

PROCESSING AND ANALYSIS OF FOETAL PHONOCARDIOGRAPHIC SIGNALS

James Thomas Edward McDonnell

**A thesis submitted for the degree of
Doctor of Philosophy
to the Faculty of Science, of the University of Edinburgh.**

Department of Electrical Engineering

1990



ABSTRACT

There has been renewed interest in foetal phonocardiography since the recent advent of wide-bandwidth (0.5-250 Hz) phono-transducers which yield higher signal-to-noise ratios than were hitherto obtainable. With this passive and non-invasive transducer there is the potential to realize the goal of long-term continuous foetal heart monitoring. In order to achieve this goal, signal processing and analysis methods are required to automatically extract from the phonocardiogram (PhCG) the beat-to-beat temporal parameters of foetal cardiac function. The research detailed in this thesis investigates the temporal and spectral morphology of the principal heart sounds, and proposes a method which detects and identifies the first and second heart sounds in the PhCG on a beat-to-beat basis.

Examination of the principal heart sounds in 252 PhCGs, each of 4.1 seconds duration, from 19 subjects has found that wide variations exist in the temporal morphology of these sounds. It was also observed that the PhCG is susceptible to persistent contamination by many adventitious sounds. These latter sounds often coalesce with the principal heart sounds and obscure the end points of both. Second heart sounds were noted to be subject to sudden and severe attenuation, but were more morphologically stable than first heart sounds.

The frequency analysis of the principal heart sounds was performed using Fourier transformation, and a high-resolution parametric technique based on the Burg algorithm. In all the 741 principal heart sounds examined from 12 subjects, it was observed that the majority of the energy in their spectra is concentrated in the previously unmonitored 20-40 Hz band. In this sample set of first and second heart sounds, it was found that each set of valves does not produce a spectral signature with which it may be uniquely associated. Results are presented to substantiate this observation.

A procedural knowledge-based system (KBS) was developed to process and analyse the PhCG with the objective of detecting and identifying individual first and second heart sounds. In this the KBS has been successful. The KBS analyses PhCGs over a wide range of FHRs (80-220 bpm), irrespective of either short-term variability in FHR or contamination by adventitious sounds. Results are presented which illustrate the performance of this system over the spectrum of PhCG types.

DECLARATION OF ORIGINALITY

The research recorded in this thesis, and the thesis itself, is the original and sole work of the author.

James Thomas Edward McDonnell

ACKNOWLEDGEMENTS

I am most grateful to Dr. James Dripps for having introduced me to this area of research and for his continual support. To Professor Peter Grant who has been a constant source of encouragement throughout the course of the work, I am especially indebted.

I should like to thank Dr. Nigel Colley for providing the recordings of foetal phonograms, and the Department of Education for Northern Ireland for the award of a research studentship.

CONTENTS

Abstract	ii
Declaration of Originality	iii
Acknowledgements	iv
Contents	v
Abbreviations	ix
Glossary	x
1 INTRODUCTION	1
1.1 FOETAL MONITORING	1
1.2 SIGNAL PROCESSING AND ANALYSIS	3
1.3 BIO-MEDICAL SIGNALS	4
1.4 CONSPECTUS OF THE THESIS	4
2 AUSCULTATION AND PHONOCARDIOGRAPHY	7
2.1 AUSCULTATION	8
2.2 PHONOCARDIOGRAPHY	12
2.3 ORIGIN OF HEART SOUNDS	20
2.4 FOETAL HEART	21
Heart Valves	23
Valvular Production of Heart Sounds	23
2.5 IDEALIZED PHONOCARDIOGRAM	25
2.6 SUMMARY AND CONCLUSION	26
3 ACQUISITION AND CONDITIONING OF THE FOETAL PHONOGRAPHIC SIGNAL	28
3.1 TRANSDUCER	29
3.2 METHOD OF SIGNAL ACQUISITION	32
3.3 SIGNAL RECORDING	33
3.4 SIGNAL CONDITIONING	33
3.5 INSTRUMENTATION	35
3.6 DATA SELECTION	35

3.7	FILTERING OF THE PHONOGRAM	37
3.8	DOMAIN PROBLEMS	39
3.9	VISUAL IDENTIFICATION OF HEART SOUNDS	45
3.10	SUMMARY	51
4	FREQUENCY ANALYSIS OF FOETAL HEART SOUNDS	52
4.1	ADULT HEART SOUND ANALYSIS	55
4.2	HEART SOUND SELECTION	57
4.3	DISCRETE FOURIER TRANSFORM	58
4.4	LIMITATIONS OF DFT PROCESSING	58
	Aliasing	58
	Windowing	59
	'Picket-Fence' Effect	61
	Frequency Resolution	63
4.5	HAMMING WINDOW	64
4.6	SPECTRUM INTERPRETATION 1	64
4.7	DOLPH-CHEBYSHEV WINDOW	70
4.8	MAIN-LOBE WIDTH OF HAMMING AND DOLPH-CHEBYSHEV WINDOWS	71
4.9	SPECTRUM INTERPRETATION 2	73
4.10	PARAMETRIC METHODS	73
4.11	RELATIONSHIP OF THE AR PROCESS TO LINEAR PREDICTION	79
4.12	BURG (HARMONIC) ALGORITHM	81
	Optimum Taper for Burg Algorithm	84
4.13	SELECTION OF MODEL ORDER	84
4.14	SPECTRUM INTERPRETATION 3	85
4.15	SUMMARY AND CONCLUSION	87
5	KNOWLEDGE-BASED SYSTEMS AND SIGNAL UNDERSTANDING	88
5.1	KNOWLEDGE-BASED SYSTEMS	91
	Knowledge Base	92
	Data Base	92
	Control Mechanism	94
	User Interface	94
5.2	KNOWLEDGE ACQUISITION AND REPRESENTATION	95
	Knowledge Acquisition	95
	Knowledge Representation	96

5.3	KNOWLEDGE-BASED SIGNAL UNDERSTANDING	97
5.4	THE BLACKBOARD MODEL	98
	Knowledge Sources	99
	Global Data Base	100
	Control	101
	Signal Understanding and the Blackboard Model	103
5.5	KNOWLEDGE-BASED UNDERSTANDING OF PHONOCARDIOGRAMS	104
	Knowledge Acquisition	104
	Knowledge Organization and Representation	107
	Signal Abstraction and Knowledge Application Strategy	110
	System Architecture and Implementational Framework	113
	Evaluation of System Performance	115
5.6	SUMMARY	115
6	PHONOCARDIOGRAM PROCESSING AND ANALYSIS: SYSTEM 1	117
6.1	PHONOCARDIOGRAM PROCESSING	119
	Cross-Correlation	120
	Template Selection	121
	Correlator Output	123
	Further Processing	126
	Segmentation	127
	Phonocardiogram Boundary Conditions	134
	Limits on the Number of Peak Clusters	134
6.2	PHONOCARDIOGRAM ANALYSIS	136
	Checks Prior to Analysis	136
	Location and Analysis of Solution Islands	138
	Expansion From Solution Islands	140
6.3	PRESENTATION OF RESULTS	147
6.4	DISCUSSION OF RESULTS	153
6.5	SUMMARY AND CONCLUSION	181
7	PHONOCARDIOGRAM PROCESSING AND ANALYSIS: SYSTEM 2	183
7.1	PHONOCARDIOGRAM PROCESSING	185
	Excursion Processing	185
	Removal of Insignificant Excursions	189
	Segmentation	193

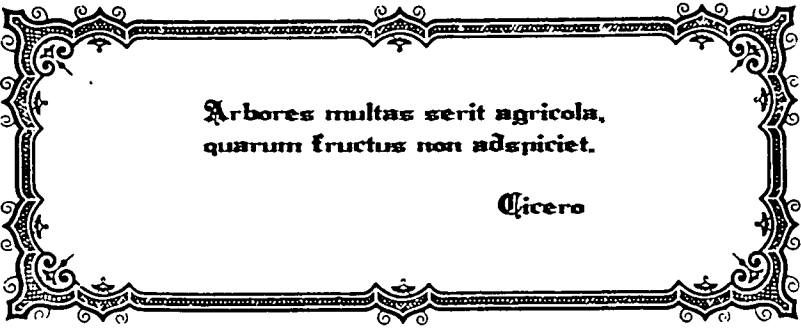
Removal of Insignificant Events	199
Further Processing	201
Pairing of Events	204
7.2 PHONOCARDIOGRAM ANALYSIS	206
Removal of Unlikely Event Couplings	207
Location and Analysis of Solution Islands	209
Expansion from Solution Islands	210
Phonocardiogram Boundary Conditions	213
Conflict Resolution	214
Missing Second Heart Sounds	217
7.3 PRESENTATION OF RESULTS	220
7.4 DISCUSSION OF RESULTS	225
7.5 SUMMARY AND CONCLUSION	242
 8 SUMMARY AND CONCLUSIONS	 244
8.1 SUMMARY AND CONCLUSIONS	244
8.2 EXTENSIONS TO THE RESEARCH	249
 REFERENCES	 253
APPENDIX (Original Publications)	268

ABBREVIATIONS

ADC	Analogue-to-digital converter
bpf	Band-pass filter
bpm	Beats per minute
DFT	Discrete Fourier Transform
DSP	Digital signal processing
ECG	Electrocardiogram
FFT	Fast Fourier transform
FHR	Foetal heart rate
FHS	Foetal heart sounds
hpf	High-pass filter
KA	Knowledge area
KBS	Knowledge-based system
KS	Knowledge source
lpf	Low-pass filter
PhCG	Phonocardiogram
uFCG	Ultrasound foetal cardiogram

GLOSSARY

Adventitious sounds	Those sounds and infra-sounds in the phonocardiogram which are not principal heart sounds.
Arrhythmia:	Abnormal alternations in the rhythm of the heart.
Atrium:	The upper chambers of the heart that receive blood directly from the veins.
Auscultation:	The act of listening to the heart, lungs, <i>etc.</i> as a medical diagnostic aid.
Bradycardia:	Relatively slow heart action.
Cardiohaemic:	Concerning the heart and the blood.
Diastole:	Widening of the chambers of the heart during which they fill with blood. The duration of mechanical diastole is the time interval between the beginning of a second heart sound and the beginning of the first heart sound of the following heartbeat.
Dysfunction:	An impairment or abnormality in functioning.
Gravida:	A pregnant woman.
Phonocardiogram:	A graphic record of the sounds produced by the functioning of the heart.
Phonocardiograph:	An instrument for detecting and recording the acoustic activity arising from the heartbeat.
Phonogram:	A composite graphic record of the sounds produced by the functioning of the heart, and the infra-sounds arising from foetal breathing movements and foetal body movements.
Principal heart sounds:	A generic name for the first and second heart sounds which are of valvular origin.
Systole:	The contraction of the heart by which blood is forced onwards and the circulation kept up. The duration of mechanical systole is the time interval between the beginning of a first heart sound and the beginning of the following second heart sound of the same heartbeat.
Tachycardia:	Relatively rapid heart action.
Ventricles:	The lower two chambers of the heart that pump blood into the arteries.



Arbores multas serit agricola,
quarum fructus non adspiciet.

Cicero

CHAPTER 1

INTRODUCTION

The common and distinctive feature of previous attempts to automatically determine instantaneous foetal heart rate using phonocardiography is that they have all been unsuccessful. In retrospect this is attributable to the design of the phono-transducers used, poor choice of filtering threshold frequency and the high levels of noise intrinsic to the phonocardiogram. These factors led to the abandonment of phonocardiography in the 1960s in favour of ultrasound-based foetal heart rate monitoring. However, since the mid-1980s there has been renewed interest in phonocardiography brought about by the development of superior phono-transducers [1,2]. In itself, improved signal acquisition is not enough — there must also be effective processing and analysis of the phonocardiographic signal to exploit its full potential. It is to the latter requirement that this thesis is addressed.

The following sections briefly introduce the background to the field of foetal heart rate monitoring, and elaborate on the general computational aspects associated with the interpretation of this particular bio-medical signal. The chapter ends with a conspectus of the work contained in this thesis.

1.1 FOETAL MONITORING

The aims of foetal monitoring are primarily to prevent foetal death and

morbidity, and to deter unnecessary intervention during the intrapartum period (conception to the onset of labour). However, there are few physiological parameters which can be obtained non-invasively to indicate the state of the foetus. The evaluation of foetal heart rate (FHR) has been the mainstay of foetal monitoring [3-7] since the discovery by Kegardec [8] in 1822 that foetal cardiovascular function is a sensitive indicator of distress.

There are currently three non-invasive techniques which provide records of the FHR:

- (1) the electrocardiogram (ECG) which registers the electrical depolarization of the heart
- (2) the Doppler ultrasound cardiogram (uFCG) which registers the openings and closings of the cardiac valves
- (3) the phonocardiogram (PhCG) which registers the sounds and infra-sounds produced by the shutting of the cardiac valves.

Ultrasound cardiography is, by far, the most commonly applied of the above techniques. This has arisen because the foetal ECG is inordinately difficult to record non-invasively, and the PhCG reputedly had a very poor signal-to-noise ratio. The main disadvantage with the uFCG is that there is concern about the possible biological consequences of irradiating the developing embryo with ultrasound energy [9]. This has consequently confined its use to brief and intermittent periods of monitoring.

The development of a new generation of phonographic transducers [1,2] which are superior to their predecessors has brought renewed interest in this neglected FHR monitoring technique. The advantage of phono-transducers is that their transduction principle is completely passive, *i.e.* it depends solely on the energy generated by the measurand. Thus phono-transducers, unlike ultrasound transducers, have the potential

to realize the goal of long-term continuous foetal heart rate monitoring, which is highly desirable in difficult pregnancies. The wide-bandwidth of these new phono-transducers has enabled the recording of the complete spectrum occupied by the foetal heart sounds. This is in contrast to earlier phono-transducers which had severely attenuated the valvular sounds and contributed to the poor and undeserved reputation of phonocardiography.

1.2 SIGNAL PROCESSING AND ANALYSIS

If a phonocardiographic signal were being recorded onto paper, at a paper speed of 53 mm/s, a record of length 190 metres would result after one hour. Clearly this is not a practical way to monitor FHR, both from the point of view of the sheer volume of data to be analysed and the time it would take to manually perform the task. In these situations it is not enough to process the phonocardiographic signal and then display the results of the processing; what is additionally required is phonocardiographic signal analysis.

The processing of a signal is intended to accentuate certain pre-determined features in that signal over the background noise. The result of this processing is ultimately another signal. A signal analysis system, however, takes a signal as input but has a symbolic description of the signal at its output. Signal analysis allows the information conveyed by the original signal to be more easily interpreted because the amount of data has been reduced to a few salient symbolic descriptors.

A signal processing and analysis system consists of the three sub-systems of detection, segmentation, and classification which are not necessarily independent modules. Detection is the process whereby the signal-to-noise ratio of the signal is enhanced; it is usually achieved by numerical computation. Segmentation localizes

individual signal events in time by specifying their end points. Finally, classification is concerned with the transformation of the information in the original waveform from a signal-based to a symbol-based representation. For example, the result of processing and analysis on a foetal phonocardiographic signal might consist of a list of temporally located and identified first and second heart sounds.

1.3 BIO-MEDICAL SIGNALS

Bio-medical signals present many challenges to any automatic processing and analysis system. Unlike signal processing and analysis in other fields, the characteristics of bio-medical signals are often not well specified, and they are subject to considerable non-stationarities. Compounded with this is the frequent requirement that such transduction be performed by non-invasive means, *i.e.* without piercing the skin. Thus the transducer is often remote from the point of origin of the signal. This separation subjects the signal to additional contamination by phenomena from other bio-medical processes which are simultaneously occurring in the vicinity. It has been a frequent observation that the peculiarities of bio-medical signals often limit the extent of achievable signal processing and analysis. As Cohen points out:

the accuracies and confidence limits that come out of bio-medical signal processing are usually not very high, at least in terms used in other engineering disciplines. [10]

1.4 CONSPECTUS OF THE THESIS

The research recorded in this thesis investigates methods for the automatic processing and analysis of foetal phonocardiographic signals. The objective of the research was to detect, segment, and classify the principal foetal heart sounds. Once these events have been located in time, they provide markers to allow the measurement of the instantaneous FHR, beat-to-beat differences in FHR, and the

duration of systole (cardiac action phase) and diastole (cardiac relaxation phase). These measures are sensitive indicators of cardiac function, which in turn reflect foetal well-being or distress.

Chapter 2 introduces the historical background to the monitoring of the foetal heart, from its beginning in auscultation to the present status of phonocardiography. The origin of heart sounds is discussed and the controversy which has surrounded this issue is noted. Finally, the constituent events in one cycle of an idealized PhCG are described along with their relative temporal positions.

Chapter 3 is concerned with the acquisition and conditioning of the phonocardiographic signal. Each stage in this process is introduced, from signal transduction to the filtering which ultimately produces the phonocardiogram. As with the origin of heart sounds, there has been disagreement among investigators regarding the filtering threshold frequency to apply to obtain the best PhCG. However, it is demonstrated that the region of the spectrum where the energy of the foetal heart sounds is concentrated falls below the region which was previously monitored. Examples are presented to illustrate the variability in every characteristic of the foetal heart sounds, and the persistent occurrence of adventitious sounds in the PhCG. It is noted that despite the high levels of noise in the signal, the eye can often discern the presence of the foetal heart sounds with comparative ease.

The analysis of the frequency content of both principal heart sounds in the PhCG is described in chapter 4. The spectra of these sounds were obtained with the intention of deriving certain spectral features which could be uniquely associated with each sound. These would then have been used to classify the principal heart sounds in a given PhCG. Two methods of spectral analysis are presented: discrete Fourier transformation, and a high-resolution parametric technique based on the Burg

algorithm. These methods were applied to 741 principal heart sounds. Representative examples are presented of the results obtained.

The concept of knowledge-based systems is introduced in chapter 5, and in particular its application to signal understanding. This chapter constructs the theoretical framework for the systems presented in chapters 6 and 7. A procedural knowledge-based system is adopted which uses meta-knowledge to control the application of rules. The method by which a phonocardiographic signal is analysed begins with a search of the abstracted versions of the signal to ascertain areas of minimal ambiguity where the identity of signal events is apparent. From these 'solution islands' the system expands the analysis into the ambiguous regions which may lie on either side. This expansion continues until either the complete signal has been analysed or the analysis is abandoned because of too much uncertainty regarding the location and identity of events.

The implementation of the knowledge-based system concept in two different realizations is detailed in chapters 6 and 7. Results are presented which illustrate the performance of both systems over the spectrum of PhCG types.

Chapter 8 summarizes the methods and conclusions of the thesis. Finally, suggestions are proposed for possible extensions to the research.

CHAPTER 2

AUSCULTATION AND PHONOCARDIOGRAPHY

As a consequence of the natural shielding afforded to the foetus, and its inaccessibility, there are few non-invasively determinable parameters available which can be used to assess the physical condition of the foetus. The foetal heart rate (FHR) was the earliest parameter to be measured, and its paramount position has been maintained to this day. Originally the FHR was determined by listening to the beating of the foetal heart (auscultation) either with the unaided ear, or through a stethoscope. Later, with the advent of recording techniques, it became possible to obtain a 'sound-picture' of these acoustic vibrations with time (phonocardiography). This graphic record has allowed the continuous monitoring of the foetal heart over extended periods of time and further insights into the condition of the foetus than were previously possible with auscultation.

This chapter reviews the historical development of the technique of FHR monitoring using the sounds produced by the beating of the heart itself, from its beginnings in direct auscultation to the automated methods of the present day. The chapter concludes with a description of those dynamic events within the foetal heart which produce an acoustic signal and the significance of the temporal separation

between these sounds in the derivation of the parameters of cardiac function.

2.1 AUSCULTATION

The sounds produced by the beating of the foetal heart have now been listened to, observed, and studied for well over a century and a half. The first recorded specific reference to obstetric auscultation is generally thought to be a transcript of a statement attributed to François-Issac Mayor of Geneva in 1818 [8,11,12]. This record is contained in an editorial footnote in a Swiss Medical Journal [13] to a report on a meeting of the Académie Royale des Séances, Paris, at which Laennec (1781-1826) presented his findings on cardiac sounds, uncovered with the aid of the newly invented stethoscope. This footnote reports Mayor's claim that the life or death of a foetus could be determined by listening for the foetal heart sounds by direct application of the ear to the maternal abdomen [14]. However, Gunn [14] and Pinkerton [8] trace the origin of the first recorded account of the hearing of the foetal heart sounds back to Phillipe Le Goust, a physician in Niort, France, in 1650. This earlier account written by Le Goust in a satirical poem directed against a certain other physician, Marsac, derides his colleague for claiming that:

the foetus jumps, turns, twirls, and moves about freely and is able to change its position frequently and that its heart beats "comm'un traquet" (like a mill clapper) [15].

As Gunn [14] points out, this discovery was fruitless insofar as neither Marsac nor Le Goust appears to have applied it to diagnosis or prognosis.

Over one hundred and fifty years later came the most notable advancement in the auscultation of both the adult and the foetus with the invention by René Laennec in 1816 of the stethoscope — at first only a wooden cylinder twelve inches long [8]. It was with this new instrument in 1821 that the French physician, Jacques Alexandre Le Jumcau, Vicomte de Kegaradec (1788-1877), while attempting to hear the noise which

he hypothesized to be produced by the movements of the foetus in the amniotic fluid, accidentally rediscovered the foetal heart sounds. From all accounts it appears he was unaware of Mayor's previously reported discovery [11, 16]:

One day whilst examining a patient near term and trying to follow the movements of the foetus with the stethoscope I was suddenly aware of a sound that I had not noticed before; it was like the ticking of a watch. At first I thought I was mistaken, but I was able to repeat the observation over and over again. On counting the beats I found that these occurred 143-148 times per minute and the patient's pulse was only 72 per minute [8].

Kegaradec applied this discovery — which he correctly deduced to be the beat of the foetal heart — to other gravidae, and published his findings [17] in a Memoir to the Royal Academy of Medicine, Paris in 1821. It is reported that the Academy received the discovery with approbation [16]. In spite of Mayor's earlier record of hearing the foetal heart sounds, it is only with the work of Kegaradec that the significance of obstetric auscultation in the monitoring of the foetus can be said to have been established.

Such was the quality of Kegaradec's treatise that Hellman in 1958 [12] stated:

So assiduously did he pursue his discovery and so meticulously did he record his investigations that subsequent observations have revealed little new.

Although, surprisingly, Kegaradec was not an obstetrician either before or subsequent to his discovery [18], he proposed the following as the practical implications of his discovery: the diagnosis of pregnancy; an indication of foetal life when movements are not detected; the diagnosis of multiple pregnancy; the determination of foetal position; and a way of judging the state of health or disease of the foetus by changes in the strength and frequency of the heart sounds [8, 17].

The first man in the British Isles to hear the foetal heart sounds using Laennec's stethoscope was John Creery Ferguson (1802-1865) [19], who was later to become

Professor of Medicine at Queen's College, Belfast. Pinkerton [20] reports that:

His [*i.e.* Ferguson's] first patient was an unfortunate young woman who presented herself to his somewhat unsympathetic scrutiny at the Dublin General Dispensary in November 1827. This patient is of great historical interest as she was the first in whom the foetal heart was heard by mediate auscultation in the British Isles. On being told that she was pregnant the patient received the news, says Ferguson, with "extreme indignation, indeed the young lady's histrionic talent was of the first order and such was her well-feigned agony at the idea of her innocence being ever suspected, that had I not the positive evidence of my senses to confirm the opinion I had expressed I should have felt extremely uncomfortable."

This new technique of obstetric auscultation, with or without a stethoscope, was not universally acknowledged, nor was its benefit immediately understood. Indeed, one French obstetrician wrote an open letter to Kegaradec in which he strongly advised him to 'abandon these toys of ignorance truly prejudicial to science and to the well-being of an amiable and interesting sex' [8].

In 1833, within twelve years of Kegaradec's treatise to the French Royal Academy, Evory Kennedy[†], Assistant Master at the Rotunda Hospital, Dublin, published a monumental monograph on the subject of auscultation in obstetrics [19], the title of which was 'Observations of Obstetric Auscultation with an Analysis of the Evidences of Pregnancy and an Inquiry into the Proofs of the Life and Death of the Foetus in utero' [21]. In writing the book it was Kennedy's stated aim to convince his colleagues of the value of obstetrical auscultation [18]. In this he succeeded, for as a result of his reports from the Rotunda the notion of obstetric auscultation spread throughout the British Isles. In common with Kegaradec, the vast improvements in obstetrical medicine over the last one hundred and fifty years have not invalidated Kennedy's findings nor his explanation of them. As Goodlin, [18] writing in 1979 on the subject matter addressed in Kennedy's book, states:

[†]It is interesting to note that it was in the Rotunda, Dublin in 1835 that Kennedy established the first maternity unit of any hospital.

most of his [*i.e.* Kennedy's] clinical impressions regarding the significance of various foetal heart rate patterns appear compatible with current concepts.

The remainder of the nineteenth century brought the further codification of the normalities and abnormalities of the foetal heart rate and their causes. In 1833, Hohl [16] found that a rapid heart rate (tachycardia) was symptomatic of foetal distress. In the same year, Kennedy [16] had ascertained the significance of a slow heart rate (bradycardia). By 1836, von Hoefft [22] had described the normal range of the foetal heart rate, and the fact that the rate decreased with gestational age. Just before the turn of the century, von Winkel [18] suggested definite criteria by which foetal distress could be discerned on the basis of foetal heart rate alone; namely, a FHR under 120 beats per minute (bpm) or over 160 bpm.

The reader is directed to the extensive lists of references contained in the works of Goodlin [18], Gultekin-Zootzman [17] and Solum [16] for a more detailed exposition of the significance of abnormal FHR patterns identified by auscultation in the nineteenth century which, however, lie beyond the compass of this brief review.

By the end of the century the value and importance of obstetric auscultation throughout the period of confinement and, especially, during labour, had been firmly established. Even to this day the discoveries of these pioneers in auscultation remain substantially unchanged [16]. When one considers that there were no means to record the sounds heard with auscultation, these early observations of the foetal heart rate become all the more remarkable [18].

Modern methods of electronic foetal heart monitoring have not displaced auscultation as an option available to the obstetrician. Auscultation is routinely used to ascertain whether continuous long-term monitoring by electronic means is called

for. The obstetrical stethoscope and aural auscultation are still both very much in evidence.

2.2 PHONOCARDIOGRAPHY

The development and refinement of photography and electrical measurement techniques in the latter half of the nineteenth century brought the means to record the acoustic signal from the foetal heart. This recording in graphical form of the heart sound vibrations with time is referred to as a phonocardiogram (PhCG). There is an obvious advantage from an experimental and clinical point of view in having a permanent record of the foetal heart sounds (FHS). Furthermore, depending on the bandwidth of such a recording system, it would also be possible to present acoustic information inaudible to the ear — particularly as the FHS lie on and below the threshold of audibility.

The making of the first visual record of the sounds audible through auscultation (figure 2.1) of the foetal heart is attributed to Hofbauer and Weiss [23] of Jena in 1908.

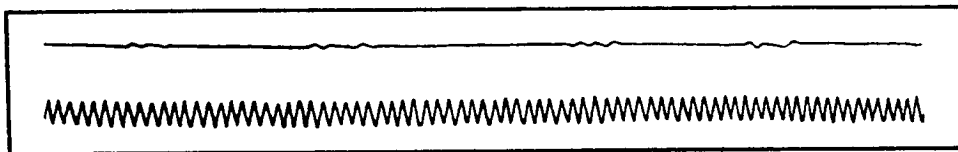


Figure 2.1 Phonocardiogram recording (upper trace) by Hofbauer and Weiss (time progresses from right to left) [24].

Such was the clarity of the recording, reports Sampson [24] in 1926, that the durations of the systolic phases could be measured. According to Gunn the apparatus used by Hofbauer and Weiss, which they called a phonoscope, consisted of a 'soap bubble in the centre of which was a thread coated with silver for the movements to be recorded

on a photographic plate' [14].

Foetal phonocardiography had now brought exactness to the art of auscultation and facilitated foetal monitoring. It was now possible with this permanent visible record of the foetal heart sounds to determine accurately the instantaneous FHR and its progression over extended periods of time. This will be appreciated when it is considered that even in its 'normal' range the slowest beat of the foetal heart is twice per second, and in the extreme over three times per second. Auscultation, however, was only ever able to give the average rate over 15 or 30 seconds [25], and its accuracy was very much dependent on the experience of the auditor [26].

A major problem facing foetal phonocardiography was, and has been, the difficulty in registering the very weak heart sounds — something it has in common with auscultation. In 1923 the idea of using thermionic wireless valves to electronically amplify the foetal heart sounds was introduced by Jacobsohn [27], it having previously been demonstrated in the case of adult heart sounds in 1921 [14]. This idea was then successfully realized in the same year by various workers: Falls and Rockwood in the USA [28], as well as Keipner [11], and Schaffer [17] in Germany.

1924 saw the introduction of electrical filtering in place of the usual acoustical filtering for the elimination of adventitious sounds [29,30] and also, according to Gamble to:

transmit only such sounds as have a frequency above or below a chosen point. This has given us definite information in regard to absolute pitch of the sounds in which we are now interested from a diagnostic standpoint, and has enabled the new apparatus to be designed with a view to marked efficiency in the frequency ranges in question.

Gamble also devised apparatus which enabled the heart sounds to be distributed to multiple earphones for use in teaching classes of medical students.

Until now the foetal heart sounds had been used merely as markers for FHR measurement with very little attention paid to the actual sounds themselves. This situation continued until the extensive study by Sampson, McCalla, and Kerr [24] in 1926 into the characteristics of the foetal heart sounds themselves, and the duration of the systolic and diastolic phases. Sampson *et al.* were the first to note that the systolic phase varies little across the range of FHRs, and that variations in FHR are effected by changes in the duration of the diastolic phase. Their apparatus, which is illustrated in figure 2.2, consisted of: a stethoscope, valve amplifiers, various selectable fixed filters, and a string galvanometer/photographic film arrangement for recording the heart sounds.

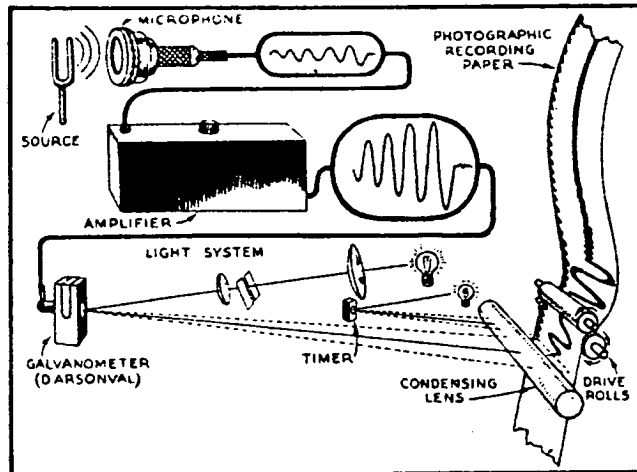


Figure 2.2 Schematic diagram of the electrostethograph [31].

In 1930 Hyman, using apparatus similar to Sampson's, investigated various irregularities in the rhythm of the foetal heart beat, and provided documentary evidence of these irregularities with PhCGs [32]. He also described instances when auscultation was apt to give an incorrect assessment of the FHR, particularly in cases of arrhythmia (abnormal alternations in the rhythm of the heart). However, when these periods were phonocardiographically recorded, the examination of the PhCG

displayed, quite plainly, the abnormality of the beat.

Although phonocardiographic equipment had advanced in the twenty years since its inception, its imperfections still prevented it from accomplishing the continuous auscultation some were proposing [11,33]. It was realized that intermittent observations of a phenomenon which continually varied could but provide incomplete information, and hindered the exploitation of the diagnostic potential of phonocardiography for early warning of foetal distress. Matthews in 1937 complained that 'the uncertainty and shortcomings of present-day methods are well known and keenly appreciated', and again that the recording systems were 'too complicated and expensive for general use' [33]. In the former comment he is most likely referring to the sound reproduction ability of the then available equipment. One problem he specifically mentions is the 'howling' produced by acoustic feedback from the speaker to the microphone.

A proposal to avoid the often-mentioned problem of acoustic feedback was made by De Costa in 1938 [34]. Here he replaced the speaker — which relayed the heart sounds — with a neon light, and named the device a photostethoscope. This photostethoscope would translate the heart sounds into flashes of light, and, according to De Costa, reproduce 'the rate and regularity of the fetal heart sounds quite as faithfully as the speaker' [34]. It would seem that this device introduced more problems than it solved. The difficulty of relating rapid light flickers to specific heart sounds, differentiating noise from arrhythmiae, and intermittent triggering caused by variable sound intensities, are problems which even its inventor alludes to. There appear to be no further references to the photostethoscope in the literature — yet the problem of acoustic feedback still remained.

Pommerenke reported in 1938 that 'we have been able to obtain sufficient output

from the amplifier, with the fetal [*sic*] heart sounds, to cut satisfactory phonograph discs' [11]. These records were to be used to compile a library of unusual FHS for teaching purposes. Pommerenke carefully records the details of each amplifying stage in his apparatus (figure 2.3) — which is unusual in papers of the time — but neglects to mention the characteristics of the microphone.

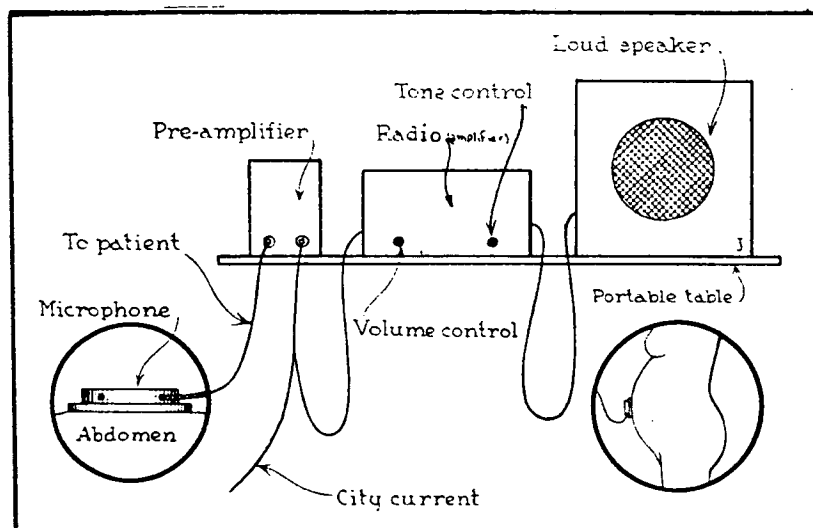


Figure 2.3 Schematic diagram of the apparatus used by Pommerenke [11] to amplify the foetal heart sounds.

Other methods of the period for detecting and recording FHS are presented in Smith [35] and Dressler [36]. It is interesting to note that both workers were making use of the stethograph [37] in which filtering was still being effected acoustically using various bell-shaped attachments to the microphone. The clarity of the sounds in the PhCG suffered as a result. Among these workers, Dressler succeeded in simultaneously recording both the foetal electrocardiogram (ECG) and PhCG. This permitted the first and second heart sounds to be identified in the PhCG by using the ECG as a reference. Previously, the principal heart sounds had been distinguished in the PhCG by the relatively longer duration of the diastolic phase in comparison to the systolic phase. At faster heart rates, however, the systolic and diastolic phases are

approximately of the same duration, and considering the variability in the intensity of the FHS, these two factors combine to hinder the identification of the principal heart sounds using the PhCG alone.

Wartime and post-war developments in sound transduction, recording, and reproduction brought great improvements in all aspects of the phonocardiographic process. This period also saw the beginnings of the automatic processing of heart sounds to extract rate information.

1957 brought the development of the first foetal heart ratemeters by Corner [38] in the USA, and Smith [39] in Britain. Both of these devices reflect the growing sophistication of phonocardiographic monitoring equipment. The foetal heart rate was derived in both cases using the first heart sound to trigger a monostable multivibrator, the output of which was fed through a frequency-to-voltage converter (RC low-pass filter) to a chart recorder. Other similar designs for ratemeters followed: Cox [40] and Maxwell [41] both in 1963, and Cornwall [42] in 1964. These ratemeters all suffered from the same flaw, namely triggering on a fixed voltage threshold, which was based on the assumption that one of the two principal heart sounds always has a greater amplitude than the other. Cornwall goes some way to correcting this using a so-called 'protected' ratemeter. In this design, succeeding pulses from the monostable must be within a preset time of one another to be accepted; if not, the ratemeter continues to register the last 'acceptable' rate until the condition is fulfilled.

It is a common feature of all the preceding papers to make reference to the technical problems which phonocardiography posed. Urbach [43] suggests that the reason for the paucity of literature on foetal heart sounds is because they are 'inordinately difficult to record'. The two problems which feature prominently among these papers are: the weakness of the foetal heart sounds, and the omnipresence and

strength of interfering sounds of whatever origin, and the subsequent problems these pose for the clarity of either the acoustic presentation of the FHS or the phonocardiogram. Hellman praised the phonocardiograph for performing as well as it did with 'satisfactory records obtainable in nearly 50 per cent of patients' [12]. Not all obstetricians were so quick to praise. In a reply to Hellman's paper, a Dr. Charles McLennan retorts:

elaborate tools of various sorts have been used for over 50 years in pursuing fetal heart activity, yet there is today very little evidence that these complex machines can outperform the clinician with a stethoscope. [44]

The most notable advancement came with the invention in 1962 by Hammacher and Gentner [45, 46] of FHS processing circuitry which addressed the problems of extraneous noise in the PhCG and the unpredictable variation in the intensity of the FHS which had always plagued auscultation and phonocardiography alike. Using logic gates to compare the temporal relationship between sounds, the system could determine when heart sounds had gone undetected, and by the same method suppress artefacts. This system formed the basis of that adopted by the Hewlett-Packard Company in their 8020 series of foetal cardiotocographic monitors. Not only did this equipment record instantaneous FHR but also the instantaneous systolic and diastolic times.

Besides the technical problems of signal acquisition, there remained the task of processing the information that the transducer produced. Serr in 1969 says:

Monitoring the fetal heart over prolonged periods, however, may prove impracticable, because of the enormous amount of data obtained, requiring manual interpretation. [47]

Serr suggests that the solution to this problem is in the use of computerized analysis of the PhCG. What follows — quoting Serr's work — appears to be the first account of the digital processing of PhCGs by computer:

A programme [*sic*] reads the converted digital values from disc [*sic*] to memory, counts the number of peaks in a specific time interval, measures the peak-to-peak intervals and prints the result.

In the 1960s foetal phonocardiography fell out of favour because of its inherent problems with interfering sounds; its place was taken by ultrasound cardiography antenatally, and the direct foetal ECG (scalp electrode) during labour. Ultrasound cardiography — which registers the rapid movements of the heart valves — has a superior signal-to-noise ratio in comparison to phonocardiography and is not susceptible to so many interferences. However, concern has always been expressed about the possible biological consequences of irradiating the developing embryo with ultrasound energy — albeit minute [7,9,18]. Phonocardiography has no such drawback; its transduction is performed non-invasively, and with no output of energy into the foetus, hence its potential for long-term continuous monitoring.

This period of neglect of phonocardiography in favour of ultrasound cardiography ended with the renewed interest in the method during the 1980s. Notably, the introduction by Talbert and Southall [1,48] of their 'TAPHO' transducer heralded a great improvement in transducer design with a consequent effect on the quality of the foetal phonocardiographic signal. This is due to the relatively much wider bandwidth of this transducer (0.5-250 Hz) and more efficient sound coupling mechanism [1]. With this new generation of wide-bandwidth transducers, Nagel [49], demonstrated that the majority of the energy contained in the FHS lay outside the previously monitored band of 80-110 Hz, specifically in the 15-60 Hz band:

Conventional narrow-band (typically 80-110 Hz) fetal phonocardiography suppresses the low-frequency components of the acoustic signal; however these contain most of the physiological information. [49]

As the nineteenth century saw the pioneering of obstetric auscultation, so the

twentieth century witnessed the pioneering of foetal phonocardiography. Phonocardiography has the potential to become the technique by which the goal of long-term continuous monitoring of the foetal heart rate will be achieved. This new generation of phono-transducers coupled with sophisticated signal processing will establish the usefulness of monitoring the acoustic output of the heart for the extraction of heart rate and the duration of the systolic phases.

2.3 ORIGIN OF HEART SOUNDS

The exact mechanism by which heart sounds are produced has been a contentious issue. Although the debate has always centred on the origin of adult heart sounds, there is no reason to suppose a different production mechanism for foetal heart sounds. It is important to know the origin of the sounds if any significance is to be attached to the temporal relationships between them and the underlying cardiac events. Two theories have been proposed to account for the sounds: the valvular theory [50], and the cardiohaemic system theory [51]. The valvular theory, which had always been the accepted one up until the mid-1960s, proposes that the sounds heard by auscultation arise directly from the closing of the leaflets in the heart valves as a result of pressure differentials across them. The cardiohaemic system theory, on the other hand, ascribes the sounds to the vibration of the valve leaflets, adjacent cardiac structures, and the blood contained therein, when the blood flow is arrested by valve closure. This latter theory was widely supported in the United States since its inception by Luisada [52] in the mid-1960s. It arose as a result of experimental work done on the simultaneous recording of intracardiac pressures using catheter tip micromanometers, and external phonocardiograms. These recordings showed that the reversal of the pressure gradient across the valves did not coincide with the associated heart sound in the PhCG — which occurred 30-35 ms later. This prompted Luisada to conclude that valve closure, in itself, was not the effect that was registered in the

PhCG. He immediately discounted the valvular theory and proposed that the sounds originated in the vibrations of the cardiohaemic system. Leatham [53] who had always been a proponent of the valvular theory pointed out the fundamental weakness in the cardiohaemic theory which lay in the expectation that a pressure gradient reversal across a valve should coincide directly with valve closure. Leatham's objection was that:

blood flowing through an open valve has momentum and the reversed pressure gradient takes a finite time to halt flow before the valve can close. Gradient crossover thus precedes valve closure by a variable interval influenced by the flow rate and the impedance characteristics of the downstream vessels [50].

To substantiate his argument, Leatham [50, 53] provided evidence using simultaneous echophonocardiography (sonar) and phonocardiography on implanted prosthetic heart valves. This showed that the onset of a heart sound as displayed by the PhCG was coincident with the moment of apposition of the valve leaflets. Since then, the valvular theory has been in the ascendant.

In the following section, before looking in detail at the origin of, and time relationships between the foetal heart sounds, the structure of the heart and the cardiac valves is described.

2.4 FOETAL HEART

The anatomy of the foetal heart is illustrated in figure 2.4. The heart is divided into four chambers of which the upper and lower pairs are called the atria and ventricles, respectively. The atria are connected to the veins which convey blood to the heart, while the ventricles are connected to the arterial network which transports blood away from the heart. When viewed as a mechanical system, these four chambers are actually pumps of which the atria are the primer pumps for the ventricles, which are the power pumps.

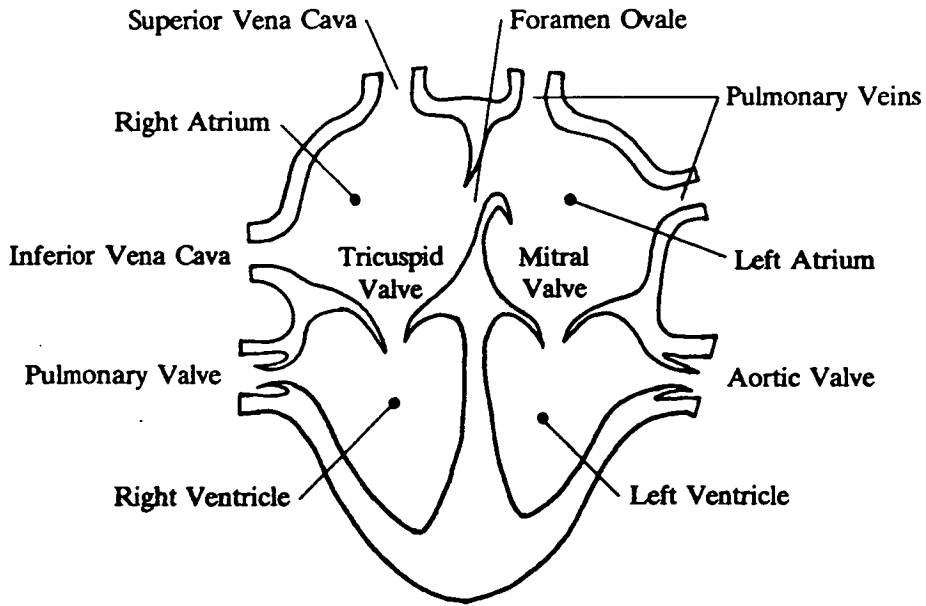


Figure 2.4 Foetal heart (diagrammatic).

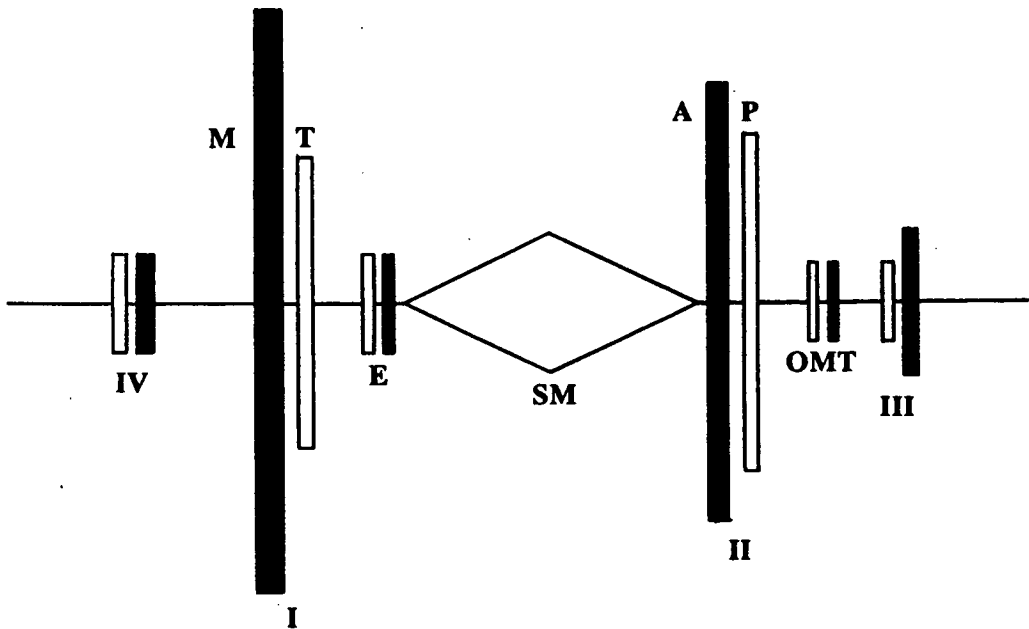


Figure 2.5 Idealized 'normal' phonocardiogram. Events represented by solid vertical bars are associated with the left side of the heart.

I	First heart sound	A	Aortic valve component
II	Second heart sound	P	Pulmonary valve component
III	Third heart sound	E	Ejection sound
IV	Fourth heart sound	SM	Systolic murmur
M	Mitral valve component	OMT	Opening of M/T valves (usually silent)
T	Tricuspid valve component		

One anatomical difference between the foetal and new-born/adult heart is the orifice which shunts the atria together. This orifice, termed the *foramen ovale*, causes most of the oxygenated blood from the placenta which enters the right atrium, to pass directly through into the left atrium and thence, by cardiac action, into the arteries.

2.4.1 Heart Valves

The heart has four valves: the mitral, the tricuspid, the aortic, and the pulmonary (figure 2.4). The larger mitral and tricuspid (M/T) valves are interposed between the atria and the ventricles, while the smaller aortic and pulmonary (A/P) valves are situated at the juncture of the ventricles and the arteries. These valves are each comprised of three cusps — two in the mitral valve — symmetrically arranged around the circumference of the valve orifice and secured at their base by a fibrous ring. Attached to the free periphery of the M/T valve leaflets are fine, strong filaments which exist to prevent eversion of the valve cusps — rather like sails tethered by clew lines. These valves are composed of tough but flexible leaflets of fibrous tissue, whose movements are essentially passive; that is, their openings and closings are effected by the direction of blood flow rather than any voluntary movement on their part. The valves — by virtue of their structure, restraining tendons, and orientation — permit only a unidirectional flow of blood through, and away, from the heart.

2.4.2 Valvular Production of Heart Sounds

The functioning of the heart produces two principal sounds (I, II), and two subsidiary sounds (III, IV) per cardiac cycle (figure 2.5). The origin of the principal heart sounds, as discussed in section 2.3 above, is in the sudden halting — through closure — of valve leaflets which had been moving at high speed [54]. The third and fourth sounds are caused by vibrations of the cardiac walls and are usually inaudible.

Before the onset of cardiac depolarization blood returning to the heart through the great veins, and oxygenated blood from the placenta continually flows into the atria. However, when ventricular evacuation is in progress, the M/T valves are shut which prevents this blood gaining access to the lower cardiac chambers, therefore, blood accumulates in the atria. Once ventricular contraction has ended, the pressure differential across the M/T valves forces them to open, allowing the blood to flow rapidly into the flaccid ventricles. This valvular opening is usually silent as it evolves relatively slowly. At the onset of cardiac depolarization, the atria contract and thereby force the remaining blood in the atria into the ventricles. This atrial contraction is associated with barely audible vibrations arising from the resulting pressure increase in the blood in the ventricles. These vibrations produce the fourth heart sound (IV). At this point the ventricles begin to contract, which causes the blood to attempt to flow back into the lower pressure atrial chambers. This reverse flow of blood is caught and arrested by the shutting of the M/T valves which gives rise to the first heart sound (I). The M/T valve leaflets are tightly sealed by the rising ventricular pressure. This valvular closure approximately marks the beginning of systole (cardiac action phase), or conversely the end of diastole (cardiac relaxation phase). The ventricular walls continue to contract and the pressure in the enclosed blood rises. Whenever the pressure becomes too great for the A/P valves to withstand, they open, and the pressurized blood is rapidly ejected into the arteries. While the ventricles are being evacuated, the pressure in the remaining blood decreases with respect to that in the arteries. This pressure gradient causes the arterial blood to attempt to flow back into the ventricles. This reflux catches the cusps of the A/P valves and causes them to balloon out and seal the orifice thereby abruptly stopping the flow and causing the second heart sound (II). This event approximately marks the end of systole or conversely the beginning of diastole. The third heart sound (III) occurs early in the diastolic phase, and is thought [54] to be the result of vibrations caused by the rapid

filling of the left and right ventricles. This sequence of events, called the cardiac cycle, then repeats.

Besides these sounds, various other sounds/noises are produced during the cardiac cycle, these include: murmurs, blood ejection and entry sounds; and background sounds originating in foetal and maternal movements, abdominal-muscle tremor and environmental noise.

2.5 IDEALIZED PHONOCARDIOGRAM

Shown in figure 2.5 is a diagrammatic representation of the heart sounds produced during one cardiac cycle in an idealized PhCG. In practice the heart sounds would resemble damped oscillations whose end points would not be so clearly delimited in time. Indeed, the variability and variety of these oscillations prevents any illustration of an archetype for a particular sound. The heart sound 'bars' in figure 2.5 are therefore only intended to be position indicators which show the temporal arrangement of the various sounds.

The first heart sound is shown with distinct mitral and tricuspid sound components. This 'splitting' or disassociation of the sound occurs whenever breathing is taking place [54,55] which causes a slight asynchrony between the shutting of the valves on either side of the heart. However, the separation between these components is never very marked, if it is present at all. Normally, both components coalesce to form an apparently homogeneous event. The same explanation applies to the second heart sound, shown with distinct aortic and pulmonary components.

Interspersed between the principal heart sounds are the third and fourth heart sounds, and certain murmurs. These sounds are generally inaudible to the unaided ear but are visible on low-frequency PhCGs.

In practice these acoustic events will be superimposed upon a background of extraneous sounds and noise. These interferences, which can be many times greater in amplitude than the heart sounds, have the effect of obscuring the presence of the heart sounds and generally complicating the analysis of the PhCG. At worst these adventitious sounds may obliterate any trace of heart sounds. The effect of noise on the PhCG is addressed in chapter 3 when actual PhCGs are examined.

The pattern of principal heart sounds in figure 2.5 does not replicate itself exactly nor approximately for each cardiac cycle. Not all the illustrated acoustic events may be present at once and this also applies to the first and second heart sounds. However, assuming that both principal sounds are present for the moment, their onset can be used as a marker for the measurement of the time lapse from one cardiac contraction to the next. The reciprocal of this duration is known as the instantaneous foetal heart rate and is usually expressed in units of beats per minute (bpm). Other parameters such as the systolic time (time separation between a I and the following II), and diastolic time (time separation between a II and the following I) are available for measurement, as are the beat-to-beat differences of these quantities.

2.6 SUMMARY AND CONCLUSION

This chapter has introduced the subject of auscultation and phonocardiography. The new generation of wide-bandwidth phonocardiographic transducers yields heart sound signals with much higher signal-to-noise ratios than were hitherto obtainable. This has aroused renewed interest in this passive and non-invasive method of monitoring the foetal heart which had previously been neglected on account of the apparently low signal-to-noise ratio of the acoustic signal. To fully realize the potential of phonocardiography for long-term continuous monitoring, signal processing and analysis techniques are required to automatically extract the latent parameters of

cardiac function contained in the signal. It is this subject which the following chapters address.

CHAPTER 3

ACQUISITION AND CONDITIONING OF THE FOETAL PHONOGRAPHIC SIGNAL

The function of the signal acquisition stage is to detect and record the acoustic and infrasonic output from the foetal heart. The analogue part of this system consists of an electro-acoustic transducer which converts the physical parameter — displacement — into an electrical signal, and a reel-to-reel FM tape recorder which preserves the form of this transduced signal for later reproduction. As the subsequent processing and analysis is to be performed digitally, the recorded analogue signal must be suitably pre-processed and converted into its equivalent digital representation. The signal conditioning circuitry used to achieve this comprises an anti-aliasing filter, and an analogue-to-digital converter. The final stage of conditioning then filters that part of the spectrum which is occupied by the FHS from the other acoustic and infrasonic phenomena.

Capturing and processing such a physiological signal presents several unique problems. Foremost, is the requirement that the detection be obtained non-invasively, *i.e.* without entering the body. Purely from a measurement point of view, the ideal would be to place the transducer directly onto the foetus but this is undesirable and impossible before the onset of labour, therefore the required information can only be

obtained indirectly. This introduces the second problem which is that the phenomenon to be measured cannot be decoupled from its environment during observation. Thus the transduced signal is influenced by the activity of the surrounding physiological processes both in the foetus and in the mother. Finally, large variations exist in every aspect of the FHS signal across populations and within individuals.

The following sections present the method and instruments of signal acquisition and conditioning. Also, the particular problems posed by this signal both from a visual recognition and automated processing perspective are outlined.

3.1 TRANSDUCER

The function of the transducer is to convert the acoustic and infrasonic signal from the foetal heart into a proportionate electrical signal. The transducer used throughout signal acquisition was the 'TAPHO' transducer (figure 3.1) designed by Talbert and Southall [1, 48, 55]. It was with this transducer that Talbert and Southall introduced wide-bandwidth phonocardiography and the idea of matching the compliance of a transducer to that of the abdominal wall in order to obtain maximum power transfer. Compliance, in this particular situation, is the specific displacement of the abdominal wall resulting from the application of a given force; it is a measure of the flexibility of the skin tissue and its units are metres per Newton. The increased bandwidth of the TAPHO transducer was a significant advancement in the design philosophy of phonocardiographic transducers. It was, as Talbert suggests [1], the narrow bandwidth of previous transducers that had produced the poor quality of PhCG recordings. However, as the sensitivity of aural perception rapidly diminishes below 50 Hz this loss of signal information would have gone unnoticed to the human ear during auscultation [49]. Such aural insensitivity had a direct effect on phonocardiography insofar as previous investigators had always endeavoured to make

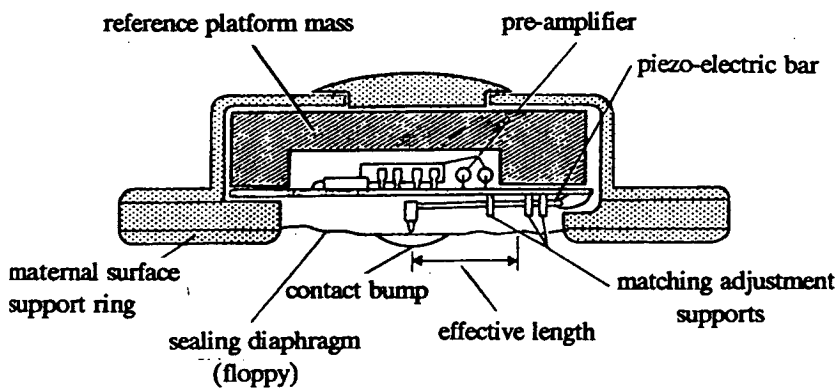
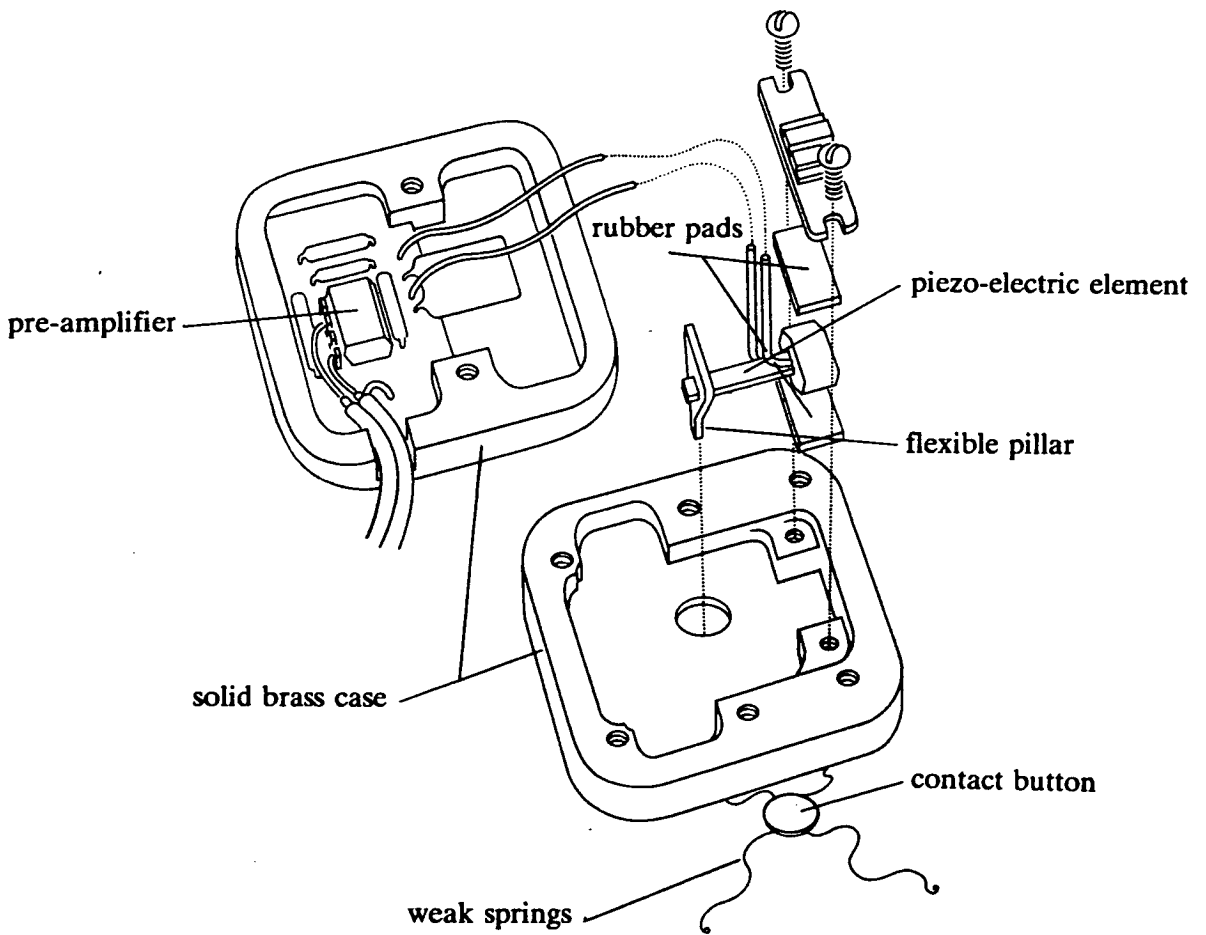


Figure 3.1 Construction of the TAPHO phonographic transducer [58][48].

the transducer produce an output similar to that heard with a stethoscope. This entailed severe attenuation of the low frequency components of the heart sounds.

The electro-acoustic transducer described above has the epithet 'TAPHO' which is an acronym for Total Acoustic PHOnography. The word 'phonography' is a coinage of the inventors [55], and is an adaptation of phonocardiography. This coinage was necessitated by the wider bandwidth of the TAPHO transducer (0.5-250 Hz) as compared to conventional transducers which registered heart sounds only. The TAPHO transducer not only detects heart sounds but can also sense the much larger amplitude and lower frequency infrasonic phenomena such as foetal breathing movements (0.3-1.5 Hz) [1] and foetal movement — although with a reduced sensitivity.

The transduction principle is based on the piezo-electric effect which occurs in certain crystals when they are subjected to mechanical stress. Under such stress these materials polarize which causes the faces of the crystal to become electrically charged. A bar of this material is clamped at one end to the reference mass of the transducer and the free end is connected via a short pillar onto a contact button which is placed against the abdominal wall (figure 3.1). Acoustic and infrasonic waves impinge upon the button causing the bar to bend which in turn produces a charge in the piezo-electric material. The charge is sensed electronically and converted into a proportionate electrical voltage. This is achieved with a voltage pre-amplifier which is situated within the body of the transducer. However, when the nature of the piezo-electric transduction principle is considered, a charge amplifier would have been more appropriate. It was not possible to rectify this short-coming as the recordings of the foetal phonogram had already been made elsewhere (section 3.3).

The compliance of the abdomen is not uniform across gravidae nor during the

period of gestation in the same gravida [56], so the compliance of the transducer must be altered to suit each case. This is achieved by varying the effective length of the piezo-electric bar — which changes its compliance — until optimum response is obtained as judged by the intensity of the foetal heart sounds on an oscilloscope [48].

Physically, the TAPHO transducer is relatively small (38mm × 38mm × 12mm) and lightweight (100g). It is attached to the maternal abdomen using double-sized adhesive tape.

3.2 METHOD OF SIGNAL ACQUISITION

Ambulatory recording of the foetal phonogram was initially attempted by Colley [57] but the amount of movement-sponsored artefact was too great for this to be a viable method of signal acquisition. However, during ambulatory recordings it was observed that the best tracings occurred whenever the subject was resting. To perform long-term continuous FHS monitoring it would have been impracticable to restrict the activity of the subject for long periods during the day, particularly since the recordings were made on low-risk pregnancies. It was decided instead to make overnight recordings in the subject's home despite the inconvenience of installing the recording equipment on that evening and collecting it the following morning [58]. The transducer was fixed onto the site of maximum FHS intensity but was not electrically connected to the recorder — this was done by the subject herself upon retiring that evening. This arrangement provided extended records of the FHS for up to eight hours. The main drawback with this unsupervised recording arose from foetal movement which would displace the optimum recording location for FHS from the initial recording site [57].

With this method of unsupervised signal acquisition, Southall [59] claims that the

FHS, hence FHR, could be visually determined in 65% of subjects at 28 weeks gestation, and 87% at 39 weeks.

3.3 SIGNAL RECORDING

The pre-amplified analogue signal from the TAPHO transducer was fed into one channel of a four channel Racal FM reel-to-reel tape recorder. Using magnetic tapes of 3,600 ft at a tape speed of 15/16 inches per second (ips) allows a maximum recording time of 12.8 hours. When the tape is recording at the above speed, the passband of the recorder is from D.C. to 300 Hz per channel [58] which is more than adequate for the recording of FHS. After recording, each reel of tape was marked with the subject's name, hospital number, and the date of the recording.

The majority of the recordings were made in 1986 as part of research conducted by Nigel Colley [58] towards a doctorate in medicine at University College Hospital, London. His objective was to improve the collection of data from the foetus and to use this data to further the understanding of foetal physiology. No attempt was made to automate the processing of the phonogram — all measurements were made by visual inspection from chart records of the phonocardiogram. The tapes were subsequently transferred to the Electrical Engineering Department of this University for an investigation into automatic methods of processing the signal.

3.4 SIGNAL CONDITIONING

Upon playback at 15/16 ips, the signals were passed through a fourth-order, anti-aliasing, low-pass filter — cut-off frequency 100 Hz — and into an analogue-to-digital converter (ADC) sampling at 500 samples/second (figure 3.2). The anti-aliasing filter was of the Bessel-Thomson type [60,61] which has optimum phase linearity (maximally flat time delay with frequency). Phase linearity is necessary to

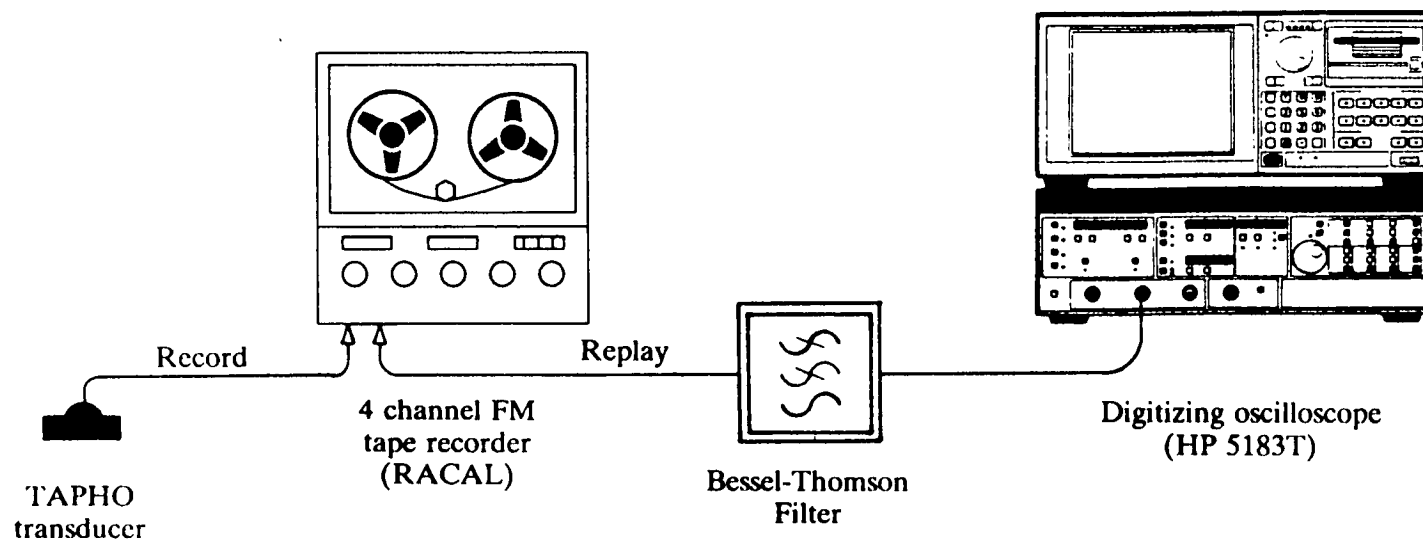


Figure 3.2 Apparatus of signal acquisition and conditioning.

preserve the shape of signal events to allow accurate inter-event timings. The amplitude response of the Bessel-Thomson filter is poor, in fact it is only slightly better than the same order of RC filter. However, there was no necessity for a steeper roll-off as the phonogram possesses little energy above 100 Hz, and the significant frequency content of the FHS is centred around 20-40 Hz [49] with insignificantly small components extending only as far as 100 Hz.

3.5 INSTRUMENTATION

The analogue-to-digital conversion was performed by the Hewlett-Packard HP5180U digitizing oscilloscope. This instrument performs the functions of waveform capture, and analysis. The capture unit is basically an ADC with memory, and the analyser unit is a combined system controller and computation engine, equipped with a display.

Each one of the four channels of the waveform recorder has a software programmable 12 bit ADC which can sample at rates of up to 12 million samples/second. The samples from the ADC are stacked in the channel memory (maximum 16K words per channel) of the digitizing oscilloscope and then transferred en bloc to the memory of the analyser. Here the signal may be displayed directly, or analysed by the built-in signal processing functions. The whole system is controlled either from a touch-key menu on the display or by a resident personal computer (PC).

3.6 DATA SELECTION

The length of a basic data block was standardized at 2048 samples (4.1 seconds at 500 samples/s) for ease of processing and display, and also because this time interval would capture a sufficient number of heart beats. Within the 'normal' FHR range of 120-160 bpm each 2048 sample block should contain between 8 and 11 heart beats.

The sampling rate was set at 500 samples/s — two-and-a-half times oversampled — which is sufficient to provide adequate time-domain resolution.

The quality of the phonocardiogram varies widely within a tape and across tapes (section 3.9), therefore some form of quality control was needed in the selection of suitable portions of data for further processing. The selection criterion used was that the FHS should be discernible by visual inspection and that, overall, the trace should be *moderately* difficult to analyse. This is to avoid those cases when even to a visual inspection the trace is obscure and in the other extreme where the sounds are strong and there is little or no noise.

Data blocks were selected as the tape was replaying. This was achieved using another ADC channel of the recorder. The signal from the anti-aliasing filter would be fed directly into the first channel, and the sampled version of this would be displayed on the screen of the digitizing oscilloscope. Simultaneously, into the second channel would be fed the same signal except that it was high-pass filtered at 40 Hz using a second-order analogue Butterworth filter [60]. The resulting signal was also displayed on the screen[†]. With this 40 Hz high-pass filtered PhCG as a guide, severely noise-contaminated data was rejected. Representative data blocks were then selected from the remaining sections at sites evenly distributed throughout the whole tape. The selected blocks, *i.e.* those not high-pass filtered at 40 Hz, were stored as binary data files on the mini floppy disk of the digitizing oscilloscope. In all, 252 data blocks of 2048 samples each, were captured from 19 subjects.

Performing a visual quality control of waveforms is acceptable at this stage of

[†] Although 40 Hz is not the optimum cut-off frequency to accentuate the FHS (section 3.7) it gives an approximation to the quality of the PhCG. Such a cut-off frequency is the lowest frequency with a second-order filter to sufficiently attenuate the relatively large amplitude low frequency infra-sounds which would otherwise produce a baseline wander on the PhCG.

algorithm development. However, if the system were to be run unsupervised, an automatic quality check would be needed to reject severely contaminated signals. A running measure of energy over a short time interval might prove effective in this respect as artefacts are generally characterized by large amplitudes or persistent signal events.

The selected binary data files were converted into ASCII format in the resident PC controller, and then transferred to a Sun 3/80 workstation where all subsequent processing and analysis was performed.

3.7 FILTERING OF THE PHONOGRAM

The phonogram (section 3.1) contains both the FHS and the intermittently occurring infra-sounds such as foetal breathing movements, and foetal trunk and limb movements. From this composite signal, the phonocardiogram is to be extracted. Fortunately, the infra-sounds do not usually spectrally overlap to a significant extent onto the FHS so that linear filtering may be used to separate the two. Large foetal movements, however, are registered across the whole spectrum occupied by the FHS and severely corrupt the PhCG.

The frequency band which the FHS occupy had always been a source of disagreement among researchers, to which table 3.1 testifies. In the past, only two concerted efforts have been made to determine the frequency characteristics of the FHS: Shelley in 1969 [62] and Nagel in 1986 [49]. Only the more recent work by Nagel could benefit from the introduction of wide-bandwidth phono-transducers. Previous sensors were of the narrow-bandwidth type which produced phonocardiograms with low signal-to-noise ratio and poor spectral fidelity with respect to the FHS. Combined with this there was the perceived need to have the circuitry

Principal Researcher	Year	Filter Spec.
Frederick [30]	1924	lpf 130 Hz
Sampson [24]	1926	lpf 130 Hz
Pommerenke [11]	1938	bpf 30-10,000 Hz
Lockhart [37]	1938	bpf 75-550 Hz
Steer [63]	1951	bpf 35-65 Hz
Smith [39]	1957	bpf 60-80 Hz
Sawyer [64]	1959	bpf 30-100 Hz
Alment [65]	1962	bpf 80 Hz
Leonard [66]	1963	bpf 60-80 Hz
Cornwall [42]	1964	bpf 60-80 Hz
Weill [67]	1964	bpf 70-140 Hz
Gentner [45]	1967	bpf 80-110 Hz
Urbach [43]	1968	bpf 130-200 Hz
Shelley [62]	1969	bpf 60-120 Hz
Melchior [68]	1969	bpf 70-85 Hz
Hewlett-Packard [48]	1970	bpf 80-110 Hz
Luisada [52]	1972	bpf 50-100 Hz
Goodlin [18]	1979	bpf 75-105 Hz
Talbert [1]	1986	hpf 50 Hz
Colley [55]	1986	hpf 40 Hz
Dripps [69]	1986	lpf 100 Hz
Holburn [70]	1989	bpf 40-125 Hz
Bassil [71]	1989	bpf 40-65 Hz

Table 3.1 Filter type and filtering threshold frequency used by various researchers to obtain the foetal phonocardiogram. This table does not include those systems which used acoustical filtering. Only Urbach [43] reports the frequency response of the transducer used. lpf: low-pass filter; bpf: band-pass filter; hpf: high-pass filter.

produce sounds that were like those heard on auscultation. These factors provide an explanation for the use of relatively higher frequency narrowband filtering (60-120 Hz) which most of the previous investigators employed (table 3.1). Undoubtedly, the resulting severe attenuation of the FHS would account for the poor reputation of phonocardiography in heart rate monitoring [16].

Although the inventors of the TAPHO transducer have used a second-order analogue high-pass filter with cut-off frequency between 40 and 50 Hz depending on which published source is referenced [1, 48, 55, 57], this was mainly for the reasons outlined in the preceding footnote. Using higher order high-pass digital filters with lower cut-off frequencies, it became apparent that the clarity of the FHS improved as the filter threshold frequency was lowered. After having filtered many phonograms from various subjects at different cut-off frequencies, it was the author's experience that the best signal-to-noise ratios were obtained, to a visual inspection, with a cut-off set at 20 Hz. Figure 3.3 illustrates the improvement achieved with this revised filter threshold for the case of an arbitrarily chosen phonogram. That the majority of the energy in the FHS lies in the previously unmonitored band from 20-40 Hz is quantitatively confirmed by the results of chapter 4, and by the work of Nagel [49].

The digital filter which was used to extract the PhCG from the phonogram is of the transversal type as there is a requirement to preserve the relative phasing of the spectral components. This filter was designed by the window method [72] and employed a Hamming window; it comprises forty-one tap weights.

3.8 DOMAIN PROBLEMS

In contrast to adult phonocardiography where a strong sound generator is close to the transducer, foetal phonocardiography has to contend with a weak sound

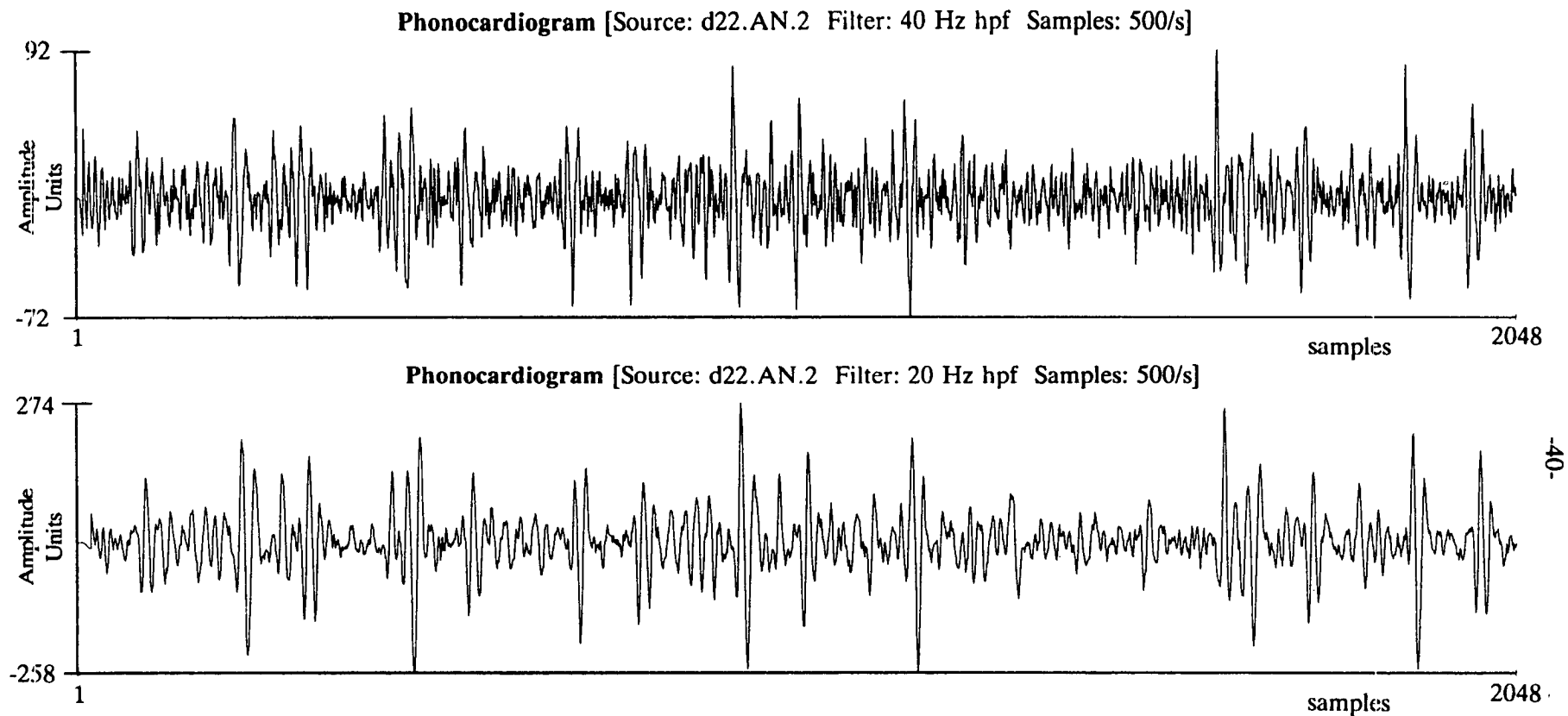


Figure 3.3 Both phonocardiograms have originated from the same phonogram and differ only in the filtering threshold of the high-pass filter used. The improvement in signal-to-noise ratio with a 20 Hz threshold is apparent.

generator separated from the sensor by up to ten heart diameters.

Whenever the closing of the cardiac valves creates a sound, this energy packet must travel through a complex and dynamic system up to the maternal abdominal surface. At the outset the intensity of a principal sound depends on factors such as: pressure differential across the valve, size of the valve leaflets, the initial extent of separation of the leaflets, the volume of blood, and the modulating effect on blood volume caused by foetal breathing [55]. All of these factors, except leaflet size, could be expected to change from beat to beat. The inherent variability of the foetal heart rate — the irregularity of which is an indication of normality — and its wide range (60-220 bpm [4, 73, 74]) originates in the variability of the durations of the systolic and diastolic times. This again is reflected in the instants of closure of the valves.

The sound transmission path is made up of the foetal chest and/or abdominal wall, the fluid-filled amniotic cavity, the muscular wall of the uterus, layers of fat, tissue, and possibly bony and cartilaginous material. Each of these substances will attenuate the forward-going sound energy both in the bulk of the material, and by reflection arising from the impedance mismatch which occurs at the boundary of each of these layers. This natural screen surrounding the foetus has the effect of diminishing the energy reaching the phonographic transducer. In addition to this, there will be extraneous sounds produced by the umbilical cord, foetal movements, and environmental noise which will superimpose themselves onto the acoustic signal from the heart. Finally, at the maternal abdominal surface, the position of the transducer will also considerably alter the shape of the sound waveform [7, 36].

This complex system would be expected to modify every characteristic of the original signal. Not only will this system be varying on a short-term basis, but also in the longer term as the gestation progresses and the foetus grows in size.

Overall these factors combine to produce 'heart sounds' which can change markedly from one beat to the next in amplitude, duration, frequency, morphology, and in time of occurrence relative to the previous beat, and even to the complete attenuation of a sound. Along with this are the other sounds/noises produced by the mechanical functioning of the heart: blood circulation sounds, ejection sounds, entry sounds, and eddies; these all contribute to further complicate the already difficult task of identifying the principal cardiac sounds. This variability in the characteristics of the principal heart sounds precludes compiling a catalogue of types for each sound either for all gravae, or for a particular grava even over a short time-interval.

Figures 3.4 to 3.7 illustrate representative examples of some of the more frequently occurring of these anomalies in the phonocardiogram. Each data segment has been extracted from a longer PhCG and represents approximately one-third of the total length. The identity of heart sounds, the duration of systole and diastole, and the FHR are first established in the less ambiguous sections surrounding the segment of PhCG; this information is then used to extrapolate the analysis into the adjacent anomalous regions. Not all the data segments are from the same subject, nor are they of the same duration. The vagaries illustrated are not restricted to any particular subject but are manifest across all the subjects examined.

Figure 3.4a clearly shows the first and second heart sounds of three beats against a background of very low noise. Although this is not to imply that the principal sounds will always be so pronounced when there is no noise — it may happen that either of the sounds, but most often the second sound, disappears even when there is no noise contamination. Figure 3.4a also demonstrates a common feature of the principal heart sounds where a first sound is of longer duration than a second and contains more cycles. In this instance the second sounds exhibit an instantaneous

frequency which tends to fall with time. It will also be noticed that the time interval between a first sound and the following second (systole) is shorter than between a second sound and the following first (diastole).

Figure 3.4b was captured from a different subject than figure 3.4a and although the FHR is the same and background noise is negligible, there is, however, a marked difference between the two in the structure of the FHS. The frequency of the FHS in figure 3.4b is higher, and there are more cycles per sound. The second sound, in the latter case, does not exhibit a decrease in instantaneous frequency across its duration. Figures 3.4a and 3.4b are exemplary PhCGs where the well-defined end points of the FHS, the stability of the structure within the sounds, the very low noise, and the absence of interfering sounds are rarely seen in combination.

If figures 3.4a and 3.4c are compared, the similarity between the second sounds in figure 3.4a and the first sounds in figure 3.4c will be noticed. This poses an obvious problem for any detection algorithm based on a template matching scheme. In figure 3.4c the second sound has very low amplitude compared to the first sound, and its form is not dissimilar. Only the low level of noise prevents the second sounds from being inextricably absorbed into the background.

Figure 3.5a shows a higher level of noise than was evident in figures 3.4a, 3.4b and 3.4c. Here again is found the non-uniqueness of sound structure where the first sounds resemble the second sounds of figure 3.4a. The presence of adventitious sounds occurring in the systolic phase (systolic sounds) of the first and third beats has obliterated the second heart sounds, although the second sound is quite plain in the second beat when the systolic sound is absent. Preceding the beginning of the first sounds can be seen the emergence of adventitious sounds in the diastolic phase (diastolic sounds) which obscure the exact start of the first sounds.

Figure 3.5b illustrates occurrences of both systolic and diastolic sounds, with the diastolic sound stretching the full length of the diastolic phase in the first beat. The first and second sounds of the first beat are replicas of one another albeit for a phase reversal. Likewise the second pair of principal heart sounds which closely resemble one another except for an amplitude scaling factor.

Figure 3.5c shows the sudden appearance of a systolic sound in the second beat which completely absorbs the second sound. There is no trace of this systolic sound in the beats on either side, where the second sound is clearly visible.

Suddenly appearing and disappearing phenomena are common, as is complete attenuation of principal sounds — particularly the second sound. Figures 3.6a and 3.6b illustrate this latter point. The first beat of figure 3.6a has a strong second sound which is severely attenuated in the following beat. Figure 3.6b shows a sustained attenuation of the second sound across three beats although there are remnants of the second sound in the first beat.

Diastolic sounds coalescing with first sounds (figure 3.6a) are a problem for two reasons: firstly, they obscure the beginnings of the first sound, and secondly if one occurs very close to a first sound, the diastolic sound and the first sound coalesce. This gives the appearance of a homogeneous entity which could easily be mistaken for a long first sound. It is difficult to detect this occurrence, and the problem is exacerbated by the persistence with which this anomaly occurs. Later in the PhCG (not shown) from which figure 3.6a is extracted, the diastolic sound disappears completely to reveal the true nature of the first sound. Another way this anomaly reveals itself is in the relative durations of systole and diastole. If the first sound plus diastolic sound is taken as an entity, this often makes the apparent duration of systole longer than diastole — which is physiologically impossible as diastole must always be

longer than, or just equal to, systole [75].

Figure 3.6c shows impulsive first and second sounds which are indistinguishable. Even the relative separation of repeating sounds gives no indication as to which phase is systole or diastole as they are both of the same duration. The principal sounds in this segment also contradict the generalization that first sounds are longer and contain more cycles than second sounds although this generalization holds in the vast majority of cases.

The attenuation of a first sound, although a rare occurrence, does happen (figure 3.7a). In the first beat, it is difficult to find the exact position of the second sound because of the level of systolic and diastolic sounds. The late diastolic sound attached to the third, first sound is not evidenced at all in the first sound of the first beat.

Figures 3.7b and 3.7c demonstrate those occasions when the analysis of the PhCG is very difficult (figure 3.7b) and completely impossible (figure 3.7c). The best that can be done on encountering a segment such as figure 3.7c is that the detection algorithm would completely neglect it. Figure 3.7b is just analysable but only after working back from clear sections of the PhCG once the pattern of events has been established.

3.9 VISUAL IDENTIFICATION OF HEART SOUNDS

Previous investigators [24, 35, 41] have sought to discover visual characteristics of particular heart sounds with which they could distinguish the acoustic events in the PhCG. Their conclusions were often contradictory [41, 75]. Others identified the FHS not from their appearance but rather from the longer duration of diastole compared to systole. However, it has become apparent from the more recent work in the area [45, 64] that the characteristic of the FHS themselves, is that there is no

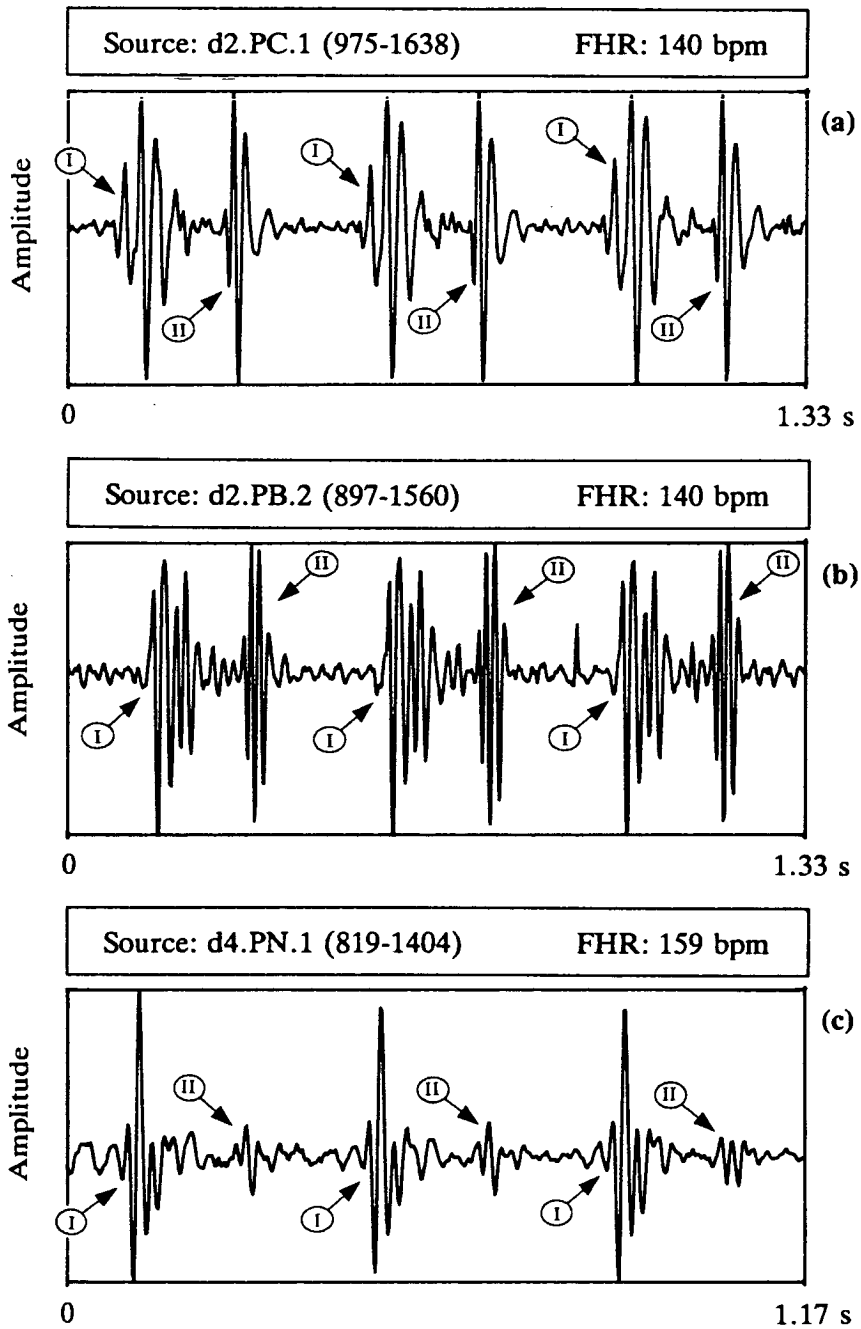


Figure 3.4 (a) and (b) Exemplary phonocardiogram segments; (c) Anomalous phonocardiogram segment exhibiting weak second heart sounds.

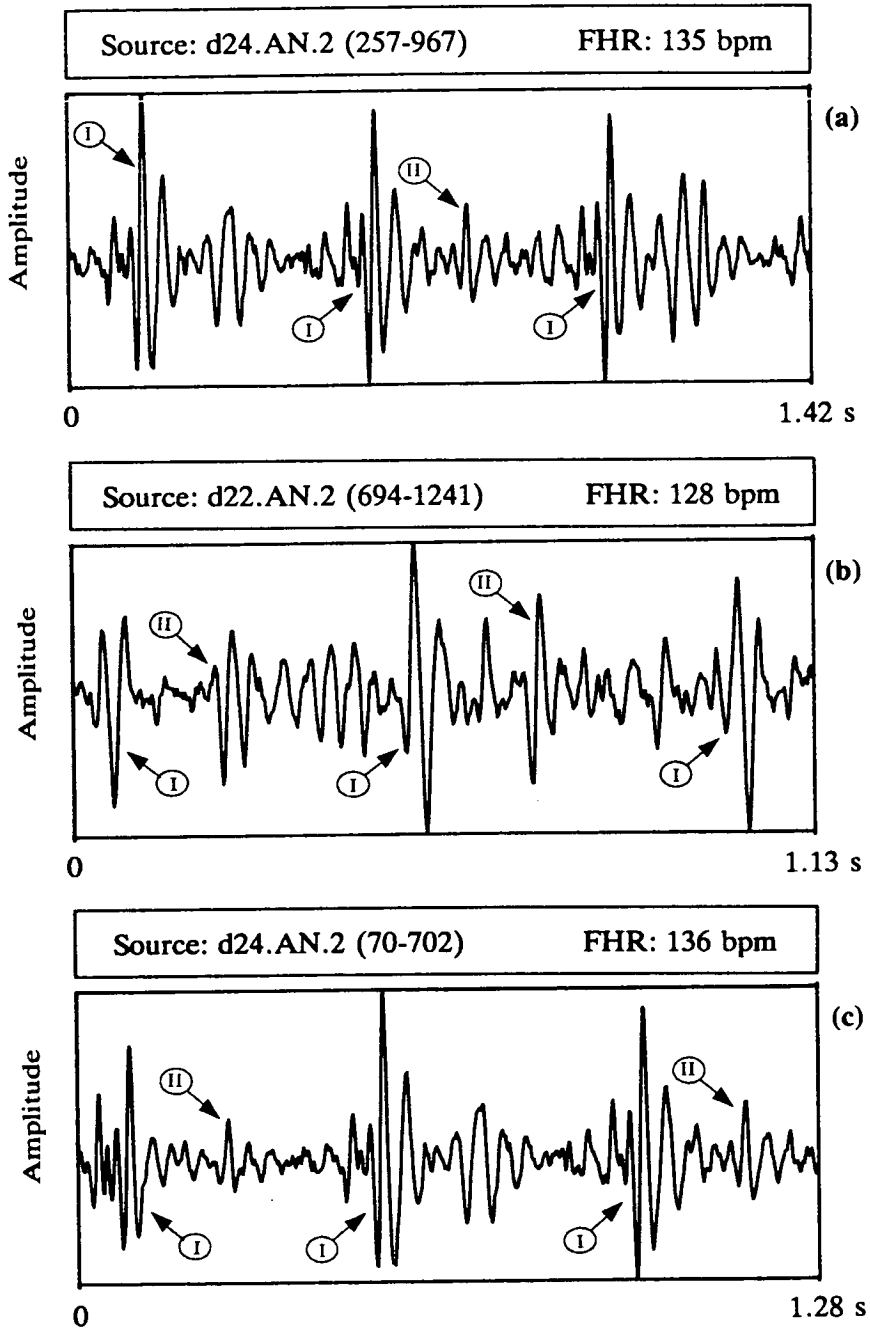


Figure 3.5 Anomalies in the phonocardiogram. (a) large systolic sounds; (b) similarity of first and second heart sounds, long diastolic sound; (c) sudden appearance of systolic sound obliterating second heart sound.

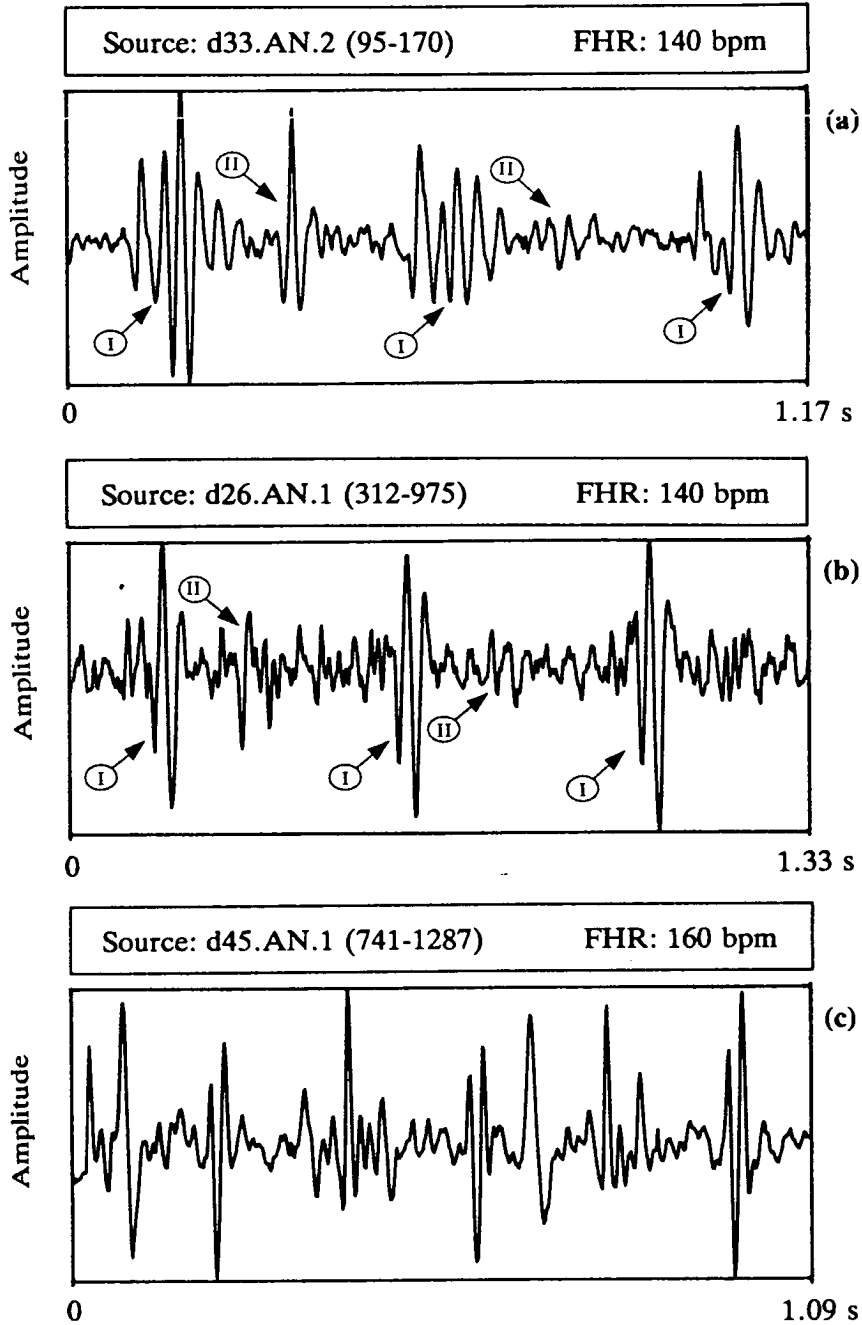


Figure 3.6 Anomalies in the phonocardiogram. (a) severe attenuation of a second heart sound, late diastolic sounds present; (b) relatively short first heart sounds, complete attenuation of second heart sound; (c) impulsive first and second heart sounds which are indistinguishable, large adventitious sound.

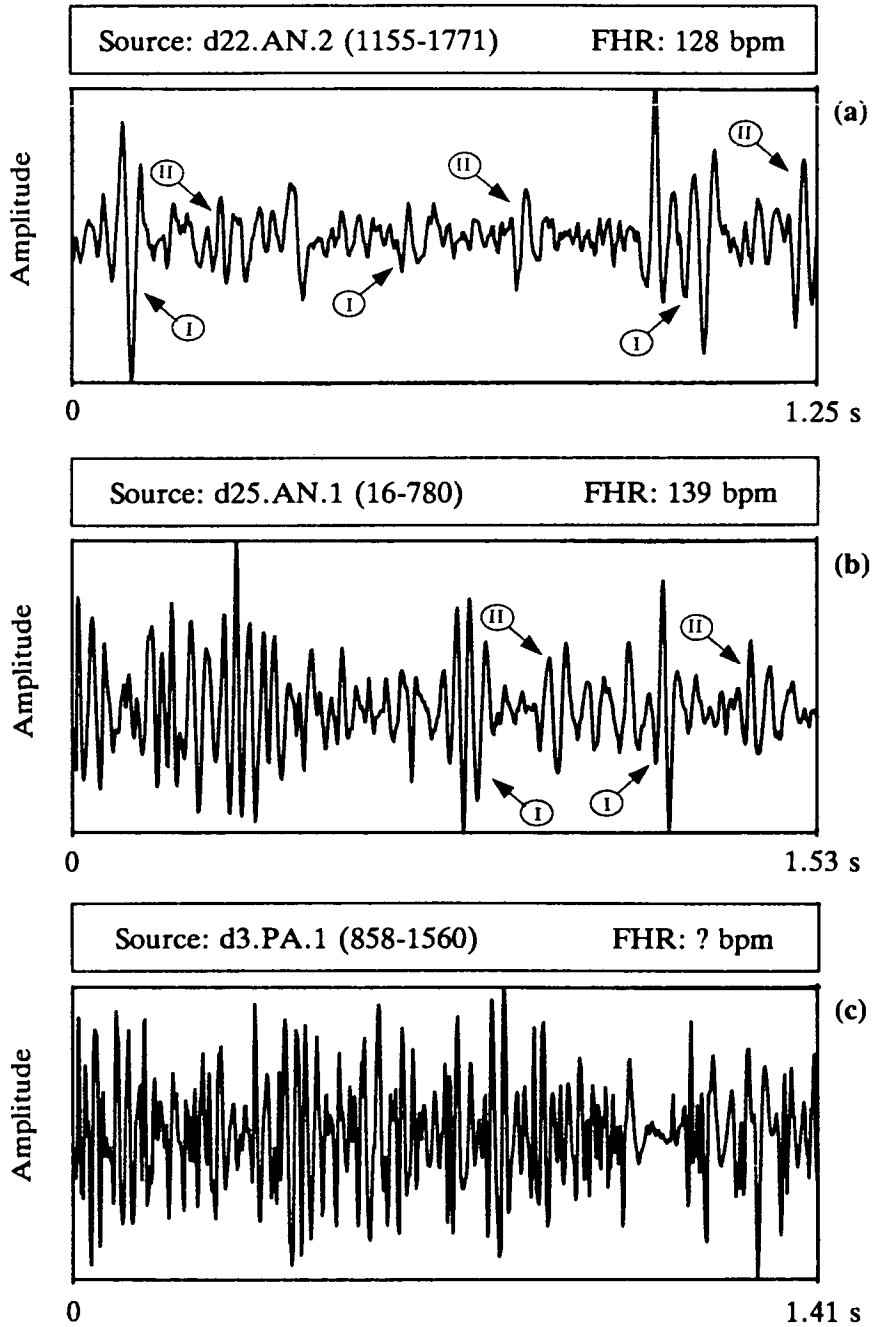


Figure 3.7 Anomalies in the phonocardiogram. (a) severe attenuation of a first heart sound, large diastolic sounds; (b) large and continuous systolic and diastolic sounds; (c) all trace of the principal heart sounds completely obscured by noise.

general characteristic. From the PhCGs examined in this study no unique characteristic was observed which could be associated with any particular heart sound. However, the first sound is generally longer than the second and contains more cycles. The second sound has a more regular shape and bears a similarity to two-and-a-half cycles of a sine wave, amplitude modulated by a bell-shaped pulse. No heart sound was consistently greater in amplitude than the other, although the second heart sound exhibits the greater amplitude variability. This phenomenon may well be due to some physiological activity such as foetal breathing rather than attenuation in the path to the transducer. Certainly from the results published by Colley [55] there is a concomitant rise in the background noise in the PhCG whenever foetal breathing is present and an attenuation of second heart sounds.

Without specifically addressing the problem of how the automatic identification of the FHS is to be achieved, it is obvious from the preceding section that it is a very difficult task. Even with the TAPHO transducer and its reputed ability to yield a PhCG with a high signal-to-noise ratio, the level of noise and artefacts pose many challenges to any form of automatic processing and analysis. Despite this difficulty, a visual inspection can locate and label the principal heart sounds from among many competing events in even the noisiest PhCG. How does the eye/brain combination achieve this? Certainly it fixes upon certain salient features: the higher amplitudes, event homogeneity, event/background dissimilarity and structure, and particularly the repetition of the principal heart sounds across the PhCG. No amount of introspection will reveal more than this. However, there are other sources of information brought into the process of identification of the FHS within the PhCG besides what is immediately visually apparent in the trace. This additional information includes knowledge of the relative durations of systole and diastole, the extent of the FHR, and the accumulated experience derived from examining many PhCGs and observing their

variability. The combination of sight and experience is the basis of the facility of the eye and brain in analysing the PhCG.

This intuitive facility is difficult to grasp at a conscious level — to which both vision and speech analysis research testify. Moreover, the comparative ease with which the task is accomplished might lead one to think that the whole operation is trivial; rather it is a testament to the acuity of the cognitive process.

In all the following chapters the interpretation of the PhCG arrived at through visual inspection is taken to be the standard which all attempts at automatic processing and analysis should strive to attain.

3.10 SUMMARY

This chapter describes the acquisition and conditioning of the acoustic signal from the foetal heart, and the processing of this signal to obtain the phonocardiogram. The indirect way by which the signal is obtained exposes it to contamination by its surroundings in its path from source to transducer. Although the particular electro-acoustic transducer used, may faithfully transform all the signal reaching the maternal abdomen, that particular aspect of the operation of the heart which is being recorded does not consistently produce an unambiguous signal. The variability in all aspects of both the heart sounds and the background events has been outlined and it is seen that this variability makes the objective of an automatic analysis of the PhCG difficult to achieve.



CHAPTER 4

FREQUENCY ANALYSIS OF FOETAL HEART SOUNDS

The overwhelming majority of the research undertaken in the area of heart sound analysis is directed towards the non-invasive detection of valve dysfunction, *i.e.* impairment or abnormality, in implanted prosthetic valves in the adult heart. The main cause of failure in artificial valves is the deposition of calcium on the leaflets which leads to valve stiffening and stenosis [76]. It has been observed [77] that with increased calcification of the leaflets, the acoustic signature of the valve on shutting has a characteristic shift in the frequency spectrum away from a defined 'normal' position towards the higher end of the spectrum. The analysis of adult heart sounds is, therefore, concerned with the spectral signature of particular valves during the various stages of their degeneration.

The most significant paper to date which specifically addresses the spectral analysis of foetal heart sounds was published by Nagel [49] in 1986. His work was the first in this field to use wide bandwidth electro-acoustic transducers. Previous results by other researchers are rendered suspect or invalidated through the use of narrow bandwidth transducers which significantly attenuate the received sounds (chapter 3, section 3.1) [1, 49]. It was Nagel's intention to characterize 'normal' foetal heart sound spectra to allow diagnostic information to be derived from the spectra in very

much the same way as the adult case. In consideration of the extent of the sample set of subjects (30-50) investigated, his contribution to the field would have been even more significant had he described the method with which the transformation to the frequency domain was achieved.

There are obvious similarities between adult and foetal phonocardiography. However, in the case of foetal heart sounds the situation is rather more complicated. Firstly, the foetal heart is a much weaker sound generator than the adult heart and it is further separated from the closest external point which would allow non-invasive recording. Both of these factors have a detrimental effect on the signal-to-noise ratio of the received foetal heart sounds. Secondly, the foetal heart is continually growing and enlarging such that the transducer reception of the foetal heart sounds improves with gestational age. However, as the whole foetus grows and with it the uterus, so the dynamics of the acoustic transmission path — which modify the detected sound — and the valves themselves cannot be assumed to remain constant for an individual and certainly not across a population. Foetal growth is also reflected in the progressively lower shift of the frequency content of foetal heart sounds with gestational age [49]. This is in contrast to the adult case where these effects would be constant and the nature of the transmission path would not be expected to vary widely within and across individuals. These factors coupled with the susceptibility of the foetal phonocardiogram to contamination by noise make it difficult to obtain characteristic signatures of the foetal heart sounds, either in time or frequency, for single subjects or populations.

Although the title of this chapter may suggest that the frequency analysis of foetal heart sounds was an end in itself, this is not the case. The sounds were analysed with a view to extracting parameters of significance from the spectrum which would

uniquely characterize particular heart sounds. It was envisaged that such parameters would be: frequency, bandwidth of the dominant peaks and relative intensity of the frequency components. The strategy was then to use this information to locate and classify the sounds by template matching in any record of the phonocardiogram thereby enabling the measures of cardiac function to be obtained. However, in examining the content and spread of the frequency components of foetal heart sounds, it was not the intention to complete a survey of the spectral changes in foetal heart sounds through pregnancy, nor across subjects of the same gestational age. Rather the objective was to investigate on a per record basis:

- (a) *if the principal heart sounds could be distinguished by their respective spectra* — this would form the first step towards an automatic recognition system. The rationale behind this thesis was threefold. Firstly, when the time-domain morphology of foetal heart sounds is considered, there appears to be no particular feature or features which distinguish one group from another. An obvious extension of the search for identifying attributes is in the frequency domain. Secondly, several auditors of foetal heart sounds have observed that the second heart sound ‘appears’ to be of higher frequency than the first heart sound [35,43]. It must be borne in mind that this might well be an artefact arising from the response of the human auditory system and/or the apparatus used in the auscultation. Thirdly, it was hoped that the differences in anatomical structure between the two sets of valves (chapter 2, section 2.4.1) and the different stresses to which each set is subjected might have been registered in their respective spectra.
- (b) *if and how the spectra of the respective heart sounds change through the course of a record.* That spectra change across subjects of the same gestational age or in

the same subject through gestation was not an issue. The main concern was the nature of the spectra solely on a per record basis as the original eight-hour phonocardiogram was to be suitably segmented for processing and analysis. Although the sound generators and the transmission path represent a complex and variable system, it might be assumed that on a per-record basis the dynamics would not change significantly. This would then allow a characterization of the spectra at least on a per record basis. However, this makes the prior assumption that a unique spectrum is produced by each group of principal heart sounds as hypothesized by (a) above.

- (c) *the similarity of the spectra of the adventitious sounds particularly third and fourth sounds, with those of valve closing sounds.* The implication here is that if the spectra are similar, then frequency information in itself would not be sufficient to distinguish between them.
- (d) *techniques for the transformation of the extracted foetal heart sounds to the frequency domain.*

The following sections describe firstly the procedure adopted for the isolation of foetal heart sounds from the phonocardiogram, and then two techniques, namely, the Fourier transform and a parametric method for the spectral analysis of these sounds. The derived spectra are then examined in accordance with the conditions outlined in (a)-(c) above to ascertain whether the sounds may be distinguished using a frequency spectrum representation of their information. Before this, however, a brief description is given of the methods used by previous investigators of adult heart sounds.

4.1 ADULT HEART SOUND ANALYSIS

It is the primary aim of adult heart sound analysis, especially in the case of

implanted bioprosthetic valves, to obtain a set of diagnostic parameters which would reveal abnormal functioning of the valve. It is in the frequency domain that these parameters have been sought.

All previous investigators with one exception separated the heart sounds from the other acoustic phenomena by visual inspection, and manually distinguished between the two sounds; no attempt was made to perform this operation automatically. This procedure was a convenience and a necessary first step to obtaining the spectral characteristics of the sounds. However, the automatic isolation of the principle heart sounds from the phonocardiogram is, in itself, a major task. Iwata [78] — the one exception — took advantage of the temporal relationship between the ECG and the PhCG to isolate fixed-length segments of the PhCG. Within these segments he then sought a dominant peak in the frequency domain to indicate the position of a sound.

Until the 1970s few studies were conducted into the frequency content of heart sounds. The methods that were applied were approximate and usually involved sweeping the phonocardiogram with bandpass filters [79-81].

Yoganathan in 1976 [82-84] was the first to apply the Fourier transform to the analysis of heart sounds. He investigated both first and second heart sounds of healthy subjects to determine the 'normal' spectrum. His results show, for both principal heart sounds, energy in the band ranging from 10 Hz to 400 Hz, with the majority concentrated in the 10-80 Hz band. The averaged spectra of first and second sounds differed only in the presence of a peak in the first heart sound spectrum around 30 Hz.

In the 1980s there was a move away from the non-parametric modelling of heart sounds to parametric procedures such as auto-regressive (AR) [85] and auto-regressive

moving average (ARMA) [86] modelling. This move was precipitated by the coarse spectral resolution of the FFT when analysing such a short duration phenomenon to obtain spectral parameters [76,87]. These spectral parameters were either the two dominant frequencies in the spectrum [86] or these frequencies plus various measures of energy distribution [76,77]. The features were chosen to reflect the clinical observation of a shift to higher frequency associated with valve dysfunction[†]. The extracted spectral parameters would then be formed into a feature vector for input into an automatic classification system.

4.2 HEART SOUND SELECTION

In the absence of an automatic technique, individual heart sounds were manually extracted from phonocardiograms. These phonocardiograms were chosen on the basis that the valvular sounds they contained were clearly defined and that there was little interference from adventitious sounds and noise. This was done to ensure that the spectra of the sounds would be as uncontaminated as possible. A problem facing any extraction technique is to estimate the extent of a sound in time. In all the literature on the subject, whether concerning the adult or the foetal case, no reference is made to the criteria by which the principal sounds are isolated from the background. This problem which is more fully addressed later, becomes all the more acute whenever automatic extraction of sounds is attempted especially when those sounds are not clearly defined. For consistency, the following criterion was adopted throughout the manual extraction: the end points of sounds are taken to be those points on either side of the valvular sound where the extremities of the sound became indistinguishable from the background noise level. First and second heart sounds were identified on the basis of the relative durations of the systolic and diastolic times (chapter 3, section

[†]The extent of the shift from the 'normal' spectrum is dependent on the type of prosthetic valve used [77].

3.9). In cases where there was any doubt regarding the integrity and/or homogeneity of a sound, *e.g.* a possible coalescence between a fourth sound and a first sound, that particular valvular sound was disregarded. The particular set of sounds used in the following experiments were extracted from arbitrary sites spaced through the length of recordings from twelve foeti of varying gestational age; in all, 404 first sounds and 337 second sounds were analysed.

4.3 DISCRETE FOURIER TRANSFORM

The *Discrete Fourier Transform* (DFT) is an operation for transforming a sequence $x(n)$ of length N into a sequence of frequency samples $X(k)$ of length N ; it is defined to be:

$$X(k) = \sum_{n=0}^{N-1} x(n) \exp(-j \frac{2\pi kn}{N}) \quad k = 0, 1, 2, \dots, N-1 \quad (4.1)$$

The *Fast Fourier Transform* (FFT) [88] is an efficient algorithm for the evaluation of (4.1). It may be shown [88] that the computation time for an N sample transform is approximately proportional to $N \log_2 N$ when the FFT is employed, in contrast to N^2 when the elimination of redundant calculations is not exploited.

4.4 LIMITATIONS OF DFT PROCESSING

The following sections consider the limitations associated with the use of the DFT for the transformation of a signal event. These limitations mostly arise from the finite nature of the signal event and, in this application, its relatively short duration. The steps commonly employed to alleviate these shortcomings are also described.

4.4.1 Aliasing

Aliasing is caused by the undersampling of a continuous signal in time and

produces an overlap of spectra in the frequency domain with the result that higher frequencies appear as lower frequencies. Potentially, aliasing could be a problem were it not for the steps taken to eliminate the possibility of its occurrence (chapter 3, section 3.4). These precautions were to interpose an analogue low-pass filter between signal source and digitizer to bandlimit the signal and to have a sufficiently high sampling rate.

4.4.2 Windowing

Whenever a heart sound is extracted from the phonocardiogram, an implicit windowing operation is performed whereby samples lying on either side of the sound are ignored. Mathematically, this operation may be represented as multiplication of the PhCG by a rectangular pulse centred on the sound which has unit amplitude for the duration of the sound and zero amplitude before and after the sound. Expressing this symbolically: for an observed heart sound sequence $x_0(n)$ of N points:

$$x_0(n) = x(n) \text{rect}(n) \quad (4.2)$$

where $x(n)$ is an infinite PhCG sequence, and $\text{rect}(n)$ is:

$$\text{rect}(n) = \begin{cases} 1 & 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases} \quad (4.3)$$

A rectangular window in the time domain is equivalently represented in the frequency domain by a complex sinc function (figure 4.1a and 4.1b). Thus the time-domain multiplication in (4.2) corresponds to a frequency-domain convolution of the Fourier transforms of the respective functions. The result of this operation is that the observed transform of a heart sound will be the convolution of the spectrum of the heart sound with the complex sinc function giving the composite spectrum:

$$X_0(f) = X(f) * D_n(f) \quad (4.4)$$

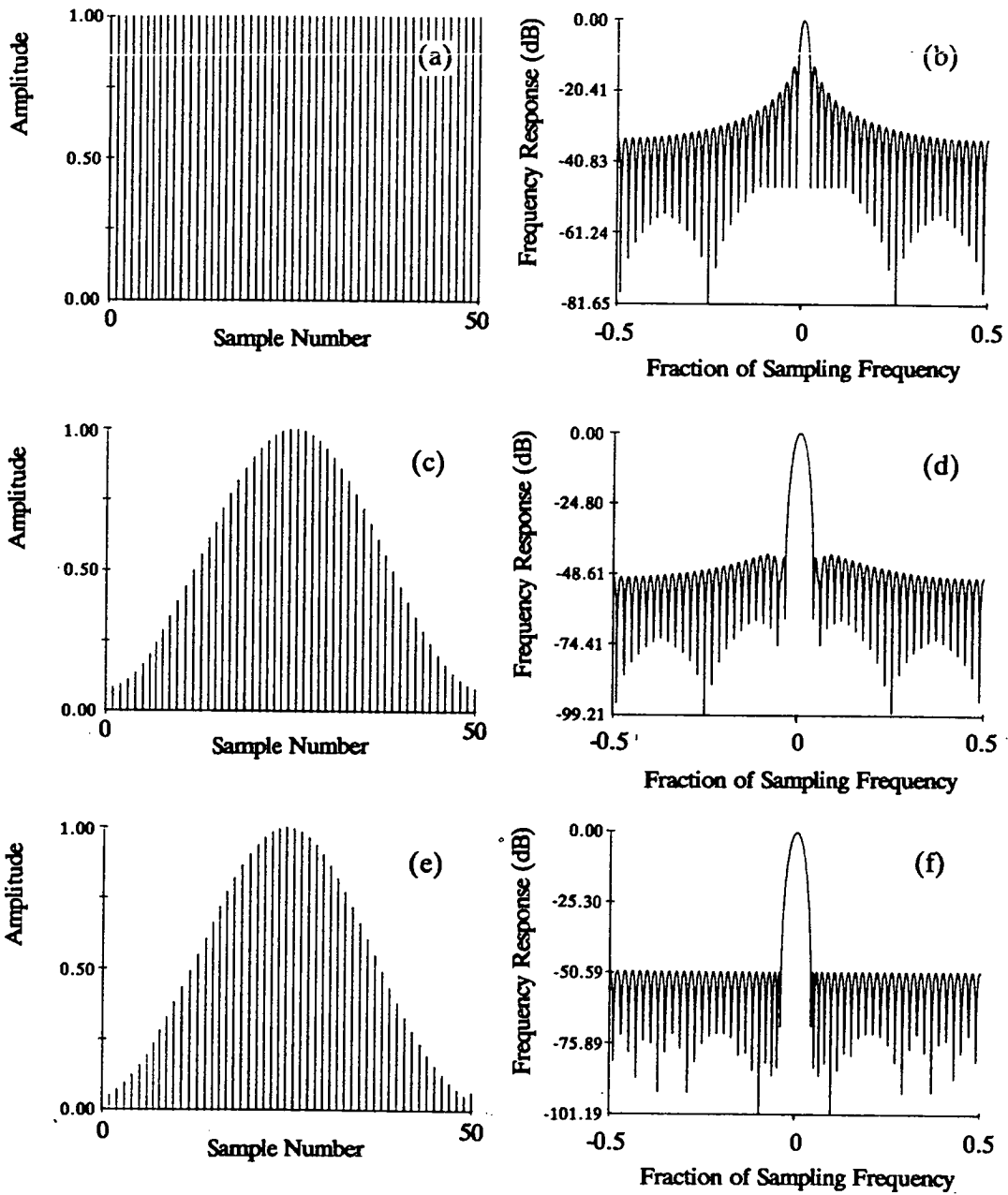


Figure 4.1 Time and frequency representations of: the rectangular window (a,b); the Hamming window (c,d); the Dolph-Chebyshev window with -50 dB sidelobes (e,f).

where convolution is denoted by “*” and $D_n(f)$ is the transform of the rectangle function known as the digital sinc function. Due to the side lobes of the window function, the convolution in frequency spreads or ‘leaks’ the energy from the heart sound across many bands. This energy might otherwise have been concentrated at a single point. Such an effect, which is called *leakage*, has a further consequence where a small-amplitude frequency component may be obliterated by leakage from a stronger component, particularly if the small component is located close to the stronger component.

The effect of leakage may be mitigated by applying a weighting function, which has a gradual termination (taper), to the extracted signal event before it is Fourier transformed. To improve the situation in the frequency domain a finite time-domain window must be chosen which has smaller spectral side lobes than those of the implicit rectangular window. However, concomitant with a decrease in the level of the side lobes of a window is a broadening of the main-lobe width, which results in a reduction of spectral resolution. A compromise must be reached therefore, between main-lobe width and the extent of side-lobe suppression. Although the implicit rectangular window has the narrowest main lobe of any window, it also has the highest side-lobe level. There is a wide selection of weighting or window functions to choose from [89]. Every window, without exception, has a broader main-lobe width relative to the main-lobe width of the rectangular window. However, this is normally a small price to pay for the alleviation of the problem of leakage.

4.4.3 ‘Picket-Fence’ Effect

A consequence of the DFT transformation of a signal is that the windowed transform is sampled in the frequency domain. The sample points in the spectrum — or bins — are spaced at integer multiples of f_s/N , where f_s is the temporal sampling

frequency. If the signal being analysed has an integer number of cycles per block of N samples, no error in the measurement of magnitude or frequency occurs. However, if the signal frequency is not one of these 'discrete' frequencies, this gives rise to the so-called 'picket-fence' effect or scalloping loss [90,91]. The spectral component between two discrete frequencies will be seen by the adjacent bins because of the leakage caused by the spectrum of the data window. In the worst case, which is that of a spectral component falling midway between two bins, the amplitude is attenuated by 3.9 dB in adjacent bins in the case of a rectangular window. This is judged to be analogous to viewing the spectrum through the gaps in a 'picket-fence' where the pickets represent the inter-bin separation — hence the name.

This effect can be minimized by decreasing the separation between bins in the DFT spectrum. This is commonly achieved by increasing the number of points in the transform by appending zeros to the data record. To illustrate this, suppose a $2N$ -point DFT is performed on a $2N$ point sequence: x_0, \dots, x_{2N-1} then

$$X(k) = \sum_{n=0}^{2N-1} x(n) \exp(-j \frac{2\pi kn}{2N}) \quad k = 0, 1, \dots, 2N-1 \quad (4.5)$$

If the data sequence had been zero padded with N zeros over x_n for $n = N, \dots, 2N-1$ then (4.5) becomes:

$$X(k) = \sum_{n=0}^{N-1} x(n) \exp(-j \frac{2\pi}{N} \frac{k}{2} n) \quad (4.6)$$

This is the same as the N -point transform of (4.1), but evaluated at twice as many frequencies. This, of course, does not yield additional signal information. rather it is an interpolation between the N original transform values. In the frequency analysis of the foetal heart sounds, the initial time sequence of typically 30-50 samples was padded out with zeros to 512 samples and then transformed to the frequency domain.

4.4.4 Frequency Resolution

The extent to which the main lobes of two tones may be distinguished from one another in the frequency domain, is called the frequency resolution, Δf , and is defined to be:

$$\Delta f = \beta \left(\frac{f_s}{N} \right) \quad (4.7)$$

where f_s is the sampling frequency, N the number of samples in the data block, and β is a coefficient which reflects the fact that a window imposes an effective bandwidth on a spectral line. The factor β will always be greater than unity [89]. If $x(n)$ contains frequency components closer together than Δf Hz, then the DFT does not represent these as separate and distinct frequencies. The term f_s/N in (4.7) represents the smallest frequency interval available in the spectrum. An equivalent formulation of (4.7) is:

$$\Delta f = \frac{\beta}{T_d} \quad (4.8)$$

where T_d is the duration of the signal event. In other words, the resolution of the DFT is inversely proportional to the duration of the captured signal event. In the case of the foetal heart sounds, the duration of the sound is dependent on its origin. The first heart sounds are typically 40-50 samples long (80-100 ms), whereas the second heart sounds are between 30-40 samples long (60-80 ms) and in the latter case sometimes as short as 20 samples (40ms). The frequency resolution, therefore, cannot be expected to be better than 10 Hz and 13 Hz respectively, depending on the type of sound. The resolution is also dependent on the choice of windowing sequence applied to the sounds. Irrespective of which tapered window is applied, it will always further decrease the resolution of the analysis.

Two different window types were applied to the DFT in the frequency analysis of the foetal heart sounds. These windows — the Hamming and the Dolph-Chebyshev — are introduced in the following sections and the resulting heart sound spectra are presented and discussed.

4.5 HAMMING WINDOW

In a first attempt to obtain the broad spectral characteristics of the foetal heart sounds, the general-purpose Hamming window was applied to the time-domain sequence prior to zero padding. This window is essentially a specific instance of a family of raised cosine pulses of the form:

$$w(n) = \alpha - (1.0 - \alpha) \cos \left[\left(\frac{2\pi}{N} \right) n \right] \quad \text{for } n = 0, 1, 2, \dots, N-1 \quad (4.9)$$

where α is an adjustable parameter which allows side-lobe cancellation to be tailored. Whenever α is set to 25/46 [92], there is perfect cancellation of the first side lobe. The commonly used approximation to this value of α is 0.54 for which the raised cosine pulse is known as the Hamming window. This approximation provides good attenuation, but not elimination, of the first side lobe. The time and frequency representations of this window are shown in figure 4.1c and 4.1d.

4.6 SPECTRUM INTERPRETATION 1

Figures 4.2b to 4.6b illustrate the spectra obtained using the DFT with a Hamming window on selected principal heart sounds. These, and the other spectra of figures 4.2 to 4.6 have been plotted with a linear vertical scale. The more usual logarithmic vertical scaling was not used because it accentuated too much of the superfluous detail in the spectra.

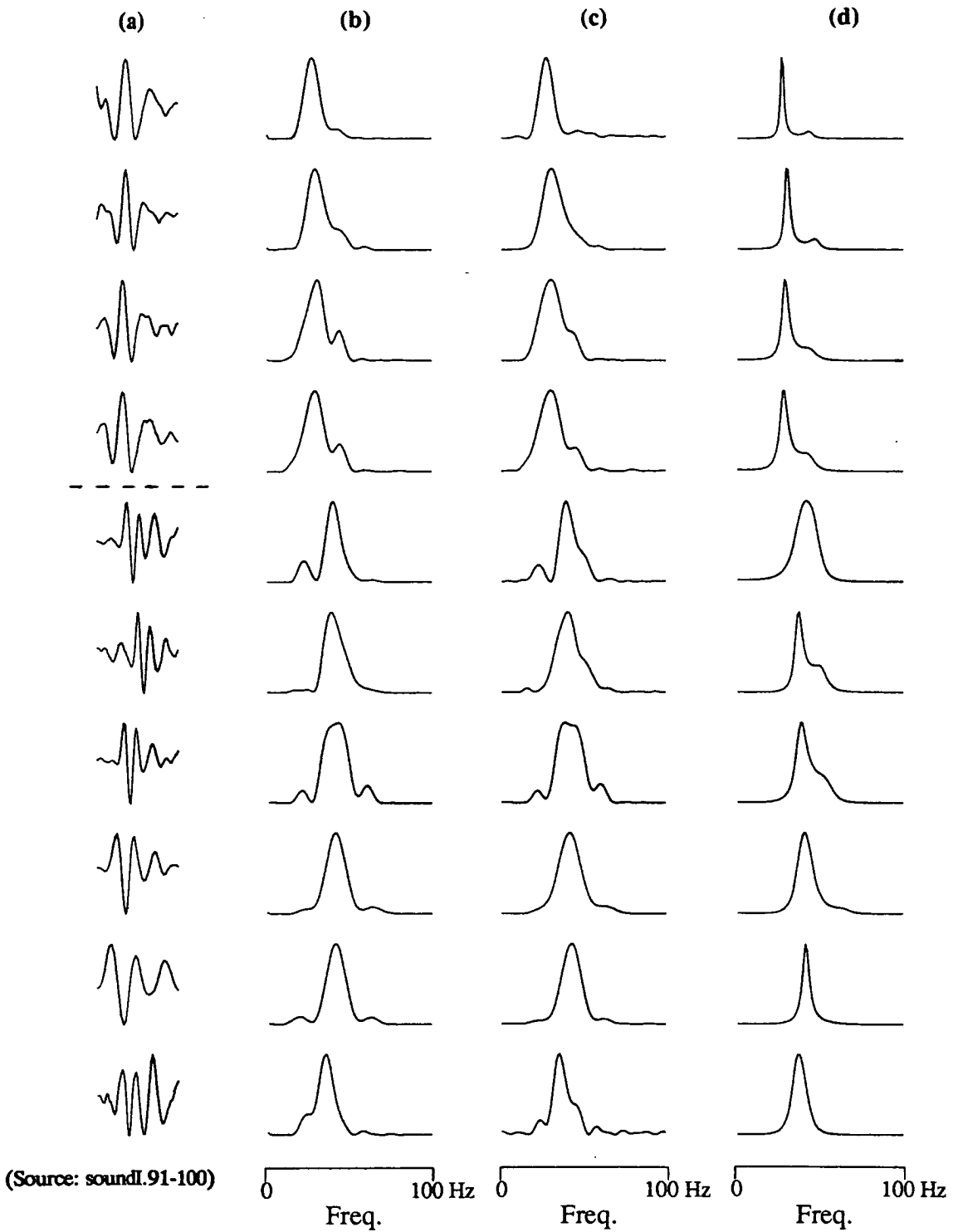


Figure 4.2 (a) Foetal first heart sounds from same subject but both sets captured 12 s apart [sounds as shown are compressed in amplitude and time]; (b) Spectra of (a) using Hamming window; (c) Spectra of (a) using Dolph-Chebyshev window with -15 dB sidelobes; (d) AR spectra of (a) using Burg algorithm with optimum taper.

Note:

The magnitude of the spectra in (b), (c), and (d) are normalized, and are plotted on linear horizontal and vertical scales.

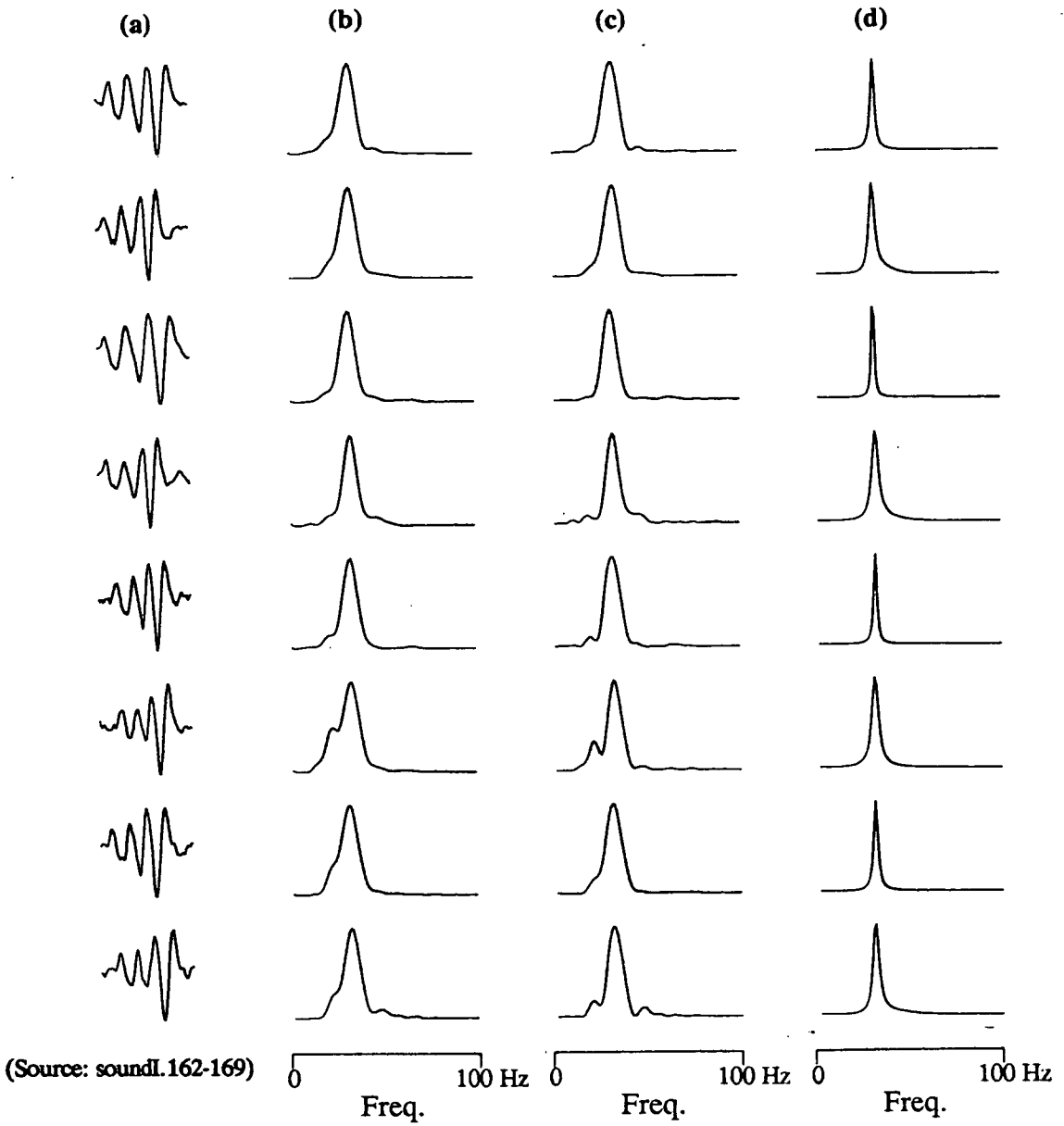


Figure 4.3 (a) Foetal first heart sounds all from same 4 s phonocardiogram [sounds as shown are compressed in amplitude and time]; (b) Spectra of (a) using Hamming window; (c) Spectra of (a) using Dolph-Chebyshev window with -15 dB sidelobes; (d) AR spectra of (a) using Burg algorithm with optimum taper.

Note:

The magnitude of the spectra in (b), (c), and (d) are normalized, and are plotted on linear horizontal and vertical scales.

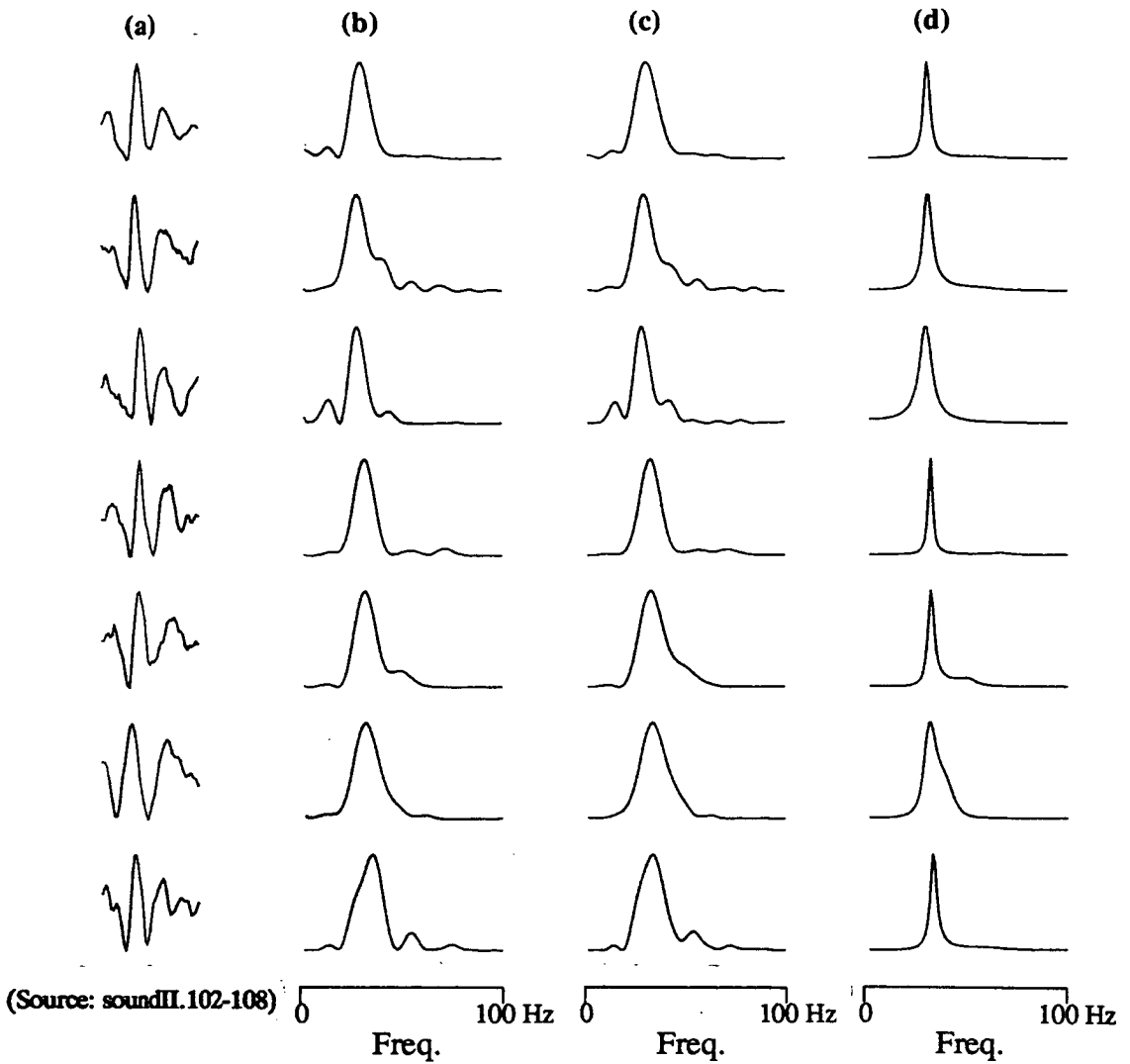


Figure 4.4 (a) Foetal second heart sounds all from same 4 s phonocardiogram [sounds as shown are compressed in amplitude and time]; **(b)** Spectra of (a) using Hamming window; **(c)** Spectra of (a) using Dolph-Chebyshev window with -15 dB sidelobes; **(d)** AR spectra of (a) using Burg algorithm with optimum taper.

Note:

The magnitude of the spectra in (b), (c), and (d) are normalized, and are plotted on linear horizontal and vertical scales.

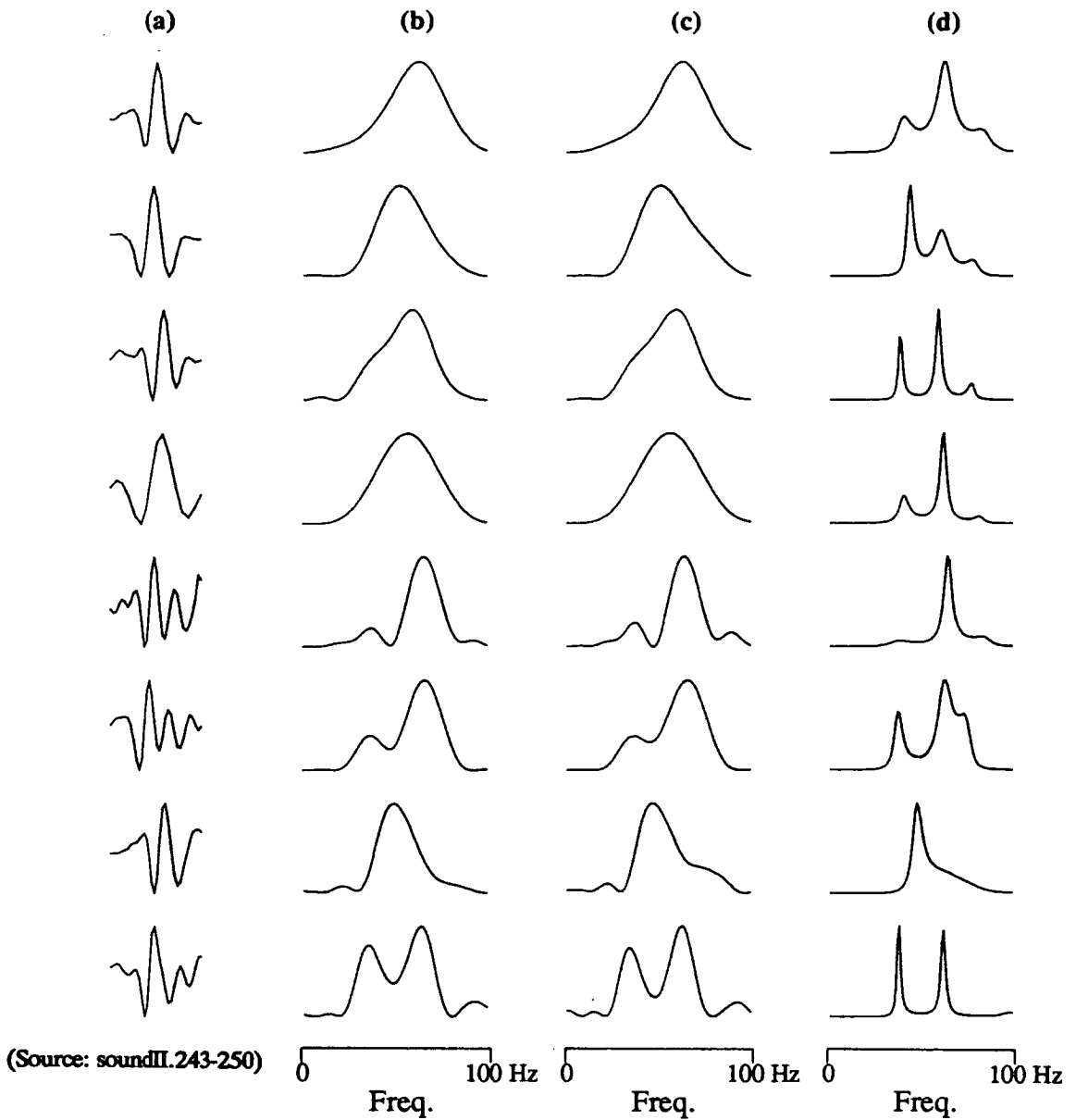


Figure 4.5 (a) Foetal second heart sounds all from same 4 s phonocardiogram [sounds as shown are compressed in amplitude and time]; (b) Spectra of (a) using Hamming window; (c) Spectra of (a) using Dolph-Chebyshev window with -15 dB sidelobes; (d) AR spectra of (a) using Burg algorithm with optimum taper.

Note:

The magnitude of the spectra in (b), (c), and (d) are normalized, and are plotted on linear horizontal and vertical scales.

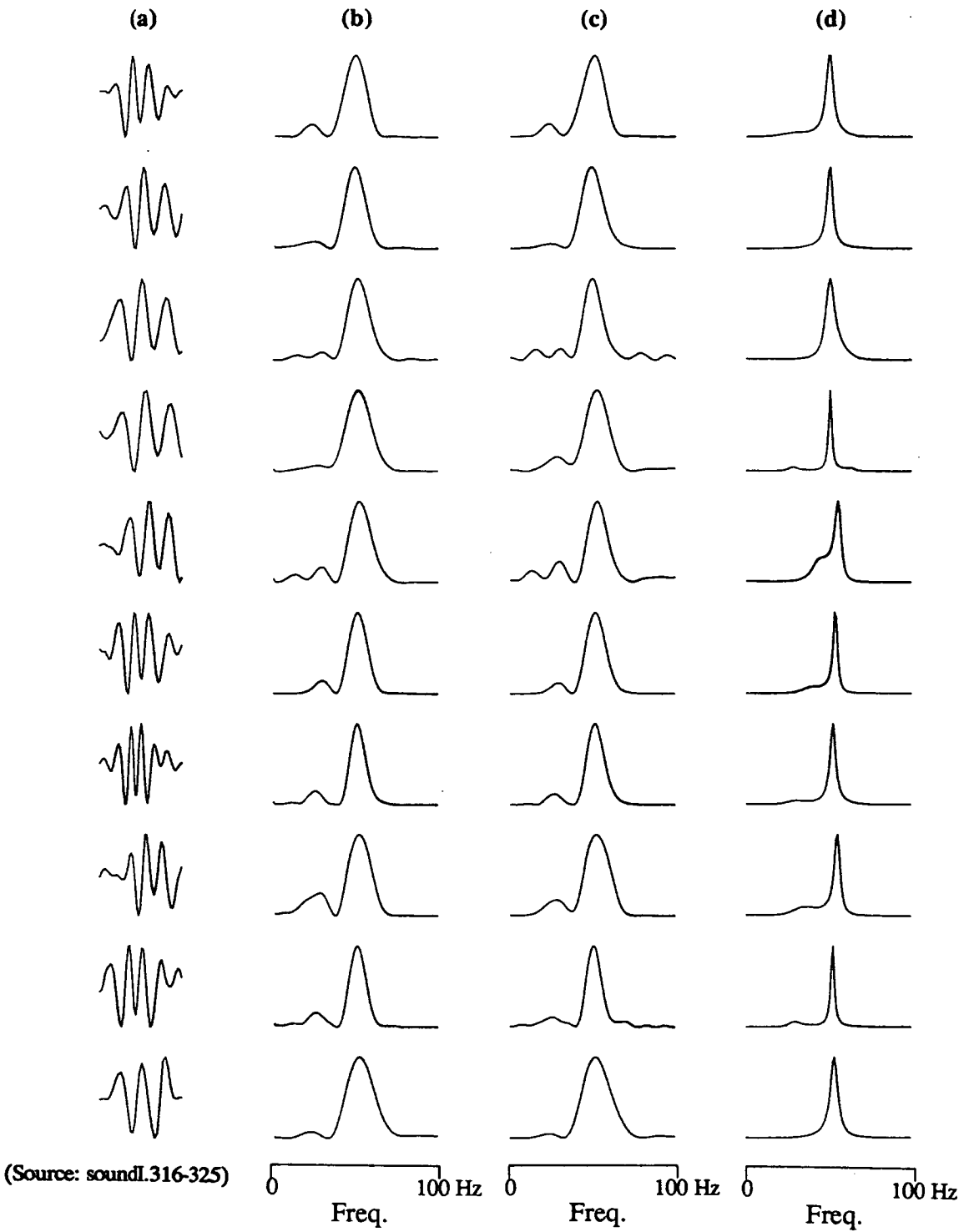


Figure 4.6 (a) Foetal first heart sounds all from same 4 s phonocardiogram [sounds as shown are compressed in amplitude and time], (b) Spectra of (a) using Hamming window, (c) Spectra of (a) using Dolph-Chebyshev window with -15 dB sidelobes, (d) AR spectra of (a) using Burg algorithm with optimum taper.

Note:

The magnitude of the spectra in (b), (c), and (d) are normalized, and are plotted on linear horizontal and vertical scales.

When first and second heart sound spectra of the twelve subjects were examined separately, it became apparent that a unique spectral signature for each sound does not emerge. This is also the case in the same gravida as shown for example in figure 4.2b where both records were captured within seconds of one another. On a per record basis — which is the primary interest — the respective spectra of first and second heart sounds are relatively constant throughout. A unusual occurrence is a marked dissimilarity in frequency (*e.g.* figure 4.5b).

When the spectra of first and second sounds were compared, there was little to distinguish the two, either for the same gravida or on a per record basis (*e.g.* figures 4.3b, 4.4b). The proposition that second sounds are of higher frequency than first sounds was generally unverifiable in the sample set examined.

The dominant peak in the spectra of respective principal sounds usually differs by only a few hertz. However, there are cases, although uncommon, where the first sounds are of higher frequency (*e.g.* compare figures 4.3b and 4.6b). Generally, the spectra of sounds are characterized by a single dominant peak around which the energy is symmetrically distributed. The position of this peak varies with subject and instance of recording in the same subject.

Thus, the frequency content and distribution of the principal heart sounds appears to be similar to within a few hertz on a per record basis. This may be due to the resolution of the DFT/Hamming window combination, or to the simple fact that both sets of valves produce a similar spectrum. In an attempt to clarify this issue, the Hamming window is substituted with a Dolph-Chebyshev window.

4.7 DOLPH-CHEBYSHEV WINDOW

Of all the N -point discrete windows, this particular window has the property that

it yields the narrowest main-lobe width for a given maximum side-lobe level [92]. The level of the peaks of the side lobes is uniform and selectable. It has been shown [72] that the coefficients of the Dolph-Chebyshev window may be generated recursively for a fixed value of the window parameter x_0 using:

$$w_M(k) = 2(x_0^2 - 1)w_{M-1}(k) + x_0^2[w_{M-1}(k-1) + w_{M-1}(k+1)] - w_{M-2}(k)$$

where M is the order of the window, $N = 2M + 1$ is the length of the window, k is the window coefficient index and

$$x_0 = \cosh \left[\frac{\cosh^{-1}(10^\alpha)}{N-1} \right]$$

where α is the logarithm of the ratio of main-lobe level to side-lobe level. The above recursive equation is subject to the following initial conditions:

$$w_0(k) = \begin{cases} 1 & k = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$w_1(k) = \begin{cases} x_0^2 - 1 & k = 0 \\ \frac{x_0^2}{2} & k = \pm 1 \\ 0 & \text{otherwise} \end{cases}$$

A realization of the Dolph-Chebyshev window for $\alpha = 2.5$ (50 dB side lobes) and 51 points is illustrated in figures 4.1e and 4.1f.

4.8 MAIN-LOBE WIDTH OF HAMMING AND DOLPH-CHEBYSHEV WINDOWS

Any window imposes an effective bandwidth on a spectral line. The sole exception to this is the case of a rectangular window on a block of data within which there are an integer number of cycles. The imposed bandwidth, in turn, affects the minimum frequency separation of tones so that their main lobes can be resolved. Some windows have the main-lobe width and side-lobe level already set, *e.g.* the Hamming

window; whereas others, like the Dolph-Chebyshev window, require that either the side-lobe level or main-lobe width be specified beforehand. The reason for the choice of the Dolph-Chebyshev window in this instance, is that an increase in the level of the side lobes may be exchanged for a decrease in the bandwidth of the main lobe (section 4.4.2). Thus resolution may be traded for the detectability of spectral components. The Dolph-Chebyshev window also offers good 'close-in' visibility to a large signal component. This is important in a small transform where most bins are close to any given large component.

In the following table (table 4.1) various main-lobe widths and the corresponding side-lobe levels are given for the two windows under discussion.

Window type	Sidelobe level max (dB)	-6 dB B/W (Hz)
Hamming	-43	17.0
Dolph-Chebyshev	-15	10.9
	-20	12.2
	-25	13.4
	-30	14.6
	-40	16.6
	-50	18.5

Table 4.1 Bandwidth of mainlobe of Hamming and Dolph-Chebyshev windows at -6 dB relative to the peak height of the main lobe (51 point window with zero padding to 1024 points).

The choice of the -6 dB main-lobe bandwidth to be the limit for the resolution of two closely spaced equi-strength spectral lines was suggested by Harris [89]. This was prompted by the desire to counteract the effect of the coherent addition of the

extraneous spectral components and to provide good resolution.

A side-lobe level of -15 dB for the Dolph-Chebyshev window was finally chosen to resolve the ambiguities present in the number and position of the *dominant* frequency components apparent in the Hamming windowed DFT. This was achieved at the expense of the suppression of the smaller components.

4.9 SPECTRUM INTERPRETATION 2

Figures 4.2c to 4.6c illustrate the spectra obtained using the DFT with a Dolph-Chebyshev window on the same sounds previously analysed in section 4.6. Compared with the spectra produced with the Hamming window, there is not a significant improvement in the separation of components. In these plots the energy distribution is still concentrated in a broad smooth lobe.

In order to decompose this lobe to ascertain what frequencies it contains, a analysis method is needed which yields higher resolution. In the past two decades, techniques have been developed which are capable of making high-resolution spectral estimates when the data length is short. These parametric methods are the subject of the next section and in particular the auto-regressive model which is the most widely applied of these techniques.

4.10 PARAMETRIC METHODS

Parametric methods of spectral analysis are based on the model of a linear signal generator system which synthesizes the observed signal from an input of white noise. The spectrum estimation problem may then be restated as the estimation of the transfer function of the model $G(z)$, using

$$\Phi_u(z) = \Phi_v(z)G(z)G(z^{-1})$$

where $\Phi_x(z)$ is the power spectral density (PSD) of the output sequence and $\Phi_w(z)$ is the PSD of the input sequence (figure 4.7).

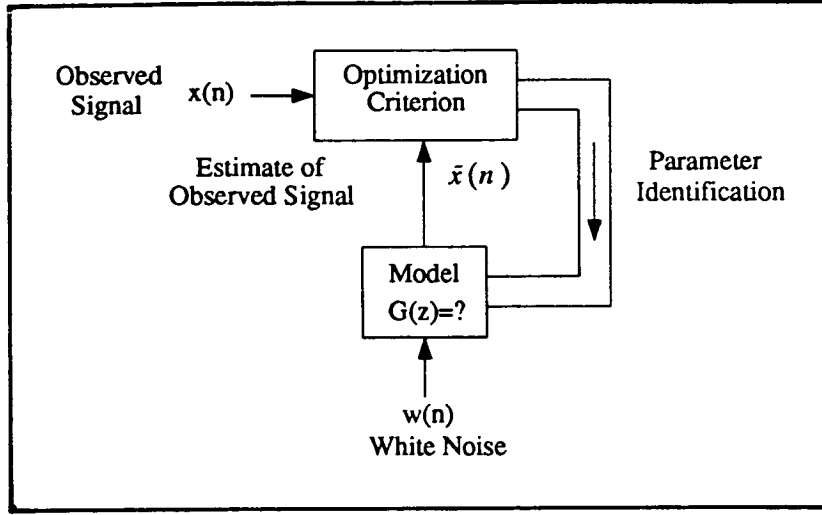


Figure 4.7 Parametric method of signal characterization. The observed signal $x(n)$ is reproduced by a linear system with transfer function $G(z)$ which is fed with white noise $w(n)$.

Parametric spectral analysis uses the auto-regressive moving average (ARMA) time-series model which approximates many discrete-time deterministic and stochastic processes [93]. In this model, the input sequence $w(n)$, and the observed data $x(n)$ which represents the output sequence are related by the linear difference equation

$$x(n) = -\sum_{k=1}^p a_k x(n-k) + \sum_{k=0}^q b_k w(n-k)$$

This model is illustrated in figure 4.8a. The transfer function, $H(z)$, for such a system is given by

$$H(z) = \frac{B(z)}{A(z)}$$

where

$$A(z) = 1 + \sum_{k=1}^p a_k z^{-k}$$

$$B(z) = \sum_{k=0}^q b_k z^{-k}$$

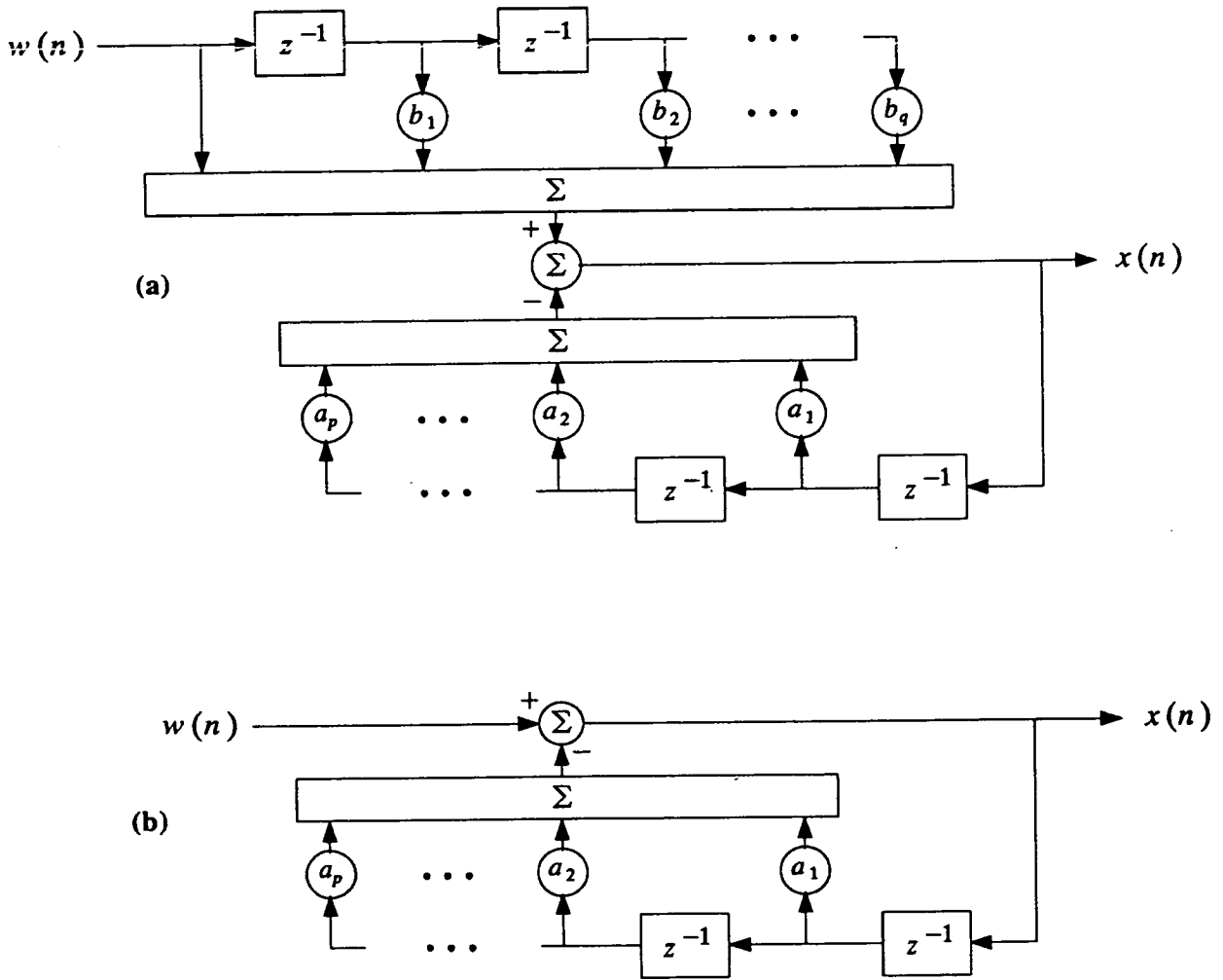


Figure 4.8 (a) Auto-regressive moving average model of order (p, q) ;
(b) auto-regressive model of order p .

The power density spectrum of the input sequence $\Phi_w(z)$ is related to the transfer function $H(z)$, and the PSD of $x(n)$, $\Phi_x(z)$, by

$$\Phi_x(z) = H(z) H(z^{-1}) \Phi_w(z)$$

The input, or driving, sequence $w(n)$ is not assumed to be available for analysis, however, in the ARMA model it is assumed to be a white noise process of zero mean with autocorrelation

$$\phi_{ww}(m) = \sigma_w^2 T \delta(m)$$

where σ_w^2 is the variance, and T is the sampling interval. Thus, the power density spectrum of the observed data is

$$\Phi_x(f) = \sigma_w^2 T \left| \frac{B(f)}{A(f)} \right|^2$$

where the polynomials $A(f)$ and $B(f)$ are defined as

$$A(f) = 1 + \sum_{k=1}^p a_k \exp(-j2\pi f k T)$$

$$B(f) = 1 + \sum_{k=1}^q b_k \exp(-j2\pi f k T)$$

If all the $\{b_k\}$ are set to zero, except $b_0=1$ (figure 4.8b), then the process is strictly an autoregression of order p , *i.e.* the sequence $x(n)$ is a linear regression on itself, *viz.*

$$x(n) = - \sum_{k=1}^p a_k x(n-k) + w(n) \quad (4.10)$$

The autoregressive process model is the most widely adopted of the parametric models. This is mainly because the AR model is suitable for representing spectra with sharp peaks (resonances), and secondly, the AR parameters are obtained from linear equations. Furthermore, the Wold decomposition theorem [94] states that an ARMA process may be approximated by a higher order AR model. This is advantageous because ARMA models require the solution of non-linear equations.

The PSD of the AR process is therefore

$$\Phi_x(f) = \frac{\sigma_w^2 T}{\left| 1 + \sum_{k=1}^p a_k \exp(-j2\pi f k T) \right|^2} \quad (4.11)$$

From (4.11) it is seen that an estimate of the PSD may be found from the model parameters $\{a_k\}$ and σ_w^2 . To determine the AR parameters, a relationship between $\{a_k\}$ and the autocorrelation function will now be developed. If (4.10) is multiplied by the complex conjugate of $x(n-m)$, and the expectation ($E[\cdot]$) is then taken of both sides of the equation, the following results

$$\begin{aligned} E[x(n)x^*(n-m)] &= -\sum_{k=1}^p a_k E[x(n-k)x^*(n-m)] + E[w(n)x^*(n-m)] \\ \phi_{xx}(m) &= -\sum_{k=1}^p a_k \phi_{xx}(m-k) + E[w(n)x^*(n-m)] \end{aligned} \quad (4.12)$$

where $\phi_{xx}(m)$ is the autocorrelation of $x(n)$ at m lags, and conjugation is denoted by a superscripted asterisk. If the cross-correlation term between input and output in (4.12) is examined separately and the assumption is made that the filter is both stable and causal then

$$\begin{aligned} \phi_{wx}(m) &= E[w(n+m)x^*(n)] \\ &= E\left[w(n+m) \sum_{k=0}^{\infty} h^*(k)w^*(n-k)\right] \\ &= \sum_{k=0}^{\infty} h^*(k) \sigma^2 \delta(m+k) \\ &= \sigma^2 h^*(-m) \\ &= \begin{cases} 0 & m > 0 \\ \sigma^2 h^*(0) & m = 0 \end{cases} \end{aligned}$$

where $h(k)$ are the filter impulse response coefficients. Substituting for the cross-correlation term in (4.12) yields the following relationships known as the *Yule-Walker*

equations or normal equations

$$\phi_{xx}(m) = \begin{cases} -\sum_{k=1}^p a_k \phi_{xx}(-k) + \sigma^2 & m = 0 \\ -\sum_{k=1}^p a_k \phi_{xx}(m-k) & m > 0 \end{cases} \quad (4.13)$$

The AR parameters are obtained from the solution of:

$$\begin{bmatrix} \phi_{xx}(0) & \phi_{xx}(-1) & \cdots & \phi_{xx}(-(p-1)) \\ \phi_{xx}(1) & \phi_{xx}(0) & \cdots & \phi_{xx}(-(p-2)) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{xx}(p-1) & \phi_{xx}(p-2) & \cdots & \phi_{xx}(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} \phi_{xx}(1) \\ \phi_{xx}(2) \\ \vdots \\ \phi_{xx}(p) \end{bmatrix} \quad (4.14)$$

and the variance σ_w^2 may be determined from

$$\sigma_w^2 = \phi_{xx}(0) + \sum_{k=1}^p a_k \phi_{xx}(-k) \quad (4.15)$$

Equations (4.14) and (4.15) may be combined into the following single augmented matrix equation:

$$\begin{bmatrix} \phi_{xx}(0) & \phi_{xx}(-1) & \cdots & \phi_{xx}(-(p)) \\ \phi_{xx}(1) & \phi_{xx}(0) & \cdots & \phi_{xx}(-(p-1)) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{xx}(p) & \phi_{xx}(p-1) & \cdots & \phi_{xx}(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (4.16)$$

Since the correlation matrix of (4.16) is Toeplitz[†], there exists a computationally efficient algorithm, known as the Levinson-Durbin recursion [95], which provides a means for sequentially solving (4.16) without matrix inversion. This algorithm has a computational complexity proportional to p^2 whereas standard techniques such as Gauss elimination have a complexity proportional to p^3 where p is the number of AR

[†]Elements along each diagonal are identical.

coefficients to be determined [93]. The Levinson-Durbin algorithm (figure 4.9) recursively calculates the AR parameter sets for successively higher-order models: $\{a_{11}, \sigma_1^2\}$, $\{a_{21}, a_{22}, \sigma_2^2\}$, \dots , $\{a_{p1}, a_{p2}, \dots, a_{pp}, \sigma_p^2\}$ where the leading subscript denotes the iteration number, and p is the model order.

Initialization:

$$a_{11} = -\frac{\phi_{xx}(1)}{\phi_{xx}(0)}$$

$$\sigma_1^2 = (1 - |a_{11}|^2)\phi_{xx}(0)$$

Recursion for $k = 2, 3, \dots, p$

$$a_{kk} = -\frac{\left[\phi_{xx}(k) + \sum_{l=1}^{k-1} a_{k-1,l} \phi_{xx}(k-l)\right]}{\sigma_{k-1}^2}$$

$$a_{ki} = a_{k-1,i} + a_{kk} a_{k-1,k-i}^*$$

$$\sigma_k^2 = (1 - |a_{kk}|^2)\sigma_{k-1}^2$$

Figure 4.9 Levinson-Durbin algorithm.

4.11 RELATIONSHIP OF THE AR PROCESS TO LINEAR PREDICTION

A complementary view of the problem of determining the AR coefficients is to find the *all-zero* filter of order p which will produce $w(n)$ from $x(n)$. The transfer function of such a filter is $A(z)$, and it is appropriately termed the whitening or inverse filter for the AR process. The filter coefficients are derived from consideration of the theory of linear prediction [96], where the aim is to remove the colouration from $\Phi_x(z)$. Thus, the problem can now be reformulated in terms of finding the

p th order linear predictor $A(z)$, and the power of the prediction error sequence σ_w^2 .

If $x(n)$ is a sample sequence from a stationary random process, then a linear predictor of order p for the data point $x(n)$ has the form

$$\hat{x}(n) = -\sum_{k=1}^p a_k x(n-k) \quad (4.16)$$

where the circumflex denotes an estimate and $\{a_k\}$ are the predictor coefficients for the p th order predictor. $\hat{x}(n)$ in (4.16) is referred to as the *forward predictor* since it 'predicts' forward in time on the basis of the previous samples. The forward prediction error, $e_p(n)$, is then

$$e_p(n) = x(n) - \hat{x}(n) \quad (4.17)$$

The minimization of the mean-square value of $e_p(n)$ with respect to the predictor coefficients leads to a set of equations identical to those of (4.13) [95]. The Levinson-Durbin algorithm (figure 4.9) may then be used to solve these for the predictor coefficients. If the Levinson-Durbin recursion for a_{ki} is substituted into (4.16) and the result is applied to (4.17), then an order-recursive equation for the *forward* prediction error is obtained:

$$\begin{aligned} e_p(n) &= x(n) + \sum_{k=1}^{p-1} (a_{p-1,k} + a_{pp} a_{p-1,p-k}^*) x(n-k) + a_{pp} x(n-p) \\ &= e_{p-1}(n) + a_{pp} b_{p-1}(n-1) \end{aligned} \quad (4.18)$$

where

$$b_p(n) = x(n-p) + \sum_{k=1}^p a_{pk}^* x(n-p+k) \quad (4.19)$$

The term $b_p(n)$ is the *backward* prediction error which represents the error arising when $x(n-p)$ is 'predicted' from the later samples: $x(n-p+1), \dots, x(n)$. If the Levinson-Durbin recursion is substituted into (4.19), the following is obtained

$$b_p(n) = b_{p-1}(n-1) + a_{pp}^* e_{p-1}(n) \quad (4.20)$$

The relationships in (4.18) and (4.20) define a lattice filter structure [97, 98] which is illustrated in figure 4.10. The transfer function of such a filter is

$$A(z) = 1 + \sum_{k=1}^p a_{pk} z^{-k}$$

which is the inverse of the transfer function $H(z)$ of the AR generator model. The implementation of the prediction error filter as a lattice filter has been found to suffer from less round-off noise and is less sensitive to coefficient value perturbations than the equivalent transversal filter realization [95].

Although (4.13) could be solved using autocorrelation estimates directly obtained from the data, better AR parameter estimates are obtained using least squares estimation techniques. The most popular of these is the method introduced by Burg.

4.12 BURG (HARMONIC) ALGORITHM

A method devised by Burg [99] for estimating the AR parameters minimizes the sum of the forward and backward prediction error energies, ϵ_p , where the prediction errors are defined for $p \leq n \leq N-1$ on a data set $x(n)$, for $n = 0, 1, 2, \dots, N-1$

$$\epsilon_p = \sum_{n=p}^{N-1} |e_p(n)|^2 + \sum_{n=p}^{N-1} |b_p(n)|^2 \quad (4.21)$$

subject to the constraint that the AR parameters satisfy the Levinson-Durbin recursion for all orders from 1 to p which in turn ensures a stable AR filter. If the derivative of (4.21) with respect to a_{ii} for $i = 1, 2, \dots, p$ is set to zero then

$$a_{ii} = \frac{-2 \sum_{k=i}^{N-1} b_{i-1}^*(k-1) e_{i-1}(k)}{\sum_{k=i}^{N-1} (|b_{i-1}(k-1)|^2 + |e_{i-1}(k)|^2)} \quad (4.22)$$

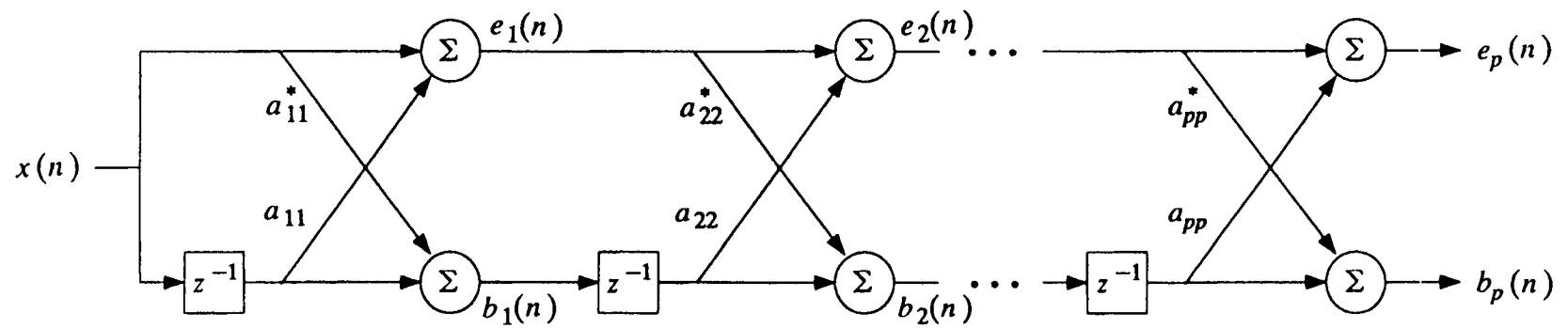


Figure 4.10 Linear predictor error filter realized in a lattice filter structure.

A further recursion discovered by Andersen [100] simplifies the evaluation of the denominator of (4.22).

Define

$$DEN(i) = \sum_{k=i}^{N-1} (|b_{i-1}(k-1)|^2 + |e_{i-1}(k)|^2)$$

then

$$DEN(i) = \left[1 - |a_{i-1,i-1}|^2 \right] DEN(i-1) - |b_{i-1}(N-i)|^2 - |e_{i-1}(i)|^2$$

The major advantages of Burg's algorithm for the estimation of the AR coefficients are: high frequency resolution in the resulting spectrum, stable AR models, and the computational efficiency of the algorithm [93]. However, Burg's method also has several limitations [101, 102]. Firstly, spectral line splitting: where one actual peak in the spectrum is split into two or more closely spaced peaks; secondly, spurious peaks: for high model orders this method may introduce detail into the spectrum which is not actually present; and, finally, frequency bias: for sinusoidal signals in noise this method exhibits a shift from the true frequency of a spectral component which is dependent on the initial phase of the signal. These limitations have prompted modifications to the Burg algorithm [101, 103] mostly in the introduction of a weighting sequence $w_p(n)$ to the energies of the forward and backward prediction errors, *viz.*

$$\epsilon_p = \sum_{n=p}^{N-1} w_p(n) \left[|e_p(n)|^2 + |b_p(n)|^2 \right]$$

upon minimization, this yields

$$a_{ii} = \frac{-2 \sum_{k=i}^{N-1} w_{i-1}(k) b_{i-1}^*(k-1) e_{i-1}(k)}{\sum_{k=i}^{N-1} w_{i-1}(k) (|b_{i-1}(k-1)|^2 + |e_{i-1}(k)|^2)}$$

4.12.1 Optimum Taper for Burg Algorithm

The particular weighting sequence used in this instance is that proposed by Kaveh and Lippert [101, 104]. Their criterion for the selection of a weighting sequence was to minimize the mean frequency error variance of a real sinusoid with respect to the weighting sequence. This resulted in the following optimum window which is parabolic in form:

$$w_p(n) = \frac{6(n-p+1)(N-n)}{(N-p)(N-p+1)(N-p+2)} \quad n = p, \dots, N-1$$

which may be generated recursively from:

Optimum Taper for Burg Algorithm
$w_p(p-1) = 0$
$w_p(p) = (N-p) \frac{\lambda}{2}$
$w_p(n) = 2w_p(n-1) - w_p(n-2) - \lambda \quad n = p+1, \dots, N-1$
$\lambda = \frac{12}{(N-p)(N-p+1)(N-p+2)}$

Both frequency bias and line splitting are alleviated with this window without a degradation in spectral resolution [101].

4.13 SELECTION OF MODEL ORDER

A vital consideration in the discussion of parametric modelling is the choice of the model order p which will not be known beforehand. Generally, if the order is too low, there will be an insufficient number of poles to adequately model the spectrum which will in consequence be highly smoothed. However, if the choice of order is too high, this will produce spurious peaks in the estimated spectrum neglecting the associated increase in the computation time. Various objective criteria have been

proposed for selecting an appropriate model order [10, 93]. However, these have been shown [105] not to be useful whenever the data length is short — which is the case in this particular application. As Marple [95] remarks, the final determination of a suitable model order is a subjective judgement particularly in the analysis of actual data which originates from an unknown process.

In selecting a model order, the following empirical criteria were considered: (a) to model sinusoidal components in the data at least two poles should be assigned to each sinusoid [106, 107]; (b) An ARMA process, *i.e.* sinusoids in white noise, is being approximated by an AR process which by the Wold decomposition theorem *q.v.* requires a higher-order AR model; (c) the model order should not be greater than half the number of sample points in the data [93]; (d) the spectrum obtained with a particular model order should be consistent with the spectrum derived from the Fourier transform method.

Finally, a model order of 10 was selected for data segments which were on average 30-50 samples long. This produced estimates which were consistent with, and better resolved than, those obtained with the windowed Fourier transform.

4.14 SPECTRUM INTERPRETATION 3

The spectra obtained using the optimum taper on the Burg algorithm are presented in figures 4.2d to 4.6d. The broad lobe apparent in the windowed DFT has been decomposed, in certain cases, to reveal detail which was hidden by the previous methods of analysis. However, in the majority of cases, this parametric analysis method has reinforced the conclusions of section 4.6. Whenever latent detail was revealed, it was seen to resemble the spectral characteristics of either mono or dichromatic tones.

One feature that was observed across all the principal heart sounds examined was that the spectrum and the frequency shifts of the spectral pattern of first and second sounds were common to both sounds. Thus, it may be supposed that first and second sounds are similarly affected by the generating mechanism and/or the transmission medium.

In conclusion, for the sample set of 741 heart sounds, there is no evidence to suggest that the two principal sounds each exhibit a unique spectral signature across subjects or on a per record basis. Once this had been ascertained, it invalidated the subsequent investigation of the frequency excursions of the heart sounds and the frequency content of the adventitious sounds.

This raises the question that if each sound has a similar spectrum, how can auscultators of foetal heart sounds have reported that the second sound is higher pitched than the first? This is a very difficult claim to falsify given the nature of the sound generator and the transmission medium. It may well be attributable to an artefact of the human auditory system arising from the shorter duration and more abrupt beginning and end of second heart sounds which gives the apparent sensation of a higher frequency. As the spectral content of the principal heart sounds is on and below the threshold of audibility, the fidelity of the human ear in exactly registering these sounds would also be called into question.

The spectra of figures 4.2 to 4.6 justify the assertion made in chapter 3, section 3.7, where it was proposed that the filtering threshold which is commonly used to obtain the PhCG should be set lower. From the spectra, it is evident that the majority of the energy in the foetal heart sounds, which is infrasonic, was filtered out in previous attempts to replicate with electronic processing what was heard on auscultation.

4.15 SUMMARY AND CONCLUSION

This chapter investigates the methods and results of the spectral analysis of foetal heart sounds. The aim of this work was to see if first and second sounds each produce a unique spectral signature, which would then have been used in an automatic classification system to distinguish between them. Using two methods — conventional Fourier transformation and auto-regressive analysis — 741 principal heart sounds from twelve subjects were investigated. The results showed that this particular sample set exhibited spectra which bore a similarity to those produced by mono and dichromatic tones. Overall, for both principal sounds no uniquely differentiating features were found in the frequency domain. In general, the spectra of both principal heart sounds are similar to within a few Hertz. It was also observed that the majority of the energy in the foetal heart sounds lies below the range that previous investigators have analysed.

CHAPTER 5

KNOWLEDGE-BASED SYSTEMS AND SIGNAL UNDERSTANDING

In chapter 3, section 3.9 the comparative ease with which the eye discerns the foetal heart sounds in any PhCG has been noted. This, of course, was the means by which it was initially possible to manually extract and identify sounds from the PhCG prior to the frequency analysis which is detailed in chapter 4. In itself, this manual operation preempted a major task in computer processing and analysis. Now that it has been established that it is not possible to distinguish between the two principal heart sounds on the basis of their frequency content, attention is turned to the time-domain, and the latent human cognitive processes which so adeptly identify the foetal heart sounds.

As the frequency analysis of the foetal heart sounds showed, there must be other factors which are brought into the identification process other than the nature of the sounds themselves when taken in isolation. These additional factors, it is proposed, are, firstly, the *context* within which a sound is found, and, secondly, specific knowledge about the domain. By context is implied the temporal and causal relationships of the principal heart sounds, not only to themselves but also to the other surrounding signal events in a segment of the PhCG. In contrast to contextual knowledge, domain or background knowledge uses information which does not derive

from what is contained in the signal itself, but rather from what is known *about* that type of signal, for example, the operational characteristics of the signal-producing mechanism — the heart.

The incorporation of these external sources of knowledge into the processing and analysis of the PhCG is an attempt to compensate for the inherently low signal-to-noise ratio of the signal which arises mainly from the presence of many non-valvular sounds and possibly, poor transducer placement. A search for an expected signal event in the signal space is made more effective when what is being sought and the region in which it must appear are both known beforehand.

It is the nature of the PhCG and the principal heart sounds which makes conventional analytic techniques of signal processing inappropriate for such a volatile environment. Traditionally, signal processing has been concerned with the application of numerical algorithms to improve the signal-to-noise ratio or to transform the domain of the signal thereby accentuating certain of its features. In both cases the characteristics of the signal and the noise are known beforehand. These numerical techniques rely on a mathematical model of the system and certain idealizing assumptions. Such an analytical approach fails when the signal characteristics are not well defined, or when there is insufficient understanding of the underlying mechanism of signal production, or again, when the system is too complex to be modelled mathematically. In these situations a more pragmatic approach is required. Such an approach is provided by the knowledge-based formalism as applied to signal processing [108].

Knowledge-based systems (KBS) represent a programming paradigm which is meant to mimic the problem-solving ability of humans. KBSs are a departure from conventional computer-based techniques of information processing insofar as they

'reason' about data using facts and heuristics which are selectively applied by an inference mechanism. Leaving aside the implementation details, the rationale behind this methodology is essentially inductivism and empiricism, insofar as the heuristic knowledge used in such a system has been induced from accumulated observations of specific instances. This 'domain knowledge' represents the captured experience of someone who has expertise in analysing that particular problem, in this case the phonocardiographic signal.

The power of this formalism lies in its empiricism and the very specificity of the domain knowledge which are both inherently close to reality. This is contrasted with conventional techniques which, of necessity, make idealizing assumptions about the nature of the signal and its generating processes. Due to the variability in the phonocardiographic signal and its low signal-to-noise ratio, conventional processing methods have met with little success which in turn led to the demise of foetal phonocardiography. However, the use of knowledge-based methods of signal analysis does not preclude the deployment of numerical signal processing algorithms, as these may be accommodated within the system as yet another aspect of the expert's knowledge.

The goal of the processing in this domain is to automatically identify certain events in the foetal cardiac cycle from the recorded acoustic signal produced by the functioning of the foetal heart. This task is doubly complicated because, firstly, the clinician is more concerned with the extremes of variation and irregularities than the norms and, secondly, the signal generators, and transmission medium are non-linear, non-stationary, non-deterministic, time-varying, noisy, and prone to transients. Although recent transducer designs have improved the achievable signal-to-noise ratio, there still exists the problem of interference from the many adventitious sounds

associated with the operation of the heart. Considering these aspects of the problem, a recourse was to acquire and imitate the procedures used by a *domain expert* while analysing foetal phonocardiograms. This was embodied within the framework of a knowledge-based system. The flexibility offered by such a formalism allows the wide variability in the signal to be accommodated.

There are various problems associated with the use of KBSs in any application. First among these is the difficulty of acquiring the necessary knowledge about the domain, and secondly the issue of how the knowledge is to be represented in a computing machine. Once these problems have been dealt with, there remains the question of how this knowledge is to be organized within the KBS and the mechanism by which it is to be applied.

This chapter describes the concept of KBSs, and discusses the previously outlined problems and provides reasons for the particular choice of representation and organization used in this application. More specifically, the work of other researchers in the field of KBS as applied to signal understanding is presented, and also the applicability of their techniques to the processing and analysis of the foetal PhCG.

5.1 KNOWLEDGE-BASED SYSTEMS

A KBS[†] may be defined as a computer program that uses large amounts of knowledge about a single domain to achieve a high level of competence in that domain [109,110]. The most important characteristic of a KBS is precisely its dependence on a 'large amount' of domain-specific knowledge. Constructors of KBS [111,112] stress the knowledge component itself over any formal reasoning method employed by such a system since 'human experts achieve outstanding performance

[†]The term knowledge-based system is broadly synonymous with expert system.

because they are knowledgeable. If computer programs embody and use this knowledge, then, they too should attain high levels of performance' [112].

The general architecture of a KBS is illustrated in figure 5.1. It is here in the separation of control and knowledge that the distinction between KBSs and conventional programming paradigms is apparent. The essential elements of this system are the knowledge base, the data base, the control mechanism or inference engine, and user interface.

5.1.1 Knowledge Base

The knowledge base contains the relationships and constraints that characterize the particular domain. These entities consist of not only facts and rules but also heuristics or rules of thumb. The knowledge base is kept distinct from the other components of the system so that it can easily change and grow through modifications and additions to the existing knowledge without the need for amendments to the other components of the system. Such a structure also allows the creation of a general rather than specific-purpose KBS where a variety of application-specific knowledge bases can be attached onto the same inference engine[†]. An important consideration in constructing the knowledge base is that it should be consistent and as complete as possible [114].

5.1.2 Data Base

The data base consists of two parts: a static part which holds the input data, whether numeric or symbolic, and a dynamic part which holds the results of the computations of the system on that data. Thus at any time the content of the data base

[†]This is the basis of commercially available expert system shells [113].

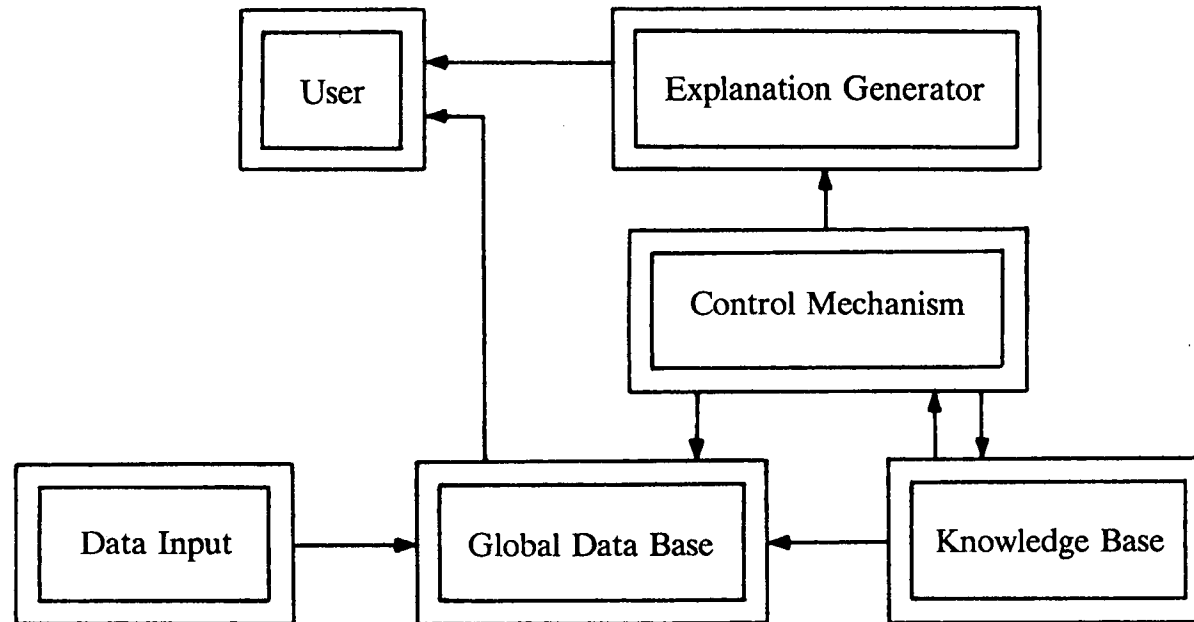


Figure 5.1 Architecture of knowledge-based systems.

will reflect the current state of progress towards the solution of the problem.

5.1.3 Control Mechanism

The control mechanism regulates the interaction of the knowledge base with the data base by the selection and application of knowledge which may then act to change or modify the contents of the data base. It also maintains a record of knowledge activations and conclusions reached which may be interrogated by the user. In its simplest form, the control mechanism searches in the knowledge base for a piece of knowledge which is applicable to the current state of the data base, then executes it. This *select-execute* cycle [115] is then repeated until a completion condition is reached. The control can either be 'forward-chaining' (inductive) or 'backward-chaining' (deductive). Forward-chaining reasons from the data to establish conclusions whereas backward-chaining searches in the data to find supporting evidence for a hypothesis. It is usually the nature of the domain which determines which of the two strategies is more appropriate, although backward chaining is the more common in existing systems [116]. However, the most effective strategy, in terms of speed in reaching a conclusion, is achieved when both methods are combined [117].

5.1.4 User Interface

The user interface is a further component in a KBS and another distinguishing feature from conventional programming paradigms. Its function is to communicate with the user by providing a justification for a particular course of action taken, or prompting the user to supply information whenever difficulties are encountered or accepting volunteered information. The ability of a KBS to justify its conclusions becomes more important when the system is operating on large data sets using approximate reasoning, especially if the user is to base important decisions on the

advice of the KBS. In these cases the user must be confident in the reasoning and assumptions employed [118]. However, among current implementations of KBSs, this is the least well-developed facility.

5.2 KNOWLEDGE ACQUISITION AND REPRESENTATION

In all the preceding discussion it was conveniently assumed that the domain knowledge had in some way been acquired from the human expert and transformed into a representation suitable for manipulation by computer. The acquisition of knowledge and its representation are fundamental to the construction of KBSs and to the ability of the system to solve actual problems. The following two sections describe the stages involved in the transfer of knowledge from man to machine.

5.2.1 Knowledge Acquisition

The task of eliciting the expertise and incorporating it into the knowledge base is the responsibility of the knowledge engineer. What might at first appear a straightforward task is in fact the most problematic and time-consuming aspect of the building of KBSs. It has been a frequent observation by knowledge engineers that this difficulty arises from the fact that experts do not structure their decision making in any formal way, and have difficulty in describing the reasoning path that they follow [110] — which may well be below the level of their conscious awareness. A further problem is the expert's hesitancy in disclosing his expertise for fear of being superceded by the very machine that he is helping to build [119].

There are three complementary methodologies used in the process of knowledge acquisition: dialogue with the expert, observing the expert at work, and consulting relevant technical literature [111, 120, 121]. Once this stage has progressed to give the knowledge engineer a satisfactory level of understanding of the domain and the

problem, an appropriate scheme is selected for representing the problem-solving knowledge — this is referred to as *knowledge representation* (section 5.2.2).

When it is judged that sufficient knowledge has been elicited to permit an expert performance by the KBS, the task of programming the system begins [122]. Once a prototype KBS has been constructed, its performance on test situations is evaluated by the expert in conjunction with the knowledge engineer, and its shortcomings are identified. Then begins a cycle of knowledge refinement where inconsistencies, errors, and omissions in the knowledge base are located and resolved.

5.2.2 Knowledge Representation

The most common knowledge representation formalism is the *production rule* or *if-then* rule [123]. KBSs which employ this representation are known as rule-based systems or production systems [124]. In general, such rules consist of an antecedent and a consequent, *viz.*

if conditions C1 and C2 hold then action A

where C1 and C2 must be satisfied by the current state of the data base before the rule can be applied and the action taken to update the data base. An example of a production rule taken from the XCON KBS (formerly known as R1) [110,114] is shown below in figure 5.2. This system which is used to configure DEC VAX mini-computers has more than 3,300 rules in its knowledge base.

The advantage of the production-rule format is in its similarity to the way in which human knowledge is expressed [115], also the convenience with which knowledge can be mapped into such a scheme [125]. Other knowledge representation formalisms are frames [110], semantic nets [110], and predicate logic [110, 126].

DISTRIBUTE-MB-DEVICES-3	
IF:	The most current active context is distributing massbus devices & there is a single port disk drive that has not been assigned to a massbus & there are no unassigned dual-port disk drives & the number of devices that each massbus should support is known & there is a massbus that has been assigned at least one disk drive and that should support additional disk drives & the type of cable needed to connect the disk drive to the previous device on the disk drive is known
THEN:	Assign the disk drive to the massbus

Figure 5.2 Transliteration of a rule in the XCON system [110].

5.3 KNOWLEDGE-BASED SIGNAL UNDERSTANDING

Knowledge-based systems find application in many areas [127] where the manipulation of symbolic information is called for. This includes, for example, intelligent tutoring [128], law [129], risk assessment in finance [130], and medical diagnosis [110,131,132]. An area to which KBSs have recently been applied — and increasingly so since then — is in the simultaneous manipulation of numeric and symbolic descriptions of data [133-135] — this is the realm of the knowledge-based understanding of signals and one which has particular relevance to the processing and analysis of foetal phonocardiograms.

Signal understanding is defined to be:

the process of constructing, using a suitable vocabulary, an explicit description of the salient qualitative and quantitative features of a signal relevant to a particular application. [136]

Signal understanding differs from signal processing in that the result of the operation is not another signal which has been suitably enhanced, but rather a symbolic description

of the information conveyed by the signal, or the environment from which the signal originates. For example, in the former case, the production of the written text equivalent of a speech input signal. The understanding of signals, as Nii [108] points out, often requires using information not present in the signals themselves. This external information is the domain knowledge referred to in previous sections. It is this domain knowledge which ultimately makes the signal understanding problem tractable especially in noisy environments. Domain knowledge may also be applied to guide data-level signal processing operations as to the type of signal event sought and its likely region and time of occurrence. This facility may be used to particular advantage in real-time signal understanding systems where domain knowledge can optimally deploy the use of computationally intense algorithms.

The understanding of a signal — the signal-to-symbol transformation [137] — must necessarily be accomplished over various levels of abstraction (figure 5.3). On the lowest level there exists the unprocessed signals from the transducers, and on the highest level a symbolic description of the synthesized information derived from the signal and knowledge from the domain; for example, in a speech-to-text system the abstraction hierarchy might be: speech signal, syllables, words, phrases [138].

A very flexible framework for accomplishing this signal-to-symbol transformation, which is widely applied, is a modification of the general structure of the classical KBS illustrated in figure 5.1; this is known as the blackboard model [139, 140].

5.4 THE BLACKBOARD MODEL

The blackboard framework (figure 5.4) was originally developed for the signal understanding problem in the Hearsay-II speech-to-text system [138, 141] and was also

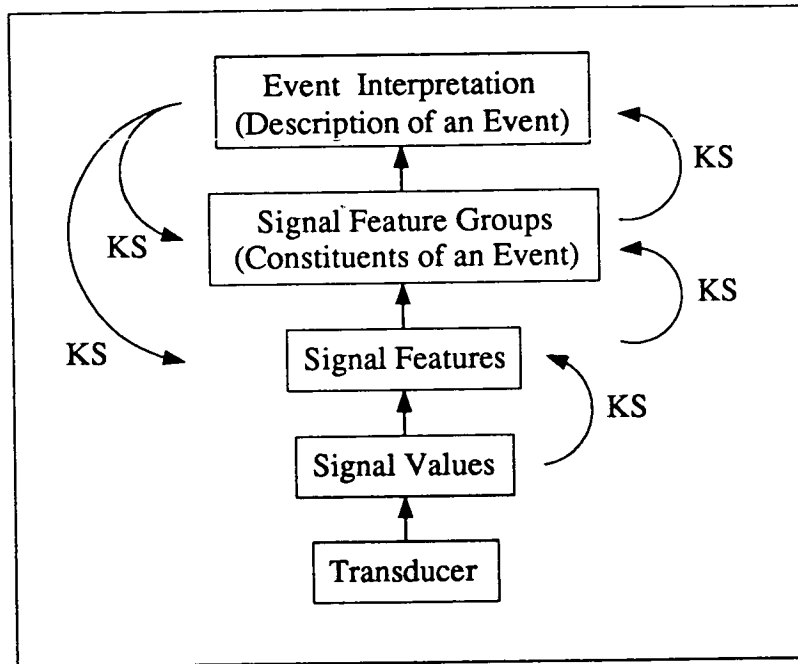


Figure 5.3 Levels of abstraction in the signal-to-symbol transformation hierarchy. Knowledge sources (KS) transform elements between levels.

applied to sonar interpretation for ocean surveillance in the HASP/SLAP project [137]. Although the blackboard model has been proposed as a general-purpose problem solving framework, it is particularly suited for the manipulation of signals. This aspect will be discussed after the components of such a system have been briefly described.

5.4.1 Knowledge Sources

The computational model of figure 5.4, in contrast to the classical KBS, has the knowledge base partitioned into separate and independent modules containing information entities that naturally belong together [142]; these are termed *knowledge sources* (KS). Thus, each KS may be regarded as a local expert in some aspect of the problem decomposition. A KS has a separate inference engine incorporated into it which has independent access to the global data base. This allows the system to accommodate different inference mechanisms and knowledge representations as the

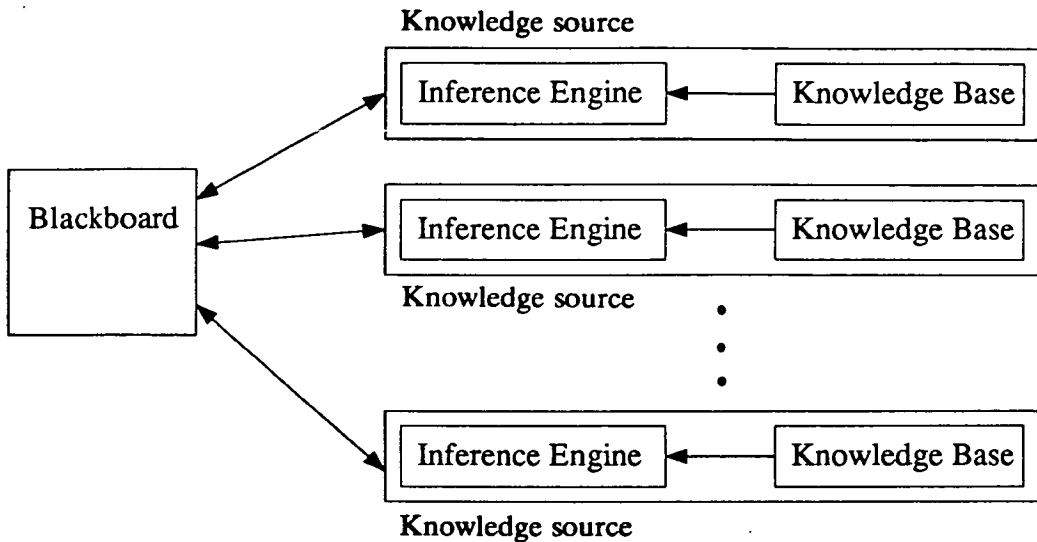


Figure 5.4 The blackboard model. \longleftrightarrow represents data flow.

internal details of KSs are hidden from one another. KSs are self-activating insofar as an individual inference engine determines the applicability of the corpus of knowledge it controls to further the problem-solving process.

5.4.2 Global Data Base

The data base itself is divided into regions so that there is no longer a requirement for a homogeneous representation of entries — this data-holding structure is known as the *blackboard*. Communication and interaction among the KSs occur only through the entries posted on the blackboard. Each KS takes its input from one region of the blackboard, transforms the information using its knowledge, and outputs into another region. In signal understanding tasks these regions are generally levels of abstraction of the sensor signal in an analysis hierarchy and may contain partial or intermediate results.

5.4.3 Control

In the blackboard model no control component is specified, as each KS is intended to respond 'opportunistically' to changes on the blackboard. Practically, however, there is need for control to 'focus attention' [143]. This is the process whereby the system determines how its resources should be allocated to advance the solution. The focus-of-attention can either be (a) which KS to activate from a list of potential contributors comprised of applicable KSs [138], or (b) which part of the unfolding situation on the blackboard to tackle next [137].

A solution to the focus-of-attention problem has been proposed by Terry [144,145] in the CRYNALIS blackboard system which infers the three-dimensional structure of proteins from X-ray crystallographic data. In CRYNALIS the control element has itself been represented as KSs on what is termed the meta-knowledge level [146]. Generally in KBSs the use of domain knowledge is dictated by the inference engine which is usually domain-independent. However, this scheme neglects the fact that experts possess knowledge about the application of their knowledge, that is, domain-specific control knowledge. This meta-knowledge has been included in CRYNALIS as explicitly stated control sequences.

Control in the CRYNALIS system is organized hierarchically (figure 5.5) on two levels: strategy or policy level, and task level. The strategy KS and the task KSs are on the upper and lower levels respectively. The strategy KS has overall control of the system and directs the task-level KSs to accomplish some function on particular regions of the blackboard. The task KSs operate under the control of the strategy KS and guide the application of the non-control KSs which act directly on the data within the context of the chosen strategy.

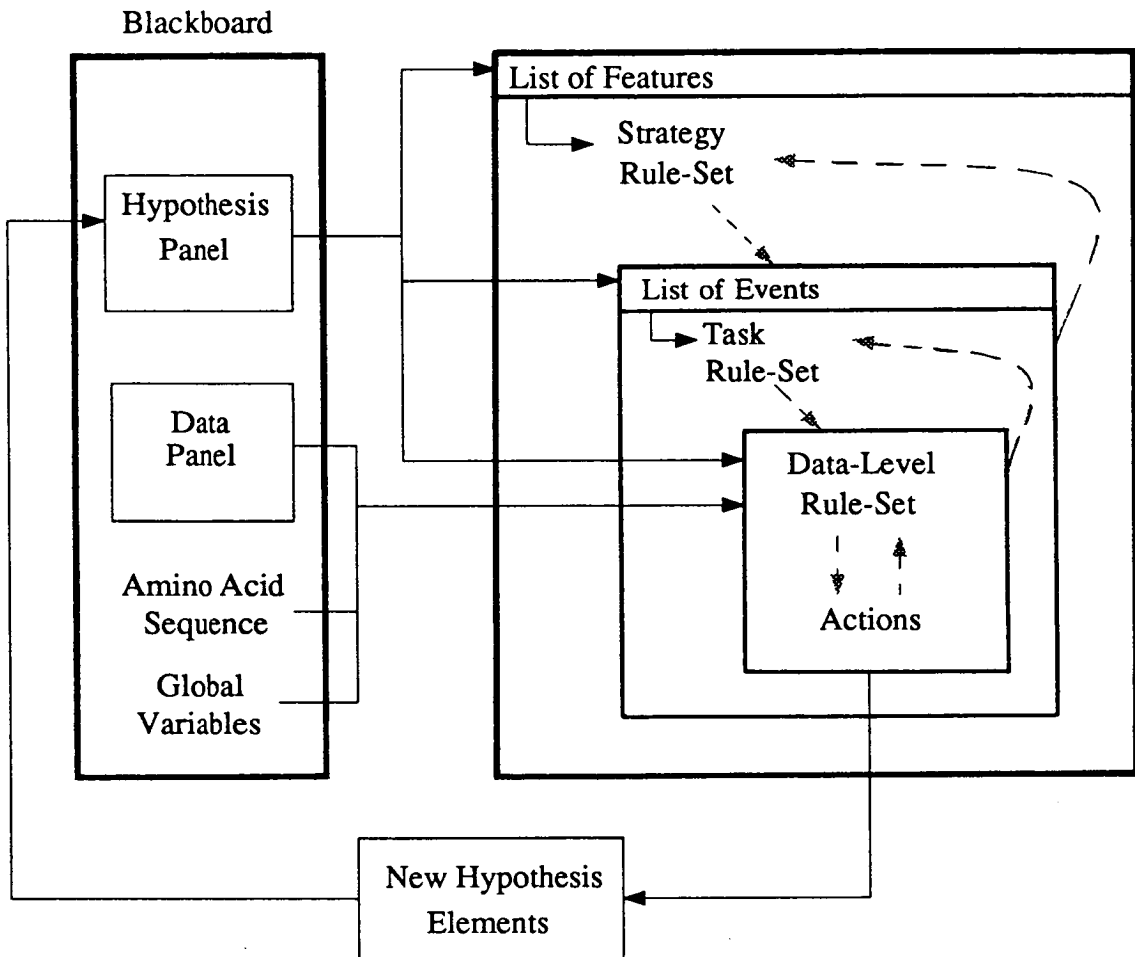


Figure 5.5 Control flow in CRYBALIS. Each of the nested boxes on the right-hand side is a control knowledge source. The data-level rule-set changes the blackboard by creating hypotheses. The other rule-sets invoke rule-sets on the next level down. - - - represents control flow, — represents data flow.

Control originates at the strategy level and proceeds down the hierarchy. The strategy level elects one or more of the task-level KSs to execute some policy. The power of control now resides in the selected task KS, and the strategy KS is suspended. The task KS will autonomously attempt to realize its function until it either succeeds or fails whereupon, in either case, control resumes to the strategy KS. When the task KS causes an applicable data-level KS to fire, control will pass down to this — the lowest level. When the data-level KS reports back to the task level, it also transfers control. This brings the task KS out of suspension. This passing of control between the two lowest levels will continue until the task has been accomplished or it terminates its execution. The power of control will then return to the strategy KS which then shifts attention to a new task.

5.4.4 Signal Understanding and the Blackboard Model

The blackboard framework is well suited to the problem of signal understanding for the following reasons. Firstly, the blackboard supports multiple levels of abstraction which is a requirement of the symbol-to-signal transformation process. Within these levels, there is no restriction that the recognition of signal events proceeds causally, *i.e.* in time sequential order; rather, events are sought in areas of minimal ambiguity from where causal and non-causal processing may take place. Secondly, the KSs allow many different types of knowledge to be represented and particularly in this context KSs which are numerical signal processing routines. Thirdly, the flexibility of the control structure supports the integration of model-directed and data-driven reasoning strategies which are central to signal understanding. Model-directed reasoning, coupled with the signal abstraction hierarchy on the blackboard, allows the generation of expectations for events on lower levels from what has been inferred on a higher level. In the noisy signal environments encountered, model-directed interpretation of this kind can significantly enhance the results obtained despite poor signal-to-noise

ratios. Referring to the HASP/SLAP system, Nii [137] has extolled the benefit of knowledge-based reasoning in such situations

It makes little sense to use enormous amounts of expensive computation to tease a little signal out of much noise, when most of the understanding can be readily inferred from the symbolic knowledge surrounding the situation. ... Sensor parameters can then be 'tuned' to the expected signals and signal directions; not every signal in every direction needs to be searched for.

Conversely, the control structure also supports data-driven reasoning, *i.e.* inferred from the input data, which is employed to suggest or reinforce hypotheses on higher levels.

5.5 KNOWLEDGE-BASED UNDERSTANDING OF PHONOCARDIOGRAMS

The preceding sections have presented the structures and mechanisms with which human expertise may be captured and used by a computer. Now the specific problem is addressed of how this methodology is to be applied to the problem of finding the first and second heart sounds in the foetal PhCG. Two systems, based on the KBS principle, have been developed to achieve this goal. The first system (system 1) is described in the following chapter, and the second system (system 2) in chapter 7. System 2 originated in the lessons learned from the building of system 1 and represents a significant improvement in the standard of analysis achieved.

Although there are distinct differences in the two systems, there is sufficient commonality in the matters of knowledge acquisition and representation, system architecture, and the dynamic application of knowledge to allow a description of these to be recorded in one place.

5.5.1 Knowledge Acquisition

The knowledge acquisition stage of the building of both KBSs is unusual insofar

as it did not involve direct input from what is conventionally thought of as a 'domain expert'. Knowledge of the domain was acquired by the less direct, but equally valid means of consulting the extensive literature relating to adult and foetal heart sounds.

The reason why this proved to be necessary has to do with the way obstetricians determine the foetal heart rate. This task, in the majority of cases, is accomplished by direct auscultation with a stethoscope and consists of counting the occurrences of a particular sound over a certain interval [26, 73]. Thus, he who would initially appear to be the natural choice for domain expert, is an expert in the aural processing of the heart sound signal rather than in its visual processing when presented as a PhCG. In addition, the information presented in a PhCG contains significant infrasonic components, which implies that there is not a direct correspondence between the aural and visual sensory perceptions. However, auscultation has a place in the knowledge acquisition as it provides the bounds on the operational characteristics of the foetal heart and various systematic rather than detailed parameters. Auscultation has yielded such values as the limits of the FHR, the range of beat-to-beat variability of the FHR, and the relative durations of systole and diastole [75]. This knowledge is enlisted whenever a PhCG is proving difficult to analyse visually and in these circumstances quantitative rather than qualitative methods need to be used.

The literature relating to heart sounds which is referred to above, especially Kelly [75] and Maki [147] presents a catalogue of exemplars of the various manifestations of foetal heart sounds extracted from PhCGs. These exemplars are intended to be an aid to *auscultation* in 'visualizing' acoustically complex sounds. These sources do not address the problem of isolating the heart sounds from the PhCG which, it would be tacitly assumed, would not pose any difficulty especially in the relatively noise-free PhCGs they chose. However, this textbook-type information is sufficiently complete

to act as a 'domain expert' whenever such assistance is required.

It is not a prerequisite that one should be skilled in cardiology or obstetrics to be able to identify two repetitious, but non-periodic, signal events in the foetal PhCG. In fact, having established the fundamental nature of heart sounds and their production (chapter 2, section 2.5), a layman could then analyse a PhCG with a high degree of confidence. Even in the presence of adventitious sounds and high levels of noise, a visual inspection can locate and classify the foetal heart sounds from among many competing events. How does the eye achieve this? Certainly there are latent cognitive processes at work in discerning the heart sounds in the PhCG, which are general faculties rather than specific to this application and which no amount of introspection could reveal. What is apparent, however, is that the eye fixes upon certain salient features: higher amplitudes, event homogeneity, event/background dissimilarity and structure, and particularly the repetition of the principal events across the PhCG. These factors have been incorporated into the signal processing KSs of both systems, but are more developed in system 2 where no conventional numerical signal processing was used.

The heuristics employed in the low-level signal processing stages had to be assembled from a basis of little *a priori* knowledge. This is attributable to the rapid decline in the use of phonocardiography in obstetrics since the advent, in the early 1960s, of Doppler ultrasound methods for foetal heart rate monitoring. Besides, even before the introduction of ultrasound techniques, visual examination of long records of PhCGs would have been rare, and these occasions have not been documented in the literature. From examination of many PhCGs recorded in various subjects of differing gestational age, certain characteristics emerge from the data regarding the heart sounds. This process was carried out by the system developer himself, who

subsequently became the 'domain expert'. This phenomenon of the knowledge engineer himself becoming his own domain expert is not unknown [112,120]. The accumulation of expertise in this way was a painstaking process which was very time-consuming, but in the absence of an alternative, it was the only recourse. It has the advantage, however, that knowledge elicitation is easier than would be the case if the expert were someone other than the system developer himself.

5.5.2 Knowledge Organization and Representation

The knowledge organization scheme used in the KSs of both systems is in the form of decision or knowledge trees [148,149]. This is an example of a *compiled* knowledge base where the initial knowledge representation using production rules has been transformed into another representation that has greater computational efficiency in terms of speed and memory space [112,150]. The choice of a decision tree format was motivated by a number of reasons, but primarily the representation suggested itself because of the naturally procedural structure of the domain knowledge, the constrained solution space of the domain, and the completeness and determinism of the knowledge. Completeness and determinism imply that there are no gaps in the knowledge, and that only one piece of knowledge is applicable at any given time.

Declarative knowledge representations are important when the knowledge base is ill-defined and will need constant revision. On the other hand, procedural knowledge representations are used when the task is well specified and efficiency is required [114]. Often the knowledge that experts possess is most naturally represented by sequences of operations rather than by independent sets of rules in which the order of application is not explicitly expressed [151]. This kind of knowledge is awkward to express in a declarative way [152].

The decision tree is constructed from the production rules that define the domain. In effect such a realization builds the search space explicitly, rather than representing the tree implicitly in the rules and then allowing the search strategy to generate explicitly only those parts that are to be explored. The following set of production rules:

if C1 and C2 and not C3 then A1
if C1 and C3 then A2
if C3 and not C1 then A3

may be equivalently represented as a decision tree of tests and actions with the following structure (figure 5.6):

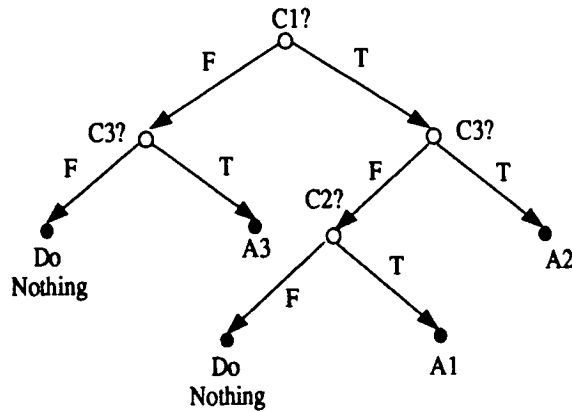


Figure 5.6 Decision tree.

In the decision tree the decision process begins at the root node (C1) and proceeds up the tree along single branches. The path taken depends on the logical outcome of the predicates encountered at non-terminal nodes (hollow circles), and ends when a terminal leaf node is reached (solid circle). The resultant action taken is that specified at the particular terminal node and is a reflection of all the predicates encountered up to that point.

This formulation, in essence, presents the concept of a procedural knowledge-

based system in which procedural knowledge is explicitly represented [151-154]. Such a system has its knowledge grouped into knowledge areas (KA) which are similar to the KSs described previously except that they hold 'large amounts of domain-specific procedural knowledge' [155]. These KAs consist of an *invocation part* and a *body*. The invocation, which resembles the antecedent of a production rule and may itself be another KA, is used to trigger the KA. The invocation expression may thus be viewed as domain-specific meta-knowledge (section 5.4.3). The body of the KA, which performs a function similar to the consequent of a production rule, is then executed if the invocation evaluates to *true*. The body of the KA is a procedure that establishes sub-goals to be achieved and draws conclusions on the basis of achieving or not achieving these sub-goals.

Extending the knowledge base of traditional production-rule systems is a matter of adding new rules; whereas extending a particular decision tree is not such a straightforward task. This arises because the rules in the latter case are closely coupled, *i.e.* there is an order of application to correctly achieve the desired result. However, although production systems are apparently modular, rules cannot be added into the knowledge base without strict attention being paid to possible interactions and inconsistencies with rules already resident. This is an example of coupling which is not explicit and is a situation which may give rise to unforeseen consequences [155].

The advantages of the decision trees structure are: (a) it is a computationally efficient way of implementing a decision making process based on static control sequences [156], (b) it facilitates the determination of the consistency and completeness of the knowledge-base, (c) it is more reliable and demands less searching than implicit decision trees [155], and (d) such algorithmic reasoning is difficult to implement in a declarative framework [152,155]. The disadvantages of the decision

tree representation are that it can take long to generate and may be complex [157]. Both of these factors arise because the developer has to explicitly anticipate all eventualities that the system might encounter and when such occasions might arise.

Such decision tree structures have found application in many areas [158] from fault diagnosis [159, 160] to real-time control systems in the NASA space shuttle [154].

5.5.3 Signal Abstraction and Knowledge Application Strategy

There exists two levels of data abstraction which are maintained in the data-base along with the primary data. The lower level of the abstraction consists of the signal *events*[†] that are considered to be *significant*[‡] and the higher level these same events — segmented or partially reconfigured where applicable. Once the primary data has been processed it will not be revisited. However, it is retained for records of possible problem areas and sounds to await later investigation by the domain expert.

The processing of a PhCG takes place on and between three levels on a scale of increasing data abstraction. On the lowest level there exists the signal processing, followed by the morphological processing, and on the highest level the semantic processing or analysis. The information arising directly from or derived from each level is passed down from the higher levels to the lower as expectations or is passed up from the lower to the higher levels in the form of hypotheses.

The main difficulty in classifying a signal event — once certain features of the signal have been clustered to form an event — as either heart sound or noise, is the non-uniqueness of event attributes across both data records and gravidae. Two events possessing identical characteristics may, depending on the context, in one instance be

[†]The definition of what constitutes an *event* differs in system 1 and 2.

[‡]The measure of the significance of an event also differs in both systems.

noise and in the other a valvular sound. This ambiguity cannot be resolved by further signal processing operations alone. What is required is an analysis by a synthesis of every source of information available about the events that are to be analysed. Such information derives from the context within which the signal events are embedded and knowledge about the domain. It is for this reason that the data is captured in sufficiently long time frames to allow the temporal constraints imposed by the mechanisms of principal heart sound production to be fully exploited. However, to take full advantage of the context, the analysis must lag behind the processing and may only begin once the complete time frame has been captured.

The strategy used to locate and identify heart sounds in the PhCG is the one which a domain expert would use and is incorporated into the highest level of the control structure of both systems. This approach involves a data-driven examination of the PhCG to find areas where, with no ambiguity, it is possible to identify certain of the significant signal events as being of valvular origin. Once these 'solution islands' or 'islands of certainty' [139] have been identified, the goal is to expand outwards from these 'islands' using a model-directed strategy into those regions where the principal heart sounds are not so immediately obvious.

The model-directed expansion is effected using the symbolic information in figure 2.5 and the physical limitations of the sound-generating process, *e.g.* the extremes of FHR. This information serves as a constraint in the search process and so reduces the number of possibilities that must be considered at any given time. This process is also aided by the redundancy of information inherent in the repetitious but non-periodic generation of the principal heart sounds over the PhCG.

An example of the power of model-directed reasoning in interpreting noisy data is illustrated — for the case of the human cognitive process — in figure 5.7. Here a

partially obscured sentence (equivalent to noisy data) may be understood by using contextual information (missing main verb and semantic clues such as 'ear' and 'sound'). It is apparent from this example that model-directed reasoning need not occur in a time-sequenced order. Information from the solution islands lying on either side of the ink spillage both contributed to the interpretation of the obscured word. This facility was also included in the problem-solving strategy of both knowledge-based systems.

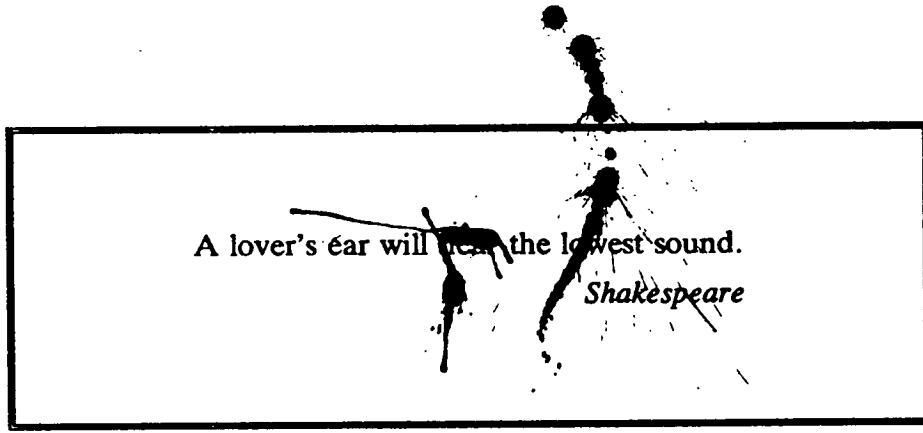


Figure 5.7 Example of model-directed reasoning in analysing noisy data.

Ultimately the expansion from the solution islands will, depending on the quality of the data, result in the merging of the expansions from individual solution islands and the complete analysis of the PhCG. Although it is not always possible to achieve a complete analysis, a partial analysis, *i.e.* resolution of certain areas of the PhCG, will still provide the required parameters of cardiac function albeit over a shorter time interval. If a region should prove to be difficult to analyse because of very high levels of noise or many missing expected events — it is abandoned. For the sake of confidence in the analysis it is better to concentrate computational effort in the more easily resolved regions where solution integrity is high than where there is an increased possibility of an incorrect assessment.

5.5.4 System Architecture and Implementational Framework

The architecture adopted for both system 1 and 2 is illustrated in figure 5.8. A modified version of the type of control system used by CRYSLIS implements the knowledge application strategy of solution island expansion. The strategy KA marshals the control sequence of task KAs to expand the analysis of the PhCG from the already identified heart sounds into those regions of the signal where the decomposition is not so immediately obvious. Each task KA reexamines the global data base on each control cycle and selects where to apply the data-level KAs. The control system as shown, implements solely the island-driving approach. This strategy is not altered during an analysis and is dogmatically applied such that if the expansion of one solution island is irretrievably terminated, that region is abandoned and the system searches for another solution island to expand.

Control messages communicated between levels in the structure consist of a binary valued statement signifying either *true* or *false*. *True* if the intended action succeeded in achieving its purpose or *false* if the KA failed to trigger or failed in its objective.

The global data base is stratified on three levels as described in the preceding section. It is realized as arrays of data structures. Each structure is divided into elements which represent some attribute of the event to which it refers, *e.g.* start time, end time.

The system has been realized in the 'C' programming language running under UNIX on a Sun 3/80 workstation. A realization in C rather than a declarative language (LISP, Prolog) incurred no disadvantage as the knowledge held in the KAs had found a natural expression in a procedurally orientated and tree-like structure.

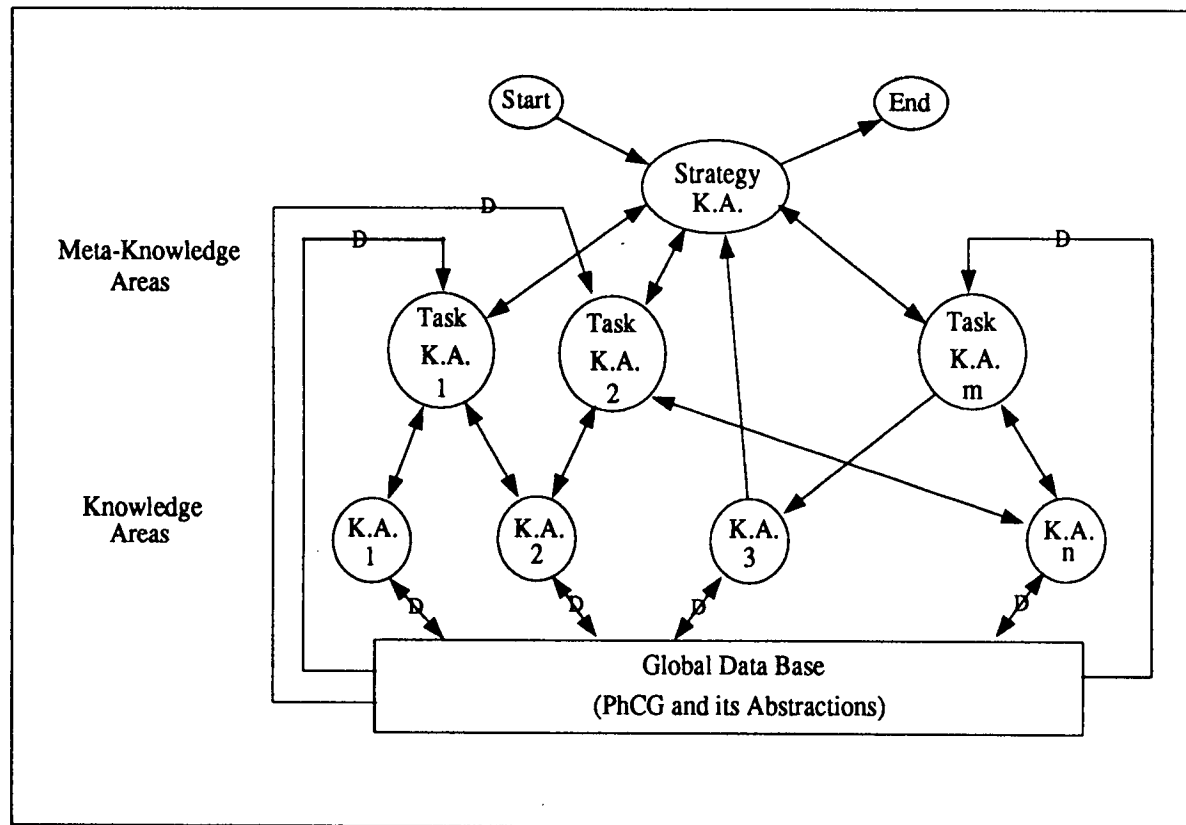


Figure 5.8 Architecture of the phonocardiographic signal understanding system. Lines marked:—D—are data lines, all others are control lines.

Indeed, the fast and efficient executable code produced by the C compiler undoubtedly made the system run much faster than a realization in an AI-specific language. The time taken by the system to analyse PhCGs is an important consideration in such an application as the amount of processing required for long-term continuous monitoring is considerable, and the nature of the task, *i.e.* warning of foetal distress, needs speed of response.

5.5.5 Evaluation of System Performance

The only standard against which an analysis of a PhCG produced by either system may be judged is a visual comparison with that achieved by the domain expert himself. This arises because there is no other technique presently available which can automatically detect individual principal heart sounds either with phonocardiography or ultrasonography. A quantitative assessment of system performance on a simulated PhCG with modelled adventitious and spurious sounds is unattainable for the very reason that a knowledge-based approach is required in this application, *i.e.* the PhCG is too complex to be represented mathematically.

5.6 SUMMARY

Processing the PhCG on the basis of the information content of the signal alone is insufficient to detect the first and second heart sounds because of the low signal-to-noise ratio, and the temporal variability of heart sound characteristics. Nevertheless, a visual inspection of the PhCG can effortlessly succeed in isolating the events representing valve closures from a background of adventitious sounds and noise when the signal-to-noise ratio is good. It is proposed that the cognitive processes are synthesizing other sources of information into the analysis task rather than relying solely on the information contained in the signal. This additional symbolic

information is composed of contextual clues from the PhCG and background knowledge of the domain.

To incorporate these sources of external knowledge into the processing of the PhCG, a system design using the knowledge-based methodology is presented. The structure and implementation of such knowledge-based systems is outlined and also the task of knowledge acquisition from a domain expert and the representation of that knowledge in a machine usable form.

Two systems were subsequently developed to apply these ideas to the analysis of the PhCG. Both systems are similar on a general level, and these areas of commonality are described, namely: architecture, knowledge representation, and knowledge application. The specific details of both systems are introduced in chapters 6 and 7.

CHAPTER 6

PHONOCARDIOGRAM PROCESSING AND ANALYSIS: SYSTEM 1

Although the inclusion of knowledge into a signal understanding system will undoubtedly yield better results than a processing component alone, it is obvious that effective processing can greatly facilitate the task of analysis and produce more accurate results overall.

The processing aspect of the system consists of every component, whether hardware or software, in the chain from the first point of contact with the signal at the transducer to where the data is transferred to the signal analysis sub-system. The wide bandwidth transducer [1] used in the signal acquisition phase is a significant improvement on previous designs, and is assumed to provide the best possible phonographic signal. To produce the phonocardiographic signal from this transduced signal is a matter of high-pass filtering at some cut-off frequency f_c . The prevailing consensus at the time when system 1 was being developed (1988) was that f_c should be set at between 40 Hz and 50 Hz (chapter 3, section 3.7). Indeed, the inventors of the TAPHO transducer used frequency cut-offs in this range in all their published work on the subject [1, 48, 55, 57, 59]. Thus, a filter threshold frequency of 40 Hz is used with system 1 as there seemed no reason to question this assumption. However, hindsight would suggest the use of a high-pass filter with a lower filtering threshold to

fully exploit those spectral components of the FHS which lie below 40 Hz.

All subsequent processing of the PhCG reflects the initial assumption about the value of the filter threshold. The knowledge contained within the non-numerical parts of the processing sub-system is specifically tuned to the PhCG which is produced by high-pass filtering at 40 Hz. Adjusting the filtering threshold to any other value would introduce features into the PhCG specific to that threshold. It is not possible, therefore, to expect the system to have the same performance irrespective of the spectral content of the input signal.

Although system 1 was contending with a non-optimum signal, it performed very well in terms of detecting first and second heart sounds in a signal with a lower signal-to-noise ratio than might otherwise have been, had a lower filtering threshold frequency been used (chapter 3, figure 3.3). This ability to counteract the inadvertently lower information content of the signal is attributable to the synthesis of knowledge about the signal events and the mechanism producing them, and is an indication of the capability of this methodology.

The knowledge encoded in the processing and analysis sub-systems has been accumulated from empirical observation of many instances of heart sound manifestation, and the phenomena and the anomalies to which the PhCG is susceptible. These observations derive from a sample set of 252 PhCGs from nineteen subjects (chapter 3, section 3.6). Unless stated to the contrary, all the rules given in the following sections have originated from these accumulated observations.

This chapter describes the overall PhCG understanding system as divided into its two constituent sub-systems of processing and analysis. Results are presented and discussed which illustrate system performance over the spectrum of PhCG record types

ranging from relatively noise free, to heavily noise contaminated. The problems created by the omnipresence and morphology of adventitious sounds for this, or any analysis system, are also described.

6.1 PHONOCARDIOGRAM PROCESSING

The first step in the processing of the phonographic signal is to extract the phonocardiogram from it. This operation effectively separates the foetal infra-sounds from the valve-closing sounds. As the spectra of these two phenomena are sufficiently non-overlapping, they may be separated using a linear frequency filter. The frequency at which to set the threshold of this filter has already been addressed in chapter 3 and in the introduction to this chapter. For the historical reasons outlined, all PhCGs illustrated in *this* chapter were produced by a 21 coefficient, 40 Hz high-pass, digital FIR filter. The linear phase characteristic of such a filter type was required to preserve the relative phasing of spectral components in the signal to allow accurate timings between signal events.

It is possible that the PhCG, even at this stage of minimal processing, could form the input to an analysis system. However, this would be to the detriment of the analysis process because it ignores that other essential element in signal understanding systems which is the ability of processing techniques to reduce the amount of uncertainty in a signal. With effective signal processing, the task of analysis becomes less difficult through the prior enhancement of signal events and the reduction of noise.

In the standardized PhCG record length of 4.1 seconds, it has been observed (chapter 3, section 3.9) that second heart sounds are less shape and structure variant than first heart sounds in the same record. The second sound, however, is much more

susceptible to amplitude variability. A signal processing method which is specifically designed to take advantage of the constancy of shape and structure in an event in relation to other competing events is cross-correlation [161]. Such an approach, in this instance, demands that there is a stored archetypal second heart sound with which to form a template. Defining and finding such an archetypal template is an issue which must be addressed.

Cross-correlation will transform the PhCG into another signal in which certain events will be emphasized. How these events are to be extracted from the signal is the final task of the processing system. It is to be expected that adventitious sounds will, in some circumstances, also be enhanced by this form of processing. It is a requirement that these interfering sounds be kept to a minimum to prevent the subsequent analysis from being overwhelmed by many possible candidates for first and second heart sound status.

These foregoing issues and the details of the implementation are introduced in the following sub-sections.

6.1.1 Cross-correlation

The cross-correlation function, $\phi_{xy}(l)$, relating two signal sequences $x(n)$ and $y(n)$, both of finite energy, may be defined as:

$$\phi_{xy}(l) = \sum_{n=-\infty}^{\infty} x(n) y(n-l) \quad l = 0, \pm 1, \pm 2, \dots \quad (6.1)$$

The index l is a time-displacement or lag imposed upon one sequence, $y(n)$, relative to the other, and the subscripts xy on the cross-correlation function indicate the sequences being correlated. When both sequences are of finite duration, of length N , and causal, the range of the sum of products in (6.1) is altered to:

$$\phi_{xy}(l) = \sum_{n=0}^{N-|l|-1} x(n)y(n-l)$$

However, in the case of the PhCG where the length of the template sequence is different from that of the finite PhCG record, a further modification is required. Suppose the template is $y(n)$, $0 \leq n \leq M-1$, and $x(n)$, $0 \leq n \leq N-1$, is the PhCG where $M < N$. The cross-correlation of these two sequences for positive values of the lag index l is:

$$\phi_{xy}(l) = \begin{cases} \sum_{n=l}^{M-1+l} x(n)y(n-l) & 0 \leq l \leq N-M \\ \sum_{n=l}^{N-1} x(n)y(n-l) & N-M < l \leq N-1 \end{cases} \quad (6.2)$$

The function $\phi_{xy}(l)$ is a quantitative measure of the similarity between the template and the PhCG as the template is shifted through the PhCG.

The cross-correlation equation in (6.2) bears a close resemblance to the finite, discrete version of the convolution integral — the only difference being that convolution requires that the displaced sequence be reflected about the point $n = l$. Furthermore, if the convolution integral is implementing a matched filter [161], the similarity is more evident. Indeed, as the second heart sound is almost symmetric, the cross-correlation may be interpreted as matched filtering. The benefits of the matched filter in signal detection are well known as it is the optimum filter realization when the signal is corrupted by additive white Gaussian noise, and if the noise is not Gaussian, it is still the optimum linear filter [162].

6.1.2 Template Selection

The disadvantage with cross-correlation, in this context, is that the operation requires the prior selection of a template. There are three options available for this.

The first is to use a universal template for all subjects, the second to choose a new template, with operator assistance, in a single subject each time the foetus is monitored. The third option is to start the cross-correlation with a universal template but update it after each PhCG record is processed so that the template tracks the variations of the second sound over the monitoring period. The first and third options have the attraction that operator intervention is unnecessary and both obviate the attendant inconvenience of having to initialize the system each time it is used.

Originally the second option was adopted in the specific case of a single subject. A second sound was chosen from a PhCG which had very low background noise, on the basis that it resembled an ideal realization. As noted in chapter 3, the second sound often has the form of two-and-a-half cycles of a sinusoid, amplitude modulated by a bell-shaped pulse. After having processed many PhCGs originating in the same subject from which the template was drawn, the results were rather poor and often the signal would be more degraded after the correlation than before.

Once the utility of a universal template for a given subject, and by induction, a single universal template for all subjects, had been dismissed because of the unsatisfactory results it produced, the alternative was the automatic template-updating scheme from a viewpoint of both system performance and ease of operation. However, such a facility is not without its difficulties. Foremost, a system must first be in operation — which can identify particular heart sounds automatically — before template-updating becomes a viable proposition. Secondly, what criteria are to be employed in choosing which detected second sounds — if not all — to incorporate into the updating, how the updating is to be effected, and how best are mis-alignment errors avoided when performing the updating. There is also the question of mistaken identifications of second heart sounds which would severely affect the template. To postpone these questions until such times as there was an operational analysis system

was a necessary expedience for both experimental purposes and during system development. Therefore, a second heart sound was manually extracted from each PhCG under investigation on the basis that it was visually the most distinct realization in that record, and most closely resembled an archetypal second heart sound.

Figures 6.1 and 6.2 show examples of an instance from both of the template selection schemes examined. The poorer results with the universal template in a single subject are evident.

6.1.3 Correlator Output

The cross-correlation of the PhCG with a second heart sound template taken from the PhCG under investigation, has the effect of accentuating the second heart sounds and reducing the background noise level (figure 6.2). As figure 6.2 also illustrates, this generic template correlates strongly with the first heart sounds. Such an outcome is to be expected as both sounds share common frequency spectra (chapter 3). A concomitant advantage of using a second heart sound template is that it produces less of a smearing effect than would be the case with a first sound template — if one could be specified. Whenever the template is being correlated with a second sound in the PhCG, the non-zero correlation output — in the noiseless case — extends over the sum of the durations of each event. The use of a template drawn from the second heart sounds, which have a shorter duration than first heart sounds, minimizes this smearing effect. When adventitious sounds are closely adjacent to the principal heart sounds, the output produced by the correlator tends to merge the correlations together such that the distinctions between these events are less clear. Considering that the objective of the overall system is the measurement of inter-event times, it is imperative that the beginning of events is preserved and the smearing effect is kept to a minimum.

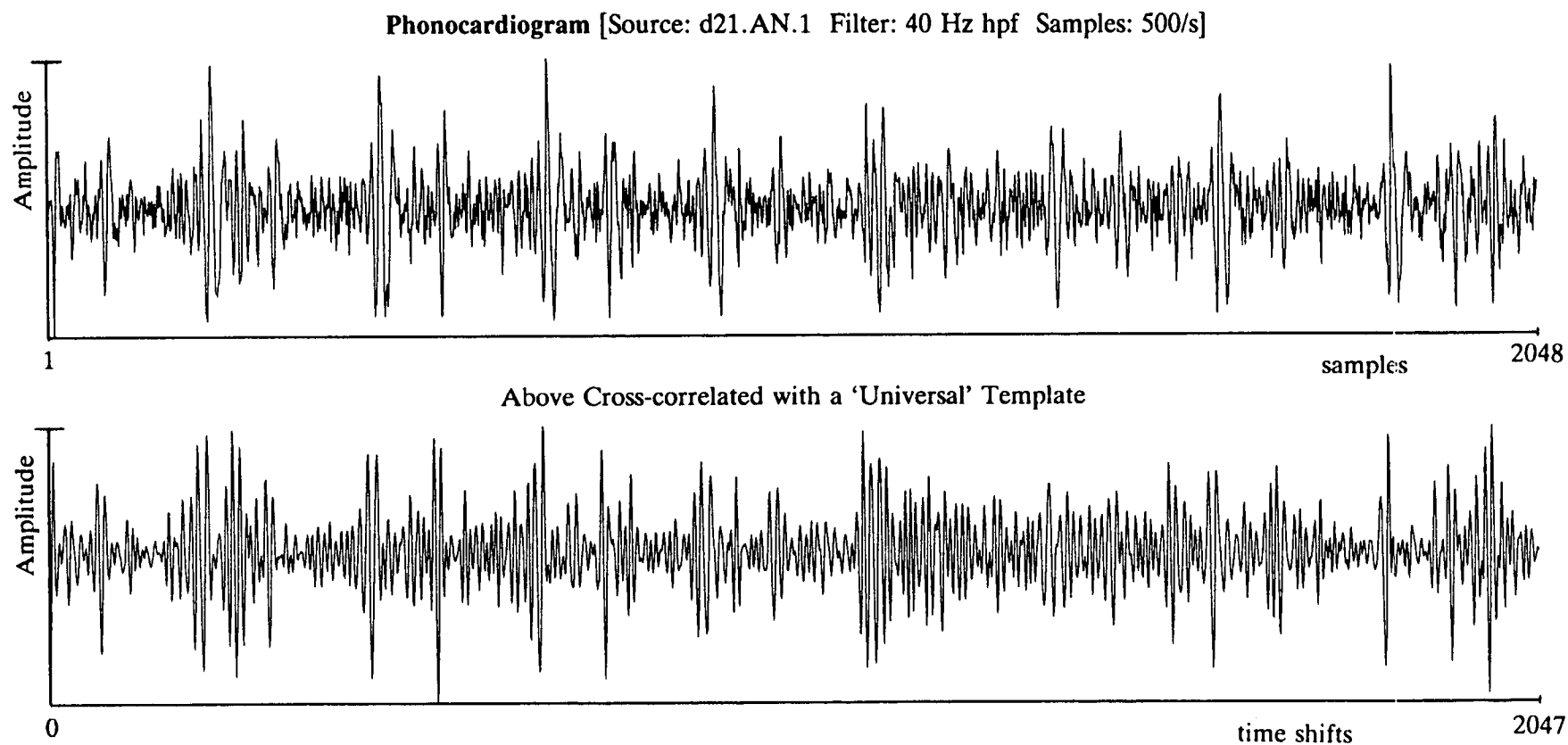


Figure 6.1 An arbitrarily chosen phonocardiogram and its cross-correlation with a second heart sound template which was extracted from an earlier phonocardiogram in the same subject. The poor performance in accentuating the principal heart sounds and suppressing background noise is apparent (compare with figure 6.2).

Cross-correlation of d21.AN.1 with an Indigenous Template

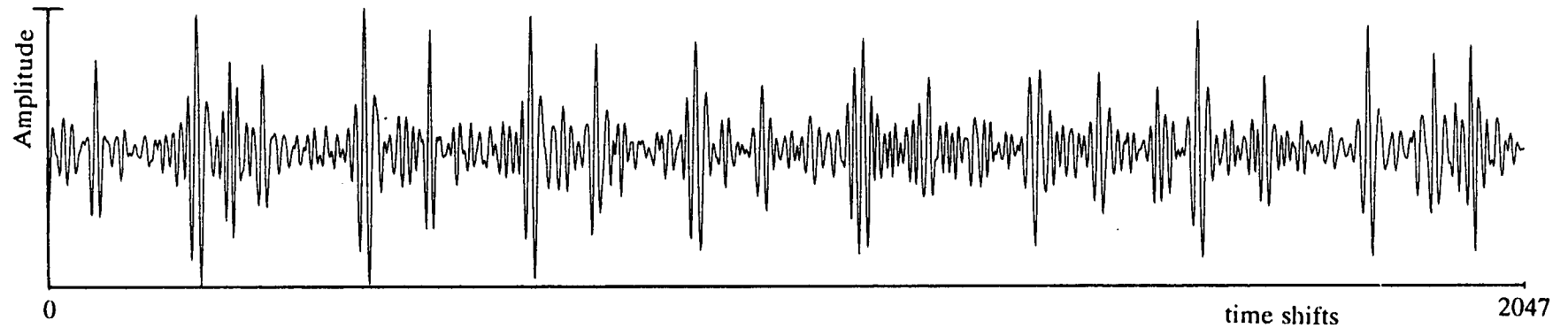


Figure 6.2 There is a marked improvement in the performance of the correlation when the template is drawn from the phonocardiogram itself (compare with figure 6.1). The principal heart sounds are accentuated and the background noise level is reduced.

Phonocardiogram [Source: d2.AN.1 Filter: 40 Hz hpf Samples: 500/s]

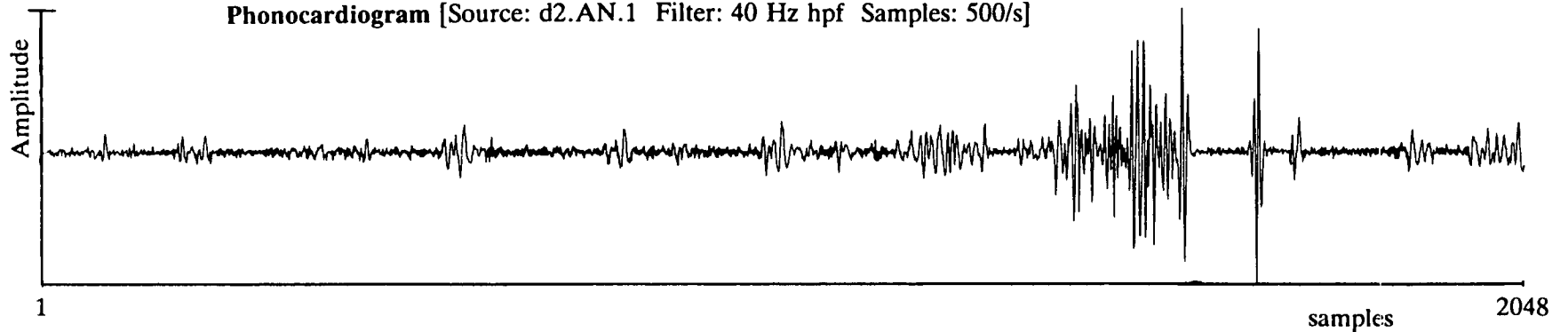


Figure 6.3 The principal heart sounds of the earlier part of this phonocardiogram have been swamped by the burst of intense signal activity which may have resulted from a movement of the transducer or the foetus.

From figure 6.2 it is apparent that correlation also accentuates the adventitious sounds and this effect has been observed across all the PhCGs examined. Such a high correlation would suggest that the adventitious sounds share a common frequency range with the principal heart sounds. There is no way to distinguish between these two very similar groups of sounds at this juncture. All events producing significant correlations must be initially treated as if they were plausible principal heart sounds. The ultimate separation of these events must await investigation by the later stages of processing and analysis.

6.1.4 Further Processing

Having thus accentuated the principal heart sounds of a particular PhCG, the next processing operation is to identify those points in the correlation sequence which indicate a high similarity between the template and PhCG events. To facilitate the search for such occurrences, the correlation sequence is rectified to produce the modulus of $\phi_{xy}(l)$. This modulus sequence is then low-pass filtered (*e.g.* figure 6.7a). A value of 30 Hz was empirically determined for the threshold of this low-pass FIR filter on the basis that the output followed the undulations of the input waveform sufficiently closely. The resulting filtered-modulus waveform is input to the data base of the system as the first level of data abstraction.

It is a common observation in this filtered-modulus sequence that the peaks which correspond to first heart sounds in the PhCG are often fragmented (*e.g.* figure 6.9a). This is due to the longer duration of the first heart sound relative to the correlating template, and variations in first heart sound homogeneity.

To find the position of the positive turning points, and positive-going points of inflexion (for brevity, both will be referred to as peaks) in the filtered-modulus

sequence, $s_{xy}(n)$ $0 \leq n \leq N - 1$, the following criterion is applied:

$$s_{xy}(n) \in \text{'peak'} \text{ if } s_{xy}(n-1) < s_{xy}(n) \geq s_{xy}(n+1) \quad 1 \leq n \leq N-2 \quad (6.3)$$

The equality is required in comparing the present value of $s_{xy}(n)$ with the immediately following one in time to detect those occurrences where the sequence $s_{xy}(n)$ climbs to a plateau. No such equality was applied to the comparison with the point immediately preceding because the determination of the beginning of an event is of more importance than finding the end of it — this also reduces the number of peaks generated. Each peak is represented as an ordered pair which consists of time of occurrence and magnitude. These ordered pairs are assembled into a linear array which forms the second level of signal abstraction in the data base.

6.1.5 Segmentation

By this stage in the processing, the original PhCG has been superseded by the filtered-modulus sequence and the peaks which derived from this sequence. From a combination of these two sequences, the PhCG will be segmented. Segmentation — by analogy with speech processing — describes that operation which groups together the constituents of a sound and determines its terminal points.

From observation of many filtered-modulus sequences from various subjects, it became apparent that those peaks which lie below a threshold level set at one-quarter of the magnitude of the global maximum peak of a given filtered-modulus sequence are nearly always associated solely with noise and interferences. This heuristic also provides a convenient means with which to delimit those peaks which may be regarded as *significant*, i.e. all those above the one-quarter level.

Grouping significant peaks to form a cluster which relates to an event in the PhCG is relatively straightforward except in those cases where other sounds are lying

close by on one or both sides. It is at this point that knowledge-based processing is required to ascertain which peaks belong to which cluster. Once this segmentation is performed, each cluster should relate by a one-to-one correspondence to a single event in the original PhCG. However, before the segmentation knowledge area (KA) is activated, a check must be made on the peaks and their relative dispositions to determine if that PhCG is amenable to analysis.

There are occasions when certain regions in the PhCG exhibit very weak foetal heart sounds which are barely detectable yet in other parts of the record there is strong signal activity. This phenomenon is typical of a large transient, the amplitude of which dominates the duration of the PhCG record and is often the result of transducer or foetal movement. The policy adopted in such situations is that the PhCG is neglected. On other occasions there may be intermittent registrations of heart sound activity in the PhCG. Here the policy is to find if the activity extends over a sufficiently long interval to make an attempt at a partial analysis worthwhile. Knowledge area 22 (ka_22) (figure 6.5) endeavours to ascertain, in an approximate way, the presence of these two conditions.

ka_22 achieves this by dividing up the total length of the signal into sectors of 290 samples (580 ms) each. The lowest 'normal' FHR is 120 bpm [4, 73, 74], thus each block of 290 samples should contain at least one first and one second heart sound, but not necessarily in that order. In the operation which divides the sequence into sectors, the first x samples at the start of the record are neglected to avoid considering the first peak cluster. This cluster of significant peaks must represent the first sound of the record. However, as its position is so close to the record boundary it may in fact actually straddle the boundary, whereupon the beginning of the sound will have been lost. The buffer region of x samples is included to discount such a

possibility which would have low integrity. The value of x is dependent on the presence of a significant peak within the first 48 samples[†] (96 ms). If such a peak does not exist, the start of the first sector is set at 96 samples (192 ms). Otherwise the start is set at the time of occurrence of such a peak plus an additional 48 samples (96 ms). Once the sequence has been sectorised, the significant peaks in each sector must be segmented. ka_22 performs only an approximate segmentation to ascertain whether it is worthwhile continuing with the particular sequence under examination. A more accurate segmentation is left to ka_1 *q.v.* However, the first significant peak encountered in a sector and all succeeding significant peaks up to 48 samples later are considered to form a cluster. The next significant peak beyond the first cluster is found and the grouping is performed again and this procedure continues until the end of the sector is reached. This approximate clustering processes is performed for the length of the sequence. Once this has been completed, for the PhCG to be passed on to the next stage of processing rather than being immediately rejected, there must be a minimum of two contiguous sectors in which there are at least two clusters of peaks. The reason for this requirement will become apparent once the task of finding solution islands has been discussed (section 6.2.2). This approximate measure of suitability in terms of having a sufficient number of clusters of significant peaks across set intervals is used to reject that class of signals of which figure 6.3 is an example.

Once the record has been through this first check it is passed over to the segmentation KA (ka_1) *per se*. ka_1 has access to the list of significant peaks and the processed correlation sequence from which it produces a list of events which are represented by the sole parameter of their respective starting time. It might seem that the strength of the correlation at the significant peak should also be included in a

[†]The value of 48 samples is the average duration of a first heart sound which has been observed over many subjects and PhCGs.

parametric description of an event along with the starting time. However, the strength of the correlation is not a feature which may uniquely be associated with either of the principal heart sounds or the adventitious sounds in conjunction with the principal sounds (figure 6.7a). As previously mentioned, it is in the class of second heart sounds that amplitude variability is most pronounced, along with a tendency to be lower in amplitude than the corresponding first heart sound. It is precisely this tendency, and the amplitude variability evident in all sounds in the PhCG — irrespective of origin — that makes any measure of amplitude or magnitude of correlation unreliable. If this were not so, the task of analysis would be considerably simplified. The segmentation module (ka_1) is the last occasion when amplitude is used as an indicator of (in)significance.

ka_1 clusters the significant peaks progressing sequentially in the direction of increasing time. To simplify the explanation of the workings of ka_1 assume that the focus of attention is sufficiently far removed from the ends of the data record that their affect may be neglected. The first significant peak encountered after the last cluster of peaks is initially assumed to be the beginning of the next event (labelled *event_start* in figure 6.4). The corresponding point is also located in the filtered-modulus sequence, wherein the extent of the sound is provisionally assumed to be 48 samples later[†] (*prov_end* in figure 6.4). The first *significant trough* which occurs beyond this point in the filtered-modulus sequence is then located. A *significant trough* is defined to be the first point after *prov_end* which satisfies:

$$s_{xy}(n-1) > s_{xy}(n) \leq s_{xy}(n+1) \quad prov_end < n \leq N-2 \quad (6.4)$$

providing $s_{xy}(n)$ is below the one-quarter level. The equality in the second comparison in (6.4) is required for those cases when the trough is the leading edge of

[†]48 samples (96 ms) is the average duration of a first heart sound.

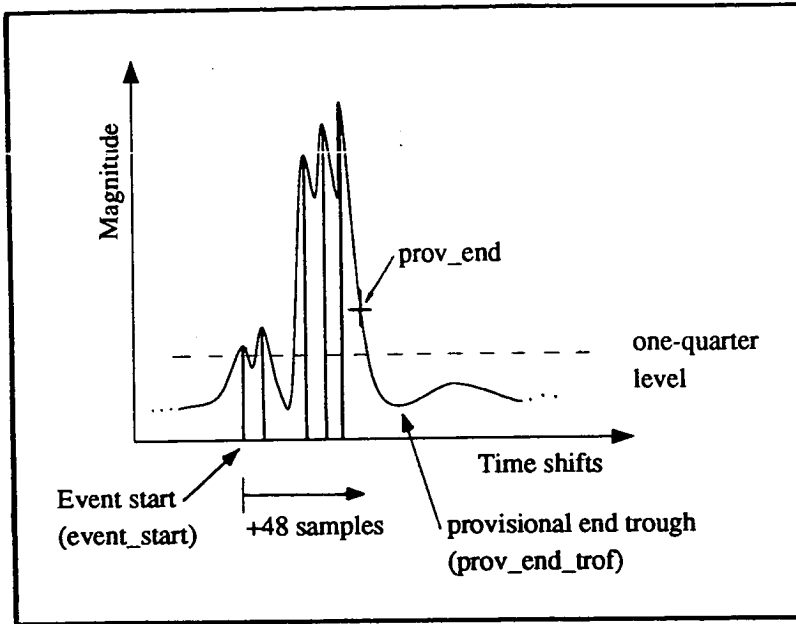


Figure 6.4 Features of a signal event used by the segmentation knowledge area to group the significant peaks in the filtered modulus of the cross-correlation sequence.

a flat-bottomed depression. If the conditions in (6.4) are satisfied, the point $s_{xy}(n)$ is then established as the trough which provisionally marks, as *prov_end_trof* in figure 6.4, the end of the event which starts at *event_start*. The trough at *prov_end_trof* represents the point at which the correlating template has exited from the event in the PhCG, thus the end of the event. Once the terminal points have been established, the number of significant peaks lying between them is counted and the enclosed local maximum significant peak is determined.

This provisional cluster is then tested for homogeneity, that is, the cluster encompasses only one sound event, whether that be an adventitious or principal heart sound. The converse case, that of an inhomogeneous cluster occurs whenever sounds are in such close proximity that they coalesce. However, before examining these latter occurrences, those cases which represent a straightforward homogeneous group are outlined. A cluster is homogeneous if it consists of:

- (a) one significant peak, or
- (b) adjacent significant peaks which are not separated by more than 29 samples *and* a significant valley does not exist between any two of these significant peaks, or
- (c) adjacent significant peaks between which there is no significant valley *and* any significant peak which occurs earlier than the local maximum significant peak is not less than 45% of the magnitude of it, or
- (d) the cluster has not been re-segmented by the procedure for checking for inhomogeneous clusters.

A *significant valley* is defined to exist between two adjacent significant peaks if the sequence $s_{xy}(n)$ satisfies the conditions in (6.4) except that the point $s_{xy}(n)$ must be less than 70% of the magnitude of the immediately preceding significant peak rather than be below the global one-quarter level. Such a criterion has been found to indicate the presence of two events in the PhCG which are so close together that the template is still correlating with the former event when it begins to correlate with the next event. The valley is caused by the low correlation in the short noise-filled gap which exists between the two events.

When an inhomogeneity has been found in a cluster, providing certain conditions hold, there is an opportunity to reject certain of the significant peaks as belonging to an adventitious sound. If these conditions do not hold, the cluster is partitioned into its separate events. The facility which allows adventitious clusters to be deleted is primarily employed to remove late systolic and late diastolic sounds which are the most frequently occurring and the most troublesome. The following conditions outline cases of inhomogeneity in a cluster:

- (a) if the first significant peak in a cluster is separated from the next significant peak in the same cluster by a significant valley *and* more than 29 samples[†] *then* put the first peak into the data base *and* start regrouping from the next significant peak after the significant valley (*i.e.* the initial grouping was composed of two or more sounds).
- (b) if in the cluster there is a significant valley *and* the significant peaks on either side are separated by more than 29 samples (58 ms) then set *prov_end_trof* to the lowest point in the valley *and* re-examine the grouping from the first significant peak of the original grouping to the new *prov_end_trof* (*i.e.* the initial grouping is composed of two or more sounds).
- (c) if all significant peaks in a cluster are not significantly far apart, *i.e.* more than 29 samples (58 ms) *and* there is a significant valley *and* the height of the significant peak preceding the valley is less than 45% of the magnitude of the local maximum significant peak *and* it occurs earlier in time than this local maximum *then* remove it from consideration *and* begin re-grouping from the next significant peak in the cluster (*i.e.* those significant peaks occurring earlier than the significant valley belong to an adventitious sound).
- (d) if all the significant peaks in a cluster are not significantly far apart and two significant peaks are separated by more than 15 samples (30 ms) *and* this occurs earlier than the local maximum significant peak *and* there is a significant valley between these two significant peaks *and* no other of the significant peaks which occur later in the cluster are separated by 15 samples *and* these later peaks are at least three in number *then* remove from consideration the significant peak which just precedes the valley *and* begin regrouping from the next significant peak in the cluster (*i.e.* those significant peaks occurring earlier than the significant

[†] A time separation greater than 29 samples (58 ms) is defined to be significant.

valley belong to an adventitious sound).

An example of the segmentation produced by these rules is illustrated in figure 6.9a.

6.1.6 Phonocardiogram Boundary Conditions

In section 6.1.5 when discussing the division of the record into sectors containing the significant peaks, it was noted that there is a potential problem when events straddle the beginning of a record. This problem is also present at the end of a record although here the start of the event is preserved. However, there is no way to determine if this start is the start of an adventitious sound which lies close to a principal heart sound or a principal heart sound itself. Rather than risk a possible incorrect analysis at the boundaries of the record, the first homogeneous cluster is always removed, and likewise the last cluster if its beginning is closer than 70 samples (140 ms) to the end.

6.1.7 Limits on the Number of Peak Clusters

ka_4 now checks the number and relative positions of the peak clusters to ensure that there are not too few groups which would indicate the presence of some anomaly in the PhCG. At the other extreme ka_4 also checks that the number of clusters of peaks is not prohibitively large which would indicate that there is too much uncertainty in the PhCG for the analysis system to cope with. Such a condition often results from the presence of transients in the PhCG whose appearance obliterates all trace of the foetal heart sounds. These transients are mostly attributable to movement arising from any one or all of: maternal, or foetal body movements, or transducer re-positioning — whether intentional or accidental. In these cases it is best to abandon the PhCG and wait for the transient to pass before recommencing.

Excluding peak clusters produced by adventitious sounds, the number of clusters is proportional to the FHR and the duration of the PhCG. This dependence allows a bound to be placed on the maximum number of clusters that are amenable to analysis. The maximum 'normal' FHR is 160 bpm [4, 73, 74], however, to provide a margin for error this will be raised to 175 bpm in all subsequent analysis. Thus, in a PhCG of length N samples where the sampling rate is 500 samples/second, there are $7N/1200$ beats with each beat comprising two principal heart sounds. If the worst case condition is taken, *i.e.* there is one adventitious cluster per beat[†], each beat now has three clusters which implies overall there are $21N/1200$ clusters in the record. If such a case is found that PhCG is rejected.

ka_4 also checks for a missing cluster. As the lowest 'normal' FHR is 120 bpm this gives an indication of the maximum separation between principal heart sounds in a noiseless PhCG. Kelly [75] has empirically observed in 144 foeti that the diastolic time never exceeds the systolic time by more than twice, so the longest separation in a beat is between the second heart sound and the first heart sound of the following beat at the lowest considered FHR. This separation is equivalent to 167 samples (334 ms). If any clusters are found further separated than 167 samples, irrespective of whether they are principal heart sound or adventitious sound clusters, a principal heart sound(s) is missing in that vicinity. A missing principal heart sound is a difficult situation to counter and would reduce confidence in the overall analysis. To resolve this problem, all those clustered significant peaks are removed which lie either earlier or later in time with respect to the missing sound depending on whether the missing sound is above or below half-way in the record, *i.e.* the objective is to retain as much of the PhCG as possible. A notification of this action, and a printed list of all deleted

[†]If there are consistently three peak clusters per beat, that PhCG is just beyond the limit of analysability in system 1. See section 6.2.1.

clusters is then issued to the user.

Having made such an alteration to the information content of the record, this curtailed record is again checked by ka_22 to ensure that the operation did not leave too few significant peaks.

6.2 PHONOCARDIOGRAM ANALYSIS

Once the information extracted from the original PhCG has reached this stage, it is represented as a linear array of the starting times in samples, of each cluster of significant peaks relative to the beginning of the record (figure 6.7a). From this point on, the analysis sub-system takes control of the manipulation of the abstracted data. The processing sub-system may now begin to prepare the next PhCG record for analysis.

6.2.1 Checks Prior to Analysis

For the purposes of the analysis, the FHR is assumed to lie within the range 117-175 bpm[†] which is slightly wider than the accepted 'normal' range of 120-160 bpm [4, 73, 74]. In a similar manner to ka_4 (section 6.1.7), these limits allow a bound to be set around the minimum and maximum number of clusters in a given interval. This will be used to notify the presence of adventitious sounds in an interval but not *where* in that interval.

It is known [46, 75, 163] that the systolic time varies only slightly even over large variations in FHR and that the diastolic time is always longer or just equal to the systolic time. If it is assumed that the minimum considered FHR is 117 bpm, and the

[†]In system 2, the allowable range for the FHR is extended to 80-220 bpm.

PhCG record length — for ease of data manipulation and presentation — is taken to be 2048 samples (4.1 seconds), then each record will contain at least eight beats. This interval will thus contain sufficient contextual clues for the analysis. The assumption of a constant systolic time is further validated by the brevity of the interval over which the processing and analysis is performed. Moreover, the other assumption is made that the maximum number of adventitious sounds remaining per beat after the segmentation is one. This figure was based on observation of the available PhCGs.

The maximum FHR considered is 175 bpm which translates to 171 samples per beat. Thus, within any given region of length just less than 171 samples, there should only be a maximum of two cluster starting times. `ka_11` makes use of this information by sequentially examining groups of three starting times of clusters to find noise-free areas. The presence of an adventitious sound is indicated, therefore, if any more than two clusters are present in the interval. If an adventitious sound is found to be present, a question mark is placed beside that group of three clusters, and the next contiguous block of three clusters is likewise examined. If an adventitious sound is not detected, no action is taken, and the checking procedure is repeated on the following three contiguous clusters beginning at the immediately adjacent cluster.

When nearing the end of the array of cluster starting times, it may happen that the remaining number of clusters may not be the requisite multiple of three. To circumvent this situation, the analysis perfunctorily questions the integrity of the remaining single or double cluster — whether this is justified or not. This boundary condition is resolved at a later stage in the analysis.

The final check, performed by `ka_9`, inspects all those clusters passed by `ka_11` to find if there are too few clusters in a given interval. The system is alerted to such a condition if there are fewer than two clusters in 256 samples. This interval represents

the length of a beat, consisting of the two principal heart sounds, at the lowest 'normal' rate. ka_9 overlaps with the test performed by ka_4 (section 6.1.7), however, it is deployed to find a single cluster between groups suspected of having an adventitious component. These single clusters, to be rigorous, are treated as being suspect on account of there being no available evidence to the contrary at this stage. In other words, those groups of clusters passed *prima facie* as not containing adventitious sounds, may in fact have missing principal heart sounds which go undetected because of the presence of a close-by adventitious sound. Of course, there will be no alteration in the status of all those groups of clusters which are passed by ka_11 to ka_4, and which are not suspected of having an adventitious cluster.

6.2.2 Location and Analysis of Solution Islands

The finding of solution islands and the labelling of their constituent components — the signal-to-symbol transformation — is the most crucial part of the analysis of a PhCG. If no solution islands are discovered then the analysis of that particular PhCG must be abandoned. Solution islands form the bridgehead which allows the decomposition of those ambiguous regions where adventitious sounds are known to be present.

Naturally, solution islands are sought from among those groups of clusters which ka_11 has declared free of adventitious sounds. The search for solution islands, performed by ka_5, makes use of the fact that the diastolic time is always longer than, or just equal to, the systolic time [75]. The systolic time is the interval between a first heart sound and the following second heart sound, whereas the diastolic time is the interval between a second heart sound and the first heart sound of the next beat. The starting time of those clusters which are presented by ka_9 and which are not suspect, must then be the starting times of unidentified principal heart sounds. Thus, if any

contiguous group of three starting times is extracted, and the time difference between the starting times is calculated, these time differences must be, depending on their relative durations, the systolic and diastolic times. If an interval can be positively identified as being systolic, then the earlier cluster starting time is a first heart sound, and the later is a second heart sound. Likewise, whenever an interval is identified as diastolic, the earlier and later cluster starting times must be second and first heart sounds respectively, where the second heart sound is the same one previously identified as marking the end of the systolic phase.

However, there is a caveat with this approach which is that only diastolic/systolic intervals, in that order, may be unambiguously identified. Supposed systolic/diastolic intervals have a potential ambiguity. This arises from the possibility that over two adjacent beats the heart may have decelerated greatly with a consequent slight lengthening of the systolic interval of the second beat. Thus there may arise the situation that the diastolic time of the first beat is shorter than the systolic time of the second beat. From a comparison based solely on the time differences which result from such an occurrence, it would appear that such an arrangement of cluster starting times suggests a first, second, first heart sound pattern (I-II-I) whereas it actually is a second, first, second pattern (II-I-II).

ka_5 successively examines contiguous groups of three valid principal-heart-sound-cluster starting-times using the above criteria with the added proviso that the first time difference must exceed the second time difference by greater than 20 ms (10 samples) for the intervals to be identified as diastolic/systolic. This is to allow for 'jitter' on the starting times of the clusters caused by noise in both the original PhCG and errors introduced by the processing of the PhCG.

If ka_5 completes its task without finding any solution islands, it assumes that there is too much uncertainty in the PhCG and reports this conclusion, which immediately terminates any further exploration of the corresponding data block.

All the preceding KAs operate under the guidance of meta-knowledge area 7 (meta_ka_7) in the control sequence represented in figure 6.5.

6.2.3 Expansion From Solution Islands

The highest level of data abstraction in the data base now contains positively identified first and second heart sounds along with groups of three starting times. These latter groups are either *bona fide* I-II-I heart sounds but which cannot be unambiguously substantiated as such, or groups within which there are latent adventitious sounds. Given the identified principal heart sounds of the solution islands, knowledge of the limits of the FHR, and the restrictions on the composition of the foetal heart beat, it is now possible to expand the analysis outwards from the solution islands into the ambiguous regions which lie on either side.

The expansion of analysis is controlled by the strategy meta-KA (figure 6.6) and is conducted in three regions: from the first solution island back to the beginning of the record (meta_ka_2 & 5), from the end of the last solution island to the end of the record (meta_ka_3), and the area which lies between solution islands (meta_ka_4). The explanation of the procedure used to expand the analysis arbitrarily takes, as an example, the region from the beginning of the record to the first encountered solution island which at this stage is still a II-I-II type. There is no special significance to commencing the explanation here, as the procedure is basically similar in all regions.

In order to commence the expansion towards the start of the record, the missing I is sought which is the pair of the II of the solution island. The closest this I can be to

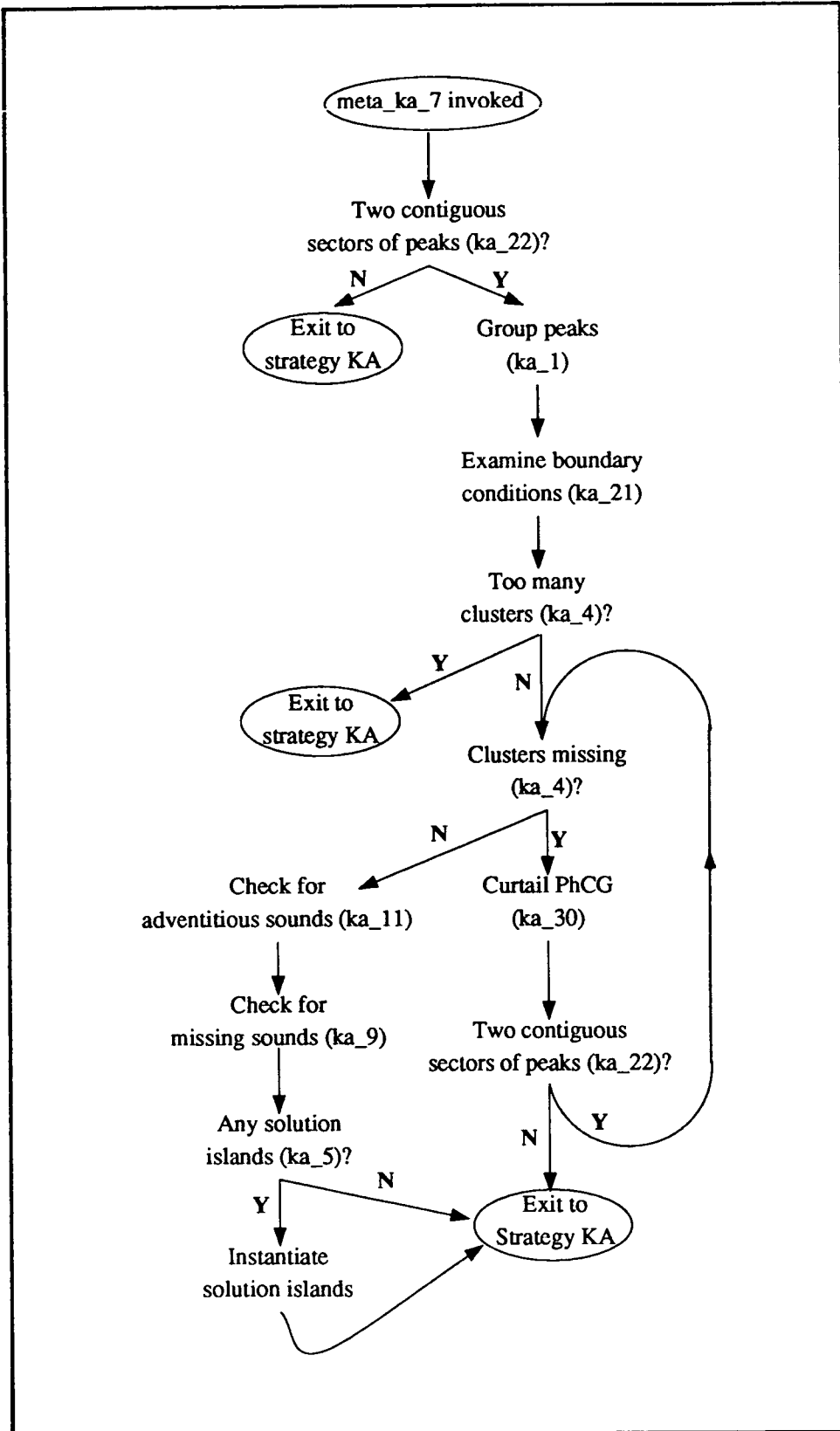


Figure 6.5 Control sequence of meta-knowledge area 7.

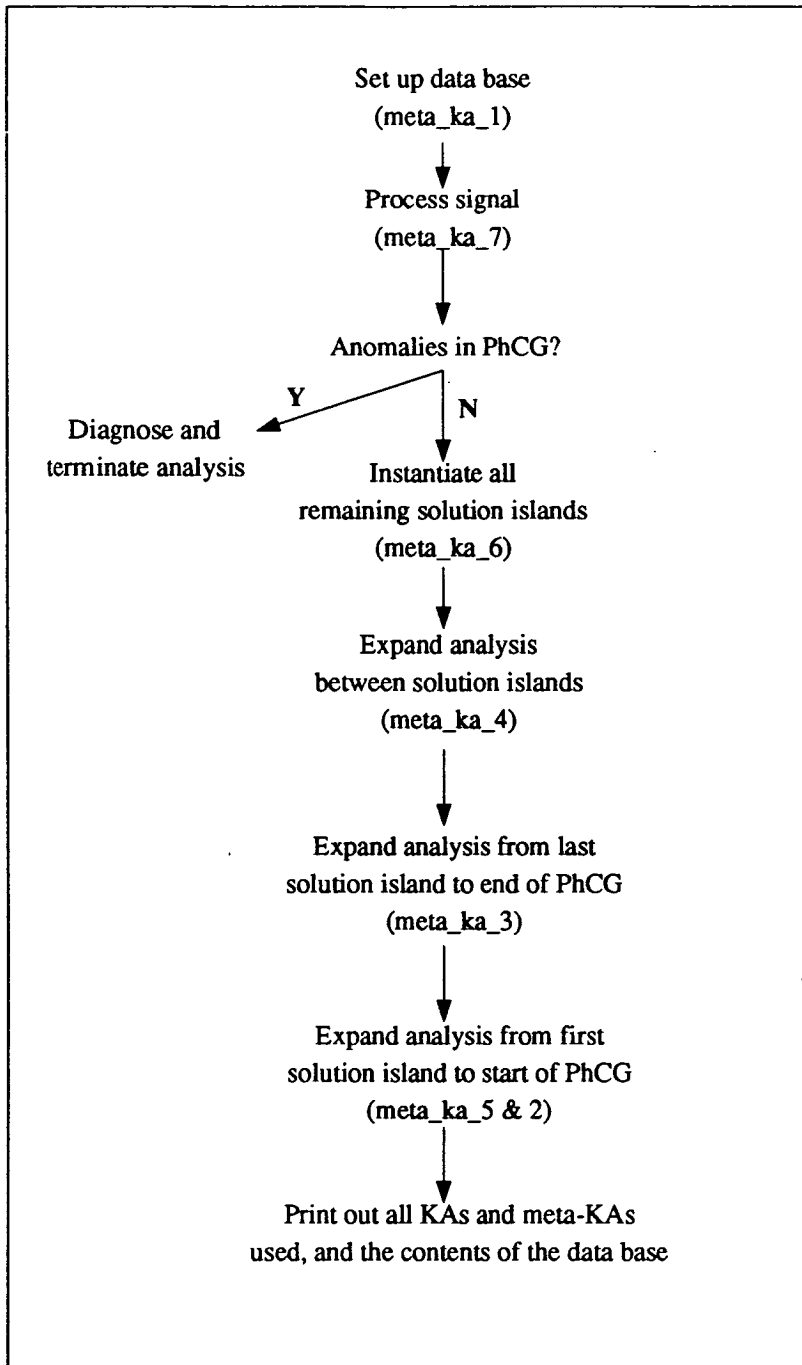


Figure 6.6 Control sequence of the strategy knowledge area.

the corresponding II is when the FHR is at 175 bpm and the systolic/diastolic time ratio is one-to-two. The furthest the I can be from the II is when the FHR is at 117 bpm and the systolic/diastolic time ratio is unity[†]. This information gives a time window within which the I must lie. The lower and upper limits of this window are set at 57 and 128 samples (114-256 ms) earlier than the II, and are established by ka_15. Any event lying between the II and 57 samples earlier must be an adventitious sound and is removed. If only one event is found in the 'allowed' window it is assumed to be the sought-after I. However, if two events are found in the window, this implies that one of these two is an adventitious sound and the other the I[‡]. An additional piece of knowledge which may resolve this case is the requirement that the systolic time must always be less than or equal to the diastolic time (ka_20a). Thus if only one of the events in the 'window' is less than or equal to a time equivalent to the diastolic time away from the II — which is calculated from the separation of the II and I in the solution island — it is accepted as the I and the competing event is deleted. Assuming that it has now been established which of the events is the I and which the adventitious sound, the remaining event in the original group of three starting times must be the II of the previous beat — subject to verification. These new facts are entered into the data base. If, however, the ambiguity remains unresolved by these tests, the expansion is halted at this point and a diagnostic message is supplied to the user.

Assuming that the expansion has been successfully accomplished, there is a possibility that the solution island now begins with a I. An expansion backwards, in this case, follows the same procedure outlined above except that the lower and upper

[†]These systolic/diastolic time ratios make the assumption that at any given FHR the systolic time interval may vary between the theoretical limits of half to one-third the time of the beat. In practice, however, with lower FHRs the systolic time never approaches half the duration of the beat.

[‡]The assumption was made (section 6.2.1) that there is a maximum of one adventitious sound per beat.

limits of the window are widened to 85 and 171 samples (170-342 ms) earlier than the I (ka_16a). These figures are arrived at when it is considered that the closest a II can be to a I is when the FHR is 175 bpm and the ratio of systolic to diastolic time is unity; the furthest a II can be from a I occurs when the FHR is 117 bpm and the systolic time comprises one-third of the beat. Contrary to the expansion backwards from a II, there is no indication now of the actual diastolic time for the beat under investigation. However, it is possible to fix the maximum duration of the systolic time. If two events are competing in the window interval for II status, one must be an adventitious sound. So the third and undisputed member of the group of three starting times must be a I. Knowing then the beginning and end of the beat, the mid-way point between these two Is is the limit of the systolic time. A check is then made to determine if one of the potential IIs falls beyond this limit whereupon, if this is verified, it is rejected and the I and II are established. Otherwise the expansion is forced to terminate.

This procedure for examining groups of three clusters to find the adventitious sound continues until the beginning of the record is reached, or the expansion is terminated by an unresolvable ambiguity. In expanding from the last solution island to the end of the record, the same method is followed except that the expansion is made in advancing rather than 'receding' time. When attempting to analyse the region between solution islands, advantage is taken of the ability to attack ambiguous regions bilaterally rather than unilaterally as previously. It may happen that regions which are unresolvable in one direction, are resolvable when the analysis is tried from the other direction.

Previously, when discussing the expansion from solution islands, it was assumed that apart from the solution islands themselves, the only other events existing were

cluster starting times which contained an adventitious sound as a latent member of a group of three starting times. In fact the expansion proposed above is attempted only after the following expansion has been undertaken (meta_ka_6).

The previous method of expansion ignores the existence of the potential I-II-I solution islands and other single and double groups of starting times which ka_11 has declared not to be adventitious sounds. To take the most straightforward case as an example: suppose a II-I-II solution island has been located, following this in time is a group of three events of which group one unknown member is an adventitious sound, and finally a group of three events which satisfy the conditions to be a I-II-I but which lack of supporting evidence prohibits from being declared as such. The intervening group which contains the adventitious sound, irrespective of the location of this sound, must comprise a I and a II in that order as the solution island ends on a II. This is sufficient evidence to positively identify the following potential I-II-I solution island as such.

This type of expansion which skips over ambiguous regions without resolving them must necessarily only be employed when the condition still holds that ambiguous regions comprise events whose number is a multiple of three. This will be the case immediately after the II-I-II solution islands have been discovered, but not when the ambiguous regions have partially been resolved.

Once both these methods of expansion have concluded, the PhCG record may have been completely analysed. If this is so, a list of all the KAs and meta-KAs which have contributed to the analysis, plus the state of the data base at each significant step in the analysis is presented to the user. An analysis of the next PhCG is then begun.

On the other hand, at this stage there may yet remain certain areas of the PhCG

which have not yielded to a complete analysis. It is to these that attention is now turned. The method used by meta_ka_4 and 5 will be illustrated by means of an example. Suppose two solution islands are separated by an ambiguous region which comprises six events, the solution island which is earlier in time will be assumed to be of the II-I-II type and the later solution island a I-II-I type. In the ambiguous region between these, there must be a I, II, I, II in that order and two adventitious sounds. Both groups of three each contribute a I, II pair. Subject to causal ordering, there are nine possible ways that these events may distribute themselves over the interval. However, certain of these permutations will have been partially resolved by the first expansion method described. These cases are those where the adventitious sound lies immediately adjacent to either or both of the solution islands. To take an instance of the remaining permutations, it is supposed that the unresolved group is literally represented by: a, b, c, d, e, f and that — unknown at this stage — the actual pattern in the overall group is:

(II, I, II), I, X, II, X, I, II, (I, II, I)

where *X* represents an adventitious sound, and the bracketed terms are the solution islands. Suppose that expansion from both lower and upper solution islands has recovered *a* and *f* — the immediately adjacent I and II respectively. Thus the remaining group of four (*b*, *c*, *d*, *e*) must contain a II, and a I in that order and it is presumed that unilateral expansion from both solution islands has been unable to resolve the conflict. Providing that the adventitious sounds are not too close to the actual principal heart sounds, the ambiguity may be resolved using both solution islands working together in a hypothesize-and-test manner. Initially *b* is assumed to be a I; this leaves *d* or *e* as the corresponding II. *b* is tested first in conjunction *d* and then in conjunction with *e*, to ascertain if the window constraints on systolic and diastolic times permit this. Suppose both permutations are proved false. Then *c* and *d*, and, *c* and *e* are separately tested in a like manner. With the non-proximity

proviso holding, c and e should be the only pair to satisfy the requirements, whereupon they are instantiated to a Π and a I respectively. This approach only works when one pair of events alone satisfies the constraints. If more than one permutation is satisfied the test is inconclusive. This region is then abandoned and will not be revisited.

The maximum number of ambiguous events which is exhaustively tested in this way is restricted to nine. All combinations had to be explicitly represented as a control sequence of KA application. Any larger number of combinations would not have justified the increased computation overhead. As this application is 'data-rich' and it is not necessary to determine every beat, larger areas of ambiguity are best avoided, not only from the perspective of computation costs, but also an increased possibility of an incorrect assessment of its constituents.

6.3 PRESENTATION OF RESULTS

Figures 6.7 to 6.19 illustrate results of the processing and analysis achieved by the system described in this chapter. Figure numbers with the suffix 'a' show a PhCG and the result of the processing performed on it. These figures comprise, from top to bottom: a PhCG, the filtered modulus of the cross-correlation of the PhCG with the template drawn from the PhCG as shown, this latter sequence overlaid with vertical lines which identify the position and magnitude of the peaks of the waveform, and finally, markers indicating the starting point — as judged by the processing sub-system — of the significant signal events in the PhCG. A figure number with the suffix 'b' presents the analysis performed by the system on the PhCG of the same figure number but with the suffix 'a'.

The results of the analysis are presented as a list of the applied KAs and any comments associated with the computations. The table of columns of numbers presents the various states of the data base in the progression from processing to final analysis. Each number represents the starting time, in samples, of a significant event. For ease in locating an event in the PhCG, the equivalent spatial starting point of each event, scaled in millimetres, is shown alongside and is measured from the beginning of the PhCG.

The first state of the data base comprises all the peaks found in the filtered modulus of the cross-correlation sequence but is too long to be conveniently displayed in the figure. The second state of the data base lists all the starting times of those events which are judged to be significant in the original PhCG. State 3 results from the processing performed by ka_11 and ka_9. A '?' is a token holding the position of the starting time of the laterally corresponding event in state 2 and represents an event which has not been confirmed to be either a principal heart sound or an adventitious sound. A number in the column represents the starting time of an event which is confidently assumed to be a principal heart sound but its particular identity has not yet been established. State 4 shows the emergence of the solution islands from the events represented as numbers in state 3. At this stage all solution islands are of the II-I-II type. The final state shows the extent of the analysis of the original PhCG. Events still represented as a '?' could not be positively identified as either principal heart sounds or adventitious sounds. Again the distance from the beginning of the PhCG in millimetres is shown alongside for convenience in locating the corresponding event. The list presented at the final stage is often shorter than that of state 2. This arises from the removal of events which were found to be adventitious sounds.

The PhCGs and the corresponding analysis of them, presented in figures 6.7 to

6.19, were selected from a sample set of 124. These thirteen examples were chosen from the span of PhCG types on the basis that they are representative of the processing and analysis capability and limitations of system 1.

Figure 6.7a is an exemplary PhCG in which the principal heart sounds are pronounced, no adventitious sounds are present, and the background noise level is low. A visual analysis of such a PhCG is immediately obvious. The correlation processing has clearly accentuated the valvular sounds. Unusually, the Is are of a similar duration to the IIs which produces single peaks in the filtered-modulus sequence corresponding to the Is and makes the task of segmentation much easier. The clarity of the PhCG has meant that the analysis was almost complete after the identification of the II-I-II solution islands. All such PhCGs present few problems for processing and analysis, and are always completely resolved.

Figure 6.8a is an example of a PhCG which is similar to that in figure 6.7a except that the background noise level has risen considerably and adventitious systolic sounds are becoming apparent in the latter stages of the PhCG. Nevertheless, the cross-correlation has succeeded in suppressing those effects and events which are not of valvular origin. The longer duration and structure of the Is relative to the correlating template has caused the Is to be represented as a split peak in the filtered-modulus sequence. The segmentation KA has taken account of this and groups both peaks into the same clusters. This is demonstrated in the position of the markers in the lowest plot which indicate the start of significant clusters. The system has achieved a complete analysis of this PhCG except for the last group of three clusters. This is explained by the proximity of the group to the end of the record which limits the deployment of contextual information, and the presence of a latent adventitious sound.

Early and late diastolic sounds are clearly in evidence in figure 6.9a which in turn make the task of grouping the significant peaks not as straightforward as in the two previous examples. However, there is very close agreement between system-determined principal heart sound starting times and those observed in the PhCG. The segmentation KA has also removed some obviously adventitious sounds during the processing, *e.g.* the sound adjacent to the I at the end of the PhCG.

The analysis of figure 6.10a is not immediately revealed to a visual inspection as there is a continuum of noise across the systolic and diastolic phases of all beats. In fact only one II-I-II solution island was initially discovered because of the distribution of mid-diastolic adventitious sounds. However, the expansion of the analysis from this solution island succeeded in resolving the whole PhCG.

The PhCG of figure 6.11a presents many difficulties to an analysis as evident in the relatively long reasoning chain involved and the partial analysis achieved. This is again due to the high levels of background noise and the very low amplitude of the IIs in the latter half of the PhCG. The missing II between 1527 and 1724 samples forced the analysis to disregard all subsequent peak clusters. The two unresolved areas (131-216 samples and 792-824 samples) are both due to the close proximity of late systolic sounds to the following IIs. As both sounds occur within the allowed 'window' there is no means by which they may be differentiated.

The PhCG of figure 6.12a initially consists of quite clear principal heart sounds but this rapidly degenerates towards the end of the record. Paradoxically, it is the Is which suffer complete attenuation in this example, while the IIs remain very strong throughout. The system produced an analysis of the first half of the PhCG but abandoned any further expansion because of the severe degradation of the signal and the many competing events.

The particular susceptibility of IIs to sudden and complete attenuation is exemplified in figure 6.13a. This phenomenon is apparently associated with periods of foetal breathing and may be caused by the modulating effect of this on the venous return of blood to the heart [55]. Late diastolic sounds are present in the former half of the PhCG. In this instance, these sounds are attenuated by cross-correlation, but this is not always the case. The coalescence of diastolic sounds and Is is a difficult occurrence to detect and is exacerbated by the smearing effect of cross-correlation which is inclined to run the approximated events together. Here, however, the two events are sufficiently inhomogeneous that the diastolic sound component was detected and dismissed from consideration by the segmentation KA. As would be expected, the analysis of the latter half of the PhCG was abandoned, after the more straightforward former half had been completely resolved.

The situation in figure 6.14a illustrates the case where Is and IIs are present in the PhCG but the high noise level in the former half of the PhCG causes a plethora of 'significant' peaks among which the principal heart sounds are indistinguishable. The analysis stops short of this section.

Figure 6.15 presents a curious example of a case where the system might be expected to attain a full analysis but, in fact, cannot find any solution islands. Although Is and IIs correlate well, the presence of an adventitious sound, whether systolic or diastolic, in every beat causes the limit of analysability of a PhCG by this system to be reached.

An example of a similar type of situation to the aforementioned is provided in figure 6.16. Here, two solution islands are uncovered from the PhCG but as they are of the I-II-I type, they cannot be instantiated until a II-I-II verifies their status. A particular distribution of adventitious sounds prevents any II-I-II solution islands from

being discovered, thus the analysis must be abandoned.

The presence of an adventitious sound very close to the beginning of the first and third principal heart sounds in figure 6.17a has caused the segmentation KA to mistakenly assume that the start of the clusters is more advanced in time. This arose because the adventitious sounds in both cases fall within the limits for acceptability as a constituent of the following cluster. The result is that the group of three clusters is taken to be a II-I-II solution island whereas it is actually a I-II-I type. However, the system has detected the inconsistency with the other valid II-I-II solution islands in the record and immediately abandons the analysis but not before issuing a statement to that effect.

Whenever a PhCG is so noisy that many adventitious sounds are present and principal heart sounds are missing there is a risk of an incorrect analysis — such is the case in figure 6.18. Around the middle of the PhCG, the II has been completely attenuated, but two adventitious sounds are sufficiently close to the region where the II should be, that no suspicion is aroused. The expansion back from the II-I-II solution island proceeds correctly until this anomaly is encountered. As one of the adventitious sounds meets the criteria for consideration as a II, it is accepted. This circumstance, combined with the number of adventitious sounds in the region, unfortunately, ensures that some event will always satisfy the criteria for principal heart sound status. Before the analysis is abandoned, the system has incorrectly instantiated three beats. The situation prevailing in the PhCG which caused this to happen, contradicts an assumption upon which the system 1 depends for its operation, that is, all principal heart sounds in a PhCG are assumed to be present in that record.

The final example, figure 6.19a, illustrates a PhCG which is heavily contaminated by noise. Although the segmentation KA has accurately determined the

starting points of the principal heart sounds, there are too many clusters arising from noise and adventitious sounds for the system to make any inroads into the decomposition of the PhCG; the analysis is then abandoned.

6.4 DISCUSSION OF RESULTS

The results presented in this section are representative of the capability of system 1 to analyse the range of PhCG types. In the relatively noise-free PhCGs, it is always guaranteed that a full and correct analysis is achieved. With increasing noise, and the presence of adventitious sounds, the analysis still performs well, and is able to abandon expansions into regions where there is too much uncertainty, and recognize inconsistencies in regions which are under analysis. Performance degrades only when the assumptions upon which the system is founded are violated.

Overall, the system illustrates the power of knowledge-based signal understanding in the analysis of a signal which is corrupted by high levels of noise. Not only is the system capable of analysing anomalous regions of the PhCG, it can also determine when the total PhCG, or parts of it, are too noise-contaminated to warrant investigation. Furthermore, this system has the ability to recognize inconsistencies in its own analysis which arise mainly when the PhCG is too degraded, and to justify the abandonment of an analysis.

The limitations of system 1 mostly occur in the signal processing stages. Firstly, there is the matter of the filtering threshold used to obtain the PhCG. The PhCG produced by a 40 Hz high-pass filter has subsequently been found to severely attenuate the principal heart sounds. In consideration of the already poor signal-to-noise ratio of the PhCG, this further degradation in signal strength greatly hinders the complete analysis of a PhCG. However, this shortcoming may be easily overcome once the

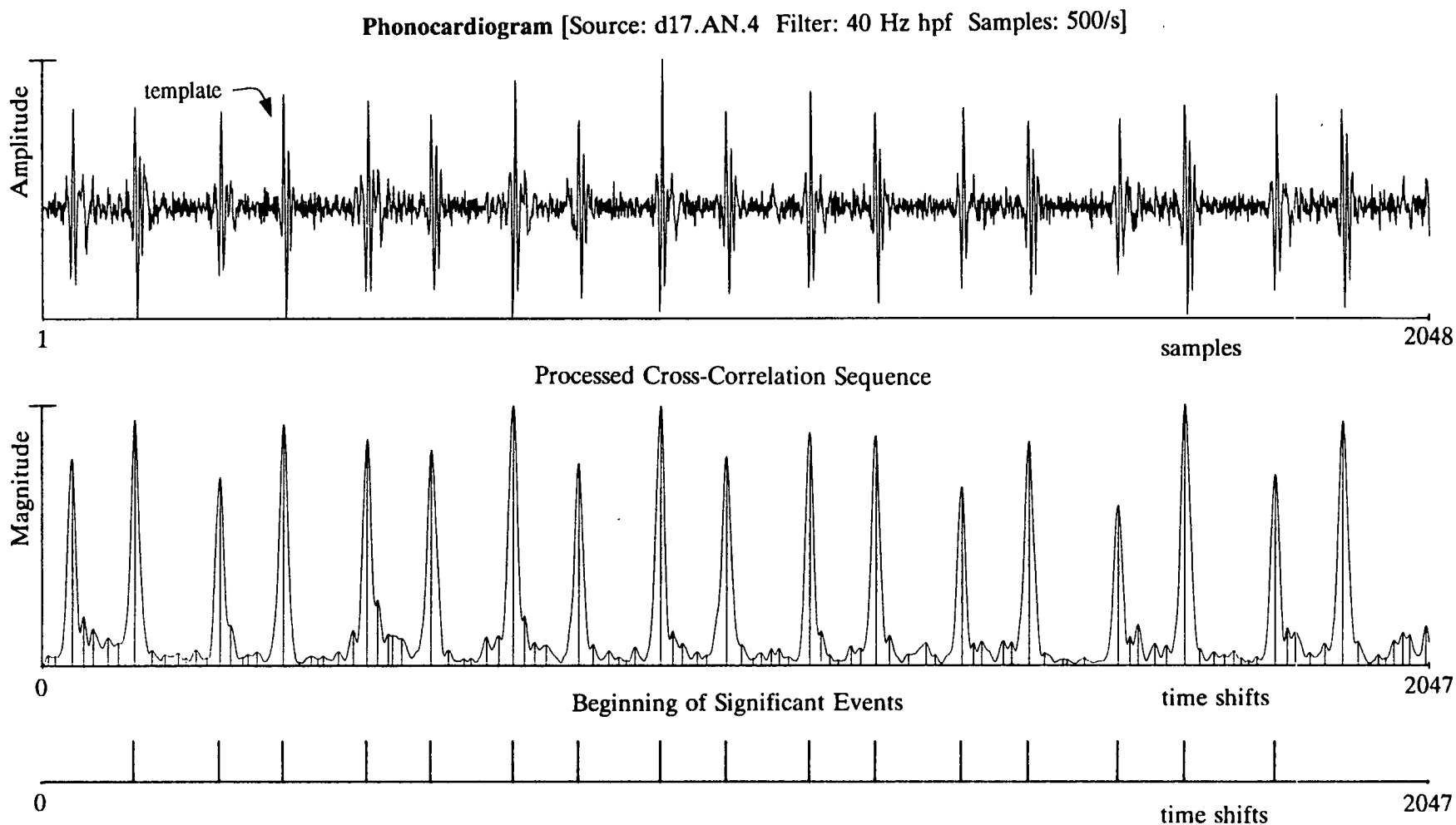


Figure 6.7a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d17.AN.4

meta_ka_1
 meta_ka_7
 ka_22
 Two contiguous sectors of peaks detected.
 ka_1
 No. of samples: 2048
 ka_21
 ka_4
 ka_11
 ka_9
 ka_5

 meta_ka_6
 interpolate

 meta_ka_4
 ka_10
 (meta_ka_4)
 Solution island search completed.

 meta_ka_3
 ka_2
 ka_10

 meta_ka_5

S 2	mm	S 3	S 4	Final	mm
137	14.6	137	II	II	137 14.6
263	28.1	263	I	I	263 28.1
357	38.1	357	II	II	357 38.1
480	51.2	480	480	I	480 51.2
575	61.3	575	II	II	575 61.3
696	74.3	696	I	I	696 74.3
793	84.6	793	II	II	793 84.6
914	97.5	914	914	I	914 97.5
1011	107.9	1011	II	II	1011 107.9
1134	121.0	1134	I	I	1134 121.0
1232	131.4	1232	II	II	1232 131.4
1359	145.0	1359	1359	I	1359 145.0
1458	155.6	1458	1458	II	1458 155.6
1590	169.6	1590	1590	I	1590 169.6
1688	180.1	?	?	II	1688 180.1
1821	194.3	?	?	I	1821 194.3

Figure 6.7b Analysis of phonocardiogram presented in figure 6.7a.

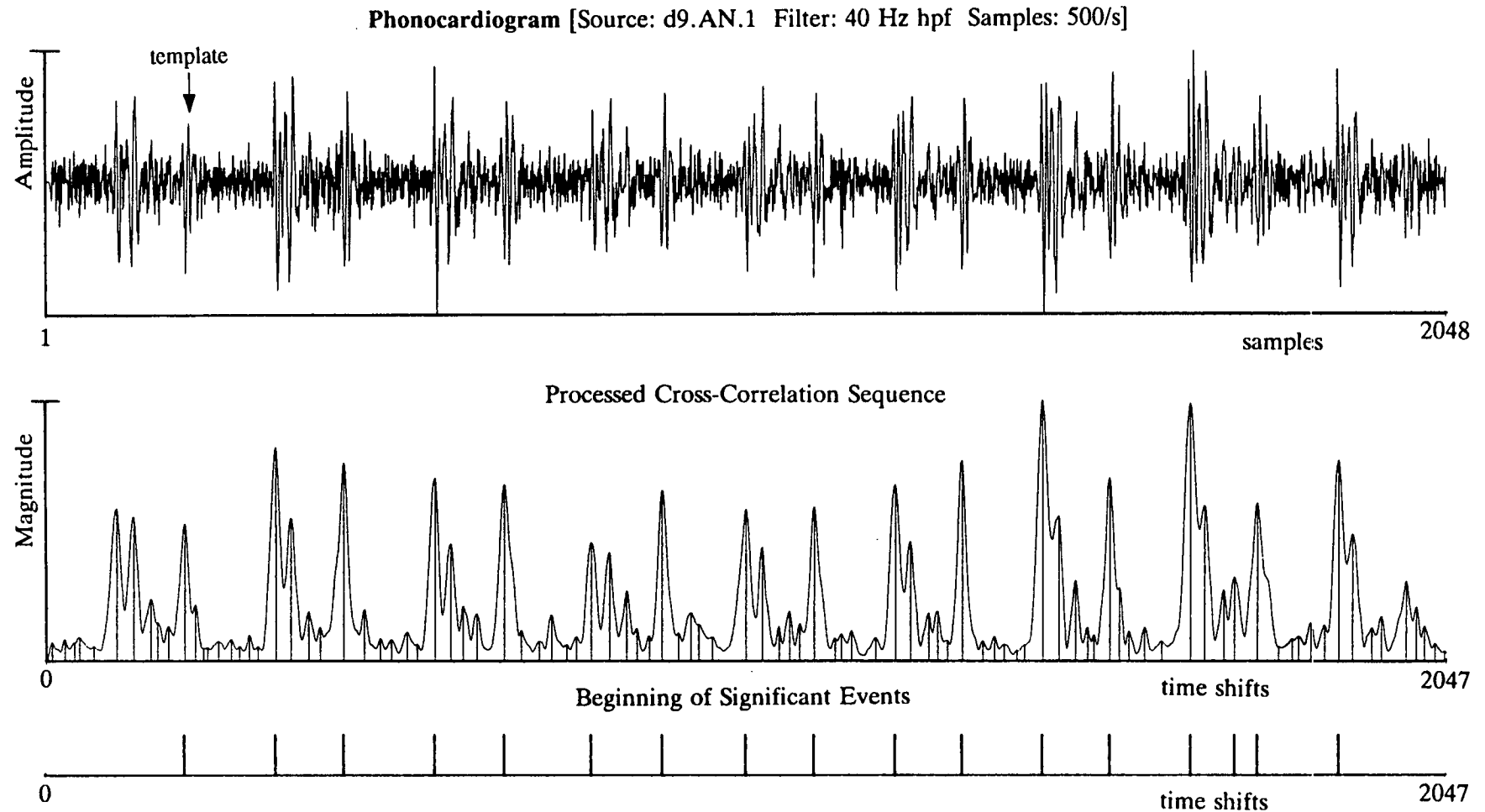


Figure 6.8a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d9.AN.1

meta_ka_1

meta_ka_7

ka_22

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

ka_11

ka_9

ka_5

meta_ka_6

interpolate

meta_ka_4

ka_10

ka_10

ka_2

(meta_ka_4)

Solution island search completed.

meta_ka_3

ka_10

ka_2

ka_20b

(meta_ka_3)

Peaks 1739 and 1772 lie from min. limit to mid-way for a II.

Conflict cannot be resolved.

meta_ka_5

S 2	mm	S 3	S 4	Final	mm
204	21.8	204	II	II	204 21.8
337	36.0	337	I	I	337 36.0
436	46.5	436	II	II	436 46.5
570	60.8	570	570	I	570 60.8
671	71.6	671	II	II	671 71.6
798	85.1	798	I	I	798 85.1
902	96.2	902	II	II	902 96.2
1024	109.2	1024	1024	I	1024 109.2
1124	119.9	1124	II	II	1124 119.9
1243	132.6	1243	I	I	1243 132.6
1341	143.1	1341	II	II	1341 143.1
1459	155.7	1459	1459	I	1459 155.7
1557	166.1	1557	1557	II	1557 166.1
1675	178.7	?	?	I	1675 178.7
1739	185.5	?	?	?	1739 185.5
1772	189.1	?	?	?	1772 189.1
1891	201.7	?	?	?	1891 201.7

Figure 6.8b Analysis of phonocardiogram presented in figure 6.8a.

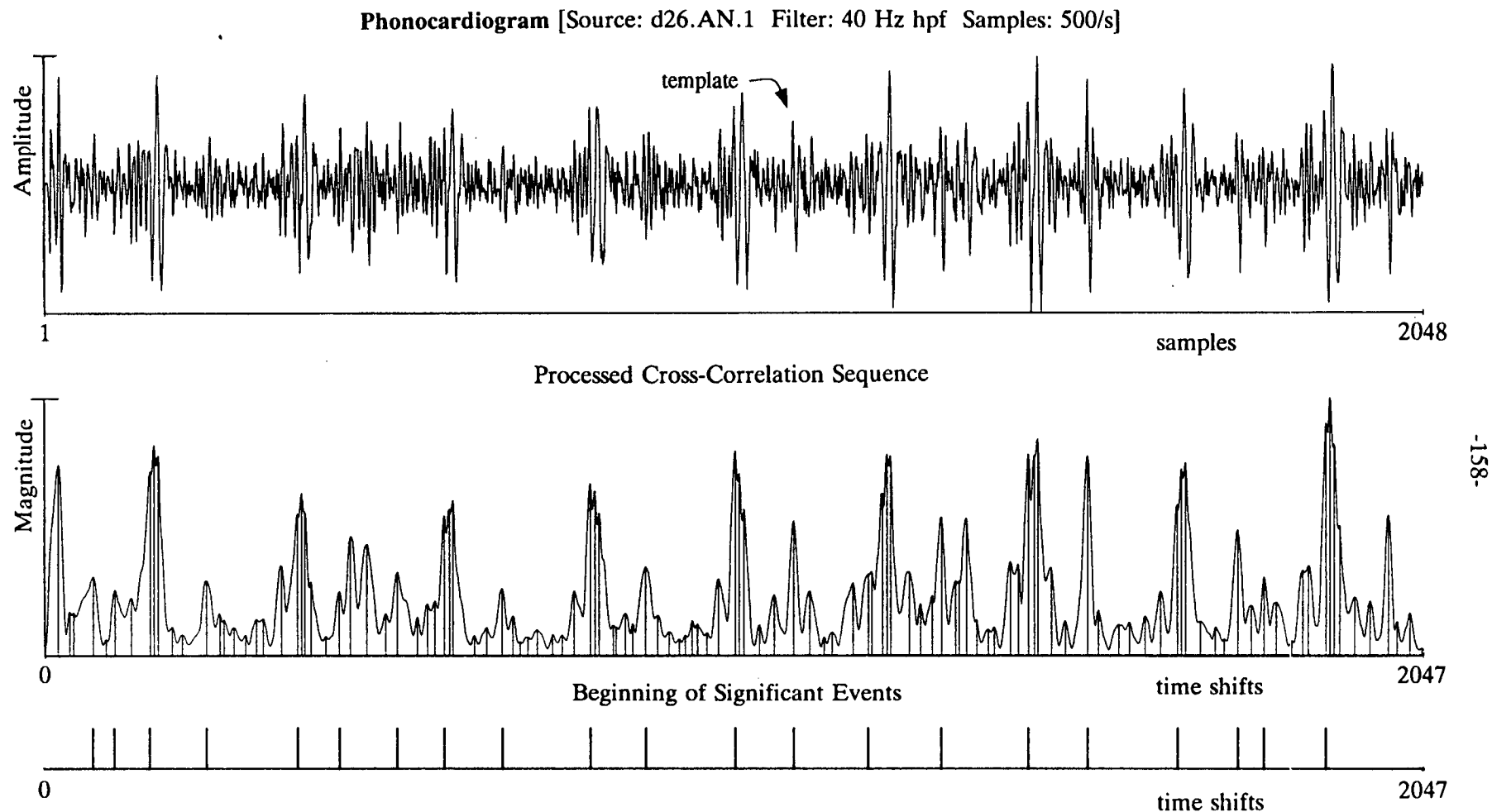


Figure 6.9a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d26.AN.1

meta_ka_1

meta_ka_7

ka_22

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

ka_11

ka_9

ka_5

meta_ka_6

interpolate

meta_ka_4

(meta_ka_4)

Solution island search completed.

meta_ka_3

ka_10

ka_2

ka_20b

ka_10

meta_ka_5

meta_ka_2

ka_15

ka_16

ka_5a

ka_16

ka_5a

ka_16

ka_16a

S 2	mm	S 3	S 4	Final	mm
73	7.8	?	?	I	157
105	11.2	?	?	II	241
157	16.8	?	?	I	376
241	25.7	?	?	II	438
376	40.1	?	?	I	594
438	46.7	?	?	II	680
524	55.9	?	?	I	811
594	63.4	594	594	II	893
680	72.5	680	II	I	1026
811	86.5	811	I	II	1113
893	95.3	893	II	I	1224
1026	109.5	1026	1026	II	1333
1113	118.7	1113	1113	I	1462
1224	130.6	1224	1224	II	1550
1333	142.2	1333	II	I	1683
1462	156.0	1462	I	II	1772
1550	165.4	1550	II	I	1903
1683	179.6	?	?	0	0
1772	189.1	?	?	0	0
1811	193.2	?	?	0	0
1903	203.0	?	?	0	0

Figure 6.9b Analysis of phonocardiogram presented in figure 6.9a.

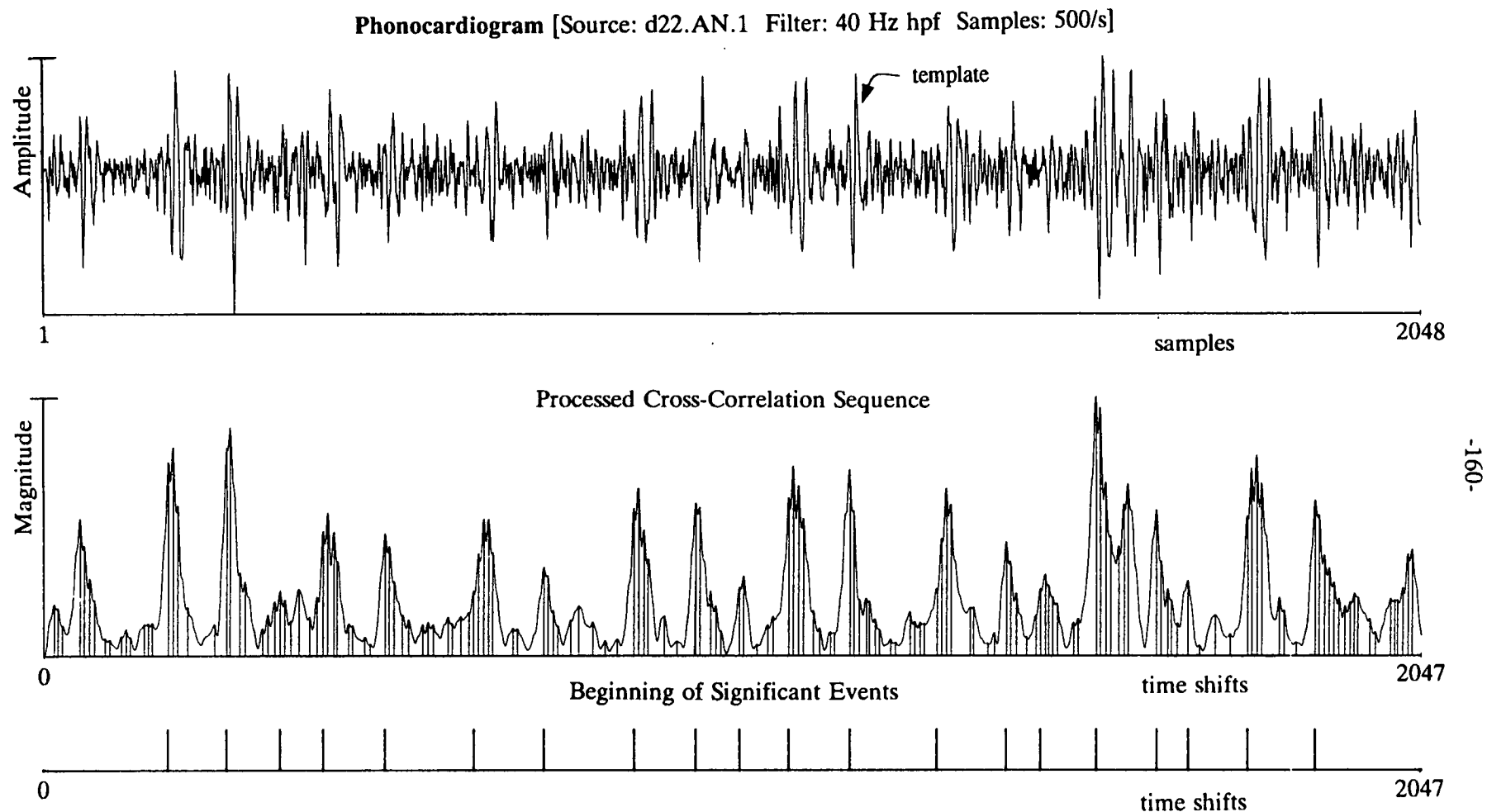


Figure 6.10a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d22.AN.1

meta_ka_1

meta_ka_7

ka_22

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

ka_11

ka_9

ka_5

meta_ka_6

meta_ka_4

ka_33

ka_13

ka_13a

ka_19

ka_27

ka_19v

ka_27v

(meta_ka_4)

Solution island search completed.

meta_ka_3

ka_2

ka_10

ka_2

ka_10

ka_2

ka_10

ka_2

meta_ka_5

meta_ka_2

ka_15

ka_16

ka_16a

S 2	mm	S 3	S 4	Final	mm
186	19.8	?	?	I	186 19.8
273	29.1	?	?	II	273 29.1
352	37.6	?	?	I	416 44.4
416	44.4	416	416	II	508 54.2
508	54.2	508	II	I	640 68.3
640	68.3	640	I	II	744 79.4
744	79.4	744	II	I	878 93.7
878	93.7	?	?	II	970 103.5
970	103.5	?	?	I	1108 118.2
1035	110.4	?	?	II	1199 127.9
1108	118.2	1108	1108	I	1329 141.8
1199	127.9	1199	1199	II	1432 152.8
1329	141.8	?	?	I	1566 167.1
1432	152.8	?	?	II	1655 176.6
1483	158.2	?	?	I	1790 191.0
1566	167.1	?	?	II	1890 201.6
1655	176.6	?	?	0	0 0.0
1702	181.6	?	?	0	0 0.0
1790	191.0	?	?	0	0 0.0
1890	201.6	?	?	0	0 0.0

Figure 6.10b Analysis of phonocardiogram presented in figure 6.10a.

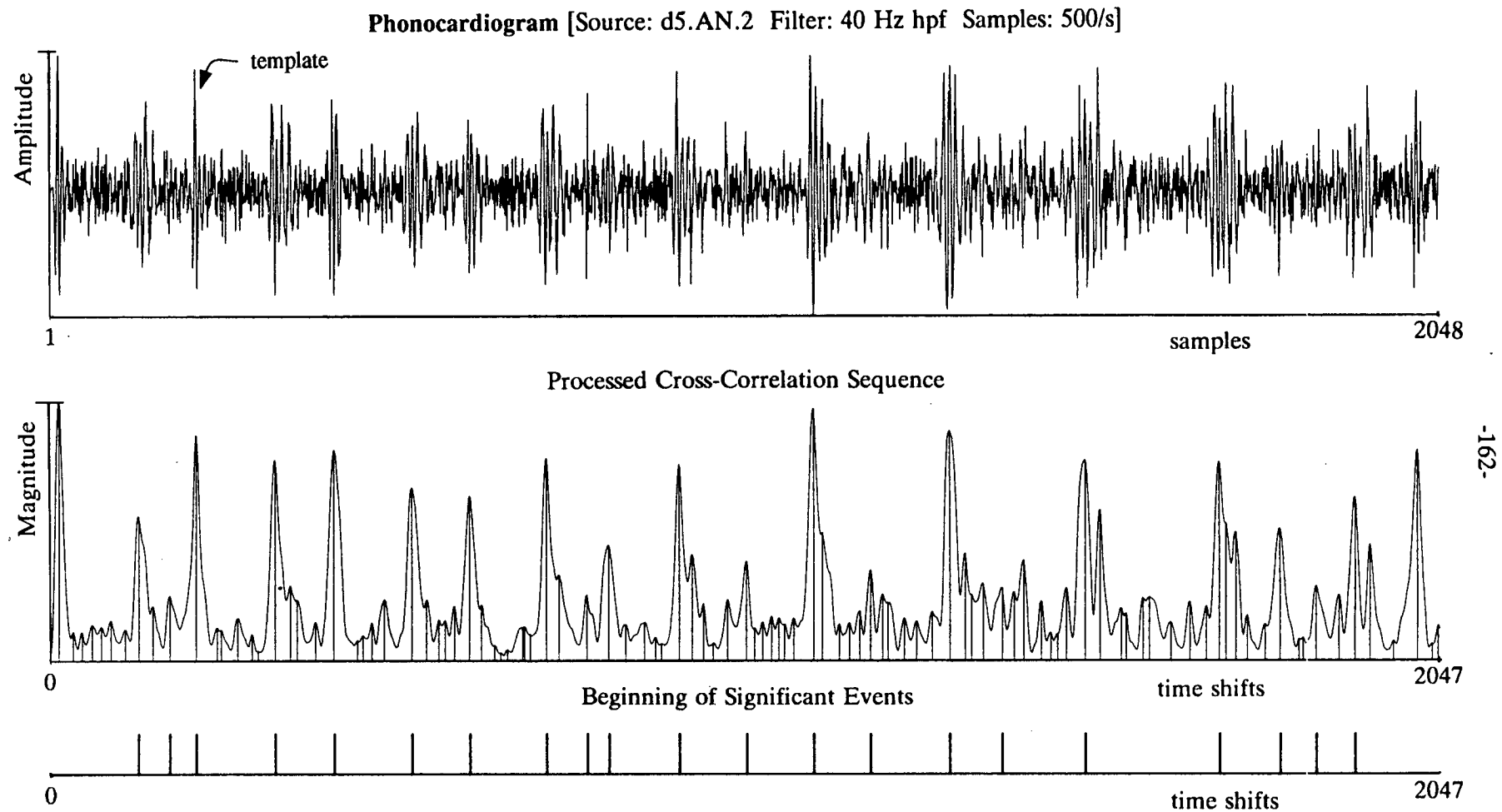


Figure 6.11a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d5.AN.2

meta_ka_1

meta_ka_7

ka_22

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

Missing peak(s) between 1527 and 1724

ka_30

Revised sample count: 1607

No. of grouped sig. peaks reduced by 4

Removed grouped peaks are:

start	mm
1724	183.9
1813	193.4
1866	199.1
1923	205.2

ka_22

Two contiguous sectors of peaks detected.

ka_11

ka_9

ka_5

meta_ka_6

interpolate

meta_ka_4

ka_33

ka_13

ka_13a

ka_19

ka_27

ka_19v

ka_27v

ka_19

(meta_ka_4)

Hypothesize and test failed to resolve conflict.

Both returned TRUE.

ka_10

(meta_ka_4)

Solution island search completed.

meta_ka_3

ka_2

ka_10

meta_ka_5

meta_ka_2

ka_15

ka_16

ka_20

(meta_ka_2)

Peaks 177 and 216 satisfy conditions for II.

(ka_20)

S 2	mm	S 3	S 4	Final	mm
131	14.0	?	?	?	131
177	18.9	?	?	?	177
216	23.0	?	?	?	216
332	35.4	332	332	I	332
419	44.7	419	II	II	419
534	57.0	534	I	I	534
619	66.0	619	II	II	619
732	78.1	?	?	I	732
792	84.5	?	?	?	792
824	87.9	?	?	?	824
928	99.0	928	928	I	928
1027	109.6	1027	II	II	1027
1126	120.1	1126	I	I	1126
1210	129.1	1210	II	II	1210
1327	141.6	1327	1327	I	1327
1404	149.8	?	?	II	1404
1527	162.9	?	?	I	1527

Figure 6.11b Analysis of phonocardiogram presented in figure 6.11a.

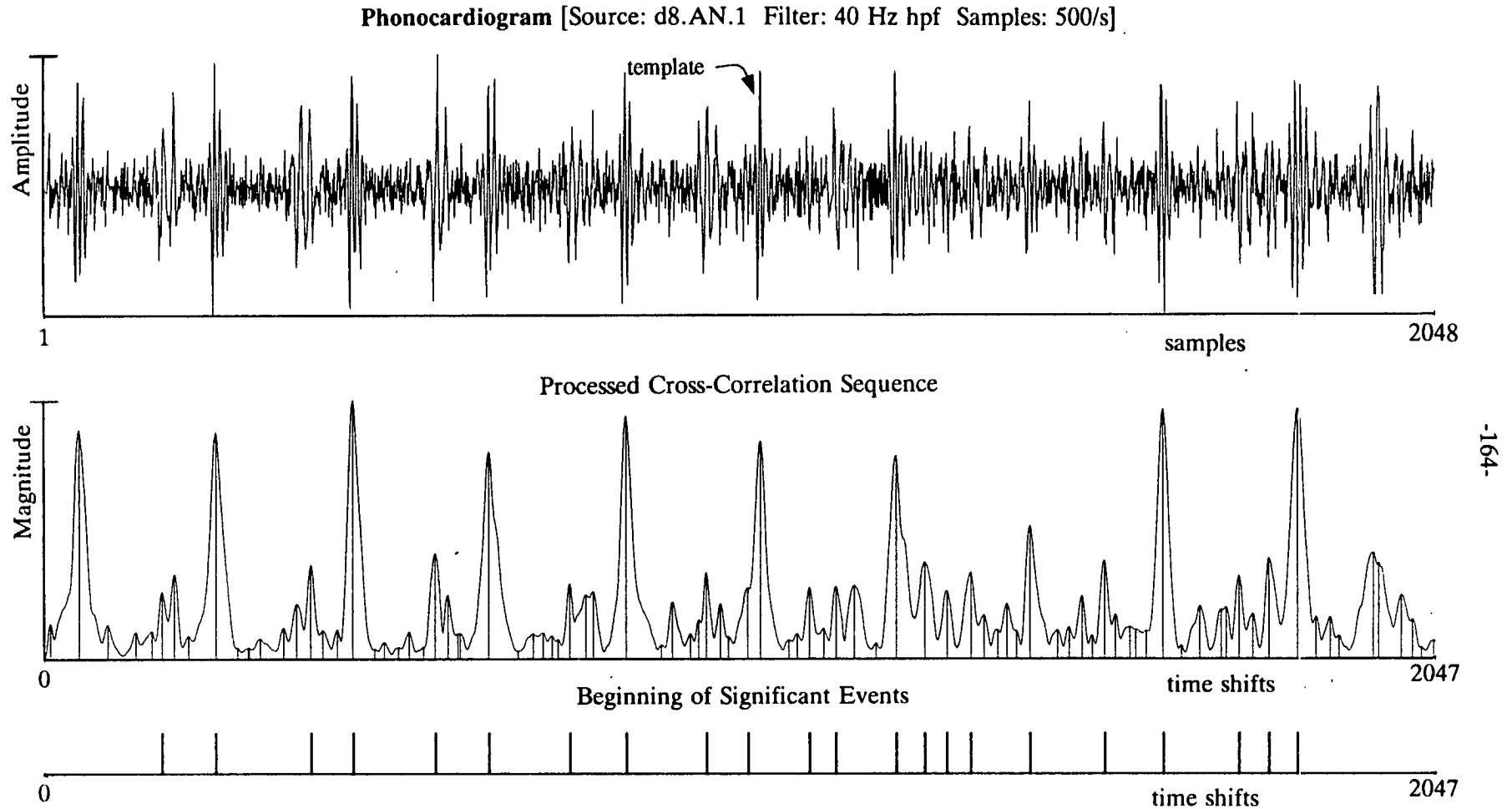


Figure 6.12a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d8.AN.1

meta_ka_1

meta_ka_7

ka_22

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

ka_11

ka_9

ka_5

meta_ka_6

interpolate

meta_ka_4

(meta_ka_4)

Solution island search completed.

meta_ka_3

ka_10

ka_2

ka_10

ka_20c

(meta_ka_3)

Peaks 1127 and 1166 lie mid-way to max. limit for I.

Conflict cannot be resolved.

meta_ka_5

S 2	mm	S 3	S 4	Final	mm
174	18.6	174	174	174	18.6
253	27.0	253	II	II	253
393	41.9	393	I	I	393
455	48.5	455	II	II	455
576	61.5	576	576	I	576
655	69.9	655	II	II	655
774	82.6	774	I	I	774
857	91.4	857	II	II	857
975	104.0	?	?	I	975
1036	110.5	?	?	II	1036
1127	120.2	?	?	?	1127
1166	124.4	?	?	?	1166
1255	133.9	?	?	?	1255
1297	138.4	?	?	?	1297
1330	141.9	?	?	?	1330
1365	145.6	?	?	?	1365
1452	154.9	?	?	?	1452
1562	166.6	?	?	?	1562
1648	175.8	?	?	?	1648
1759	187.7	?	?	?	1759
1803	192.4	?	?	?	1803
1845	196.8	?	?	?	1845

Figure 6.12b Analysis of Phonocardiogram presented in 6.12a.

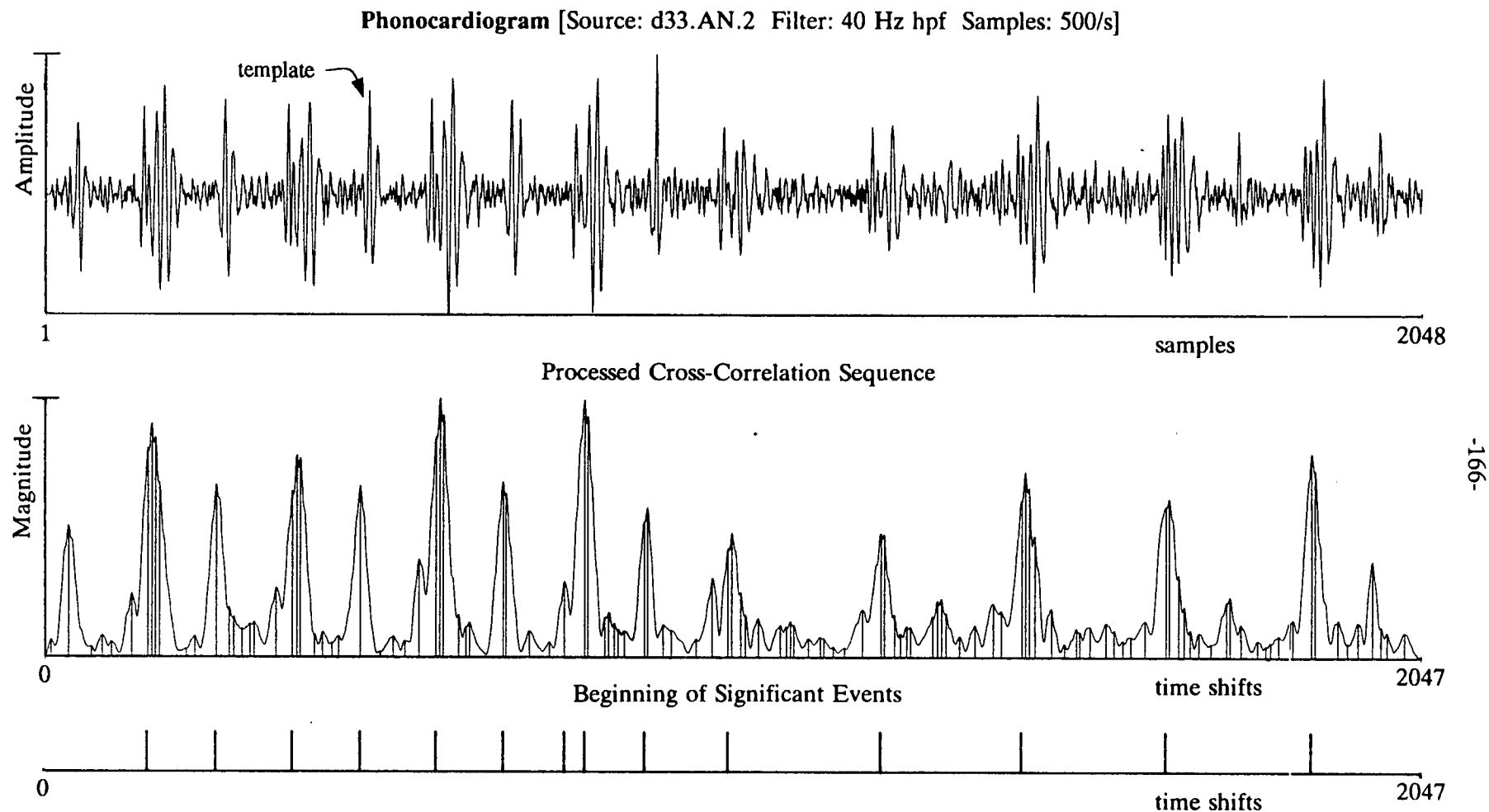


Figure 6.13a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d33.AN.2

meta_ka_1

meta_ka_7

ka_22

Only one peak (at 1455) detected in interval 1322 to 1614.

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

Missing peak(s) between 1017 and 1245

ka_30

Revised sample count: 1097

No. of grouped peaks reduced by 4

Removed grouped peaks are:

start	mm
1245	132.8
1455	155.2
1669	178.1
1885	201.1

ka_22

Two contiguous sectors of peaks detected.

ka_11

ka_9

ka_5

meta_ka_6

meta_ka_4

ka_10

(meta_ka_4)

Solution island search completed.

meta_ka_3

ka_2

ka_10

ka_20c

ka_2

ka_10

meta_ka_5

S 2	mm	S 3	S 4	Final	mm
154	16.4	154	154	154	16.4
256	27.3	256	II	II	256
369	39.4	369	I	I	369
470	50.1	470	II	II	470
583	62.2	583	583	I	583
683	72.9	?	?	II	683
774	82.6	?	?	I	804
804	85.8	?	?	II	893
893	95.3	?	?	I	1017
1017	108.5	?	?	0	0

Figure 6.13b Analysis of phonocardiogram presented in figure 6.13a.

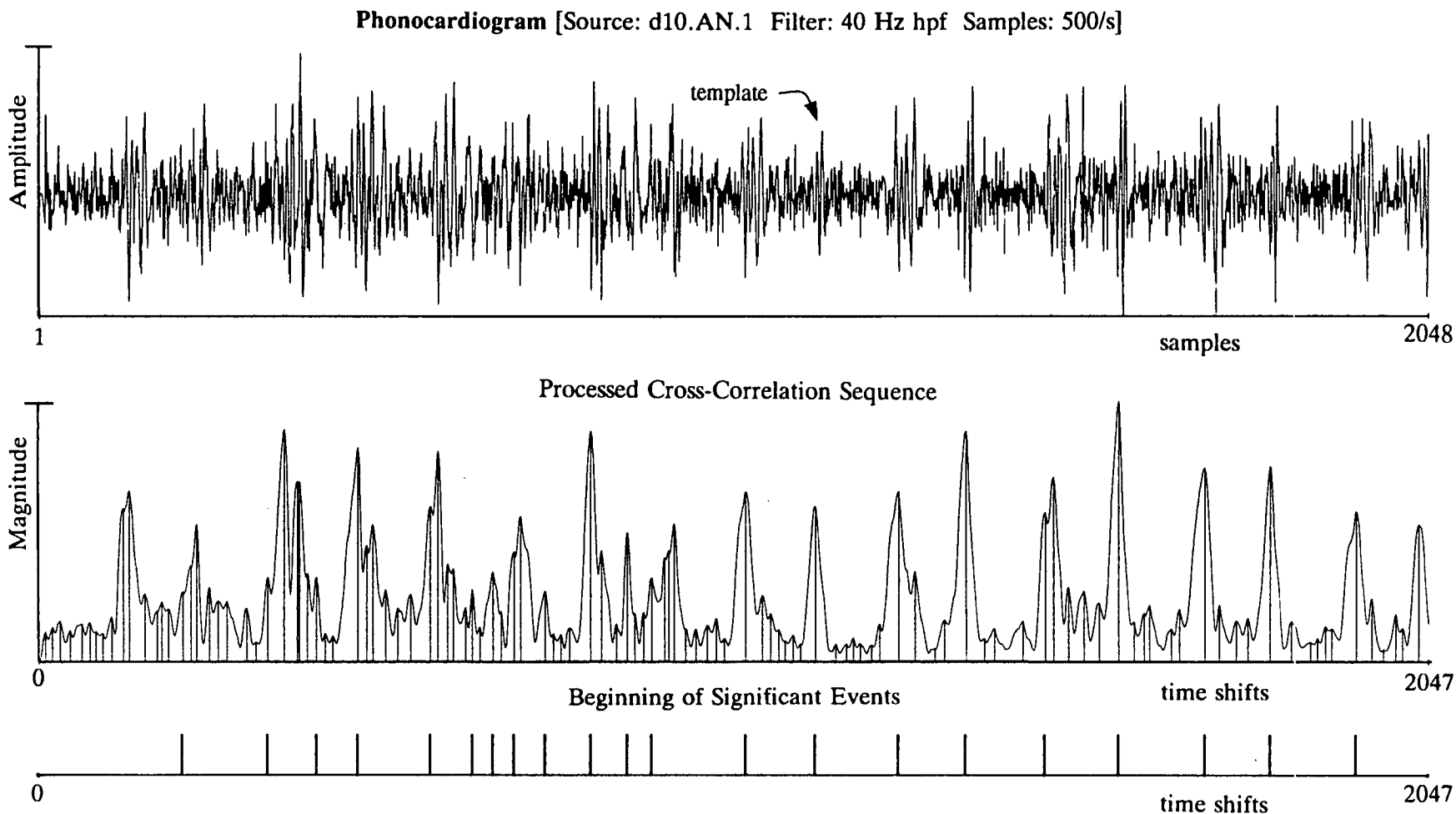


Figure 6.14a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d10.AN.1

meta_ka_1

meta_ka_7

ka_22

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

ka_11

ka_9

ka_5

meta_ka_6

meta_ka_4

ka_2

ka_10

(meta_ka_4)

Solution island search completed.

meta_ka_3

ka_2

ka_10

meta_ka_5

S 2	mm	S 3	S 4	Final	mm
213	22.7	213	213	213	22.7
338	36.1	?	?	?	338
410	43.7	?	?	?	410
471	50.3	?	?	?	471
578	61.7	?	?	?	578
640	68.3	?	?	?	640
670	71.5	?	?	?	670
701	74.8	?	?	?	701
747	79.7	?	?	?	747
814	86.8	?	?	?	814
868	92.6	868	868	868	92.6
904	96.4	904	II	II	904
1042	111.2	1042	I	I	1042
1145	122.2	1145	II	II	1145
1268	135.3	1268	1268	I	1268
1367	145.8	1367	1367	II	1367
1484	158.3	1484	1484	I	1484
1592	169.8	1592	1592	II	1592
1719	183.4	1719	1719	I	1719
1815	193.6	?	?	II	1815
1941	207.1	?	?	I	1941

Figure 6.14b Analysis of phonocardiogram presented in figure 6.14a.

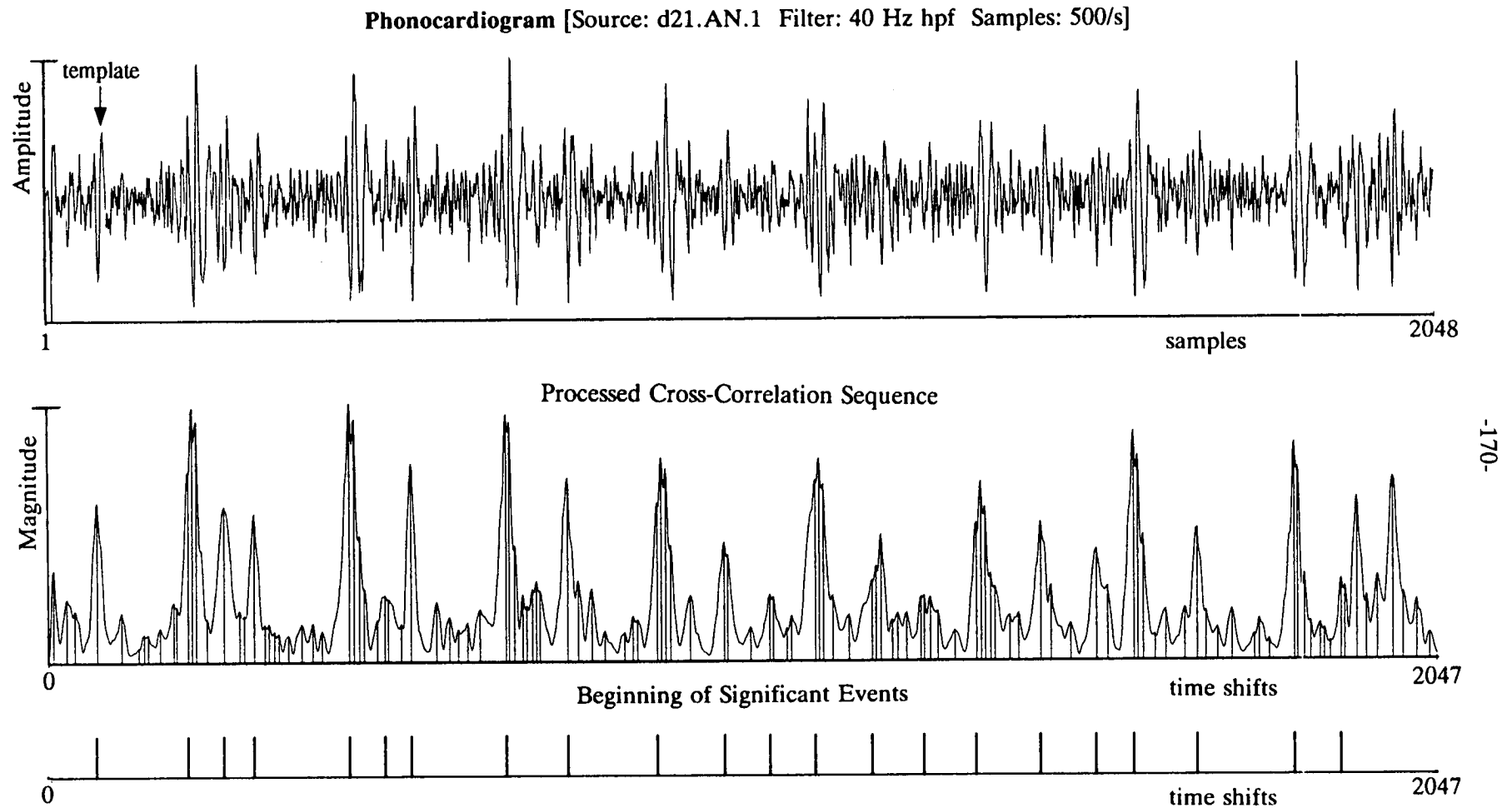


Figure 6.15a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d21.AN.1

meta_ka_1

meta_ka_7

ka_22

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

ka_11

ka_9

ka_5

No solution islands existing (Data noisy).

S 2	mm	S 3	S 4	Final		mm
73	7.8	73	73	0	0	0.0
207	22.1	?	?	0	0	0.0
260	27.7	?	?	0	0	0.0
304	32.4	?	?	0	0	0.0
445	47.5	?	?	0	0	0.0
497	53.0	?	?	0	0	0.0
536	57.2	?	?	0	0	0.0
676	72.1	676	676	0	0	0.0
767	81.8	767	767	0	0	0.0
899	95.9	?	?	0	0	0.0
998	106.5	?	?	0	0	0.0
1065	113.6	?	?	0	0	0.0
1132	120.8	?	?	0	0	0.0
1216	129.7	?	?	0	0	0.0
1293	137.9	?	?	0	0	0.0
1370	146.2	?	?	0	0	0.0
1465	156.3	?	?	0	0	0.0
1547	165.0	?	?	0	0	0.0
1602	170.9	?	?	0	0	0.0
1695	180.8	?	?	0	0	0.0
1838	196.1	?	?	0	0	0.0
1907	203.5	?	?	0	0	0.0

Figure 6.15b Analysis of phonocardiogram presented in figure 6.15a.

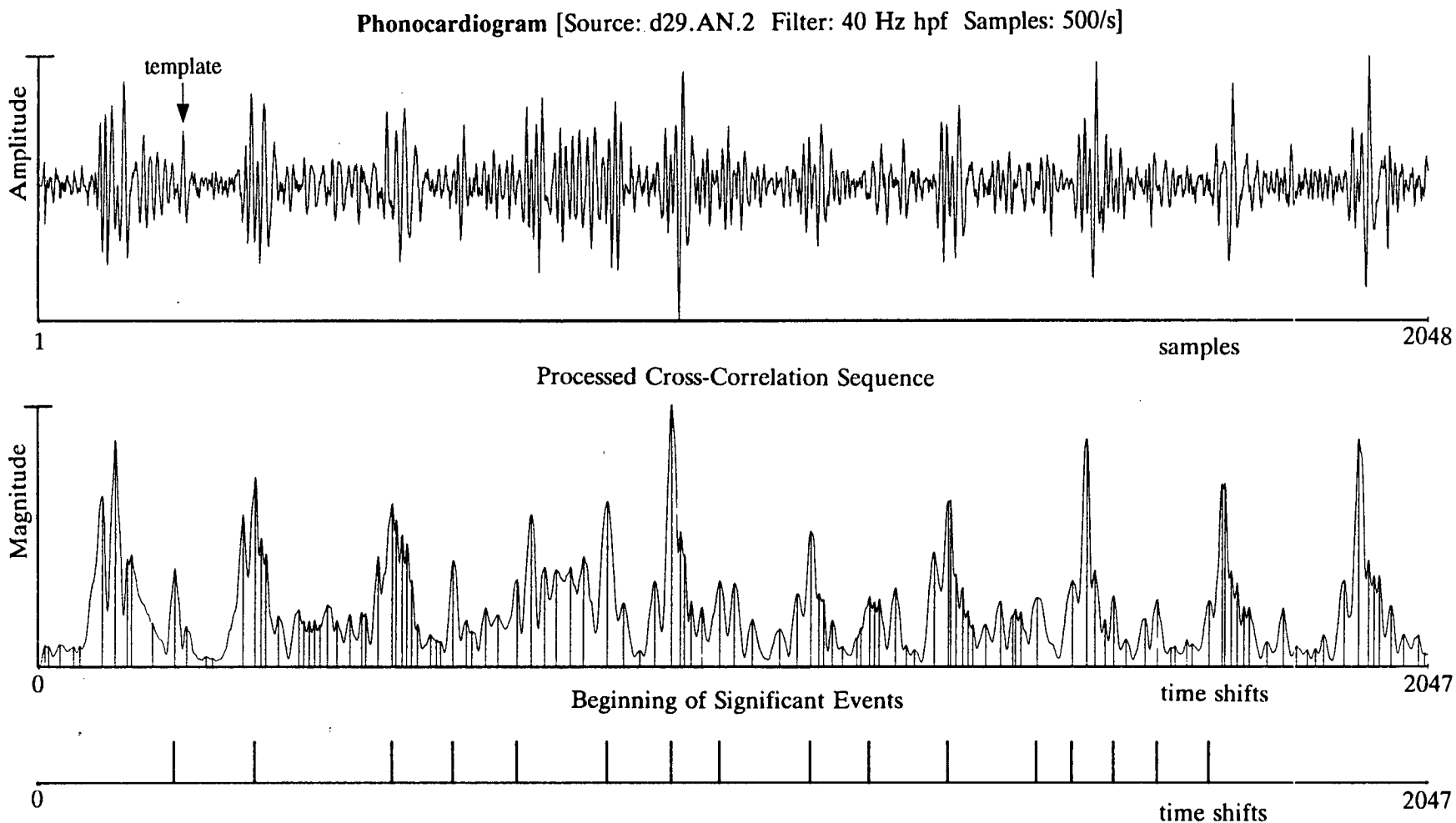


Figure 6.16a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d29.AN.2

meta_ka_1

meta_ka_7

ka_22

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

Missing peak(s) between 320 and 522

ka_30

Revised sample count: 1607

No. of grouped sig. peaks reduced by 2

Removed grouped peaks are:

start	mm
202	21.6
320	34.1

ka_22

Two contiguous sectors of peaks detected.

ka_11

ka_9

ka_5

No solution islands existing (Data noisy).

S 2	mm	S 3	S 4	Final	mm
522	55.7	522	522	0	0 0.0
612	65.3	612	612	0	0 0.0
706	75.3	706	706	0	0 0.0
839	89.5	?	?	0	0 0.0
934	99.6	?	?	0	0 0.0
1005	107.2	?	?	0	0 0.0
1139	121.5	1139	1139	0	0 0.0
1226	130.8	1226	1226	0	0 0.0
1342	143.2	1342	1342	0	0 0.0
1473	157.2	?	?	0	0 0.0
1525	162.7	?	?	0	0 0.0
1586	169.2	?	?	0	0 0.0
1650	176.0	?	?	0	0 0.0
1726	184.1	?	?	0	0 0.0

Figure 6.16b Analysis of phonocardiogram presented in figure 6.16a.

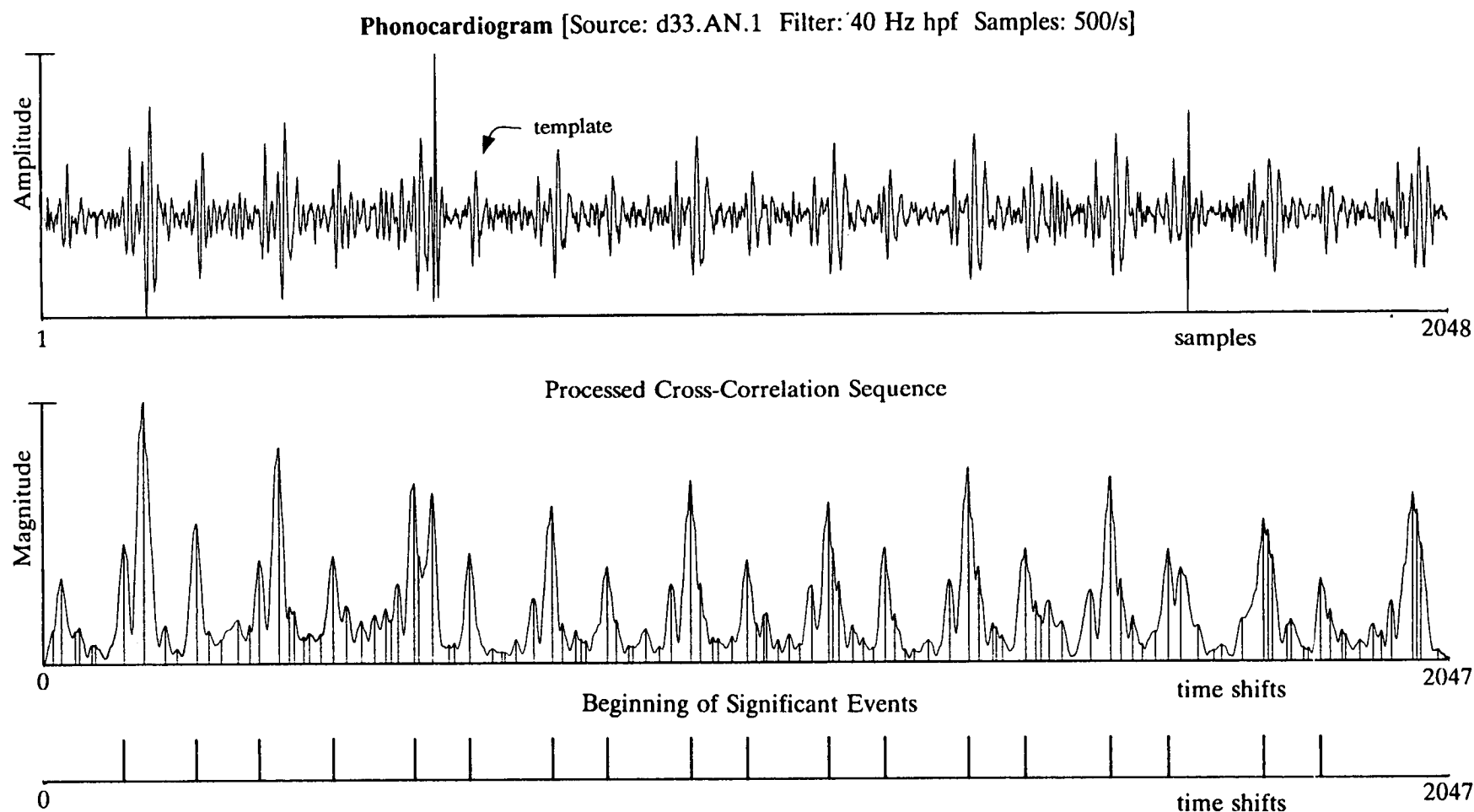


Figure 6.17a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d33.AN.1

meta_ka_1

meta_ka_7

ka_22

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

ka_11

ka_9

ka_5

meta_ka_6

interpolate

(meta_ka_6)

Inconsistency between solution islands.

(interpol)

S 2	mm	S 3	S 4	Final	mm
118	12.6	118	II	II	118 12.6
224	23.9	224	I	I	224 23.9
315	33.6	315	II	II	315 33.6
423	45.1	423	II	II	423 45.1
541	57.7	541	I	I	541 57.7
622	66.4	622	II	II	622 66.4
742	79.2	742	742	742	742 79.2
822	87.7	822	II	II	822 87.7
944	100.7	944	I	I	944 100.7
1026	109.5	1026	II	II	1026 109.5
1145	122.2	1145	1145	1145	1145 122.2
1227	130.9	1227	II	II	1227 130.9
1349	143.9	1349	I	I	1349 143.9
1432	152.8	1432	II	II	1432 152.8
1556	166.0	1556	1556	1556	1556 166.0
1640	175.0	1640	1640	1640	1640 175.0
1778	189.7	?	?	?	1778 189.7
1861	198.5	?	?	?	1861 198.5

Figure 6.17b Analysis of phonocardiogram presented in figure 6.17a.

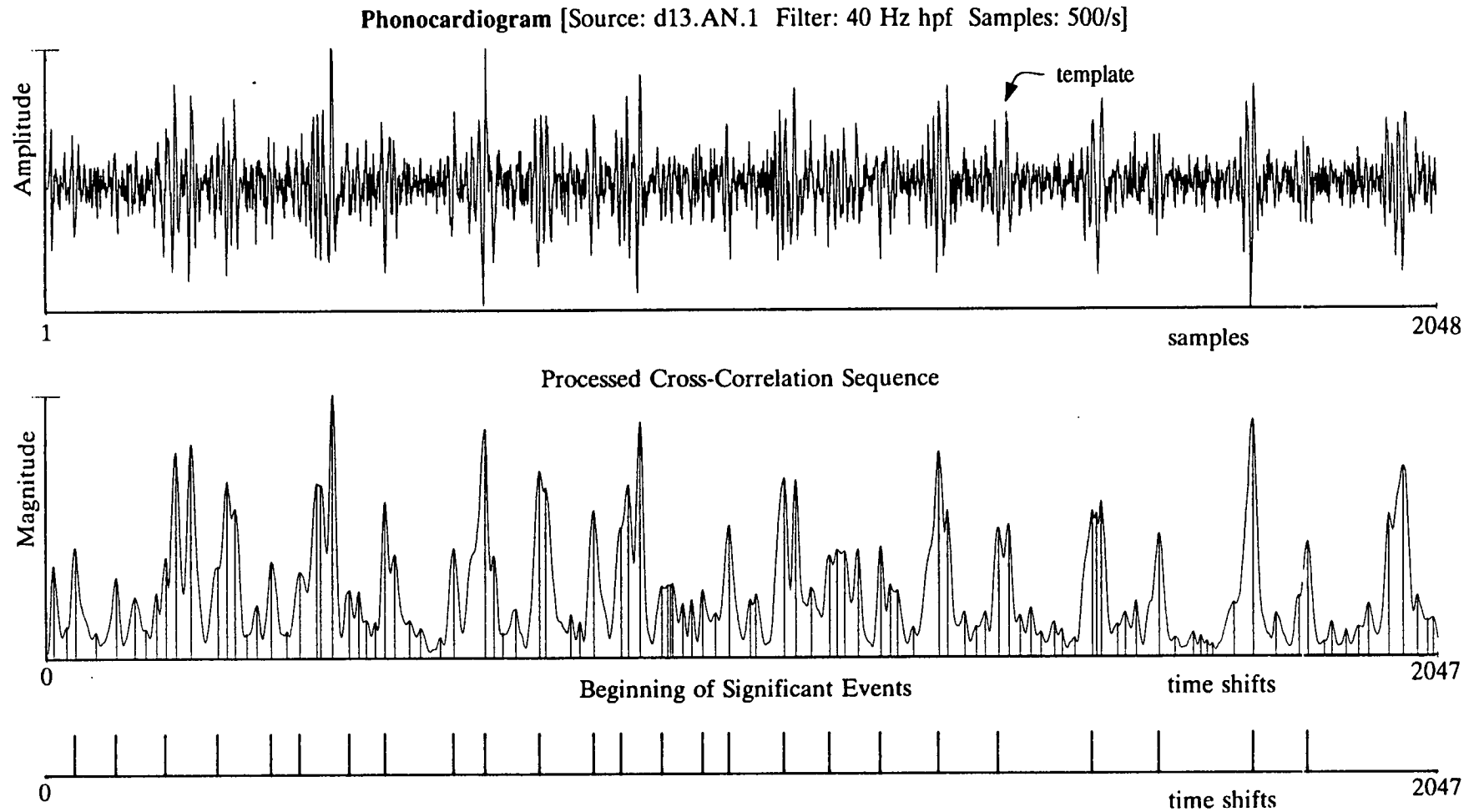


Figure 6.18a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d13.AN.1

meta_ka_1
 meta_ka_7
 ka_22
 Two contiguous sectors of peaks detected.
 ka_1
 No. of samples: 2048
 ka_21
 ka_4
 ka_11
 ka_9
 ka_5
 meta_ka_6
 meta_ka_4
 (meta_ka_4)
 Solution island search completed.
 meta_ka_3
 ka_10
 ka_2
 meta_ka_5
 meta_ka_2
 ka_15
 ka_16
 ka_20
 ka_5a
 ka_16
 ka_5a
 ka_16
 ka_5a
 ka_16
 ka_5a
 ka_16
 ka_20
 (meta_ka_2)
 Peaks 332 and 374 satisfy conditions for II.
 (ka_20)

S 2	mm	S 3	S 4	Final	mm
44	4.7	?	?	?	44 4.7
104	11.1	?	?	?	104 11.1
177	18.9	?	?	?	177 18.9
254	27.1	?	?	?	254 27.1
332	35.4	?	?	?	332 35.4
374	39.9	?	?	?	374 39.9
447	47.7	?	?	I	499 53.2
499	53.2	?	?	II	601 64.1
601	64.1	?	?	I	727 77.6
647	69.0	?	?	II	807 86.1
727	77.6	?	?	I	907 96.8
807	86.1	?	?	II	967 103.2
847	90.4	?	?	I	1087 116.0
907	96.8	?	?	II	1154 123.1
967	103.2	?	?	I	1315 140.3
1006	107.3	?	?	II	1403 149.7
1087	116.0	?	?	I	1541 164.4
1154	123.1	?	?	II	1639 174.9
1229	131.1	1229	1229	I	1777 189.6
1315	140.3	1315	1315	II	1857 198.1
1403	149.7	1403	II	0	0 0.0
1541	164.4	1541	I	0	0 0.0
1639	174.9	1639	II	0	0 0.0
1777	189.6	?	?	0	0 0.0
1857	198.1	?	?	0	0 0.0

Figure 6.18b Analysis of phonocardiogram presented in figure 6.18a.

Phonocardiogram [Source: d16.AN.1 Filter: 40 Hz hpf Samples: 500/s]

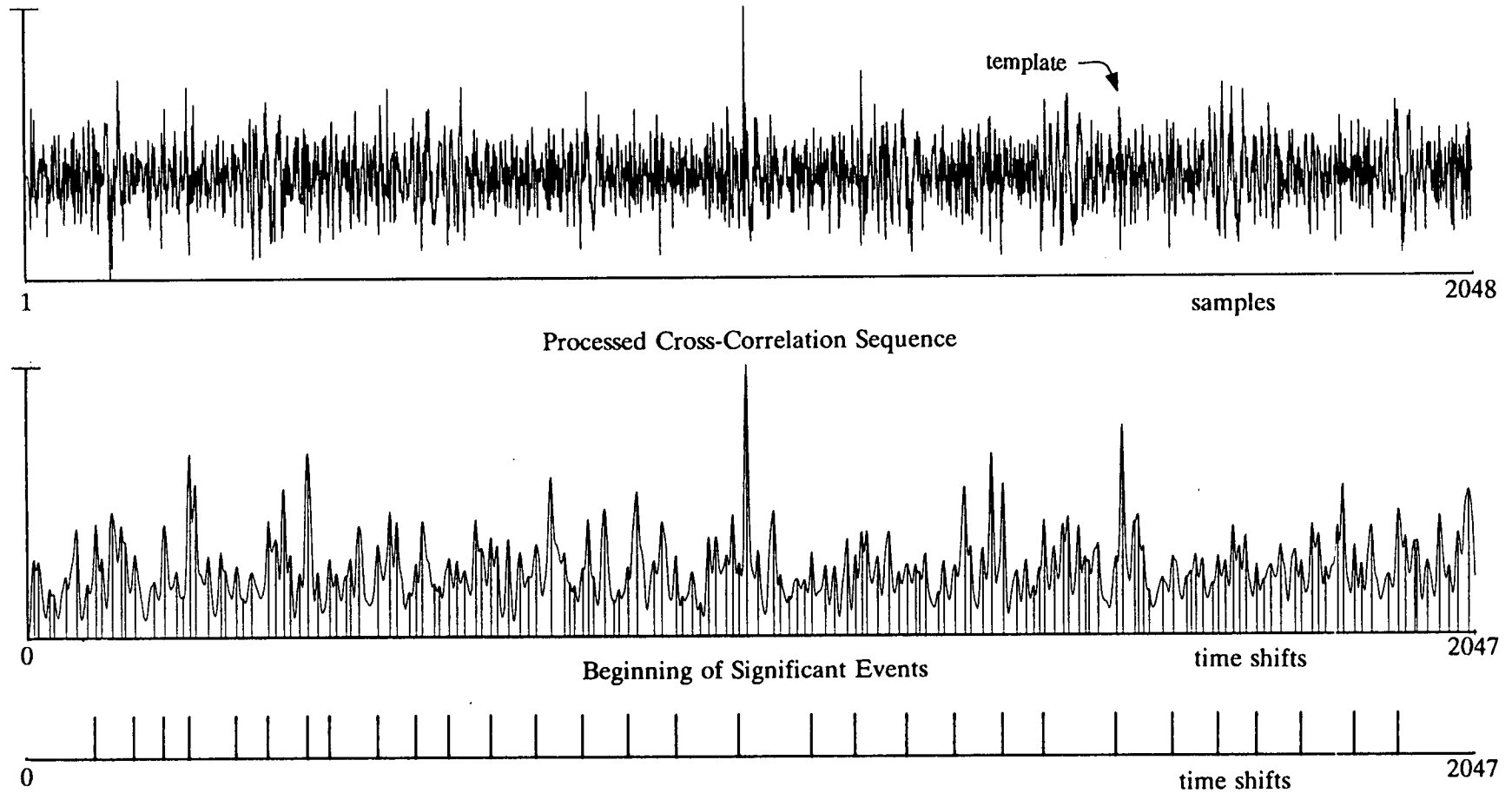


Figure 6.19a A phonocardiogram and its levels of abstraction. The processed cross-correlation sequence consists of the rectified and low-pass filtered output of the correlation of the phonocardiogram with the second heart sound template indicated. The vertical lines show the position of all the peaks present in the waveform. The bottom trace marks the *beginning* of those signal events in the phonocardiogram which are considered to be significant by the system.

PhCG: d16.AN.1

meta_ka_1

meta_ka_7

ka_22

Two contiguous sectors of peaks detected.

ka_1

No. of samples: 2048

ka_21

ka_4

ka_11

ka_9

ka_5

No solution islands existing (Data noisy).

S 2	mm	S 3	S 4	Final		mm
98	10.5	?	?	0	0	0.0
153	16.3	?	?	0	0	0.0
195	20.8	?	?	0	0	0.0
231	24.6	?	?	0	0	0.0
297	31.7	?	?	0	0	0.0
342	36.5	?	?	0	0	0.0
398	42.5	?	?	0	0	0.0
429	45.8	?	?	0	0	0.0
497	53.0	?	?	0	0	0.0
551	58.8	?	?	0	0	0.0
598	63.8	?	?	0	0	0.0
657	70.1	?	?	0	0	0.0
721	76.9	?	?	0	0	0.0
787	84.0	?	?	0	0	0.0
851	90.8	?	?	0	0	0.0
919	98.0	?	?	0	0	0.0
1008	107.5	?	?	0	0	0.0
1111	118.5	?	?	0	0	0.0
1173	125.1	?	?	0	0	0.0
1246	132.9	?	?	0	0	0.0
1314	140.2	?	?	0	0	0.0
1382	147.4	?	?	0	0	0.0
1440	153.6	?	?	0	0	0.0
1542	164.5	?	?	0	0	0.0
1622	173.0	?	?	0	0	0.0
1686	179.9	?	?	0	0	0.0
1740	185.6	?	?	0	0	0.0
1803	192.4	?	?	0	0	0.0
1878	200.4	?	?	0	0	0.0
1940	207.0	?	?	0	0	0.0

Figure 6.19b Analysis of phonocardiogram presented in figure 6.19a.

optimum filter threshold frequency has been found which maximizes the signal-to-noise ratio of the PhCG (chapter 3, section 3.7). Secondly, although cross-correlation accentuates the principal heart sounds, it does so at the expense of also accentuating the adventitious sounds. This is especially troublesome when late diastolic or late systolic sounds are present. The smearing effect of correlation and the proximity of the aforementioned events to the principal heart sounds interact to produce what appears to be a homogeneous group. When it is considered that the quality of the analysis depends to a large extent on the accurate determination of cluster starting times, this shortcoming needs to be rectified. Segmentation goes a considerable way to easing this problem (*e.g.* figure 6.9a), but is limited by the non-uniqueness of the attributes of such coalescences. In one instance such attributes may signify a coalescence, however, in another instance they may be indicative of a slight disassociation of the mitral and tricuspid components of a I. Thirdly, cross-correlation demands that a template of a second heart sound is available which must, in some way, be continuously updated to keep track of the evolution of the principal heart sounds. The problems associated with an updating scheme and an injudicious choice of template have been explored (section 6.1.2). As an expedience, the templates used to process the PhCGs in system 1 were all manually extracted from each PhCG. After an analysis system was in operation, an investigation of updating schemes was not pursued for the reason that the whole issue of processing the PhCG was being reexamined. This reexamination ultimately led to the deficiency in the 40 Hz filtering threshold becoming apparent, and the introduction of a new processing scheme which is described in chapter 7.

Another limitation, although subsidiary, is the restriction on the allowed range of the FHR to be 117-175 bpm. Although this range encompasses the 'normal' FHR of 120-160 bpm, the FHR, in exceptional circumstances, may range from 60 bpm to 220 bpm [7, 164, 165]. Any FHR which were to fall outside this set range in system 1

could not be tracked by the analysis. However, in all the PhCGs examined no such abnormal FHR was observed.

A common assumption with commercially available ultrasound FHR monitors is that the current instantaneous FHR is dependent on the FHR of the previous beat within certain limits called the short-term variability (STV). Whenever cardiac accelerations or decelerations exceed these limits false FHR estimates are given [166-168]. No such assumption was made in system 1 regarding short-term variability; all rate changes which fall within the permissible FHR range are accommodated irrespective of the STV.

6.5 SUMMARY AND CONCLUSION

This chapter describes the implementation of a system for the processing and analysis of foetal PhCGs which uses the knowledge-based methodology outlined in chapter 5. The various stages in the processing of the PhCG are introduced from the initial extraction of the PhCG from the phonogram to the segmentation of signal events. The analysis of a PhCG begins with a search for solution islands from which the system attempts to expand the analysis into those regions where the interpretation of events is not immediately apparent. A complete analysis is reached when all the principal heart sounds have been identified. In cases where the expansion is irretrievably terminated by high levels of noise, these solution islands, which are complete partial analyses rather than intermediate steps in an analysis, can still provide the required parameters of cardiac function albeit over a shorter time interval.

Although this system analysed PhCGs which had an inadvertently lower signal-to-noise ratio than was subsequently found to be achievable — owing to the choice of initial filtering threshold frequency — it performed very well over the range of PhCG

types presented. The ability to successfully analyse PhCGs which were corrupted by high levels of noise is an indication of the applicability and power of the knowledge-based methodology in this particular application.

CHAPTER 7

PHONOCARDIOGRAM PROCESSING AND ANALYSIS: SYSTEM 2

System 1 (chapter 6) — despite its limitations — has demonstrated the effectiveness of the application of knowledge-based techniques to the processing and analysis of the PhCG. Even in high levels of adventitious noise, system 1 is able to achieve an analysis. This has, therefore, encouraged the idea of a successor to system 1 which will overcome the limitations of the latter. This new system — system 2 — shares the same conceptual framework as system 1 but is a superior realization, and introduces facilities which are not found in commercial FHR monitors.

The high-pass filter threshold frequency used to obtain the phonocardiograms presented in chapter 6 had inadvertently reduced the signal-to-noise ratio of the signals. In spite of this, system 1 was able to produce an analysis of PhCGs in all but the most degraded cases. Of course, this degradation is simultaneously attributable to the underlying quality of the phonogram itself and the filtering used. The results presented in this chapter were obtained with a 20 Hz filter threshold frequency consequently the PhCG suffers less from background noise, and the principal heart sounds are more distinct, although the problem posed by the presence of adventitious sounds still remains. These latter sounds have similarly increased in amplitude as a result of the reduced filter threshold frequency.

The main limitation of system 1 has been identified as the tendency of the correlation processing to produce artificial mergers of events which are separate in time but very close together. The observed artificial coalescences have been between principal heart sounds and late diastolic or late systolic sounds. This problem, although aggravated by cross-correlation, is fundamentally attributable to the presence of adventitious sounds in the PhCG. These adventitious sounds arise from the very nature of the physiological signal which is being monitored and are present irrespective of the filter pre-processing applied. The number and location of adventitious sounds are the limiting factors in any PhCG event understanding system. Indeed the main objection to foetal phonocardiography, after poor signal quality, has always been the excess of adventitious events in the signal. Not only do the adventitious sounds obscure the PhCG but they are also the source of the frequent mistaken FHR estimates from commercially available monitoring equipment [16,45]. As the adventitious sounds are very similar to the principal heart sounds — especially second heart sounds — there is a restriction on what processing alone can do to eliminate them from the analysis. When presented with the choice of either attempting to identify and remove adventitious sounds at the processing stage or retaining all events as potential principal heart sounds, the latter option is to be preferred. The former option runs the risk of biasing the information presented to the analysis, whereas the latter makes the task of analysis more difficult.

The other potential limitation of system 1, namely the restriction which confines the FHR to lie within 115-175 bpm, has been relaxed, and now all FHRs which fall between the limits of 80-220 bpm are accommodated — irrespective of the rate of change. The assumption made in system 1 that there may only be one adventitious sound per beat has also been superseded. This significant relaxation of the restrictions imposed by the previous analysis system has been made possible by both the better

quality of the 20 Hz high-pass filtered PhCG, and the incorporation of more domain knowledge into the processing and analysis.

The objectives of the processing stages in system 2 are: to accurately identify and preserve the end points of all sounds, to overcome the smearing effect produced by the cross-correlation processing in system 1, and to eliminate any reliance on operator-assisted initialization schemes. The approach presented in this chapter meets the aforementioned objectives. A significant modification to the method of PhCG processing is presented. The method of signal analysis, however, still relies on the concept of expansion from solution islands. Results which illustrate the performance of system 2 on a range of PhCG types are presented and discussed with reference to system 1.

7.1 PHONOCARDIOGRAM PROCESSING

A visual inspection of the PhCG locates the principal heart sounds with little difficulty when the signal-to-noise ratio is good — as is suggested in chapter 3, section 3.9. This is achieved by initially fixing attention on the more ‘prominent’ events in the record. In system 2, an unconventional method of processing the signal is proposed which incorporates the notion of the *prominence* of a signal event. Segmentation of these ‘prominent’ events is then achieved on the basis of both a prominence measure and a pseudo instantaneous frequency spectrum across each event. In the following sections this new processing scheme and its implementation is described.

7.1.1 Excursion Processing

In the introduction reference has been made to the ‘prominence’ of an event without defining what is meant by the term in this context. Qualitatively, an event may be said to be ‘prominent’ if, to a visual inspection, it is more noticeable than

other events in that PhCG. The features which define 'prominence', it is proposed, are a relatively higher amplitude and an extended zero-crossing interval, although amplitude is the more significant attribute of the two. The time interval between zero-crossings is incorporated into the prominence criterion because empirical observations by the author have revealed that an increase in the value of this parameter is usually associated with the presence of a heart sound. These proposed attributes of prominence permit a formulation of a quantitative measure of prominence which is defined, on a per-excursion basis, to be the product of the duration of the zero crossing and the square of the amplitude of the signal. In this formulation, an excursion is taken to be that part of the PhCG bounded between two immediately adjacent transitions across that horizontal line which lies on the zero-amplitude level of the PhCG. The processing interval has deliberately been based on the excursion rather than the sound complex so that processing may be more localized. This strategy should result in the better preservation of event end points in the original PhCG. Although such an excursion-based measure of prominence has a similarity with the energy contained in a half cycle of a sinusoid, this is merely coincidental. Overall, this heuristic has been an effective means of accentuating the principal heart sounds, especially the stronger first heart sounds.

As the PhCG is by this stage a digital waveform, the temporal limits of an excursion may fall between points in the sequence. The determination of the end points is achieved by linear interpolation between those samples which have been found to lie immediately adjacent on either side of the horizontal zero-amplitude line. The beginning (or end) of the excursion is set at that point where the interpolation cuts the horizontal axis. The time difference between these limits is then the duration of the zero-crossing. The amplitude factor in the prominence criterion of an excursion is defined to be the largest sample value which occurs during that excursion —

providing there is only one turning point over the interval. Once both factors have been found, the 'prominence' is then calculated, and this, combined with the starting time of the excursion form the parametric description of the excursion. The KA which performs the processing and parameterization of excursions is `prominence_ka`[†]. All processing KAs operate under the control of `process_excur_meta_ka`.

The excursion processing outlined above is for the case when the PhCG has only one turning point between the beginning and the end of the excursion. It sometimes happens that two independent events in the PhCG, usually a first heart sound and a late diastolic sound, are so close together in time that the ending phase of one event coincides with the beginning of the other. Such an occurrence would have been smoothed by the cross-correlation processing employed in system 1 to produce an apparently homogeneous continuum. However, there is often a discontinuity in the amplitude between the two events which indicates such an approximation (figure 7.1). As the processing is effected locally, this discontinuity may first be searched for over an excursion before the excursion is processed. If it is found, it marks the vicinity of the beginning of another event.

Such a discontinuity is located by counting the number of peaks in the waveform over the excursion. A peak, in this instance, is defined to be a point in the PhCG, $p(n)$, $0 \leq n \leq N-1$, which satisfies for $1 \leq n \leq N-2$, one of the following conditions:

$$p(n-1) < p(n) \geq p(n+1) \quad \text{if} \quad p(n) > 0$$

or

$$p(n-1) > p(n) \leq p(n+1) \quad \text{if} \quad p(n) < 0$$

If more than one peak exists, then a discontinuity has occurred. The peak which is

[†]The KAs of system 2 are distinguished by name, *e.g.* `prominence_ka`, to avoid confusion with the notation used in system 1 (chapter 6).

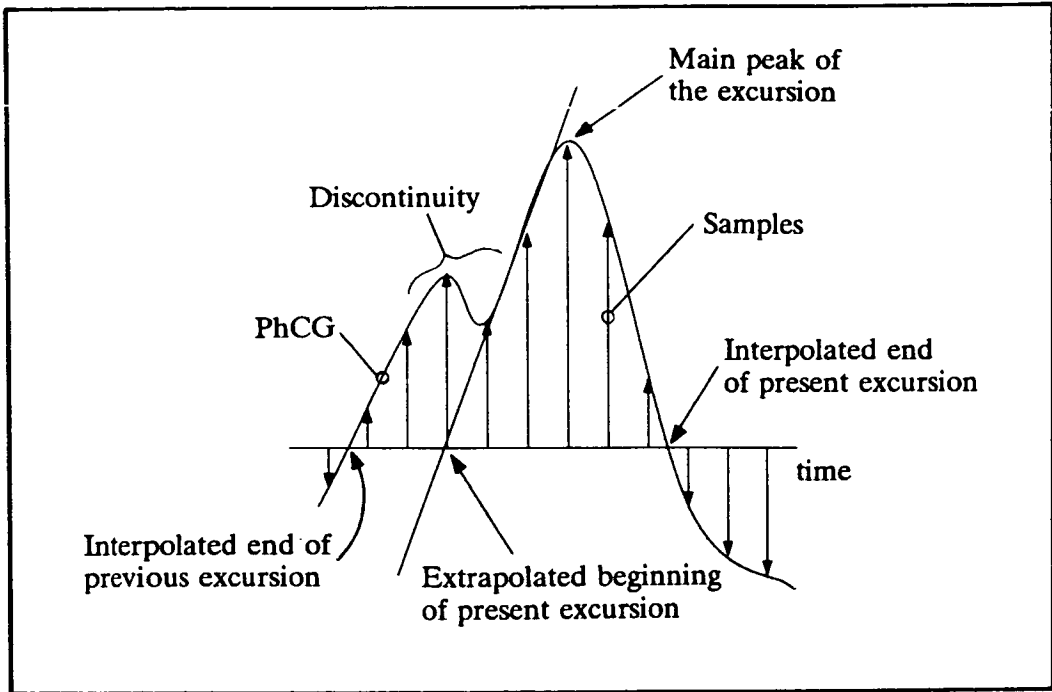


Figure 7.1 Significant points of an excursion.

chosen to represent the amplitude of the excursion is the one which is last to occur in that excursion.

Since the two sounds are so close together that the trailing and leading edges of both essentially overlap, the disassociation must be made more explicit. This is achieved at the expense of a slight curtailment of the trailing edge of the preceding event and is justified by the more pressing requirement to preserve the beginning of an event rather than its precise end. The approach adopted in these circumstances is to let the start of the excursion (figure 7.1) mark the end of the preceding event by assigning to it a prominence value of zero. As an approximation to the start of the following excursion — in the absence of its precursor — the points between the discontinuity and the last peak are extrapolated to the horizontal axis. This extrapolation is effected by a least-squares fit of a straight line to the points.

Occasionally, the extrapolation would seek to place the proposed beginning of the excursion earlier in time than the end of the previous excursion. This mostly arises from a low value of the differential of amplitude with time between the discontinuity and the peak. In this case, the extrapolated value is adjusted so that it falls one sample later than the end of the previous excursion. This adjustment is required to give a tangible indication of the presence of the discontinuity despite the overshoot.

Once an excursion has been processed, three parameters are produced: the starting time of the excursion, the prominence measure, and a flag which signifies whether the excursion contained a discontinuity, and if so, whether the extrapolated value was adjusted. This latter parameter is needed in the next stage of processing. An array of these ordered triples forms the first level of data abstraction. Examples of the processing of excursions in an arbitrarily chosen PhCG are illustrated in figure 7.2.

7.1.2 Removal of Insignificant Excursions

Although the excursion processing assigns prominence values to each excursion, within this class of prominence, there will be varying degrees of significance. The background noise level, which is higher in frequency and much lower in amplitude than the principal heart sounds, produces prominence values which are relatively low and which form a small fluctuating baseline just above the zero-prominence level (figure 7.2). These insignificant excursions may be removed by thresholding (`decimate_ka`). The level of this threshold has been set at 4.5% of the largest prominence value in the record under examination. Again from empirical observation, the excursions which lie below such a level are usually associated with the background noise level except for those instances detailed below. The threshold has been kept purposely low so that even very weak second heart sounds would not be discarded.

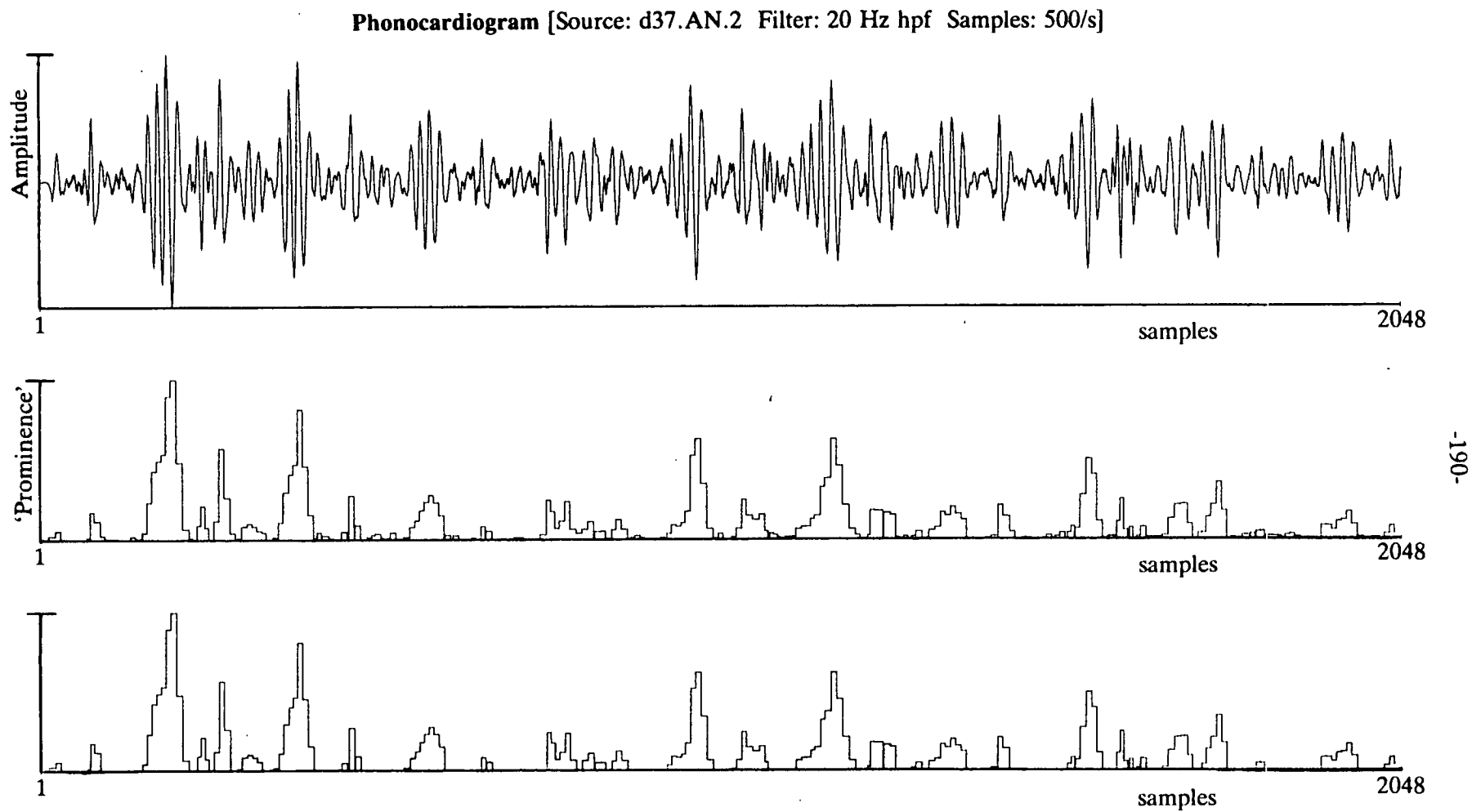


Figure 7.2 A phonocardiogram and the results from the first two stages of processing; *top trace*, an arbitrarily chosen phonocardiogram which has been high-pass filtered at 20 Hz; *middle trace*, sequence of processed excursions; *bottom trace*, sequence of significant excursions.

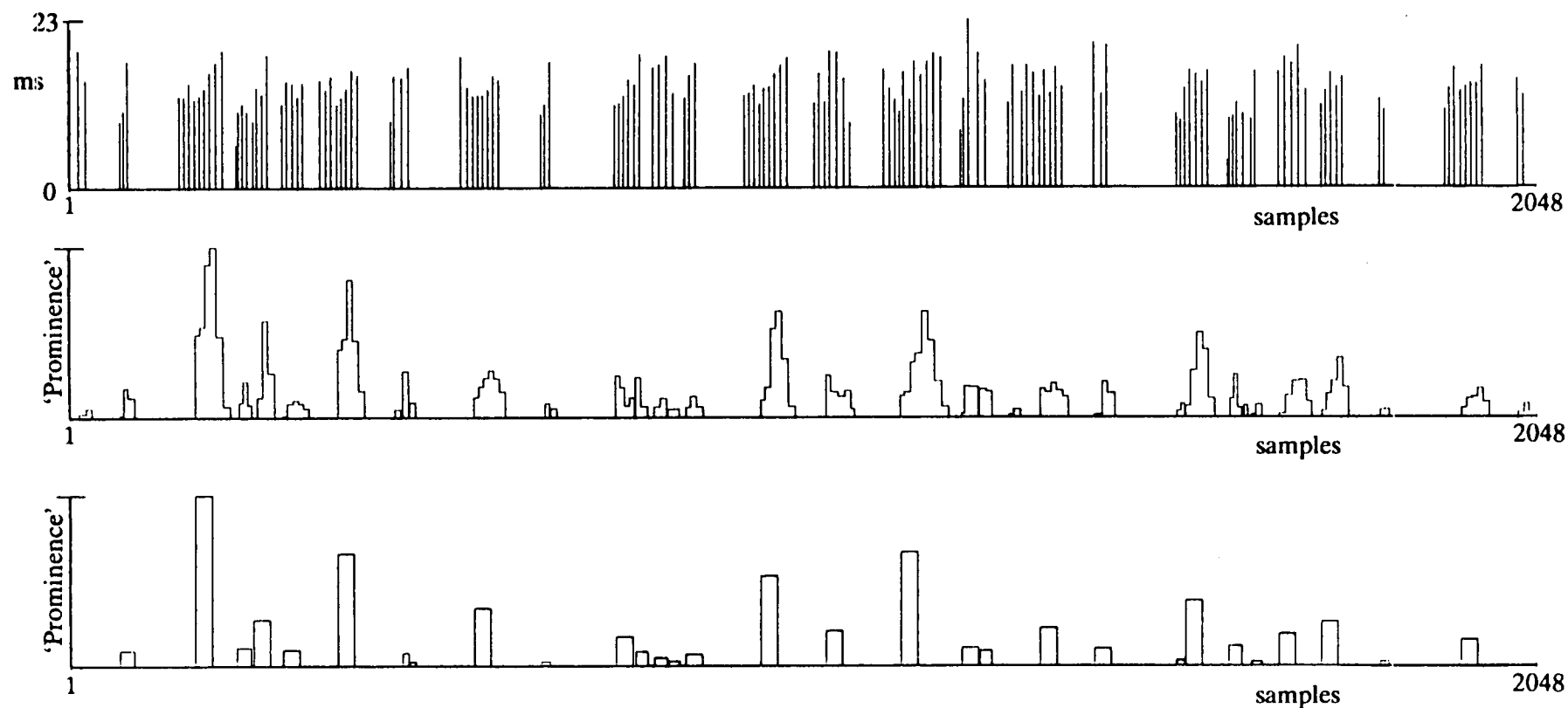


Figure 7.3 Further stages in the processing of the phonocardiogram presented in figure 7.2; *top trace*, durations of the significant excursions: each vertical line marks the beginning of a significant excursion, and its height is proportional to the duration of that excursion; *middle trace*, sequence of segmented events; *bottom trace*, sequence of 'reconstructed' events, top thirty ranks only.

Even if an excursion falls below the threshold, it may still be retained depending on the prominence values of bordering excursions and whether the excursion itself arose from the processing of a discontinuity. This facility is used to preserve low-lying precursors to potential principal heart sounds. At this stage, the beginning of potential heart sounds is taken to be every excursion, the prominence value of which exceeds the threshold. The following conditions define when an excursion is to be retained, even though its prominence value falls below the set threshold:

- (a) if the excursion did not arise from a discontinuity *and* the following excursion is above the prominence threshold *and* it too did not arise from a discontinuity
- (b) if the excursion arose from an unadjusted extrapolation *and* the following excursion is above the prominence threshold *and* it did not arise from a discontinuity
- (c) if the excursion arose from an adjusted extrapolation *and* the previous excursion is not above the prominence threshold *and* the following excursion is above the prominence threshold *and* the following excursion did not arise from a discontinuity *and* the excursion duration is longer than that of the following excursion (the latter condition reflects the observation that the instantaneous frequency across an event decreases with increasing time).

All those excursions below the prominence threshold which do not meet the requirements of (a), (b) or (c) above, are removed. Once all the insignificant excursions have been removed, each remaining excursion is represented by the ordered pair of starting time, and prominence — the discontinuity parameter is now discarded. An example of the results produced by the removal of insignificant excursions is illustrated in figure 7.2.

The variability in all aspects of an actual valvular sound prevents stringent

discriminating criteria with which to delimit significance, with the consequence that many more clusters of excursions are taken to be significant than are principal heart sounds in the PhCG record. This overestimate will depend on the quality of the signal. The advantage of a weak discriminant is that the probability of rejecting a principal sound as *insignificant* is small. Consequently the analysis stage of the system does not have to contend with biased information. This deferral of discrimination until more knowledge can be brought to bear prevents premature decisions. Its disadvantage is that the subsequent analysis is more taxed to ascertain which of all the events presented are the principal heart sounds.

7.1.3 Segmentation

Not all coalescences between adventitious sounds and principal heart sounds are identified by a discontinuity during an excursion. However, there are other clues which have been empirically discovered to reveal these coalescences, namely the distribution of prominence values and zero-crossing rate across an interval. The function of the segmentation knowledge area (*segment_ka*) is to detect such inhomogeneities in a group and to separate the group into its homogeneous parts, and, providing certain conditions hold, to remove an adventitious precursor to a potential principal heart sound.

In the explanation of the workings of the segmentation KA, it is assumed that the area under investigation is sufficiently far removed from the ends of the record that boundary conditions may be neglected. When all insignificant excursions have been removed, the processed data sequence consists of groups of prominent excursions which are separated from each other by excursions whose prominence values are zero, *i.e.* featureless gaps (figure 7.2). These groups of significant excursions are initially assumed to be homogeneous entities and are tested for evidence of inhomogeneity. If

a group fails the tests for inhomogeneity it is considered to be homogeneous. In the following description, reference is made to figure 7.4 which illustrates the control sequence of the segmentation KA.

A group of significant excursions is delimited by the first insignificant excursion encountered on either side of the group. These contiguous excursions are grouped together to form what is termed an *event* without any consideration being given to the possibility that the resulting event may be inhomogeneous. That is, the event may be composed of more than one sound. All subsequent application of processing KAs by the segmentation KA on this event takes place within the interval occupied by the event. Firstly, the number of excursions in the group is determined. If the number of excursions exceeds six, *i.e.* three cycles, the group is assigned to that part of the segmentation KA which deals with potential first heart sounds, otherwise it is assigned to that part which deals with potential second heart sounds. This test reflects the observation that second heart sounds often consist of two-and-a-half cycles of a quasi-sinusoidal wave (chapter 3, section 3.9).

It is assumed, for illustrative purposes, that the present event is a potential second heart sound. The next test determines if there is a trough in the distribution of prominence values across the event. For the excursion prominence values, $e_p(m)$, over the interval occupied by the event, a trough is defined to be:

$$e_p(m-1) > e_p(m) < e_p(m+1) \quad (7.1)$$

or

$$e_p(m-1) = e_p(m) < e_p(m+1)$$

where the latter conditional statement defines the beginning of the rising edge from a plateau. If no such trough is found, the prominence values were convex over the event. For these relatively short groups of less than seven excursions, this pattern has

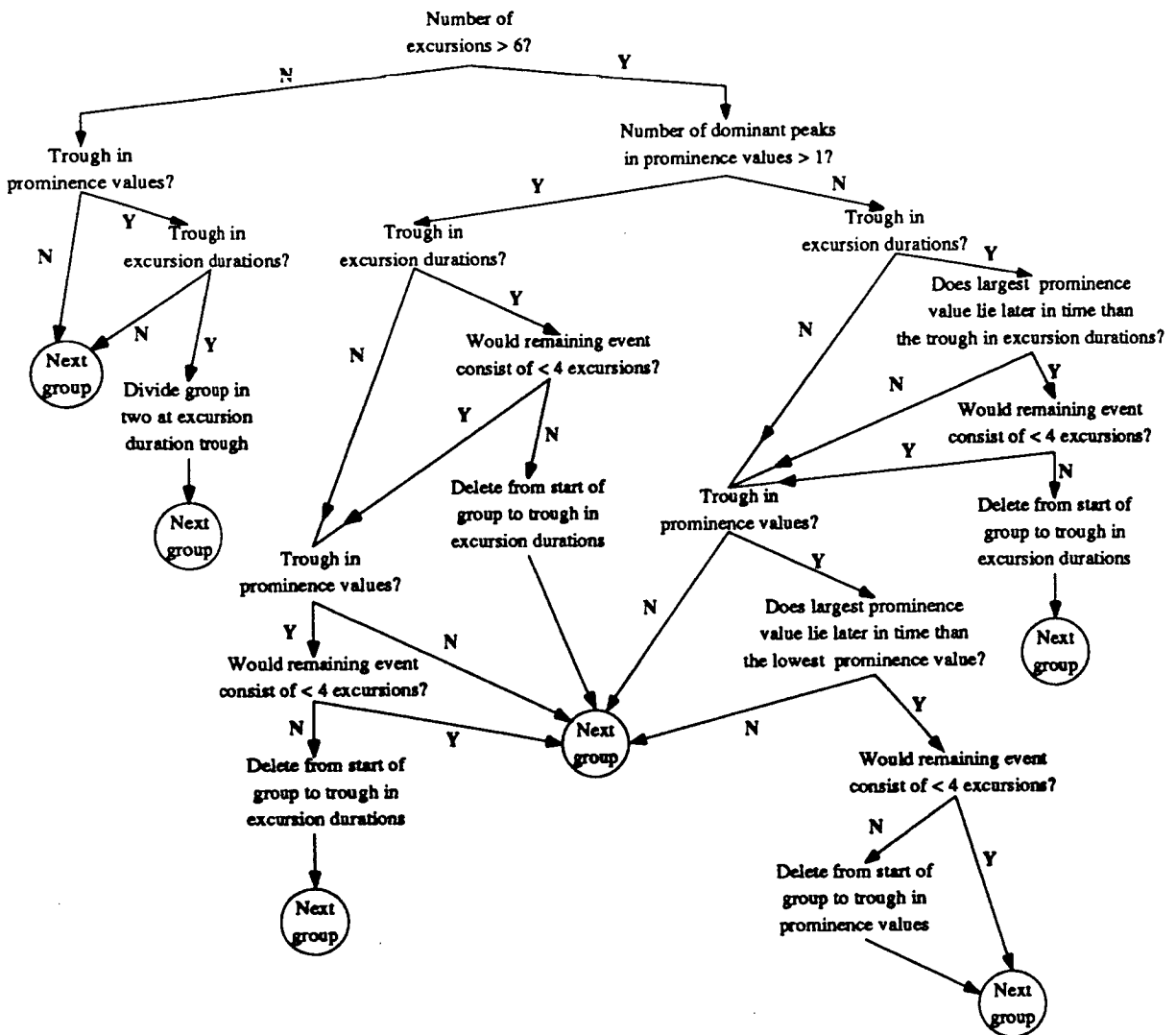


Figure 7.4 Control sequence of the segmentation knowledge area.

been found to be indicative of homogeneity. If a trough had been located, the relative durations of the excursions are then compared. It has been observed of all sounds registered in the PhCG, irrespective of origin, that each exhibits excursion durations which increase with time, *i.e.* sounds start at a relatively higher frequency and tend to fall to a lower frequency over the duration of the event. This phenomenon is also present in percussive musical instruments [169]. Thus if two sounds have coalesced, two such high-to-low frequency sweeps would be expected.

The same conditions as (7.1) are applied (except that duration values are substituted for prominence values) to search for the lowest excursion duration, *i.e.* the highest frequency, which occurs after the first high-to-low frequency transition. If no second such transition is found, despite the trough in the prominence values, the group is taken to be homogeneous. If such a transition is found, the event is segmented at the start of the excursion which has the lowest duration in the second high-to-low frequency transition. This is effected by curtailing the previous excursion by one sample point and assigning a prominence value of zero to an artificially created new excursion interposed between the previous excursion and the immediately following excursion. This artificial excursion has thus a duration of one sample interval. No attempt is made to ascertain if the coalescences arose from the approximation of a late systolic sound and a second heart sound, or of a second heart sound and an early diastolic sound. First heart sounds are generally of larger amplitude than late diastolic sounds which safely allows the precursor to be removed. However, this is not the case when late systolic or early diastolic sounds coalesce with second sounds, as the former and the latter are often indistinguishable in appearance. Once this group has been segmented, the processing moves on to the next group.

The segmentation of potential first sounds relies on similar tests to those applied

to potential second sounds. However, further tests are needed to confirm the hypothesis that an adventitious sound has coalesced with a first heart sound, before the adventitious sound is removed. Although the remaining event after the removal of an adventitious sound is termed a potential first heart sound, there is insufficient proof at this stage to instantiate it as such. Long diastolic sounds may well exhibit coalescences and similar characteristics. Thus sounds are only identified at the analysis stage when more global information may be applied. The potential first heart sound branch of the segmentation KA (figure 7.4) produces many permutations of predicate sequences which will not be explicitly described here. Rather each test — which is in effect a data-level KA — will be outlined below. Each test is based on empirical observations of associations of both prominence value and excursion duration attributes across all the 252 PhCGs of the 19 subjects investigated.

(a) *Is there more than one dominant peak among the prominence values of the group?*

Whenever there is one peak in an event this is naturally taken to be the dominant peak. If two or more peaks exists, there is potentially more than one dominant peak. The locally global peak is found, and if any of the other peaks in the event are greater than 75% of this value, they too are classed as dominant. The locally global peak in the prominence values, $e_p(m)$, is defined over the duration of the group to be that excursion which has the largest prominence value. If there are other excursions within the event having such a value then these are also classed as dominant peaks. Subsequent non-global peaks are defined by:

$$e_p(m-1) < e_p(m) > e_p(m+1)$$

for m ranging over the duration of the group. This test reflects the observation that a single sound usually produces one dominant peak. If two or more dominant peaks exist, this is indicative of a fall in the amplitude of the event

somewhere between its extremities and has often been empirically observed to be associated with a coalescence.

- (b) if this test follows a trough-in-prominence-values test, then: *Does the largest prominence value lie later in time than the position of the lowest prominence value in the group?* or if this test follows a trough-in-excursion-durations test, then: *Does the largest prominence value lie later in time than the position of the excursion which has the shortest duration?* This test is invoked when a potential coalescence with an adventitious sound has been detected. It is included to guard against the removal of potential first heart sounds which have an uncharacteristic distribution of excursion durations or prominence values which would otherwise appear to suggest a coalescence with another sound. If the initially assumed coalescing sub-event were to be removed, it would also remove that part of the event which is most locally prominent. This would imply that the first heart sound has a lower overall prominence than a late diastolic sound. It is known from empirical observation that adventitious sounds rarely exceed first heart sounds in prominence, therefore, if the outcome of the test is false, the event is assumed to be homogeneous.
- (c) *Would the remaining event consist of fewer than four excursions?* This test is invoked once a potential coalescence with an adventitious sound has been detected and a request is pending to delete the precursor. The intention of this test is to ensure that the original group which consisted of seven or more excursions (greater than three-and-a-half cycles) is not reduced to an event which has three excursions (one-and-a-half cycles) or less. Such a short event would be unlikely to be a first heart sound.

Those events which were found to be homogeneous, and those events the extent of which was reduced by the segmentation process, are posted in the data base as the

second level of data abstraction. An example of the segmentation produced by the above rules is illustrated in figure 7.3.

7.1.4 Removal of Insignificant Events

All homogeneous and 'homogenized' events presented by the segmentation KA are processed by a further operation (*preamble_ka*) which reduces the number of potential principal heart sounds before the events are transferred to the analysis subsystem. In advance of this, however, consideration is given to the possibility of a boundary effect produced at the beginning of the PhCG by the extraction of the PhCG from a much longer sequence. That event which occurs first in the PhCG is removed from consideration if its beginning lies within 60 samples (120 ms) of the start of the PhCG. The duration of this 'buffer zone' is sufficiently long to prevent a sound being admitted into the processing and analysis which may have been curtailed by the beginning of the record. A description is deferred until section 7.2.3.1 of the steps taken to avoid a potential boundary effect arising at the end of the PhCG.

Each event in the second level of data abstraction consists of a group of contiguous and homogeneous excursions which are each represented by individual prominence and excursion duration parameters. Such a detailed parametric description is no longer required as the processing is now moving to assume a more global perspective. Each event is reduced to three parameters: the starting time of the first excursion of the event, the time of occurrence of the end of the last excursion of the event, and the sum of all the constituent excursion prominence values. This ordered triple forms the third level of data abstraction; each entry on this level is still referred to as an event. The transformation of an event from an excursion-based representation to an event-based representation, is the final step in the erasing of the variabilities in the morphology of signal events in the PhCG. Each event is now primarily

distinguished by its overall prominence — all amplitude and excursion variabilities within an event have been absorbed into this global and homogeneous representation.

By this stage there may well be more than fifty events representing the original PhCG. This large number of potential heart sounds has resulted from the intentionally weak discriminating criteria by which is defined a significant prominent excursion and its derivative, the event. To reduce the uncertainty in the abstracted data, a restriction is placed on the maximum number of events which are to be presented to the analysis. In system 2, the fastest FHR considered is 220 bpm. Thus in a PhCG of length 2048 samples (4.1 seconds) at the fastest FHR there will be 15 heart beats. Each heart beat has two principal heart sounds, so in 2048 samples, the maximum possible number of principal heart sounds will be 30. Usually the FHR lies in the range 120-160 bpm (chapter 6, section 6.2.1) which would produce a maximum of 22 principal heart sounds. As a result of the weak discriminants used, the system may propose up to twice as many adventitious sounds as principal heart sounds for the case of a 'normal' FHR. Such uncertainty makes the task of analysis extremely difficult particularly since many of the proposed events are insignificantly small. To reduce the number of contenders to principal heart sound status is a necessity, if the analysis is not to be overwhelmed. Therefore, all the events are ranked in descending order in accordance with the sum of their constituent prominence values. All those events falling below the thirtieth rank are disregarded. This filtering operation is based on the assumption that first and second heart sounds have prominence values which would place them among the top thirty ranks in any PhCG — which is the case. If the FHR were at the extreme value of 220 bpm, the heart sounds are reported [36] to be very strong relative to the background noise level. It is assumed, therefore, that should the fastest FHR occur, the principal heart sounds would be above the level of adventitious sounds. Hence no principal heart sounds would be lost by this filtering operation.

7.1.5 Further Processing

More explicit use is made in system 2 than system 1 of the constancy of the systolic time interval over wide variations in FHR. The basis for assuming a constant systolic time has been introduced elsewhere (chapter 6, section 6.2.1). Using this relationship between first and second heart sounds is an attempt to limit the degrees of freedom allowed to the events by exploiting certain features which are unique to the principal heart sounds and which are not usually shared by adventitious sounds. The diastolic time interval was not used in this regard as it is relatively more variable in duration from beat to beat and may be twice as long as the systolic time. Of course situations arise when adventitious sounds, because of their relative positions, are also separated by a time equivalent to a systolic time interval. However, further constraints applied by the analysis should resolve these cases. The systolic time is to be used by the next processing stage to find all those events separated by this time interval (section 7.1.6). The majority of these pairs of events should be the principal heart sounds of each beat.

The systolic time is determined by the autocorrelation of events on the third level of data abstraction. This function is performed by the meta-KA called `sys_time_meta_ka`. A complete autocorrelation over the full duration of the signal is not required as physiological limits exist to the extent of the systolic time. This in turn defines the limits within which a 'reduced' autocorrelation may be performed. From Kelly's [75] extensive research on 144 foeti, it is known that the minimum and maximum systolic times are 150 ms and 210 ms. Thus, the autocorrelation need only start at a discrete time delay most closely equivalent to 150 ms relative to itself, and end 60 ms later. This is a 'reduced' autocorrelation in comparison with an autocorrelation over the full extent of the heartbeat at the lowest considered FHR (80

bpm) which would be 750ms long. Such a reduced autocorrelation also shortens the computation time in which to achieve an estimate of the systolic time. The autocorrelation function [88], $\phi_{e_h e_h}(l)$, (l is a time displacement or lag imposed upon a replica of the original sequence) of the events on the highest level of data abstraction, $e_h(n)$, $0 \leq n \leq N-1$, is defined to be:

$$\phi_{e_h e_h}(l) = \frac{1}{N - |l|} \sum_{n=0}^{N-|l|-1} e_h(n) e_h(n-l) \quad \text{for } l_{\min} \leq l \leq l_{\max} \quad (7.2)$$

The unbiased autocorrelation used in (7.2) takes account of the fact that the autocorrelation estimate for each lag is formed over an ever-decreasing number of samples. The lower and upper extent of the relative time displacements between the two sequences (l_{\min}, l_{\max}) should idealistically be set at the limits of the systolic time which were empirically determined by Kelly [75]. to be, respectively, 150 ms and 210 ms (75 samples and 105 samples at 500 samples/s). However, to allow a margin for error in the estimated location of the principal heart sounds, this range has been expanded to cover the interval from 62 samples to 117 samples (124 ms to 234 ms).

The events on the highest level of data abstraction have had all the original morphological information removed which relates to the corresponding events in the PhCG. Each event is represented by a beginning and end time, and the sum of the prominence values across the event. To convert these parameterized events into signal events upon which autocorrelation may be performed, necessitates the reconstruction of the PhCG from these events in a highly abstracted version. This reconstruction is performed by taking each event in turn, and converting it into a rectangular pulse which begins and ends at the same point as the event. The amplitude of this pulse is set at the value which represents the sum of the prominence values of the event. The intervening gap between events is instantiated to zero at every sample point. These operations gives rise to a signal comprised of rectangles with different heights and

widths (figure 7.3). By extrapolating the prominence value across the duration of the whole event, this has given added emphasis to those events which are relatively longer in duration, and have a higher overall prominence value, namely, the first heart sounds. In an autocorrelation of these rectangular events, the correlation of first and second heart sounds will occur at the value of temporal lag index representing the systolic time.

A further piece of domain knowledge is incorporated into the autocorrelation to decrease the computational load and to form a better estimate of the duration of the systolic time. It is known from the nature of first heart sounds, and empirical observation, that the largest events existing in the reconstructed signal are first heart sounds. Thus, only those larger events which are shifted by the autocorrelation processing onto smaller events (potential second heart sounds) are admitted into the calculation of the autocorrelation function at each value of lag l . A smaller event shifted onto a larger event is either a second heart sound just beginning to correlate with the following first heart sound or an adventitious sound which is displaced by the equivalent of a systolic interval from the position of the following first heart sound. These phenomena give rise to a non-monotonic autocorrelation function. It is difficult to uniquely associate the peaks of this function with any of the three possible underlying causes. At the higher FHRs, when the systolic and diastolic times are approximately equal, this reduced autocorrelation will reject second heart sounds which are correlating with first heart sounds. This is of no consequence, as the autocorrelation still accepts the correlation produced by first heart sounds being shifted onto second heart sounds — which was the original intention.

The correlation of two rectangles produces a triangle-shaped autocorrelation function. Thus, when examining the autocorrelation function resulting from the

reconstructed signal, a distinct and global peak is produced in the correlation function at a temporal lag value equivalent to the systolic time. To further enhance this peak, and the monotonic increase and decrease of the autocorrelation function, all events in the reconstructed waveform are limited to a maximum duration of 23 samples (46 ms). The temporal curtailment of events is effected by altering the end of an event rather than its beginning. This empirically-derived value of 46 ms is the minimum duration of a second heart sound. If the longer duration events (usually first heart sounds) are restricted to this duration, at a lag equivalent to the systolic time, they must fully overlap with the potential second heart sounds. If the longer events were to retain their original duration, then a fragmented peak would arise in the autocorrelation function which could not be uniquely associated with the systolic time interval.

7.1.6 Pairing of Events

Once an estimate of the systolic time has been obtained from `sys_time_meta_ka`, all pairings of potential first and second heart sounds may be obtained. This is a first step towards finding solution islands (chapter 6, section 6.2.2) in the PhCG. Although this matter is fully addressed in the next section, solution islands in system 2 are composed of I-II pairs rather than II-I-II triples as in system 1. The pairing of first and second heart sounds is a consequence of a more explicit reliance on the constancy of the systolic time interval, and the fewer assumptions made in system 2 about the nature of adventitious sounds and the limits of FHR.

The pairing of events is based on the estimate of the systolic time and the relative displacements of the events in the 'reconstructed' signal at the instant when the autocorrelation function reaches its peak in the interval 62-117 sample lags (section 7.1.5). To explain the methodology used by the event pairing meta-KA (`pair_meta_ka`), it is assumed that the current focus-of-attention is sufficiently far

removed from the boundaries of the record that end effects may be ignored. To find the pair of an event involves searching in an interval which is 23 samples long (46 ms) situated further back in time by a duration equal to the systolic time. The width of the interval to be searched, is the maximum duration of any event in the reconstructed signal after event curtailment has taken place. This interval represents the limits within which the unshifted event must lie which correlated with the shifted event, and which contributed to the peak in the autocorrelation function. In this interval there may be more than one event. Figure 7.5 illustrates the eleven possible ways that the potential pairs may distribute themselves over the interval of length 23 samples.

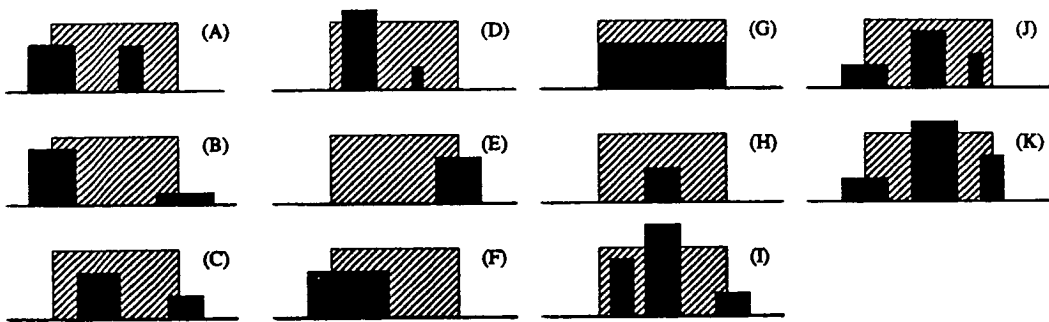


Figure 7.5 All eleven possible allowed orderings of one, two or three events in an interval of length 23 samples (46 ms). Signal events are represented by arbitrarily dimensioned black rectangles; the shaded region defines an interval of length 23 samples.

The maximum number of events to be considered in the interval is three — it is unlikely that more than this could be accommodated in such a short duration.

Each event in the interval is assessed for suitability as the pair of the event a systolic time away, on the basis of the contribution it singularly made to the peak in the autocorrelation function. In the case where only one event is present in the interval, it is immediately paired. The situation, denoted by the letter 'A' in figure 7.5, is arbitrarily chosen to illustrate the case when two events are simultaneously in the interval. Only that part of each event which falls within the interval of length 23

samples is considered (shaded region in figure 7.5). Thus the beginning part of the first event is neglected (events are represented by black rectangles). As the event which was shifted has uniform amplitude, each unshifted event will have contributed to the autocorrelation peak an amount equal to the sum of the products of both amplitudes over that part of the event which falls within the shaded region. Without loss of generality, if the amplitude of the shifted event is assumed to be unity, then each unshifted event contributes an amount equivalent to the product of its amplitude and the extent of the overlap with the shifted event. Whichever of the two events makes the greater contribution is chosen to be the other member of the pair. If both blocks have made the same contribution, then that event, the starting time of which is closer to the systolic time, is chosen. If both events are temporally equidistant from that point which marks the systolic time ahead of the shifted event when the latter event is in the unshifted position, then whichever of the two events has the higher amplitude is chosen. This approach is used for all cases, including those with three events within the interval. The pairing meta-KA progresses across the 'reconstructed' signal and pairs each event in turn — providing a suitable corresponding event is located. Those events which have already been paired with an event earlier in time may themselves be the originator of a further pairing with a later event.

7.2 PHONOCARDIOGRAM ANALYSIS

The events presented to the analysis sub-system are in pairs. Each event in a pair is represented by its beginning time, and the sum of its constituent prominence values. Unlike system 1, the analysis of system 2 has direct access to the lower levels of data abstraction in the data base. These levels, specifically the second and third levels, present the processed signal after segmentation, and after all but the top thirty ranked events have been removed. Allowing the analysis to revisit the lower levels of abstraction is a means by which the analysis can compensate for the deficiencies in the

processing of excursions and events. This is a powerful concept, which although in a rudimentary form in system 2, allows the analysis to specifically tune the processing stages to either search more diligently for an expected but missing occurrence in a particular region, or to acquire more information either in particular or in general. Thus the whole system converges towards a complete analysis by way of the synthesizing of information extracted from the analysis into the reprocessing of the original signal. This information-feedback mechanism from the higher to the lower levels of data abstraction allows the weight of contextual information on the highest level to be brought to bear on areas where initially local processing was insufficiently effective.

The stages in the analysis of the PhCG are the finding of solution islands and the expansion of the analysis from these resolved regions into the ambiguous areas which lie on either side. Before solution islands are located, however, all those pairs of events are removed which are unlikely to be first and second heart sounds couples.

7.2.1 Removal of Unlikely Event Couplings

As a result of the pairing procedure undertaken by pair_meta_ka (section 7.1.6), all those events separated by a time approximately the duration of the systolic time have been formed into couples. These couples, because they are separated by approximately a systolic time, are plausible first and second heart sounds. However, in the pairing meta-KA, no consideration was given to the relative prominence of each event which forms a couple. It is the function of uncouple_meta_ka to reexamine these pairs to ascertain an unlikely I, II pairing which may be indicated by the relative prominence of each member of the pair. It is unlikely that a couple, the earlier member of which has a small prominence and the later member of which has a large prominence, could be a I and a II respectively. As first heart sounds have consistently

large prominence values because of their large amplitude and relatively long durations, only those couples which originate in a large-small event prominence pairing are admitted. 'Smaller' in this context is defined to be 50% or less of the prominence value of the earlier event in the pair. The brief of the KA which performs the investigation of these pairs (single_pair_ka) is to remove the pairing of adventitious sounds in the diastolic region with first heart sounds.

Another observed coupling phenomenon is referred to as a *continuously linked triple*. This arises, for example, in the case of a I pairing with a II which also pairs with a diastolic sound. It is assumed, in this configuration, that the events are not an instance of a large to small prominence pairing as described above. In these circumstances, if the first pair has a relative prominence of large-small, and likewise the second pairing, then the latter pairing is removed while still retaining the first pair, *i.e.* the third member of the triple is disregarded (double_pair_ka). Such a continuously linked triple has often been empirically associated with a I, II, and mid-to-late diastolic sound.

The final type of pairing investigated is the case when pairings cross over one another. This type of occurrence is nearly always associated with (a) a diastolic sound paired with a later systolic sound, and between the diastolic and systolic sound, the first heart sound is paired with the second heart sound which follows the systolic sound, or (b) a systolic sound paired with a late diastolic sound, the pairing of which crosses over the pair created by a first heart sound and the second heart sound, the latter of which lies between the systolic and diastolic sound. This situation may be resolved in certain cases by using the criterion that adventitious sounds often have smaller overall prominence values than the principal heart sounds. If the later pair in the crossed pairing is examined, a threshold level is set at 50% of its overall

prominence value. If both events in the earlier pair which cross over it are below this threshold, then the earlier pair is removed (`crossed_pairs_ka`).

7.2.2 Location and Analysis of Solution Islands

The initial solution islands in system 1 (chapter 6, section 6.2.2) were of the II-I-II type. This particular configuration arose as a result of the assumptions made, namely, after segmentation a maximum of one adventitious sound would remain per beat, and the maximum considered FHR was 175 bpm. In system 2 no such assumptions were made, so the technique by which solution islands are located and instantiated in system 2 cannot use the previous methodology. The only assumption that is made in system 2 concerning the duration of the heart beat, is that the systolic time is constant — which is a most valid assumption. This, combined with the relatively larger overall prominence values of the first heart sounds, would suggest that the I-II unit, in that order, is the natural choice for the basis of solution islands.

A potential solution island is located by `sol_isl_meta_ka` by finding two events which are paired but do not form pairings with any other event(s). In itself, this is insufficient proof to instantiate a I-II solution island. To verify I-II status, an 'exclusion zone' is established around this particular pair. This zone extends to a duration equivalent to twice the systolic time before the first member of the pair, and twice the duration of the systolic time after the second member of the pair. By this is defined the maximum separation between one heartbeat and the next which occurs at the lowest FHR. In all, therefore, this zone extends for a duration equivalent to five systolic times. Irrespective of the FHR, one heartbeat must occur in any interval of duration equivalent to three systolic times[†]. Thus, within an interval equivalent to five systolic times, there must be *at least* one complete beat and the II of the previous

[†]The maximum duration of the diastolic time is twice the systolic time [75].

beat and the I of the following beat. At the other extreme, when the FHR is at its maximum, the systolic and diastolic times are equal [75]. Therefore, within an equivalent of five systolic times there are, at the very limit, a maximum of three heart beats. When both these cases are combined, they provide conditions which a potential event pair must satisfy before it can be instantiated to a I-II solution island.

If the later part of the 'exclusion zone' is examined first, then within this interval only the first member of another pair is allowed to occur, *i.e.* another potential I-II couple may not begin and end within the interval. The converse applies in the initial part of the 'exclusion zone'. Here, only the second member of another pair is allowed to occur, *i.e.* another potential I-II couple may not begin and end within the interval. Providing both these conditions are satisfied, the event pair is instantiated to I-II. This process continues until the end of the record is reached. If no solution islands are found, that PhCG is rejected because of the presence of too much uncertainty. Processing then begins on the next PhCG presented to the system.

7.2.3 Expansion from Solution Islands

Once solution islands have been found and instantiated, the strategy meta-KA (figure 7.6) attempts to expand the analysis from the solution islands into the unresolved regions which may lie on either side of the island. The process of expansion is conducted in a similar way to that of system 1 (chapter 6, section 6.2.3), *i.e.* by establishing a time window within which an expected event must occur, and searching for that event between the limits of the window. To illustrate the slightly different method employed in system 2, it is assumed that the analysis expansion meta-KA (`expand_pairs_meta_ka`) is expanding the analysis backwards in time from a I-II solution island. From Kelly [75], it is known that the closest a following I can be to a II is at the fastest FHR, when the systolic and diastolic times are equal. As the

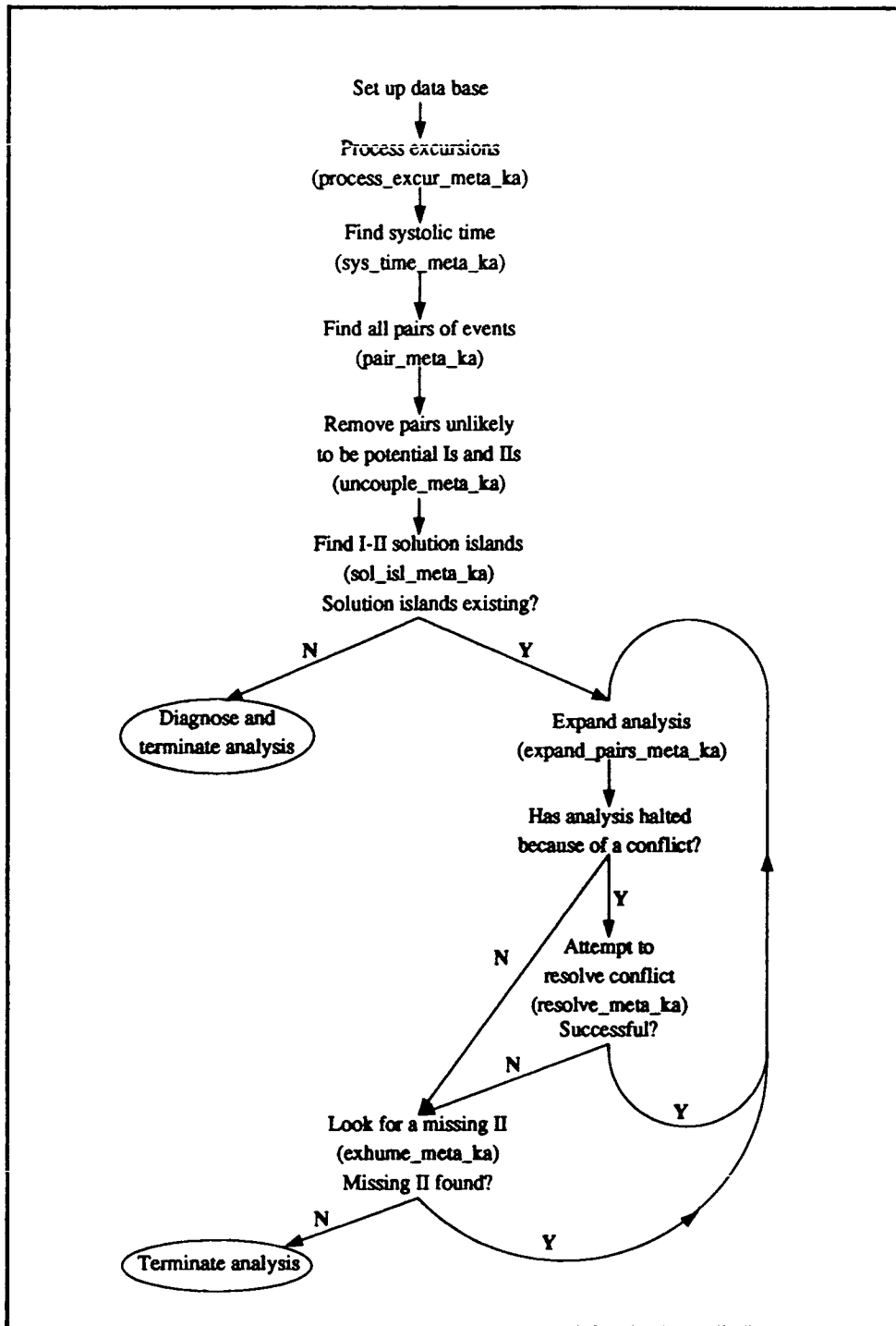


Figure 7.6 Control sequence of the strategy knowledge area (system 2).

systolic time has already been determined (sys_time_meta_ka), this effectively establishes the lower limit of the window at the time of occurrence of the II of the solution island plus an additional time equivalent to the systolic time. Again from Kelly, the maximum separation of a following I from a II occurs at the slowest FHR, when the diastolic time is twice the systolic time. Thus the upper limit of the window is at a time instant equivalent to a systolic time further back from the lower limit of the window. The window, therefore, occupies an interval equivalent in duration to the systolic time. This is a narrower window than that employed in system 1 which used the theoretical limits of principal heart sound separation, irrespective of the value of systolic time in that record. In system 2, it is possible to employ a narrower window because of the constancy of the systolic time; this fact was not exploited in system 1.

With the instantiated I-II as a bridgehead, the uncertainty in the data may be reduced. Any pair is removed, the first member of which originates between the II and the lower limit of the window. Such a pairing must have arisen between a diastolic sound and one of the following: a late diastolic sound, a first heart sound, a systolic sound. If one pair alone originates within the interval occupied by the window, it is instantiated to I-II. In system 2, expansion of analysis is achieved two events at a time, *i.e.* either by deletions or instantiations. This has the advantage that the uncertainty in the record is reduced more quickly than in system 1, and is a result of events being treated as couples rather than single entities.

Expansion continues in this way until either two pairings of events originate in the window interval, or the end of the record is approached. The latter case is examined below; the former is deferred until section 7.2.4. Expansion of analysis may also proceed in a direction towards the beginning of the record. The same expansion procedure is adopted, except that the limits of the window fall earlier in time with

respect to the I of the I-II solution island. The lower and upper limits of the window are set earlier than the I of the solution island by a time equivalent to a systolic time, and two systolic times respectively.

7.2.3.1 Phonocardiogram Boundary Conditions

When the limits of the observation window are established, a check is made to ensure that the sample number representing the upper limit of the window does not exceed the total number of samples in the record. If the window does extend beyond the end of the record, the expansion of analysis terminates without considering those events occurring after the last I-II pair.

If the upper limit of the window falls completely within the record, but is close to the end of the record, the analysis must consider the possibility that events may be missing as a result of the imposed artificial termination of the PhCG. A 'buffer zone' is established at the end of the record and is of a length equivalent to a systolic time plus 70 samples, back from the end of the record. Any pairs which originate in the observation window and terminate in this 'buffer zone' are considered to be sufficiently far from the end of the record that they are completely intact. To explain the origin of the duration of this interval, it is assumed that the upper limit of the window just falls on the earlier edge of this zone. The upper limit of the window defines the position of maximum separation of a potential I from the II of the solution island. If a potential I does lie at this extreme limit, the corresponding potential II must occur later by a time approximately equivalent to the systolic time — hence the systolic time component in the observation window. The additional 70 samples are included to provide a sufficient time interval for the potential II to completely manifest itself. A I falling on the upper limit of the window is the extreme case which occurs at the lowest FHR, when the diastolic time is at its maximum extent. A 'buffer zone'

of the length indicated, is required to accommodate this eventuality. Unlike system 1, the processing of the potential boundary effect at the end of the PhCG is delayed until contextual information is available. This is done to preserve as much of the original PhCG as possible. The compensation for a possible end effect at the beginning of the PhCG has already been described (section 7.1.4).

7.2.4 Conflict Resolution

The expansion of analysis is halted whenever more than one event pair originates within the window extended from the last identified I-II solution island. In these circumstances, control is transferred to the meta-KA, `resolve_meta_ka`, which attempts to resolve the conflict. Each event pair which originates within the limits of the window is a potential I-II. This is a difficult conflict to resolve as the pairs are apparently equally valid first and second heart sounds. The test which is applied to the pairs which originate in the window, takes each pair separately and hypothesizes that it is an actual I-II pair. Then, for each hypothesized pair, the test investigates the position of the next supposed I-II pair which would result, were the hypothesis to be verified. To illustrate the methodology, it is assumed that the expansion is progressing back in time and that there are two potential I-II pairs in the window (figure 7.7). The first pair is hypothesized to be a I-II, and another window, of duration equivalent to a systolic time, is established at a time equivalent to a systolic time later than that of the II of the hypothesized I-II. The number of pairs originating in this new window is determined. The same procedure is then applied to the second hypothesized I-II pair in the original window. Again, the number of pairs is counted which originate in the window which is established at a time equivalent to a systolic time interval later than the hypothesized II of this second potential I-II in the original window. If both counts are zero, then events are missing in the region. A diagnostic message is issued to this effect. If one count is zero, and the other non-zero, this is taken to be sufficient

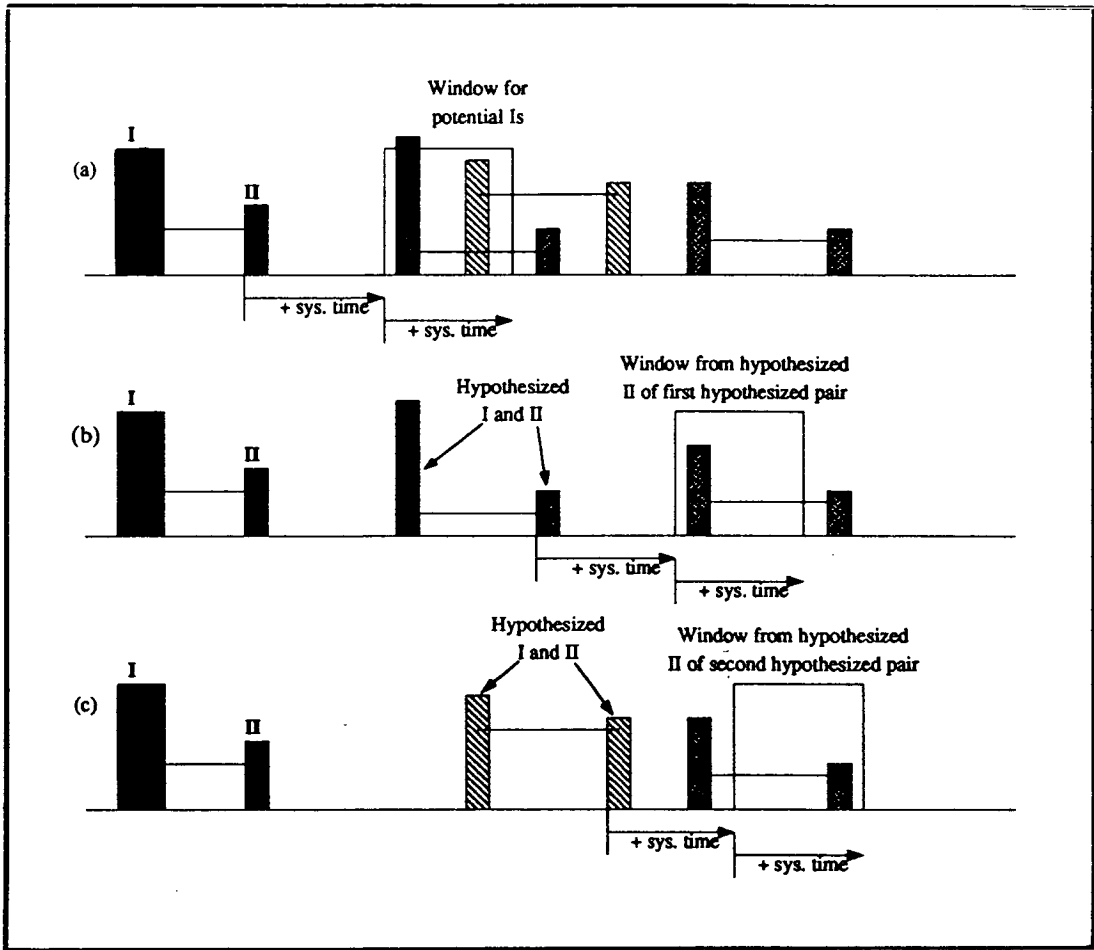


Figure 7.7 Example of the method of conflict resolution. A link between rectangles (events) indicates a pairing. (a) Initial situation with two potential first heart sounds competing in the window; (b) first pair originating in the window is hypothesized to be a I-II; (c) second pair originating in the window is hypothesized to be a I-II. Only the first pair of events in the conflict set satisfy the requirements for I-II status.

evidence to instantiate one of the hypothesized pairs to be a I-II. A count of zero indicates that if the corresponding hypothesized I-II were to be instantiated to a I-II, this would then imply the non-existence of the originator of a pair in the following allowed window region for a I. This is an unlikely situation to arise as actual first heart sounds are rarely missing from a PhCG[†]. Therefore, the window established by that hypothesized I-II pair is situated in a region where first and second heart sounds do not exist. However, in the window established by the other hypothesized I-II pair, there exist originators of pairs implying that one of these is the originator of the next I-II pair. This latter hypothesized I-II pair is instantiated to be an actual I-II, and the invalidated pair is removed.

The test described above is proved inconclusive in those situations when both windows erected by the hypothesized first and second heart sounds, return counts which are both greater than or equal to unity. The policy adopted in this circumstance, is to apply a prominence criterion. If one of the originators of pairs, in the window erected by the last positively identified I-II pair, has an overall prominence value which is greater than twice the next highest overall prominence value in the window, this pair is then instantiated to be a I-II and the adventitious pair is removed. This test reflects the large prominence values arising from first heart sounds because of their higher amplitudes and longer durations.

The conflict resolution strategy outlined above is also applied to expansions which progress towards the beginning of the record. If a conflict has been resolved in either direction, `expand_pairs_meta_ka` is recalled to expand the analysis further.

[†]The only two observed cases when a first heart sound was severely attenuated are illustrated in figures 3.7a and 6.12a. These are the only occurrences in 252 PhCGs of 4 seconds each from 19 subjects.

7.2.5 Missing Second Heart Sounds

Despite the precautions taken by the processing stages to retain as many of the signal events as possible, the case may arise that a second heart sound is missing on the third level of data abstraction. This situation may arise from a missing II in the original PhCG, or a particularly weak II which does not fall within the top thirty ranked events. This latter case is often the result of a poor signal-to-noise ratio. A missing second heart sound[†] is identified by the non-existence of the corresponding I-II pair in a region where the pair is expected. Such a region falls between two I-II solution islands which are separated from one another by an interval which is greater than, or equal to, a time equivalent to three systolic times. Within an interval of three systolic times, one heart beat must occur. This is a corollary of Kelly's observation [75] that the maximum duration of the diastolic phase is twice the systolic phase. If a missing II is found, the corresponding I, which is its pair, must also be missing from the array of event pairs. This is not to imply, however, that the I is also missing in the PhCG, rather, because the corresponding II of the pair is missing, the I was not included among the pairs of events. To resolve this situation, the excluded I must be found before the missing II is searched for (`exhume_meta_ka`).

The I is initially sought on the third level of data abstraction (section 7.1.4). It is known that the I must arise in a window of duration equal to the systolic time, which begins at the equivalent of a systolic time later than the II of the earlier I-II solution island. If no events are found in this window interval, it is assumed that this indicates the unusual occurrence of a missing I. This region is abandoned by `exhume_meta_ka`, which then searches for missing events in other regions. On the other hand, if one event alone is found within the limits of the window, and the window is not close to

[†]First heart sounds are assumed to be present on all levels of data abstraction because of their large prominence values.

the end of the record, there are two possible underlying causes: (a) the II pair of the I is included in a I which apparently persists for the duration of the systolic time, or (b) the II is missing on the third level of data abstraction.

The first possibility, (a), may indeed indicate an elongated I, or a systolic sound which has formed an apparently homogeneous event with the I, and the corresponding II. `exhume_meta_ka` is alerted to both these possibilities by examining the duration of the one event which occurs in the window interval. If the end of the event exceeds its beginning by more than the equivalent duration of a systolic time, then such a elongated event has occurred. In this case, the beginning of the event is assumed to indicate the start of the I, and because the II is indistinguishable, a synthetic II is created at a time equivalent to a systolic time following the I. This instantiation is only performed once a check has been made that this probable I-II satisfies the 'allowed-window-for-a-II' condition as instigated by the immediately following I-II solution island. If this condition is not met, the event is abandoned by `exhume_meta_ka`.

The second possibility, (b), indicates that the II is not present on the current level of data abstraction. The search for the missing II is extended by `exhume_meta_ka`, to the second level of data abstraction (section 7.1.3)[†]. As the once-excluded I has been located, this allows a window to be established, within which the II must lie, on the second level of abstraction. This window is 46 samples long (92 ms), *i.e.* twice the window width used in `pair_meta_ka` (section 7.1.6), and begins at 23 samples less than a time equivalent to a systolic time away from the once-excluded I. The width of this window has been extended to increase the likelihood that the II will fall within its limits. If no event is found in this window, then the II is either missing or has been severely attenuated. In this case, a synthetic II is created at the

[†]The stage of processing before the ranking of events to admit only the thirty most prominent events.

point where the II would most likely have been, *i.e.* a systolic time interval behind the I. This is done to allow the analysis to continue on the highest level of data abstraction. Before instantiation takes place, however, this hypothesized I-II is tested against the 'allowed-window-for-a-II' condition by the immediately following I-II solution island. A statement is issued to the effect that a synthetic event has been created. If events do exist within the window on the second level of data abstraction, that event closest to a systolic time duration from the once-excluded I is taken to be the missing II. Again, this is verified before instantiation takes place.

Conditions (a) and (b) above examined the case of one event in the allowed window interval which was extended from the earlier I-II solution island. The other possibility is that there is more than one event in this interval is now considered. There are two cases which may give rise to this situation: (1) the II has been absorbed by a systolic sound and is apparently homogeneous to the segmentation KA, or, (2) the II is missing on the third level of data abstraction but not necessarily on the yet lower levels. This latter case is neglected because there would be too many possible events competing to be the II. In the worst case, the combined window from both potential Is would extend over 92 samples (184 ms), *i.e.* 2×46 samples. The former case, (1), represents a scenario which has occasionally been observed. The systolic sound coalescing with the II has formed an apparently homogeneous event. This has brought the effective start of the II so far forward in time that it falls earlier than the lower limit of the window erected by both potential Is in pair_meta_ka (section 7.1.6). In considering the coalesced event, it is assumed that this is the only systolic sound which occurs in the systolic interval. This assumption is based on empirical observation. Each of the two potential Is is, in turn, hypothesized to be a I. A window is then established from this hypothesized I. This window begins at a time equivalent to 23 samples less than a systolic time later, and extends for 46 samples (92 ms), *i.e.*

twice the window width used in `pair_meta_ka`. The II must fall within this interval. A further condition to be satisfied is that the time instant equivalent to a systolic time away from the hypothesized I must occur within the body of the supposed II. That is, the earlier part of this event represents the coalescence with the systolic sound. Once this test has been completed, only one of the hypothesized Is must satisfy the conditions which will then allow an instantiation. If both Is simultaneously satisfy the conditions, the test is inconclusive. In the case of the former outcome, a I-II is instantiated after verification by the following I-II solution island. In the case of the latter outcome, the region is henceforth disregarded by `exhume_meta_ka`.

The meta-KA `exhume_meta_ka` progresses through the record solely in the direction of increasing time. Although `exhume_meta_ka` could also be made to operate non-causally, this was avoided because there is a large solution space to search (an interval equivalent to a systolic time) for a missing II. On the lower levels of data abstraction, this window would encompass many noise events, each one of which is a potential II. Hypothesized I-II couples would be too numerous, and many would pair with the same potential I event. In these circumstances there is increased risk of an incorrect assessment. Expanding from a I to search for a II, *i.e.* causally, has a search space of 46 samples which is much shorter than the minimum systolic time (62 samples). Furthermore, because of the large prominence of Is there is less likelihood of a missing I, which allows the generation of synthetic I-II pairs — subject to validation.

7.3 PRESENTATION OF RESULTS

Figures 7.2, 7.3 and 7.8 to 7.14 are a representative selection of PhCG types and the corresponding processing and analysis of them is indicative of what system 2 can achieve. These example PhCGs have been chosen from among the nineteen subjects

examined. However, the particular anomalies that each PhCG presents are not confined to that particular subject from which it is drawn, rather these anomalies are present in varying degrees in all the subjects examined. Half the number of example PhCGs illustrated have been taken from the subject identified as 'AN' in the figure titles. This is mainly because that subject displayed the spectrum of anomalies in the PhCG with which the analysis was designed to contend. These anomalies may not have been present in the other subjects either to such a marked degree, or to a similar extent within the same data record.

The display of the processing and analysis of each PhCG is divided between two figures. Those figure numbers with a suffix 'a' show three traces which represent from top to bottom: the phonocardiogram which resulted from a phonogram which was high-pass filtered at 20 Hz; the results of the processing of excursions; and the 'reconstructed' sequence of significant events (section 7.1.5). The middle trace has been artificially enhanced to illustrate more clearly the extent of an event as displayed by its prominence values. Here, the prominence value of an excursion is horizontally extrapolated until the next excursion is encountered. This extrapolation graphically delimits more clearly the ends of events where the prominence of excursions falls to a low value relative to those in the event. The figure numbers with a suffix 'b' present the analysis performed by system 2 on the PhCG which shares the same figure number but with the suffix 'a'.

The presentation of the results of an analysis is in a different format to that of system 1 (chapter 6, section 6.3). This is a reflection of the dissimilarities in the way events are parameterized in both systems, and the iterative way in which system 2 achieves an analysis (figure 7.6). Not all of the states of the data base may be displayed within a single page figure. Therefore, the number of states that are

displayed has been reduced to two, namely the state just before the commencement of analysis, and the final state of the data base once the analysis sub-system has terminated. Each signal event in the data base is represented as an ordered pair which comprises event starting time, and overall prominence value. The equivalent spatial position in the corresponding PhCG of these temporal events is shown alongside in millimetres, and is measured from the beginning of the PhCG. If an event pair has been identified as a first and second heart sound, a I-II token is placed adjacent to that pair, otherwise, uncertainty about the identity of an event pair is symbolically indicated by a '?' token.

Figure 7.2 illustrates a PhCG, the analysability of which would have been made more difficult by the cross-correlation processing used in system 1 (chapter 6). The persistent late diastolic sound is latently present until the latter half of the PhCG. The presence of this sound is more explicit in the sequence of excursion durations (figure 7.3) where the expected high-to-low frequency sweep of the first heart sounds is prefixed with that of the adventitious sound. This 'instantaneous' frequency information has been used by the segmentation KA to detect and remove these coalescences. This PhCG exemplifies the ability of the excursion-based processing method to preserve the end points of sounds, and to keep separate even those events which are very close together. System 2 was able to achieve a complete and correct analysis of this PhCG, although the results are not presented here.

Figure 7.8a illustrates a PhCG which has well defined first and second heart sounds, a low level of background noise, and only slight contamination by adventitious sounds which mostly occur in the late diastolic phase. Such is the clarity of the principal heart sounds, and the absence of interferences, that the analysis was almost complete at the stage of solution island identification (not shown).

In comparison with figure 7.8a, the PhCG of figure 7.9a has an increased background noise and, systolic and diastolic sounds are more in evidence. However, these adventitious sounds are sufficiently far removed from the principal heart sounds that the segmentation KA has not had to alter the configuration of any event. Excursion processing, as illustrated by the middle trace, has accurately delimited each event in the PhCG. The background noise has been successfully suppressed because of its relatively lower amplitude and shorter excursion duration which combine to produce relatively much lower prominence values. The analysis of this PhCG presented few problems and was, like the PhCG of figure 7.8a, almost complete once the I-II solution islands had been identified (not shown). It is interesting to compare this PhCG (filtered at 20 Hz) with the PhCG (filtered at 40 Hz) (figure 6.15a) which originated from the same phonogram. The better quality of the former PhCG has enabled a complete analysis to be accomplished by system 2, whereas system 1 was forced to abandon the PhCG at the analysis stage.

In figure 7.10a, the first and second heart sounds are not immediately obvious to a visual inspection because of the presence of many large-amplitude adventitious sounds. This difficulty is compounded by the proximity of these sounds to the principal heart sounds, particularly during the systolic phase. This PhCG exemplifies the power of excursion-based processing to accurately delimit each signal event despite the numerous coalescences. Had cross-correlation processing been used on such a PhCG, the segmentation of events would have been very problematic, if not unattainable. The quality of the knowledge-based processing is reflected in the subsequent complete analysis of this PhCG. The exploitation of the constancy of the systolic time across the PhCG has reduced the number of potential I-II pairs to nine, in a PhCG which originally presented thirty events to the analysis sub-system.

The PhCG illustrated in figure 7.11a is the immediate precursor to that of figure 7.10a. In the present case, the level of systolic and diastolic sounds, particularly in the middle section of the PhCG, has obscured the principal heart sounds. The uncertainty in this PhCG is shown in the number of potential I-II pairs presented to the analysis. Despite this, however, the analysis has succeeded in the relatively less noise-corrupted regions which lie at either end of the PhCG. The burst of indistinguishable signal activity in the middle section of the PhCG has been disregarded by the analysis. In this region, despite the coalescences of many events, the processing has been able to accurately delimit each event. The analysis sub-system has neglected this region because of the numerous event pairings which it cannot resolve.

The difficulties posed by the PhCG of figure 7.12a are: the unusually high frequency content of the second heart sounds and their uncharacteristic shape. These effects have combined to cause the fifth, second heart sound to be absent from the events in the 'reconstructed' sequence. Although this latter sound was present after the excursion processing stage, its small prominence value did not fall within the top thirty ranked events. The absence of a I-II pair in this region was detected by `exhume_meta_ka` which then located the excluded I on the third level of data abstraction and the small remnant of the II on the second level of data abstraction.

Figure 7.13a presents another example where second heart sounds are missing. In this case, however, the missing IIs are not discernibly present in the corresponding PhCG. In these circumstances, `exhume_meta_ka` has found the excluded Is, and created two synthetic IIs. In the expansion of analysis from the solution island which immediately precedes the region where the IIs are missing, a window is established from the II of the solution island within which the following I must lie. When the expansion of analysis halted because of the absence of expected events, the strategy

KA transferred control from `expand_pairs_meta_ka` to `exhume_meta_ka`. This latter meta-KA examined the third level of data abstraction with the objective of finding the missing I. This I had been previously excluded from the sequence of events which were transferred to the analysis because of the missing II in its pair. Once the excluded I had been located, a search for the II was not instigated on account of the absolute confidence in the identity of the single event which fell within the allowed window for a I, and the obviously small prominence (zero) of the missing II. A synthetic II was created to allow I-II instantiation in this region, and the analysis to proceed. The prominence value arbitrarily assigned to the missing II, is one-quarter of the prominence of the corresponding once-excluded I.

To a visual inspection, the PhCG of figure 7.14a does not appear more difficult to analyse than any of the other previous PhCGs illustrated. However, system 2 has not been able to analyse it. The difficulty arises from the proximity of diastolic sounds to Is. This has caused the IIs of previous beats to pair with these diastolic sounds. Furthermore, both diastolic sounds and Is are observed to pair with the same II because that II falls within the window of length 23 samples (section 7.1.6) extended from both of these latter events. The number of I-to-II-to-diastolic sound continuous pairings is evident in the list of paired events presented by the analysis (figure 7.14b). Such a situation may have been resolvable had the prominence values of the Is been significantly different to those of the other events. This was not so, therefore, the analysis had to be abandoned without any I-II instantiations.

7.4 DISCUSSION OF RESULTS

The unconventional but superior PhCG processing method advocated in this chapter has been visibly demonstrated in the above figures. Although the cross-correlation processing used in system 1 (chapter 6) would have benefited from the

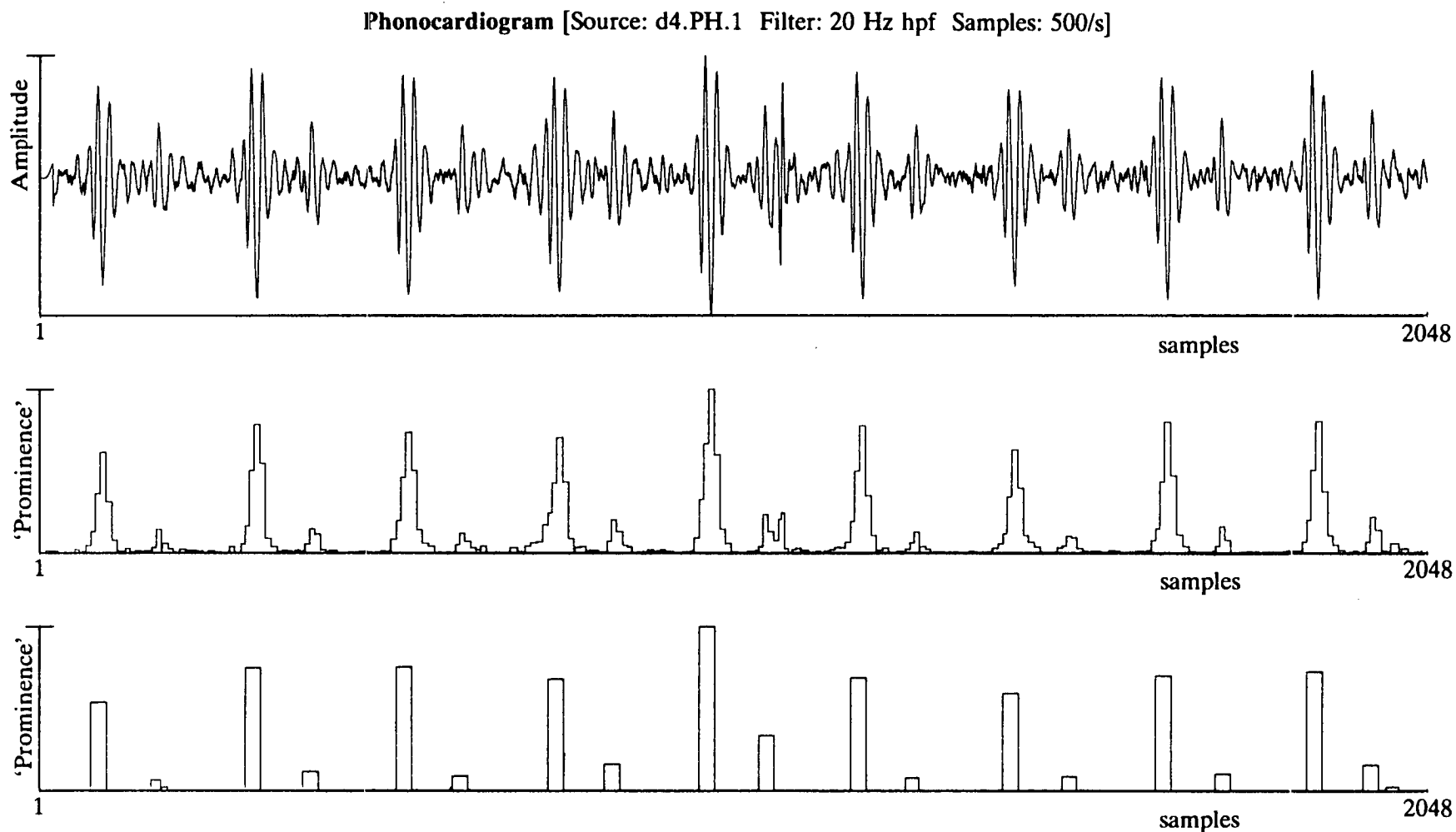


Figure 7.8a A phonocardiogram and its levels of abstraction. *top trace*, phonocardiogram; *middle trace*, sequence of processed excursions (for clarity of illustration, the prominence value of an excursion is horizontally extrapolated until the next excursion is encountered); *bottom trace*, sequence of events.

PhCG: d4.PH.1

Initial State of the Data Base

All pairs (start, prominence)		Start (mm)
(77, 984948)	& (166, 118156)	8.2 & 17.7
(304, 1370979)	& (389, 210988)	32.4 & 41.5
(527, 1378887)	& (610, 161825)	56.2 & 65.1
(751, 1247315)	& (834, 298979)	80.1 & 89.0
(974, 1833601)	& (1062, 615263)	103.9 & 113.3
(1198, 1263008)	& (1279, 142456)	127.8 & 136.5
(1423, 1089631)	& (1511, 160102)	151.8 & 161.2
(1648, 1288059)	& (1736, 186592)	175.8 & 185.2
(1870, 1334461)	& (1953, 289259)	199.5 & 208.4

Final State of the Data Base

All pairs (start, prominence)		ID	Start (mm)
(77, 984948)	& (166, 118156)	I-II	8.2 & 17.7
(304, 1370979)	& (389, 210988)	I-II	32.4 & 41.5
(527, 1378887)	& (610, 161825)	I-II	56.2 & 65.1
(751, 1247315)	& (834, 298979)	I-II	80.1 & 89.0
(974, 1833601)	& (1062, 615263)	I-II	103.9 & 113.3
(1198, 1263008)	& (1279, 142456)	I-II	127.8 & 136.5
(1423, 1089631)	& (1511, 160102)	I-II	151.8 & 161.2
(1648, 1288059)	& (1736, 186592)	I-II	175.8 & 185.2
(1870, 1334461)	& (1953, 289259)	?	199.5 & 208.4

Figure 7.8b Analysis of phonocardiogram presented in figure 7.8a.

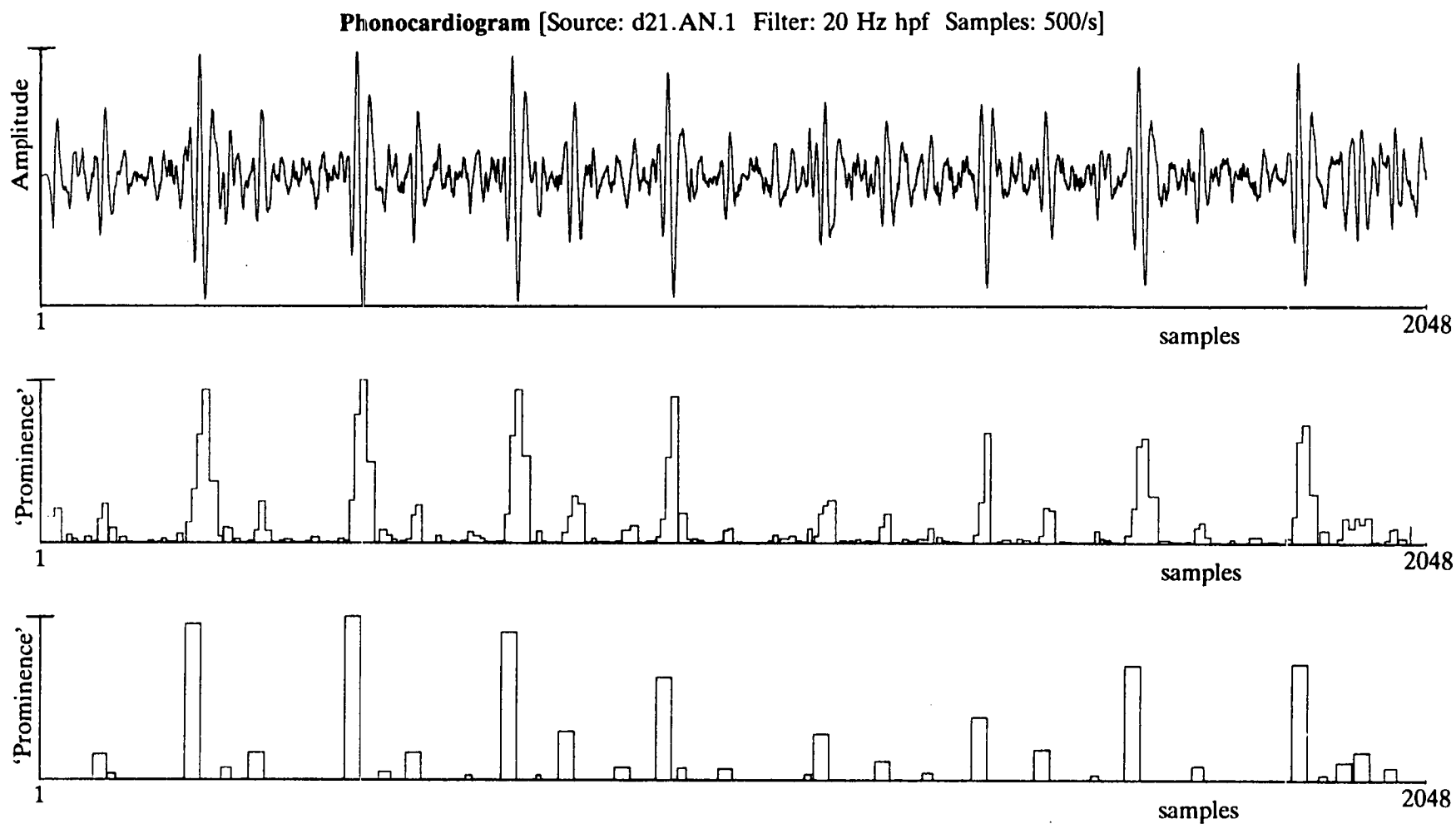


Figure 7.9a A phonocardiogram and its levels of abstraction. *top trace*, phonocardiogram; *middle trace*, sequence of processed excursions (for clarity of illustration, the prominence value of an excursion is horizontally extrapolated until the next excursion is encountered); *bottom trace*, sequence of events.

PhCG: d21.AN.1

Initial State of the Data Base

All pairs (start, prominence)			Start (mm)	
(215, 844771)	&	(309, 146190)	22.9	& 33.0
(451, 884586)	&	(541, 147446)	48.1	& 57.7
(541, 147446)	&	(630, 23788)	57.7	& 67.2
(630, 23788)	&	(734, 25635)	67.2	& 78.3
(682, 801291)	&	(767, 264989)	72.8	& 81.8
(767, 264989)	&	(850, 68051)	81.8	& 90.7
(850, 68051)	&	(943, 64258)	90.7	& 100.6
(911, 557424)	&	(1003, 61273)	97.2	& 107.0
(1131, 31312)	&	(1235, 101614)	120.7	& 131.8
(1144, 250572)	&	(1235, 101614)	122.1	& 131.8
(1306, 39595)	&	(1378, 341272)	139.3	& 147.0
(1378, 341272)	&	(1471, 166216)	147.0	& 156.9
(1471, 166216)	&	(1555, 27537)	156.9	& 165.9
(1604, 622531)	&	(1704, 76051)	171.1	& 181.8
(1850, 630705)	&	(1942, 151153)	197.4	& 207.2
(1890, 27077)	&	(1987, 67994)	201.6	& 212.0

Final State of the Data Base

All pairs (start, prominence)			ID	Start (mm)	
(215, 844771)	&	(309, 146190)	I-II	22.9	& 33.0
(451, 884586)	&	(541, 147446)	I-II	48.1	& 57.7
(682, 801291)	&	(767, 264989)	I-II	72.8	& 81.8
(911, 557424)	&	(1003, 61273)	I-II	97.2	& 107.0
(1144, 250572)	&	(1235, 101614)	I-II	122.1	& 131.8
(1378, 341272)	&	(1471, 166216)	I-II	147.0	& 156.9
(1604, 622531)	&	(1704, 76051)	I-II	171.1	& 181.8
(1850, 630705)	&	(1942, 151153)	I-II	197.4	& 207.2

Figure 7.9b Analysis of phonocardiogram presented in figure 7.9a.

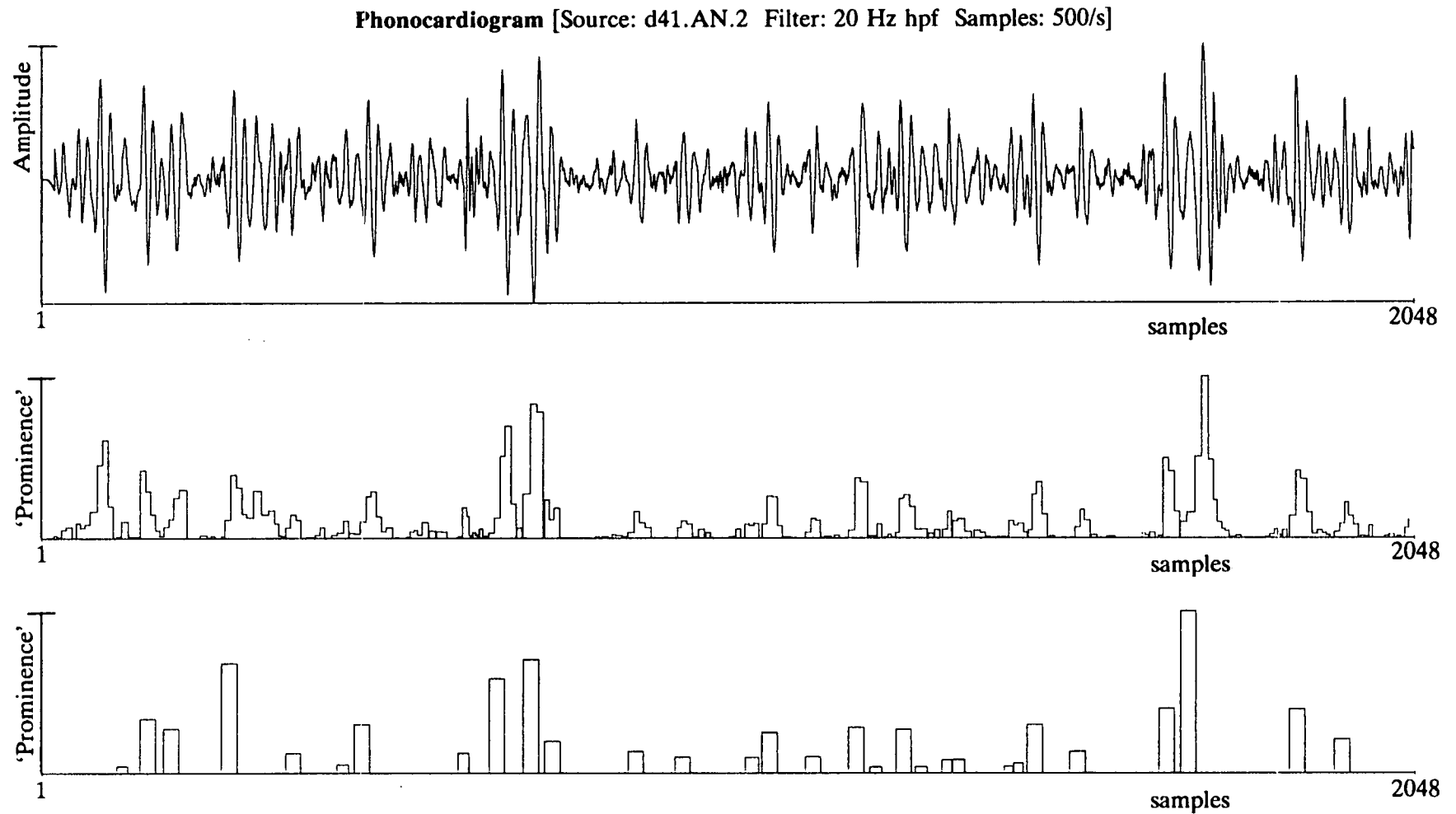


Figure 7.10a A phonocardiogram and its levels of abstraction. *top trace*, phonocardiogram; *middle trace*, sequence of processed excursions (for clarity of illustration, the prominence value of an excursion is horizontally extrapolated until the next excursion is encountered); *bottom trace*, sequence of events.

PhCG: d41.AN.2

Initial State of the Data Base

All pairs (start, prominence)		Start (mm)
(114, 52570)	& (183, 334307)	12.2 & 19.5
(183, 334307)	& (269, 818864)	19.5 & 28.7
(365, 150080)	& (441, 66305)	38.9 & 47.0
(668, 706880)	& (751, 242327)	71.3 & 80.1
(875, 165662)	& (945, 122095)	93.4 & 100.8
(1050, 119524)	& (1140, 123280)	112.0 & 121.6
(1075, 306657)	& (1140, 123280)	114.7 & 121.6
(1140, 123280)	& (1204, 344510)	121.6 & 128.5
(1204, 344510)	& (1275, 329597)	128.5 & 136.0
(1236, 46226)	& (1304, 50135)	131.9 & 139.1
(1275, 329597)	& (1344, 99445)	136.0 & 143.4
(1344, 99445)	& (1437, 52653)	143.4 & 153.3
(1360, 102148)	& (1437, 52653)	145.1 & 153.3
(1451, 77888)	& (1535, 164273)	154.8 & 163.8
(1471, 365282)	& (1535, 164273)	156.9 & 163.8
(1861, 480652)	& (1928, 255690)	198.5 & 205.7

(exhume_meta_ka)

Synthetic II inserted: 1739 samples (185.5 mm).

Final State of the Data Base

All pairs (start, prominence)		ID	Start (mm)
(365, 150080)	& (441, 66305)	?	38.9 & 47.0
(668, 706880)	& (751, 242327)	I-II	71.3 & 80.1
(875, 165662)	& (945, 122095)	I-II	93.4 & 100.8
(1075, 306657)	& (1140, 123280)	I-II	114.7 & 121.6
(1275, 329597)	& (1344, 99445)	I-II	136.0 & 143.4
(1471, 365282)	& (1535, 164273)	I-II	156.9 & 163.8
(1667, 485359)	& (1739, 121339)	I-II	177.9 & 185.5
(1861, 480652)	& (1928, 255690)	I-II	198.5 & 205.7

Figure 7.10b Analysis of phonocardiogram presented in figure 7.10a.

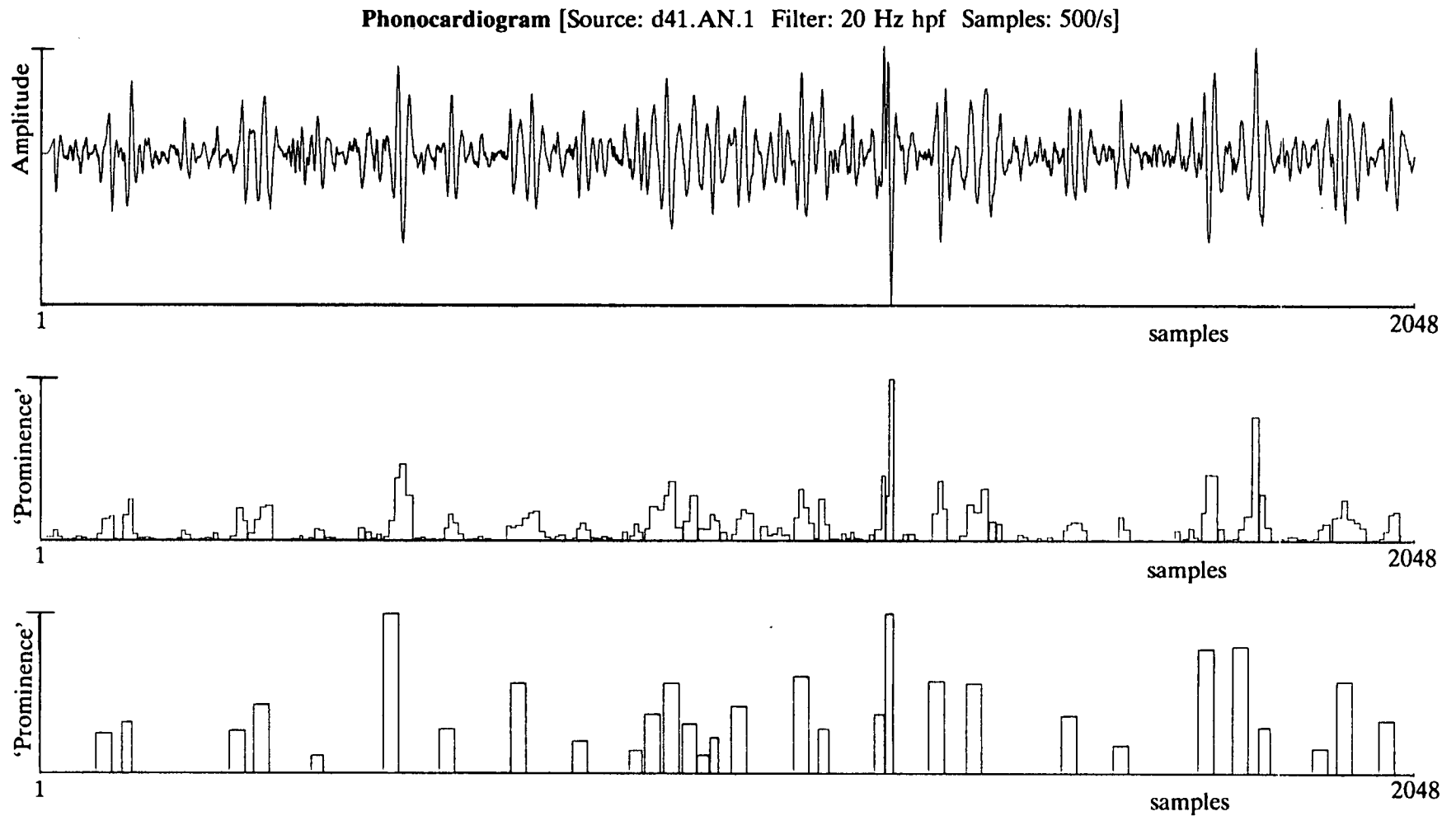


Figure 7.11a A phonocardiogram and its levels of abstraction. *top trace*, phonocardiogram; *middle trace*, sequence of processed excursions (for clarity of illustration, the prominence value of an excursion is horizontally extrapolated until the next excursion is encountered); *bottom trace*, sequence of events.

PhCG: d41.AN.1

Initial State of the Data Base

All pairs (start, prominence)		Start (mm)
(319, 275455)	& (405, 70542)	34.0 & 43.2
(512, 641085)	& (596, 177917)	54.6 & 63.6
(702, 362323)	& (794, 130971)	74.9 & 84.7
(794, 130971)	& (879, 92156)	84.7 & 93.8
(879, 92156)	& (958, 199393)	93.8 & 102.2
(902, 238832)	& (980, 72724)	96.2 & 104.6
(930, 363677)	& (999, 144493)	99.2 & 106.6
(958, 199393)	& (1031, 270470)	102.2 & 110.0
(1031, 270470)	& (1124, 390813)	110.0 & 119.9
(1161, 180179)	& (1245, 238505)	123.9 & 132.8
(1245, 238505)	& (1326, 372068)	132.8 & 141.5
(1261, 643051)	& (1326, 372068)	134.5 & 141.5
(1524, 233409)	& (1601, 113034)	162.6 & 170.8
(1727, 502115)	& (1817, 187207)	184.3 & 193.9
(1817, 187207)	& (1897, 100623)	193.9 & 202.4
(1897, 100623)	& (1996, 214110)	202.4 & 213.0
(1933, 372189)	& (1996, 214110)	206.2 & 213.0

Final State of the Data Base

All pairs (start, prominence)		ID	Start (mm)
(319, 275455)	& (405, 70542)	I-II	34.0 & 43.2
(512, 641085)	& (596, 177917)	I-II	54.6 & 63.6
(702, 362323)	& (794, 130971)	I-II	74.9 & 84.7
(902, 238832)	& (980, 72724)	?	96.2 & 104.6
(930, 363677)	& (999, 144493)	?	99.2 & 106.6
(958, 199393)	& (1031, 270470)	?	102.2 & 110.0
(1031, 270470)	& (1124, 390813)	?	110.0 & 119.9
(1161, 180179)	& (1245, 238505)	?	123.9 & 132.8
(1245, 238505)	& (1326, 372068)	?	132.8 & 141.5
(1261, 643051)	& (1326, 372068)	?	134.5 & 141.5
(1524, 233409)	& (1601, 113034)	I-II	162.6 & 170.8
(1727, 502115)	& (1817, 187207)	I-II	184.3 & 193.9
(1933, 372189)	& (1996, 214110)	?	206.2 & 213.0

Figure 7.11b Analysis of phonocardiogram presented in figure 7.11a.

Phonocardiogram [Source: d53.AN.1 Filter: 20 Hz hpf Samples: 500/s]

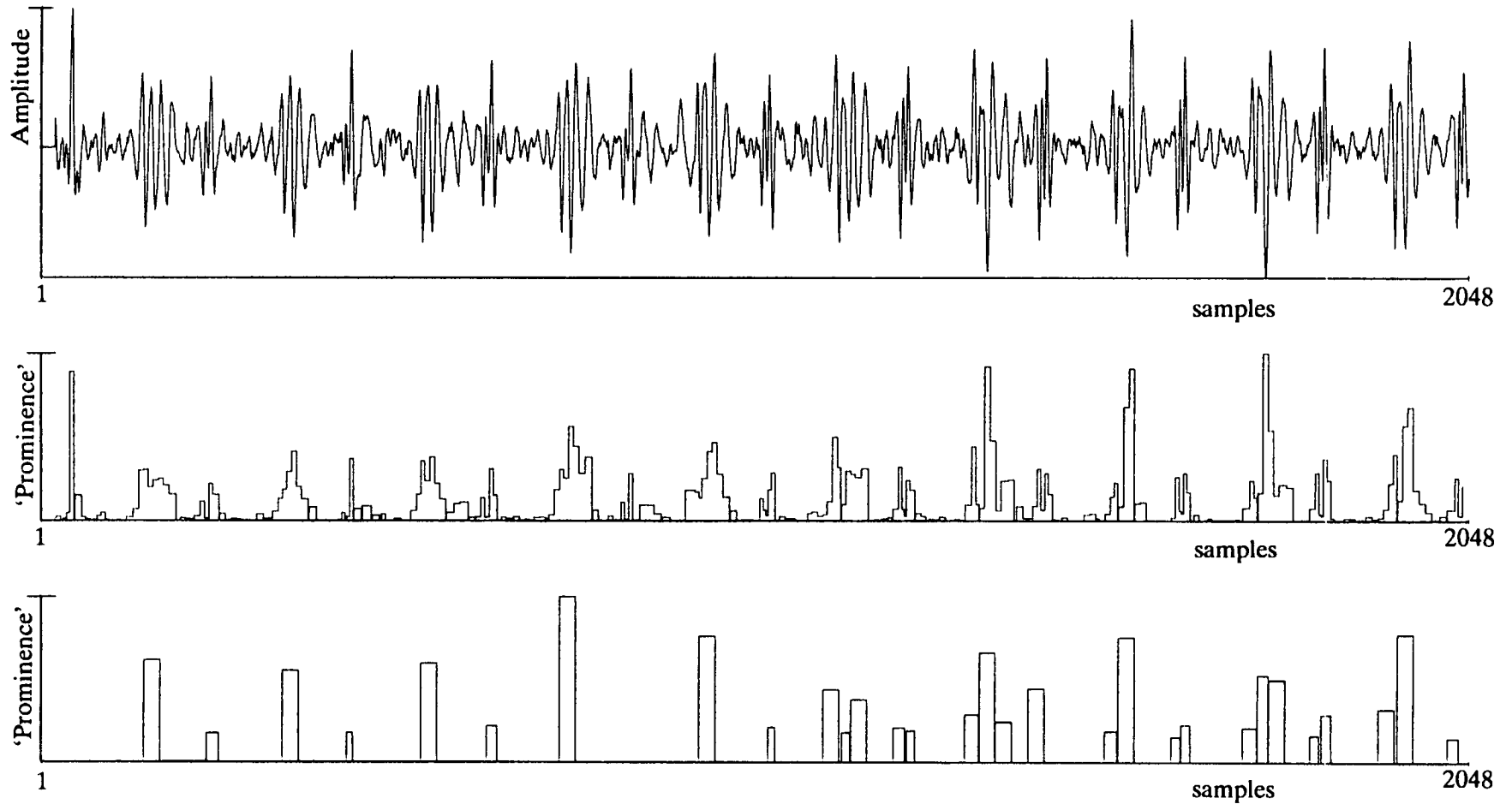


Figure 7.12a A phonocardiogram and its levels of abstraction. *top trace*, phonocardiogram; *middle trace*, sequence of processed excursions (for clarity of illustration, the prominence value of an excursion is horizontally extrapolated until the next excursion is encountered); *bottom trace*, sequence of events.

PhCG: d53.AN.1

Initial State of the Data Base

All pairs (start, prominence)		Start (mm)
(148, 482895)	& (238, 137327)	15.8 & 25.4
(346, 433479)	& (438, 139063)	36.9 & 46.7
(545, 467974)	& (639, 171392)	58.1 & 68.2
(944, 596890)	& (1043, 166046)	100.7 & 111.3
(1043, 166046)	& (1122, 343859)	111.3 & 119.7
(1122, 343859)	& (1223, 164308)	119.7 & 130.5
(1149, 140621)	& (1242, 149639)	122.6 & 132.5
(1162, 298166)	& (1242, 149639)	124.0 & 132.5
(1223, 164308)	& (1326, 227756)	130.5 & 141.5
(1242, 149639)	& (1326, 227756)	132.5 & 141.5
(1326, 227756)	& (1417, 351643)	141.5 & 151.2
(1347, 520595)	& (1417, 351643)	143.7 & 151.2
(1527, 147473)	& (1622, 121023)	162.9 & 173.0
(1546, 591975)	& (1636, 177178)	164.9 & 174.5
(1622, 121023)	& (1724, 163825)	173.0 & 183.9
(1636, 177178)	& (1724, 163825)	174.5 & 183.9
(1724, 163825)	& (1820, 127966)	183.9 & 194.2
(1745, 411760)	& (1836, 226147)	186.2 & 195.9
(1761, 390990)	& (1836, 226147)	187.9 & 195.9
(1820, 127966)	& (1918, 252597)	194.2 & 204.6
(1836, 226147)	& (1918, 252597)	195.9 & 204.6
(1918, 252597)	& (2017, 114305)	204.6 & 215.2

(exhume_meta_ka)

Missing II found on second level of data abstraction.

I-II pair installed: 744 & 824 samples (79.4 mm & 87.9 mm).

Final State of the Data Base

All pairs (start, prominence)		ID	Start (mm)
(148, 482895)	& (238, 137327)	I-II	15.8 & 25.4
(346, 433479)	& (438, 139063)	I-II	36.9 & 46.7
(545, 467974)	& (639, 171392)	I-II	58.1 & 68.2
(744, 785010)	& (824, 42685)	I-II	79.4 & 87.9
(944, 596890)	& (1043, 166046)	I-II	100.7 & 111.3
(1162, 298166)	& (1242, 149639)	I-II	124.0 & 132.5
(1347, 520595)	& (1417, 351643)	I-II	143.7 & 151.2
(1546, 591975)	& (1636, 177178)	I-II	164.9 & 174.5
(1745, 411760)	& (1836, 226147)	I-II	186.2 & 195.9
(1918, 252597)	& (2017, 114305)	?	204.6 & 215.2

Figure 7.12b Analysis of phonocardiogram presented in figure 7.12a.

Phonocardiogram [Source: d3.PL.1 Filter: 20 Hz hpf Samples: 500/s]

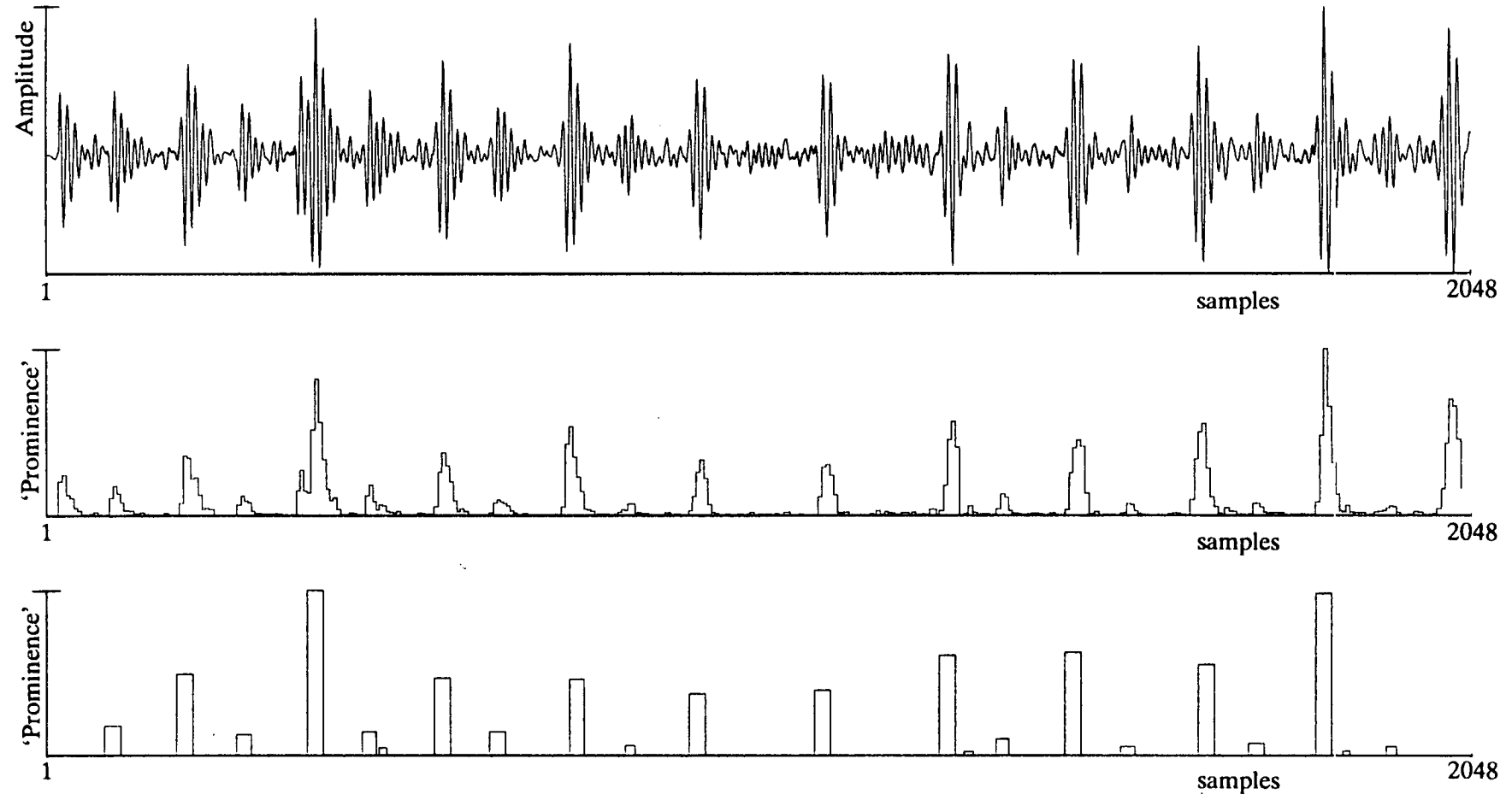


Figure 7.13a A phonocardiogram and its levels of abstraction. *top trace*, phonocardiogram; *middle trace*, sequence of processed excursions (for clarity of illustration, the prominence value of an excursion is horizontally extrapolated until the next excursion is encountered); *bottom trace*, sequence of events.

PhCG: d3.PL.1

Initial State of the Data Base

All pairs (start, prominence)			Start (mm)	
(188, 2)	&	(274, 0)	20.1	& 29.2
(274, 0)	&	(375, 4)	29.2	& 40.0
(375, 4)	&	(454, 0)	40.0	& 48.4
(478, 0)	&	(558, 2)	51.0	& 59.5
(558, 2)	&	(637, 0)	59.5	& 68.0
(752, 2)	&	(832, 0)	80.2	& 88.8
(832, 0)	&	(924, 1)	88.8	& 98.6
(1284, 2)	&	(1366, 0)	137.0	& 145.7
(1366, 0)	&	(1465, 3)	145.7	& 156.3
(1465, 3)	&	(1545, 0)	156.3	& 164.8
(1656, 2)	&	(1728, 0)	176.7	& 184.4
(1728, 0)	&	(1824, 4)	184.4	& 194.6
(1824, 4)	&	(1925, 0)	194.6	& 205.4

(exhume_meta_ka)

Synthetic II inserted: 1004 samples (107.1 mm).

Synthetic II inserted: 1184 samples (126.3 mm).

Final State of the Data Base

All pairs (start, prominence)			ID	Start (mm)	
(188, 2)	&	(274, 0)	I-II	20.1	& 29.2
(375, 4)	&	(454, 0)	I-II	40.0	& 48.4
(558, 2)	&	(637, 0)	I-II	59.5	& 68.0
(752, 2)	&	(832, 0)	I-II	80.2	& 88.8
(924, 1)	&	(1004, 0)	I-II	98.6	& 107.1
(1104, 1)	&	(1184, 0)	I-II	117.8	& 126.3
(1284, 2)	&	(1366, 0)	I-II	137.0	& 145.7
(1465, 3)	&	(1545, 0)	I-II	156.3	& 164.8
(1656, 2)	&	(1728, 0)	I-II	176.7	& 184.4
(1824, 4)	&	(1925, 0)	I-II	194.6	& 205.4

Figure 7.13b Analysis of phonocardiogram presented in figure 7.13a.

Phonocardiogram [Source: d1.PB.2 Filter: 20 Hz hpf Samples: 500/s]

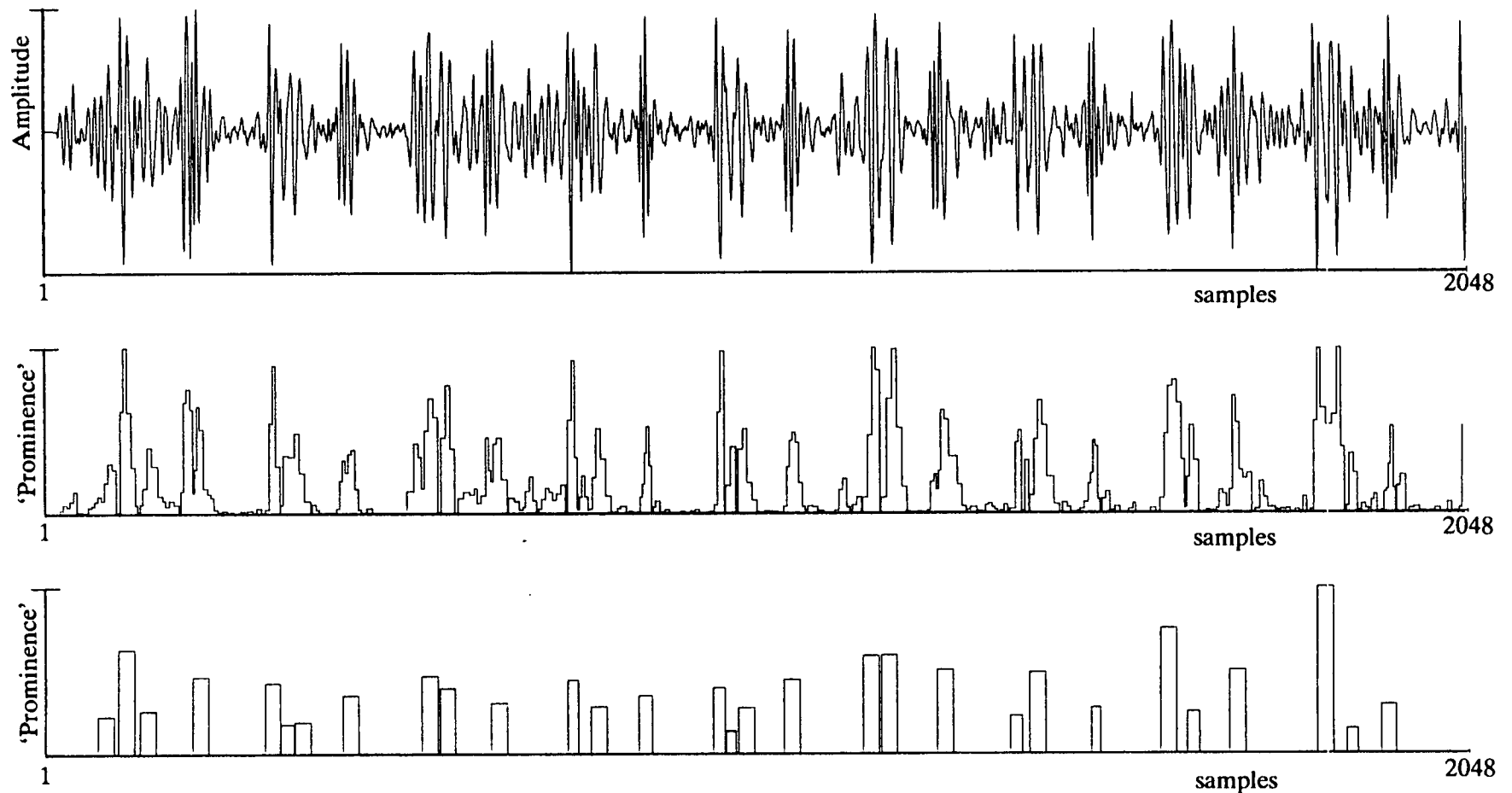


Figure 7.14a A phonocardiogram and its levels of abstraction. *top trace*, phonocardiogram; *middle trace*, sequence of processed excursions (for clarity of illustration, the prominence value of an excursion is horizontally extrapolated until the next excursion is encountered); *bottom trace*, sequence of events.

PhCG: d1.PB.2

Initial State of the Data Base

All pairs (start, prominence)			Start (mm)	
(106, 6)	&	(213, 5)	11.3	& 22.7
(213, 5)	&	(317, 4)	22.7	& 33.8
(317, 4)	&	(428, 3)	33.8	& 45.7
(339, 1)	&	(428, 3)	36.2	& 45.7
(428, 3)	&	(542, 5)	45.7	& 57.8
(542, 5)	&	(642, 3)	57.8	& 68.5
(642, 3)	&	(752, 4)	68.5	& 80.2
(752, 4)	&	(854, 3)	80.2	& 91.1
(854, 3)	&	(961, 4)	91.1	& 102.5
(961, 4)	&	(1063, 4)	102.5	& 113.4
(979, 1)	&	(1063, 4)	104.4	& 113.4
(1063, 4)	&	(1177, 6)	113.4	& 125.6
(1177, 6)	&	(1284, 5)	125.6	& 137.0
(1203, 6)	&	(1284, 5)	128.3	& 137.0
(1284, 5)	&	(1389, 2)	137.0	& 148.2
(1389, 2)	&	(1506, 3)	148.2	& 160.7
(1417, 5)	&	(1506, 3)	151.2	& 160.7
(1506, 3)	&	(1605, 8)	160.7	& 171.2
(1605, 8)	&	(1704, 5)	171.2	& 181.8
(1830, 11)	&	(1922, 3)	195.2	& 205.1

Final State of the Data Base

All pairs (start, prominence)			ID	Start (mm)	
(106, 6)	&	(213, 5)	?	11.3	& 22.7
(317, 4)	&	(428, 3)	?	33.8	& 45.7
(339, 1)	&	(428, 3)	?	36.2	& 45.7
(428, 3)	&	(542, 5)	?	45.7	& 57.8
(542, 5)	&	(642, 3)	?	57.8	& 68.5
(642, 3)	&	(752, 4)	?	68.5	& 80.2
(752, 4)	&	(854, 3)	?	80.2	& 91.1
(854, 3)	&	(961, 4)	?	91.1	& 102.5
(961, 4)	&	(1063, 4)	?	102.5	& 113.4
(1063, 4)	&	(1177, 6)	?	113.4	& 125.6
(1177, 6)	&	(1284, 5)	?	125.6	& 137.0
(1203, 6)	&	(1284, 5)	?	128.3	& 137.0
(1389, 2)	&	(1506, 3)	?	148.2	& 160.7
(1417, 5)	&	(1506, 3)	?	151.2	& 160.7
(1605, 8)	&	(1704, 5)	?	171.2	& 181.8
(1830, 11)	&	(1922, 3)	?	195.2	& 205.1

Figure 7.14b Analysis of phonocardiogram presented in figure 7.14a.

better signal-to-noise ratio of the 20 Hz high-pass filtered PhCGs, this would not have alleviated the smearing effect which correlation imposes (chapter 6, section 6.1.3). It was the objective of the excursion-based processing method that the end points of events should be accurately preserved. This objective has been met by an unconventional method of signal processing. Despite the approximation of adventitious sounds to the principal heart sounds, the excursion processing has been able to delimit each. This accurate event delimitation and the method of segmentation, which relies on prominence values and excursion durations across an event, have contributed to the successful detection of coalescences of adventitious sounds with principal heart sounds.

The method by which events are analysed has also been modified from that used in system 1. In system 2, there is a reliance on events being separated by a time equivalent to a systolic time to remove unlikely first and second heart sound pairs. This has also reduced the number of events which are finally transferred to the analysis sub-system. The 'allowed-window-of-occurrence' method by which the analysis is expanded from solution islands into the more ambiguous regions of the PhCG, has also benefited from knowledge of the constancy of the systolic time. In system 2, the upper and lower limits of the window are directly dependent on the systolic time interval of the PhCG under investigation. This has significantly narrowed the extent of the window from that used in system 1, where the theoretical limits of principal heart sound separation were employed to calculate the upper and lower limits of the window. The assumption made in system 2, with strong empirical evidence [46,75,163], regarding the constancy of the systolic time, provides a more realistic window than that used in system 1. The superior processing, the narrower width of the 'allowed window of occurrence', and the event-pairing approach, have resulted in system 2 being able to completely analyse PhCGs which would have been either totally or partially unanalysable by system 1. This latter statement would still be

true, even if system 1 were to be tailored for the analysis of 20 Hz high-pass filtered PhCGs.

Although the FHR has been approximately constant over all the 252 PhCGs examined, system 2 has the capability to accommodate all FHRs which lie in the range 80-220 bpm irrespective of the short-term variability (chapter 6, section 6.4). Again, this is made possible by knowledge of the constancy of the systolic time interval.

In system 2 the limit of PhCG analysability is reached whenever adventitious sounds *persistently* appear at what may be termed the critical points of a PhCG, *i.e.* those points immediately adjacent on either side of the principal heart sounds. In these cases, the adventitious sounds have the appearance of *bona fide* I-II pairs. This problem is unresolvable whenever the adventitious sounds have similar prominence values to the principal heart sounds, particularly the first heart sounds. The ability of system 2 to analyse a PhCG which has many adventitious sounds, depends therefore, on the relative temporal separation of these from the principal heart sounds. Figure 7.10a is an example of the presence of numerous adventitious sounds in a PhCG which was, nevertheless, amenable to a complete analysis.

The number and temporal location of adventitious sounds, notwithstanding those cases of which figure 7.10a is an example, effectively impose the limit of analysability in both system 1 and system 2. In PhCGs which are heavily contaminated by adventitious sounds, the best analysis that can be achieved by a PhCG understanding system, should result in the outright rejection of that PhCG. A deficiency of system 2 is that it does not have the battery of tests which were incorporated into system 1 to reject anomalous PhCGs before they reached the analysis stage. There were occasions, although not illustrated, when the PhCG was corrupted by foetal or transducer movements to the extent that foetal heart sounds were completely obliterated. When

system 2 was applied to these cases, there may have been at most one or two event pairs, which by a chance combination of circumstances satisfied all the conditions necessary for instantiation as a I-II couple. A defence against these types of PhCG was not incorporated into system 2, therefore, the application of such PhCGs is not a valid test of system 2. This deficiency may be easily rectified by employing similar tests to those of system 1.

7.5 SUMMARY AND CONCLUSION

This chapter describes a knowledge-based system which automatically identifies first and second foetal heart sounds for all FHRs which lie in the range 80-220 bpm, irrespective of the short-term variability. The ability to encompass such a range of FHRs arises from the exploitation of the constancy of the systolic time interval over wide variations in FHR.

A novel method is introduced for the processing of PhCGs which ensures that the end points of signal events are accurately preserved. In this scheme, individual excursions of the PhCG across the horizontal axis are separately processed and then returned to their original temporal location. This has prevented the entrance of processing-associated effects such as phase shifts and the smearing of events. After further processing, contiguous excursions are grouped together to form events which are segmented, where appropriate, to ensure the homogeneity of each event.

The analysis of these events commences with a search for solution islands which comprise I-II pairs. From these partial analyses, the system expands the overall analysis into those regions of the PhCG where the location of subsequent I-II pairs is ambiguous. Such ambiguity results from the presence of signal events which have originated from adventitious sounds.

As a consequence of the superior processing and analysis employed by system 2 in comparison with system 1 (chapter 6), system 2 is able to completely and correctly analyse a broader spectrum of PhCG types. The ability of system 2 to analyse PhCGs, the visual interpretation of which is not immediately apparent, is again an indication of the effectiveness of the knowledge-based signal analysis methodology in this application.

CHAPTER 8

SUMMARY AND CONCLUSIONS

The objective of the research described in this thesis was the automatic detection and identification of the principal heart sounds in the foetal phonocardiogram. Once these events have been located in time, certain beat-to-beat temporal parameters of cardiac function may be inferred. To achieve this objective, two methods of signal processing and analysis were applied. Firstly, the frequency-domain representations of both principal heart sounds were derived using Fourier and parametric models. The spectral signatures of both sounds were examined for features which could be uniquely associated with each. Secondly, a knowledge-based methodology was applied in an effort to distinguish between the principal heart sounds in the time domain using a synthesis of contextual information and background knowledge. The following section summarizes the methods employed in this work for the acquisition, conditioning, processing, and analysis of the phonocardiographic signal. The conclusions drawn from the research are also given. The final section suggests possible extensions to the research detailed in the thesis.

8.1 SUMMARY AND CONCLUSIONS

The acquisition and conditioning of the phonographic signal is described in chapter 3. The wide-bandwidth TAPHO phono-transducer which was used for all

signal acquisition represents a significant advancement over the existing narrow-bandwidth phono-transducers. The latter had severely attenuated the majority of the energy contained in the foetal heart sounds which the author has observed to lie below the previously monitored band of 50-120 Hz. It would appear, therefore, that previous investigators had endeavoured to reproduce with transducers and signal conditioning circuitry, the sounds that were heard on auscultation. The combined effect of narrow-bandwidth transducers and band-pass filters incorrectly situated in the spectrum, produced PhCGs with low signal-to-noise ratios. Undoubtedly this was a major contributory factor in the premature abandonment of phonocardiography in foetal heart rate monitoring.

The TAPHO transducer allows the recording of the whole spectrum occupied by the foetal heart sounds (20-100 Hz). The spectral content of the foetal heart sounds (FHS) has been a source of disagreement among previous investigators. However, in this thesis it has been demonstrated that the majority of the energy in the FHS lies in the 20-40 Hz band. In the conditioning circuitry of the apparatus used, strict attention was paid to the phase linearity of both analogue (Bessel-Thomson) and digital filters to preserve the relative phasing of the spectral components. This is an important point which has not been addressed in any of the previous published works in this field. In all, 252 phonograms, each of 4.1 seconds, were captured from 19 subjects. The sampling rate for the ADC was 500 samples/s — two-and-a-half times oversampling — which provides sufficient time-domain resolution. Phonocardiograms were produced from these phonograms using a 20 Hz high-pass transversal digital filter with 41 taps designed by the window method with a Hamming window. It was observed in this sample set of 252 PhCGs that the FHS can change markedly from one beat to the next in amplitude, duration, frequency, morphology, and in time of occurrence relative to the previous beat. These effects combined with the susceptibility of the PhCG to

contamination by adventitious sounds, make the task of the automatic detection and identification of the FHS a difficult one to achieve. Generally, however, the variability in the PhCG is considerably less than that of the signals obtained with Doppler ultrasound FHR monitors.

The frequency analysis of the first and second foetal heart sounds is presented in chapter 4. These FHS were manually extracted from 20 Hz high-pass filtered PhCG recordings. In all, 404 first and 337 second heart sounds from 12 foeti were analysed. Initially, the discrete Fourier transform (DFT) with alternatively the Hamming window and the Dolph-Chebyshev window was used to obtain the spectra of these sounds. As a result of the brevity of the FHS, the resulting spectra obtained by Fourier transformation were highly smoothed, irrespective of which window was applied. Despite this, however, the previous qualitative observation that the energy of the FHS lies in the 20-40 Hz band was quantitatively substantiated. To decompose the smoothed spectra obtained with the DFT, a high-resolution parametric spectral analysis technique was employed. The particular one chosen was the optimum tapered Burg algorithm which alleviates the problems associated with the unwrapped Burg algorithm. Although this technique succeeded in revealing the latent detail in the smoothed spectra produced by the DFT, it was observed that both first and second heart sounds produce very similar spectra. In the majority of cases, the FHS could not be distinguished from each other on the basis of their respective frequency content. With the failure of the frequency-domain representation of the FHS to provide uniquely identifying characteristics for each sound, attention was turned to the time-domain.

The knowledge-based system (KBS) approach to the processing and analysis of signals is introduced in chapter 5. In this methodology, the analysis of signal events is

achieved by the synthesis of contextual information from the immediate signal environment, and knowledge about the signal domain. This represents a departure from conventional techniques of signal processing and analysis which have always relied on numerical computation. In the particular KBS described, 'if-then' rules are applied by other controlling 'if-then' rules to 'reason' about the signal events and their interrelationships. This reasoning may function to either propose the presence of a signal event from the input data, or to hypothesize an expectation of an event from what has already been deduced from the signal. Such a methodology is very suitable for the case of PhCG analysis because a substantial amount of information is known about the foetal heart and the FHS. Synthesizing this information into an analysis of the phonocardiographic signal allows the inherent variability of the signal events and the presence of adventitious sounds to be accommodated.

In chapters 6 and 7 the KBS methodology is applied to the understanding of the PhCG. System 1, described in chapter 6, and system 2, described in chapter 7, are similar on a conceptual level and differ only in implementational aspects. For historical reasons, system 1 was designed to analyse PhCGs which were produced by high-pass filtering phonograms at 40 Hz. System 2, however, was designed for PhCGs which result from the same phonograms as above except that the high-pass filter threshold frequency is reduced to 20 Hz. In the latter case this yields a much better signal-to-noise ratio.

To accentuate the principal heart sounds prior to analysis, system 1 uses cross-correlation processing of the PhCG with a second heart sound template drawn from that PhCG. The signal events which are extracted from the processed sequence are segmented to ensure homogeneity, and then transferred to the analysis sub-system. This sequence of events is then tested for suitability for analysis, *i.e.* there are no

transients in the signal which have resulted from foetal or transducer movement. The analysis of a PhCG begins with a search for regions of the signal where there is no ambiguity and the identity of events is apparent. These 'solution islands' form a bridgehead for the expansion of analysis into the ambiguous regions of the PhCG. System 1 can determine when parts of the PhCG are too noise contaminated to warrant analysis, and whenever there are inconsistencies in its analysis. There are four limitations associated with the design of system 1. Firstly, system 1 is 'tuned' to the analysis of 40 Hz high-pass filtered PhCGs. This restricts the scope of analysable PhCG types because of the inherently low signal-to-noise ratio with this sub-optimum filtering threshold frequency. Secondly, the FHR is restricted to lie in the range 117-175 bpm. However, within this range all FHR beat-to-beat differences are accommodated irrespective of the short-term variability. Thirdly, the cross-correlation processing employed to accentuate the principal heart sounds, also accentuates the adventitious sounds, and is inclined to smear the end of one signal event into the beginning of the next. This further complicates the task of segmentation. Finally, the template used in the cross-correlation must be manually extracted from the PhCG beforehand. Despite these limitations, system 1 performed very well over the range of PhCG types presented.

System 2 (chapter 7) was specifically designed to overcome the limitations of system 1. To a certain extent the improvements over system 1 were made possible by the better signal-to-noise ratio of the 20 Hz high-pass filtered PhCGs which system 2 analysed. The range of FHR which could be accommodated was expanded to cover 80-220 bpm irrespective of either the short or long-term variability. This latter facility is a significant advancement over commercially available instruments which usually set a limit of between 15 and 20 bpm on the short-term variability. System 2 employs an unconventional method of signal processing which is based on the excursions of the

amplitude of the PhCG across the horizontal (0 volt) axis. This has allowed the end points of events to be accurately specified, and overcomes the disadvantages associated with correlation processing. The method of analysis used in system 2 is conceptually the same as that in system 1, except that the implementation is superior. The limit of PhCG analysability in system 2 is determined by the number and temporal location of the adventitious sounds. However, system 2 has been able to analyse a wide variety of PhCG types, ranging down as far as those which are not immediately obvious to a visual inspection. The ability of this system to detect and identify individual principal heart sounds on a beat-to-beat basis is another significant advancement achieved in this work. There are presently no commercial or other research systems which can provide beat-by-beat analysis.

The objective of all the foregoing processing and analysis schemes has been the automatic detection and identification of the first and second foetal heart sounds in a given PhCG. The proposition of this thesis is that the knowledge-based system methodology is a technique which can successfully meet these objectives. A knowledge-based system has been developed which automatically detects and identifies foetal heart sounds across the spectrum of PhCG types, and over a wide range of foetal heart rates (80-220 bpm), irrespective of the short-term heart rate variability.

8.2 EXTENSIONS TO THE RESEARCH

A prototype system for the processing and analysis of foetal PhCGs is now operational. For this system to be more comprehensively tested, and its routine clinical value assessed, there is a requirement for a real-time implementation of the algorithm. Although computational efficiency was not an issue which was specifically addressed during system development, the algorithm can process and analyse a 4.1 second PhCG in 30-40 seconds on a SUN 3/80 workstation. The duration of run-time is dependent

on the length of the reasoning chain which is, in turn, dependent on the uncertainty in the PhCG. A significant proportion of the computation time is absorbed by high-pass filtering the 2048 sample (4.1 seconds) phonogram with a 41 tap filter to produce a PhCG. All subsequent computation is centred on additions and comparisons. With the availability of high-speed DSP chips, it would be possible to achieve real-time operation of this algorithm.

Once the algorithm is operating in real-time, the analysis it produces needs to be compared to an accepted standard. There are presently no commercial instruments which identify first and second heart sounds. However, the derived FHR over a certain interval could form the basis of a comparison with that obtained with either a commercial ultrasound FHR monitor or a scalp electrode (ECG) during labour.

A more substantial extension to the research in this thesis, is the possibility of simultaneously synthesizing the analysis of a PhCG with the analysis obtained from other different types of transducers which also monitor the foetal heart. It is envisaged that the ultrasound foetal cardiograph (uFCG) may be a suitable choice. The ultrasound transducer relies on the Doppler principle for its operation. A beam of ultrasound energy is directed towards the foetal heart and the 'backscatter' is frequency shifted in proportion to the velocity of the moving structures it encounters in its path. As the cardiac valves move much more quickly than any of the other major components of the foetal heart, this permits the separation of the valvular movements from the extraneous signal generators. The uFCG has the potential to register the closing of the valves and also their opening. Although there are four principal events per cardiac cycle, some of these may not be registered by the Doppler ultrasound technique if the valve movement is across the beam. Furthermore, the signal is very variable and suffers from scintillation effects whenever the reflectors shift

relative to the transducer. Its advantage is that the returns do have a high signal-to-additive-noise ratio.

The ultrasound cardiograph used in conjunction with the phonocardiograph provides a different perspective on certain cardiac events which are registered in common. The more different sensory information that can be integrated into the system the more confidence there will be in the ultimate analysis. Using two different types of transducer also minimizes the amount of information required from each independently. This information redundancy can be exploited to: prompt the analysis of the other transducer signal, act as an independent source of verification for the already-identified events in the other signal, and to combat noise which may be uncorrelated in the two signals. Furthermore, incorporating the uFCG allows the instant of opening of the aortic valve (figure 8.1) to be determined. The occurrence of this event, which is registered only in the uFCG, is necessary for measuring the various phases which occur during the systolic time interval [6, 163, 170].

The systolic time has two constituent phases: the ventricular ejection time (VET), and the iso-volumetric contraction time (IVCT). The VET is the time the heart spends in evacuating the ventricles, and the IVCT is the duration of the contraction of the ventricular walls which precedes the VET. Although the exact significance of these fundamental parameters has yet to be assessed it is thought that they may be better indicators of cardiac function than the more general parameters which relate to heart rate.

The complete system which incorporates both transducer signals is shown in figure 8.2. The *information synthesizer* performs the function of exploiting the redundancy between the transducer signals to facilitate complete analysis of each signal. Ultimately, the information synthesizer should output the instants of closure of both the mitral/tricuspid and the aortic/pulmonary valve pairs and the instant of

opening of the latter valve pair. From these can be derived all the temporal parameters of cardiac function.

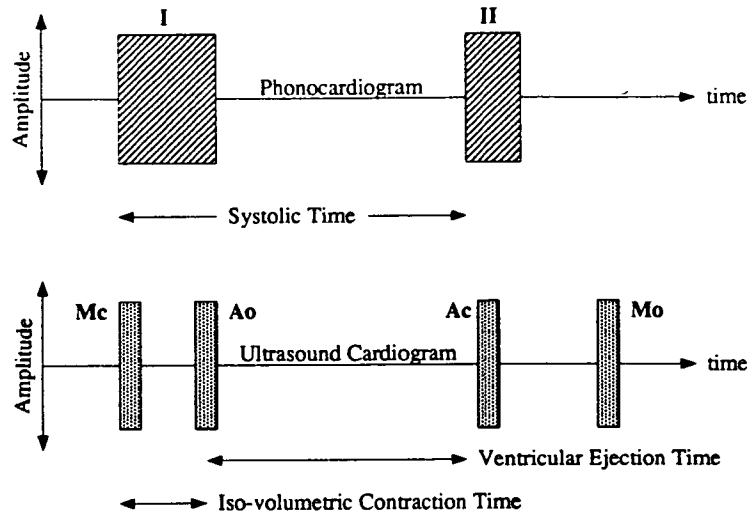


Figure 8.1 Phases of the systolic time interval.

I	First heart sound	Ao	Aortic valve opening
II	Second heart sound	Ac	Aortic valve closure
Mc	Mitral valve closure	Mo	Mitral valve opening

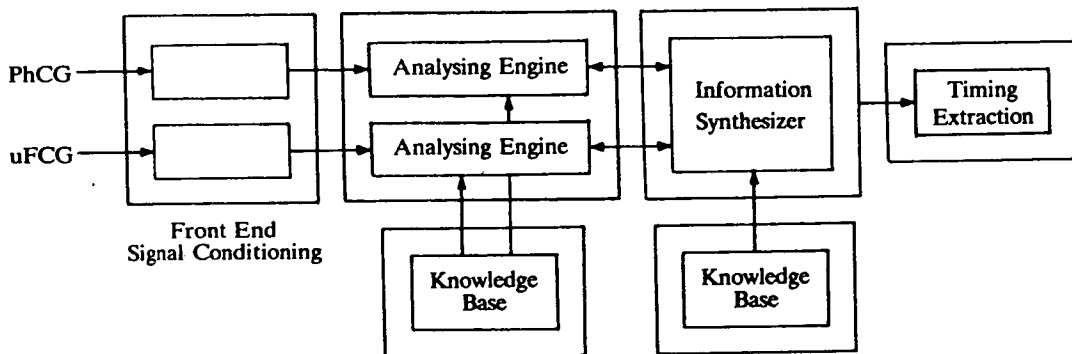


Figure 8.2 Transducer and analysis fusion system.

References

1. D.G. Talbert, W.L. Davies, Frank Johnson, Nicholas Abraham, Nigel Colley, and David Southall, "Wide Bandwidth Fetal Phonography Using a Sensor Matched to the Compliance of the Mother's Abdominal Wall," *IEEE Trans. Biomed. Eng.*, vol. BME-33, no. 2, pp. 175-180, Feb. 1986.
2. Henk G. Goovaerts, Otto Rompelman, and Herman P. Van-Geijn, "A Transducer for Detection of Fetal Breathing Movements," *IEEE Trans. Biomedical Engineering*, vol. 36, no. 4, pp. 471-478, April 1989.
3. Yuji Murata, "Advances on the Horizon," *Clinics in Perinatology — Symposium on Fetal Monitoring*, vol. 9, no. 2, pp. 433-441, W.B. Saunders Company, Philadelphia, June 1982.
4. J. Malcolm Pearce and Keith Willson, "Fetal Heart Rate Monitors — Some Clinical and Technical Aspects," *British Journal of Hospital Medicine*, pp. 123-131, August 1983.
5. J.E.E. Fleming, "Fetal Monitoring," in *Physics in Medicine and Biology Encyclopedia*, ed. T.F. McAinsh, pp. 332-339, Pergamon Press, Oxford.
6. S.P.W. Raymond and C.R. Whitfield, "Systolic Time Intervals in the Fetal Cardiac Cycle," *Balliere's Clinical Obstetrics and Gynaecology*, vol. 1, no. 1, pp. 185-201, March 1987.
7. D.M. Serr, "Methods for Recording Fetal Heart Rate and Uterine Contractions," in *Fetal Physiology and Medicine*, ed. Richard W. Beard, and Peter W. Nathanielsz, pp. 510-530, W.B. Saunders Co. Ltd., London, 1976.
8. J.H.M. Pinkerton, "Kegaradec, Friend of Laennec and Pioneer of Foetal Auscultation," *Proc. Roy. Soc. Med.*, vol. 62, pp. 477-483, 1969.
9. Margaret B. McNay and John E.E. Fleming, "The Safety of Diagnostic Ultrasound," *Brit. J. Hos. Med.*, pp. 35-38, July 1984.
10. Arnon Cohen, *Biomedical Signal Processing*, 1, p. 5, CRC Press, Inc., 1986.
11. W.T. Pommerenke and Francis W. Bishop, "Amplification of Fetal Heart Sounds," *Am. J. Obstet. Gynec.*, vol. 35, pp. 851-854, 1938.

12. Louis H. Hellman, Morton A. Schiffer, Schuyler G. Kohl, and Walter E. Tolles, "Studies in Fetal Well-Being: Variations in Fetal Heart Rate," *Am. J. Obstet. Gynec.*, vol. 76, pp. 998-1010, 1958.
13. François-Issac Mayor, *Bibliothèque Universelle des Sciences, Belles-lettres, et Arts*, vol. 9, p. 249, Geneva, 1818. Cited in Brigette Gultekin-Zootzman, "The History of Monitoring the Human Fetus," *J. Perinat. Med.*, vol. 3, pp. 135-144, 1975.
14. Alistair L. Gunn and Michael C. Wood, "The Amplification and Recording of Foetal Heart sounds," *Proc. Roy. Soc. Med.*, pp. 85-91, 1952.
15. E.J. Quilligan and J. Fitzsimmons, "Fetal Monitoring," in *Clinical Obstetrics*, ed. Carl J. Pauerstein, pp. 249-266, John Wiley, 1987.
16. Thore Solum, "Antenatal Cardiotocography," *Acta Obstet. Gynec. Scand. Suppl.*, vol. 96, pp. 3-31, 1980.
17. Brigitte Gultekin-Zootzmann, "The History of Monitoring the Human Fetus," *J. Perinat. Med.*, vol. 3, pp. 135-144, 1975.
18. Robert C. Goodlin, "History of Fetal Monitoring," *Am. J. Obstet. Gynec.*, vol. 133, no. 3, pp. 323-352, 1979.
19. O'Donel T.D. Browne, *The Rotunda Hospital 1745-1945*, E. & S. Livingstone Ltd., Edinburgh, 1947.
20. J.H.M. Pinkerton, "John Creery Ferguson," *Br. J. Obstet. Gynaec.*, vol. 87, no. 4, pp. 257-260, 1980.
21. Evory Kennedy, *Observations on Obstetric Auscultation*, Hodges and Smith, Dublin, 1833.
22. Naren Patel, *Obstetrics*, pp. 373-382, Church Livingstone, Avon, 1989.
23. J. Hofbauer and O. Weiss, "Photographische Registrierung der Fotalen Herztone," *Zentralbl. Gynaecol.*, vol. 13, p. 429, 1908. Cited in Thore Solum, "Antenatal Cardiotocography," *Acta Obstet. Gynec. Scand.*, supp. 96, pp. 3-31, 1980.
24. J.J. Sampson, R.L. McCalla, and W.J. Kerr, "Phonocardiography of the Human Fetus," *Am. Heart J.*, vol. 1, pp. 717-734, 1926.

25. C. Sureau, "The Clinical Significance of Electronic Methods for Monitoring the Fetal Heart," in *Perinatal Medicine*, ed. Peter J. Huntingford, pp. 109-112, Georg Thieme Verlag, Stuttgart, 1969.
26. C. Wood and P. Renou, "Fetal Heart Rate Monitoring," in *Fetal Physiology and Medicine: The Basis of Perinatology*, ed. Richard W. Beard, and Peter W. Nathanielsz, pp. 454-475, W.B. Saunders Co. Ltd, London, 1976.
27. Leo Jacobsohn, "Amplified Audibility of Heart Sounds," *JAMA*, vol. 80, no. 7, pp. 492-93, 17 Feb. 1923. Cited in the Berlin entry of the foreign letters section.
28. Frederick H. Falls and Alan C. Rockwood, "Use of Microphonic Stethoscope in Demonstration of Fetal Heart Tones," *JAMA*, vol. 81, no. 20, pp. 1683-84, 17 Nov. 1923.
29. C.J. Gamble and D.E. Replogie, "A Multiple Electrical Stethoscope for Teaching," *JAMA*, vol. 82, no. 5, pp. 387-88, 2 Feb. 1924.
30. H.A. Frederick and H.F. Dodge, "The Stethophone," *The Bell Sys. Tech. J.*, vol. 3, pp. 531-549, 1924.
31. W.L. Bierring, H.C. Bone, and M.L. Lockhart, "Use of the Electrostethograph for Recording Heart Sounds (Preliminary Report)," *Jour. Amer. Med. Assoc.*, p. 628, February 1935.
32. Albert S. Hyman, "Irregularities of the Fetal Heart," *Am. J. Obstet. Gynec.*, vol. 20, pp. 332-347, 1930.
33. Harvey B. Matthews, "The Continuous Auscultation of the Fetal Heart by Means of an Amplifying Stethoscope," *Am. J. Obstet. Gynec.*, vol. 34, pp. 898-900, 1937.
34. Edwin J. DeCosta, "The Photostethoscope," *JAMA*, vol. 111, no. 26, pp. 2008-2009, 26 Nov. 1938.
35. Arthur L. Smith and W.J. Hervert, "A Method for Recording and Reproducing Fetal Heart Sounds," *Am. J. Obstet. Gynec.*, vol. 40, pp. 102-110, 1940.
36. Morris Dressler and Samuel N. Moskowitz, "Fetal Electrocardiography and Stethography," *Am. J. Obstet. Gynec.*, vol. 41, pp. 775-791, 1941.

37. Marshall L. Lockhart, "The Stethograph," *Am. Heart J.*, vol. 16, pp. 72-78, 1938.
38. George W. Corner and Herbert M. Stran, "A Fetal Heart Ratemeter," *Am. J. Obstet. Gynec.*, vol. 73, pp. 190-195, 1957.
39. D.H. Smith, "A Foetal Pulse Rate Recorder," *Electronic Engineering*, pp. 132-135, March 1957.
40. L.W. Cox, I.B. Wall, and A.E.R. Wood, "The Design of a Foetal Heart Ratemeter," *Aust. N.Z.J. Obstet. Gynaec.*, vol. 3, pp. 60-62, 1963.
41. Ian Maxwell, "Phonocardiography as a Foetal Heart Monitor," *Aust. N.Z.J. Obstet. Gynaec.*, vol. 3, pp. 63-72, 1963.
42. J.B. Cornwall and F.G. Tattam, "An Instrument for Recording the Fetal Pulse Rate," *Trans. IEEE Biomed. Eng.*, vol. BME-11, pp. 24-28, 1964.
43. J.R. Urbach, E.R. Carrington, M.A. Friel, J.F. Rojas, H.L. Zweizig, M.W. Loveland, C.M. Cavoto, M.D. Tin, and R.L. Lambert, "Characteristics of Fetal Heart Sounds and Murmurs in the Second Half of Gestation," *Biol. Neonat.*, vol. 13, pp. 90-108, 1968.
44. Charles McLennan, *Am. J. Obstet. Gynec.*, vol. 76, pp. 1010-1011, 1958. Discussion of "Studies in Fetal Well-Being: Variations in Fetal Heart Rate," by Louis H. Hellman, *Am. J. Obstet. Gynec.*, vol. 76, pp. 998-1010, 1958.
45. O. Gentner and K. Hammacher, "An Improved Method for the Determination of the Instantaneous Fetal Heart Frequency from the Fetal Phonocardiogram," *7th Int. Conf. on Med. and Biol. Eng.*, Stockholm 1967, p. 140, 1967.
46. K. Hammacher, "The Clinical Significance of Cardiotocography," *Perinatal Medicine*, pp. 80-93, Georg Thieme Verlag, Stuttgart, 1969.
47. D.M. Serr, S.M. Mannor, A. Ron, E. Rabau, H. Zakut, and L. Schaudinshky, "Phonocardiography in Fetal Heart Monitoring," in *Perinatal Medicine*, ed. Peter J. Huntingford, pp. 99-107, Georg Thieme Verlag, Stuttgart, 1969.
48. D.G. Talbert, John Dewhurst, and D.P. Southall, "New Transducer for Detecting Fetal Heart Sound: Use of Compliance Matching for Maximum Sound Transfer," *The Lancet*, pp. 426-27, 25 Feb. 1984.

49. Joachim Nagel, "New Diagnostic and Technical Aspects of Fetal Phonocardiography," *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 23, pp. 295-303, 1986.
50. Aubrey Leatham and Graham J. Leech, "Auscultation of the Heart," in *The Heart*, ed. J. Willis Hurst, pp. 203-249, McGraw-Hill, 1982.
51. Robert F. Rushmer, in *Cardiovascular Dynamics*, p. 300, Saunders, Philadelphia, 1970.
52. Aldo A. Luisada, *The Sounds of the Normal Heart*, Warren H. Green Inc., St. Louis, Missouri, 1972.
53. G.J. Leech and A. Leatham, "Correlation of Heart Sounds and Valve Motions," in *Evaluation of Cardiac Function by Echocardiography*, ed. W. Bleifeld, pp. 89-96, Springer-Verlag, Berlin, 1980.
54. Aubrey Leatham, in *Auscultation of the Heart and Phonocardiography*, J. & A. Churchill, London, 1970.
55. N. Colley, N.G. Abraham, P. Fayers, D.G. Talbert, W.L. Davies, and D.P. Southall, "The Fetal Phonogram: A Measure of Fetal Activity," *The Lancet*, pp. 931-935, 26 April 1986.
56. Doren Cohen, Arthur E. Timbs, and Kevin J. Dalton, "Measurement of Compliance of the Maternal Abdominal Wall in Pregnancy," *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 23, pp. 267-272, 1986.
57. N. Colley, D.G. Talbert, and D.P. Southall, "Biophysical Profile in the Fetus from a Phonographic Record," *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 23, pp. 261-265, 1986.
58. Nigel Colley, *The Fetal Phonogram*. M.D. Thesis, University College Hospital, London. 1988.
59. D.P. Southall, N. Colley, and D.G. Talbert, "Overnight Recordings of Fetal Heart Sounds and Fetal Movements," *Europ. J. Obstet. Gynec. Reprod. Biol.*, vol. 20, pp. 85-86, 1985.
60. M.E. Van-Valkenburg, *Analog Filter Design*, Holt-Saunders, 1982.

61. Horowitz and Hill, *The Art of Electronics*, pp. 155-158, Cambridge University Press, Cambridge, 1987.
62. Trevor Shelley, "Detection of Foetal Heart Sounds: An Analysis," *Am. J. Obstet. Gynec.*, vol. 105, no. 4, pp. 597-603, 1969.
63. Charles M. Steer and George J. Hertsch, "Continuous Observation of the Fetal Heart," *Am. J. Obstet. Gynec.*, vol. 62, pp. 1139-1142, 1951.
64. H.S. Sawyer, "The Design of a Foetal Phonocardiometer," *IRE National Convention Record*, no. 9, pp. 64-72, 1959.
65. E.A.J. Alment, "Low Frequency Foetal Phonocardiography," *J. Obstet. Gynaec. Brit. Comm.*, vol. 69, p. 1033, 1962.
66. J.L. Leonard and W. Farrer, "A Foetal Heart Monitor," *Electronic Engineering*, pp. 83-85, Feb. 1963.
67. G. Weill and R. Schick, "L'electro et la Phonocardiographie Simultanee dans la Surveillance de la Grossesse," *Revue Francaise de Gynecologie et d'Obstetrique*, vol. 59, pp. 57-66, 1964.
68. J. Melchior, "Comments on Fetal Heart Rate Monitoring," in *Perinatal Medicine*, ed. Peter J. Huntingford, pp. 107-108, Georg Thieme Verlag, Stuttgart, 1969.
69. J.H. Dripps, G.K. Manning, and X. Zhu, "A Signal-Processing Research Facility and its Application to the Processing of Fetal Phonocardiographic Signals for Heart Rate Estimation," *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 23, pp. 281-288, 1986.
70. D.M. Holburn and T.D. Rowsell, *Real Time Analysis of Fetal Phonography Signals Using the TMS320*. IEE Colloquium on, "Biomedical Applications of Digital Signal Processing", Digest No. 1989/144, pp. 7/1-7/12, London, 30 November 1989.
71. Hany E. Bassil and James H. Dripps, "Real Time Processing and Analysis of Fetal Phonocardiographic Signals," *Clin. Phys. Physiol. Meas.*, vol. 10, pp. 67-74, 1989.

72. R.A. Roberts and C.T. Mullis, *Digital Signal Processing*, p. 183, Addison Wesley Pub. Co., 1987.
73. Michael R. Neuman, "Fetal Monitoring Techniques," *CRC Critical Reviews in Bioengineering*, pp. 132-148, January 1975.
74. T. Solum, I. Ingemarsson, and N.O. Sjöberg, "Selection Criteria for Antenatal Cardiotocography," *Z. Geburtsh. u. Perinat.*, vol. 183, pp. 212-217, 1979.
75. John V. Kelly, "The Fetal Heart: Comparison of its Electrical and Mechanical Events," *Am. J. Obstet. Gynec.*, vol. 91, no. 8, pp. 1133-1140, 1965.
76. L.G. Durand, M. Blanchard, G. Cloutier, H.N. Sabbah, and P.D. Stein, "Pattern Recognition of Normal, Lightly and Heavily Calcified Bioprosthetic Valves Implanted in the Mitral Position," *Proc. Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 11, pp. 53-54, 1989.
77. L.G. Durand, M. Blanchard, H.N. Sabbah, M.S. Hamid, S.R. Kemp, and P.D. Stein, "A Bayes Model for Automatic Detection and Quantification of Bioprosthetic Valve Degeneration," *Mathl. Comput. Modelling*, vol. 11, pp. 158-163, 1988.
78. A. Iwata, N. Ishii, N. Suzumura, and K. Ikegaya, "Algorithm for detecting the first and second heart sounds by spectral tracking," *Medical and Biological Engineering and Computing*, pp. 19-26, January 1980.
79. R.J. Adolph, J.F. Stephens, and K. Tanaka, "The clinical value of frequency analysis of the first heart sound in myocardial infarction," *Circulation*, vol. 41, p. 1003, 1970. Cited in Ajit P. Yoganathan *et al.*, "Use of the Fast Fourier Transform for the Frequency Analysis of the First Heart Sound in Normal Man," *Medical and Biological Engineering*, pp. 69-73, January 1976.
80. A. Sakai, L.P. Feigen, and A.A. Luisada, "Frequency distribution of the heart sounds in normal man," *Cardiovas. Res.*, vol. 5, p. 358, 1971. Cited in Ajit P. Yoganathan *et al.*, "Use of the Fast Fourier Transform for the Frequency Analysis of the First Heart Sound in Normal Man," *Medical and Biological Engineering*, pp. 69-73, January 1976.

81. Trevor Shelly, "Detection of Fetal Heart Sounds: An Analysis," *Am. J. Obstet. Gynec.*, vol. 105, no. 4, pp. 597-603, 1969.
82. Ajit P. Yoganathan, Ramesh Gupta, Firdaus E. Udwadia, J. Wayen Miller, William H. Corcoran, Radha Sarma, John L. Johnson, and Richard J. Bing, "Use of the Fast Fourier Transform in the Frequency Analysis of the First Heart Sound in Normal Man," *Medical and Biological Engineering*, pp. 69-73, January 1976.
83. Ajit P. Yoganathan, Ramesh Gupta, Firdaus E. Udwadia, William H. Corcoran, Radha Sarma, and Richard J. Bing, "Use of the Fast Fourier Transform in the Frequency Analysis of the Second Heart Sound in Normal Man," *Medical and Biological Engineering*, pp. 455-459, July 1976.
84. Ajit P. Yoganathan, Ramesh Gupta, and William H. Corcoran, "Fast Fourier transform in the Analysis of Biomedical Data," *Medical and Biological Engineering*, pp. 239-244, March 1976.
85. Akira Iwata, Nobuo Suzumura, and Kazuo Ikegaya, "Pattern Classification of the Phonocardiogram Using Linear Prediction Analysis," *Med. & Biol. Eng. & Comput.*, vol. 15, pp. 407-412, July 1977.
86. Tae. H. Joo, James H. McClellan, Rodney A. Foale, Gordon S. Myers, and Robert S. Lees, "Pole-Zero Modelling and Classification of Phonocardiograms," *IEEE Trans. Biomed. Engg.*, vol. BME-30, no. 2, pp. 110-117, February 1983.
87. Louis G. Durand, Jacques De Guise, Guy Cloutier, Robert Guardo, and Maurice Brais, "Evaluation of FFT-Based and Modern Parametric Methods for the Spectral Analysis of Bioprosthetic Valve Sounds," *IEEE Trans. Biomedical Engineering*, vol. BME-33, no. 6, pp. 572-578, June 1986.
88. Alan V. Oppenheim and Ronald W. Schaffer, *Digital Signal Processing*, Prentice-Hall International, Inc., London, 1975.
89. Fredric J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51-83, Jan. 1978.
90. Alan Sowards, "Introduction to DSP," *Electronics and Wireless World*, vol. 94, no. 1630, pp. 741-746, August 1988.

91. R. B. Randall, *Frequency Analysis*, K. Larsen & Son, Denmark, 1987.
92. Frederic J. Harris, *Handbook of Digital Signal Processing*, pp. 253-287, Academic Press, Inc..
93. Steven M. Kay and Stanley L. Marple, "Spectrum Analysis — A Modern Perspective," *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1380-1419, Nov. 1981.
94. D.A. Linkens, "Short-Time-Series Analysis of Biomedical Data," *Proceedings of the IEE part A*, vol. 129, no. 9, pp. 663-672, 1982.
95. S. Lawrence Marple, *Digital Spectral Analysis with Applications*, Prentice-Hall Inc., 1987.
96. John Makhoul, "Linear Prediction: A Tutorial Review," *Proc. IEEE*, vol. 63, no. 4, April 1975.
97. John Makhoul, "Stable and Efficient Lattice Methods for Linear Prediction," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-25, pp. 423-428, Oct. 1977.
98. John Makhoul, "A Class of All-Zero Lattice Digital Filters: Properties and Applications," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-26, no. 4, pp. 304-314, Aug. 1978.
99. J.P. Burg, Maximum Entropy Spectral Analysis. Ph.D. Thesis, Dept. of Geophysics, Stanford University, California. 1975. Cited in John. G. Proakis and Dimitris G. Manolakis, *Introduction to Digital Signal Processing*, Macmillan Pub. Co., 1988.
100. N. Andersen, "Comments on the Performance of Maximum Entropy Algorithms," *Proc. IEEE*, vol. 66, no. 11, pp. 1581-1582, Nov. 1978.
101. Mostafa Kaveh and Geo. Lippert, "An Optimum Tapered Burg Algorithm for Linear Prediction and Spectral Analysis," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-31, no. 2, pp. 438-444, April 1983.
102. John. G. Proakis and Dimitris G. Manolakis, *Introduction to Digital Signal Processing*, Macmillan Pub. Co., 1988.

103. D.N. Swingler, "A Modified Burg Algorithm for Maximum Entropy Spectral Analysis," *Proc. IEEE*, vol. 67, no. 9, pp. 1368-1369, Sept. 1979.
104. D.N. Swingler, "On the Optimum-Tapered Burg Algorithm," *Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-32, pp. 185-186, February 1984.
105. T.J. Ulrych and R.W. Clayton, "Time Series Modelling and Maximum Entropy," *Phys. Earth Planetary Interiors*, vol. 12, pp. 188-200, April 1976. Cited in S. Kay and S. Marple, "Spectrum Analysis — A Modern Perspective," *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1380-1419, Nov. 1981.
106. Lloyd J. Griffiths, "Rapid Measurement of Digital Instantaneous Frequency," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-23, no. 2, pp. 207-222, April 1975.
107. Douglas O'Shaughnessy, "Linear Predictive Coding," *IEEE Potentials*, vol. 7, no. 1, pp. 29-32, February 1988.
108. B.P. Lathi, H.P. Nii, and E.A. Feigenbaum, "Rule-Based Understanding of Signals," in *Pattern-Directed Inference Systems*, ed. D.A. Waterman, and F. Hayes-Roth, pp. 483-501, Academic Press, 1978.
109. Martin A. Fischler and Oscar Firschein, *Intelligence: The Eye, the Brain, and the Computer*, Addison-Wesley, 1987.
110. Peter Jackson, *Introduction to Expert Systems*, Addison-Wesley, 1986.
111. J.G. Holman and M.J. Cookson, "Expert Systems for Medical Applications," *J. of Medical Engineering and Technology*, vol. 11, no. 4, pp. 151-159, July/August 1987.
112. *Building Expert Systems*, ed. Frederick Hayes-Roth, Donald A. Waterman, and Douglas B. Lenat, Addison-Wesley, London, 1983.
113. Ruth Wallsgrove, "Crystal," *Personal Computer World*, pp. 172-175, November 1988.
114. Elaine Rich, *Artificial Intelligence*, McGraw-Hill, 1986.
115. Randall Davis and Jonathan J. King, "The Origin of Rule-Based Systems in AI," in *Rule-Based Expert Systems*, ed. B.G. Buchanan, and E.H. Shortliffe, pp. 20-40, Addison-Wesley, 1984.

116. Herbert Schildt, *Artificial Intelligence Using C*, pp. 71-118, Osborne McGraw-Hill, 1987.
117. Chuck Williams, "Expert Systems, Knowledge Engineering, and AI Tools — An Overview," *IEEE Expert*, pp. 66-70, Winter 1986.
118. D.M. Connah and C.A. Fishbourne, "Intelligent Knowledge-Based Systems," *J. Inst. Electronic and Radio Engineers*, vol. 55, no. 6, pp. 223-228, June 1985.
119. David Guest, "Soothing the Savage Expert," *Computer Systems Europe*, pp. 30-33, September 1988.
120. A. Betta and D.A. Linkens, "Intelligent Knowledge-Based System for Dynamic System Identification," *IEE Proceedings, Pt. D*, vol. 137, no. 1, pp. 1-12, January 1990.
121. J.H. Breuker and B.A. Wielinka, "Techniques for Knowledge Elicitation and Analysis," *Esprit Project 12, Memorandum 28*, 1984.
122. Frederick Hayes-Roth, "The Knowledge-Based Expert System: A Tutorial," *Computer*, pp. 11-28, September 1984.
123. Gordon McCalla and Nick Cercone, "Approaches to Knowledge Representation," *Computer*, pp. 12-18, October 1983.
124. Dana S. Nau, "Expert Computer Systems," *Computer*, pp. 63-85, February 1983.
125. Ronald J. Brachman, "The Basics of Knowledge Representation and Reasoning," *AT&T Technical Journal*, vol. 67, no. 1, pp. 7-24, January/February 1988.
126. Lofti A. Zadeh, "Making Computers Think Like People," *IEEE Spectrum*, pp. 26-32, August 1984.
127. Bruce G. Buchanan, "Extended Bibliography of Expert Systems," *Expert Systems*, vol. 3, no. 1, pp. 40-51, January 1986.
128. T. O'Shea, "Rule-Based Computer Tutors," in *Expert Systems in the Micro-Electronic Age*, ed. Donald Michie, pp. 226-232, Edinburgh University Press, 1979.
129. Kevin D. Ashley and Edwina L. Rissland, "A Case-Based Approach to Modelling Legal Expertise," *IEEE Expert*, vol. 3, no. 3, pp. 70-77, Fall 1988.

130. R. Duda, "Syntel: Using a Functional Language for Financial Risk Assessment," *IEEE Expert*, vol. 2, no. 3, pp. 18-31, Fall 1987.
131. Edward H. Shortliffe, Bruce G. Buchanan, and Edward A. Feigenbaum, "Knowledge Engineering for Medical Decision Making: A Review of Computer-Based Clinical Decision Aids," *Proceedings of the IEEE*, vol. 67, no. 9, pp. 1207-1224, September 1979.
132. *IEEE Trans. on Biomedical Engineering*, vol. BME-36, no. 5. Special Issue on Medical Applications of Artificial Intelligence and Information Systems, May 1989.
133. Gary E. Kopec and Alan V. Oppenheim, "Knowledge-Based Signal Processing," *Trends and Perspectives in Signal Processing*, pp. 1-6, July 1982.
134. C.T. Kitzmiller and J.S. Kowalik, "Symbolic and Numerical Computing in Knowledge-Based Systems," in *Coupling Symbolic and Numerical Computing in Expert Systems*, ed. J.S. Kowalik, pp. 3-17, Elsevier Science Publishers, 1986.
135. Evangelos E. Milios and S. Hamid Nawab, "Signal Abstractions in Signal Processing Software," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, no. 6, pp. 913-928, June 1989.
136. John Hallam, "Artificial Intelligence and Signal Understanding," Dept. of Artificial Intelligence, University of Edinburgh, Edinburgh. Research paper no. 384, 1988.
137. H.P. Nii, E.A. Feigenbaum, J.J. Anton, and A.J. Rockmore, "Signal-to-Symbol Transformation: HASP/SIAP Case Study," *The AI Magazine*, pp. 23-35, Spring 1982.
138. Lee D. Erman, Frederick Hayes-Roth, Victor R. Lesser, and D. Raj Reddy, "The Hearsay-II Speech Understanding System: Integrating Knowledge to Resolve Uncertainty," *Computing Surveys*, vol. 12, no. 2, pp. 213-251, June 1980.
139. H. Penny Nii, "Blackboard Systems: The Blackboard Model of Problem Solving and the Evolution of Blackboard Architectures," *The AI Magazine*, pp. 38-53, Summer 1986.

140. Barbara Hayes-Roth, "The Blackboard Architecture: A General Framework for Problem Solving," Heuristic Programming Project, Computer Science Dept., Stanford University. Report no. HPP-83-30, May 1983.
141. Victor R. Lesser, Richard D. Fennell, Lee D. Erman, and D. Raj Reddy, "Organization of the Hearsay-II Speech Understanding System," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-23, no. 1, pp. 11-24, February 1975.
142. H. Penny Nii and Nelleke Aiello, "AGE (Attempt to Generalize): A Knowledge-Based Program for Building Knowledge-Based Programs," *Proceedings of the Sixth IJCAI*, pp. 645-655, 1979.
143. H. Penny Nii, "Blackboard Systems: Blackboard Application Systems, Blackboard Systems from a Knowledge Engineering Perspective," *The AI-Magazine*, pp. 82-106, August 1986.
144. Robert Englemore and Allan Terry, "Structure and Function of the Crysalis System," *Proceedings of the Sixth IJCAI*, pp. 250-256, 1979.
145. Allan Terry, "Using Explicit Strategic Knowledge to Control Expert Systems," in *Blackboard Systems*, ed. Robert Englemore, and Tony Morgan, pp. 159-187, Addison-Wesley, 1988.
146. Randall Davis and Bruce G. Buchanan, "Meta-Level Knowledge," in *Rule-Based Expert Systems*, ed. B.G. Buchanan, and E.H. Shortliffe, pp. 507-526, Addison-Wesley, 1984.
147. Y. Maki, "Studies of the Electrophonocardiogram in Fetuses and Newborns," *Acta Medica (Fukuoka)*, vol. 35, pp. 218-255, 1965.
148. Simon A. Hayward, "Is a Decision Tree an Expert System?," in *Research and Development in Expert Systems*, ed. M.A. Bramer, pp. 185-192, Cambridge University Press, 1985.
149. Beverly Thompson and William Thompson, "Finding Rules in Data," *BYTE*, pp. 149-155, November 1986.
150. J.S. Aikins, J.C. Kunz, E.H. Shortliffe, and R.J. Fallat, "PUFF: An Expert System for Interpretation of Pulmonary Function Data," *Computers and*

- Biomedical Research*, vol. 16, no. 3, pp. 199-208. cited in J.G. Holman and M.J. Cookson, "Expert Systems for Medical Applications," *Journal of Medical Engineering and Technology*, vol. 11, no. 4, pp. 151-159, July/August 1987.
151. M.P. Georgeff, "A Framework For Control in Production Systems," *Proceedings of the IJCAI*, pp. 328-334, 1979.
 152. Michael Georgeff and Umberto Bonollo, "Procedural Expert Systems," *Proceedings of the IJCAI*, pp. 151-157, 1983.
 153. M.P. Georgeff, "Procedural Control in Production Systems," *Artificial Intelligence*, vol. 18, pp. 175-201, North-Holland, 1982.
 154. Michael P. Georgeff and Amy L. Lansky, "Procedural Knowledge," *Proceedings of the IEEE*, vol. 74, no. 10, pp. 1383-1398, October 1986.
 155. Ron Sauers, "Controlling Expert Systems," in *Expert System Applications*, ed. L. Bolc, and M.J. Coombs, pp. 79-197, Springer-Verlag, 1988.
 156. B.M.E. Moret, "Decision Trees and Diagrams," *ACM Computing Surveys*, vol. 14, no. 4, 1982.
 157. Joseph A. Vrba and Juan A. Herrera, "Expert System Tools: The Next Generation," *IEEE Expert*, pp. 75-76, Spring 1989.
 158. J. Ross Quinlan, *Applications of Expert Systems*, Addison-Wesley, 1987.
 159. T.M. Crawford and V. Marton, "A Machine Learning Approach to Expert Systems for Fault Diagnosis in Communications Equipment," *Computer-Aided Engineering Journal*, pp. 31-38, February 1987.
 160. Simon Hayward, "Diagnostic Expertise and its Use in Commercially Viable Expert Systems," in *Advances in Artificial Intelligence*, ed. T. O'Shea, pp. 71-75, North-Holland, 1985.
 161. Paul A. Lynn, *An Introduction to the Analysis and Processing of Signals*, Macmillan, 1984.
 162. Anthony D. Whalen, *Detection of Signals in Noise*, 1971. (ISBN: 0-12-744850-0; details of publisher missing)
 163. Robert Goodlin, John Girard, and Arno Hollmen, "Systolic Time Intervals in the Fetus and Neonate," *Obstetrics and Gynecology*, vol. 39, no. 2, pp. 295-302,

February 1972.

164. Keith R. Greene, "Fetal cardiac Assessment," in *Non-Invasive Physiological Measurements*, ed. Rolfe, vol. 2, pp. 1-60, Academic Press, London, 1983.
165. Sonicaid Ltd. Manual for FM/2 Fetal Monitor. Sonicaid Ltd., Sussex, England.
166. Jack C. Amato, "Fetal Heart Rate Monitoring," *Am. J. Obstet. Gynecol.*, vol. 147, no. 8, pp. 967-969, December 1983.
167. T. Fukushima, C.A. Flores, E.H. Hon, and E.C. Davidson, "Limitations of Autocorrelation in Fetal Heart Rate Monitoring," *Am. J. Obstet. Gynecol.*, vol. 153, no. 6, pp. 685-692, November 1985.
168. W.D. Post and G.D. Carson, "Masking of Supraventricular Tachycardia by Electronic Monitoring," *Can. Med. Assoc. J.*, vol. 131, no. 5, pp. 462-464, September 1984.
169. Murray Campbell and Clive Greated, *The Musician's Guide to Acoustics*, J.M. Dent & Sons Ltd., London, 1987.
170. J.E.E. Fleming, S.P.W. Raymond, G.C.S. Smith, and C.R. Whitfield, "The Measurement of Fetal Systolic Time Intervals: Lessons from Ultrasound," *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 23, pp. 289-294, Elsevier Science Publishers B.V., 1986.

APPENDIX

ORIGINAL PAPERS

The thesis is based on the following publications:

References

1. J.T.E. McDonnell, *The Knowledge-Based Processing and Analysis of Foetal Phonocardiograms*. [Awarded the IEE National Younger Members' Premium Award 1989], [Awarded the IEE Harry Henniker Premium 1989].
2. J.T.E. McDonnell, J.H. Dripps, and P.M. Grant, "The Knowledge-Based Processing and Analysis of Foetal Phonocardiograms," *Proceedings of the Fourth International Conference on the Applications of Artificial Intelligence in Engineering (AIENG 89)*, Cambridge, 1989.
- 3.[†] J.T.E. McDonnell, J.H. Dripps, and P.M. Grant, "Processing and Analysis of Foetal Phonocardiograms," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 11, pp. 61-62, Seattle, 1989.
4. J.T.E. McDonnell, J.H. Dripps, and P.M. Grant, "The Knowledge-Based Detection, Segmentation, and Classification of Foetal Heart Sounds," *Proceedings of the IEE Colloquium on the Application of Artificial Intelligence Techniques to Signal Processing*, vol. 1989/42, pp. 6/1-6/4, London, March 1989.
5. J.T.E. McDonnell, "Rule-Based Processing of Foetal Phonocardiograms," in *Artificial Intelligence: Concepts and Applications in Engineering*, ed. A.R. Mirzai, pp. 171-194, Chapman and Hall, London, 1990.
- 6.[†] J.T.E. McDonnell, "Knowledge-Based Interpretation of Foetal Phonocardiographic Signals," *IEE Proceedings, Part F*, vol. 137, no. 5, pp. 311-318, October 1990.

[†] Reprinted in this appendix.

The work in this thesis has also been presented at the following:

"The Knowledge-Based Processing and Analysis of Foetal Phonocardiograms," *South-East Scotland IEE Younger Members' Papers Evening*. [Awarded the IEE National Younger Members' Premium for Best Lecture Presentation], [Awarded the Faraday House Commemorative Prize for Best Lecture Presentation], [Awarded the South-East Scotland IEE Younger Members' Premium for Best Lecture].

"Recent Developments in the Analysis of Foetal Phono Signals," *Meeting of the Working Group on Fetal Phonocardiography*, The University Engineering Dept., Cambridge, January 1990.

"The Knowledge-Based Detection, Segmentation, and Classification of Foetal Heart Sounds," *Symposium on Current Trends in Knowledge-Based Systems*, Heriot-Watt University, Edinburgh, April 1989.

PROCESSING AND ANALYSIS OF FOETAL PHONOCARDIOGRAMS

E. Mc Donnell J.H. Dripps P. Grant

Dept. of Electrical Eng., University of Edinburgh,
Mayfield Road, Edinburgh. U.K. EH9 3JL.

ABSTRACT

An original technique is presented to segment the foetal phonocardiogram into the systolic and diastolic phases on a per beat basis. All the parameters of cardiac function are thereby derivable. Contrary to other methods of automated segmentation the system uses only one transducer and no other input. The technique makes no assumptions about the foetal heart rate nor the beat-to-beat variability. The system relies on external knowledge and contextual information to aid signal analysis. System performance is resistant to high levels of noise and missing heart sounds. Accuracy is good.

INTRODUCTION

Traditionally the obstetrician would listen with a stethoscope to the sounds produced by the foetal heart to determine the foetal heart rate (FHR). Although the ear is unequalled in detecting sounds over a particular frequency range, the spectrum occupied by the foetal heart sounds is on the threshold of audibility. This aural processing has always caused problems relating to accuracy and inter-auditor agreement. The foetal phonocardiogram (PhCG) presents a way of avoiding these problems. The greater low-frequency sensitivity of the phonocardiographic transducer [1] allows all the acoustic information to be registered. The visual processing of the PhCG is the only reliable way of determining the temporal measures of cardiac function from the acoustic signals of the heart. This visual processing is satisfactory when only short lengths of data are considered. However, when long term monitoring is used the amount of data produced is considerable. There is, therefore, a pressing need for the automation of the process.

DATA CAPTURE AND CONDITIONING

The 'TAPHO' [1] transducer is used to pick up the sounds produced by the functioning of the foetal heart. The comparatively wide bandwidth of this transducer (20-250Hz) allows all the components of the principal heart sounds to be registered which previous transducers, which were narrow band, attenuated [1]. For system development purposes and experiment repeatability the acoustic signal was recorded directly onto magnetic tape. On playback the signal was passed through a 5th order 100Hz low-pass analog Bessel-Thomson filter which acts to prevent aliasing in the ensuing digitizer. As the objective is the extraction of timing information from the signal, phase linearity is essential - for which this type of filter is optimal. The resulting signal was A/D converted to 12 bits/word at 500Hz - to allow time-domain resolution. Once sampled the signal is high-passed filtered at 20Hz with a 11th order digital filter. Although others have set this cut-off frequency higher, between 40Hz and 50Hz [1], the significant components of the foetal heart sounds lie yet lower in the spectrum - as far as 20Hz [5]. The signal is divided

into contiguous blocks of 4s duration for subsequent processing. It would be unrealistic to expect identification of heart sounds on arrival because of the variability in all aspects of a heart sound and its similarity to spurious sounds. The 'context' provided by the surrounding sounds in the data block allows supporting evidence to be drawn from a global perspective in the sound identification process.

PHONOCARDIOGRAM

The signal which results from the above conditioning is known as the foetal phonocardiogram (PhCG) (fig. 1). Although within a sound the S/N ratio is high, on a per data block basis it is poor. This arises from the many non-valvular events which are also registered in the PhCG. Potentially the PhCG may register many spurious sounds which complicates the task of identifying the valvular sounds. The variability of the sound generators and the complex transmission medium combine to produce 'heart sounds' which often change markedly from beat to beat in frequency, amplitude (extending to complete attenuation), morphology, duration, and in instant of occurrence relative to the previous beat.

SIGNAL PROCESSING

To combat the variability in all aspects of a heart sound some descriptor was required which was resilient to the fluctuations. Previous attempts with auto- and crosscorrelation with a known first or second sound in the latter case, resulted in enhancement of noise and non-valvular events to the extent that they exceeded the heart sounds. The assumptions which the sound detection process makes are: the arrival of a sound causes a local increase in amplitude, the sound has one critical point per excursion, and the frequency of a sound is relatively low compared to the non-valvular events. Except for a further few weak assumptions *q.v.* no other attribute of a valvular sound was assumed. The measure used a detector squares the peak amplitude per excursion, *i.e.* for one turning point or extrapolates from the last encountered peak if there is more than one, and multiplies this by the excursion duration. Although this bears a similarity to the energy per half cycle in the case of an analog sinusoid the metric was chosen so that events which have higher amplitude and lower frequency are enhanced to the exclusion of all others.

Although this metric produces good detections there remains low lying inter-sound events. These are removed with an adaptive threshold which considers how the metric is varying both locally and globally. The results for a typical PhCG are illustrated in figure 2.

The outcome of the detection processing is the formation of natural groups of excursion-blocks which are non-zero whenever a significant event is present. However, these clusters may not be comprised of one single event. A common occurrence is the

coalescence of a diastolic sound with a first sound which produces an apparently homogeneous event in the PhCG (fig. 1 a,b,c,d,e). It has been found that this may be uncovered by examining the distribution of the excursion durations which is a measure of frequency. Whenever an individual valvular sound is present it always exhibits a spectrum which tends from high to low across the sound. Whenever the above mentioned coalescence occurs there will be two frequency distributions across the event which both exhibit this characteristic. This allows the two events to be partitioned into two separate and homogeneous clusters.

SIGNAL ANALYSIS

Once the events have been enhanced and noise reduced the identification of the principal heart sounds remains. In identifying the first and second heart sounds the relative constancy of the systolic time (ST), *i.e.* the time interval between a first and a second heart sound, is exploited. It is known that the ST varies less than the diastolic time [6] and is always the shorter of the two. The limits of foetal systolic variability have been determined elsewhere [7]. Knowing the extent of the ST permits a reduction in the number of necessary shifts in the autocorrelation function. In the signal analysis each excursion cluster is 'rationalized' (fig. 3), *i.e.* the shape is reduced to a rectangle whose amplitude is the sum of its constituent excursion metrics and whose duration is reduced to be equal to the average duration of a second sound [7]. This gives a strong correlation peak whenever all the first sounds are aligned with the seconds, and makes the autocorrelation function monotonically increase and decrease which unambiguously locates the peak of the autocorrelation. The time lag to this peak is the ST. The 'rationalized' version of the PhCG is now searched to ascertain all possible first/second sound couples. The principal sounds may now be identified using contextual clues and 'domain knowledge', *e.g.* the foetal heart rate is restricted to lie between 60bpm and 220bpm. Having by now found the first and second sounds it is a simple matter to measure the time lapse between a second sound and a first to find the diastolic time, and between a first and the next first to find the instantaneous foetal heart rate and beat-to-beat difference.

SUMMARY

A technique has been presented which allows the automatic identification of the systolic and diastolic times of the foetal cardiac cycle. It makes few assumptions about the nature of the first and second heart sounds, and none about the extent of short term variability in the FHR. This overcomes reported shortcomings in the processing circuitry of other rate detectors [2][3][4]. The locating and identification of the principal heart sounds by the system correlates closely with those determined by visual inspection of the PhCG. As the transducer is passive and non-invasive it may be used for extended monitoring periods of the foetus.

- [1] Talbert, D.G., et al., "Wide Bandwidth Foetal Phonography Using a Sensor Matched to the Compliance of the Mother's Abdominal Wall", *IEEE Trans Biomed Eng*, BME-33, pp. 175-180: February 1986.
- [2] Amato, J.C., "Fetal Heart Rate Monitoring", *Am J Obstet Gynecol*, 147(8), pp. 967-9: December 1983.
- [3] Post, W.D., "Masking of Fetal Supraventricular Tachycardia by Electronic Monitoring Artefact", *Can Med Assoc J*, 131(5), pp. 462-4: September 1984.
- [4] Fukushima, T., et al., "Limitations of Autocorrelation in Fetal Heart Rate Monitoring", *Am J Obstet Gynecol*, 153(6), pp. 685-692: November 1985. Nagel, J., "New Diagnostic and Technical Aspects of Fetal Phonocardiography", *Eur J Obstet Gynecol Reprod Biol*, 23, pp. 295-303: 1982.
- [6] Goodlin, R.C., et al., "Systolic Time Intervals in the Fetus and Neonate", *Obstet and Gynecol*, 39, pp. 295-302: February 1972.
- [7] Kelly, J.V., "The Fetal Heart: Comparison of its Electrical and Mechanical Events", *Am J Obstet Gynecol*, 91(8), pp. 1133-1140: April 1965.

Author

Mr. Edward Mc Donnell, Dept. of Electrical Eng.. The King's Buildings, Mayfield Road, Edinburgh. EH9 3JL. U.K. (tel. 031 667 1081 ext. 3627).

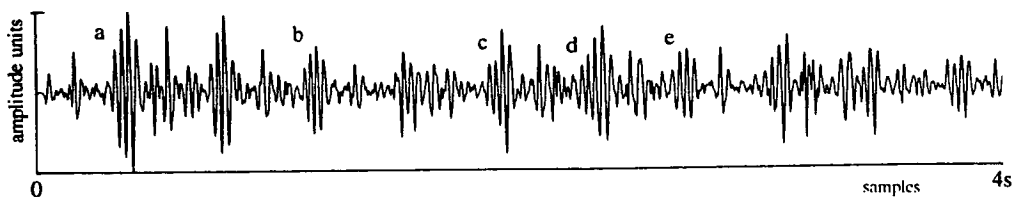


Fig. 1: Typical PhCG (letters refer to text).

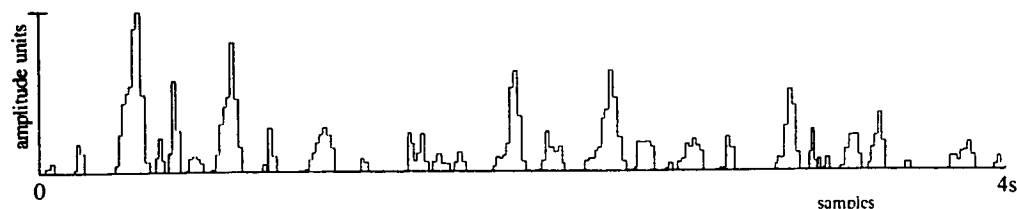


Fig. 2: Significant detected events.

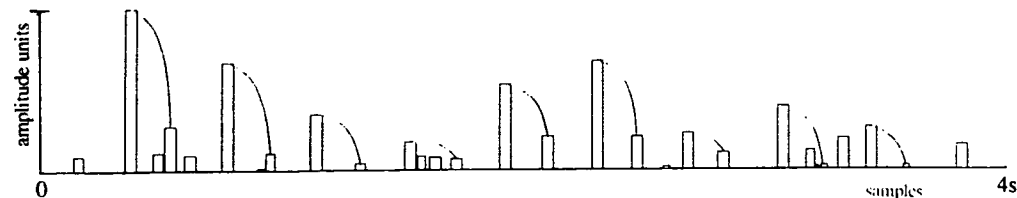


Fig. 3: 'Rationalized' version of the PhCG showing identified I. II couples.

Knowledge-based interpretation of foetal phonocardiographic signals

J.T.E. McDonnell

Indexing terms: Artificial intelligence, Biomedical applications, Signal processing

Abstract: A knowledge-based system (KBS) is presented for the processing and analysis of foetal phonocardiographic signals. Conventional analytic techniques of signal processing are inappropriate for such a volatile signal environment where little information about the signal events is known beforehand. The essential feature of this KBS is its hierarchical organisation of procedurally based knowledge areas. This formulation ensures that the knowledge is consistent and also allows an efficient search for applicable rules. An analysis of the signal is effected using a solution island driving approach. The performance of the system is comparable to that achieved by a domain expert.

1 Introduction

Signal processing may be defined as the operation or operations whereby the useful information in a set of data is extracted from the irrelevant. What constitutes useful, or conversely irrelevant, information is very much dependent on those features which are being sought from the signal.

Traditionally, signal processing has been concerned with the application of numerical algorithms to improve the signal-to-noise ratio or to transform the domain of the signal, thereby accentuating certain of its features. In both cases the characteristics of the signal and the noise are known beforehand. These numerical techniques rely on a mathematical model of the system and certain idealising assumptions. Such an analytical approach fails when the signal characteristics are not well defined, or when there is insufficient understanding of the underlying mechanism of signal production, or where the system is too complex to be modelled mathematically. In these situations a more pragmatic approach is required. Such an approach is provided by the knowledge-based formalism as applied to signal processing [1]. The rationale behind this formulation is essentially empiricism, insofar as the knowledge which represents the associations within and between events is induced from accumulated observations of specific instances. Thereafter, this knowledge about the domain, which is usually expressed as

condition-action pairs [1], will, in an individual case, provide the inference to be made whenever its conditional antecedent is satisfied. This 'domain knowledge' represents the captured experience of someone who has expertise in analysing the particular type of signal, e.g., and electrocardiographer. The power of this formalism lies in its empiricism and the specificity of the domain knowledge, which are both inherently close to reality. Obtaining this knowledge is one of the difficulties which arise in systems which use knowledge-based methodologies [2]. The use of such methods of signal processing does not preclude the deployment of numerical signal-processing algorithms, as these may be accommodated within the system as yet another aspect of the expert's knowledge.

The goal of the processing in this domain is to automatically identify certain events in the foetal cardiac cycle from the recorded acoustic signal produced by the functioning of the foetal heart. This task is doubly complicated because, firstly, the clinician is more concerned with the extremes of variation and irregularities than the norms and, secondly, the signal generators and transmission medium are nonlinear, nonstationary, non-deterministic, time-varying, noisy, and prone to transients.

Owing to such variability in the signal and its low signal-to-noise ratio, conventional processing methods have failed [3-7] and this has led to the abandonment of foetal phonocardiography in the 1960s. Since then its place in foetal monitoring has been taken by methods using Doppler ultrasound. Recently, however, improvements in the design of phonotransducers [8] have brought renewed interest in the technique especially as concern has always been expressed about the possible harmful effects of irradiating the developing embryo with ultrasound energy, albeit in minute quantities. Phonocardiography has no such drawback because its transduction principle is passive; it is eminently suitable for long-term continuous monitoring.

Although recent transducer designs have improved the achievable signal-to-noise ratio, there still exists the problem of interference from the many adventitious sounds associated with the operation of the heart. It is this, coupled with the inherently variable nature of the heart sounds themselves, which has caused the failure of previous efforts to automate the determination of heart rate.

Considering this, the only recourse was to imitate the procedures used by the domain expert while analysing foetal phonocardiograms. This was embodied within the framework of a knowledge-based procedural expert system. The flexibility offered by such a formalism allows the variability in the signal to be accommodated.

Paper 7458F (C4, E5, S9), first received 6th September 1989 and in revised form 18th June 1990

The author is with the Department of Electrical Engineering, University of Edinburgh, The King's Buildings, Mayfield Road, Edinburgh EH9 3JL, United Kingdom

Before describing the organisation of the expert system, some background information will be presented on the production of heart sounds and the problems associated with the analysis of such a signal.

2 Foetal heart sounds and the phonocardiogram

The foetal heart, like that of the adult, is divided into two pairs of chambers and has four valves (Fig. 1). The upper pairs of chambers (the atria) are connected to the lower pair of chambers (the ventricles) by the mitral and tricuspid valves (MT). The ventricles, in turn, are connected to the arterial network through the aortic and pulmonary valves (AP). Viewed as a mechanical system the atria act as primer pumps for the ventricles, which are the power pumps.

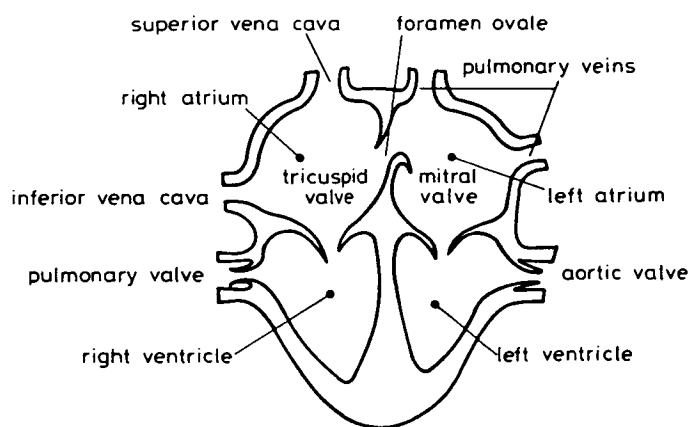


Fig. 1 Foetal heart

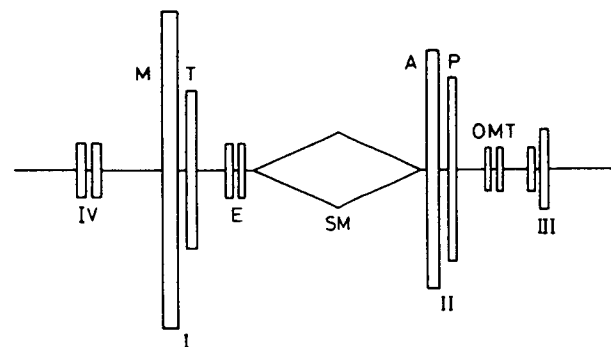


Fig. 2 Idealised normal phonocardiogram

I = first heart sound; III = third heart sound; E = ejection sound; OMT = opening sound of MT valves; T = tricuspid component; P = pulmonary component; II = second heart sound; IV = fourth heart sound; SM = systolic murmur; M = mitral component; A = aortic component

Before the onset of cardiac depolarisation, the blood returning from the veins and the oxygenated blood from the placenta accumulate in the atria, owing to the closure of the MT valves. The cardiac cycle begins with the contraction of the atrial walls, which in turn pressurises the enclosed blood. As the ventricles are both relaxed and evacuated, the pressure differential across the MT valves forces the leaflets to open and the blood to cascade into the ventricles. At this point the ventricles begin to contract, which makes the blood attempt to flow back into the lower pressure atrial chambers. The reverse flow of blood is caught and arrested by the shutting of the MT valves, which produces the first heart sound (I). The ventricular walls continue to contract and the pressure in the enclosed blood rises. Whenever the pressure becomes too great for the AP valves to withstand, they open, and the pressurised blood is rapidly ejected into the arteries.

While the ventricles are being evacuated the pressure of the remaining blood decreases with respect to that in the arteries. This pressure gradient causes the arterial blood to flow back into the ventricles. The AP valves, which perform the same function as the MT valves, arrest this reverse flow by shutting, which gives rise to the second heart sound (II). This sequence of events, called the cardiac cycle, then repeats.

A record of the variation of this acoustic signal with time is known as the phonocardiogram (PhCG) (Fig. 2) and is obtained with a phonocardiographic transducer attached to the mother's abdomen [8]. Identification of the two principal heart sounds of the cardiac cycle permits measurement of the instantaneous foetal heart rate, beat-to-beat differences, and the duration of systole and diastole. These measurements are sensitive indicators of cardiac function, and hence of foetal well-being.

3 Domain anomalies

In contrast to adult phonocardiography, where a strong sound generator is close to the transducer, foetal phonocardiography has to contend with a weak sound generator which is separated from the sensor by up to ten heart diameters.

Whenever the closing of the cardiac valves creates a sound, this energy packet must travel through a complex and dynamic system up to the maternal abdominal surface. At the outset the intensity of a principal sound depends on factors such as pressure differential across the valve, size of the valve leaflets, the initial extent of separation of the leaflets, and the volume of blood which is modulated by foetal breathing [9]. All of these factors, except leaflet size, would be expected to change from beat to beat. The inherent variability of the foetal heart rate (60–220 beats per minute), the irregularity of which is an indication of normality, originates in the variability of the durations of the systolic and diastolic phases. This again is reflected in the relative instants of closure of the valves.

The sound transmission path is made up of the fluid-filled amniotic cavity, the muscular wall of the uterus, layers of fat, tissue, and possibly bony and cartilaginous material. Each of these substances will attenuate the forward going sound energy both in the bulk of the material, and by reflection arising from the impedance mismatch which occurs at the boundary of each of these layers [8]. The natural screen surrounding the foetus has the effect of diminishing the energy of the principal heart sounds reaching the phonographic transducer. In addition to this, there will be extraneous sounds produced by the umbilical cord, foetal movements and environmental noise which will superimpose themselves onto the acoustic signal from the heart. Finally, at the material abdominal surface, the position of the transducer will also considerably alter the shape of the sound waveform.

This complex system would be expected to modify every characteristic of the original signal. Not only will this system be varying on a short-term basis, but also in the longer term as the gestation progresses and the foetus grows in size.

Overall these factors combine to produce 'heart sounds', which often change markedly from one beat to the next in frequency, amplitude, duration, morphology and in instant of occurrence relative to the previous beat, and even to the complete attenuation of a sound. Along with this are the other sounds/noises produced by the mechanical functioning of the heart: blood circulation sounds, ejection sounds, entry sounds, and eddies; these

all contribute to complicate further the already difficult task of identifying the principal cardiac sounds.

This variability in the characteristics of the principal heart sounds precludes compiling a catalogue of types for each sound either for a particular gravida or on a per data block basis.

Fig. 3 illustrates occurrences of certain of these anomalies in the phonocardiogram. They were all recorded within 60 s from the same gravida.

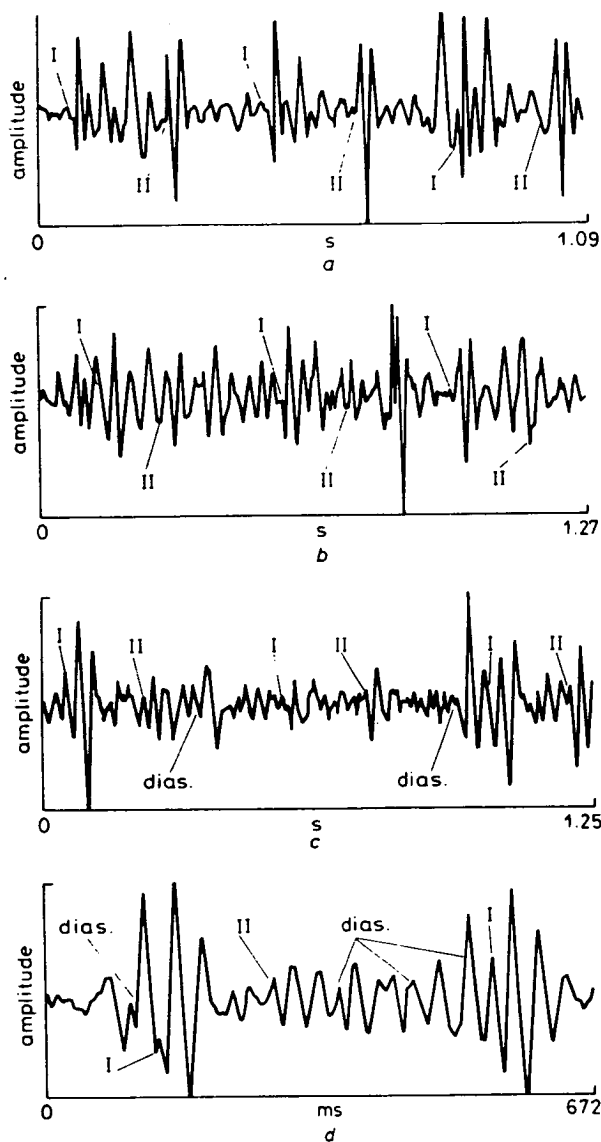


Fig. 3 Anomalies in the phonocardiogram

- a Impulsive I and II sounds; large systolic and diastolic sounds; similarity of I and II sounds
- b Continuum of events; I and II sounds barely distinguishable from background
- c Weak I and II sounds; large diastolic sound
- d Oscillatory pattern of all sounds; poorly defined end points; weak II sound; diastolic murmurs

4 System organisation

The PhCG processing and analysis expert system is organised hierarchically on three levels (Fig. 5). The upper levels of the structure contain the strategy and task modules which together comprise the control element of the system, while the signal-processing modules make up the lowest level. These latter modules are applied under the guidance of the control knowledge areas, and operate directly on the primary data and its abstractions. The task knowledge areas are themselves under the direction of the strategy control module.

The knowledge areas (KA) [10], whether control or signal-processing modules, each contain knowledge about a particular aspect of the problem. These KAs consist of an *invocation part* and a *body*. The invocation, which is similar to the antecedent of traditional production rules, is used to trigger the KA. The body of the KA, which performs a function similar to the consequent of a production, consists of a large amount of domain-specific procedural knowledge. This knowledge is held within a decision-tree structure. Fig. 4 illustrates an

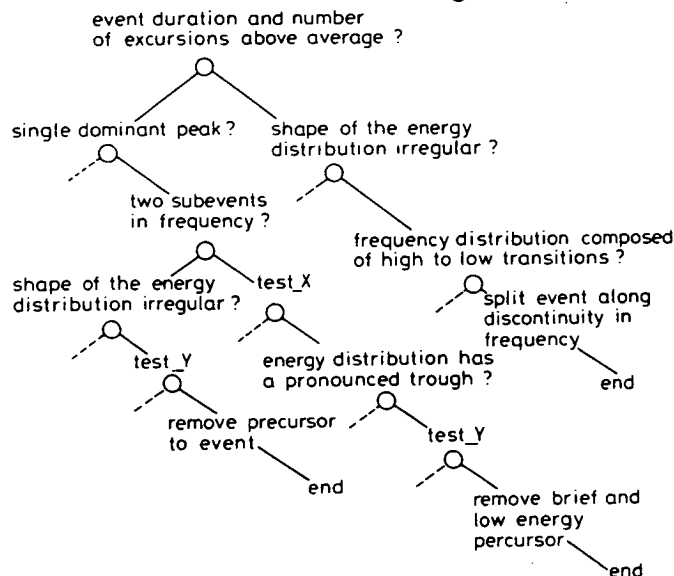


Fig. 4 Part of the segmentation knowledge area

Test_X: Would removal of suspect subevent in frequency cause the resulting event to be too brief, or have too few cycles, or to have too low an energy?

Test_Y: Would removal of suspect subevent in time cause the resulting event to be too brief, or have too few cycles, or to have too low an energy?

example of part of a KA which was taken from the event segmentation module.

The following Sections will detail how control effort is expended in applying the signal processing KAs, and how the embedded knowledge is applied in the identification and classification of signal events in the global database.

4.1 Control system organisation

The control mechanism is structured in the form of hierarchically organised meta-knowledge areas (Fig. 5). These

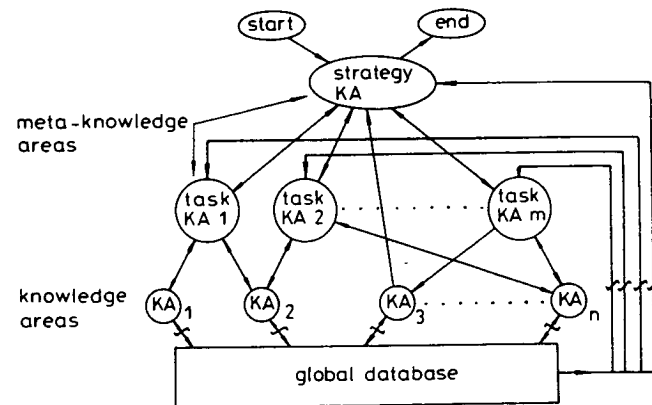


Fig. 5 Control system architecture

Lines marked with ✓ are data lines; all other lines are control lines

meta-knowledge areas are KAs which embody knowledge about the application of the more specific KAs on the lower levels.

At the uppermost level of the control structure is the strategy knowledge area. It is the function of this KA to

initialise the data-base, to determine whether *prima facie* it is worthwhile attempting an analysis, to establish solution islands, to focus attention on a particular area, to revisit the lower levels of data abstraction in the light of contextual information, and to realise when to abandon the analysis and when a total, or the best partial, analysis has been achieved.

On the intermediate level there are task KAs. Once the strategy KA has determined what to do, it will invoke one or more of these task KAs. The task KAs autonomously guide the application of the object-level KAs, using knowledge embedded as a decision tree. Each task has a certain area of expertise, e.g., finding a solution island, which it will use in the region defined by the strategy KA.

On the lowest level of the control structure are the signal-processing and analysis KAs. These object-level KAs are under the control of one or more task KAs. It is these KAs which are, of all the KAs, the only ones which may generate, modify, or delete entries in the database. Each KA at this level may be viewed as a local expert in some aspect of the analysis. The body of these KAs consists of knowledge embedded in a decision tree. Although these KAs are independent, i.e., one KA cannot trigger or fire another, they may influence one another through the entries or modifications they generate in the database.

Each level in the control structure communicates only with its calling module. This communication takes the form of either a *true* or a *false*. A *false* signifies that the KA precondition was not met, whereas a *true* signifies that the rule both triggered and fired. Depending on the outcome, the calling KA will decide, within its remit, what to do next.

4.2 Control execution cycle

Control originates at the strategy level and proceeds down the hierarchy. The strategy level elects one or more of the task-level KAs to execute some policy. The power

of control now resides in the selected task KA and the strategy KA is suspended. The task KA will autonomously attempt to realise its function until it either succeeds or fails whereupon, in either case, control resumes to the strategy KA. When the task KA causes an applicable data-level KA to fire, control will pass down to this, the lowest level. When the data-level KA reports back to the task level, it also transfers control. This brings the task KA out of suspension. This passing of control between the two lowest levels will continue until the task has been accomplished or its execution is terminated. The power of control will then return to the strategy KA.

4.3 Knowledge application

There exist two levels of data abstraction which are maintained in the database along with the primary data. The lower level of the abstraction consists of the signal events that are considered to be significant and the higher level these same events, segmented or partially reconfigured where applicable (Fig. 6). Once the primary data has been processed it will not be revisited, however it is retained for records of possible problem areas and sounds to await later investigation by the domain expert.

The processing of a PhCG takes place on and between three levels on a scale of increasing data abstraction (Fig. 6). On the lowest level there exists the signal processing, followed by the morphological processing, and on the highest level the semantic processing or analysis. The information arising directly from or derived from each level is passed down from the higher levels to the lower as expectations or is passed up from the lower to the higher levels in the form of hypotheses.

The main difficulty in classifying an event as either heart sound or noise is the nonuniqueness of event attributes across both data blocks and gravidae. Two events possessing identical characteristics may, depending on the context, in one instance be noise and in the other a valvular sound. The use of contextual information in the

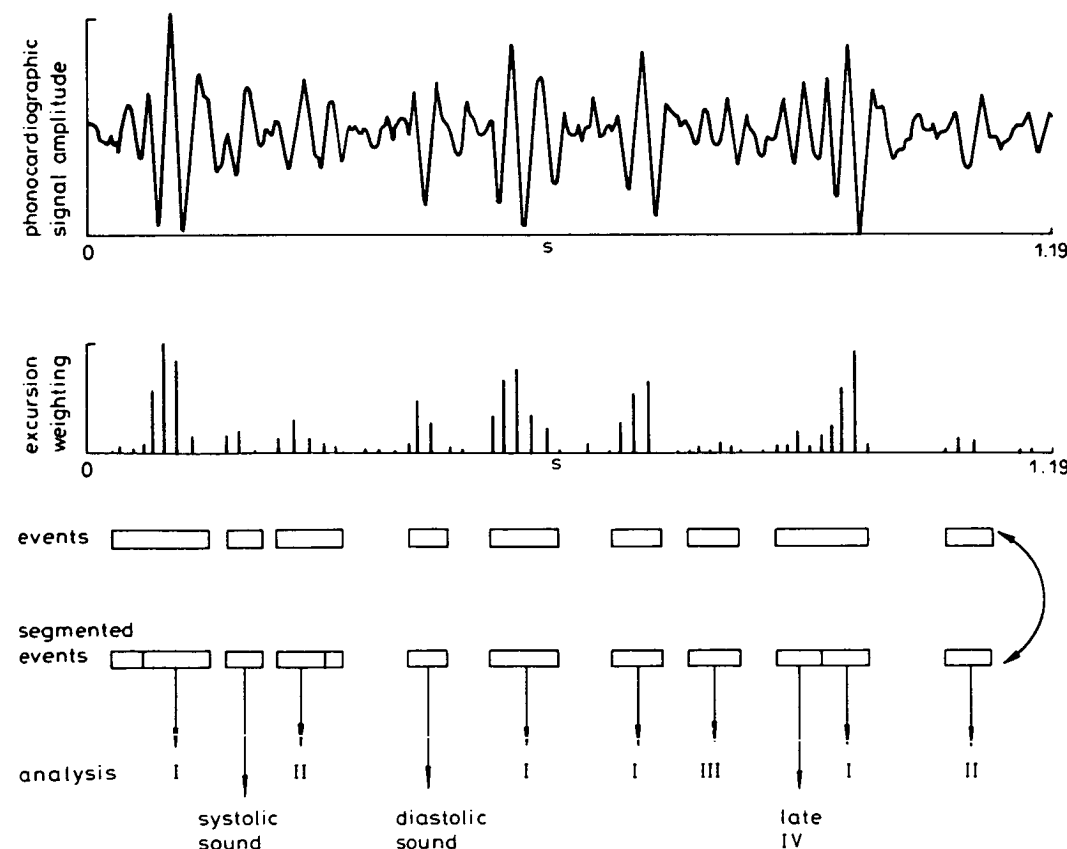


Fig. 6 Levels of signal abstraction held in the global database

identification and classification of the heart sounds is the one element which ultimately makes the problem tractable. It is for this reason that the data is captured in sufficiently long time frames to allow the constraints imposed by the mechanisms of principal heart sound production to be fully exploited. However, to best use the context the analysis must lag behind the processing and may only begin once the complete time frame has been captured.

The method used at the analysis level of the PhCG is the one which a domain expert would use. This involves examining the PhCG to find areas where, with confidence, it is possible to identify certain of the significant signal events as being of valvular origin. Once these 'solution islands' [11] have been identified, the goal is to expand outwards from these 'islands' into regions where the solution is not so immediately obvious.

This narrows the field for the model-driven search, which uses the information contained in Fig. 2 for events which satisfy the criteria for consideration as heart sounds. These events are then tested within the overall context to find whether they fit into the unfolding pattern of positively identified sounds. This process is also aided by the redundancy of information inherent in the repetitive but nonperiodic acoustic signal from the heart.

Ultimately the expansion will, depending on the quality of the data, result in the merging of the expansions from the solution islands and the complete analysis of the PhCG. The search space over which the expansion from solution islands takes place is bounded by the physical temporal and causal constraints of the physiological process which generates the cardiac acoustic signal.

5 Knowledge-based signal processors

The low-level processes of the system, i.e., those closest to the primary data in the processing hierarchy, are themselves regarded as knowledge areas within the overall signal analysis expert system. The function of these processing stages is to remove noise and to enhance the principal sounds in the PhCG prior to classification.

5.1 Detection

Detection is the process whereby certain events in the PhCG are discriminated from the background. In conventional signal-processing algorithms detection is achieved using numerical computation. This will either be some form of filtering when the signal shape or its statistics are known, or domain transformation. Such techniques are of limited use in this context because of the nondeterministic and variable nature of the data. The knowledge required to perform the task of detection is the least expert of all the knowledge embodied in the system. However, it performs a function on which depend all the other KAs, and, ultimately, the quality of the subsequent PhCG analysis. A layman on examining Fig. 3 could easily spot the significant events in this short data record. Although a simple task for the human eye, it is a difficult one to emulate by computer.

What seems to make certain events visually significant to an observer is primarily their local amplitude contrast and, secondly, their frequency and structure. The structure refers to the smooth continuity of the PhCG from the beginning to the end of an event. Contrasted with this is the ill-structuredness of the background. The background is here taken to mean not only physiological and

environmental noise but also sounds of nonvalvular origin.

The detection process cannot be discriminating between principal sounds and nonvalvular sounds as one may masquerade as the other. Furthermore, no morphological similarities have been observed empirically within a class, or consequently during the block processing time frame. In these circumstances the function of the detector cannot be to directly isolate the components of valvular sounds from the background. Rather, the detector proposes what elements could possibly be these components. Owing to the nature of the PhCG the detector always overestimates the number of candidates.

The detection process acts on a per excursion basis. An excursion is defined to be that part of the phonocardiogram bounded between two immediately adjacent transitions across the horizontal. If the excursion meets the criteria for being classed as significant, namely the combination of local amplitude contrast, frequency, and continuity, it is assigned a weighting in proportion to these three factors.

5.2 Removal of insignificant excursions

It has been found that whenever the domain expert inspects the PhCG, he subconsciously takes account of the global scenario first. Then, working at the local level, he accepts or rejects events in the context of the overall scenario while progressing causally, i.e. left to right, across the phonocardiogram.

This procedure has also been adopted here for the automated removal of insignificant excursions. Although the detection process extracts the significant excursions, within this class there will be varying degrees of significance. Once the excursion processing has been completed, which furnishes global information and corresponds to the expert's initial broad overview, the local level is revisited in the light of this extracted global knowledge. An adaptive thresholding technique has been employed to remove the insignificant excursions. The level at which the threshold is locally set, takes account of how the excursion energy levels are varying with time in the context of both the local and global energy fluctuations. As sometimes happens a very weak valvular sound may be discarded at this stage. Although undesirable, this situation can be recovered later by the higher level processes which rely on information from the model and expectations drawn from the context.

5.3 Grouping of excursions

The outcome of this global/local processing is termed the *significant* excursion of the PhCG. The term 'significant excursion' implies a locally pronounced and bounded part of the PhCG which exhibits the characteristics of a component of a valvular sound. This, however, does not imply that it necessarily is one.

Contiguous excursions are grouped together to form the so-called significant events. The terminal points of these groupings are established when an *insignificant* excursion is encountered. These groupings inherit the attributes of their constituent excursions so that the whole event assumes the quality of being heart sound like.

5.4 Feature extraction

Once the significant events have been located, the feature extraction process transforms these events into a set of features. These features are: start time, end time, duration, total energy content, number and location of the

dominant energy peaks, number of excursions, information about the rate of change of excursion duration, and energy distribution. These features have been empirically observed to be the most useful measures in a parametric description of an event and offer the most discrimination between heart sounds and noise. Each event is represented as a vector of these features. These vectors, which are data structures, are then assembled into a linear array which constitutes the first level of data abstraction.

The variability in all aspects of an actual valvular sound has the effect that the discriminating criteria which delimit a significant event cannot be stringent, with the consequence that many more events are taken to be significant than are valvular sounds in the PhCG trace. This overestimate will depend on the quality of the data; for example, in the case of a moderately noisy PhCG there will typically be twice as many spurious valvular sounds as actual ones. The advantages of a weak discriminant is that the chance of rejecting a valvular sound as insignificant is small. Consequently, the higher level processing stages do not have to contend with biased information. This deferral of discrimination until more knowledge can be brought to bear prevents premature decisions. Its disadvantage is that the higher levels are more taxed to ascertain which of all the events are the valvular sounds.

5.5 Segmentation

Whenever the grouping module proposes the concatenation of certain excursions as an event, it does not consider the possibility that the resulting 'event' may not be homogeneous. That is, the event may be composed of more than one sound. This conglomeration arises mainly from the proximity of the sounds within the 'event' (Figs. 3b and d), and from a smooth continuum of the previously outlined weights (Section 5.1) across the sounds. Segmentation is the process which attempts to decompose these coalesced sounds. The localisation of a sound depends on the determination of its end points, particularly the starting point, which in turn effectively marks the end point of the preceding event.

Segmentation is performed in two passes: initially excursions are clustered (Section 5.3), then each cluster is tested for homogeneity. When segmentation is first attempted, i.e., at database initialisation and therefore without contextual information, it is still possible to decompose some of the combined sounds, e.g., when the constituent sounds exhibit an obvious dissimilarity. Although the amplitude distribution in the primary data may give no clue to a coalescence, both the energy and the excursion duration distributions may reveal its presence. One of the commonly occurring coalescences is that of a diastolic sound and a I heart sound (Figs. 3b and d). Although it cannot be said what the particular features of a valvular sound are, there are certain characteristics which disclose a nonvalvular sound, e.g., an instantaneous frequency across the event which increases with time. In these circumstances the nonvalvular event is removed. However, nonvalvular events are not always identified at this stage, so segmentation of coalescences must be delayed until reprocessing occurs under the guidance of contextual information.

Once segmentation has been completed, the feature extraction module (Section 5.4) is recalled to process the resulting events. This refined version of the significant events is posted in the database as the second level of data abstraction.

Although the segmentation KA has been included under the section dealing with preprocessors, it may be, if

required, later recalled under interpretation control [12] to reperform its function on certain stretches of the PhCG. As such, segmentation is also a guided data reprocessor.

6 Knowledge acquisition

The knowledge acquisition stage of this project is unusual insofar as it did not involve direct input from what are conventionally thought of as 'domain experts'. Their contribution was rather more indirect, coming, as it did, from the extensive literature relating to adult and foetal heart sound analysis.

Why this is so has to do with the way obstetricians determine the foetal heart rate. This task, in the majority of cases, is accomplished by direct auscultation with a stethoscope and consists of counting the occurrences of a particular sound over a certain interval. Thus, the domain expert is an expert in the aural processing of the heart sound signal rather than its visual processing when presented as a PhCG. What is more, the information represented in a PhCG contains significant subsonic components. The literature referred to above, specifically Leatham [13] and Kelly [14], presents a catalogue of exemplars of heart sounds extracted from PhCGs which are intended to be an aid to auscultation. These sources do not address the problem of isolating the heart sounds from the PhCG which, it would be tacitly assumed, would not pose any difficulty, especially in the relatively noise-free PhCGs they chose. However, this textbook information is sufficiently complete to act as a 'domain expert' whenever such assistance is required.

It is not a prerequisite that one should be skilled in cardiology or obstetrics to be able to identify two repetitive, but nonperiodic, signal events in the foetal PhCG. In fact, having established the fundamental nature of heart sounds and their production (Section 2), a layman could then analyse a PhCG with a high degree of confidence. Even in the presence of adventitious sounds and high levels of noise, a visual inspection can locate and classify the foetal heart sounds from among many competing events. How does the eye achieve this? Certainly it fixes upon certain salient features: higher amplitudes, event homogeneity, event/background dissimilarity and structure, and particularly the repetition of the principal events across the PhCG. These factors have all been incorporated into the knowledge-based signal-processing modules.

The heuristics employed in the low-level signal-processing stages had to be assembled from a basis of little *a priori* knowledge. This is attributable to the rapid decline in the use of phonocardiography in obstetrics since the advent, in the early 1960s, of Doppler ultrasound methods for foetal heart-rate (FHR) monitoring. From examination of many PhCGs recorded in various subjects of differing gestational age, certain characteristics emerge from the data regarding the heart sounds. This process was carried out by the system developer himself, who subsequently became the 'domain expert'. The accumulation of expertise was a painstaking process which was very time consuming, but in the absence of an alternative it was the only recourse. It has the advantage, however, that knowledge elicitation is easy.

Whenever a PhCG is proving difficult to analyse, there are other sources of information brought into the analysis. These factors derive from the experience gained through auscultation regarding the limits of the FHR, the range of beat-to-beat variability of the FHR, and relative

durations of the systolic and diastolic times, which are to be found in the literature.

7 Implementation

The knowledge-based system described has been implemented in the C programming language under UNIX, mainly because C produces very fast and efficient executable code. In this application the nature of the task demands speed of response, and the amount of processing required for long-term continuous monitoring is considerable.

In those computationally intense KAs, notably the signal processors, the speed with which they perform their task is expedited by the determinism of the decision-tree structure. A further advantage of the tree structure is that it enables the designer to ascertain whether the knowledge within a task is complete and consistent. The explicit coupling between rules in the KAs means that modifying a KA may be difficult. However, it is less difficult than revising rules when the coupling is not explicit, which may lead to unintended interactions.

Realisation in C also permitted use of the data structures and pointer facilities provided in that language to implement the various elements of the system. The disadvantage of not using a knowledge representation language is the lack of the built-in facilities which these possess. Consequently, the C code tends to be rather long and complex.

8 System performance

Although the system is fully developed, it has not, as yet, been used clinically. To permit comprehensive testing, eight-hour overnight recordings of the PhCG were taken from 18 different gravidae and stored on magnetic tape. From these a representative sample set of 250, four-second blocks of data were extracted. Each of these test PhCGs was chosen on the basis that the foetal heart sounds should be identifiable by visual inspection and that, overall, the signal should be moderately difficult to analyse. This was to avoid those cases where, even to a visual inspection, the signal is obscure and in the other extreme where the sounds are strong and there is little noise.

The task of analysing the phonocardiogram is one of visual processing. It is against this standard that the analysis of the system is judged. Presently, no other

system is available which can automatically identify the instants of closure of the cardiac valves.

The analysis obtained by the system concurs with that of the domain expert in all but the most difficult cases. In these latter instances either the system is able to analyse only certain parts of the PhCG, or the analysis degrades with missed or incorrectly identified heart sounds. However, if the system were to be run unsupervised, an automatic quality check would be needed to reject severely noise-corrupted signals before they were passed to the analysis stage.

Fig. 7 shows an example of a complete analysis of a PhCG which was achieved by the system on a heavily noise-contaminated PhCG.

The computational performance of the system was not specifically addressed in the design. The primary objective was to ascertain if it were possible to achieve an automatic analysis of PhCGs. Nevertheless, an analysis of a four-second PhCG takes, on average, six seconds on a VAX 11/750.

9 Summary

This paper has presented the realisation of a knowledge-based expert system for the identification of the principal heart sounds in the PhCG. The variable nature of the signal precluded using the standard analytic methods of signal processing whose idealising assumptions about the signal and its context could not be satisfied within the reality of this particular application. The deployment of knowledge-based signal processing and analysis has enabled a task to be performed automatically which previously could only be reliably carried out through visual inspection.

The power of the knowledge-based methodology derives from the expert's knowledge of the domain, and the perspective of the system which considers the overall signal environment. Knowledge-based signal processing, of necessity, must be domain-specific and empirical. Although lacking the rigour of analytic methods it has the advantage that it remains close to reality, and possesses a flexibility which allows it to accommodate the unpredictability of the signal. This application specificity is the strength of the method.

The essential feature of this system is its organisation as a hierarchically structured procedural expert system. This formulation allows an efficient search for applicable rules in an application where speed of response is important.

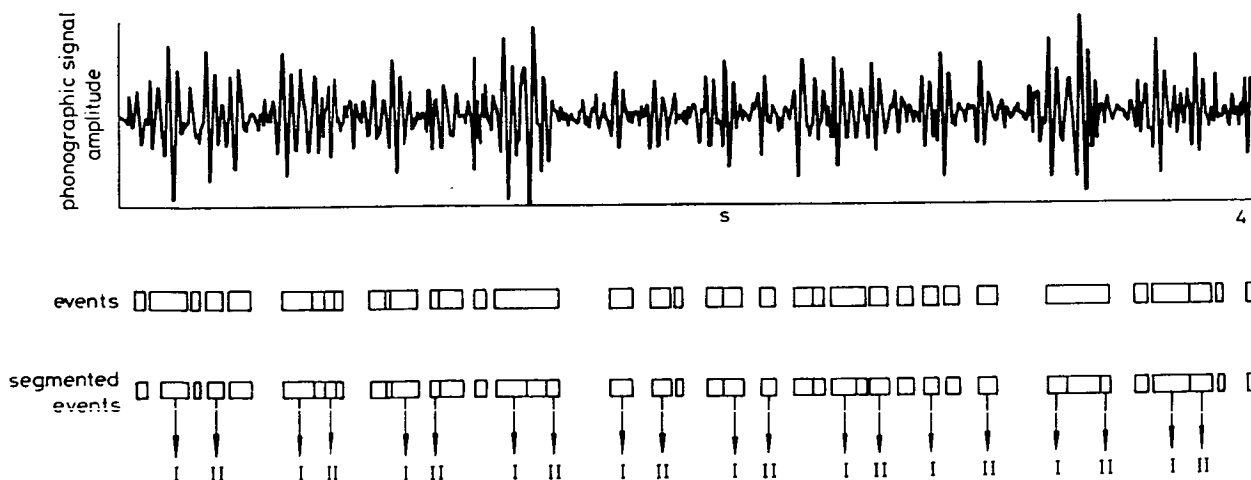


Fig. 7 Analysis of a noisy PhCG performed by the KBS

The terminal edges of the boxes on the event and segmented event levels refer directly to the extent of the event immediately above in the PhCG. The labels I, II on the analysis level represent identified first and second heart sounds, respectively

The analysis of the PhCG begins with the detection and segmentation of events. These events are then searched for solution islands. From these islands the system attempts to expand the analysis into the more ambiguous regions. A complete analysis of a data block is achieved when all the principal sounds have been identified. The extent of an analysis will be determined by the ambiguity of the PhCG which in turn depends on the noise levels and the strength of the valvular sounds.

Incorporated within the expert system is an information feedback mechanism from the higher to the lower levels of data abstraction. This allows the contextual and abstracted information on the higher level to modify the parameters of the knowledge-based processing functions which will then be directed to reprocess certain areas of the signal in the light of the prevailing conditions.

10 Acknowledgment

The funding for this research, which was provided by the Department of Education for Northern Ireland, is gratefully acknowledged.

11 References

- 1 NII, H.P., and FEIGENBAUM, E.A.: 'Rule-based understanding of signals', in HAYES-ROTH, F. (Ed.): 'Pattern-directed inference systems' (Academic Press, 1978), pp. 483-501
- 2 HAYES-ROTH, F., and WATERMAN, D.A.: 'Building expert systems' (Addison-Wesley, 1983)
- 3 CORNER, G.W., and STRAN, H.M.: 'A fetal heart ratemeter', *Am. J. Obstet. Gynec.*, 1957, 73, pp. 190-195
- 4 LEONARD, J.L., and FARRER, W.: 'A foetal heart monitor', *Electron. Eng.*, February 1963, pp. 83-85
- 5 COX, L.W., WALL, I.B., and WOOD, A.E.R.: 'The design of a foetal heart ratemeter', *Aust. N.Z.J. Obstet. Gynaec.*, 1963, 3, pp. 60-62
- 6 MAXWELL, I.: 'Phonocardiography as a foetal heart monitor', *Aust. N.Z.J. Obstet. Gynaec.*, 1963, 3, pp. 63-72
- 7 CORNWALL, J.B., and TATTAM, F.G.: 'An instrument for recording the fetal pulse rate', *IEEE Trans.*, 1964, BME-11, pp. 24-28
- 8 TALBERT, D.G., DAVIES, W.L., JOHNSON, F., ABRAHM, N., COLLEY, N., and SOUTHALL, D.: 'Wide bandwidth fetal phonography using a sensor matched to the compliance of the mother's abdominal wall', *IEEE Trans.*, 1986, BME-33, (2), pp. 175-180
- 9 COLLEY, N., ABRAHAM, N.G., FAYERS, P., TALBERT, D.G., DAVIES, W.L., and SOUTHALL, D.P.: 'The fetal phonogram: a measure of fetal activity', *The Lancet*, 26th April 1986, pp. 931-935
- 10 GEORGEFF, M.P.: 'A framework for control in production systems', *IJCAI*, 1979, pp. 328-334
- 11 NII, H.P.: 'Blackboard systems', *AI Mag.*, Summer 1986, pp. 38-53
- 12 TENENBAUM, J.M., and BARROW, H.G.: 'Experiments in interpretation guided segmentation', *Artif. Intell.*, 1977, 8, (2), pp. 241-274
- 13 LEATHAM, A.: in 'Auscultation of the heart and phonocardiography' (J. & A. Churchill, London, 1970)
- 14 KELLY, J.V.: 'The fetal heart: comparison of its electrical and mechanical events', *Am. J. Obstet. Gynec.*, 1965, 91, (8), pp. 1133-1140