20

Automatic Synthesis and Compression of Cardiological Knowledge

I. Bratko

E. Kardelj University and J. Stefan Institute

I. Mozetič and N. Lavrač

J. Stefan Institute,
Ljubjana, Yugoslavia

Abstract

The paper reports on a study into the mechanical synthesis of the operational knowledge needed for the expert task of electrocardiographic (ECG) interpretation. This knowledge-base was synthesized by means of qualitative simulation based on a causal model of the heart. The resulting (ECG) knowledge-base was subsequently compressed by using inductive learning tools.

1. INTRODUCTION

This paper reports on a study into mechanical synthesis of the operational knowledge needed to perform an expert task. The particular task in question is the interpretation of the electrical signals generated by the heart muscle, known as the electrocardiographic (ECG) interpretation. The main contribution of this research is a qualitative model of the electrical activity of the heart which was the basis for mechanical derivation of the ECG diagnostic knowledge.

The heart can be viewed as a mechanical device with an electrical control system. This electrical system works completely autonomously within the heart and is responsible for generating the rhythmical stimulation impulses that cause the contraction of the heart muscle. For proper functioning of the heart, the stimuli have to reach the atria (upper part of the heart) somewhat earlier than the ventricles (lower part of the heart). This is coordinated by the electrical control system which is shown schematically in Figure 1. The contractions of the heart muscle cause changes in the electrical potentials in the body. The changes of these potentials in time can be recorded as an electrocardiograph. Disturbances in the functioning of the heart are reflected in the ECG curves. The interpretation of ECG signals is concerned with the question: if a given ECG curve is not normal, what are the disorders in the heart which could have caused this abnormality?

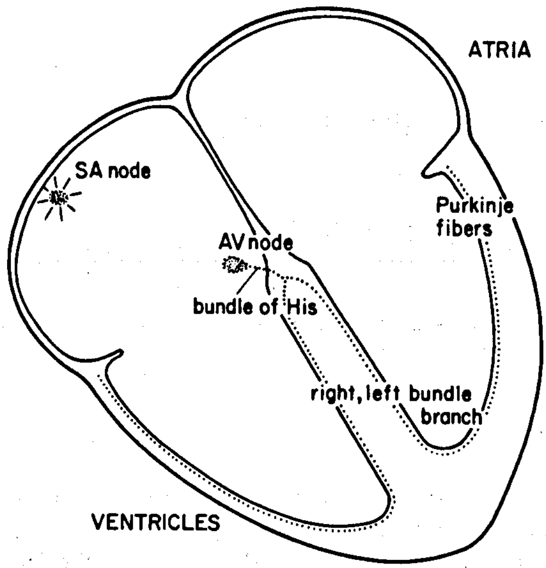

Figure 1. A scheme of the electrical control system of the heart. The nodes generate electrical impulses. The dotted lines represent conduction pathways for impulses.

Various disorders can occur in the electrical control system of the heart. For example, an impulse generator may become silent, or an extra generator may appear, or some electrical conductance may become blocked, etc. These disorders are called *cardiac arrhythmias*. There are about 30 basic disorders and each of them causes some characteristic changes in the ECG. There can be several disorders simultaneously present in the heart. Combined disorders are called *multiple arrhythmias* as opposed to *simple arrhythmias* which correspond to single, basic disorders. The combinatorial nature of arrhythmias complicates the ECG interpretation problem because of the large number of potentially possible combinations. In the medical literature on the cardiac arrhythmias (e.g. Goldman, 1976) there is no systematic description of ECG features which correspond to pairs of simple arrhythmias, let alone triple and even more complicated arrhythmias. On the other hand, these are not very rare in medical practice. In addition to this, multiple arrhythmias are hard to diagnose because there is no simple rule for combining ECGs that correspond to constituent disorders. In other words, if we know which ECGs correspond to any simple disorder, in general it is not clear how to 'sum' these ECGs into 'combined' ECGs corresponding to combinations of simple disorders.

We approached the problem of multiple arrhythmias by constructing a model of the heart. Any combination of disorders can be inserted into the model. The model is *deep* in the context of the distinction between *deep*, causal knowledge, and *shallow*, operational knowledge. By definition, the shallow-level knowledge is sufficient for performing the task itself, but typically without any understanding of the underlying causal mechanisms. The deep knowledge, on the other hand, captures this causal underlying structure and allows the system to reason from first principles.

Our model also is *qualitative* in the sense that it does not deal with electrical signals represented numerically as voltages in time, but represented by symbolic descriptions that specify qualitative features of signals. Such a qualitative modelling approach has several advantages over the conventional numerical modelling. Among the advantages are:

1. The qualitative view is closer to the actual physiological descriptions of and reasoning about the processes and failures in the heart.
2. To execute the model we do not have to know exact numerical values of the parameters in the model.
3. The qualitative simulation is computationally less complex than numerical simulation.
4. The qualitative simulation can be used as a basis for constructing explanations of the mechanism of arrhythmias.

In respect of the qualitative approach to modelling our work is related to the work of Forbus (1984), de Kleer and Brown (1984), and Kuipers (1984).

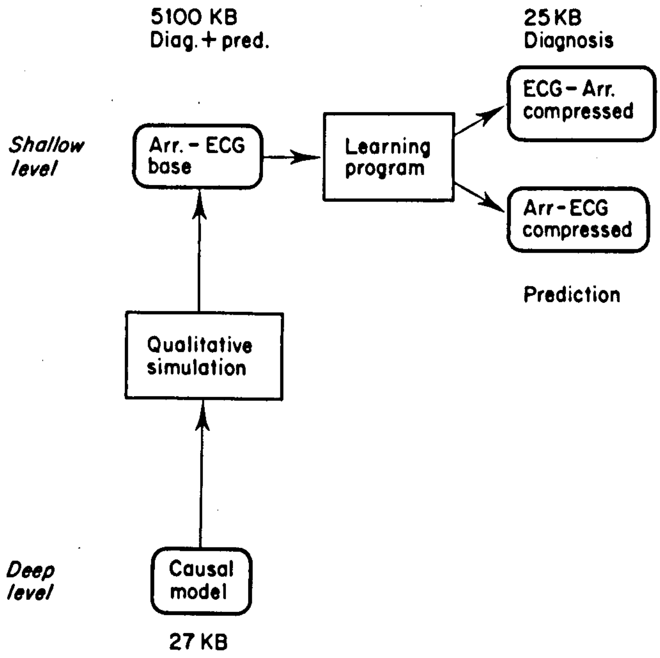

Figure 2. Deep and shallow levels of cardiological knowledge and transformations between these representations.

We used the model for the automatic synthesis (through simulation) of the shallow, operational representation of the ECG interpretation knowledge (see Figure 2). This representation facilitates fast ECG diagnosis, but is rather complex in terms of memory space (about 5 Mbytes). Therefore, as Figure 2 shows, we compressed this knowledge-base by means of inductive learning programs. The representation thus obtained is compact and diagnostically efficient.

In the remainder of the paper we describe the model of the heart, the qualitative simulation algorithm and its efficient implementation, the synthesized shallow knowledge-base and its subsequent compression.

2. THE QUALITATIVE MODEL OF THE HEART

Our qualitative model of the electrical activity of the heart specifies causal relationships between objects and events in the heart. These include electrical impulses, ECG signals, impulse generation, impulse conduction and summation. The model can be thought of as an electrical network, as shown in Figure 3. However, signals that propagate in this

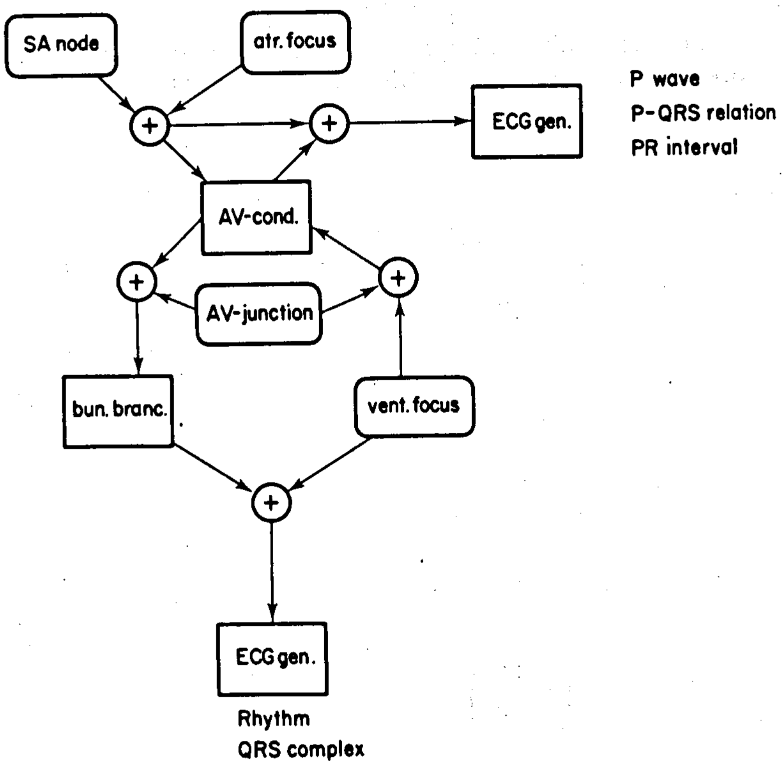

Figure 3. The model of the heart as a network composed of impulse generators, conduction pathways, impulse summators, and ECG generators.

network are represented qualitatively by symbolic descriptions rather than by voltage v. time relations.

The ingredients of the model are: nodes of the network; a dictionary of simple arrhythmias related to heart disorders; 'legality' constraints over the states of the heart; 'local' rule sets; 35 'global' rules.

These ingredients are reviewed in more detail below.

Nodes of the network

There are four types of nodes: impulse generators, conduction pathways, impulse summators, and ECG generators, illustrated in Figure 4. Recall that the word 'impulse' in this figure refers to a symbolic description, so these elements are in fact operators on descriptions. Impulse generators and conduction pathways can be in normal or abnormal functional states. For example, a generator can generate impulses or can be silent; a conduction pathway can conduct normally or it can be blocked or partially blocked in various ways: it may just cause a delay of an impulse, or it can suppress every second or third impulse, etc. These abnormal states of individual elements correspond to simple disorders of the heart.

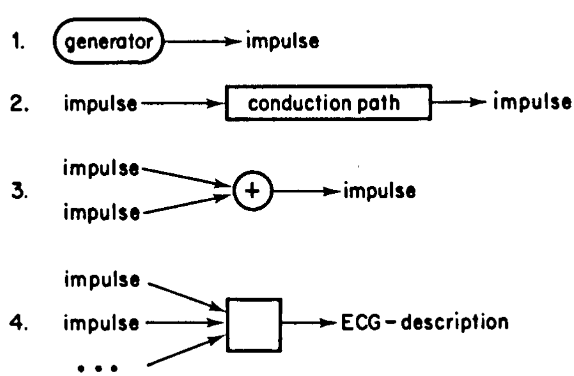

Figure 4. Building blocks for the heart model.

A dictionary of simple arrhythmias related to heart disorders

Each simple arrhythmia is defined in terms of the functional states of the components of the heart. Roughly speaking, each simple arrhythmia corresponds to a disorder in one of the heart's components.

'Legality' constraints over the states of the heart

This is a predicate on the functional states of the heart which recognizes certain categories of states that are rejected by the model as 'illegal'. These categories include: logically impossible states, physiologically impossible states, and 'medically uninteresting' states.

A state is logically impossible if one of the heart's components is in two different states at the same time. An example of a physiologically impossible state is a situation in which two generators in the atria discharge permanent impulses. An example of a 'medically uninteresting' state is one in which there is no atrial activity and the atrio-ventricular(av) conduction is blocked. In such a case the block has no effect on the function of the heart and also cannot be detected in the ECG.

'Local' rule sets

These specify the behaviour of the individual components of the heart (generators, summators and conduction pathways) in the presence of various abnormal states.

'Global' rules

These rules define causal relations between impulse generators and conduction pathways in the heart, electrical impulses and ECG features; these rules also reflect the structure of the network in Figure 3. There are 35 global rules in the model.

All the rules in the model have the syntax of the first-order predicate calculus, in particular, the syntax that is accepted by PROLOG under Edinburgh notational conventions (Pereira *et al.*, 1978). According to these conventions, the names of constant symbols and functors start with lower-case letters, and the names of variables start with capital letters.

Rules are composed of subexpressions in specialized languages for describing the state of the heart, impulses that are conducted through the heart, and ECG patterns.

For example, the term

heart(atr_focus: permanent(regular, between_100_250))

is a partial specification of the state of the heart. It says that the atrial focus is discharging permanent impulses (as opposed to periodical) with a regular rhythm at the tachycardic rate (i.e. somewhere between 100 and 250). Each statement about the state of the heart tells in what functional state a component of the heart is (the atrial focus in the example above).

The following is an example of an ECG description:

[rhythm = irreglar] &
[regular_P = abnormal] &
[rate_of_P = between_100_250] &
[relation_P_QRS = after_P_some_QRS_miss] &
[regular_PR = prolonged] &
[regular_QRS = normal] &
[rate_of_QRS = between_60_100 or between_100_250]

This specification consists of values assigned to qualitative ECG attributes that are normally used in the cardiological literature, such as the rhythm and the shape and the rate of P-waves (Figure 5). Notice that the

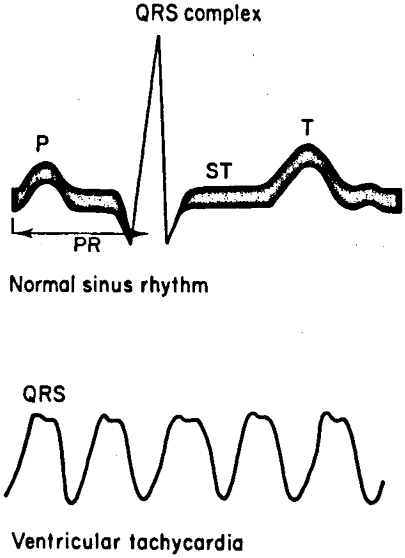

Figure 5. Upper part an ECG curve that corresponds to the normal heart. Marked are features that are normally looked at by an ECG diagnostician. Lower part: ECG curve that corresponds to the arrhythmia ventricular tachycardia. This abnormal ECG is characterized by its higher rate ('tachycardia', between 100 and 250 beats per minute), and the 'wide' shape of the QRS-complexes.

description above gives two values for the rate of QRS waves: it can be either normal (between_60_100) or tachycardia (between_100_250).

Impulses are described by expressions of the form illustrated by the following example:

impulse(atr_focus: form(unifocal, regular, between_100_250))

This says that there are unifocal regular impulses with the tachycardic rate at the atrial focus.

Figure 6 shows two examples of global rules and some rules that specify the behaviour of the individual components of the heart. The first global rule in Figure 6 says:

IF
  the atrial focus discharges permanent impulses at some rhythm Rhythm and rate Rate
THEN
  there will be impulses at the atrial focus characterized by Origin, Rhythm and Rate
WHERE
  Origin, Rhythm and Rate must satisfy the atr_focus relation.

The atr_focus relation describes the behaviour of the atrial focus. This

```
% Two global rules

[heart(atr_focus: permanent(Rhythm, Rate))] ⇒
[impulse(atr_focus: form(Origin, Rhythm, Rate))] &
atr_focus(Origin, Rhythm, Rate).

[impulse(atria: form(_, Rhythm0, Rate0)), heart(av_conduct: State)] ⇒
[impulse(av_conduct: form(State, Rhythm1, Rate1))] &
av_conduct(State, Rhythm0, Rhythm1, Rate0, Rate1).

% Some local relations

atr_focus(unifocal, quiet, zero).
atr_focus(unifocal, regular, between_60_100).
atr_focus(unifocal, regular, between_100_250).
atr_focus(wandering, irregular, between_60_100).
atr_focus(wandering, irregular, between_100_250).
atr_focus(circulating, regular, between_250_350).
  ...

av_conduct(normal, Rhythm, Rhythm, Rate, Rate) :-
  below(Rate, over_350).
av_conduct(progress_delayed, regular, irregular, Rate0, Rate1) :-
  reduced(Rate0, Rate1).

av_conduct(progress_delayed, irregular, irregular, Rate0, Rate1) :-
  reduced(Rate0, Rate1).
  ...

reduced(Rate, Rate).

reduced(Rate, Rate1) :-
  succ(Rate1, Rate).

succ(zero, under_60).
succ(under_60, between_60_100).
succ(between_60_100, between_100_250).
  ...
```

Figure 6. Two global rules and some local rules of the heart model. Global rules are, from the point of view of PROLOG, unit clauses of the form: $A \Rightarrow B$ & C, which can in the model be read: if A then B where C. A special rule-interpreter in PROLOG uses rules of this type. Local rules specify the behaviour of individual components of the heart and are directly executed by the PROLOG system as rules of a PROLOG program.

relation is partially specified in Figure 6 by 'local rules'. It tells that the atrial focus can be quiet, it can behave 'unifocally' discharging impulses at a normal or tachycardic rate with regular rhythm, or it can be 'wandering' discharging impulses with irregular rhythm at normal or tachycardic rate, etc.

The second global rule in Figure 6 can be read:

IF

in the atria there are permanent impulses of some rhythm Rhythm0 and rate Rate0, and the state of the av-conductance is State

THEN

there are impulses of type State, rhythm Rhythm1 and rate Rate1 at the exit from the av-conductance

WHERE

State, Rhythm0, Rhythm1, Rate0 and Rate1 have to satisfy the relation 'av_conduct'.

The av_conduct relation is specified by a set of local rules, as a directly executable PROLOG procedure. This procedure qualitatively defines the physiology of the av_conductance pathway. As can be seen from the definition of this relation in Figure 6, the components of the heart often behave 'non-deterministically' in the sense that they can react to the same input with different responses at the output.

Complete details of the model can be found in Mozetic *et al.* (1984).

3. THE QUALITATIVE SIMULATION ALGORITHM

Formally, the qualitative simulation consists of theorem proving and theorem generation. Although the 35 global rules have the syntax of PROLOG they are not directly executed by PROLOG's own interpreting mechanism, the main reason being the necessity for additional control in order to improve the execution efficiency. Thus the qualitative simulation is done by a special rule-interpreter implemented in PROLOG.

Each simulation run consists of the following steps:

1. Instantiate the model by a given arrhythmia, using the definitions of arrhythmias in terms of the heart disorders.
2. Check the resulting functional state of the heart against the legality constraints (logical, physiological, etc.)
3. Execute the model by triggering the rules until no more rules fire (this process is combinatorial due to the non-deterministic nature of the heart's components).
4. Collect the proved assertions about ECG signals and then construct an ECG description that corresponds to the given arrhythmia.

The complex part above is step 3. It is based on the forward chaining of global rules in the model. The simulator starts with some initial data base of facts (initially these just specify the state of the heart) and keeps firing the global rules until no more can fire. The constraint here is that no rule is repeatedly executed on the same piece of information. Execution of rules generates new assertions that are added into the data base. These new assertions are regarded as hypotheses that can later be proved false.

1. By the definition of [atrial_tachycardia, wenckebach] instantiate the state of the heart to:

   heart(sa_node: permanent(quiet, zero)) &
   heart(atrial_focus: permanent(regular, between_100_250)) &
   heart(av_conduct: progress_delayed) &
   ...

2. The assertion

   heart(atrial_focus: permanent(regular, between_100_250))

   triggers the first global rule in Figure 6. The goal

   atr_focus(Origin, regular, between_100_250)

   is evaluated, using the local relation atr_focus. This succeeds and the new assertion is added to the database:

   impulse(atr_focus: form(unifocal, regular, between_100_250))

3. After summing together the atrial impulses we get the asertion:

   impulse(atria: form(unifocal, regular, between_100_250))

4. This assertion triggers the second global rule in Figure 6. The Prolog goal to be evaluated is now:

   av_conduct(progress_delayed, regular, Rhythm1, between_100_250, Rate1)

   This can be satisfied in two ways, either by Rate1 = between_100_250 or by Rate1 = between_60_100. So two hypotheses are indicated:

   impulse(av_conduct: form(progress_delayed, irregular, between_100_250))

   or

   impulse(av_conduct: form(progress_delayed, irregular, between_60_100))

   Depending on the search strategy used, the system may now assert one of the above hypotheses into the database, and consider the second one on backtracking which corresponds to a depth-first style search; or, it may assert both which corresponds to a breadth-first style search.

Figure 7. Fragments of the qualitative simulation trace for the combination of arrhythmias atrial_tachycardia and wenckebach. Some steps were omitted from the actual trace.

Backtracking to a previous point occurs when the current content of the data base is found inconsistent, i.e. some assertion has been generated which leads to contradiction. Roughly, the rule triggering process is as follows.

Assume that there is a hypothesis A in the data base. Then apply a global rule of the form

$$A \Rightarrow B \ \& \ C$$

In general, in such rules A and B are PROLOG terms and C is a PROLOG goal which can directly be executed by PROLOG. The precondition matching is simply the logic unification. Normally, C is a call to evaluate a local relation. Thus, to apply a rule of the above form, do:

Evaluate C; if C is false then A must be false and discard it; otherwise if C is true, assert a new hypothesis B and continue firing rules.

In the case that C is false a contradiction has been detected and backtracking is indicated. This process terminates when there are no more rules to fire. At that stage, all the remaining hypotheses in the data base are accepted as true since there is now no way of showing a contradiction. Among these facts there are also statements about the ECG. The simulator collects those statements and forms an ECG description which corresponds to the arrhythmia with which the simulation process was started.

As an example, Figure 7 shows part of a simulation run when the state of the heart is a combination of two simple arrhythmias: *atrial_tachycardia* and *wenckebach.*

4. IMPLEMENTATION OF THE SIMULATION ALGORITHM

The easiest way of implementing the simulation algorithm outlined above is to use the depth-first search strategy. This is straightforward and suitable for single simulation runs, that is for answering questions of the *prediction* type: given an arrhythmia, what are its corresponding ECGs. Different possible ECGs are simply generated through backtracking. Also, an execution trace obtained in such a simulation run can be used as the basis for generating a user-oriented explanation of what is going on in the heart. This is suitable since the simulation steps follow the causal chains of events in the heart, according to the global rules of the model. These rules essentially describe the causal relations between events in the heart.

In a PROLOG implementation of depth-first simulation on the DEC-10 (Edinburgh implementation of PROLOG; Pereira 1978), each simulation run takes a few c.p.u. seconds, producing all alternative ECGs.

Diagnostic-type questions as opposed to *prediction-type* questions, are of the form: given an ECG, what arrhythmias could have caused it? To answer such questions, we could run the model in the opposite direction. Start with a given ECG and end with the possible functional states of the heart that might cause this ECG.

We can in fact run the model in this direction by the backward chaining of the rules in the model. In order to do that we reversed the global rules and used the simple depth-first search. However, the practical problem of efficiency now arises because the branching factor ('non-determinism') in the backward direction is much higher than that in the forward direction. This entails much more backtracking and rather complex search, thus rendering this approach to diagnosis impractical. Efficiency can be improved by re-writing the model so as to introduce more constraints into the rules, which helps the system recognize contradictory branches at an earlier stage. An attempt at re-formulating rules, however, revealed two drawbacks. The size of the model increases considerably, and the transparency is greatly affected. This, in turn, mars the explanation of the heart's behaviour based on the execution trace.

An alternative way to achieve efficient diagnosis is to generate from the deep model of the heart a complete shallow-level representation of the arrhythmia–ECG relation as a set of pairs of the form:

(Arrhythmia, ECG-description)

In principle this can be done by executing the depth-first simulation (forward chaining) for each possible combined arrhythmia, and storing all its ECG manifestations. It would be necessary to repeat this for all possible alternative execution paths in order to obtain all possible ECGs for each arrhythmia. This is again rather inefficient for two reasons. First, for each disjunctive solution the simulator has to backtrack to some previously used rule in the model and restore its previous state. Second, the final resulting ECG descriptions have the form of disjunctions of ECG patterns. These disjunctive expressions can be more complex than necessary and can be later simplified. This posterior simplification, however, is again a complex operation. Each disjunct is the result of an alternative execution path. The simplification can be carried out much more economically at the very moment that a disjunct (or, typically part of it) is generated, before it is further expanded and mixed in the expression with other not closely related terms.

These two factors (saving the restoration of previous states, and immediate simplification of disjunctive expressions) motivated the implementation of another type of simulation algorithm which handles alternative execution paths in a breadth-first fashion. This algorithm develops alternatives essentially in parallel and currently simplifies disjunctions. The simplification rules actually used are rather model dependent in the sense that they do not preserve logical equivalence in general but only in the special case of the properties of the heart model.

So the 'breadth-first' simulation is not general and we would possibly have to modify the simplification rules in the case of a change in the model.

This specialized simplification method proved to be rather powerful. As a typical example of the reduction effect, consider the combined arrhythmia *atrial_fibrillation and ventricular_ectopic_beats.* The depth-first simulation generates 72 ECG descriptions which corresponds to an ECG expression with 72 disjunctive terms. The breadth-first simulation results in a description comprising four disjunctive terms. There was a similar factor of improvement in general, which can be seen from the results of generating the complete arrhythmia–ECG relation.

5. GENERATION OF A COMPLETE ARRHYTHMIA–ECG KNOWLEDGE-BASE

The breadth-first simulation algorithm was executed on all mathematically possible combinations of simple arrhythmias. The majority of these combined arrhythmias were eliminated by the legality constraints over the states of the heart. The complete arrhythmia-ECG knowledge-base was thus automatically generated. Results of the generation are depicted in Figure 8.

Figure 8 reveals some interesting points. Of all possible arrhythmias, the combinations of four simple arrhythmias are the largest subset. Note the large number (140,966) of generated ECG patterns. This is indicative of the difficulty in ECG diagnosis of cardiac arrhythmias. On average, each arrhythmia has almost 60 different corresponding ECG manifestations. There are altogether 2419 'legal' combined arrhythmia within the level of

Number of disorders in the heart	Mathematically possible combinations	Numbers of generated		
		Multiple arrhythmias	Prolog clauses	ECG descriptions
1	30	18	27	63
2	435	118	286	2,872
3	4,060	407	1,207	17,551
4	27,405	759	2,679	45,939
5	142,506	717	2,867	52,707
6	593,775	340	1,164	20,322
7	2,035,800	60	84	1,512
Σ	2,804,011	2,419	8,314	140,966

Figure 8. Results of generating the arrhythmia knowledge-base. The number of generated arrhythmias for 'combinations' of simple arrhythmias is 18 which is less than the number of all simple arrhythmias (30). The reason is that some simple arrhythmias (conduction disturbances and ectopic beats) cannot occur alone, but only in combination with other arrhythmias (e.g. with sinus rhythm).

detail of the heart model. The relation to their ECG manifestations is represented by 8314 PROLOG clauses. Each clause represents a pair: arrhythmia–ECG expression. Each ECG expression specifies a number of possible ECGS, about 20 on average. This is the reduction factor due to the simplification technique used in the breadth-first simulation.

The set of 140,966 ECG patterns (the right-hand sides of the arrhythmia–ECG rules) are not unique. The same ECG patterns can occur at several places which means that several arrhythmias can have the same ECG manifestation. Consequently, arrhythmias cannot be unambigously diagnosed from a given ECG. Empirical probing showed that for a given ECG there are typically between two and four possible combined arrhythmias in the arrhythmia knowledge-base. From the medical point of view, however, these alternative diagnoses are not significantly different in the view of treatment. They would typically all require the same treatment.

The arrhythmia–ECG base generated from the model is complete in two ways. First, it comprises all physiologically possible arrhythmias at the level of detail of the model. Second, each arrhythmia is associated with all its possible ECG manifestations. In principle, the problem of diagnosing is now simple. As the rules in the knowledge-base are logical implications, we can apply *modus tollens* rule of inference on them. Consider a rule of the form

Arrhythmia ⇒ ECG_description

where ECG_description is the disjunction of *all* possible ECGS that Arrhythmia can cause. Then, if a given ECG does not match ECG_description it follows that Arrhythmia is eliminated as a diagnostic possibility. All arrhythmias that are not thus eliminated form the set of possible diagnoses with respect to the given ECG data. Any further discrimination between the set of arrhythmias thus obtained can be done only on the basis of some additional evidence (e.g. clinical data). Also, as the knowledge-base is complete the empty set of possible arrhythmias would imply that the given ECG is physiologically impossible.

6. COMPRESSION OF THE ARRHYTHMIA–ECG BASE USING INDUCTIVE LEARNING TOOLS

The main motivation for having the arrhythmia–ECG base is that it can be used for ECG diagnosis based on a simple pattern-matching rule. However, it is rather bulky for some practical application requirement. If stored as a text file, the 8314 PROLOG clauses that represent the arrhythmia–ECG relation, occupy 5.1 Mbytes of store. Also its complexity renders this knowledge-base difficult to compare with the conventional medical codifications of electrocardiographic knowledge. Therefore an attempt has been made to find a more compact representation of the arrhythmia–ECG base that would still allow efficient ECG diagnosis.

The main idea was to use the knowledge-base as a source of examples of particular features (heart disorders or ECG features) and to use an inductive learning algorithm to obtain their compact descriptions. The inductive learning programs used were GEM (Reinke, 1984) and EXCEL (Becker, 1985). Taking the complete arrhythmia–ECG base as the set of examples, the number of examples for these two learning programs would be too high. Therefore we had first to generate a subset of the knowledge-base that would retain its completeness to the greatest possible extent. The following domain-specific factorization properties facilitated the selection of a considerably reduced subset for learning, whereby the information thus lost can be recovered by a small set of additional rules.

Some disorders in the heart are of a permanent nature while some do not occur regularly; the latter are called ectopic beats. A large number of combined arrhythmias and in particular ECG descriptions are due to the unconstrained combinatorial nature of ectopic beats. If we disregard mutual combinations of different types of ectoptic beats we can substantially reduce the number of generated multiple arrhythmias and ECG descriptions. The information thus lost can easily be reconstructed from the remaining rules in the knowledge-base. Namely, different types of ectopic beats are both mutually independent and independent of permanent disorders. The presence or absence of an ectopic beat does not affect the part of the ECG description produced by other disorders. The learning subset of the knowledge-base was further reduced by disregarding three simple arrhythmias whose ECGS can easily be deduced from the behaviour of other similar arrhythmias.

To summarize—the learning subset was constructed from the original arrhythmia-ECG base by the following reductions:

1. The subset only deals with 27 simple arrhythmias instead of the original repertoire of 30. We omitted *sinus_arrhythmia, right_bundle_branch_block,* and *multi_ventricular_ectopic_beats* whose ECG description can be constructed from the descriptions of sinus_node_disorders, left_bundle_branch_block, and ventricular_ectopic_beats respectively.

2. We discarded mutual combinations of different types of ectopic beats (atrial_ectopic_beats, junctional_ectopic–beats, and ventricular_ectopic_beats).

The subset thus obtained was substantially smaller than the original arrhythmia–ECG base. There are 586 combined arrhythmias and 2405 ECGS in the subset compared with 2419 arrhythmias and 140,966 ECGS in the complete knowledge-base.

Roughly, the procedure for compressing the knowledge now proceeded as follows. The learning subset of the arrhythmia–ECG base comprised

586 rules (corresponding to 586 combined arrhythmias) of the form:

Combined_arrhythmia ⇒ ECG_description.

The goal of learning was to convert this information into rules of two forms:

1. Compressed *prediction rules* which answer the question: what ECGS may be caused by a given disorder in a heart's component?
2. Compressed *diagnostic rules* which answer the question: what heart disorders are indicated by a given isolated ECG feature?

Compressed prediction rules were synthesized by the GEM inductive learning program, and compressed diagnostic rules were synthesized by the EXCEL program. Both programs are based on the AQ11 learning algorithm (Michalski, 1983) and both generate class descriptions as APC expressions (Annotated Predicate Calculus; Michalski, 1983). Before the programs could be used, the learning subset had to be converted into rules of yet two other forms in order to obtain the right input for the learning programs required, i.e. examples of objects that belong to classes being learned. The principle of how to do that in general is described in Mozetic (1986). For synthesizing prediction rules, the proper starting form was:

Isolated_disorder ⇒ ECG_description

where an 'isolated disorder' is, for example, the atrial focus being in the tachycardic state. The starting point for the synthesis of diagnostic rules were rules of the form:

Isolated_ECG_feature ⇒ Heart_state_description

where an 'isolated ECG feature' is, for example, P-wave having abnormal shape. In general, rules of these forms are not completely logically equivalent to the original rules, so they have to be used with care. Mozetic (1986) states the conditions under which both the original rules

Total number of	Original arrhythmia knowledge-base	Subset of the knowledge-base	Arrhythmias combined	Diagnostic rules
rules	2,419	586	45	49
conjunctions	8,314	957	75	144
attributes	58,197	6,699	248	371
Kbytes	5,100	400	10	13

Figure 9. Comparison between the original arrhythmia–ECG base, the selected subset for learning, and the derived compressed rules of both types (prediction and diagnosis). A 'rule' above corresponds to a combined arrhythmia, a 'conjunction' corresponds to a PROLOG clause. 'Attributes' mean all the references to attributes in a whole rule set. The last row gives the sizes of these representations if stored as text files.

and the transformed ones are in fact logically equivalent. So this additional request had to be verified in our case as well.

Figure 9 shows the compression effects achieved in terms of the number and complexity of rules, and in terms of storage space needed when storing different representations simply as text files.

Figure 10 shows some prediction and some diagnostic rules generated

```
combined(wenckebach) ⇔ [av_conduct = wen] ⇒
  [relation_P_QRS = after_P_some_QRS_miss] &
  [regular_PR = prolonged]

combined(atrial_tachycardia) ⇔ [atr_focus = at] ⇒
  [regular_P = abnormal] &
  [rate_of_P = between_100_250] &
  [regular_PR = meaningless ∨ normal ∨ prolonged]
  ∨
  [regular_P = abnormal] &
  [regular_PR = shortened ∨ normal ∨ prolonged] &
  [regular_QRS = wide_LBBB_RBBB ∨ delta_LBBB ∨ delta_RBBB]
  ∨
  [regular_P = abnormal] &
  [rate_of_P = between_100_250] &
  [regular_PR = shortened] &
  [regular_QRS = normal ∨ wide_LBBB ∨ wide_RBBB]

[regular_P = abnormal] ⇒
  [sa_node = quiet] &
  [atr_focus = quiet ∨ at ∨ afl ∨ af ∨ aeb] &
  [reg_vent_focus = quiet ∨ vr ∨ avr ∨ vt]

[regular_PR = shortened] ⇒
  [atr_focus = quiet ∨ wp ∨ at ∨ mat ∨ aeb] &
  [av_conduct = wpw ∨ lgl]
  ∨
  [av_conduct = normal] &
  [av_junction = jb ∨ jr ∨ jt]
  ∨
  [atr_focus = at] &
  [av_conduct = normal ∨ wpw ∨ lgl]

[regular_QRS = normal] ⇒
  [av_conduct = normal ∨ avb1 ∨ wen ∨ mob2 ∨ avb3 ∨ lgl] &
  [bundle_branches = normal] &
  [reg_vent_focus = quiet]
```

Figure 10. Examples of prediction and diagnostic rules generated by the inductive learning programs.

by the induction algorithms. Some of these descriptions correspond very well to the definitions in the medical literature. For example, the descriptions of the *wenckebach* disorder corresponds precisely to the conventional medical descriptions. On the other hand, some of the synthesized descriptions were considerably more complex than those in the medical literature. The computer-generated descriptions in such cases give much more detailed specification than may be necessary for an intelligent reader with a physiological background. Such a reader can usually infer the missing detail from his background knowledge. The additional details must still be made explicit in case of a computer application in the form of a diagnostic expert system, otherwise a lot of background knowledge and inference would have to be added which would be extremely difficult and its correctness hard to verify. Mozetič (1986) describes in detail how the knowledge compression was done.

7. CONCLUSIONS

Various representations of the ECG knowledge and transformations between these representations were described. The main three representations are at different knowledge-levels in the sense of a distinction between 'deep knowledge' (causal, first principles) and 'shallow knowledge' (operational knowledge). These three representations are: deep level (the qualitative causal model of the heart); shallow level (the complete arrhythmia–ECG base); and shallow level compressed (compact diagnostic and prediction rules).

Figure 11 compares these representations from the points of view of: nature of knowledge, method of construction, representational formalism, size as text file (in kilobytes), direction of inference the representation supports, functional role.

The size of the compressed diagnostic knowledge of 25 kbytes apparently contradicts the size of compressed diagnostic rules in Figure 9. The difference stems from the fact that the compressed *rules* themselves are not sufficient for the diagnosis because of the loss of information in the selection of the learning subset of the arrhythmia–ECG base. To attain the full diagnostic equivalence with the complete shallow knowledge-base, we have to add descriptions of arrhythmias and ECG features that were eliminated when reducing the complete knowledge-base into the learning subset. Furthermore, 'legality' constraints and related testing procedures must also be included. After adding all of these, the size of the compressed diagnostic knowledge increases to 25 kbytes.

The ECG knowledge of this study is used in various forms in the KARDIO expert system for ECG interpretation (Lavrac *et al.*, 1985). The automatically synthesized shallow-level electrocardiographical knowledge-base is complete with respect to the level of detail of the model. In an

	I causal model of the heart	II arrhythmia knowledge-base	III compressed diagnostic knowledge
Nature of knowledge	deep causal	shallow operational	shallow operational
Method of construction	manual	automatic synthesis from I	automatic compression from II
Representational formalism	first-order logic	propositional logic	propositional logic
Size in kbytes	27	5,100	25
Direction of inference	ARR→ECG forward	ARR↔ECG both	ARR←ECG backward
Role	qualitative simulation generate II	diagnosis, provide examples for III	diagnosis

Figure 11. Comparison of different representations of electrocardiological knowledge.

assessment study of KARDIO (Grad and Cercek, 1984), cardiologists made the following estimates: the knowledge-base covers 90–95% of a 'non-selected' patient population suffering from cardiac arrhythmias (*non-selected* in the sense that these patients would not be referred to a specialist cardiologist on the account of previous examinations). In a *selected population* KARDIO-E (the version of KARDIO used in this assessment study) would correctly handle 75% of arrhythmia cases. In an actual test on 36 randomly selected arrhythmia cases from internal medical practice the arrhythmia knowledge-base was sufficient in 34 cases (94%). The failed cases are due to some incompleteness of the deep model, such as the present model's incapability to handle artificial pacemakers.

The main vehicles for implementing various knowledge representations and transformations were the following tools and techniques of Artificial Intelligence: logic programming (PROLOG in particular), qualitative modelling, and inductive learning tools. It should be noted, of course, that the inductive programs were used in this work as tools for compression of a representation and not for actual learning. Since the input information to the learning programs was complete no generalization could have occurred.

Further work can be directed along various lines including: elaboration of explanation capabilities based on the qualitative simulation; extending the model of the heart with treatment of mechanical failures; and stratifying the model by introducing several levels of abstraction. This could be based on hierarchical relations between components of the heart and attribute values. Such a hierarchy would have an important role in generating a good and concise explanation of the heart to provide a means of flexibly concentrating the explanation on points selected by the user.

Acknowledgements

The authors would like to thank the cardiologists: Professor M. Horvat, B. Cercek, A. Grad, and P. Rode of the University Medical Centre, Ljubljana, for their consultations as domain specialists in this research.

REFERENCES

Becker, J. (1985) Inductive learning of decision rules with exceptions. M.Sc. thesis. University of Illinois at Urbana-Champaign.

Forbus, K. D. (1984) Qualitative process theory. *Artificial Intelligence* **24**, 85–168.

Grad, A. and Cercek, B. (1984) *Evaluation of the applicability of the KARDIO-E expert system.* Issek Workshop 84, Bled, Yugoslavia.

Goldman, M. J. (1976) *Principles of clinical electrocardiography.* Lange Medical Publications, Los Altos.

Kleer, J. de and Brown, J. S. (1984) A qualitative physics based on confluences. *Artificial Intelligence* **24**, 7–84.

Kuipers, B. (1984) Commonsense reasoning about causality: deriving behaviour from structure. *Artificial Intelligence* **24**, 169–204.

Lavrac, N., Bratko, I. Mozetic, I., Cercek, B, Horvat, M. and Grad, A. (1985) KARDIO-E—an expert system for electrocardiographic diagnosis of cardiac arrhythmias. *Expert Systems* **2**, 46–50.

Michalski, R. S. (1983) A theory and methodology of inductive learning. In *Machine learning—an artificial intelligence approach* (eds R. S. Michalski, J. G. Carbonell and T. M.Mitchell) pp. 83–134. Tioga, Palo Alto.

Mozetic, I. (1986) Compression of the ECG knowledge-base using the AQ inductive learning algorithm. Report no. UIUCDCS-F-85-943, Department of Computer Science, University of Illinois at Urbana-Champaign.

Mozetic, I. (1986) *Knowledge extraction through learning from examples.* In *Machine learning; a guide to current research* (eds T. M. Mitchell, J. G. Carbonell, and R. S. Michalski) pp. 227–31. Kluwer, Boston.

Mozetic, I., Bratko, I. and Lavrac, N. (1984) *The derivation of medical knowledge from a qualitative model of the heart.* ISSEK Workshop 84, Bled, Yugoslavia. Updated version to appear as *KARDIO: a study in deep and qualitative Knowledge for expert systems,* MIT Press, Cambridge, Mass.

Pereira, F., Pereira, L. M., and Warren, D. H. D. (1978) *DecSystem-10 Prolog user guide.* Department of Artificial Intelligence, University of Edinburgh, Edinburgh.

Reinke, R. E. (1984) Knowledge-acquisition and refinement tools for the ADVISE META-EXPERT system. M.Sc. thesis, University of Illinois at Urbana-Champaign. Also appeared as UIUCDCS-F-84-921.