AN EXPERIMENT IN AUTOMATIC SYNTHESIS OF EXPERT KNOWLEDGE THROUGH QUALITATIVE MODELLING

I.Mozetič (1), I.Bratko (1,2), N.Lavrač (1)

(1) Institute Jozef Stefan, Jamova 39, Ljubljana, Yugoslavia
(2) Faculty of Electrical Engineering, Trzaska 25, Ljubljana

Abstract

We have developed a qualitative model of the heart for the simulation of its electrical behaviour. The model was used to automatically generate a knowledge-base of all physiologically possible combinations of cardiac arrhythmias and their corresponding ECG descriptions. The knowledge thus generated was verified by cardiologists and is used by a medical expert system. The model of the heart is formally expressed in a subset of the first-order logic. The qualitative simulation is carried out by a simple and efficient inference mechanism implemented in Prolog.

Introduction

We have developed the diagnostic part of an expert system for the diagnosis and treatment of patients with cardiac arrhythmias to be used at the University Medical Centre in Ljubljana. In the paper we concentrate on the ECG interpretation module which now includes a qualitative model of the heart. There were at least four reasons for deepening the system's knowledge by including the model. The physiological knowledge about the heart is of great importance for:

- finding the causes of arrhythmias,
- for choosing an appropriate treatment of diseases,
- for intelligent explanation of the system's answers,
- for automatically generating the electrocardiographic knowledge-base for the combinations of single arrhythmias already known to the system.

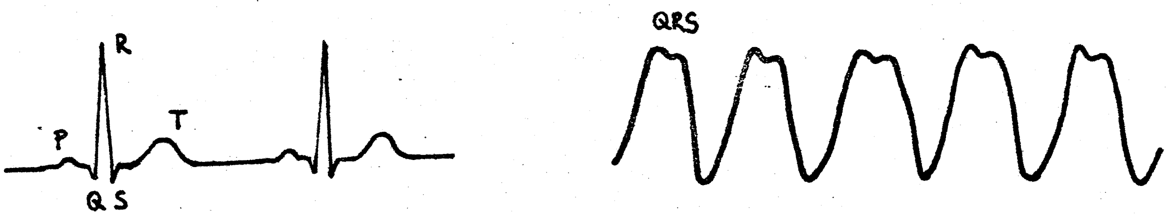
This last reason was in fact our immediate goal.

Our model is qualitative and developed along similar lines as e.g. the work of Forbus (1982) or de Kleer (1977). One reason why a qualitative model is a natural choice is that the physiological descriptions of the heart are largely qualitative. Another reason is that for a computer simulation based on a quantitative model, numerical values of the model parameters for a given patient would be needed. Such parametrs, however, practically cannot be measured. A similar aproach to medical diagnosis is exemplified in the CASNET system (Weiss, Kulikowski, Amarel, 1978).

Interpretation of ECG

Fig. 1 shows two ECG diagrams, the first for a normal heart, and the second for ventricular tachicardia, one of the arrhythmias that are handled by the system. The ECG is in the system represented by its qualitative description rather than by an actual voltage vs. time relationship. The description of a given ECG diagram consists of elementary patterns present in the ECG diagram and the relations between these patterns.

The medical literature on the relationship between various heart disorders and their corresponding ECG diagrams (e.g. Phibbs 1973, Mandel 1980) is quite indicative of the nature of these elementary patterns. However, we could not find any definite proposal, or formalisation, of a complete and compact set of such patterns. The language that we designed for describing ECG consists of a set of 10 attributes, each of which having typically 3 or 4 values. Fig. 1 shows two examples of such descriptions.



normal sinus rhythm

rhythm: regular,
frequency: between_60_100,
frequency_P: between_60_100,
regular_P: normal,
relation_P_QRS: after_P_QRS,
regular_PR: normal,
regular_QRS: normal

ventricular tachycardia

rhythm: regular,
frequency: between_100_250,
regular_P: absent,
regular_QRS: wide

Fig. 1: Two ECG diagrams and their qualitative descriptions.

The construction of a knowledge-base which covers the relation between 26 simple cardiac arrhythmias and their corresponding ECG diagrams was relatively straightforward. It was completed in consultation with cardiologists in about three months. The relation between the arrhythmias and ECG is in the system represented by rules of the form:

    if diagnosis then ECG-description

For example:

```
if    ventricular tachicardia
then  rhythm is  regular  and
      frequency is  between 100 and 250  and ...
```

Accordingly, this part of the knowledoe-base is used not to confirm some diagnosis, but to eliminate those diagnoses that contradict the patient's ECG. The remaining set of non-eliminated diagnoses (typically a few diagnoses) is then input for the differential diagnosis in which clinical data is used. The clinical knowledge ranks the remaining arrhythmias by estimating their relative likeliness.

This knowledge-base is, however, not sufficient for dealing with the more difficult problem of diagnosing the patients with multiple arrhythmias. As the number of combinatorially possible multiple arrhythmias (combined of 2, 3, 4 etc. single ones) exceeds hundred thousand, the direct specification of their ECG descriptions by exhaustive manual tabulation is practically imposible. Also, there is no systematic and exhaustive treatment of multiple arrhythmias in the medical literature.

This conclusion motivated, among other reasons, the development of a model of the heart to facilitate the automatic derivation of the relation between multiple heart failures and their corresponding ECG descriptions. With the introduction of the model, the knowledoe base was "deepened" as illustrated in Fig. 2.

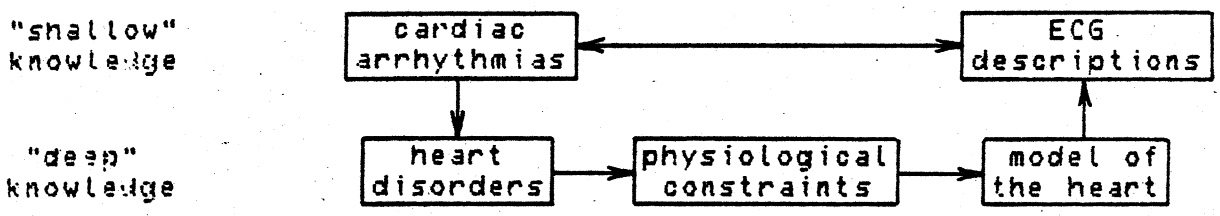


Fig. 2: "Shallow" and "deep" diagnostic knowledge. In the deep knowledge diagnoses are defined in terms of heart disorders; if a set of disorders is physiologically possible it instantiates the model of the heart; the corresponding ECG is derived by running the model.

The model of the heart

For its electrical behaviour, the heart can be represented as a network consisting of: impulse generators, impulse propagation paths and summation elements for impulses as shown in Fig. 3. In the medical literature we can find the following definition of the cardiac arrhythmias: The cardiac rhythm - be it normal or abnormal - can be characterized and classified with respect to the characteristics of impulse origin, discharge sequence and impulse conduction (WHO/ISFC Task Force 1979). These characteristics are of following types: A generator can be silent, or an extra (ectopic) generator may appear; impulse propagation paths can be partially or totally blocked, or extra paths may appear, ...

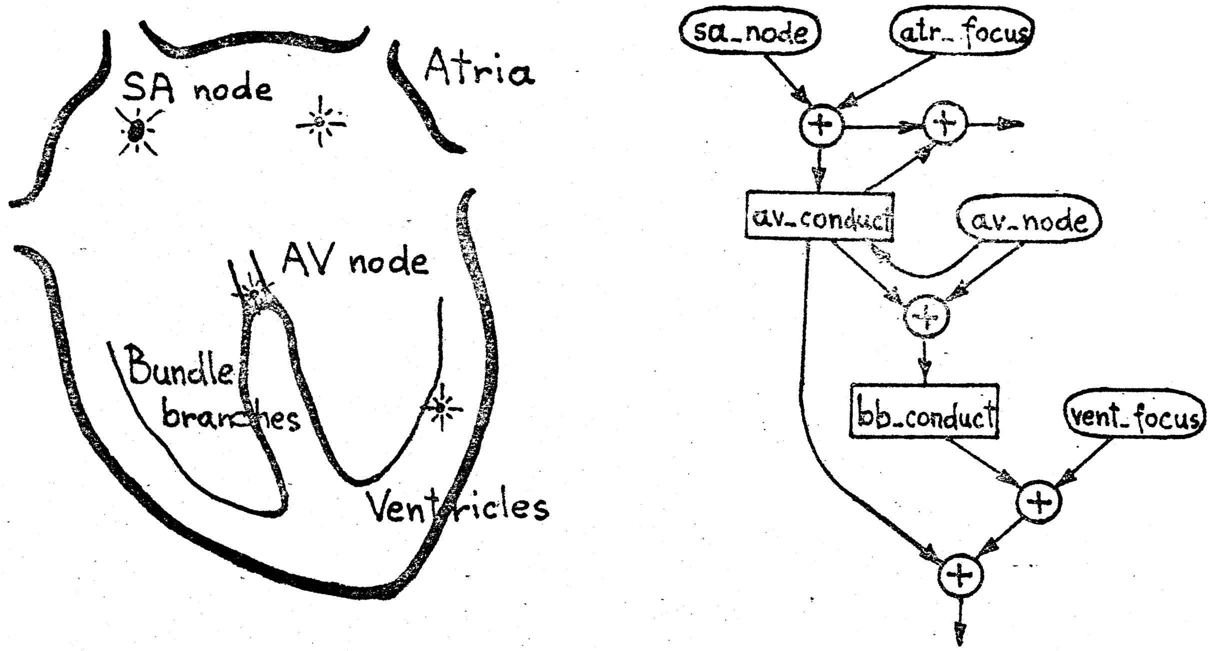


Fig. 3: A scheme of the heart and the overall logic of the model.

The state of the heart is represented by the states of its parts (4 impulse generators, 2 propagation paths and the rate of atria and ventricles). Each simple arrhythmia is defined as the abnormal state of one of the heart parts; other parts are assumed to be normal. Two or more arrhythmias can be combined if they do not contradict (e.g. are not defined by different states of the same heart part).

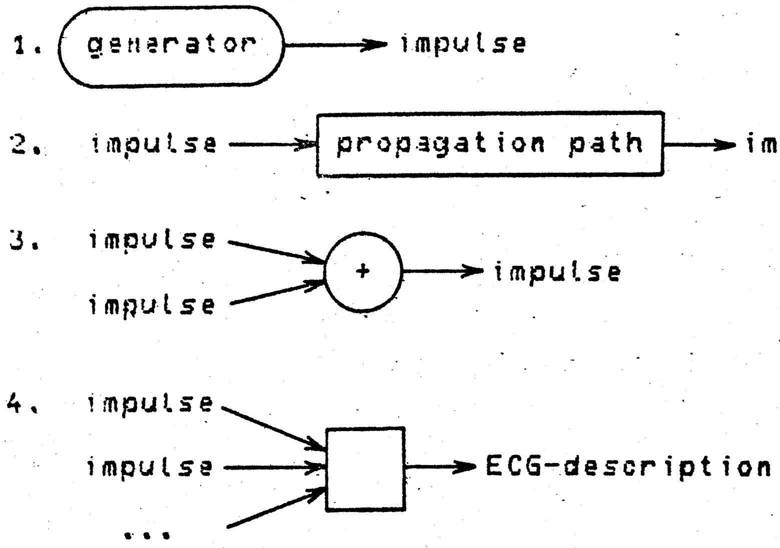
There are additional physiological constraints on the state of the heart. The constraints in the model are based on an assumption that malfunctions of impulse generators, giving permanent rhythm can be mutually combined only if there is a complete conduction block between them. Even if these malfunctions are sometimes physiologically possible, they cannot be seen on the surface ECG leads and are never considered by physicians.

The model defines relations between parts of the heart, electrical impulses and corresponding ECG descriptions. Formally, it is expressed as a set of if-then rules in a clausal form of the first-order logic. It was possible to order the list of rules according to the following principle: For each pair of rules R1 and R2, R2 may proceed R1 only if no literal in the consequent of R2 occurs in the antecedent of R1. This implies that there is no cyclic or recursive rules. This constraint on the list of rules facilitates fast, one pass execution of the model.

The inference mechanism that runs the model for a given multiple arrhythmia is relativly simple. The model of the heart is first instantiated with the state of the heart parts. The states of the heart parts are added to the set of rules as unit clauses. Then the inference mechanism sequentially passes through the rules and by applying modus ponens derives all positive facts. One pass through the ordered list of rules

suffices for generating all possible ECG descriptions which correspond to this multiple cardiac arrhythmia.

Rules, 62 of them, define relations of the following types:



An example of a rule is:

<u>if</u> there are ectopic impulses at the His bundle and in the supraventricles originating at the AV focus
<u>then</u> this results in the following ECG features: either a short PR interval, or no P wave, or P wave after the QRS complex

These three cases result from the "qualitative summation" (as also percieved in the ECG diagram) of two signals which can be relatively shifted in time in three ways as shown in Fig. 4.

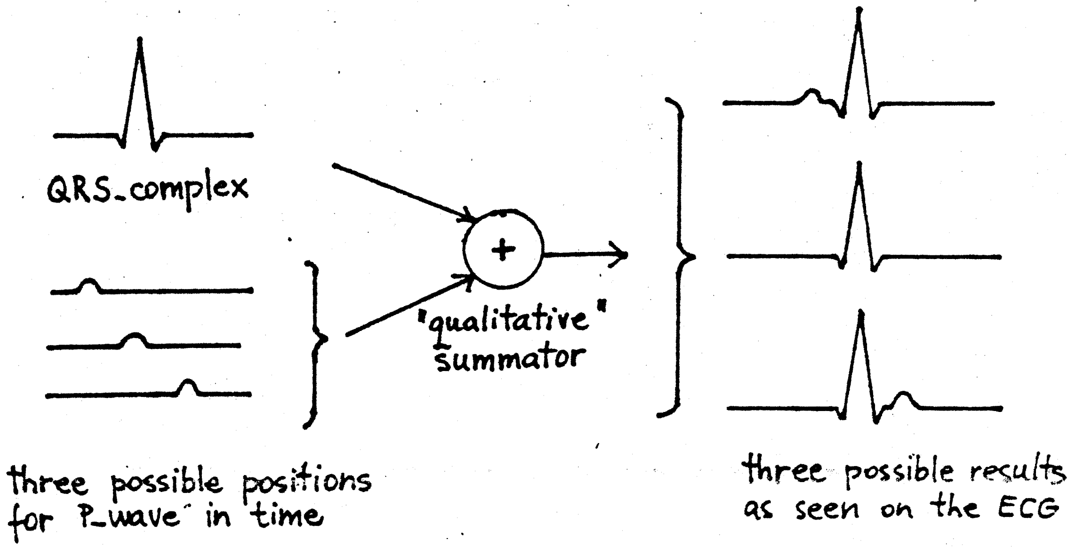


Fig. 4: The qualitative summation of ECG patterns.

Implementation and results

We ran the system for all combinations of simple cardiac arrhythmias. A large proportion of the corresponding states of the heart parts were recognized as physiologically impossible. For the physiologically possible arrhythmias the model generated corresponding ECG descriptions. The following table shows the number of mathematically and physiologically possible arrhythmias against the number of their constituent single arrhythmias.

No. of constituent arrhythmias	1	2	3	4	5	6	7
No. of mathematical combinations	23+3	253	1771	8855	33649	100947	...
No. of physiologically possible combinations	13+3	85	231	163	73	20	0

Note that some arrhythmias cannot occur alone (e.g. blocks), but only in combination with others (e.g. sinus rhythm). Three arrhythmias cannot be combined with others.

The whole system is implemented in Prolog on DEC-10 (Pereira, Pereira, Warren 1978). The compiled program generated ECG descriptions for all combinations of arrhythmias in 340 CPU seconds.

The thus obtained knowledge-base of ECG descriptions for all possible multiple arrhythmias can be used for diagnosis. If an explanation is requested then for a given ECG description the corresponding states of the heart parts are retrieved by table look-up. The model is then re-run for these states and its trace can serve as an explanation. We have not yet found an efficient implementation of the model to be run in the inverse direction, i.e. from the ECG toward the diagnoses.

Conclusion

The model facilitated, as our main result, the automatic derivation of an exhaustive catalog of multiple arrhythmias and their corresponding qualitative ECG descriptions.

The present system handles 588 combinations of arrhythmias, all of them physiologically possible and majority of them can be observed in everyday clinical praxis. The importance of recognition of these arrhythmias is very different. Sometimes the diagnosis of certain rhythm disturbances is critical for the treatment. On the other hand, some arrhythmias are only of theoretical interest without practical consequences for the patient.

The model of the heart also provides a good basis for the system's explanation of its own reasoning, and (as hoped) for the treatment decision-making. We are planning to extend the model in two directions:

- to handle the mechanical activity of the heart as well as electrical;
- to provide causal reasoning about the effects of drugs and their interaction for treatment decision-making.

Acknowledgement

The authors would like to thank cardiologists prof. M. Horvat, P. Rode, B. Čerček and A. Grad of the University Medical Centre in Ljubljana for the consultations in the development of the knowledge-base.

References

[1] de Kleer J. (1977) Multiple representations of knowledge in a mechanics problem-solver. Proceedings of IJCAI 1977.

[2] Forbus K.D. (1982) Qualitative Process Theory. MIT, AI Lab, AI Memo 664.

[3] Mandel W.J. (1980) Cardiac Arrhythmias. J.B. Lippincott Co.

[4] Pereira F., Pereira L.M, Warren D.H.D. (1978) User's Guide to DEC-10 Prolog. University of Edinburgh, Dept. of AI.

[5] Phibbs B. (1973) The Cardiac Arrhythmias. The C.V. Mosby Co.

[6] Weiss S.M., Kulikowski C.A., Amarel S. (1978) A model-based method for computer-aided medical decision-making. Artificial Intelligence, Vol. 11, pp. 145-172.

[7] WHO/ISFC Task Force, Classification of cardiac arrhythmias and conduction disturbances. American Heart Journal, Vol. 98, No. 2, Aug. 1979.