the development of knowledge representation and reasoning techniques. Quite a large number of different systems were developed in those early days (cf. [23] for a description of the most important early ideas and systems). Typically, researchers developed their own representation methods guided by thoughts on how to handle a particular medical problem. An example of how thoughts on clinical problem solving and computer-based knowledge representation can interact is the work by Pople [21] on heuristic methods for medical diagnostic problem solving. The key idea here is that one needs a kind of structure of the hypothesis space to guide the problem-solving process. In a medical context this means that one needs taxonomic knowledge, i.e. medical knowledge organised according to the principles of a subsumption taxonomy, and causal knowledge, i.e. knowledge that describes the world according to cause-effect relationships. Pople also realised that disease manifestations and the diseases themselves are linked to each other by a, possibly abstract, model of the pathophysiology, and those play a different role in the problem-solving process. Even in that early work it already clear that medicine is a semantically rich field, not only concerned with different type of knowledge of different kind, coming from different sources, but also used for different purposes.

The development of these early systems gave rise to the phrase *knowledge-based system*, or *knowledge system*, which is generally employed to denote information systems in which some symbolic representation of human knowledge of a domain

A. Hommersom and P.J.F. Lucas (eds.), *Biomedical Knowledge Representation*, LNAI 9521, DOI 10.1007/978-3-319-28007-3_2

is applied, usually in a way resembling human reasoning, to solve actual problems in the domain. As this knowledge is often derived from experts in a particular field, and early knowledge-based systems were actually developed in close collaboration with experts, the term *expert system* was the term used in the early days to refer to these systems. Knowledge, however, can also be extracted from literature, or from a datasets by using machine-learning methods. At the time of writing, the terminology of systems that employ formalised knowledge to solve problems is even less clear than it was in the past. For this book this is of little concern, as the focus is on knowledge representation and reasoning for different medical purposes.

Present generation knowledge-based systems are capable of dealing with significant (medical) problem domains. Gathering, maintaining and updating the incorporated knowledge taking into account its associated context, such as working environment, organisation and field of expertise belongs to an area referred to as *knowledge management*. The art of developing a knowledge-based system is called *knowledge engineering*, when there is emphasis on the pragmatic engineering aspects, or *knowledge modelling*, when development of domain models is emphasises. The latter is strictly speaking part of the former. The process of collecting and analysing knowledge in a problem domain is called *knowledge acquisition*, or *knowledge elicitation* when the knowledge is gathered from interviews with experts, normally using interview techniques as developed by psychologists.

Although the early papers on knowledge representation for biomedical problems are still worth reading, there has been significant progress in the techniques, i.e. languages and tools, that act as the basis for knowledge representation. In contrast to the early work, there is now a solid understanding of the importance of logical language to act as a basis for knowledge representation. At the same time, specialised logical languages, such as decoration logics, have been developed to deal with specific knowledge representation and reasoning problems. There has also been a lot of progress in the development of reasoning with uncertainty. Probabilistic graphical models, and in particular Bayesian networks, have come into play since the 1990s as a natural formalism to represent uncertain biomedical knowledge. Specific types of non-monotonic reasoning have also emerged and proven their use in the biomedical context. The theory of argumentation is a typical example. For specific biomedical problems, such as problems that can be handled by clinical guidelines, there are now languages and tools available to represent and to reason with the relevant knowledge.

In general, the significant progress in techniques for knowledge representation and reasoning render it possible to develop knowledge systems of which the foundations are well understood in such way that certainty (computational) properties are guaranteed to be satisfied. Of course, capturing and modelling biomedical knowledge is still a significant challenge. However, with the techniques available nowadays, the modelling is at least supported by sound methods and techniques.

In this chapter, we will review common knowledge representation formalisms in artificial intelligence and link these to the healthcare field.

2.2 Techniques for Knowledge Representation and Reasoning

The knowledge-representation formalism and the types of reasoning supported are of major importance for the development of knowledge-based systems. Logic, probability theory and decision theory are sufficiently general to permit describing the nature of knowledge representation, inference and problem solving without having to resort to special-purpose languages. In the next section, some of the general ideas underlying knowledge representation are summarised and illustrated by means of simple examples.

2.2.1 Horn-clause Logic

Knowledge-based systems usually offer a number of different ways to represent knowledge in a domain, and to reason with this knowledge automatically to derive conclusions. Although the languages offered by actual systems and tools may differ in a number of ways, there are also many similarities. The aspects that the languages have in common can be best understood in terms of a logical representation, as accomplished below.

A *Horn clause* or *rule* is a logical implication of the following form

$$\forall x_1 \cdots \forall x_m ((A_1 \wedge \cdots \wedge A_n) \rightarrow B) \tag{2.1}$$

where A_i, B are literals of the form $P(t_1, \ldots, t_q)$, i.e. without a negation sign, representing a relationship P between terms t_k, which may involve one or more universally quantified variables x_j, constants and terms involving function symbols. As all variables in rules are assumed to universally quantified, the universal quantifiers are often omitted if this does not give rise to confusion. If $n = 0$, then the clause consists only of a conclusion, which may be taken as a *fact*. If, on the other hand, the conclusion B is empty, indicated by $\bot$, the rule is also called a *query*. If the conditions of a query are satisfied, this will give rise to a contradiction or inconsistency, denoted by $\bot$, as the conclusion is empty. So, an empty clause means actually inconsistency.

A popular method to reason with clauses, and Horn clauses in particular, is *resolution*. Let $\mathscr{R}$ be a set of rules not containing queries, and let $Q \equiv (A_1 \wedge \cdots \wedge A_n) \rightarrow \bot$ be a query, then

$$\mathscr{R} \cup \{Q\} \vdash \bot$$

where $\vdash$ means the application of resolution, implies that the conditions

$$\forall x_1 \cdots \forall x_m (A_1 \wedge \cdots \wedge A_n)$$

are not all satisfied. Since resolution is a sound inference rule, meaning that it respects the logical meaning of clauses, it also holds that $\mathscr{R} \cup \{Q\} \vDash \bot$, or equivalently

$$\mathscr{R} \vDash \exists x_1 \cdots \exists x_m (A_1 \wedge \cdots \wedge A_n)$$

if $\mathscr{R}$ only consists of Horn clauses. This last interpretation explains why deriving inconsistency is normally not really the goal of using resolution; rather, the purpose is to derive certain facts. Since resolution is only complete for deriving inconsistency, called *refutation completeness*, it is only safe to 'derive' knowledge in this indirect manner. There exist other reasoning methods which do not have this limitation. However, resolution is a simple method that is understood in considerable depth. As a consequence, state-of-the-art resolution-based reasoners are very efficient. Resolution can also be used with clauses in general, which are logical expressions of the form

$$(A_1 \wedge \cdots \wedge A_n) \rightarrow (B_1 \vee \cdots \vee B_m)$$

usually represented as:

$$\neg A_1 \vee \cdots \vee \neg A_n \vee B_1 \vee \cdots \vee B_m$$

Rules of the form (2.1) are particularly popular as the reasoning with propositional Horn clauses is known to be possible in linear time, whereas reasoning with propositions or clauses in general (where the right-hand side consists of disjunctions of literals) is known to be NP-complete, i.e. may require time exponential in the size of the clauses. Note that allowing negative literals at the left-hand side of a rule is equivalent to having disjunctions at the right-hand side. Using a logical language that is more expressive than Horn-clause logic is sometimes unavoidable, and special techniques have been introduced to deal with their additional power.

Using logic to represent (medical) knowledge gives rise to a knowledge base that is sometimes called *object knowledge*.

Let KB be a knowledge base consisting of a set (conjunction) of rules, and let F be a *set of facts* observed for a particular problem $\mathscr{P}$, then there are generally three ways in which a problem can be solved, yielding different types of solutions. The formalisation of problem solving gives rise to knowledge that is sometimes called *meta knowledge*. Let $\mathscr{P}$ be a problem, then there are different classes of solutions to this problem:

- **Deductive solution:** S is a *deductive solution* of a problem $\mathscr{P}$ with associated set of observed findings F iff

$$\mathrm{KB} \cup F \vDash S \tag{2.2}$$

  and $\mathrm{KB} \cup F \nvDash \bot$, where S is a set of solution formulae.
- **Abductive/inductive solution:** S is an *abductive solution* of a problem $\mathscr{P}$ with associated set of observed findings F iff the following *covering condition*

$$\mathrm{KB} \cup S \cup K \vDash F \tag{2.3}$$

is satisfied, where K stands for *contextual knowledge*. In addition, it must hold that $\mathrm{KB} \cup S \cup C \nvDash \bot$ (consistency condition), where C is a set of logical constraints on solutions. For the abductive case, it is assumed that the knowledge base KB contains a logical representation of *causal knowledge* and S consists of facts; for the inductive case, KB consists of background facts and S, called an *inductive solution*, consists of rules.

- **Consistency-based solution:** S is a *consistency-based solution* of a problem $\mathscr{P}$ with associated set of observed findings F iff

$$\mathrm{KB} \cup S \cup F \nvDash \bot \tag{2.4}$$

Note that a deductive solution is a consistent conclusion that follows from a knowledge base KB and a set of facts, whereas an abductive solution acts as a hypothesis that *explains* observed facts in terms of causal knowledge, i.e. cause-effect relationships. An inductive solution also explains observed facts, but in terms of any other type of knowledge. A consistency-based solution is the weakest kind of solution, as it is neither required to be concluded nor is it required to explain observed findings.

2.2.2 Objects, Attributes and Values

Even though facts or observed findings can be represented in many different ways, in many systems facts are represented in an object-oriented fashion. This means that facts are described as properties, or *attributes*, of objects in the real world. Attributes of objects can be either multivalued, meaning that an object may have more than one of those properties at the same time, or singlevalued, meaning that values of attributes are mutually exclusive.

In logic, multivalued attributes are represented by predicate symbols, e.g.:

$$\mathrm{Parent}(\mathrm{John}, \mathrm{Ann}) \wedge \mathrm{Parent}(\mathrm{John}, \mathrm{Derek})$$

indicates that the ‘object’ John, represented as a constant, has two parents (the attribute ‘Parent’): Ann and Derek, both represented by constants. Furthermore, singlevalued attributes are represented as function symbols, e.g.

$$\mathit{gender}(\mathrm{John}) = \mathit{male}$$

Here, ‘*gender*’ is taken as a singlevalued attribute, ‘John’ is again a constant object, and ‘*male*’ is the value, also represented as a constant.

It is, of course, also possible to state general properties of objects. For example, the following bi-implication:

$$\forall x \forall y \forall z((\text{Parent}(x, y) \wedge \text{Parent}(y, z)) \leftrightarrow \text{Grandparent}(x, z))$$

defines the attribute 'Grandparent' in terms of the 'Parent' attribute.

Another typical example of reasoning about properties of objects is *inheritance* [2]. Here one wishes to associate properties of objects with the classes the objects belong to, mainly because this yields a compact representation offering in addition insight into the general structure of a problem domain. Consider, for example, the following knowledge base KB:

$$\begin{aligned} \forall x(\text{Mammal}(x) &\rightarrow \text{Endotherm}(x)) \\ \forall x(\text{Human}(x) &\rightarrow \text{Mammal}(x)) \\ \forall x(\text{Human}(x) &\rightarrow \textit{number-of-chromosomes(x)} = 46) \end{aligned}$$

Clearly, it holds that

$$\text{KB} \cup \{\text{Human}(\text{John})\} \vDash \textit{number-of-chromosomes}(\text{John}) = 46$$

as the third rule expresses that as a typical property of humans. However, the knowledge base also incorporates more general properties of humans, such as:

$$\text{KB} \cup \{\text{Human}(\text{John})\} \vDash \text{Mammal}(\text{John})$$

Now, given the fact that a human is a mammal, we can now also conclude

$$\text{KB} \cup \{\text{Human}(\text{John})\} \vDash \text{Endotherm}(\text{John})$$

The example knowledge base discussed above can also be represented as a graph, called an object *taxonomy*, and is shown in Fig. 2.1. Here ellipses indicate either classes of objects (Human and Mammal) or specific objects (John). Solid arcs in the graph indicate that a class of objects is a subclass of another class of objects; a dashed arc indicates that the parent object is an element – often the term 'instance' is used instead – of the associated class of objects. The term 'inheritance' that is associated with this type of logical reasoning derives from the fact that the reasoning goes from the children to the parents in order to derive properties.

2.2.3 Description Logics

Describing the objects in a domain, usually but not always in a way resembling a taxonomy, usually with the intention to obtain a formal description of the terminology in a domain, is known as an *ontology*. Instead of describing these properties in standard first-order logic, it is common nowadays to use specialised description logics for that purpose and in particular OWL, the Web Ontology Language [13, 16], is being used for that purpose.

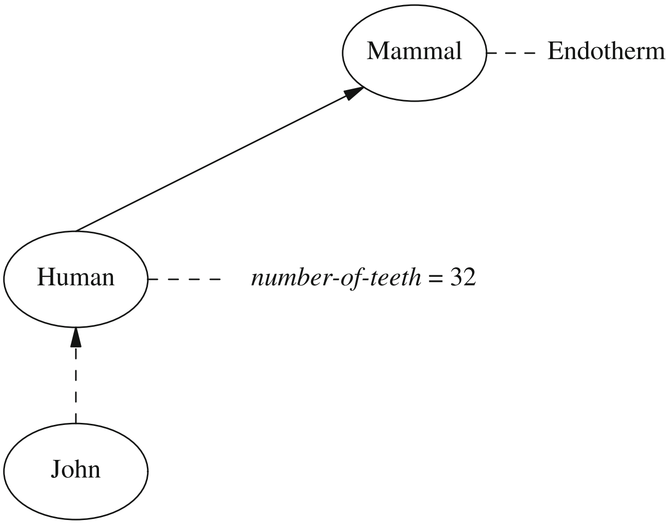

Fig. 2.1 An object taxonomy.

There are two primary ways in which knowledge is being described using OWL:

1. by *combining concepts* using Boolean operators, such as $\sqcap$ (conjunction), and $\sqcup$ (disjunction);
2. by *defining relationships* between concepts (whether primitive or obtained by combining primitive concepts) using the *subsumption* relation $\sqsubseteq$ (also called *general concept inclusion* – GCI).

Thus, a concept description is constructed from

- *primitive concepts* C, e.g., Disease, $\top$ (most general), $\bot$ (empty);
- *primitive roles* r, e.g., hasSymptom;
- *conjunctions* $\sqcap$, e.g., Cardiac_Disease $\sqcap$ Cerebral_Disease;
- *disjunctions* $\sqcup$, e.g., Hepatitis $\sqcup$ Cirrhosis;
- a *complement* $\neg$, e.g., $\neg$Hepatitis;
- a *value restriction* $\forall r.C$, e.g., $\forall$causes.Fever;
- an *existential restriction* $\exists r.C$, e.g., $\exists$likelyFatal.Metastasis.

All understood in terms of (groups of) individuals and properties of individuals.

For example, by

$$\text{Hepatitis} \sqcup \text{Cirrhosis}$$

we have combined two concepts, but we have not established how they are related to each other. By writing:

$$\text{Hepatitis} \sqsubseteq \text{LiverDisease}$$

we have established a relationship between the concepts ‘Hepatitis’ and ‘LiverDisease’, where the first is less or equally general than the latter. By combining two subsumption relations, it is possible to *define* a new concept:

$$\text{Metastatic\_Cancer} \sqsubseteq \text{Cancer} \sqcap \exists \text{hasMetastasis.Tumour\_Tissue}$$

and

$$\text{Cancer} \sqcap \exists \text{hasMetastasis.Tumour\_Tissue} \sqsubseteq \text{Metastatic\_Cancer}$$

is abbreviated to

$$\text{Metastatic\_Cancer} \equiv \text{Cancer} \sqcap \exists \text{hasMetastasis.Tumour\_Tissue}$$

Note that an expression such as ∃hasMetastasis.Tumour_Tissue is also a concept: the role hasMetastasis establishes a relationship between an instance of the concept Tumour_Tissue (the tumour discovered at a distance from the original cancer) and all concepts that participate in the role, which are then intersected with the concept ‘Cancer’, yielding a definition of ‘Metastatic_Cancer’.

General descriptions of a domain form, what is called, the TBox (Terminology Box). In a sense, the TBox restricts the terminology we are allowed to use when describing a domain. The actual domain is described by means of *assertions*, which together form the ABox (Assertion Box).

2.2.4 Temporal Logics

As soon we wish to model the execution of actions in biomedicine, we need to incorporate time into our knowledge-representation formalism, and thus also when it is based on logic.

Several temporal logics have been developed, in particular tense logics since the 1960s. Differences between logics result from different models of time and expressiveness. In linear temporal logics (e.g., Linear Temporal Logic (LTL) [19]), models form a linear trace, while in branching logics models typically (e.g., Computation Tree Logic (CTL) [1, 5, 10]) form a tree.

In LTL, propositional logic is extended with several temporal operators. The temporal operators used are **X**, **G**, **F**, and **U**. With $\mathbf{X}\varphi$ being true if φ holds in the next state, $\mathbf{G}\varphi$ if φ holds in the current state and all future states, $\mathbf{F}\varphi$ if φ holds in the current state or some state in the future, and $\varphi\mathbf{U}\psi$ if φ holds until eventually ψ holds.

For example:

$$\mathbf{G}(\text{Human} \to \text{Mammal})$$

expresses that it is always the case that humans are mammals. To specify the mortality of humans, one could model this with a rule such as:

$$\mathbf{G}(\text{Human} \rightarrow \mathbf{F}\,\text{Death})$$

Note the subtle difference with a logical rule such as the following:

$$\text{Human} \rightarrow \mathbf{F}\,\text{Death}$$

which states that only the humans existing at this moment are mortal; this does not specify that those born in the future are mortal. One could even be slightly more precise and specify that humans remain human at least until they die:

$$\mathbf{G}(\text{Human}\,\mathbf{U}\,\text{Death})$$

Of course, reasoning with temporal knowledge is supported as well. One can derive for example that:

$$\mathbf{G}(\text{Human}\,\mathbf{U}\,\text{Death}),\ \text{Human} \models \mathbf{F}\,\text{Death}$$

In contrast to LTL, CTL provides operators for describing events along a multiple computation paths (possible futures), and is therefore sometimes referred to as a 'branching' temporal logic. The path quantifiers **A** and **E**, which are always combined with one of the LTL operators, are used to specify that all or some of the paths starting at a specific state have some property. While LTL formulas describe all possible futures, in CTL we may describe what happens in some or all of the possible futures.

For example, the specify that cancer may lead to metastatic cancer, but at the same time be optimistic that there is a possibility that does not does not occur, one could write the following rules:

$$\mathbf{AG}(\text{Cancer} \rightarrow \mathbf{EF}\ \text{Metastatic\_Cancer})$$
$$\mathbf{AG}(\text{Cancer} \rightarrow \mathbf{EG}\ \neg\text{Metastatic\_Cancer})$$

Automatic reasoning methods for temporal logics have been developed, although reasoning with temporal logic is a hard problem (for example, checking satisfiability and entailment for LTL is PSPACE-complete). One practical method is to look upon temporal logics as first-order formula with a quantification over the temporal states, i.e., each predicates has an additional argument that models the state (and path for CTL) in which the predicate holds. Temporal quantification can then be mapped to ordinary first-order quantification. For example:

$$\mathbf{G}p \equiv \forall t\ p(t)$$

Reasoning methods for first-order logic can then directly be applied to reason about temporal logics, for example, resolution.

2.3 Problem-Solving Methods

Using formalisations of medical knowledge to solve problems can be seen as a form of meta-level reasoning, as discussed in Sect. 2.2.1. In the section on probabilistic logic, we already saw a form of meta-level reasoning with uncertaint knowledge. Just to illustrate the idea, we discuss various examples of diagnostic reasoning. In addition, treatment planning — one of the aspects of guideline execution — is briefly sketched.

2.3.1 Diagnostic Problem Solving

Above, the general features of knowledge representation and inference were sketched. Most of the insight that has been gained in the field, however, concerns particular methods with associated knowledge to handle classes of problems. As said above, inference or reasoning methods can be used to implement problem-solving methods. A typical example is the diagnosis of disorders in patients or faults in equipment by diagnostic methods. Many different methods have been developed for that purpose. Three well-known diagnostic methods with their associated types of knowledge will be discussed in the following.

2.3.1.1 Deductive Diagnosis

Most of the early knowledge-based systems, including MYCIN [3], were based on expert knowledge concerning the relationships among classes expressed by rules. In the reasoning process these rules were subsequently used to classify cases into categories. This problem-solving method is known as *heuristic classification*, as most of the knowledge encoded in the rules is empirical or heuristic in nature rather than based on first principles [4]. The form of the rules is:

$$(c_1 \wedge \cdots \wedge c_k \wedge \sim c_{k+1} \wedge \cdots \wedge \sim c_n) \rightarrow c$$

where c_i is either a condition on input data or on a subclass. The rules are *generalised* rules, as conditions may be prefixed by a special negation sign $\sim$, called *negation by absence*. It represents a special case of the closed-world assumption (CWA); a condition $\sim c_i$ only succeeds if there is at least one finding concerning the associated attribute. Formally:

$$\sim A(o, v) \equiv \exists x (A(o, x) \wedge x \neq v)$$

for object o and value v, where o and v are constants. If the attribute A represents a measurement or test, then negation by absence checks whether the test has been carried out, yielding a result different from the one specified.

Consider the following toy medical knowledge base KB:

$$\forall x((\text{Symptom}(x, \textit{coughing}) \wedge \sim \text{Symptom}(x, \textit{chest-pain}) \wedge \text{Sign}(x, \textit{fever}))$$
$$\rightarrow \text{Disorder}(x, \textit{flu}))$$
$$\forall x((\textit{temp}(x) > 38) \rightarrow \text{Sign}(x, \textit{fever}))$$

Then it holds that:

$$\text{KB} \cup \{\textit{temp}(\text{John}) = 39, \text{Symptom}(\text{John}, \textit{coughing})\} \vDash_{\text{NA}} \text{Disorder}(\text{John}, \textit{coughing})$$

using negation by absence (NA). Note that Sign(John, *fever*) is true, and may be viewed as a classification of the finding $\textit{temp}(\text{John}) = 39$; $\sim$ Symptom(John, *chest-pain*) holds due to negation by absence. Both rules in the knowledge base KB above are examples of heuristic classification rules.

2.3.1.2 Abductive Diagnosis

In abductive diagnosis, use is made of causal knowledge to diagnose a disorder in medicine or to determine faults in a malfunctioning device [6, 18, 20]. Causal knowledge can be represented in many ways, but a rather convenient and straight-forward way to represent causal knowledge is by taking logical implication as standing for the causal relationship. Thus, rules of the form:

$$d_1 \wedge \cdots \wedge d_n \rightarrow f \tag{2.5}$$
$$d_1 \wedge \cdots \wedge d_n \rightarrow d \tag{2.6}$$

are obtained, where d_i stands for a condition concerning a defective component or disorder; the conjunctions in (2.5) and (2.6) indicate that these conditions interact to either cause observable finding f or another abnormal condition d as effect. Sometimes uncertainty is added, usually represented in a non-numerical way as an assumption α:

$$d_1 \wedge \cdots \wedge d_n \wedge \alpha_f \rightarrow f \tag{2.7}$$
$$d_1 \wedge \cdots \wedge d_n \wedge \alpha_d \rightarrow d \tag{2.8}$$

The literals α may be either assumed to be true or false, meaning that f and d are a possible, but not necessary, consequences of the simultaneous occurrence of $d_1, \ldots, d_n$.

An *abductive diagnosis* S is now simply an abductive solution, where literals in S are restricted to d_i's and α's. The contextual knowledge may be extra conditions on rules which cannot be derived, but must be assumed and may act to model conditional causality. For simplicity's sake it is assumed here that K is empty. The set of constraints C may for instance consist of those findings f which have not been observed, and are assumed to be absent, i.e. $\neg f$ is assumed to hold.

Consider, for example, the causal model with set of defects and assumptions:

$$\Delta = \{fever, influenza, sport, \alpha_1, \alpha_2\}$$

and observable findings

$$\Phi = \{chills, thirst, myalgia, \neg chills, \neg thirst, \neg myalgia\}$$

'Myalgia' means painful muscles. The following knowledge base KB contains medical knowledge concerning influenza and sport, both 'disorders' with frequent occurrence:

$$\begin{aligned}
&fever \wedge \alpha_1 \rightarrow chills\\
&influenza \rightarrow fever\\
&fever \rightarrow thirst\\
&influenza \wedge \alpha_2 \rightarrow myalgia\\
&sport \rightarrow myalgia
\end{aligned}$$

For example, $influenza \wedge \alpha_2 \rightarrow myalgia$ means that influenza *may cause* myalgia; $influenza \rightarrow fever$ means that influenza *always causes* fever. For illustrative purposes, a causal knowledge base as given above is often depicted as a labelled, directed graph G, which is called a *causal net*, as shown in Fig. 2.2. Suppose that the abductive diagnostic problem with set of facts

$$F = \{thirst, myalgia\}$$

must be solved. As constraints we take $C = \{\neg chills\}$. There are several solutions to this abductive diagnostic problem (for which the consistency and covering conditions are fulfilled):

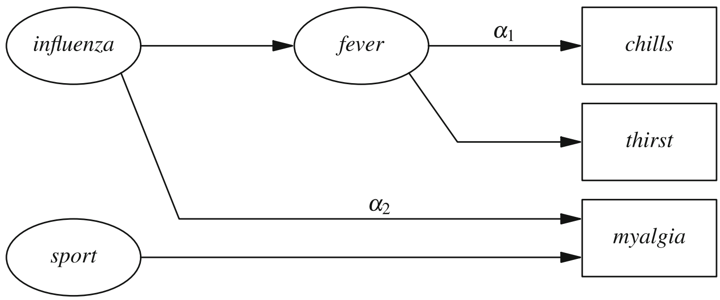

Fig. 2.2 A knowledge base with causal relations.

$$
\begin{aligned}
S_1 &= \{\mathit{influenza}, \alpha_2\} \\
S_2 &= \{\mathit{influenza}, \mathit{sport}\} \\
S_3 &= \{\mathit{fever}, \mathit{sport}\} \\
S_4 &= \{\mathit{fever}, \mathit{influenza}, \alpha_2\} \\
S_5 &= \{\mathit{influenza}, \alpha_2, \mathit{sport}\} \\
S_6 &= \{\mathit{fever}, \mathit{influenza}, \mathit{sport}\} \\
S_7 &= \{\mathit{fever}, \mathit{influenza}, \alpha_2, \mathit{sport}\}
\end{aligned}
$$

Note that $S = \{\alpha_1, \alpha_2, \mathit{fever}, \mathit{influenza}\}$ is incompatible with the constraints C.

2.3.1.3 Consistency-Based Diagnosis

In consistency-based diagnosis, in contrast to abductive diagnosis, the malfunctioning of a system is diagnosed by using mainly knowledge of the normal structure and normal behaviour of its components [8, 11, 22]. For each component COMP_j its normal behaviour is described by logical implications of the following form:

$$\forall x((\mathrm{COMP}_j(x) \wedge \neg \mathrm{Ab}(x)) \rightarrow \mathrm{Behaviour}_j(x))$$

The literal $\neg\mathrm{Ab}(x)$ expresses that the behaviour associated with the component only holds when the assumption that the component is not abnormal, i.e. $\neg\mathrm{Ab}(c)$, is true for component c. Sometimes knowledge of abnormal behaviour is added to implications of the form above, having the form:

$$\forall x((\mathrm{COMP}_j(x) \wedge \mathrm{Ab}(x)) \rightarrow \mathrm{Behaviour}_j(x))$$

These may result in a reduction in the number of possible diagnoses to be considered. Logical behaviour descriptions of the form discussed above are part of a *system description*. In addition to the generic descriptions of the expected behaviour of components, a system description also includes logical specifications of how the components are connected to each other (the structure of the system), and the names of the components constituting the system. The system description is now taken as the knowledge base KB of a system. Problem solving basically amounts to adopting particular assumptions about every $\mathrm{COMP}_j(c)$, either whether $\mathrm{Ab}(c)$ is true or false. This sort of reasoning is called *assumption-based* or *hypothetical reasoning*.

In medicine, a component may be one of the organs or structures that are part of a physiological system. For example, for the cardiovascular system the ‘blood’ might be one of the components. As for the cardiovascular system it is the blood volume that affects its physiology, we will take ‘blood volume’ as a component in the medical example below. We will describe how a description of cardiovascular physiology can be employed in diagnosis (cf. [9] for details).

The following logical implications give the steady-state equations of the cardiovascular system, i.e. when the system is stable:

$$\begin{aligned}\neg\mathrm{Ab}(C_{R_{\mathrm{sys}}}) &\rightarrow \Delta P_{\mathrm{sys}} = \mathrm{CO} \cdot R_{\mathrm{sys}}\\ \neg\mathrm{Ab}(C_{\mathrm{BV}}) &\rightarrow P_v = \mathrm{BV}/\mathrm{VC}\\ \neg\mathrm{Ab}(C_{\mathrm{VC}}) &\rightarrow P_v = \mathrm{BV}/\mathrm{VC}\\ \neg\mathrm{Ab}(C_{P_v}) &\rightarrow \Delta P_v = \mathrm{VR} \cdot R_v\end{aligned}$$

Here, the following abbreviations are used:

- R_{sys}: the systemic resistance of the cardiovascular system;
- ΔP_{sys}: the difference between arterial and venous pressure;
- CO: the cardiac output (volume of blood per minute pumped out of the heart);
- P_v: the venous pressure;
- BV: blood volume;
- VC: venous compliance (the elastic force of the vessel wall against increased internal volume);
- VR: venous return (volume of blood per minute returned to the heart); it is equal to CO;
- R_v: venous resistance.

With C_X is indicated the corresponding component X that can be malfunctioning, $\mathrm{Ab}(C_X)$. Any disturbance of the steady state may violate any of the equations in the right-hand sides of the implications above. In this case, the set of potential 'faulty' components is:

$$\mathrm{COMPS} = \{C_{R_{\mathrm{sys}}}, C_{\mathrm{BV}}, C_{\mathrm{VC}}, C_{P_v}\}$$

The cardiovascular system is controlled in such way that changes in its parameters are compensated automatically by changes in other parameters, leading to homeostatis. The following equation describes, for example, how the blood-pressure regulator (baroreceptor system) reacts to a change in arterial blood pressure (P_a) by changing the systemic resistance:

$$R_{\mathrm{sys}} = -0.17 P_a + 34 \tag{2.9}$$

Now, assume that a patient gets kidney damage. This will lead to water retention, and thus the blood volume increases. In turn increased blood volume will lead to increase of arterial pressure P_a. The barorecepter system will through Eq. (2.9) compensate for the increased arterial pressure through decrease in systemic resistance R_{sys}.

Let us assume that the following measurements are made in a patient:

$$F = \{P_a = 160\,\mathrm{mmHg},\ P_v = 15\,\mathrm{mmHg},\ \mathrm{CO} = 7\,\mathrm{l/min}\},$$

thus, $\Delta P_{\mathrm{sys}} = 160 - 15 = 145\,\mathrm{mmHg}$. $R_{\mathrm{sys}} = 6.8\,\mathrm{mmHG\ min/l}$ is the predicted effect of the regulatory baroreceptor mechanism using Eq. (2.9). Clearly, the steady-state equation for the systemic resistance is violated:

$$\Delta P_{\text{sys}} = 145 \neq \text{CO} \cdot R_{\text{sys}} = 7 \cdot 6.8 = 47.6$$

This indicates that the system is malfunctioning. This can also be verified by noting that when assuming all components to behave normally, i.e. $S = \{\neg\text{Ab}(c) \mid c \in \text{COMPS}\}$, it follows that

$$\text{KB} \cup S \cup F$$

is inconsistent.

Diagnosing the problem simply consists of assuming particular components to be abnormal (Ab(c) is true for those components), and checking whether the result is still inconsistent. If it is not, a diagnosis has been found. So, a *consistency-based diagnosis* is a consistency-based solution S consisting of a conjunction of Ab literals, one for every component.

Consider again the example above. Here,

$$S = \{\text{Ab}(C_{R_{\text{sys}}}), \text{Ab}(C_{\text{BV}}), \neg\text{Ab}(C_{\text{VC}}), \neg\text{Ab}(C_{P_v})\}$$

is a consistency-based diagnosis as

$$\text{KB} \cup S \cup F \nvDash \perp$$

Note that R_{sys} and BV are, thus, possibly faulty, as assuming them to be abnormal yield no output for this components. There are other solutions as well, such as

$$S' = \{\text{Ab}(C_{R_{\text{sys}}}), \neg\text{Ab}(C_{\text{BV}}), \text{Ab}(C_{\text{VC}}), \neg\text{Ab}(C_{P_v})\}$$

2.3.2 Treatment Planning

As medical management is a time-oriented process, diagnostic and treatment actions described in guidelines are performed in a temporal setting. It is assumed that two types of knowledge are involved in detecting the violation of good medical practice:

- Knowledge concerning the (patho)physiological mechanisms underlying the disease, and the way treatment influences these mechanisms. The knowledge involved could be causal in nature, and is an example of *object-knowledge*.
- Knowledge concerning good practice in treatment selection; this is *meta-knowledge*.

Below we present some ideas on how such knowledge may be formalised using temporal logic (cf. [15] for early work).

We are interested in the prescription of drugs, taking into account their mode of action. Abstracting from the dynamics of their pharmacokinetics, this can be formalised in logic as follows:

$$(\mathsf{G}\, d \wedge r) \rightarrow \mathsf{G}(m_1 \wedge \cdots \wedge m_n)$$

where d is the name of a drug or possibly of a group of drugs indicated by a predicate symbol (e.g. $\mathrm{SU}(x)$, where x is universally quantified and 'SU' stands for sulfonylurea drugs, such as Tolbutamid, which are prescribed in diabetes mellitus type 2), r is a (possibly negative or empty) *requirement* for the drug to take effect, and m_k is a mode of action, such as decrease of release of glucose from the liver, which holds at all future times.

The modes of action m_k can be combined, together with an *intention* n (achieving normoglycaemia, i.e. normal blood glucose levels, for example), a particular patient *condition* c, and *requirements* r_j for the modes of action to be effective:

$$(\mathsf{G}m_{i_1} \wedge \cdots \wedge \mathsf{G}m_{i_m} \wedge r_1 \wedge \cdots \wedge r_p \wedge \mathsf{H}c) \rightarrow \mathsf{G}n$$

Good practice medicine can then be formalised as follows. Let $\mathscr{B}$ be background knowledge, $T \subseteq \{d_1, \ldots, d_p\}$ be a set of drugs, C a collection of patient conditions, R a collection of requirements, and N a collection of intentions which the physician has to achieve. A set of drugs T is a *treatment* according to the theory of abductive reasoning if [20]:

(1) $\mathscr{B} \cup \mathsf{G}T \cup C \cup R \nvDash \bot$ (the drugs do not have contradictory effects), and
(2) $\mathscr{B} \cup \mathsf{G}T \cup C \cup R \vDash N$ (the drugs handle all the patient problems intended to be managed)

If in addition to (1) and (2) condition

(3) $O_\varphi(T)$ holds, where O_φ is a meta-predicate standing for an optimality criterion or combination of optimality criteria φ, then the treatment is said to be *in accordance with good-practice medicine*.

A typical example of this is subset minimality $O_\subset$:

$$O_\subset(T) \equiv \forall T' \subset T : T' \text{ is not a treatment according to (1) and (2)}$$

i.e. the minimum number of effective drugs are being prescribed. For example, if $\{d_1, d_2, d_3\}$ is a treatment that satisfies condition (3) in addition to (1) and (2), then the subsets $\{d_1, d_2\}$, $\{d_2, d_3\}$, $\{d_1\}$, and so on, do not satisfy conditions (1) and (2). In the context of abductive reasoning, subset minimality is often used in order to distinguish between various solutions; it is also referred to in literature as *Occam's razor*. Another definition of the meta-predicate O_φ is in terms of minimal cost O_c:

$$O_c(T) \equiv \forall T', \text{ with } T' \text{ a treatment: } \; c(T') \geq c(T)$$

where $c(T) = \sum_{d \in T} \mathit{cost}(d)$; combining the two definitions also makes sense. For example, one could come up with a definition of $O_{\subset,c}$ that among two subset-minimal treatments selects the one that is the cheapest in financial or ethical sense.

2.4 Reasoning with Uncertainty

Uncertainty is another essential aspects of much medical knowledge and data. Here again, there has been a lot of research in artificial intelligence.

2.4.1 Bayesian Networks

Up until now, it has been assumed that in representing and solving a problem in a domain dealing with uncertainty is not of major importance. As this does not hold for many problems, the possibility to represent and reason with the uncertainty associated with a problem is clearly of significance. There have been a number of early attempts where researchers have augmented rule-based, logical methods with uncertainty methods, usually different from probability theory, although sometimes also related. However, those methods are now outdated, and have been replaced by methods which take probability theory as a starting point. In the context of knowledge-based systems, in particular the formalism of Bayesian (belief) networks has been successful [7, 12, 14, 17].

A *Bayesian belief network* $\mathscr{B} = (G, \Pr)$, also called *causal probabilistic network*, is a directed acyclic graph $G = (V(G), A(G))$, consisting of a set of nodes $V(G) = \{V_1, \ldots, V_n\}$, called *probabilistic nodes*, representing discrete random variables, and a set of arcs $A(G) \subseteq V(G) \times V(G)$, representing causal relationships or correlations among random variables. Consider Fig. 2.3, which shows a simplified version of a Bayesian belief network modelling some of the relevant variables in the diagnosis of two causes of fever. The presence of an arc between two nodes denotes the existence of a direct causal relationship or other influences; absence of an arc means that the variables do not influence each other directly. The following knowledge is represented in Fig. 2.3: variable 'FL' is expressed to influence 'MY' and 'FE', as it is known that flu causes myalgia (muscle pain) and fever. In turn, fever causes a change in body temperature, represented by the random variable TEMP. Finally, pneumonia (PN) is another cause of fever.

Associated with a Bayesian belief network is a joint probability distribution Pr, defined in terms of conditional probability tables according to the structure of the graph. For example, for Fig. 2.3, the conditional probability table

$$\Pr(\text{FE} \mid \text{FL}, \text{PN})$$

has been assessed with respect to all possible values of the variables FE, FL and PN. In general, the graph associated with a Bayesian belief network mirrors the (in) dependences that are assumed to hold among variables in a domain. For example, given knowledge about presence or absence of fever, neither additional knowledge of flu nor of pneumonia is able to influence the knowledge about body temperature, since it holds that TEMP is conditionally independent of both PN and FL given FE.

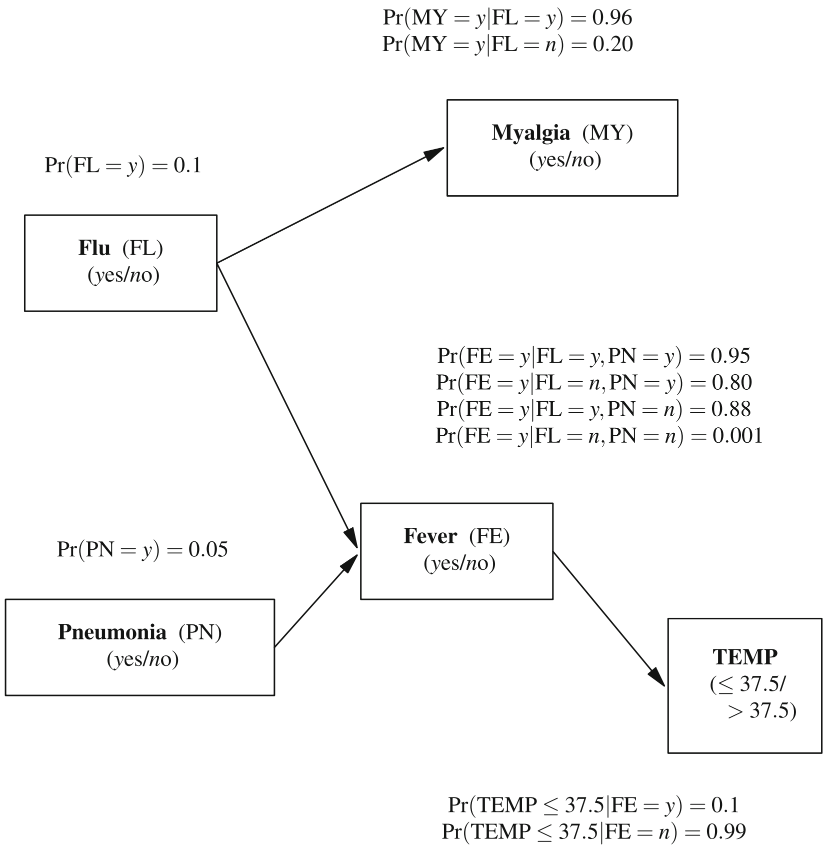

Fig. 2.3 Bayesian network $\mathscr{B} = (G, \Pr)$ with associated joint probability distribution Pr (only probabilities $\Pr(X = y \mid \pi(X))$ are shown, as $\Pr(X = n \mid \pi(X)) = 1 - \Pr(X = y \mid \pi(X))$).

For a joint probability distribution defined in accordance with the structure of a Bayesian network, it, therefore, holds that:

$$\Pr(V_1, \ldots, V_n) = \prod_{i=1}^{n} \Pr(V_i \mid \pi(V_i))$$

where V_i denotes a random variable associated with an identically named node, and $\pi(V_i)$ denotes the parents of that node. As a consequence, the amount of probabilistic information that must be specified, exponential in the number of variables in general when ignoring the independencies represented in the graph, is greatly reduced.

By means of special algorithms for probabilistic reasoning – well-known are the algorithms by Pearl [17] and by Lauritzen and Spiegelhalter [14] – the marginal probability distribution $\Pr(V_i)$ for every variable in the network can be computed; this is shown for the fever network in Fig. 2.4. In addition, a once constructed Bayesian belief network can be employed to enter and process data of a specific case, i.e. specific values for certain variables, like TEMP, yielding an updated network. Figure 2.5 shows the updated Bayesian network after entering evidence about a patient's body temperature into the network shown in Fig. 2.3. Entering evidence in a network is also referred to as *instantiating* the network. The resulting probability distribution of the updated network, $\Pr^E(V_i)$, which is a marginal probability distribution of the probability distribution $\Pr^E$, is equal to the posterior of the original probability distribution of the same variable, conditioned on the evidence E entered into the network:

$$\Pr^E(V_i) = \Pr(V_i \mid E)$$

Bayesian belief networks have also been related to logic by so called *probabilistic Horn clauses*. This formalism offers basically nothing else then a recipe to obtain a logical specification of a Bayesian belief network. Reasoning with probabilistic Horn clauses is accomplished by logical abduction; the axioms of probability theory are used to compute an updated probability distribution.

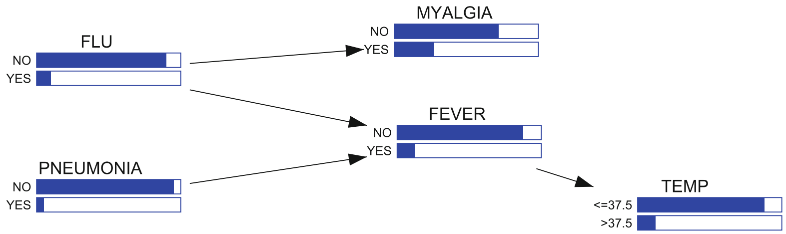

Fig. 2.4 Prior marginal probability distributions for the Bayesian belief network shown in Fig. 2.3.

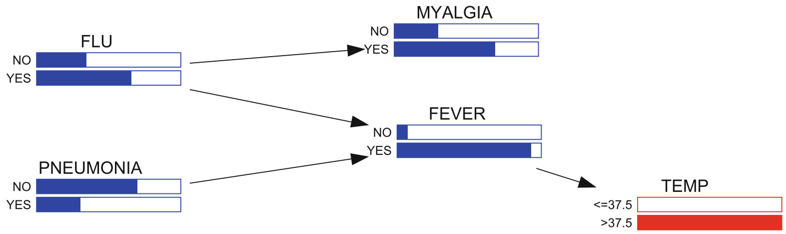

Fig. 2.5 Posterior marginal probability distributions for the Bayesian belief network after entering evidence concerning body temperature. Note the increase in probabilities of the presence of both flu and pneumonia compared to Fig. 2.4. It is also predicted that it is likely for the patient to have myalgia.

2.4.2 Probabilistic Logic

There have been different recent proposals in the AI literature to combine logic and probability theory, where usually predicate logic is combined with probabilistic graphical models. David Poole has developed so-called *independent choice logic* (which later was integrated into AIlog). It combined Prolog-like logic with Bayesian networks. Another approach, developed by Williamson et al. makes use of *credal networks*, which are similar to Bayesian networks but reason over probability intervals instead of probabilities. The last few years *Markov logic* has had an enormous impact on the research area. The idea is to use predicate logic to *generate* Markov networks, i.e., joint probability distributions that have an associated *undirected* graph. Formalisms such as independent choice logic and Markov logic are examples of what is called *probabilistic logic*.

Various probabilistic logics, such as the independent choice logic, are based on logical abduction. The basic idea of these kind of logics is to define the probability of a query in terms of the probability of its *explanations* (sometimes called a *prediction* in theory of logical abduction) of a certain query (cf. Sect. 4.5) given a logic program. Probability of the explanations are defined by a very simple distribution, namely by a set of independent random variables, which makes it possible to (relatively) efficiently compute a probability. The nice thing about this approach is that it truly combines logical reasoning (finding the explanations) with probabilistic reasoning (computing the probability of the set of explanations).

Defining the probability distributions over the explanations is done by associating probabilities to hypotheses in a set Δ. In order to make sure that we end up with a valid probability distribution, we require a partitioning of this set into subsets $\Delta_1, \ldots, \Delta_n$, i.e., such that it holds that:

$$\bigcup_{i=1}^{n} \Delta_i = \Delta$$

and $\Delta_i \cap \Delta_j = \emptyset$ for all $i \neq j$. Each possible grounding of Δ_i, i.e. $\Delta_i\sigma$ with σ a substitution, is associated to a random variable $X_{i,\sigma}$, i.e., $dom(X_{i,\sigma}) = \Delta_i\sigma$. While you could imagine that every random variable is different, here we will assume that every grounding of $h \in \Delta$ has to have the same probability, i.e., for all substitutions σ, σ':

$$P(X_{i,\sigma} = h\sigma) = P(X_{i,\sigma'} = h\sigma')$$

whereas each pair of random variables as we have just defined is assumed to be independent, the hypotheses *in the same partition* are dependent. Suppose for example, we have a random variable X with three possible hypotheses:

$$dom(X) = \{\mathit{influenza}, \mathit{sport}, \mathit{not\_sport}\_\text{or}\_\mathit{influenza}\}$$

In each possible state (element of the sample space), each random variable is exactly in one state at the time, i.e., in this case, we assume that we either have influenza, or we sport, or neither, but we do not sport while we have influenza. In other words: sport and influenza are considered to be inconsistent.

To understand the space of explanations that we may consider is by picking a possible value for each random variable. In the language of the *independent choice logic*, this is called a *choice* (hence, the name). In order to make this work probabilistically, we need some slight restrictions on our logic program. First, it is not allowed to have two hypotheses in Δ that unify. Further, it is not allowed that an element from Δ unifies with a head of one of the clauses. Finally, mostly for convenience here, we will restrict ourselves to acyclic logic programs consisting of Horn clauses and substitutions that can be made using the constants in the program.

The probability distribution over Δ is now used to define a probability for arbitrary atoms. As mentioned earlier, this will be defined in terms of explanations, which are slightly different than we have seen before due to the probabilistic semantics. Given a causal specification $\Sigma = (\Delta, \Phi, \mathscr{R})$, a (probabilistic) explanation $E \subseteq \Delta\sigma$ for some formula $F \in \Phi$ is:

$$\mathscr{R} \cup E \models F$$
$$\mathscr{R} \cup C \cup E \not\models \bot$$

where

$$C = \{\bot \leftarrow h_1, h_2 \mid \Delta_i \text{ is one of the partitions of } \Delta, h_1, h_2 \in \Delta_i\}$$

and $\Delta\sigma$ grounded. Note that the consistency condition entails that we only pick at most one value for each random variable. The intuitive assumption that is now being made is that an atom is true if and only if at least one of its (grounded) explanations is true. Suppose $\mathscr{E}(F)$ is the set of all explanations for F, then we define:

$$F = \bigvee_{E_i \in \mathscr{E}(F)} E_i$$

Notice that this definition is equivalent to assuming Clarke's completion of the given theory (cf. Sect. 4.3.1).

Recall that an explanation E is called minimal if there does not exist an explanation E' such that $E' \subset E$. It is not difficult to see that we can restrict our attention to the set of *minimal explanations* $\mathscr{E}_m(F)$: by logical reasoning it holds that, if $E' \subset E$ then $E' \vee E = E'$, so it can be shown that $\mathscr{E}(F) = \mathscr{E}_m(F)$. We then have:

$$F = \bigvee_{E_i \in \mathscr{E}_m(F)} E_i$$

Again, there is a close connection to the semantics of abduction, as $\bigvee_{E_i \in \mathscr{E}_m(F)} E_i$ is sometimes referred to as the *solution formula*. Of course, if two things are equal, then their probability must be equal:

$$P(F) = P(\bigvee_{E_i \in \mathscr{E}_m(F)} E_i)$$

It is now clear how we can solve the problem of computing the probability of F: first we find the (minimal) explanations of F and then we use the probability distribution defined over the hypotheses to compute the disjunction of the explanations.

Consider the causal specification $\Sigma = (\Delta, \Phi, \mathscr{R})$, with

$$\Delta = \{influenza, sport, not\_sport\_or\_influenza, \alpha_1, not\_\alpha_1, \alpha_2, not\_\alpha_2\}$$

and

$$\Phi = \{chills, thirst, myalgia\}$$

and the set of logical formulae $\mathscr{R}$ as presented in Fig. 2.2.

First we need to define a probability distribution over Δ. For example, we may assume to have three independent random variables X, Y, Z, such that:

$$\begin{aligned}
&P(X = sport) = 0.3\\
&P(X = influenza) = 0.1\\
&P(X = not_sport_or_influenza) = 0.6\\
&P(Y = \alpha_1) = 0.9\\
&P(Y = not_\alpha_1) = 0.1\\
&P(Z = \alpha_2) = 0.7\\
&P(Z = not_\alpha_2) = 0.3
\end{aligned}$$

Note that explanations containing e.g., *sport* and *influenza* are inconsistent with this probability distribution, as X can only take the value of one of them (they are mutually exclusive).

Suppose we have interested in the probability of myalgia, i.e., $P(myalgia)$. The set of all minimal explanations for myalgia, i.e., $\mathscr{E}_m(myalgia)$ is $\{E_1, E_2\}$, where:

$$\begin{aligned}
E_1 &= \{influenza, \alpha_2\}\\
E_2 &= \{sport\}
\end{aligned}$$

Clearly, there are many more explanations, e.g.,

$$\begin{aligned}
E_3 &= \{influenza, sport, \alpha_2\}\\
E_4 &= \{influenza, \alpha_1, \alpha_2\}\\
E_5 &= \{influenza, not_\alpha_1, \alpha_2\}\\
&\vdots \qquad \vdots
\end{aligned}$$

Note that for example, the set:

$$E' = \{influenza, \alpha_1, not\_\alpha_1, \alpha_2\}$$

is inconsistent, because α_1 and not_α_1 cannot both be true. Therefore, it is not an explanation.

Since we assumed that a formula is true if only if at least one of its explanations is true, the probability of myalgia is defined it terms of influenza and sport:

$$P(myalgia) = P((influenza \wedge \alpha_2) \vee sport)$$

Since $influenza \wedge \alpha_2$ and *sport* are mutually exclusive, the probability of the disjunction is the sum of the disjuncts, i.e.:

$$\begin{aligned} P(myalgia) &= P(influenza \wedge \alpha_2) + P(sport) \\ &= P(influenza)P(\alpha_2) + P(sport) \\ &= 0.1 \cdot 0.7 + 0.3 = 0.37 \end{aligned}$$

2.5 Conclusions

In this introductory chapter we have briefly reviewed the most important languages for knowledge representation as using in medicine. It is not possible given the scope of this chapter to be complete, but since logic and probability theory act as the core of the majority of the modern work on knowledge representation, this introduction will at least pinpoint the most important ideas.

References

1. Ben-Ari, M., Manna, Z., Pnueli, A.: The temporal logic of branching time. Acta Inform. **20**, 207–226 (1983)
2. Brachman, R.J.: What IS-A is and isn't: an analysis of taxonomic links in semantic networks. IEEE Comput. **16**(10), 30–36 (1983)
3. Buchanan, B.G., Shortliffe, E.H.: Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Addison-Wesley, Reading (1984)
4. Clancey, W.J.: Heuristic classification. Artif. Intell. **27**, 289–350 (1985)
5. Clarke, E.M., Emerson, E.A.: Design and synthesis of synchronization skeletons using branching time temporal logic. In: Kozen, D. (ed.) Logic of Programs. LNCS, vol. 131, pp. 52–71. Springer, Heidelberg (1981)
6. Console, L., Theseider Dupré, D., Torasso, P.: A theory of diagnosis for incomplete causal models. In: Proceedings of the 10th International Joint Conference on Artificial Intelligence, pp. 1311–1317 (1989)
7. Cooper, G.F.: A method for using belief networks as influence diagrams. In: Proceedings of the 4th Workshop on Uncertainty in Artificial Intelligence, pp. 55–63 (1988)
8. de Kleer, J., Williams, B.C.: Diagnosing multiple faults. Artif. Intell. **32**, 97–130 (1987)
9. Downing, K.: Physiological applications of consistency-based diagnosis. Artif. Intell. Med. **5**(1), 9–30 (1993)

10. Emerson, E.A., Clarke, E.M.: Characterizing correctness properties of parallel programs using fixpoints. In: de Bakker, J., van Leeuwen, J. (eds.) Automata, Languages and Programming. LNCS, vol. 85, pp. 169–181. Springer, Heidelberg (1980)
11. Forbus, K.D., De Kleer, J.: Building Problem Solvers. The MIT Press, Cambridge (1993)
12. Glymour, C., Cooper, G.F.: Computation. Causation & Discovery. MIT Press, Menlo Park (1999)
13. Hitzler, P., et al. (eds.): OWL 2 Web Ontology Language: Primer. W3C Recommendation, 27 October 2009. http://www.w3.org/TR/owl2-primer/
14. Lauritzen, S.L., Spiegelhalter, D.J.: Local computations with probabilities on graphical structures and their application to expert systems. J. R. Stat. Soc. (Ser. B) 50, 157–224 (1987)
15. Lucas, P.J.F.: The representation of medical reasoning models in resolution-based theorem provers. Artif. Intell. Med. **5**(5), 395–414 (1993)
16. W3C OWL Working Group. OWL 2 Web Ontology Language: Document Overview. W3C Recommendation, 27 October 2009. http://www.w3.org/TR/owl2-overview/
17. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, Palo Alto (1988)
18. Peng, Y., Reggia, J.A.: Abductive Inference Models for Diagnostic Problem Solving. Springer, New York (1990)
19. Pnuelli, A.: A temporal logic of concurrent programs. Theor. Comput. Sci. **13**, 45–60 (1981)
20. Poole, D.: Explanation and prediction: an architecture for default and abductive reasoning. Comput. Intell. **5**(2), 97–110 (1989)
21. Pople, H.E.: Heuristic methods for imposing structure on ill-structured problems: the structure of medical diagnostics. In: Szolovits, P. (ed.) Artificial Intelligence in Medicine. AAAS Selected Symposium, vol. 51, pp. 119–190. Westview Press, Boulder (1982)
22. Reiter, R.: A theory of diagnosis from first principles. Artif. Intell. **32**, 57–95 (1987)
23. Szolovits, P.: Artificial Intelligence in Medicine. AAAS Selected Symposium, vol. 51. Westview Press, Boulder (1982)