

# An Intelligent System for Pacemaker Reprogramming\*

Peter Lucas

Department of Computer Science, Utrecht University  
P.O. Box 80089, 3508 TB Utrecht, The Netherlands  
E-mail: lucas@cs.uu.nl

Astrid Tholen and Geeske van Oort

Research and Development Department, Vitatron Medical B.V.  
P.O. Box 76, 6950 AB Dieren, The Netherlands  
E-mail: {Astrid.Tholen,Geeske.van.Oort}@vitatron.com

## Abstract

The process of reprogramming a cardiac pacemaker can be described in terms similar to those used for describing diagnostic problem solving. In this paper, the process of reprogramming a pacemaker is formalised as a special form of abductive diagnostic reasoning, where observable findings are interpreted with respect to results obtained from diagnostic tests. The dynamics of this process is cast as a diagnostic strategy, where information is gathered in a structured fashion. This abductive theory of pacemaker reprogramming has been used as the basis for an actual system that in its present form is capable of assisting cardiologists in dealing with problems in atrial sensing and pacing. The performance of the system has been evaluated using data from actual patients.

*Keywords:* medical decision support, medical expert systems, pacemaker programming, model-based reasoning, abductive reasoning.

## 1 Introduction

Cardiac disease in a patient may cause the excitatory and conductive system of the heart to fail, resulting in an abnormal cardiac rhythm, usually referred to as *arrhythmia*. Some arrhythmias are very dangerous, and may lead to death of the patient; others may be the origin of symptoms and signs that are less threatening, but for which medical treatment is nevertheless required. One of the possible treatments for these patients is assistance by a cardiac pacemaker.

Modern pacemakers are sophisticated electronic devices, capable of providing assistance on demand, i.e. when the excitatory and conductive system of the heart fails to operate normally. In order to accommodate specific patient needs, modern pacemakers can be programmed by setting particular parameters in such way that the resulting pacemaker therapy is optimal for the patient. Changing previously set pacemaker parameters is called *reprogramming*.

---

\*Published as: P.J.F. Lucas, A.M. Tholen and G. van Oort (1999). An intelligent system for pacemaker reprogramming. *Artificial Intelligence in Medicine*, **17**: 249–269.

Unfortunately, reprogramming a pacemaker is not an easy task: both sufficient time and knowledge of pacemaker functionality and possible pacemaker therapy must be available. It has been observed that due to a lack of one or both of these factors, in many patients, a given pacemaker therapy is suboptimal: in many implanted pacemakers even the original factory settings are kept unchanged. On the one hand, pacemaker technology is moving fast, and the role of software in pacemaker functioning is increasing, yielding pacemaker devices that are almost annually enhanced in their capabilities. On the other hand, pacemaker equipment producers are confronted with the limitations of what clinicians can and are willing to do; they are beginning to realise that some form of intelligent decision support is needed in order to let patients benefit from further advances in pacemaker technology.

The process of reprogramming a pacemaker consists of observing symptoms and signs in the patient, and collecting information of past and present electrical behaviour of heart and pacemaker, stored within the device. Based on this information, decisions can be made with respect to desirable changes to pacemaker settings. The entire process has much in common with the process of diagnostic problem solving, which consists of: (1) observing findings, (2) finding possible causes of observed abnormality, and (3) suggesting tests, that, when carried out, accomplish to discriminate between the causes of the problems suspected so far. In this article, it is shown that the problem of reprogramming a pacemaker can indeed be viewed in such diagnostic terms.

Both cardiac pacemakers and the excitatory and conductive system of the heart have a clear physical structure, which seems to make the problem amenable to model-based techniques. Model-based techniques are characterised by using explicit representations of structure and behaviour. In particular the theory of model-based diagnosis seems to offer an attractive set of tools. In this article, we present a theory of model-based diagnosis, based on earlier work in the field of abductive diagnosis, that has been adapted to deal with problems of pacemaker reprogramming. Based on this theory, a system has been implemented in order to assist in reprogramming a pacemaker. It is currently being extended for use in a pacemaker programmer. The resulting system appears to be one of few medical, model-based intelligent systems that have moved from academia to industry.

The structure of this paper is as follows. In following section, we summarise some basic (patho)physiology of the heart, so far as needed for the understanding of this article. Next, in Section 3, the principles of the structure and function of pacemakers are briefly discussed. In Section 4, we present a theory of model-based diagnosis used for pacemaker reprogramming. Section 5 contains a description of the implementation of the actual system. Preliminary evaluation results of the system are discussed in Section 6. Finally, in Section 7 we summarise what has been achieved by the research discussed in this paper, and mention some future work.

## 2 Structure and function of the heart

The heart can be viewed as a pump, responsible for maintaining blood pressure and flow within the vascular system. A continuous blood flow is required in order to deliver essential nutrients, like oxygen and glucose, to the tissues and to remove waste products, like carbon dioxide and urea, from the tissues to various organs for further degradation and clearance. Pressure and flow are the results of a rhythmic contraction of the cardiac muscle, the *myocardium*, under control of specialised excitatory and conductive cardiac tissue. We shall briefly review the structure and function of this tissue, and mention some of the disorders associated with it.

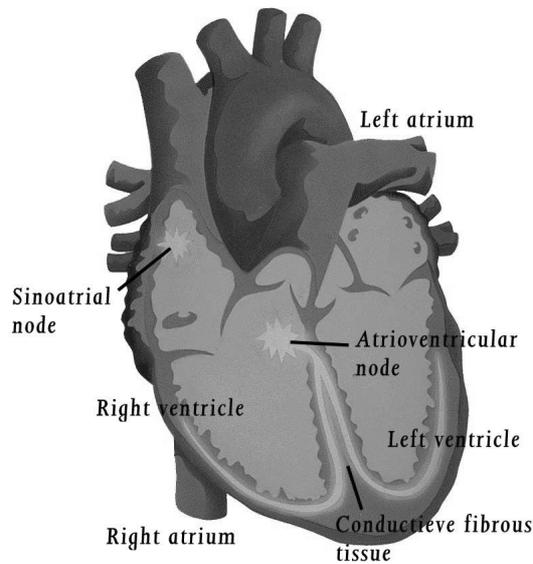


Figure 1: The excitatory and conductive system of the heart.

## 2.1 The excitatory and conductive system of the heart

The excitatory and conductive system of the heart, as schematically shown in Figure 1, consists of several parts, with separate, but related, functionality. The normal human heart beat is under control of the *sinoatrial node*, or *sinus node*, a small strip of specialised self-excitatory tissue, located in the wall of the right atrium. The sinoatrial node fires rhythmically at a rate of approximately 75 beats per minute. It generates an electrical impulse, called an *action potential*, that spreads through the atrial muscular wall, causing the atrial muscle to contract. Next, the impulse travels through a muscular fibre pathway to the right ventricle, where the impulse causes the *atrioventricular node* to fire. The generated impulse travels fast through the right and left bundle branches (indicated in the figure as two lightly shaded curves, originating in the atrioventricular node), which consist of specialised conductive tissue, to the muscular tissue of the right and left ventricles through so-called Purkinje fibres. The ventricular muscular tissue responds by a contraction (*systole*).

## 2.2 Bradycardia

There are a number of diseases that may cause the excitatory and conductive system of the heart to fail. This may give rise to a heart rate that is too low, referred to as *bradycardia*. Bradycardia may be caused by a variety of disorders, such as *sinus node dysfunction*. The term *sick sinus syndrome* refers to a combination of symptoms, like dizziness, fatigue, fainting and heart failure, due to sinus node dysfunction. There are a number of other disorders of the excitatory and conductive system of the heart, causing different parts of the system to fail, giving rise to bradycardia as well. Failure of the atrioventricular pathway to conduct electrical impulses from the atrium to the ventricle, called *atrioventricular block*, is a cause of bradycardia that is rather common. In that case, only the frequency of ventricular contraction is decreased. Long-term treatment of bradycardia is accomplished by an implantable pacemaker, although there is some place for pharmacological therapy in the early stages [5].



Figure 2: The Diamond II pacemaker.

### 3 Cardiac pacemakers

A pacemaker is capable of taking over control of the rhythmic contraction of the cardiac muscle, thus replacing the function of the natural pacemaker, i.e. the sinoatrial node, or of parts of the conductive pathways.

#### 3.1 Structure of a pacemaker

A pacemaker consists of a *can*, which contains a microprocessor, a battery and an impulse generator. Impulses are transmitted to the heart by means of a *lead*, which is attached to the can's connector. A lead is either unipolar or bipolar; a *unipolar* lead contains one insulated coil, whereas a *bipolar* lead contains two coils, separated by an inner insulation. An outer insulation shields a lead from the environment. The tip of a lead, which contains an electrode, is implanted into the inner, endocardial surface of the heart; the actual location depends on the type of pacemaker. The pacemaker can is usually implanted in the pectoral region, with the lead running through the right subclavian vein to the internal surface of the heart. An example of a modern, advanced pacemaker is Vitatron's *Diamond II* pacemaker (See Figure 2).

A pacemaker is programmed by means of a *programmer*, a computer with a special user interface for data entry and display, and with special software to control an external magnet that communicates with the pacemaker. Figure 3 shows an example of a pacemaker programmer. The magnet is placed above the location of the pacemaker; information from the programmer to the pacemaker, and back, is transmitted by means of telemetry.

#### 3.2 Functions of a pacemaker

A modern pacemaker is not only capable of stimulating, or pacing, the heart, but also of sensing the intrinsic activity of the heart. Sensed activity is used as information for the pacemaker to adopt appropriate stimulating activity as therapy. Modern pacemakers are even capable of adapting their pacing rate, dependent on the patient's needs; this capability is called *rate responsiveness*.

A pacemaker, like the Diamond II, stores and collects a lot of information, called *diagnostics*, that may be used to diagnose problems:

- *patient-specific information*, like the patient's name, age, and date of implantation;



Figure 3: Pacemaker programmer.

- *counters* collect information on the frequency of occurrence of certain events;
- *histograms* offer graphical information about the distribution of certain events;
- *holters* collect information about certain events over a particular period of time.

In addition to the diagnostics, the programmer also shows the programmed *pacemaker settings*, which determine the operation of the pacemaker. For example, the counter *AV synchrony percentage* expresses the percentage of cardiac events in which an atrial contraction is followed by a ventricular contraction; the *P-wave histogram* offers information of the frequency distribution of P-wave amplitudes, i.e. information of electrical events associated with atrial contractions; the *24-hour holter* contains information about the mean heart rate, measured every 6 minutes over the last 24 hours.

### 3.3 Pacemaker problems

Cardiac symptoms and signs in a patient with an implanted pacemaker can be due to medical problems, inappropriate pacemaker settings, or pacemaker faults. In this section we focus on problems with atrial sensing and pacing, although this is just a selection of problems that may occur.

There are three main causes of atrial sensing and pacing problems:

- *atrial undersensing*: impulses generated by the sinoatrial node have not been sensed, causing the pacemaker to give inappropriate therapy;
- *atrial oversensing*: the pacemaker senses and reacts to a signal, which, however, has not been generated by the sinoatrial node, for example muscle noise;

- *loss of atrial capture*: the pacemaker produces an electrical impulse, which fails to result in an atrial contraction.

In turn, atrial under- and oversensing, and loss of atrial capture are caused by a number of different problems. For example, atrial undersensing may be due to problems such as atrial lead dislocation (the lead’s tip has lost contact with the surface), inner or outer insulation break, a lead conductor fracture – conduction of the pacemaker impulse is then impossible –, or a connector problem due to the lead not being attached sufficiently tight to the connector part of the pacemaker. A typical cause of atrial oversensing is electromagnetic interference by an external source. One of the possible causes of loss of atrial capture is an increase in the energy threshold, due, for example, to a fibrous tissue barrier around the lead’s tip, which obstructs impulse conduction.

The three types of problems mentioned above may give rise to what is called the *pacemaker syndrome*: the patient feels a beat in the neck, due to the regurgitation of ventricular blood through the atrioventricular valves back into the atria, caused by asynchrony.

### 3.4 Diagnostic tests

When particular problems with a pacemaker are suspected by the cardiologist, diagnostic tests may be carried out in order to obtain further information about possible misprogrammings. Some tests can be performed directly, yielding immediate test results, whereas others results are available only after some delay.

An example of a test that offers immediate results is the measurement of the atrial lead impedance. An increase in the lead impedance may be due to an atrial lead conductor fracture; a decrease in the lead impedance may be due to an insulation break.

A test that takes more time, yielding results that are usually available only at follow-up, is, for example, an X-ray of the chest requested to confirm a conductor break.

## 4 Model-based pacemaker reprogramming

Pacemaker reprogramming can be viewed as the process of finding appropriate values for pacemaker settings that avoid the occurrence of abnormal symptoms and signs in the patient. This process can be seen as a form of diagnostic problem solving: when there are particular symptoms or signs in the patient, indicating suboptimal pacemaker settings, pacemaker faults or a medical disorder, the possible causes should be determined and dealt with. The theory of model-based diagnosis offers several ways in which such a diagnostic process can be described. Conceptually, this diagnostic process may be described in terms of *matching abnormal behaviour* (MAB) diagnosis, as schematically shown in Figure 4 [8]. In MAB diagnosis, there is a model of abnormal behaviour available, which is used to predict abnormal findings that may or must be observed; these predictions are next matched with findings actually observed. A collection of causes described in the model and associated with predictions that best match the findings observed, is taken as a diagnosis. MAB diagnosis is a conceptual model; it is typically formalised in terms of abductive reasoning.

### 4.1 Theory of abductive diagnosis

The design of our theory of abductive diagnosis was inspired by earlier work by L. Console and P. Torasso [1, 2]. In their theory of diagnosis, the abnormal behaviour of a system is

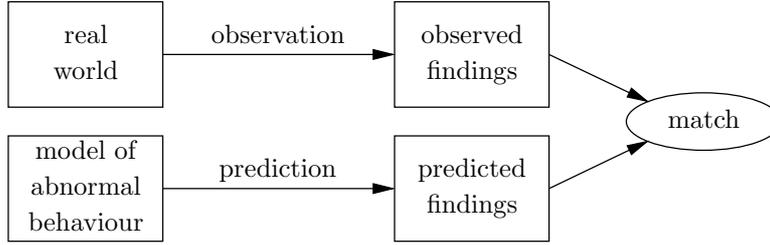


Figure 4: Matching-abnormal-behaviour diagnosis.

represented as causal knowledge, relating abnormal states and resulting abnormal findings. We shall refer to abnormal states in this article as *defects*, which may be anything, varying from medical disorders, to incorrect pacemaker settings and pacemaker faults.

In this paper, it is assumed that causal knowledge can be represented in Horn formulae of the following two forms:

$$d_1 \wedge \dots \wedge d_n \rightarrow f \quad (1)$$

$$d_1 \wedge \dots \wedge d_n \rightarrow d \quad (2)$$

where  $d, d_i, i = 1, \dots, n$ , are positive literals representing defects, and  $f$  is a positive literal representing an *observable finding*.

Console and Torasso have also proposed a simple mechanism to weaken the causality relation, by means of literals  $\alpha$ . These literals represent incompleteness of knowledge with respect to the underlying causal mechanisms relating causes and effects. They can be used to block the deduction of a finding  $f$  or defect  $d$  if the defects  $d_i, i = 1, \dots, n$ , hold true, but the literal  $\alpha$  is assumed to be false. The weakened Horn formulae have the following form:

$$d_1 \wedge \dots \wedge d_n \wedge \alpha_f \rightarrow f \quad (3)$$

$$d_1 \wedge \dots \wedge d_n \wedge \alpha_d \rightarrow d \quad (4)$$

The literals  $\alpha$  are called *incompleteness-assumption literals*, abbreviated to *assumption literals*; every assumption literal occurs uniquely within a Horn formula. In the following, the convention is adopted that present findings are denoted by positive finding literals; absent findings are denoted by negative finding literals. A similar convention is used for defect literals.

Now, let  $\mathcal{C} = (\Delta, \Phi, \mathcal{R})$  stand for a *causal specification*, where:

- $\Delta$  denotes a set possible (positive and negative) defect and assumption literals;
- $\Phi$  denotes a set of possible (positive and negative) finding literals;
- $\mathcal{R}$  stands for a set of Horn formulae of the form (1) – (4), representing a causal model of abnormal behaviour.

A causal specification can be employed for the prediction of observable findings in the sense of Figure 4.

**Definition 1.** If  $\mathcal{C} = (\Delta, \Phi, \mathcal{R})$  is a causal specification then, a set  $H \subseteq \Delta$  is called a *prediction* for a set of observable findings  $F \subseteq \Phi$  if

- (1)  $\mathcal{R} \cup H \models F$ , and

(2)  $\mathcal{R} \cup H$  is satisfiable.

Obviously, the notion of prediction formalises the arrow in the lower half of Figure 4; the resulting set of findings  $F$  corresponds to the predicted (observable) findings in the same figure.

An *abductive diagnostic problem*  $\mathcal{A}$  is now defined as a pair  $\mathcal{A} = (\mathcal{C}, E)$ , where  $E \subseteq \Phi$  is called a *set of observed findings* if  $E$  is consistent. A set of observed findings corresponds to the box in the upper half of Figure 4.

Formally, a solution to an abductive diagnostic problem  $\mathcal{A}$  can be defined as follows.

**Definition 2.** Let  $\mathcal{A} = (\mathcal{C}, E)$  be an abductive diagnostic problem, where  $\mathcal{C} = (\Delta, \Phi, \mathcal{R})$  is a causal specification with  $\mathcal{R}$  a set of Horn formulae of the form (1) – (4), and  $E \subseteq \Phi$  a set of observed findings. A set of defect and assumption literals  $H \subseteq \Delta$  is called a *solution* to  $\mathcal{A}$  if:

- (1)  $\mathcal{R} \cup H \models E$  (*covering condition*);
- (2)  $\mathcal{R} \cup H \cup C \not\models \perp$  (*consistency condition*)

where  $H$  is minimal with respect to set inclusion, and  $C$ , called the *constraint set*, is a set of formulas in first-order logic, consisting of defect and finding literals only.

There are many possible ways for defining a constraint set  $C$ . In [1], the constraint set  $C$  is defined as follows (we take  $C = C_1$ ):

$$C_1 = \{\neg f \in \Phi \mid f \in \Phi, f \notin E, f \text{ is a positive literal}\}$$

i.e. the constraint set stands for findings assumed to be false, because they have not been observed (and are therefore assumed to be absent); this is an application of the closed world assumption (CWA) [12], restricted to observable findings. Note that in this case, the sets  $E$  and  $C$  are disjoint, and it holds that if  $f \in E$  then  $\neg f \notin C$ .

The covering condition (1) ensures that sufficient defect and assumption literals are assumed to account for all given observed findings. The consistency condition (2) helps to ensure that not too many defect and assumption literals are assumed. It may only be necessary to include an assumption literal  $\alpha$  in a solution for implications  $d \wedge \alpha_f \rightarrow f$  and  $d \wedge \alpha_{d'} \rightarrow d'$  if the defect  $d$  is deducible from the assumed (initial) defects and assumption literals. The condition of minimality with respect to set inclusion of a solution  $H$  implies that no more defect and assumption literals will be assumed as part of a solution than are required to predict findings observed.

We have as yet not defined the concept of ‘diagnosis’. An entire solution  $H$  may be taken as a diagnosis, but it is more natural to take a *diagnosis* to consist of the defect literals in a solution  $H$  only; we shall do so accordingly.

In the consistency condition as defined above, it is assumed that if a finding is not included in the set of observed findings, it is assumed to be absent. However, it may not always be justified to assert negative findings in this manner; sometimes, it is more natural to take the findings as being unknown. Hence, the definition of the constraint set  $C$  given above may be too strong for practical purposes. When we use predicate logic and let particular predicates associated with findings stand for tests, the definition of  $C$  above might be replaced by the following definition:

$$C_2 = \{\neg\pi(t) \in \Phi_N \mid \pi(s) \in E, t \neq s, \text{ or } \neg\pi(t) \text{ has been observed}\}$$

where  $\pi$  stands for predicate symbols, and  $t$  and  $s$  are constants. The consistency condition remains the same, but its effects on the computation of a diagnosis differs, because of the altered definition of the constraint set  $C$ . For example, when we have tests  $p$ ,  $q$  and  $r$ , then with the first definition of a constraint set, the set consists of all negative finding literals  $\neg p(t)$ ,  $\neg q(t')$ ,  $\neg r(t'')$  that have not been observed, even if particular tests have not been done. With the second definition, only negative literals concerning tests done are included, if not positively observed, supplemented with findings explicitly observed to be absent. Other findings are assumed to be unknown.

Now, if some defect  $d$  is included in a solution  $H$  and

$$\mathcal{R} \cup \{d\} \models f$$

where  $f \notin E$ , this means that the model predicts that if the test is actually carried out, the finding  $f$  will be observed. This, in fact, may be used as a basis for diagnostic problem solving by suggesting to the user to carry out particular tests.

In the domain of pacemaker reprogramming it was known beforehand that particular combinations of defects cannot occur. Such impossible combinations can be represented as a set of additional domain constraints  $D$ , e.g.  $\neg(d_1 \wedge d_2)$  indicates that  $d_1$  and  $d_2$  may not occur together, imposing a further limitation on the number of possible solutions. The final definition of the constraint set that appeared suitable in the present case was therefore as follows:

$$C = C_2 \cup D$$

## 4.2 Interpretation of observables

The description above suggests a dynamic, diagnostic process where preliminary diagnoses and the proposal of additional tests are generated by the system; when new test information becomes available, the old diagnoses may need revision. The manner in which old diagnoses are revised, is the subject of the present section.

### 4.2.1 Proposing tests

Using the definition of abductive diagnosis given above, and using the last definition of a constraint set as consisting of absent findings, either observed or inferred, it is straightforward to add test selection to the abductive reasoning scheme considered so far. Let  $\mathcal{A} = (\mathcal{C}, E)$  be an abductive diagnostic problem, with  $E$  the set of observed findings and  $\mathcal{C} = (\Delta, \Phi, \mathcal{R})$  a causal specification, then if  $H$  is a solution to  $\mathcal{A}$  it may hold that  $\mathcal{R} \cup H \models F$ , where  $F \supset E$ . In this case, the solution  $H$  predicts findings that have not, as yet, been observed. Nevertheless, all findings resulting from tests are included, either positively in  $E$  or negatively in  $C$ . Hence, the observable findings in the difference set  $F \setminus E$  pertain to tests that have not yet been carried out. Using this information, the system may suggest to the user to perform particular diagnostic tests.

### 4.2.2 Hypothesis revision

Now, suppose that a particular test, as suggested by the system, has been carried out, and that a (positive) test result has been entered into the system. Of interest are the effects of this additional information on the diagnostic solutions. The following lemma concerns the

situation in which the test result corresponds exactly to the finding previously predicted by a diagnostic solution.

**Lemma 1.** *Let  $\mathcal{A} = (\mathcal{C}, E)$  be an abductive diagnostic problem with causal specification  $\mathcal{C} = (\Delta, \Phi, \mathcal{R})$ , and let  $\mathcal{A}' = (\mathcal{C}, E')$  be an abductive diagnostic problem, such that  $E' = E \cup \{\pi(t_i)\}$ ,  $\pi(t_i) \notin E$ . Furthermore, let the constraint set of  $\mathcal{A}'$  be equal to  $C' = C \cup \{\neg\pi(t_1), \dots, \neg\pi(t_{i-1}), \neg\pi(t_{i+1}), \dots, \neg\pi(t_n)\}$ , where  $C$  is the constraint set of  $\mathcal{A}$ . Finally, let  $H$  be a solution to  $\mathcal{A}$ , such that  $\mathcal{R} \cup H \models \pi(t_i)$ . Then,  $H$  is also a solution to  $\mathcal{A}'$  if  $\mathcal{R} \cup H \cup C' \not\models \perp$ .*

**Proof.** The proof is a straightforward check against the definition of a solution (Definition 2).

This lemma suggests that it is sufficient to check the satisfiability of the consistency condition as soon as information corresponding to a suggested diagnostic test result becomes available.

The next lemma concerns the situation where the actually observed finding turns out to be different from the one previously predicted.

**Lemma 2.** *Let  $\mathcal{A} = (\mathcal{C}, E)$  be an abductive diagnostic problem with causal specification  $\mathcal{C} = (\Delta, \Phi, \mathcal{R})$ , and let  $\mathcal{A}' = (\mathcal{C}, E')$  be an abductive diagnostic problem, such that  $E' = E \cup \{\pi(t_j)\}$ ,  $\pi(t_j) \notin E$ , and  $C' = C \cup \{\neg\pi(t_1), \dots, \neg\pi(t_i), \dots, \neg\pi(t_{j-1}), \neg\pi(t_{j+1}), \dots, \neg\pi(t_n)\}$ , where  $C$  and  $C'$  are the constraint sets of  $\mathcal{A}$  and  $\mathcal{A}'$ , respectively. Furthermore, let  $H$  be a solution to  $\mathcal{A}$ , such that  $\mathcal{R} \cup H \models \pi(t_i)$ ,  $i \neq j$ . Then, for any solution  $H'$  it holds that  $H' \not\supseteq H$  and  $H' \not\subseteq H$ .*

**Proof.** According to the premise, it must hold that  $\mathcal{R} \cup H' \models \pi(t_j)$ , with  $j \neq i$ , and  $\mathcal{R} \cup H' \models E$ , whereas  $\mathcal{R} \cup H \models \pi(t_i)$  and  $\mathcal{R} \cup H \models E$ . From this and the monotonicity of the entailment relation  $\models$ , it follows that  $H' \not\supseteq H$ , otherwise  $\mathcal{R} \cup H' \cup C \models \perp$ . However, since  $\mathcal{R} \cup H \models E$ , with  $H$  minimal with respect to set inclusion,  $H'$  cannot be a subset of  $H$  either. Hence, it holds as well that  $H' \not\subseteq H$ .

This lemma is rather weak, but note that from this lemma, it follows that a solution  $H'$  should be different at least in one defect  $d$  from a previous solution  $H$  if the test result observed is different from the one predicted.

**Lemma 3.** *Let  $\mathcal{A} = (\mathcal{C}, E)$  be an abductive diagnostic problem with causal specification  $\mathcal{C} = (\Delta, \Phi, \mathcal{R})$ , and let  $\mathcal{A}' = (\mathcal{C}, E')$  be an abductive diagnostic problem, such that  $E' = E \cup \{\pi(t_i)\}$ ,  $\pi(t_i) \notin E$ . Furthermore, let  $C$  be the constraint set of  $\mathcal{A}$  and  $C'$  be the constraint set of  $\mathcal{A}'$ , where  $C' = C \cup \{\neg\pi(t_1), \dots, \neg\pi(t_{i-1}), \neg\pi(t_{i+1}), \dots, \neg\pi(t_n)\}$ . If  $H$  is a solution to  $\mathcal{A}'$ , then  $H$  is also a solution to  $\mathcal{A}$  if for each  $H' \subset H$  it holds that  $\mathcal{R} \cup H' \not\models E$ .*

**Proof.** If  $\mathcal{R} \cup H \models E \cup \{\pi(t_i)\}$ , then  $\mathcal{R} \cup H \models E$ . Furthermore, if  $\mathcal{R} \cup H \cup C' \not\models \perp$ , then  $\mathcal{R} \cup H \cup C \not\models \perp$ . However, the set of defect and assumption literals  $H$  need not be minimal with respect to set inclusion. Hence,  $H$  may not be a solution to  $\mathcal{A}$ , although the covering and consistency conditions are fulfilled. This explains the inclusion of the extra condition in the premise.

Practically spoken, this last lemma implies that solutions computed by taking a test result into account will be either identical to old solutions or be supersets of old solutions.

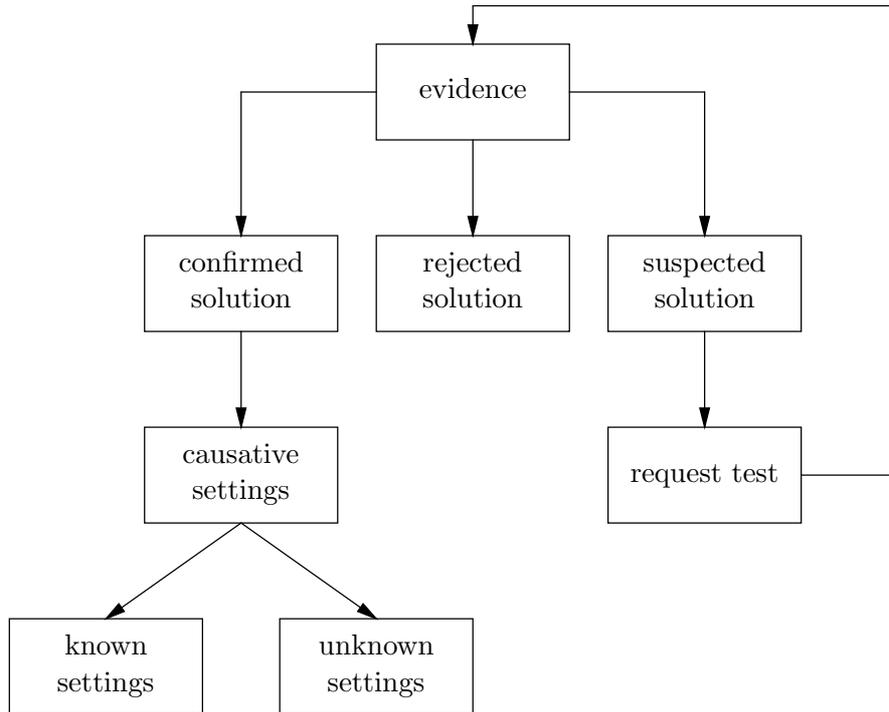


Figure 5: Problem-solving strategy.

### 4.3 Diagnostic strategy

Until now, we have refrained from making assumptions about the order in which diagnostic tests may be performed, as suggested by the system to the user. In reality, it is usually mandatory to go through the consecutive steps of the diagnostic cycle, more in particular of test selection and hypothesis generation, in a structured fashion, taking information already gathered into account.

As has been discussed above, diagnostic solutions that are causally related to findings that have as yet not been observed, may give rise to requests for further information. Since new, incoming information may or may not affect the validity of previous solutions, it seems natural to make a distinction between various types of solution. The theoretical background of this distinction has been developed above. *Suspected* solutions predict at least one finding that may be obtained by a test that has not yet been carried out. *Rejected* solutions are solutions to a previous problem, which are now rejected because of the availability of new, additional evidence, that somehow refutes the previous solution. Finally, in the case of a *confirmed* solution, all predicted, associated tests have been carried out, and the observed results appear to correspond to the results predicted. Incorporating such a distinction between various solutions in a diagnostic reasoning method is part of a *diagnostic strategy*. The overall structure of this diagnostic strategy, as applied to reprogramming a pacemaker, is depicted in Figure 5.

Strategic control of diagnostic reasoning not only concerns the process of hypothesis generation and testing, but the process of gathering relevant *evidence* as well. In diagnostic probabilistic systems, the gathering of evidence is frequently guided by taking the expected contribution of that evidence into account. A popular measure of the expected contribution

of evidence is the notion of ‘value of information’ [4]. Such a measure is not available for qualitative systems. However, as stated above, in many domains information is collected in a structured fashion, and this structure can also be used as a basis for evidence gathering. In the case of reprogramming an artificial pacemaker, the gathering of evidence may be structured in such a way that pacemaker settings and diagnostics, which are readily available from a pacemaker device, are always requested first. Next, information from the follow-up, i.e. information that requires some extra tests, yielding results that are also immediately available, are requested. Finally, information obtained from additional tests, such as a chest radiograph, could be taken as a last source of evidence.

It is possible to order resulting multiple solutions by taking the number of assumption literals occurring in each individual solution into account. This number may be taken as a simple, qualitative measure of uncertainty of a given solution. Obviously, solutions with no assumption literals included will be more likely than those with one or more assumption literals included. Since it is undesirable to neglect even unlikely solutions, this rather crude approach to the ordering of diagnostic solutions may be adequate in this problem domain. Furthermore, solutions with an equal number of assumption literals could be ordered according to the number of defect literal elements, indicating that solutions that include many defect literals are less likely than those with a few defect literals.

## 5 Resulting model and system

The resulting system consists of a causal model, an abductive inference engine and a graphical user interface. We shall not go into details concerning the user interface.

### 5.1 Causal model

Based on an analysis of causal graphs, constructed with the help of a number of pacemaker engineers and clinicians, a model of abnormal atrial behaviour was developed. The resulting causal graph is shown in Figure 6. Ellipses in the graph represent ‘defects’; tests with associated test results are denoted by rectangular vertices. Textual labels outside ellipses stand for pacemaker settings. The direction of each arc in the graph mirrors a cause–effect relationship. The interaction between vertices with outgoing arcs entering the same effect vertex is disjunctive: each of the causes is considered sufficient for producing the effect. No explicit distinction between weakly and strongly causal relationships has been indicated in the graph, but some weakly causal relations were included in the causal logical specification.

### 5.2 Abductive inference engine

The abductive inference engine of the system was built on top of the Theorist system. *Theorist* is a Prolog program, developed by David Poole and colleagues, that supports a form of hypothetical reasoning quite similar to reasoning with defaults [9, 11]. The key to mapping abductive diagnosis to reasoning in Theorist appeared to be that defect and assumption literals may be assumed during the diagnostic process; they, therefore, are interpreted in a way similar to defaults [13]. *Defaults* are rules of the form

$$\frac{P : J}{C}$$



where  $P$  is called the *prerequisite*, denoting a formula that must hold true,  $J$  is called the *justification*, standing for a formula that must be consistent with a given background logical theory, and  $C$ , called the *consequent*, will be assumed only if the prerequisite and the justification both hold. In the case of *normal* defaults there is an empty prerequisite, i.e. the prerequisite is always satisfied. For example, the assumption literals  $\alpha$  in a causal specification could be represented as a normal default:

$$\frac{: \alpha(x)}{\alpha(x)} \tag{5}$$

meaning that the literal  $\alpha(x)$  may be assumed when no contradiction arises. Suppose, for example, that we have the following Horn formula

$$d \wedge \alpha(1) \rightarrow f$$

then using the default rule (5) the literal  $\alpha(1)$  may be assumed if it is consistent with the extension of the default theory under consideration, including  $d$ , and if this assumption is required in order to account for the observed finding  $f$ .

Knowledge is represented in Theorist in terms of a set of facts  $\mathcal{F}$ , consisting of formulas in first-order logic, a set of defaults  $\Delta$ , and a set of logical constraints  $C$ . Let  $D$  be a set of ground instances of consequents of  $\Delta$ , i.e. all variables have been replaced by constants or terms only containing constants, and let  $g$  be a closed formula (a formula without free variables), then  $\mathcal{F} \cup D$  is called an *explanation* of  $g$  if:

$$(1) \mathcal{F} \cup D \models g, \text{ and}$$

$$(2) \mathcal{F} \cup D \cup C \not\models \perp$$

The logical structure of the abductive theory of diagnosis as developed in Section 4.1 appears almost identical to the theory of hypothetical reasoning underlying Theorist; however, abductive diagnosis is much less general [10, 8]. Simply by mapping a set of Horn formulas  $\mathcal{R}$  to a set of facts  $\mathcal{F}$ , and mapping a set of defects and assumption literals  $\Delta$  to defaults, the constraint set  $C$  to constraints, and by taking the set of observed findings  $E$  as  $g$ , Theorist may be used for computing diagnostic solutions. We just had to add an engine for strategic inference control to Theorist, in order to implement the reasoning strategy as described in the previous section.

## 6 Evaluation

It almost goes without saying that an intelligent system that is aimed at assisting cardiologists in reprogramming a pacemaker must be highly reliable. Part of the reliability requirements may be fulfilled by the adoption of sound engineering principles, among others, by using mathematically sound techniques, and by taking conceptually clear models as a starting point of design. We have chosen for formal, logical techniques as a basis for the system. Furthermore, by following a model-based approach to system design, it was relatively easy to ensure the correctness of the collected domain knowledge.

Even so, a study of the performance of a system using real-life data remains indispensable. The results of such a study are discussed next.

## 6.1 Performance measurement

Most measures in use to assess the performance of a diagnostic system derive from early work on the evaluation of probabilistic diagnostic systems [3]. These measures cannot be used here directly, because of differences between the output produced by a probabilistic and qualitative diagnostic system. Probabilistic systems are always capable of producing a diagnosis – except when a probability threshold is used, because then cases may be unclassified, as in the case of qualitative systems – even when no (patient) findings are available. If no findings are available, the system typically generates a prior probability distribution. Hence, common performance measures, as developed for diagnostic probabilistic system, were used only after adaptation.

### 6.1.1 Case-specific performance measures

Let DB denote a database comprising  $N$  cases, where, for each case, one or more diagnostic results are available. Recall that a diagnostic system need not always produce a diagnostic conclusion for each case in a given database DB. For each case  $k$  in the database,  $1 \leq k \leq N$ , we define:

- $c_k$  to be the number of results correctly diagnosed by the system;
- $u_k$  to be the number of results not diagnosed (or unclassified) by the system;
- $i_k$  to be the number of incorrect diagnostic conclusions by the system.

The number of correct diagnostic conclusions for case  $k$  produced by the system is equal to the number of correctly diagnosed results included in the database. Hence, it is not necessary to distinguish between these two numbers. The number of results available in the database DB for each case  $k$  is equal to  $r_k = c_k + u_k$ , and the number of conclusions produced by the system for case  $k$  is equal to  $e_k = c_k + i_k$ .

Based on these numbers, several performance measures can be defined. The measures defined below were inspired by those proposed in [6]. The *accuracy* of a diagnostic system, denoted by  $\alpha_m$ , is defined as the weighed fraction of results that have been interpreted correctly:

$$\alpha_m = \frac{1}{N} \sum_{k=1}^N \frac{c_k}{c_k + u_k} \quad (6)$$

This measure expresses how well the system is capable of correctly diagnosing results associated with cases in the database. The subscript  $m$  indicates that this measure is suitable to interpret *multiple* conclusions associated with cases. The complementary measure,

$$\epsilon_m = 1 - \alpha_m \quad (7)$$

is called the *error rate*. It gives the false impression that all disorders that have not been diagnosed correctly, have been inferred to be absent. However, a system may establish no conclusions at all for certain results mentioned in the database; the error rate produces an overestimation of incorrectly diagnosed results.

A limitation of the accuracy measure is that it does not provide precise information regarding how well a diagnosis produced by the system predicts a disorder to be present or

absent. This information depends on the number of incorrect conclusions produced by the system for a particular case. The *predictive value* of a system, denoted by  $\pi_m$ , conveys such information. It is defined as follows:

$$\pi_m = \frac{1}{N} \sum_{\substack{k=1, \\ c_k + i_k > 0}}^N \frac{c_k}{c_k + i_k} \quad (8)$$

If  $c_k + i_k = 0$ , for case  $k$ ,  $1 \leq k \leq N$ , no diagnosis has been produced by the system. If the cases for which no diagnosis is produced are ignored, a new measure is obtained, called the *adjusted predictive value* of the system, denoted by  $\pi_m^a$ :

$$\pi_m^a = \frac{1}{N^a} \sum_{\substack{k=1, \\ c_k + i_k > 0}}^N \frac{c_k}{c_k + i_k} \quad (9)$$

where  $N^a$  is the number of cases in the database, ignoring unclassified cases.

Formula (6) can be simplified if only a single diagnostic result is available for each case, as follows

$$\alpha_s = \frac{1}{N} \sum_{k=1}^N c_k = \frac{C}{N} \quad (10)$$

where  $C = \sum_{k=1}^N c_k$  is the total number of cases for which a correct diagnostic conclusion has been established by the system. The subscript  $s$  indicates that this measure concerns a single result. This simplification is allowed, because  $c_k + u_k = 1$ ,  $k = 1, \dots, N$ . Similarly, if we assume that only a single diagnostic result will be produced by the system for a given case, it holds that  $(c_k + i_k) \in \{0, 1\}$ ,  $k = 1, \dots, N$ . Then, the adjusted predicted value can be simplified to the following expression:

$$\pi_s^a = \frac{1}{N^a} \sum_{k=1}^N c_k = \frac{C}{N^a} \quad (11)$$

Note that we now have that  $\pi_s = \alpha_s$ , and if all cases have been classified, in addition it holds that  $\pi_s = \pi_s^a$ .

The accuracy and predictive-value measures only provide information concerning correct classification. Similar measures can be defined for incorrect or unclassified cases. Together, these measures offer a full, global description of the diagnostic performance of a diagnostic system.

### 6.1.2 Problem-specific measures

If there are many alternative problems in a diagnostic system, measurement of the performance with respect to individual problems or diagnoses  $D$  yields useful information as well. Let  $\Delta_{DB}$  denote the set of combinations of defects included in the database DB, and let  $\Delta_{DS}$  be the number of diagnoses produced by the diagnostic system. Furthermore, let  $N_{DB,D}$ ,  $D \subseteq \Delta_{DB}$ , be the number of cases concerning the combination of defects  $D$  included in database DB, and let  $N_{DS,D}$ ,  $D \subseteq \Delta_{DS}$ , denote the number of diagnoses equal to  $D$  produced by the system.

The *accuracy with respect to diagnosis  $D$*  of the system is now defined as follows:

$$\alpha_D = \frac{C_{DB,D}}{N_{DB,D}}$$

$D \subseteq \Delta_{DB}$ , where  $C_{DB,D}$  denotes the number of cases with combination of defects  $D$  that have been diagnosed correctly by the system. Similarly, the *predictive value with respect to diagnosis  $D$*  of the system is defined as follows:

$$\pi_D = \frac{C_{DS,D}}{N_{DS,D}}$$

$D \subseteq \Delta_{DS}$ , where  $C_{DS,D}$  denotes the number of correct diagnoses  $D$ . Note that where cases with a known combination of defects  $D$  can be either diagnosed correctly, incorrectly or may be unclassified, a diagnosis  $D$  produced by a system is either correct or incorrect.

It is usually instructive to compute measures for incorrectly diagnosed and unclassified cases, and incorrectly produced diagnoses by the system. These measures, however, are quite similar to those defined above regarding correctness.

The measures defined above can be taken as point estimators of statistical parameters. The uncertainty with respect to their value can be determined by computing their associated confidence intervals. However, when only a few cases are available for validation, confidence intervals will vary widely, and, therefore, convey little useful information. In particular, computation of confidence intervals for the problem-specific measures will not be very informative.

## 6.2 Results

For the evaluation study of the pacemaker reprogramming advisory system, 126 patients were selected retrospectively from the files, having a total of 143 problems. There were 4 patient cases included with some evidence of a pacemaker problem, but for whom the cause had remained unknown. There were also 43 patient cases without any clinical evidence of pacemaker problems. Only 19 of the 126 patients had sensing or pacing problems of the atrium, which concern problems currently covered by the system. The remaining patients experienced problems unrelated to atrial sensing or pacing, such as problems with the rate-response system or problems with ventricular sensing or pacing. Data of these patients, including those with unknown or absent causes, were used to assess the risk of obtaining incorrect diagnoses by the system, for problems outside the causal model's scope. The various problems covered in the database are summarised in Table 1.

Only one of the 19 patients appeared to have symptoms caused by a combination of problems (far field R-wave sensing combined with a too low P-wave amplitude). Results in which the conclusions of an expert were checked against the conclusions of the system are shown in Table 2. Note that the only incorrect advice concerned the single patient who had a combined problem. Actually, in this particular case the system diagnosed that there was a far field R-wave sensing problem, but the system did not reach the conclusion that the P-wave amplitude was too low in this patient. The overall accuracy of the system is equal to  $\alpha_m = 0.82$ , with  $\alpha_s = 0.79$ .

The results obtained by comparing diagnoses produced by the system with the expert's conclusions, are summarised in Table 3. The single case where only one of two problems was diagnosed has been classified as being incorrect. The overall predictive value of the system is equal to  $\pi_m = 0.79$ , with  $\pi_m^a = 0.94$ .

<b>Problem type</b>	<i>n</i>	%
No problem	43	(30)
Ventricular	4	(3)
Rate-response system	16	(11)
Atrial problem	19	(13)
Atrial threshold	11	(8)
Ventricular threshold	27	(19)
AV synchrony	4	(3)
Mode problem	15	(10)
Unknown	4	(3)
<b>Total</b>	<b>143</b>	<b>(100)</b>

Table 1: Frequency distribution of pacemaker problems.

<b>Expert conclusion</b>	<b>Total</b>	<b>Correct</b>	<b>Incorrect</b>	<b>Unclassified</b>
	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
Far field R wave sensing, and P amplitude low	1	0	1	0
Far field R wave sensing	10	8	0	2
P-wave amplitude low	5	5	0	0
Retrograde conduction	3	2	0	1
<b>Total</b>	<b>19</b>	<b>15</b>	<b>1</b>	<b>3</b>

Table 2: Accuracy of the system.

<b>System conclusion</b>	<b>Total</b>	<b>Correct</b>	<b>Incorrect</b>
	<i>n</i>	<i>n</i>	<i>n</i>
Far field R-wave sensing	9	8	1
P-wave amplitude low	5	5	0
Retrograde conduction	2	2	0
<b>Total</b>	<b>16</b>	<b>15</b>	<b>1</b>

Table 3: Predictive value of the system.

<b>Total</b>	<b>Correctly Unclassified</b>	<b>Suspected diagnosis</b>	<b>Confirmed diagnosis</b>
<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
107	78	21	8

Table 4: Results for normal cases or problems not covered by the system.

Finally, in only 8 of the remaining 107 patient cases (7.5%) without atrial pacing or sensing problems, the diagnosis produced was incorrect, as shown in Table 4. Of course, these diagnoses were still sensible, because the clinical evidence also pointed in the direction of atrial pacing and sensing, although it appeared that the actual problem was different from the diagnosis. These results offer information of the risk of diagnosing problems that in reality are absent in the patient.

## 7 Discussion

In this paper, we have described a system that is capable of assisting a cardiologist in reprogramming an implanted pacemaker. In designing the system we have used formal, logical techniques. As a consequence, it is feasible, even straightforward, to precisely describe the underlying representation and reasoning methods, and to prove properties of the system. A further contribution in the pursuit of achieving correctness of the system derives from the use of model-based techniques in the design of the system’s knowledge base. For the problem of reprogramming a pacemaker, it appeared to be sufficient to use causal models of abnormal behaviour as a basis for domain knowledge. Our experience is that models of normal behaviour are insufficient for advising appropriate pacemaker settings, as well as for diagnosing medical and pacemaker problems.

At present, the system only covers part of the entire domain. In the near future, the system will be extended by including other parts of the problem domain as well. Furthermore, the current implementation will be replaced by a program that can be integrated with the pacemaker programmer software, so that the user will view the system as a functional addition to the programmer software. Integration also has the advantage that pacemaker data need not be entered by hand, but can be extracted automatically from the pacemaker device.

## References

- [1] L. Console, D. Theseider Dupré and P. Torasso, A theory of diagnosis for incomplete causal models, in: *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, Los Angeles, USA (1989) 1311–1317.
- [2] L. Console and P. Torasso, A spectrum of logical definitions of model-based diagnosis, *Computational Intelligence* 7(3) (1991) 133–141.
- [3] J.D.F. Habbema, J. Hilden and B. Bjerregaard, The measurement of performance in probabilistic diagnosis I: the problem, descriptive tools, and measures based on classification matrices, *Methods of Information in Medicine* 17 (1978) 217–226.

- [4] D.E. Heckerman, E.J. Horvitz and B.N. Nathwani, Toward normative expert systems: Part I, The Pathfinder project, *Methods of Information in Medicine* 31(2) 90–105.
- [5] K.J. Isselbacher, et al., *Harrison's Principles of Internal Medicine*, 13th edition, McGraw-Hill, New-York, 1994.
- [6] N. Indurkha and S.M. Weiss, Models for measuring performance of medical expert systems, *Artificial Intelligence in Medicine* 1 (1989) 61–70.
- [7] P.J.F. Lucas, Logic engineering in medicine, *The Knowledge Engineering Review* 10(2) (1995) 153–179.
- [8] P.J.F. Lucas, Symbolic diagnosis and its formalisation, *The Knowledge Engineering Review* 12(2) (1997) 109–146.
- [9] D. Poole, R. Goebel and R. Aleliunas, Theorist: a logical reasoning system for defaults and diagnosis, in: N. Cercone and G. Mc Calla, eds., *The Knowledge Frontier* (Springer-Verlag, Berlin, 1987).
- [10] D. Poole, Representing diagnosis knowledge, *Annals of Mathematics and Artificial Intelligence* 11 (1994) 33–50.
- [11] D. Poole, *Local Users Guide to Theorist*, Report, Department of Computer Science, University of British Columbia, Vancouver, 1990.
- [12] R. Reiter, On closed world data bases, in: H. Gallaire and J. Minker, eds., *Logic and Databases* (Plenum-Press, New York, 1978) 55–76.
- [13] R. Reiter, A logic for default reasoning, *Artificial Intelligence* 13 (1980) 81–132.